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

Katarina Richardson and Gerard J. van den Berg*

Summary

■ The vocational employment-training program is the most ambitious and expensive amongst the active labor market policy programs for the unemployed in Sweden. 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 exploiting variations in the timing of training and exit to work. The approach involves estimating 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 the program. Its magnitude is very large shortly after finishing the course but it then diminishes. If we also take the time spent in the program into account, the net effect of participation in the program on the mean unemployment duration is close to zero.■

JEL Code: J64, J24, I28, J68.
Keywords: Vocational training, program evaluation, duration analysis, Swedish labor market, selectivity bias, treatment effect.

* *Katarina Richardson is a research fellow at IFAU in Uppsala. Gerard J. van den Berg is professor of economics in the Department of Economics at the Free University Amsterdam. He is also affiliated with IFAU in Uppsala, the Tinbergen Institute and CEPR.*

SWEDISH ECONOMIC POLICY REVIEW 8 (2001) 175-213

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

Katarina Richardson and Gerard J. van den Berg*

Training programs for the unemployed have constituted the cornerstones of the Swedish labor market policy for many decades. In fact, such programs have been used since 1918. Nowadays, the so-called employment-training program (here simply called the AMU program) is the most prestigious. It aims at improving the chances of unemployed job seekers to obtain a job, through courses where their skills are substantially improved. In 1997, on average 37,000 individuals per month participated in AMU, which corresponds to over 10 percent of the total number of unemployed.[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 evaluational studies is rather small, and most of these analyze the effect of AMU on the participants' earnings and/or use data from the 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 and the results are of obvious importance for evaluating the AMU program and the underlying "Swedish model". In addition, they are important in the light of recent policy shifts in other European countries towards an increased use of active measures for returning the

* *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.*

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

unemployed to work, notably through reschooling unemployed workers with low skills or obsolete qualifications (see e.g. Fay, 1996).

We use a unique fresh set of longitudinal administrative data containing the complete population of unemployed individuals in Sweden within the period January 1, 1993 to June 22, 2000. This dataset matches detailed records from employment offices with records from unemployment insurance agencies. The employment office data provide detailed information on the types of training and the corresponding dates of entry and exit. A large number of individuals have experienced multiple unemployment spells within our observation window.

The empirical analysis applies a methodology where information on the *timing* of events (like the moment when the individual enrolls in training and the moment he finds a job) is used to estimate the individual training effect. This method takes into account that the training decision and the decision to accept a job are affected by unobserved factors. If individuals receiving training were to have found a job relatively fast anyway, and if this is ignored in the analysis, 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 requiring the selection effects to be completely captured by observed variables (like the so-called matching approach), in particular if the set of variables observed 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 changes in the individual training effect over time and we can learn something about the reasons why training works or not.

The empirical approach involves estimation of models simultaneously explaining the duration of unemployment before obtaining work or participating in training programs. The approach has been used in a number of empirical studies, like Eberwein et al. (1997) and Van den Berg et al. (2002) (on training programs), and Abbring et al. (1997) (on punitive sanctions for unemployment insurance recipients; see also Gritz, 1993; and Bonnal et al., 1997, for extensions focused on training programs; and Van den Berg, 2001, for a survey). Abbring and Van den Berg (2000) provide an underpinning of this approach, and a systematic account of the behavioral assumptions required for valid use.[2] For example, individuals are not allowed to anticipate the moment when they go into training a long time in advance, although they will know the distribution of this moment over time. We argue that AMU fits rather well into this methodological framework, whereas other active labor market programs in Sweden may not. To demonstrate this, we rely to some extent on evidence from in-depth interviews with caseworkers, and to some extent on existing studies on unemployment, unemployment insurance, and active labor market programs in Sweden. These include Eriksson (1997a,b), 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 the expiration of benefits entitlement on the other. We return to this issue in Sections 1 and 2.

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 finished a program in the final quarter of 1996. They use a subset of our data and match it to data from a postal survey conducted in late 1997. Harkman and Johansson estimate a bivariate probit model on the employment probability, for different programs one year after the end of the program. Furthermore, 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 rela-

[2] Here we just provide some clarification as to what drives the identification of the training effect. Consider individuals who enter training at time t. The natural control group consists of individuals 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 t. The duration model framework allows for such randomization because it specifies assignment by the *rate* of entering training. In addition, we must deal with the selection issue that the unobserved heterogeneity distribution differs 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 of explaining identification emphasizes the importance of the timing of events. If training and outcome are typically realized very shortly after each other, notwithstanding the length of the elapsed duration in the state of interest before the training, this is evidence of a positive causal training effect. The selection effect does not give rise to the same type of quick succession of events.

tion between the composition of programs and the probability of exit to work indicates a training effect. However, it is not clear whether the assumption is justifiable. Their results indicate that there is a higher probability for individuals in AMU to get a job. Subjective responses to the perceived importance of program participation correspond with the estimation results.[3]

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

The paper is organized as follows. Section 1 describes the AMU program. In Section 2, we briefly discuss the model framework and we argue that AMU fits into it. The technical details are relegated to the appendix. Section 3 describes the data. Section 4 contains the main estimation results. Section 5 concludes.

