MEMORANDUM

No 24/2008

Norwegian Vocational Rehabilitation Programs: Improving Employability and Preventing Disability?

Lars Westlie

ISSN: 0809-8786

Department of Economics
University of Oslo

This series is published by the
University of Oslo
Department of Economics

P. O.Box 1095 Blindern
N-0317 OSLO Norway
Telephone: + 47 22855127
Fax: + 47 22855035
Internet: http://www.oekonomi.uio.no
e-mail: econdep@econ.uio.no

In co-operation with
The Frisch Centre for Economic Research

Gaustadalleén 21
N-0371 OSLO Norway
Telephone: +47 22 95 88 20
Fax: +47 22 95 88 25
Internet: http://www.frisch.uio.no
e-mail: frisch@frisch.uio.no

Last 10 Memoranda

No 23/08	Geir B. Asheim and Bjart Holtsmark *Pareto Efficient Climate Agreements Can Always Be Renegotiation-proof*
No 22/08	Simen Gaure, Knut Røed and Lars Westlie *The Impacts of Lavor Market Policies on Job Search Behavior and Post-Unemployment Job Qualilty*
No 21/08	Snorre Kverndokk and Adam Rose *Equity and Justice in Global Warming Policy*
No 20/08	Halvor Mehlum *Samfunnsøkonomen Schweigaard*
No 19/08	Dag S. Holen *It Ain't Necessarily So*
No 18/08	Dag S. Holen *Pension Incentives, Labor Supply and Heterogeneous Pension Systems*
No 17/08	Dag S. Holen *Disability Pension Motivated Income Adjustment*
No 16/08	Florian K. Diekert, Dag Ø. Hjermann, Eric Nævdal and Nils Chr. Stenseth *Optimal Age- and Gear-Specific Harvesting Policies for North-East Arctic Cod*
No 15/08	Karine Nyborg *I don't Want to Hear About it: Rational Ignorance among Duty-Oriented Consumers*
No 14/08	Fedor Iskhakov *Pension reform in Norway: evidence from a structural dynamic model*

A complete list of this memo-series is available in a PDF® format at:
http://www.oekonomi.uio.no/memo/

Norwegian Vocational Rehabilitation Programs:

Improving Employability and Preventing Disability?

Lars Westlie*
Ragnar Frisch Centre for Economic Research

Abstract

This paper investigates the effects of five different vocational rehabilitation (VR) programs on the hazard rates into e*mployment, disability* and *temporarily withdrawals from the labor market* for persons who face severe problems in re-entering the labor market, mostly due to medical problems. One of the main findings is that re-education into a new profession is an effective way to improve employability and prevent disability. Work training produces varying results and is more effective the more it resembles a real job. All programs, and in particular re-education, comes with a cost of increased VR duration. Finally, those with the worst initial employment prospects are the ones who benefit most from participation.

Keywords: Vocational rehabilitation, program evaluation, disability, heterogeneous treatment effects, multivariate hazards
JEL classification: C14, C15, C41, I21, J24, J64

* This research has received financial support from the Norwegian Ministry of Labor and Social Inclusion and Ministry of Finance through the "Strategic Institute Program on Labor Market and Pension Research". I would like to thank Erik Biørn, Bernt Bratsberg, Tyra Ekhaugen, Morten Henningsen, Simen Markussen, Knut Røed and Oddbjørn Raaum for their valuable input. Correspondence to: Lars Westlie, The Ragnar Frisch Centre for Economic Research, Gaustadalléen 21, 0349 Oslo, Norway. E-mail: lars.westlie@frisch.uio.no.

1. Introduction

In Norway, as in many other European countries, the number of disability pension recipients has increased rapidly over the past decades. This has serious consequences, not only for the individuals concerned, who often suffer large income reductions and are excluded from an important social arena, but also for the overall national economy due to the loss of a valuable contribution to the labor force. In an attempt to reduce the inflow to disability benefit status, the government has increased its funding of the Norwegian Vocational Rehabilitation (VR) program. The VR program is designed to improve the employability of persons who face a risk of permanent withdrawal from the labor market. Depending on the needs of the participants this includes increasing general skills as well as learning a new profession.

Some studies have contributed to our understanding of how these programs affect the employment probability. Aakvik (2001); (2003) and Aakvik et al. (2005) estimate how VR programs affect the employment probability up to five years after the participants applied to the program. The main finding in these papers is that the least employable participants are the ones who benefit most from participating in terms of increased employment probability. In addition, they also report that these persons are least likely to be enrolled in a program, indicating that the effectiveness of the VR sector can be improved by reversing the selection rule into programs. The VR sector provides a variety of different training programs. In a large meta-analysis of ordinary labor market programs, Kluve (2006) shows that there is a wide range of program effects among different types of programs. An additional explanation of the pattern reported by the different Aakvik papers may be that the programs provided to the least employable individuals are the most effective ones, making this a question of program composition as well.

Based on Swedish data, Frölich et al. (2004) estimate how six different types of programs affect the employment probability for persons with histories of long-term sickness, three years after program application. Their findings are in line with the Norwegian studies, showing that program effects are either non-existent or negative, particularly as regards educational programs. However, since educational programs tend to be long-lasting, a restrictive time window may influence the estimated effects of these

programs more than others. The Swedish study concludes with a call for further research based on "more sophisticated nonparametric regression methods" due to the statistically insignificant though non-negligible treatment effects.

This paper concentrates on three main questions: How the different VR programs affect the probability of employment and disability, how they affect the total time spent in the VR system, and how these effects vary according to observable individual characteristics? In order to address these questions, individual spells from several register-datasets are constructed, containing detailed information about welfare transfers and labor market status over a ten-year period in addition to a number of individual characteristics. A well-known problem in the treatment literature is separating the causal treatment effect from the spurious correlation that may originate from unobserved heterogeneity across participants and non-participants (see Heckman et al. (1999) for a comprehensive introduction to these problems). I will attempt to overcome this problem by using longitudinal and cross-sectional variation in treatment capacity (the supply and demand of VR programs) as an exogenous source of variation in participation propensity.

The existing literature focuses mainly on whether or not program participants are employed at a given point in time after VR entrance, treating all non-employed similarly. Even though a return to employment is the main objective of these programs, it could be interesting to study the composition of the group that remains non-employed. The non-employed group will typically consist of some disability pension recipients (who have more or less permanently left the labor force), and some who may still be looking for work.

In order to see how programs affect the time spent in the VR system, I will split the treatment effect into an *on-program effect* and an *after-program effect*, see Røed and Raaum (2006). The former captures how programs affect search behavior during participation (the lock-in effect), whereas the latter will capture the effect on employability after program completion. Based on these effects I find that the VR programs increase the employment probability for an average VR client by 8.4 percentage points. However, this effect varies greatly by program type and participant characteristics. In addition, VR programs increase the time spent in the VR regime by 7.4 months.

An oft-mentioned policy proposal has been to transfer persons into the VR system at an early stage of the sickness recovery process. This group of participants has had a higher employment rate and lower disability rate than participants with a longer sickness history. Although this paper does not directly estimate the effect of an early program start, it investigates whether or not the early starters experience larger program effects than others. The findings indicate that this higher employment rate is due to individual characteristics rather than treatment effect. In fact, the paper shows that individuals who enter the VR system at an early stage experience a higher participation probability, higher employment probability and lower treatment effect than other participants. These results are in line with the main results of Aakvik et al. (2005).

2. The Norwegian vocational rehabilitation system

VR programs are established to help individuals who experience severe difficulties in (re-)entering the labor force. While most problems are of a medical nature, a substantial share of program participants has no medical diagnosis, but participates due to problems of a more social nature. This paper distinguishes between three groups of participants: *long-term ill*, *short-term ill* and the *previously unemployed*.

Most of the participants have some kind of health problem or injury that made it impossible to continue in their previous job. In Norway all employees are entitled to 12 + 12 months on sick leave benefits[1]. Some may even get extended periods if further medical treatment seems necessary to restore their work capacity. Those who become healthy enough to re-enter the labor market, but remain unable to take up their former job, may apply to the local labor market office for a place in a VR program. As a result, most participants here labeled *long-term ill*, have quite long period of inactivity before they start on a program. This group of participants is also the main target group of the VR program. The second group of participants is labeled *short-term ill*, and consists of persons who enter the VR system before the end of the first 12-months sickness period.

[1] The benefits are conditional on a medical certificate issued by a physician. During the first 12 months recipients are ensured full pay (workers in the private sector without collective agreements have an upper limit of 400.000 NOK, 1 Euro ≈ 8 NOK). In the second 12-month spell, called medical rehabilitation, payment is reduced to around 64 percent of full pay.

More often than not, they have lost the ability to do their previous job, but are still able to perform in other professions. The third group has no previous sickness history, but consists of *unemployed* persons for whom VR programs are considered to be a better alternative than ordinary labor market programs, e.g. persons with learning difficulties, behavioral problems or drug addictions. Released prisoners are also included herein.