1. Swedish labor market training

1.1. The AMU program

The purpose of the AMU program is to improve the chances for job seekers of obtaining a job, and making it easier for employers to find workers with suitable skills. The aim is to increase unemployed indi-

[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 several variables in 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; in our empirical analyses, we restrict the attention to individuals aged over 25 (see Subsection 2.2). 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.

viduals' transition rate to work, which is achieved by the participation of individuals in training and educational courses.[4]

The program is targeted at unemployed as well as employed individuals running the risk of becoming unemployed. The individuals must 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 younger individuals are nowadays entitled to participate if they are disabled or receive unemployment insurance (UI) benefits.

During the 1980s, the yearly average number of individuals in AMU was about 40,000 per month. During the large Swedish recession of the early 1990s, this number increased to 85,000, with seasonal peaks of about 100,000 participants. After 1992, this number once more decreased to about 30,000-40,000, which is about 1 percent 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 (USD 1.2 billion), half of which was for training procurement and half for training grants. This equals about USD 10,000 for procurement and USD 10,000 for grants per participant, on a yearly base (AMS, 1997).

There is strong evidence that in 1991 and 1992, participation in AMU was often used to extend benefits entitlement (Regnér, 2002: and Edin et al., 1998). This requires a brief exposition. A commonly recognized problem with Swedish labor market programs is that they are used to extend an individual's entitlement to unemployment benefits (300 working days (≈ 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 the completion of the program; in fact, if the program runs over more than a few months, 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 the expiration of benefits entitlement on the other, but do not consider differences across programs. They find that many unemployed workers move into programs

[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).

shortly before expiration. Carling et al. (1996) use data from 1991-1992 to study these issues, 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 at individuals whose benefits entitlements expire. Participation usually amounts to performing tasks in the non-profit private sector that would otherwise not be carried out. In 1993, the size of other non-AMU programs also increased, and other new programs were designed. Once more, these programs are much cheaper than AMU.

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

Non-vocational training (basic general training) means participating in courses within the regular educational system, i.e. at adult educational centers and universities. Non-vocational training specifically aims at preparing the individual for other types of training (the goal of an increased transition rate to work is thus less direct). Before 1997, a substantial part of AMU consisted of non-vocational training. In 1997, a new program of adult education (called the Adult Education Initiative, or Knowledge Lift) was introduced which is, amongst other things, supposed to replace the non-vocational training part of AMU (see Brännäs, 2000). Nevertheless, since January 1995, non-vocational training amounts to approximately 40 percent of all AMU courses. For 2000, this number is even higher (about 50 percent).

Concerning UI, it should be mentioned that entitlement also requires registration at the employment office. In the mid-1990s, about 40 percent of the inflow into unemployment and about 65 percent of

[5] Note that this also suggests that workers do not enjoy training a great deal, since they would otherwise 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.

the stock of unemployed qualified for UI (Carling et al., 2001). Part of the remaining 60 percent received "cash assistance" benefits, which are typically much lower than UI benefits. The average replacement rate for UI recipients is about 75 percent (Carling et al., 2001).

During the training, the participants' income is called a training grant. Those entitled to UI receive a grant equal to their UI benefits level, with a minimum of SEK 240 per day (about USD 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 must submit 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 the 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 programs and changes in program participation over time, respectively. In 1997, on average 191,000 individuals (4.5 percent of the total labor force) participated in one of the programs. The government's part of the total costs has exceeded 3 percent 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 above 55 differ from those aged between 25 and 55. Young persons must participate in a program after 100 days of unemployment, otherwise they lose their unemployment benefits. They may use special programs not available for other age groups. Persons above 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, but they find no significant crowding out effects of AMU.

1.2. 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 from documents from the Swedish National Labour Market Board (AMS) (see e.g.

AMS, 1998) and 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 results from the interviews.

Usually the employment offices advertise the availability of AMU courses at the office and in the newspapers. Most of them advertise one or two months before the scheduled starting date and they invite those interested to an information meeting. At this meeting, the individuals are informed about the contents of the course and the eligibility rules. The personal caseworkers are usually also available 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 as concerns the educational level. However, most courses are at a fairly basic level, so these requirements are not very 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 applicants undergo a test to determine whether they will benefit from the course. One may, for example, test the skills in mathematics or the Swedish language and there may also be some ability testing. Another way of determining whether the individual's skills can be enhanced is by examining his "expected" unemployment duration, which is considered to be high in case of a low 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 explaining why he wishes to participate in a specific AMU-course. If the person has work experience in the relevant 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 individual's original occupation. If an individual rejects a caseworker's offer of an AMU course, his unemployment benefits may, in principle, be completely cut off, but this does not seem to happen in practice.

Occasionally, caseworkers may work closely with firms requiring certain skill categories. These firms may have an influence on who is

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

accepted in the program. Training (of the unemployed individual) and job search effort (done by his caseworker) may go hand in hand, so that the effect of AMU may consist of a skill enhancing as well as a search effort effect.

If the number of applicants is insufficient, 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, 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 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 for dealing with such cases, like workfare programs, and efforts are instead made to push the individual into those programs. Similarly, AMU is generally not offered to individuals who, in the caseworker's opinion, need practical experience in order to be able to get a job, or just "need something to do". In such cases, the individual is offered another active labor market program, like a work experience program.

There is approximately one month between the first information meeting and the first day of the course. On average, the period from application to acceptance is 2-3 weeks, while the period from acceptance to the start of the course is 1-2 weeks. An individual may try the AMU-course before actually starting the course. For example, if he is interested in welding, he can make a one-week visit to the school offering welding courses. Further, 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 leading to an individual's enrolment in AMU is confirmed by existing empirical studies. Eriksson (1997a,b) 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 also constitute the major administrative data set for our own analyses. It is shown that the personal characteristics observable in HÄNDEL cannot give a very precise prediction of actual participation in AMU versus non-participation. The predictive performance can be substantially enhanced if accounting for self-reported (by the unemployed) measures of the extent to 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. (Naturally, they may also reflect an ex-post rationalization of actual choices in the past.) Eriksson (1997a) notes that informal interviews with caseworkers reveal that the motivation of the unemployed is a very important criterion for placing an unemployed individual in AMU.

Eriksson (1997b) exploits survey data obtained by letting caseworkers give AMU-advice on the basis of actual files of unemployed individuals supplied by the survey agency. The allocation of files to caseworkers is completely random. 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 the heterogeneity of 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. Thus, there is a great deal of variation in the caseworkers' decisions which cannot be attributed to the unemployed individuals' identities but to the caseworkers' identities. When making selections on the basis of observable personal characteristics, officials seem to use rules of thumb which are often not in accordance with the stated goals of AMU on priority groups. If the caseworkers consider that an individual would benefit a great deal from participation, the individual is also more likely to be an actual participant. But the actual participation also depends on the unemployed individual and unexplained factors.

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

According to Eriksson (1997b), caseworkers are reluctant to let current participants in non-AMU programs enter AMU. Further, work experience programs and public temporary employment are substitutes for each other, but not for AMU. Caseworkers consider AMU to be a fundamentally different kind of program. Thus, 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), AMU is nowadays typically not used for UI entitlement extensions.

2. The model framework

2.1. A class of bivariate duration models for treatment evaluation

We normalize the point in time when the individual enters unemployment to zero. 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 consider the period in AMU as part of the unemployment spell. At the moment, we also ignore time spent in other training programs. As a result, T_u also denotes the duration of unemployment. The population considered concerns the inflow into unemployment.

We assume that the individual distribution of T_p can vary with the observed and unobserved explanatory variables x and v_p, respectively. Similarly, we assume that the individual distribution of T_u can vary with the observed and unobserved explanatory variables x and v_u and 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 previously been completed, as a function of t. It provides a full characterization of the duration distribution (see Lancaster, 1990; and Van den Berg, 2001). Somewhat loosely, this is the speed at which the duration is realized.[7] We the use notations $\theta_p(t|x,v_p)$ and $\theta_u(t|T_p,x,v_u)$, respectively. We do not require

[7] For a nonnegative random (duration) variable T, the hazard rate is defined as $\theta(t) = \lim_{dt\downarrow 0} \Pr(T \in [t,t+dt) \mid T \geq t)/dt$. 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 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 conditioned upon.

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 from unemployment. Treatment and exit are characterized by the *moments* when 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. This 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 hazard rates $\theta_u(t|t_p,x,v_u)$ and $\theta_p(t|x,v_p)$,

$$\theta_p\left(t|x,v_p\right) = \lambda_p(t)\cdot \exp\left(x'\beta_p\right)\cdot v_p \tag{1}$$

$$\theta_u\left(t|T_p,x,v_u\right) = \lambda_u(t)\cdot \exp\left(x'\beta_u\right)\cdot \delta\left(t|T_p,x\right)^{I\left(t>T_p\right)}\cdot v_u \tag{2}$$

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 that $\delta(t|T_p,x)$ is equal to a constant larger than one. If T_p is realized, the level of the individual exit rate to employment increases by a fixed amount, which will reduce the remaining unemployment duration in comparison to the case where the individuals enter AMU at a later point in time. In more general, we allow the effect of AMU to vary with the moment T_p of entry into AMU and 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 time $t-T_p$, since entry into AMU (in fact, in the empirical analysis we only consider the variation of the effect with $t-T_p$ and not with t or T_p separately). The effect of $t-T_p$ may capture the fact that the exit rate is low during the training course or high immediately after 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 with the realized unemployment durations of other individuals. If the individuals entering AMU at t_p have relatively short unemployment durations, this might 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, 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 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. we may observe more than one unemployment spell for individuals in the sample. 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 larger the number of 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 a 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 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 at t+100. This basically rules out the anticipatory effects of the training. If an individual does anticipate participation in AMU at a particular future date t_p, he may want to wait for the training by reducing his search intensity for jobs, i.e. he may change strategies which may decrease the probability that T_u is quickly realized. If this is ignored in the empirical analysis, the training effect may be over-estimated. However, if the time span between the moment when the anticipation occurs and the actual training is short relative to durations T_p and $T_u – T_p$, and if the anticipatory effect is not very large, then the estimation results may be rather insensitive to the assumption of no anticipation.