When admitted into the VR system, they are assigned a caseworker. This stage is what this paper refers to as *Declared for Program* (DP) and consists of eligible potential participants who are waiting for the appropriate program. Persons in the DP state may also conduct job search activities and apply for a disability pension.

Participants may attend five types of programs; *Work Training in Ordinary firms* (WTO), *Work Training in Protected firms* (WTP), *education provided by the local employment service* (AMO), *Public Education* (EDU) and *Wage Subsidies* (WS). This categorization is in line with previous studies of labor market programs (see for instance Kluve (2006) and also in accordance with the suggestion of Aakvik et al. 2005[2]. Figure 1 sum up the Norwegian VR system.

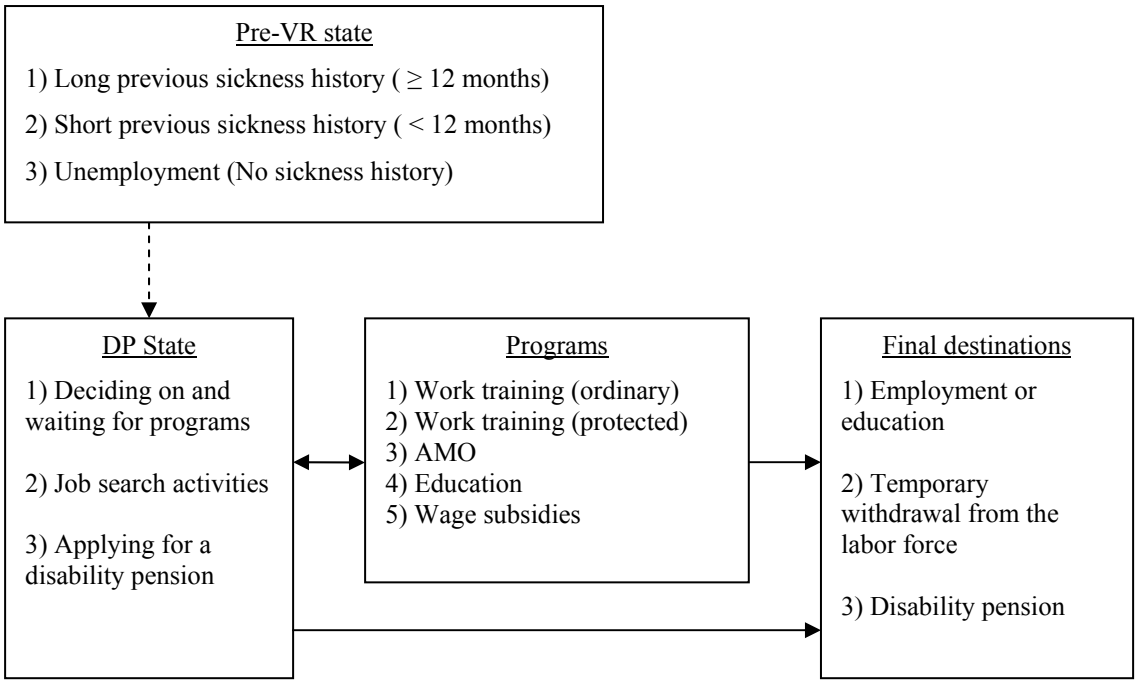
In WTO, participants work in ordinary firms performing regular tasks under some kind of supervision. The firm has to be approved in advance by the local labor market office and has no influence on which participants are directed to them by the caseworker. Even so, we might suspect that the caseworker will seek to maintain a good relationship with the firm and hence be reluctant to allocate people lacking skills or motivation. Participants receive rehabilitation benefits that are approximately 64 percent of their previous labor market income. These firms may be either public or private. WTO program participation is limited to three years.

In WTP, participants work in firms established specifically to provide people with extraordinary needs with work training combined with education and improved social abilities. The work is done under close supervision. The maximum duration is approximately two years. For more on the largest group of WTP firms, see Aakvik and Dahl (2006).

[2] Aakvik et al. (2005) recommends using a three-program structure, with wage subsidies, education and on-the-job training. However, this paper splits the last two program groups into two due to the large within-group difference among the programs.

EDU and AMO provide different forms of classroom training. EDU programs include all kinds of ordinary public and private schooling while AMO programs are courses provided by the local employment service (see Raaum and Torp (2002) for more on AMO courses). While AMO courses have a total upper limit of ten months, there was no maximum duration for ordinary education until 2002, when a limit of three years was introduced.

Figure 1. The basic structure of the Norwegian vocational rehabilitation system.



In WS, participants work for a regular firm, but the employment office finances part of the salary (up to 60%). The fact that the employer pays at all illustrates that participants are expected to be somewhat more productive than WTO participants, though these programs may be rather similar in content. Also, in WS the employer is not obliged to accept any given participant. One of the intentions of the WS program is to enable participants to continue working for the firm after the funding ends. This is less common in the WTO program. The program has a maximum duration of two years[3].

[3] Reduced to 18 months from January 2002.

3. Data and descriptive statistics

The data are constructed from a rich set of Norwegian administrative data, containing detailed information about the labor market status of each inhabitant at the end of each month in addition to several individual characteristics. From these registers, we select all "fresh"[4] VR entrances between January 1994 and September 2003 for persons below 56 years of age. Then, individual spells are constructed, which contains information about the current state (i.e. the five programs and the DP state) on a monthly basis. Spells that do not meet the *fresh* requirement are removed. This is done to ensure that all spells are recorded from the actual beginning.

The spell ends when a person starts receiving a disability pension or drops out of all relevant[5] public registers for three consecutive months. For the latter group, we observe who has landed a job or begun an education that is not a part of the VR program during these three months. These persons will be referred to as *employed* in the rest of the paper. Those who find themselves outside all these states, i.e. any relevant public registers, employment, education or disability, are treated as *temporary withdrawals from the labor market*. These three final destinations are the ones referred to in figure 1.

All ongoing spells at the end of the time window (September 2003) are treated as censored. The same goes for people who die or migrate and for women who give birth. Spells containing more than three programs will also be treated as censored, as it may be difficult to allocate the right treatment effect to the right program for such a small group[6]. Finally, persons starting on programs where ordinary employment is no longer regarded as the ultimate goal, are also treated as censored.

The data consists of 177,353 spells. From table 1, these spells have a mean duration of 19.5 months. Spells that actually contain a program (70.4 percent do) have an average duration of 26.4 months. The employment frequency is roughly equal to the

---

[4] By fresh we mean people who enter the VR system registered as DP at the end of a given month and who have no VR record for the previous twelve months.

[5] The relevant registers include all health related payments such as sickness- and medical rehabilitation benefits and the unemployment register.

[6] As a consequence of the three-program limit, long-lasting spells may be very selective. Therefore all spells with a duration of more than seven years are treated as censored.

temporary withdrawal and the disability frequencies put together, at around 30 percent. The remaining spells (around 40 percent) are censored, mostly due to the end of the time window.

Table 1.
Descriptive VR spell statistics

Number of spells	177,353
Mean duration of all spells (months)	19.5
Mean duration of spells with programs	26.4
Mean duration of spells without programs	9.7
Share of spells containing programs (percent)	70.4
Share of spells containing multiple programs	31.2
Share of spells ending in (percent)	
Employment (work or education)	28.3
Temporary withdrawals	17.0
Disability	13.9
Censored (due to time window)	30.5
Censored (other reasons)	10.3

Table 2 contains descriptive statistics for each of the five programs. We see that EDU is the most common program, included in 40.5 percent of all spells, followed by WTO with a 31.7 percent share. WS, on the other hand, is only included in 6.9 percent of spells. The duration of the programs varies greatly. EDU is not only the most common one but also the one with longest average duration at 15.8 months. The other programs have average durations of around 8 to 10 months, except for AMO which has an average duration of 4.7 months.

Table 2
Descriptive VR program statistics

Variables	All programs	WTO	WTP	AMO	EDU	WS
Percent of spells containing this program	72.4	31.7	13.9	12.3	40.5	6.9
Average program duration (months)	10.91	7.4	10.3	4.7	15.8	8.2
Average waiting time before program entry (months)	4.78	4.9	4.8	4.7	4.8	4.5
Percent of spells combining the following programs*						
WTO	62.4	-	12.8	13.4	40.0	7.8
WTP	45.2	28.5	-	8.3	15.6	3.4
AMO	75,5	36.0	10.1	-	46.7	5.1
EDU	44.0	30.2	5.3	13.1	-	3.6
WS	69.0	43.7	8.4	10.6	26.5	-

* Note that the sum of participation frequencies by program exceeds the overall participation frequencies due to multiple programs. Only spells containing at least one program are considered.