With well-established programs like AMU, it is plausible that the *determinants* of the training assignment affect the individual's exit rate from unemployment *before* his actual entry into training. For example, at any time before participation in AMU, 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 from unemployment. Such an effect should not be confused with anticipating the *realization* of entry into training, because in the latter case, the individual knows the stochastic outcome rather than the determinants of the process. Likewise, the lack of anticipation does not mean that ex ante effects of AMU are ruled out.

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

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 of a future job opportunity that is independent of whether training will occur, and he 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, in turn, would affect the rate at which entry into training occurs. If something similar occurs 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 spans over a long period of time, such an effect may be unimportant, as employers may be unwilling to wait for a new employee for several months. If the time span between the moment of anticipation and the moment of the actual exit to work is relatively short, and if the anticipatory effect is not very large, the estimation results may be rather insensitive to this. Once more, lack 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, 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 $\lambda_p(t)$.

A more technical aspect of the model specification follows from the fact that we specify the assignment of training by 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, it is difficult to address its effect. In the extreme case where individuals can only enter AMU, 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 also impossible to justify that entry into AMU is not anticipated.

2.2. 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 1.2 and the studies by Eriksson (1997a,b), it is obvious that unobserved (to us) heterogeneity of 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 individual's future career) are also indications of unobserved determinants of the individual exit rate to work. The empirical analysis should therefore take potentially related unobserved heterogeneity terms in θ_u and θ_p into account.

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 he knows that he may be subject to treatment, then he has a strong incentive to question the actual value of the variable. Subsequently, he will take his value of the variable into account when determining his optimal strategy which, in turn, affects the rate at which he moves to employment. The variables 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, this is generated by the variation in the moment when 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 the anticipation of the moment of entry into AMU. From Subsection 1.2, the time period between the moment when the individual is informed about the possibility of enrolling in an AMU course and when the course starts is very short. (Naturally, we allow individuals to be aware of the existence of the AMU *program*.) There are, however, two reasons why individuals may anticipate the moment of entry, both of which make us restrict the scope of the empirical analysis somewhat.

First, as shown in Section 1, AMU was often used to extend benefits entitlement in 1991 and 1992. In that case, the date of inflow into AMU is mainly determined by the date of expiration of the 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 anticipating 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 that we may expect the unobserved heterogeneity distribution to differ over time. From January 1993 and onwards, other programs took over its role as a means of extending benefits entitlement. We therefore restrict the attention to data from 1993 and onwards.

Second, recall from Section 1 that part of AMU concerns non-vocational training (in particular before 1997) which primarily aims at preparing the individual for other types of training. It is often given within the regular school system and 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 terms. As a result, it is easy for an unemployed individual to anticipate the date of inflow into such a program and we therefore restrict ourselves to vocational training. There are actually two other reasons for this. First, vocational training is relatively expensive and second, it 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 1.2 that caseworkers regard vocational AMU training as a very different type of program than other active labor market programs. The latter are regarded to be substitutable to a high degree. For persons below 25, there are programs that are more similar to AMU vocational training. Further, the similarity with courses and tracks in the regular school system may be important for these individuals. For this reason, we restrict the attention to individuals over 25. Young individuals must also 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 vocational AMU training seems to have relatively small advantages for them.

It follows from the above that our model framework may be less suited for the analysis of the effects of other active labor market programs on unemployment duration. For 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 because of the high degree of substitutability.

3. 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 HÄNDEL , which contains information on unemployed individuals' training activities and work experience activities, is the most important source. 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 participation in the program. According to Carling et al. (2001), more than 90 percent 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 and onwards and provide information on the wage level and working hours in the employment prior to the spell of unemployment, for individuals eligible for UI. The full HÄNDEL data are also informative on whether an individual in AMU obtains vocational or non-vocational training.

Our observation window runs from January 1, 1993 until June 22, 2000. The unit of observation is one individual. For each individual who appears 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

[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.

HÄNDEL and AKSTAT is used to make a list of characteristics at the beginning of the spell, and a list of dates when changes occur, including the nature of the change. It is particularly important to include 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 on the first date when he registers at the employment office as being "openly" unemployed. This eliminates registration spells that start because the individual wants to change employers 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 actually becomes unemployed. We also ignore unemployment spells that have already stated at the beginning of the observation window, because using these 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 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 differ 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. If the exit destination is employment, we observe a realization of the duration variable of interest. If the exit destination is different (e.g. "regular education", or "other reason"), this duration variable is right-censored. The duration is independently right-censored if the spell continues at the end of the observation window. If exit occurs into "wage subsidy" or "(public) sheltered employment", we remove the individual from the sample, since these programs are for handicapped people (who are typically not in open unemployment anyway). As a result, our data set contains 500,960 individuals. Note that following individuals over time, we may observe multiple unemployment spells per individual.