The average waiting time before entering a program is quite similar across the different programs, at around four and a half months. Conditional on participation in at least one program, around 65 percent have several programs within their spell. For instance, three out of four AMO participants also took part in at least one other program. WTO and WS are also often combined with other programs (62.4 and 69 percent respectively), while EDU and WTP more frequently stand alone (only 44 and 45.2 percent in combination with others). The next columns show to which extent the different programs are used in combinations with each other. For instance, 40 percent of WTO participants also participated in EDU. While only 7.8 percent of WTO participants also participated in WS, this group consists of 43.7 percent of all WS participants.

Table 3 reports descriptive VR client statistics. The first column describes the marginal distribution of some selected characteristics, i.e. the share of spells with the corresponding characteristic. We see that 63.9 percent of those entering the VR regime are suffering long-term illnesses, corresponding well with this being the main target group for VR. 25.5 percent of the individuals have short-term illness experiences prior to VR entry, while 10.6 percent arrive from unemployment. Next, 30.4 percent are below 30 years of age, while almost half of all entrants are between 30 and 44 years old. The most common educational level is vocational track high school. Finally, gender is more or less equally distributed among clients.

In the subsequent columns in table 3, participation frequencies are reported, both overall and conditioned on each of the five programs. Persons with short-term illnesses have the highest participation frequency at 77.8 percent. Next, persons with long-term illnesses have a participation frequency of 71.7 percent while only 63.3 percent of those unemployed participate in one of the programs. There are also large differences between the different programs. There is not much difference in WTO and EDU participation among those with long-term illnesses, unlike those with short-term illnesses who enter the EDU program more often. WTP is the most frequently used program by the unemployed. The participation frequency differs to some extent by age. While 77.3 percent of those below 30 years of age participate in at least one program, the corresponding frequency is 62.1 percent for those who are older than 44. This difference is mainly due to a different participation pattern in the EDU program. Educational background seems to have little impact on the participation frequency, the only exception

being persons with only compulsory education, who have a low EDU participation frequency.

Table 3
Descriptive VR client statistics (All numbers are in percent)

Characteristics	Distribution of individual characteristics	Participation frequency	Participation frequency by program*				
			WTO	WTP	AMO	EDU	WS
Pre-VR state **							
Long–term ill	63.9	71.7	34.3	13.0	12.4	39.3	6.5
Short-term ill	25.5	77.8	30.8	7.5	13.8	52.9	8.3
Unemployment	10.6	63.3	18.1	34.6	8.3	17.8	5.9
Age group							
Age below 30	30.4	77.3	30.2	16.1	11.8	49.2	6.2
Age between 30 and 44	46.8	74.2	32.3	12.3	13.5	43.6	7.3
Age between 45 and 55	22.8	62.1	32.4	14.2	10.5	22.6	6.9
Previous education							
Compulsory school	20.5	65.0	31.7	17.0	12.2	26.3	6.7
High school – General studies	11.2	73.2	30.7	14.1	11.3	43.8	5.3
High school – Vocational track	59.9	74.1	32.2	13.9	13.2	42.6	7.3
Higher education	8.5	77.3	29.5	6.3	8.22	55.8	6.2
Little previous work experience***	23.4	67.5	31.8	18.2	11.7	30.9	4.9
Men	51.1	70.9	28.4	15.7	12.6	38.0	8.9
Women	48.9	74.0	35.2	12.0	12.1	43.1	4.7

* Note that the sum of participation frequencies by program exceeds the overall participation frequencies due to multiple programs.
** See previous section for more about the pre-VR state.
*** This is defined as having less previous work experience than the 25th percentile compared to VR candidates of the same age.

Table 4 reports the share of non-censored spells that ends in employment. While the main employment rate is 47.1 percent, this number differs considerably across program groups conditioned on the last program in the spell. As many as 71.1 percent of those who have WS as their final program return to employment, while the corresponding number for EDU is 62.8. In contrast, spells ending with WTP have an employment rate of only 31.5 percent. Non-participants and WTO participants have a quite similar employment rate, at around 40 percent, while AMO scores somewhat higher at 48.5 percent. Persons with short-term illnesses have the highest employment rate compared to the other pre-VR states. The employment rate also diminishes by age and increases with the level of previous education. Finally, men have a higher employment frequency than women.

Table 4.
Employment frequency of non-censored spells conditioned on the last program (percent)

			Conditioning on the last program				
	All Spells	Non-participants	WTO	WTP	AMO	EDU	WS
All spells	47.1	38.3	39.6	31.5	48.5	62.8	71.1
Pre-VR state*							
Long-term ill	39.1	27.6	34.5	18.5	40.2	56.8	65.6
Short-term ill	65.1	63.0	54.2	40.0	58.2	73.8	79.4
Unemployed	48.6	41.2	42.7	48.2	64.7	62.6	74.9
Age group							
Age below 30	55.7	51.7	44.7	40.8	56.6	66.0	75.8
Age between 30 and 44	49.7	41.6	42.7	31.1	49.1	63.2	72.5
Age between 45 and 55	32.1	24.5	30.3	19.0	35.5	51.8	62.4
Type of previous education							
Only compulsory	37.3	28.4	33.7	26.5	47.1	54.4	69.2
High school - General studies	45.2	36.9	39.3	30.4	45.8	59.2	68.2
High school – Vocational track	49.6	42.0	41.3	34.0	49.5	63.7	72.0
Higher education	58.0	51.4	44.5	33.5	47.7	70.5	71.4
Little previous work experience**	32.8	23.2	29.5	24.9	42.4	49.1	61.5
Men	51.2	41.8	43.1	37.7	53.5	66.1	75.2
Women	42.5	33.9	37.0	22.7	42.8	59.4	63.3

* See previous section for more about the pre-VR state.
** This is defined as having less previous work experience than the 25th percentile compared to VR candidates of the same age.

There are (at least) three explanations to the large variation in employment frequencies across VR programs as seen from table 4. First, as we remember from table 3, participants in the different programs are quite different as regards characteristics that are highly correlated with employment probability, and so we would expect to observe great differences in the outcomes between these groups, even if the programs had no effect. That is, people are selected (or select themselves) into the different programs based on their employment prospects. Second, different programs may have different effects on employment probability. Third, different participants may experience different effects from the different programs (i.e. heterogeneous treatment effects). This is about as far as an inspection of frequency distribution and summary statistics can bring us. In the following sections we go one step further and analyze the data at hand by an econometric model, attempting to separate the causal treatment effects from any spurious correlation originating from selection processes.

4. The econometric model

4.1. Model setup

This section presents a formal model, explaining the transitions into the five different programs and the three final destinations, *employment*, *disability* and *temporary withdrawal*. More specifically, we use a multivariate mixed proportional hazard rate model (MMPH) with eight competing events. As we only observe labor market status at the end of each month, the econometric model is set up in terms of grouped hazard rates (Prentice and Gloeckler (1978); Meyer (1990)). The effect of all time-varying covariates, including calendar time and spell duration, is assumed to be constant within each month.

Equation (1), *the participation equation*, and equation (2), *the outcome equation*, explain the monthly integrated hazard rates into each of the five different programs (k=1,…,5) and each of the three final destinations (k=6,7,8) respectively, during month *t* for individual *i*:

$$(1) \qquad \varphi_{kit} = \exp\left(\mu_k o_{ijt} + \pi_k a_{ijt} + \sigma_k s_{it} + \nu_k r_{it} + \beta_k x_{it} + \alpha_k z_{it} + v_{ki}\right), \quad k = 1,...,5$$

$$(2) \qquad \varphi_{kit} = \exp\left(\mu_k\left(x_{it}^o\right) o_{ijt} + \pi_k\left(x_{it}^o, r_{it}\right) a_{ijt} + \sigma_k s_{it} + \lambda_k d_{it} + \beta_k x_{it} + v_{ki}\right), \quad k = 6,7,8$$

The explanatory variables, *o,a,s,r,d,x,z* and *v* are described in table 5 and further below. A more detailed description of each variable is also available at www.frisch.uio.no/docs/VR_prog.html. j denotes program type (j=WTO,WTP,AMO,EDU,WS). All explanatory variables are measured at the beginning of each month.

Programs may affect the hazard rates in two ways, i.e. while in progress (the on-program effect, *o*) and after their completion (the after-program effect, *a*). The on-program effect may be regarded as a lock-in effect since participants may have less time for job search activities. The after-program effect may reflect the increased human capital that participants are meant to gain from the program or a signaling effect. In order to allow the program-effects towards the final destinations to differ according to observed characteristics, the treatment parameters in the outcome equation (2), will be functions of some of the other explanatory variables in the model. These are gender, age, education, previous welfare history, previous work experience, medical diagnosis, current local labor market tightness and previous program experience (within the spell). The after-

program effect is also a function of time since program completion (r). This may be regarded as a depreciation rate of the program.

Table 5
Overview of explanatory variables

Explanatory variable	Description
o	10 dummies indicating current treatment status
a	10 dummies indication previous treatment status
s	22 dummies indicating calendar year and month
r	12 dummies indicating consecutive months in the DP state. In equation (2) r is assumed to be a linear function.
x	Age, previous labor market history, family status, gender, education, country of birth, previous social security history, medical diagnosis.
x^o	Gender, age, education, pre-VR state, previous work experience, medical diagnosis, current labor market tightness and previous program experience.
z	Instruments (Variables that only affect program transitions, k=1,…,5)
v	Unobserved heterogeneity components
d	36 dummies indicating spell duration
μ() and π()	Linear functions describing treatment effect

Time has two dimensions in this model; calendar time and process time. The *calendar time* dimension reflects business cycle and seasonal fluctuations in addition to government regulations and priorities towards the VR sector. In the model, the effects of calendar time are represented by 22 dummy variables (10+12), one for each calendar year (1994 - 2003) and one for each calendar month. Local labor market conditions, measured as the observed transition rate from unemployment to employment in the local district, are also included. All calendar time variables are included in s_{it}. In the outcome equation, process time is defined as time since the spell started and is represented by 36 dummies (*d*). Process time may affect the different hazard rates through discouragement and statistical discrimination. In addition, even though there is no limit on the time a person is allowed to stay in the VR system, a long duration may indicate that the VR spell is nearing its end. In the participation equation, process time is measured as consecutive months in the DP state, i.e. time since the spell started or, if the subject already has program experience, since the last program was completed. In addition to previous program experience, this is assumed to cover the main factors related to process time in the participation probability.

The x-vector includes all individual characteristics. Previous labor market history is captured by previous work experience (i.e. number of years with labor market income above approximately 130,000 NOK), average income in these working years, the

previous work profession and the public disability payment which the person is entitled to. The effect of previous work experience is allowed to differ depending on age. Persons with less work experience than the 25th percentile within a five-years age group, is labeled as having *little work experience.* Family status is represented in the model by the age distribution of the children in the family (i.e. all possible combinations of the four age groups (0-3, 4-6, 7-12, 13-16)). In addition, the effects of these dummies are allowed to differ by the gender of the parent. For married persons we also include the labor market status of the spouse (working, receiving disability pension or staying at home) and the income of the working spouses. Educational attainment is included by 7 dummies. These are only compulsory education, four different high school degrees and two levels of higher education (1-2 and more than 3 years). Dummies for gender and the pre-VR state are also included in addition to 11 dummies describing the medical diagnoses. Emigrational status is included by eight dummies based on the number of years since the day of arrival (more or less than 7 years), gender and information about the country of birth (OECD or non-OECD).

4.2. Identification

Some of the explanatory variables in the model, especially process time (d,r) and program participation (o,a), are clearly endogenous in the sense that they are determined jointly with the final outcomes. People may enter the different programs based on their motivation, expected economic gain and previous experience. Some of these attributes may be uncovered in the data, and hence lead to selection problems and biased estimates. In order to estimate the causal program effects, I need to sort out the spurious correlation that originates from these selection mechanisms. While previous (labor market) experience may capture most of these disturbances for those ordinary unemployed (which is often assumed in the matching literature; see Heckman et al. (1999) for more about matching), the VR clients have an extra source of unobserved bias, namely the occurrence that caused the need of VR programs. This occurrence may be orthogonal to all other observed individual attributes as well as having large effects on the choice of program and the final outcome probabilities, and hence be an important source of unobserved selection bias. In a related study on Swedish data, Frölich et al. (2004) demonstrate the importance of including the subjective recommendations of physicians

and caseworkers regarding the subjects health status. This information is not included in the data at hand. In order to solve this problem, a set of (time-invariant) unobserved individual characteristics (v_{ki}) is included and allowed to be correlated across transitions, i.e. a multivariate heterogeneity distribution. For instance, a person with unobserved characteristics that are favorable towards both entering a specific program and employment will not erroneously cause a bias in the estimated program effect, since these characteristics will be captured by v.

The model is non-parametrically identified based on the *timing-of-events* results of Abbring and Van den Berg (2003). They prove that modeling the time to the final outcome(s) and time to program participation simultaneously in a multivariate proportional hazard rate model solves the selection problem. In addition, McCall (1994) and Brinch (2007) show that the occurrence of time-varying covariates strengthens the identification. As pointed out by Eberwein et al. (1997), time-varying variables naturally provide an exclusion restriction in the sense that past values of these variables affect the current transition probabilities only through the selection process. Particularly the local business cycle conditions and the calendar time dummies (reflecting governmental priorities). The model is thoroughly tested by Monte Carlo procedures in Gaure et al. (2007). They conclude that it is extremely reliable, and accurately separates the causal treatment effects from sorting effects.

The identification strategy is also strengthened by introducing a set of exclusion restrictions, i.e. instruments that will have an impact on the participation hazards but are assumed to have no direct effect on the three final outcomes conditioned on the other observed covariates. These variables are included in (z) and will induce exogenous variation in the treatment probability. Aakvik et al. (2005) use the *degree of rationing*, measured as the percentage of applicants in a local district who do not participate in a program, as their instrument. However, this instrument may raise two concerns. First, the actual share of participants may reflect the long-term equilibrium in the local region. If the participants in the different regions have different needs for programs (due to different employment prospects or health problems), the share of participants will be a function of the employment probability and hence not a valid instrument. Second, as pointed out by Van den Berg (2007), the VR candidates may act on knowledge about future realizations of the instrument. For instance, knowledge about a high *degree of*

rationing in the next months may result in a higher effort in job search activities today since it seems unlikely that they will enter into a program in the nearest future.

In order to avoid these types of problems, I will construct instruments based on regional shocks in the supply and demand of programs within each local labor market offices (there are around 200 regional offices in Norway). I will assume that the VR candidate could not anticipate or make adjustment for these shocks in advance. The first instrument is called *work pressure on the caseworker* and is defined as the relative change in the inflow of new potential participants in month *t* relative to the average inflow in the three previous months. The idea here is that in months with a relatively high inflow, the caseworker will be under more work pressure and thus have less time to help each potential participant. In addition, the local employment service will not be able to adjust the number of program slots on such short notice.

The last two instruments are called the *share of new training programs* and *share of new AMO courses*. The first is calculated as the rate of new available program slots in the three programs WS, WTO and WTP in the previous month relative to the number of non-participants within a region. The second is constructed in the same way, only with new AMO slots rather than new training slots. The reason for distinguishing between programs that are provided by workplaces and programs that are given in classrooms, is that the supply elasticities may differ. While it may be hard to rapidly increase the training sector capacity, as recruitment of new firms or increasing the number of positions in existing firms may take some time, the program administrator can always find a bigger classroom for the AMO courses. The idea of these last two instruments is to capture differences in the supply of programs. In months with many new program slots relative to the number of people waiting, we expect an increase in the transition rate into the program group concerned. Both these instruments may be correlated with the local unemployment rate and seasonal cycles. However, these factors are included in x and s, and should therefore not represent any problem. New slots in EDU are not included in any of these instruments since new slots in this sector would be hard to calculate. In addition, public education follows the calendar year so this pattern is picked up by the calendar time dummies. Section 7 and appendix A2 reports some tests regarding the validity of the exclusion restrictions.

4.3. The likelihood function and estimation

Before introducing the likelihood function, an expression for the period-specific transition probability is needed. The probability of individual *i* making a transition to state *k* during period *t* is equal to:

$$p_{kit} = \left(1 - \exp\left(-\sum_{k \in K_{it}} \varphi_{kit}\right)\right) \frac{\varphi_{kit}}{\sum_{k \in K_{it}} \varphi_{kit}}, \tag{3}$$

where K_{it} is the set of feasible transitions for individual *i* in period *t*[7]. Here, φ is the monthly hazard rate presented in equation (1) and (2). An example of non-feasible transitions is that people are not allowed to make a transition to program j while attending this very program. They may, however, make transitions to all the other programs as well as the three final destinations. Let y_{kit} be an outcome indicator variable, equal to 1 if the corresponding observation ended in a transition to state *k*, and zero otherwise, and let Y_i be the complete set of outcome indicators available for individual *i*. The contribution to the likelihood function formed by a particular individual, conditional on the vector of unobserved variables v_i can then be formulated as:

$$L_i(v_i) = \prod_{y_{kit} \in Y_i} \left[\prod_{k \in K_{it}} \left[\left[\left(1 - \exp\left(-\sum_{k \in K_{it}} \varphi_{kit}\right)\right) \frac{\varphi_{kit}}{\sum_{k \in K_{it}} \varphi_{kit}} \right]^{y_{kit}} \right] \times \left[\exp\left(-\sum_{k \in K_{it}} \varphi_{kit}\right) \right]^{\left(1 - \sum_{k \in K_{it}} y_{kit}\right)} \right] \tag{4}$$