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

If we treated participation in other programs before participation in AMU as regular unemployment, the transition rate from unemployment into AMU would be extremely low during participation in the other programs. Participation in non-AMU programs most likely also reduces the transition rate into employment, so that it may be preferable to halt the timing of the duration until regular employment, during such a period of program participation. As our baseline assumption, the time spent in training (in non-AMU programs as well as in AMU) does thus 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 2.2, we restrict the attention to individuals who were at least 25 and below 55 when entering 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. As concerns nationality, we also distinguish between three categories: Eastern Europe, Africa/Asia, and others (including Sweden). As concerns the type of unemployment benefits received during unemployment, we distinguish between three categories: UI, cash allowance, and neither. For UI recipients in 1994 and onwards, the AKSTAT data contain the hourly wage earned in the employment held just before the onset of the spell of unemployment. This is almost linearly related to their UI level (see e.g. Carling et al., 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 return 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 percent random subsample of the full data set at our disposal. For each individual, we include three unemployment spells at most. 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.

656 of these spells contain a period of participation in an AMU course, and exactly 8000 do not. Some of the latter are naturally 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 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.

27 percent of the 656 spells observed to include AMU participation are also observed to include participation in another type of active labor market program before participating in the AMU program. Not surprisingly, this predominantly occurs in long spells with a high realized value of T_p. Only 12 percent of the 328 spells with T_p smaller than its median of 161 days, are also observed to contain participation in another type of active labor market program before AMU participation. 18 percent of the 8000 spells not observed to include participation at AMU, are observed to contain participation in another type of active labor market program. If we restrict the focus to spells with T_u smaller than 161 days, this figure drops to 10 percent. Participation in other programs does thus not seem to be related to participation in AMU. Spells with AMU participation relatively often also contain participation in other programs , since 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 percent sample

	All spells	No AMU	With AMU
Individual characteristics:			
log(age)	3.54 (.23)	3.54 (.23)	3.58 (.22)
short senior high school	.26	.26	.29
senior high school	.20	.20	.20
short tertiary education	.05	.05	.05
university	.16	.16	.14
female	.50	.50	.44
UI recipient	.67	.68	.70
cash allowance recipient	.07	.07	.09
from Eastern Europe	.05	.05	.08
from Africa, Asia or S. America	.05	.05	.05
log (hourly wage)	2.69 (2.31)	2.66 (2.32)	3.05 (2.14)
experience in occupation (dummy)	.64	.63	.71
education in occupation (dummy)	.62	.61	.65
professional and technical work	.15	.15	.14
health, nursing and social work care	.14	.15	.09
adm., manag. and clerical work etc.	.13	.12	.18
sales	.11	.11	.11
agriculture and mining	.07	.07	.08
services (incl. not categorized occ.)	.17	.17	.15
large city (dummy)	.52	.53	.45
needs guidance (dummy)	.08	.07	.14
willing to move (dummy)	.15	.15	.18
accepts part time work (dummy)	.05	.06	.03
fraction of spells that starts in:			
1993	.19	.19	.27
1994	.16	.16	.21
1995	.16	.15	.20
1996	.13	.13	.10
1997	.12	.12	.09
1998	.11	.11	.09
1999	.09	.10	.03
2000	.04	.05	.01
Observed labor market outcomes			
spells contains AMU	.08	0	1
spells ends in exit to work	.57	.58	.53
realized t_m	170 (214)	153 (191)	372 (342)
realized t_p	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)

Note: The unit of interval is one spell. Standard deviations in parentheses. The observed labor market outcomes are reported in fractions and days.

4. The empirical analysis

4.1. 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 which 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 for the period from 57 to 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 which has an associated probability. The possible values of the v_i as well as the probabilities are then estimated. 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.

4.2. Estimation results

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 estimating the model under the following assumptions: δ is a constant, the lengths of the time intervals spent in AMU and in other programs are set to zero, and any subsequent participation in AMU after the first course within a spell is ignored. We include data on as many as three unemployment spells per individual, if available.