Equation (4) depends on unobserved components, and can therefore not be included directly into the data likelihood. To disengage the unobserved heterogeneity v_i, the heterogeneity distribution is approximated in a nonparametric fashion by means of a discrete distribution, see Lindsay (1983). As recommended by Heckman and Singer (1984), the number of mass-points are chosen by adding new points until it is no longer possible to increase the likelihood function. Let Q be the (a priori unknown) number of support points in this distribution and let $\{v_m, q_m\}$, $m = 1, ..., M$, be the associated location

---

[7] This can be derived from the continuous time hazards - which, given the within-period constancy assumption, are equal to the unit-interval integrated hazards - as follows:

$$\int_{t-1}^{t} \varphi_{kit} \exp\left(-\sum_{k \in K} \int_{t-1}^{u} \varphi_{kit} ds\right) du = \int_{t-1}^{t} \left(\varphi_{kit} \exp\left(-\sum_{k \in K} (u - (t-1)) \varphi_{kit}\right)\right) du = \left(1 - \exp\left(-\sum_{k \in K} \varphi_{kit}\right)\right) \frac{\varphi_{kit}}{\sum_{k} \varphi_{kit}}.$$

vectors and their probabilities. Expressed in terms of observed variables, the likelihood function is then given as

$$L = \prod_{i=1}^{N} E\left[L_i\left(v_i\right)\right] = \prod_{i=1}^{N}\sum_{m=1}^{M} q_m L_i\left(v_m\right), \quad \sum_{m=1}^{M} q_m = 1 \tag{5}$$

where $L_i\left(v_i\right)$ is given in equation (4).

The estimation procedure consists of repeatedly maximizing (5) with respect to all the model parameters and the parameters in the heterogeneity distribution for alternative values of *M*. By starting out with *M*=1, the model is expanded with new support points until the likelihood can no longer be increased.[8] The scope for adding additional points is evaluated at all stages of the process evaluated by means of simulated annealing (Goffe et al., 1994) as well as by full estimation based on randomly selected heterogeneity parameters. The optimization routine is described in detail in Gaure et al. (2007). The estimation was performed using a supercomputer at the University of Oslo[9].

5. Effects on transitions

The model consists of 1,786 estimated parameters, out of which 475 characterize the treatment effects and 152 characterize the heterogeneity distribution. Due to this large number, this section will only focus on the treatment effects. The estimated effects of some individual characteristics are commented in the appendix while the full set of estimation results can be downloaded from www.frisch.uio.no/docs/vr_prog.html.

5.1. Program effects on the employment hazard

Table 9 reports the estimated after-program effects $\pi_1\left(x^o, r\right)$ on the employment hazard for the reference person[10] in the first month after program completion. The reference

[8] For practical and computational reasons, we consider this to be the case when the log-likelihood increases by less than 0.01.

[9] The program is developed by Simen Gaure at USIT and the Frisch Centre and may be studied at www.frisch.uio.no/NPMLE.html .

[10] The reference person is a female between 30 and 44 years old with a sickness history of at least 12 months, muscular/skeletal illness, more previous work experience than 25th percentile (given age), has completed high school (12 years of education) and participated during average business cycle conditions

person may be regarded as a typical VR client. All programs, except WTP, increase the employment hazard after program completion. WS comes out with the largest effect, followed by EDU, WTO and AMO. The estimated effect for the reference person of having participated in EDU or WS is 0.70 and 1.00 respectively, in the first month after program completion. This corresponds to an increase in the hazard rate of 101 and 172 percent[11]. These effects may represent both newly acquired human capital as well as signaling effects.

Participation in more than one program only has a minor impact on the estimated effects. An exception here is EDU, which seems to have less effect given that participants have already completed other programs. One explanation for this may be that people who have chosen other programs prior to EDU tend to choose different types of education than those who choose this as their first program. In fact, participants with previous program experience spend on average two months less in EDU than participants whose only program is EDU. It may also be a reflection of EDU being more effective when preceding work training. Previous studies have reported similar results. For instance, Lechner and Wiehler (2007) finds that qualification programs[12] are more effective if they precede an active job search program than vice versa. The estimated effect of WS also drops when it follows other programs. However, this difference is not statistically significant. As for EDU, the average duration of WS also drops when it is the last of several programs.

The after-program effect is a declining function of time since completion for all programs. This is especially prominent for the three work training programs. In fact, the positive after-effect of the WTO vanishes and the effect of WS is reduced by 50 percent eight months after program completion. The depreciation rate of the two classroom programs (AMO and EDU) is not that strong however, around one half of the depreciation rate of the three work training programs. One explanation is that different types of classroom training, which often results in a certificate of the newly acquired knowledge, have a more lasting effect. However, as pointed out by Gaure et al. (2007)

---

[11] This increase is calculated from eq (2) as $\Delta\phi_{kit} = \dfrac{\phi_{kit}(a=1) - \phi_{kit}(a=0)}{\phi_{kit}(a=0)} = \exp(\pi_k) - 1$

[12] Qualification programs are classroom programs that may end in a formal vocational degree.

Table 9.
Program effects on the employment hazard

	After-program effects (π)									
	WTO		WTP		AMO		EDU		WS	
	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.
The reference person*	0.41	*0.08*	-0.08	*0.14*	0.31	*0.11*	0.70	*0.06*	1.00	*0.12*
The reference person (with previous program experience) *	0.46	*0.08*	-0.24	*0.14*	0.41	*0.11*	0.38	*0.07*	0.74	*0.11*
+ Interaction with month since completion	-0.05	*0.00*	-0.06	*0.01*	-0.03	*0.01*	-0.03	*0.00*	-0.05	*0.01*
Heterogeneous effects (interaction terms)										
+ Short-term ill	-0.07	*0.04*	0.03	*0.09*	-0.20	*0.06*	-0.18	*0.03*	-0.25	*0.06*
+ Unemployed	-1.22	*0.09*	-0.65	*0.08*	-0.68	*0.09*	-1.12	*0.07*	-0.98	*0.10*
+ Below 30	-0.06	*0.05*	0.12	*0.07*	0.02	*0.06*	-0.03	*0.03*	-0.01	*0.06*
+ Older than 44	-0.09	*0.05*	-0.26	*0.09*	-0.27	*0.07*	-0.03	*0.05*	-0.12	*0.07*
+ Little prev. work exp.	0.03	*0.05*	0.15	*0.08*	0.28	*0.07*	0.14	*0.04*	0.25	*0.08*
+ Male	-0.19	*0.04*	0.08	*0.07*	0.07	*0.05*	-0.12	*0.03*	-0.14	*0.06*
+ Business cycle**	-0.09	*0.02*	-0.04	*0.03*	-0.09	*0.02*	-0.08	*0.01*	-0.15	*0.02*
+ Mental diagnoses	-0.21	*0.05*	-0.02	*0.09*	0.13	*0.07*	-0.26	*0.04*	0.17	*0.07*
+ Other diagnoses	-0.11	*0.05*	0.07	*0.10*	-0.13	*0.07*	-0.09	*0.04*	0.03	*0.07*
+ Years of education (deviation from 12 years)	0.02	*0.01*	-0.01	*0.01*	0.01	*0.01*	0.05	*0.01*	0.00	*0.01*
	On-program effects (μ)									
	WTO		WTP		AMO		EDU		WS	
	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.
The reference person*	-0.88	*0.09*	-1.11	*0.12*	-1.22	*0.17*	-1.96	*0.08*	0.79	*0.09*
The reference person (with previous program experience)*	-0.84	*0.09*	-1.58	*0.12*	-1.10	*0.17*	-2.17	*0.08*	0.59	*0.09*
Heterogeneous effects (Interaction terms)										
+ Short-term ill	-0.19	*0.05*	-0.12	*0.08*	-0.42	*0.09*	-0.32	*0.04*	-0.42	*0.05*
+ Unemployed	-0.55	*0.08*	-0.49	*0.08*	0.35	*0.10*	-0.81	*0.07*	-0.98	*0.08*
+ Below 30	-0.23	*0.05*	0.07	*0.06*	0.24	*0.09*	0.01	*0.04*	-0.08	*0.05*
+ Older than 44	0.17	*0.06*	-0.15	*0.08*	0.12	*0.11*	0.38	*0.06*	-0.05	*0.06*
+ Little prev. work exp.	-0.01	*0.06*	0.11	*0.07*	0.59	*0.10*	0.31	*0.06*	0.23	*0.06*
+ Male	0.05	*0.05*	0.30	*0.06*	0.15	*0.08*	0.21	*0.04*	-0.07	*0.05*
+ Business cycle**	0.02	*0.02*	0.00	*0.02*	-0.03	*0.04*	0.00	*0.02*	-0.09	*0.02*
+ Mental diagnoses	-0.21	*0.06*	-0.10	*0.08*	0.03	*0.11*	-0.07	*0.05*	-0.01	*0.06*
+ Other diagnoses	-0.04	*0.06*	-0.19	*0.09*	-0.09	*0.11*	-0.01	*0.05*	0.04	*0.06*
+ Years of education (deviation from 12 years)	-0.01	*0.01*	0.00	*0.01*	-0.01	*0.01*	-0.01	*0.01*	-0.01	*0.01*

* The reference person is a female between 30 and 44 years old with long-term illness, muscular/skeletal diagnosis, with more previous work experience than the 25th percentile (conditioned on age), has completed high school and participated during normal business cycle conditions. ** Evaluated at one standard deviation.
Note: The heterogeneous effects are included as interaction parts (i.e. in addition to the effect for the reference person. The relative effect on the hazard rate may be calculated as exp(π)-1.

this estimated parameter may induce spurious duration dependence. If there are heterogeneous treatment effects not accounted for in the model, those with the highest (unobserved) treatment effects are the ones who first enter into employment.

The following heterogeneous effects are reported as interaction terms between program dummies and observed covariates. Persons with short-term illnesses experience a smaller after-program effect from AMO, EDU and WS than those with long-term illnesses. For those arriving from unemployment, all the positive after-program effects disappear and, WS excepted, even become negative. These results indicate that the main target group, i.e. those with long-term illnesses, is the one with most to gain from participating in terms of increased employment hazard. It is worth noting that due to the non-linearity aspects of the model, the relative differences and the absolute differences in terms of increased hazard rates may differ substantially. In the next section, however, the program effects on the final outcome probabilities are calculated depending on three pre-VR states. The conclusion is that the long-term illness group experiences the greatest effect in both absolute and relative terms.

Persons younger than 30 experience the same program effects as the reference group (30-44). For participants older than 44 years, on the other hand, all programs, apart from EDU have smaller effects. In addition, those with little previous work experience enjoy a stronger effect from all programs. All programs, except WTP, seem to be less effective when labor market conditions are good. This is in contrast to the findings of Røed and Raaum (2006) where the after-program effects tend to increase with better business cycle conditions. AMO and WS has a greater effect, and WTO and EDU has a smaller effect, on persons with mental diagnoses than on the reference group. In fact, the effect of AMO and EDU is equal for those with mental diagnoses.

The second part of table 9 reports the on-program effect towards employment. All programs, apart from WS, have a negative effect on the employment hazard during participation. This finding is not very surprising. People are focusing on building up their human capital and hence may have less time for job search activities. One of the intentions of the WS program is for the participant to continue working for the firm, even when the public funding ends. This may explain the positive effect. EDU has the strongest lock-in effect, both as a single program and as the last in combination of others. One possible explanation is the relatively high cost of leaving the EDU program before

completion, as the participant will not get a diploma. Participants may expect a wage premium from the diploma, making them more reluctant to accept job offers at this time. Persons with short-term illnesses and the unemployed have stronger negative on-program effects than those with long-term illnesses of almost all programs, i.e. those with the highest employment probability also have the largest lock-in effect. One exception is AMO for the previously unemployed, which actually has a positive interaction part. Persons with little previous work experience have smaller negative on-program effects of AMO, EDU and WS.

The substantial effect of WS corresponds well to previous literature on ordinary labor market programs (see for instance Kluve (2006) for an overview). Criticism has been leveled of these effects failing to take into account that some employers may seek this subsidy for candidates that they would have hired anyway (Martin (1998)). This criticism will also hold for my model. However, it only applies to the on-program effect (this effect captures the direct transition to the program provider). The after-program effect should not be affected since these employers are not receiving a subsidy. The results of this paper show that WS is the most effective work training program, this criticism notwithstanding.

5.2. Program effects on the disability hazard

In the first part of table 10, the after-program effects on the disability hazard are reported. Classroom programs (AMO and EDU) have the largest negative effect. Not surprisingly, providing the participants with a new profession where their previous injuries no longer constitute a problem is the most effective way of reducing the inflow into disability. The effect of WS is also positive. These effects are not very time-persistent. Upon completion, all three have diminishing effects over time. WTP on the other hand, has a positive effect, while WTO has no effect on the disability hazard. One way of exlaining this pattern is that these program have dual objectives. In addition to providing work training, they are also used to establish the participant's degree of work capasity (screening). It may be easier to be granted a disability pension if a third party (i.e. the program provider) is able to confirm that the participant is not in a suitable condition to hold down an ordinary job.

Table 10.
Program effects on the disability hazard

	After-program effects (π)									
	WTO		WTP		AMO		EDU		WS	
	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.
The reference person*	0.04	*0.09*	0.63	*0.11*	-1.25	*0.18*	-1.60	*0.11*	-0.53	*0.21*
The reference person (with previous program experience)*	-0.13	*0.09*	0.58	*0.11*	-1.43	*0.18*	-1.40	*0.11*	-0.57	*0.19*
+ Interaction with month since completion	0.02	*0.00*	-0.01	*0.01*	0.09	*0.01*	0.09	*0.01*	0.05	*0.01*
Heterogeneous effects (interactions)										
+ Short-term ill	0.15	*0.06*	0.15	*0.08*	0.00	*0.11*	0.10	*0.07*	0.14	*0.12*
+ Unemployed	0.00	*0.11*	-1.21	*0.11*	-0.98	*0.29*	-0.08	*0.15*	-0.36	*0.26*
+ Below 30	0.17	*0.07*	0.28	*0.08*	0.15	*0.14*	0.15	*0.07*	0.04	*0.18*
+ Older than 44	-0.13	*0.04*	-0.21	*0.06*	-0.14	*0.08*	-0.19	*0.06*	-0.23	*0.10*
+ Little prev. work exp.	-0.08	*0.05*	-0.12	*0.06*	0.03	*0.09*	0.10	*0.06*	-0.08	*0.13*
+ Male	-0.07	*0.04*	-0.15	*0.05*	-0.11	*0.08*	-0.06	*0.05*	-0.27	*0.11*
+ Business cycle**	-0.07	*0.02*	-0.04	*0.02*	-0.10	*0.04*	-0.09	*0.02*	-0.02	*0.04*
+ Mental diagnoses	-0.09	*0.05*	0.00	*0.07*	0.16	*0.10*	0.13	*0.06*	0.11	*0.13*
+ Other diagnoses	-0.06	*0.05*	-0.10	*0.07*	0.05	*0.10*	0.18	*0.06*	-0.04	*0.12*
+ Years of education (deviation from 12 years)	0.01	*0.01*	0.02	*0.01*	-0.01	*0.02*	-0.01	*0.01*	0.04	*0.02*
	On-program effects (μ)									
	WTO		WTP		AMO		EDU		WS	
	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.
The reference person*	-1.67	*0.13*	-1.38	*0.15*	-1.91	*0.49*	-4.73	*0.22*	-0.61	*0.20*
The reference person (with previous program experience)*	-1.53	*0.13*	-1.12	*0.15*	-1.86	*0.47*	-4.26	*0.23*	-0.45	*0.18*
Heterogeneous effects (Interactions)										
+ Short-term ill	0.05	*0.09*	0.40	*0.10*	0.20	*0.33*	0.14	*0.13*	-0.07	*0.13*
+ Unemployed	0.32	*0.14*	-0.73	*0.13*	-0.97	*1.11*	0.94	*0.24*	-0.77	*0.31*
+ Below 30	0.34	*0.10*	0.52	*0.10*	-0.09	*0.49*	0.57	*0.15*	0.19	*0.17*
+ Older than 44	-0.24	*0.07*	-0.37	*0.08*	-0.11	*0.29*	0.68	*0.12*	-0.38	*0.11*
+ Little prev. work exp.	-0.14	*0.08*	-0.16	*0.08*	-0.72	*0.38*	-0.06	*0.14*	-0.22	*0.13*
+ Male	-0.30	*0.07*	-0.41	*0.07*	-0.49	*0.29*	-0.11	*0.11*	-0.76	*0.11*
+ Business cycle**	-0.03	*0.03*	-0.08	*0.03*	0.13	*0.12*	-0.13	*0.05*	-0.02	*0.04*
+ Mental diagnoses	-0.09	*0.08*	-0.16	*0.09*	-0.40	*0.41*	0.17	*0.14*	-0.18	*0.13*
+ Other diagnoses	0.05	*0.08*	-0.11	*0.09*	0.31	*0.31*	0.44	*0.13*	-0.10	*0.12*
+ Years of education (deviation from 12 years)	0.02	*0.01*	0.03	*0.01*	-0.04	*0.04*	0.03	*0.02*	0.00	*0.02*

* The reference person is a female between 30 and 44 years old with long-term illness, muscular/skeletal diagnosis, with more previous work experience than the 25th percentile (conditioned on age), has completed high school and participated during normal business cycle conditions. ** Evaluated at one standard deviation.
Note: The heterogeneous effects are included as interaction parts (i.e. in addition to the effect for the reference person. The relative effect on the hazard rate may be calculated as $\exp(\pi)-1$.

Participants with short- and long-term illnesses experience the same effects, while the unemployed experience stronger (negative) effects from both WTP and AMO. From table 3 we recall that WTP is the largest program for this group (used by 34.6 percent) this is an encouraging finding. The after-program effect is more negative the older the participants are, except for AMO which has no significant age effect. Since the disability hazard strongly increases by age, this is well worth noting.