Table 2. Estimation results for the baseline model

	To work, θ_u		To AMU training, θ_p	
AMU effect				
δ	.83	(.10)*		
Individual characteristics				
log(age)	-.59	(.11)*	.35	(.29)
short senior high school	-.01	(.06)	.20	(.16)
senior high school	-.10	(.07)	.05	(.18)
short tertiary education	.08	(.12)	-.13	(.29)
university	.23	(.09)*	-.14	(.22)
female	.01	(.06)	-.04	(.14)
UI recipient	.26	(.07)*	.04	(.18)
cash allowance recipient	.21	(.10)*	.43	(.23)*
from Eastern Europe	-.58	(.13)*	.35	(.25)
from Africa or Asia	-.85	(.14)*	.18	(.28)
log(hourly wage)	.00	(.03)	.16	(.09)
experience in occupation (dummy)	.17	(.05)*	.12	(.14)
education in occupation (dummy)	.23	(.05)*	.12	(.13)
professional and technical work	-.16	(.09)*	.08	(.22)
health, nursing and social work	-.08	(.08)	-.29	(.24)
adm., managerial and clerical work	-.39	(.09)*	.34	(.22)*
sales	-.29	(.09)*	.03	(.22)
agriculture and mining	.08	(.09)	.09	(.23)
services (incl. non categorized occ.)	-.17	(.09)*	-.06	(.19)
large city (dummy)	-.08	(.05)	-.31	(.12)*
needs guidance (dummy)	-.48	(.11)*	.54	(.20)*
willing to move (dummy)	.04	(.06)	.18	(.16)
accepts part time work (dummy)	.11	(.09)	-.46	(.32)
1994	.24	(.07)*	-.07	(.18)
1995	.20	(.07)*	-.04	(.17)
1996	.24	(.08)*	-.48	(.22)*
1997	.53	(.08)*	-.44	(.22)*
1998	.58	(.08)*	-.37	(.22)*
1999	.75	(.09)*	-1.07	(.22)*
2000	.75	(.15)*	-.69	(.68)

Table 2. Continued….

	To work, θ_u		To AMU training, θ_p	
Duration dependence				
λ_{i2}	.11	(.05)*	-.32	(.15)*
λ_{i3}	.05	(.06)	-.34	(.17)*
λ_{i4}	.10	(.08)	-.18	(.18)
λ_{i5}	.04	(.09)	.05	(.18)
λ_{i6}	-.17	(.12)	-5.00	(.19)*
λ_{i7}	-.13	(.13)		
λ_{i8}	-.19	(.11)*		
Unobserved heterogeneity				
log (v_1)	-5.82	(.11)		
log (v_2)	-7.41	(.16)		
log (v_3)			-8.02	(.68)
log (v_4)			-7.00	(.60)
q_{13}	-1.00		(1.19)	
q_{14}	.94		(2.64)	
q_{23}	1.32		(3.81)	
log likelihood value	-22068.5			
Number of individuals	5010			

Notes: Standard errors in parentheses. The superindex * denotes the significance at the 5 percent level (only for elements in β_i and λ_i (with $i = u, p$) and δ).

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 Eastern European, African or Asian, and type of occupation = manufacturing. Log age and log hourly wage in the previous employment are measured as the 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, which is why we normalize $\lambda_{u1} = \lambda_{p1} = 0$ and 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 by slightly more than 100 percent, which means that it more than doubles. The effect on the mean or median unemployment duration depends on the moment when training occurs. If the training is given within the first month, 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. (Naturally, 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 behalf 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 and it is higher for individuals with two years of high school and university graduates. It is also higher for UI recipients, thereby 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 is insignificant. The interpretation of the effects of the calendar year is complicated by the fact that some vocational courses have been recorded 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. Further, 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, this rate is higher which 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. Then, it becomes 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 purpose.

Table 3. Estimation results for the training effect on the transition rate to work when imposing the absence of unobserved heterogeneity

	To work, θ_u	
AMU effect		
δ	.57	(.05)*
Individual characteristics		
log(age)	-.50	(.07)*
short senior high school	.19	(.04)*
senior high school	.10	(.05)*
short tertiary education	.18	(.07)*
university	.38	(.05)*
female	.08	(.04)*
UI recipient	.65	(.05)*
cash allowance recipient	.61	(.07)*
from Eastern Europe	-.15	(.09)*
from Africa or Asia	-.39	(.09)*
log(hourly wage)	.08	(.02)*
experience in occupation (dummy)	.44	(.04)*
education in occupation (dummy)	.35	(.04)*
professional and technical work	.10	(.06)*
health, nursing and social work	.25	(.06)*
adm., managerial and clerical work	-.18	(.06)*
sales	-.01	(.06)
agriculture and mining	.22	(.06)*
services (incl. non categorized occ.)	.19	(.05)*
large city (dummy)	.04	(.03)
needs guidance (dummy)	-.20	(.07)*
willing to move (dummy)	.13	(.04)*
accepts part time work (dummy)	.23	(.06)*
1994	.52	(.05)*
1995	.47	(.05)*
1996	.51	(.06)*
1997	.70	(.06)*
1998	.71	(.06)*
1999	.91	(.06)*
2000	.96	(.10)*

Table 3. Continued….

	To work, θ_u	
Duration dependence		
λ_{i1}	-7.46	(.05) *
λ_{i2}	-7.42	(.06) *
λ_{i3}	-7.56	(.06) *
λ_{i4}	-7.58	(.07) *
λ_{i5}	-7.72	(.07) *
λ_{i6}	-8.00	(.09) *
λ_{i7}	-8.02	(.10) *
λ_{i8}	-7.99	(.11) *
log likelihood value	-17413.1	
Number of individuals	5010	

Notes: Standard errors in parentheses. The superindex * denotes the significance at the 5 percent level (only for elements in β_i and λ_i (with $i = u, p$) and δ).

As a first informal check on the robustness of the estimates, we compare them to those obtained from the misspecified model where 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 shown in Table 3. 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 3. Typically, when unobserved heterogeneity is ignored in the 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).

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 time. (Naturally, 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 1.2, there are reasons for suspecting that the effect is smaller if this time elapsed is long. Further, 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, the model could be extended 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 4 gives the estimates for δ_1 and δ_2. Clearly, the training effect is very large right after the training participation period. The individual is three times as likely to move to employment within a month after AMU training, as compared to if he had not participated in the training. After the first month, the effect is still positive, but it is much smaller.

Table 4. Estimation results for the training effect on the transition rate to work when this is allowed to depend on the time elapsed since training

	To work, θ_u	
AMU effect		
δ_1 (i.e. $\leq$ 28 days)	1.24	(.13) *
δ_2 (i.e. > 28 days)	.29	(.13) *
log likelihood value	-22057.8	
Number of individuals	5010	

Notes: Standard errors in parentheses. The superindex * denotes the significance at the 5 percent level.

For the 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 $\delta(t - t_p)$ has now also become a source of negative duration dependence. There is also a slight change in the estimates of the unobserved heterogeneity distribution.

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 4).

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. Incidentally, note 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 the 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 of finding 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 great deal from the course find a job quickly, and those who do not benefit remain unemployed longer. The heterogeneity may be due to heterogeneity of individual characteristics or heterogeneity of the characteristics of the training course. The results suggest that human capital accumulation in 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 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 a short tertiary or university education.

The main results are presented in Table 5. 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. Once more, 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 of 6.4. Under the null hypothesis, this test statistic has a chi-square distribution with three degrees of freedom, which leads to the acceptance of this null hypothesis at the 5 percent (though not at the 10 percent) level.

Table 5. Estimation results for the training effect on the transition rate to work when this is allowed to depend on gender, educational level and nationality

	To work, θ_u	
AMU effect		
main effect	.94	(.13) *
i.e. woman	-.00	(.15)
i.e. high education	-.40	(.20) *
Immigrant from Africa, Asia or Eastern Europe	.11	(.26)
log likelihood value	-22065.3	
Number of individuals	5010	

Notes: Standard errors in parentheses. The superindex * denotes the significance at the 5 percent level.

Time in training and in other programs

We now replace the rule that the length of the time intervals spent within AMU and other programs are set to zero by the rule that the timing keeps running during such periods. For T_u, this is more appropriate if individuals move into employment at the same rate within such periods as 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 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 and the effect after participation, where the latter was captured by parameter δ the basic model. 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.

The parameter estimates are reported in Table 6. Most are similar to those in Table 2. The estimate of the parameter of interest, Δ, is virtually equal to zero, however, and it is a compromise between the very low transition rate from AMU to work in the first months and the very high transition rate from AMU to work after that.

Table 6. Estimation results for the training effect on the transition rate to work when real time in programs is included and the training effect works from the moment of entering (instead of leaving) AMU

	To work, θ_u		To AMU training, θ_p	
AMU effect				
Δ	.11	(.11)		
Individual characteristics				
log(age)	-.71	(.11)*	.59	(.30)*
short senior high school	.04	(.06)	.18	(.16)
senior high school	-.08	(.07)	.04	(.18)
short tertiary education	.11	(.12)	-.12	(.30)
university	.22	(.09)*	-.14	(.22)
female	.03	(.06)	-.06	(.14)
UI recipient	.22	(.07)*	.02	(.19)
cash allowance recipient	.21	(.11)*	.34	(.23)
from Eastern Europe	-.75	(.15)*	.39	(.26)
from Africa or Asia	-.95	(.15)*	.19	(.30)
log(hourly wage)	-.01	(.04)	.17	(.09)
experience in occupation (dummy)	.20	(.06)*	.08	(.14)
education in occupation (dummy)	.22	(.05)*	.10	(.14)
professional and technical work	-.11	(.09)	.03	(.22)
health, nursing and social work	.07	(.09)	-.34	(.25)
adm., managerial and clerical work	-.38	(.09)*	.32	(.23)
sales	-.26	(.09)*	.01	(.22)
agriculture and mining	-.06	(.10)	.09	(.24)
services (incl. non categorized occ.)	-.17	(.08)*	-.09	(.20)
large city (dummy)	-.07	(.05)	-.26	(.12)*
needs guidance (dummy)	-.54	(.11)*	.51	(.20)*
willing to move (dummy)	.02	(.07)	.18	(.16)
accepts part time work (dummy)	.16	(.10)*	-.46	(.33)
1994	.22	(.07)*	-.15	(.18)
1995	.17	(.07)*	-.17	(.18)
1996	.21	(.08)*	-.63	(.22)*
1997	.46	(.08)*	-.63	(.22)*
1998	.52	(.08)*	-.59	(.22)*
1999	.81	(.09)*	-1.39	(.33)*
2000	.84	(.16)*	-1.03	(.69)*

Table 6. Continued….