The last part of table 10 indicates that all programs have a negative on-program effect on the hazard rate into disability. This is as expected. One of the requirements for being granted a disability pension is to try VR programs first. It may be difficult to obtain a disability pension before the program is completed. Again EDU has the largest lock-in effect, followed by AMO, WTO, WTP and WS. Those arriving from unemployment get a stronger negative effect from participating in WTP, AMO and WS, while their participation in WTO and EDU has less effect compared to people arriving with previous (long- or short-term) illnesses. The lock-in effect from participating in WTO and WTP increases with age, while in EDU both the youngest and the oldest ones have a weaker lock-in effect than the middle-aged.

5.3. Program effects on the temporary withdrawal hazard

Table 11 reports the program effects on the hazard rate into temporary withdrawal. All programs, except WTP, have a negative after-program effect. EDU has the strongest negative effect, while the effects of WTO, AMO and WS are quite similar to each other. The short-term illness group tends to have a somewhat smaller effect of WTO, AMO and EDU than those with long-term illness. The unemployed group, on the other hand, experiences a substantial drop in the hazard rate into temporary withdrawal relative to persons with both short- and long-term illnesses. Persons older than 44 and persons with little previous work experience also get a stronger negative effect from having participated in any of the programs than the reference person does.

All participants experience a negative on-program effect on the temporary withdrawal hazard, except for AMO participants with previous program experience, who are not affected at all. In addition, the lock-in effect of EDU drops substantially when it is the last of several programs (relative to being the only one). In other words,

Table 11.
Program effects on the temporary withdrawals hazard

	After-program effects (π)									
	WTO		WTP		AMO		EDU		WS	
	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.
The reference person*	-0.61	*0.10*	-0.07	*0.13*	-0.73	*0.14*	-1.12	*0.10*	-0.60	*0.20*
The reference person (with previous program experience)*	-0.67	*0.10*	-0.27	*0.13*	-0.31	*0.14*	-0.87	*0.10*	-0.71	*0.18*
+ Interaction with month since completion	-0.02	*0.00*	-0.04	*0.01*	0.04	*0.01*	0.04	*0.00*	0.01	*0.01*
Heterogeneous effects (interactions)										
+ Short-term ill	0.16	*0.07*	0.06	*0.11*	0.22	*0.09*	0.31	*0.06*	0.19	*0.12*
+ Unemployed	-1.17	*0.09*	-1.38	*0.08*	-1.10	*0.11*	-1.24	*0.11*	-1.48	*0.18*
+ Below 30	0.03	*0.06*	0.06	*0.07*	0.00	*0.08*	-0.03	*0.05*	0.11	*0.11*
+ Older than 44	-0.26	*0.06*	-0.25	*0.08*	-0.32	*0.09*	-0.24	*0.07*	-0.38	*0.12*
+ Little prev. work exp.	-0.32	*0.05*	-0.29	*0.07*	-0.33	*0.08*	-0.36	*0.06*	-0.47	*0.11*
+ Male	-0.20	*0.05*	-0.19	*0.06*	-0.27	*0.07*	-0.19	*0.05*	-0.50	*0.10*
+ Business cycle**	-0.14	*0.02*	-0.14	*0.03*	-0.06	*0.03*	-0.15	*0.02*	-0.13	*0.04*
+ Mental diagnoses	-0.13	*0.06*	-0.29	*0.08*	-0.13	*0.09*	0.00	*0.06*	-0.09	*0.12*
+ Other diagnoses	0.01	*0.07*	-0.27	*0.09*	-0.12	*0.09*	0.06	*0.06*	0.06	*0.12*
+ Years of education (deviation from 12 years)	-0.01	*0.01*	-0.04	*0.01*	0.01	*0.01*	-0.03	*0.01*	0.07	*0.02*
	On-program effects (μ)									
	WTO		WTP		AMO		EDU		WS	
	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.
The reference person*	-2.25	*0.11*	-2.08	*0.14*	-2.33	*0.21*	-3.29	*0.11*	-0.77	*0.19*
The reference person (with previous program experience)*	-1.07	*0.10*	-1.72	*0.14*	0.04	*0.16*	-1.77	*0.11*	-1.18	*0.17*
Heterogeneous effects (Interactions)										
+ Short-term ill	0.47	*0.06*	0.30	*0.11*	0.47	*0.10*	0.33	*0.06*	0.22	*0.11*
+ Unemployed	-0.94	*0.08*	-0.58	*0.09*	-1.44	*0.15*	-1.20	*0.10*	-1.19	*0.15*
+ Below 30	0.10	*0.06*	0.19	*0.07*	-0.16	*0.09*	-0.12	*0.05*	0.10	*0.10*
+ Older than 44	-0.37	*0.07*	-0.35	*0.09*	-0.21	*0.11*	0.14	*0.08*	-0.42	*0.12*
+ Little prev. work exp.	-0.45	*0.06*	-0.31	*0.07*	-0.49	*0.09*	-0.31	*0.06*	-0.52	*0.12*
+ Male	-0.09	*0.05*	-0.12	*0.06*	-0.31	*0.08*	-0.19	*0.05*	-0.48	*0.09*
+ Business cycle**	0.16	*0.08*	-0.06	*0.04*	-0.07	*0.02*	-0.10	*0.04*	-2.75	*0.24*
+ Mental diagnoses	-0.01	*0.06*	-0.01	*0.09*	-0.06	*0.10*	0.08	*0.06*	-0.10	*0.11*
+ Other diagnoses	-0.04	*0.07*	-0.07	*0.10*	-0.19	*0.11*	0.07	*0.07*	-0.01	*0.12*
+ Years of education (deviation from 12 years)	0.00	*0.01*	-0.01	*0.01*	0.04	*0.01*	-0.02	*0.01*	0.04	*0.02*

* The reference person is a female between 30 and 44 years old with long-term illness, muscular/skeletal diagnosis, with more previous work experience than the 25th percentile (conditioned on age), has completed high school and participated during normal business cycle conditions. ** Evaluated at one standard deviation.
Note: The heterogeneous effects are included as interaction parts (i.e. in addition to the effect for the reference person. The relative effect on the hazard rate may be calculated as exp(π)-1.

previous program experience seems to increase the probability of dropping-out from classroom training. The lock-in effect is weaker for persons with short-term illnesses and stronger for the unemployed compared to the reference person. Also, the older ones and those with little previous work experience get a stronger lock-in effect than the reference group.

6. Effects on the final outcome probabilities and VR duration

Section 5 reports how the VR programs affect the hazard rates into the three final destinations, both while participating and after program completion. The aim of this section is to sum up all these effects into one, by calculating how the VR programs affect the final outcome probabilities. In order to calculate these effects we have to take into account that each program affects the final destinations in several ways. First, we have both the on-program effect and after-program effect which cannot simply be added up together. Second, each program affects the probability of entering other programs. This leads to an additional second order effect on the final outcome. Third, a program with a negative direct effect on the hazard rate may still have a positive effect on the final outcome probability if the hazard rates towards the other destinations are reduced even further. Fourth, the programs may not only affect the outcome probabilities, but also the time before it occurs. This may be regarded as a cost of VR programs in terms of extended spell duration.

These calculations are performed using the estimated model to simulate new VR spells. First, a simulation based on all the estimated program effects is compared with a simulation where all program effects are assumed to be zero (i.e. $\pi_k = 0$ and $\mu_k = 0$, k=1,…,8). Comparing the final outcome frequencies of these two simulations results in an estimate on the overall effect of how the VR programs affect the share of spells ending in each of the three final destinations. In the rest of this section this is referred to as the "effect on outcome probabilities". In addition, the cost of programs in terms of increased spell duration is reported. Second, simulations for different types of participants are compared in order to see if some groups of participants benefit more from the VR programs than others. The third strategy is to compare two simulations conditioning on

the last program the participant attended. This is done to see how the different programs affect the outcome probabilities of those who actual participate.

The confidence intervals for the overall program effects are calculated using a parametric bootstrap procedure, i.e. parameter estimates are drawn repeatedly from their joint normal distribution[13]. In total, we make 100 simulations for the correct and counterfactual assumptions respectively, and calculate 98 percent confidence intervals for the statistics that characterize the effect of the VR programs (see Røed and Westlie (2007) for previous application of this simulation procedure). Some assumptions are made in order to simplify the simulation process. The business cycle and calendar year variables are always used at their mean values. In addition, the simulation process is ended when spell duration exceeds nine years. Due to this duration limit, approximately one percent of the spells without program effects and two percent of the spells with program effects do not end in a final destination.