	To work, θ_u		To AMU training, θ_p	
Duration dependence				
λ_{i2}	.08	(.05)	-.62	(.15) *
λ_{i3}	-.05	(.07)	-.72	(.17) *
λ_{i4}	.03	(.08)	-.70	(.19) *
λ_{i5}	.11	(.09)	-.61	(.20) *
λ_{i6}	-.11	(.11)	-5.52	(.19) *
λ_{i7}	-.06	(.12)		
λ_{i8}	-.19	(.09) *		
Unobserved heterogeneity				
log (v_1)	-5.74	(.12)		
log (v_2)	-7.09	(.14)		
log (v_3)			-8.50	(.97)
log (v_4)			-6.76	(.40)
q_{13}	-.01		(.50)	
q_{14}	2.20		(2.06)	
q_{23}	3.63		(11.81)	
log likelihood value	-22781.8			
Number of individuals	5010			

Notes: Standard errors in parentheses. The superindex * denotes the significance at the 5 percent level (only for elements in β_i and λ_i (with $i = u, p$) and Δ).

In this extension, T_u is real time spent out of work (i.e. in unemployment and in training programs and other active labor market programs) since the entry into unemployment. The treatment effect Δ also summarizes the effect of entering AMU on the total period of unemployment. 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 which reflects the fact that individuals rarely move from other programs directly into AMU. The fact that the β_i parameter estimates remain unchanged means that they are insensitive to whether we include time spent in other programs. Naturally, other programs may have their own causal effect on θ_u, but an analysis of this raises new selection problems and would be beyond the scope of this paper.

5. Conclusions

After participation in 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 he 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, 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 this policy repairs market failures or reduces the variation in individual outcomes on an aggregate level, and whether this would makes the policy socially effective. A comprehensive cost-benefits analysis must take the effects on post-unemployment wages and the duration of subsequent employment into account, but the available evidence suggests these effects to be insignificant (see earlier references and Korpi, 1994).

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

Second, in reality, participation in other programs may be affected by unobserved determinants related to the unobserved determinants of entry into AMU training and exit to work. In addition, one cannot typically 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, our results under-estimate the effect of AMU training. The analysis should then include participation in other programs, which would naturally increase the complexity of the model and the burden of estimation.

References

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

Abbring, J.H., van den Berg, G.J. and van Ours, J.C. (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, Prora 2001:2, National Labour Market Board.

Björklund, A. (1993), The Swedish experience, in Measuring Labour Market Measures, Ministry of Labour, Denmark, Copenhagen.

Bonnal, L., Fougère, D. and Sérandon, A. (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. and Richardson, K. (2001), The relative efficiency of labor market programs: Swedish experience from the 1990s, IFAU Working Paper 2001:2, Uppsala.

Carling, K., Holmlund, B. and Vejsiu, A. (2001), Do benefit cuts boost job finding? Swedish evidence from the 1990s, Economic Journal 111, 766-790.

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

Dahlberg, M. and Forslund, A. (1999), Direct displacement effects of labour market programmes: The case of Sweden, IFAU Working Paper 1999:7, Uppsala.

Eberwein, C., Ham, J.C. and LaLonde, R.J. (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. and Holmlund, B. (1991), Unemployment, vacancies and labour market programmes: Swedish evidence, in F. Padoa Schioppa (ed.), Mismatch and Labour Mobility, Cambridge University Press, Cambridge.

Edin, P.A., Harkman, A., Holmlund, B. and Söderberg, H. (1998), Escaping long-term unemployment in Sweden, Mimeo, Department of Economics, Uppsala University, Uppsala.

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 Johansson, A. (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, IFAU Working Paper 2000:1, Uppsala.

Regnér, H. (1997), Training at the job and training for a new job: Two Swedish studies, Dissertation series 29, Swedish Institute for Social Research, Stockholm University.

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

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 (eds.), Handbook of Econometrics, Volume V, North-Holland, Amsterdam.

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

Zettermark, Å., Regnér, H., Aspgren, M. and Ericson, J. (2000), Efficient training programmes for the unemployed, Ministry of Industry, Ministry Publication Series 2000:38, 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

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

It is easily shown that v_u and v_p are independent if and only if $\text{cov}(v_u,v_p) = 0$.

In the estimation procedure, we actually estimate the transformed probabilities $q_{ij,}$ implicitly defined as logistic versions of p_i:

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

Because 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 that estimating the q_{ij} instead of the p_{ij} has the advantage that no boundary restrictions must be imposed on the parameter space. Moreover, conditional on $v_1 \neq v_2$ and $v_3 \neq v_4$, it holds that $\text{corr}(v_u,v_p) = 0$ if and only if $q_{23} = q_{13}-q_{14}$.