Comparing the simulation with program effects using simulations where the program effects are left out, we get the following results, as presented in table 12. First, in the simulation with program effects, 47 percent of spells ended in employment. This number is very close to the actual share of 47.1 percent, as presented in table 3. Comparing this to the employment share of 38.5 percent as in the counterfactual world without any program effects (column 2), we find that *the VR programs increase the average employment probability by 8.4 percentage points.* This difference is highly statistically significant. The VR programs are not equally successful in reducing the disability probability, but the reduction of 2.7 percentage points is highly statistically significant. Most of the increased employment probability comes at the expense of the temporary withdrawal probability, which is reduced by 6.7 percentage points. These favorable effects come with a cost of increased spell duration. On average, the VR programs increase the expected spell duration by 7.4 months. This corresponds to a 40 percent increase relative to the simulation without program effects. The three effects do not add up to zero due to the duration limit (i.e. a few spells ends without a final

[13] Note that we make drawings from the vector (Hva betyr "make drawings from" her? Kan det evt. skrives om?) of 1,634 parameters attached to observed covariates only, since the parameters describing the unobserved heterogeneity are not normally distributed; see Gaure, Røed and Zhang (2007).

outcome). All program effects are measured per spell. However, these differences are only caused by participants. Since 78 percent of the simulated spells contain program participation, the effect per participant is 1.3 times higher than the presented effect.

Table 12.

Predicted impact of VR programs on final outcome probabilities and spell duration.

Final outcome	Outcome probabilities and differences relative to a world with all program effects						
	Estimated program effects	Zero program effects	Zero effect of				
			WTO	WTP	AMO	EDU	WS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Employment	47.0	38.5	45.5	47.0	46.0	42.1	44.3
With – without		8.4	1.5	-0.1	1.0	4.9	2.7
		[7.1 , 9.3]	[1.1 , 1.9]	[-0.5 , 0.3]	[0.7 , 1.4]	[4.2 , 5.5]	[2.1 , 3.3]
Disability	23.3	25.9	22.6	22.4	23.6	26.4	23.8
With – without		-2.7	0.5	0.9	-0.4	-3.2	-0.6
		[-3.4 ,-1.9]	[0.13 , 0.9]	[0.6 , 1.3]	[-0.8 ,-0.1]	[-3.9 ,-2.7]	[-1.0 ,-2.2]
Temporary withdrawal	28.7	35.3	30.9	29.6	29.3	31.1	30.7
With – without		-6.7	-2.2	-1.0	-0.6	-2.4	-2.0
		[-7.7 ,-5.6]	[-2.7 ,-1.8]	[-1.4 ,-0.6]	[-0.9 ,-0.3]	[-3.1 ,-1.9]	[-2.4 ,-1.5]
Spell duration	29.7	22.3	28.5	28.9	29.4	25.6	30.0
With – without		7.4	1.2	0.8	0.4	4.2	-0.3
		[7.0 ,7.9]	[1.0 , 1.4]	[0.6 , 1.0]	[0.2 , 0.6]	[3.9 , 4.5]	[-0.5 ,-0.1]

Note: The confidence interval on the estimated differences is reported in brackets [ ] and at a 98 percent level. Spell duration is measured in months. Two percent of the spells with program effect, and one percent of the spells without, have no final destination.

In columns 3 to 7 we remove the effect of one program at a time, in order to decompose the total effect of the VR programs into the contribution from each of the five programs. Note that the contribution of each program is influenced both by the effectiveness of the program and by the relative size of the program, i.e. both the program effect and the number of participants. EDU has the largest effect on the employment probability. Removing the effect of EDU leads to a decrease of 4.85 percentage points in the employment rate. EDU also has a large effect on the disability probability. The contribution from EDU of -3.2 percentage points is in fact larger than the total reduction. This is due to the positive effect of WTO and WTP. However, EDU comes with the largest cost in terms of increased spell duration which accounts for more than 50 percent of the total increase. In spite of being the smallest program, WS has a large effect on the

average employment probability. The effect on the disability probability is not that large, reducing it by 0.6 percentage points. Last, WS has no cost in terms of increased spell duration. WS in fact reduces the average spell duration by 0.3 month. WTO and AMO both have a positive effect on the employment probability, although not as large as EDU. The fact that both WTO and WTP have a positive effect on the disability probability is not that surprising. VR program participation is required of almost all disability applicants, and WTP and to a certain extent also WTO are targeted at the most vocationally disabled individuals.

Table 13 reports how VR programs affect three different groups of VR candidates depending on the pre-VR state. That is, VR clients with short- and long-term illnesses and the unemployed. This is done by comparing simulations for each of the three groups. The VR programs increase the employment probability for all three groups. Persons with long-term illnesses however, experience the largest increase in both relative and absolute terms.

Table 13.
Predicted impact of the VR programs on spell outcome, conditioned on the pre-DP situation

	Long-term ill	Short-term ill	Unemployed
Employment			
With	40.6	62.0	47.8
With – without	9.6	5.3	8.1
	[8.5 , 10.9]	[3.1 , 7.0]	[6.2 , 9.8]
Disability			
With	28.3	16.1	9.3
With – without	-3.2	-2.9	1.2
	[-4.2 , -2.2]	[-4.3 , -1.9]	[0.1 , 2.2]
Temporary withdrawal			
With	29.5	20.5	42.1
With – without	-7.5	-3.4	-10.1
	[-8.5 , -6.0]	[-5.0 , -1.8]	[-12.1 , -8.1]
Spell duration			
With	30.4	30.9	22.9
With – without	7.1	7.0	10.7
	[6.6 , 7.5]	[6.5 , 7.7]	[9.7 , 11.2]

Note: The confidence interval on the estimated differences is reported in brackets [ ] and at a 98 percent level. Spell duration is measured in months. Around two percent of the spells with program effect, and one percent of the spells without, have no final destination.

The effect on disability is identical for both groups with an illness history. The unemployed, on the other hand, increased their disability probability through the VR programs. Bearing in mind that the unemployed have no previous sickness related transfers, the disability pension is not targeted at this group. Being included in a system

where disability pension is a more common outcome however, may increase the probability of receiving this benefit. The cost in terms of increased spell duration is almost identical for both of the groups with a previous sickness history, at around 7 months. Those with a history of unemployment experience an even larger increase, some 10.7 months. Table 9 reported the estimated after-program effect for the unemployed to be either zero or negative. Nonetheless, the estimated effect on the outcome probability is positive and quite strong. This is due to the competing risks, i.e. the programs reduce the hazard rate to the other final destinations even more. However, these participants experience the largest increase in expected duration.

Table 14 reports the outcome probabilities conditional on the last program in the simulated spell. Since the program effects differ depending on observed characteristics and different persons participate in different programs, the results presented in table 14 may be regarded as the a*verage treatment effect on the treated* (TT). These results rest on the assumption that the same type of person participates in both the true and the counterfactual simulation. Since previous program experience is an important factor in the participation equation of several programs, we have to include this when constructing the counterfactual world, i.e. in contrast to the previous simulations the program effects in the participation equations will here remain their estimated values. That is, in the counterfactual world $\pi_k = 0$ and $\mu_k = 0$, for k= 6,…,8 only.

First in table 14, the simulation with program effects reports employment frequencies nearly identical to those in table 3, the exception being the spells ending with WS, with a 5.5 percentage points lower employment rate compared to the observed data. This may be due to the different programs responding differently to the simplifications in the simulation procedure (i.e. both business cycle and calendar year variables are always at their mean values). In addition, 30 percent of the original population is censored due to the end of the time window. If the composition of individual characteristics differs over time, the expected outcome will also differ.

As expected, the non-participants are not affected when the program effects are removed. Even the number of non-participants stays the same. Looking at the different programs, we recognize much of the same pattern as in table 12, column (3)-(7). However, the program effect reported here is not affected by the number of participants.

Having WTO as the last program increases the employment probability by 5.8 percentage points. AMO, EDU and WS have even stronger effects, increasing the employment probability by 11.8, 15.3 and 30.7 percentage points respectively. WTP, on the other hand, still has no effect on the employment probability.

Table 14.
Predicted impact of VR programs conditional on the last program attended

		The last program in the spell is:				
	Non-participants	WTO	WTP	AMO	EDU	WS
Employment						
With	37.8	39.8	31.9	47.2	60.2	65.5
Without	37.8	34.0	33.1	35.5	44.8	34.9
With – without	0.0	5.8	-1.2	11.7	15.4	30.6
	[-0.9 , 0.8]	[3.9 , 7.4]	[-3.3, 0.9]	[9.2 , 13,9]	[13.4 , 17.0]	[26.8 , 34.5]
Disability						
With	31.4	30.0	31.1	15.3	13.5	12.4
Without	31.3	27.8	25.1	21.9	22.9	22.4
With – without	0.1	2.2	6.0	-6.6	-9.4	-10.0
	[-0.8 , 0.7]	[0.2 , 3.7]	[3.9 , 8.0]	[-8.1 , -5.0]	[-11.2 , -7.7]	[-12.8 , -8.1]
Temporary withdrawal						
With	30.6	28.5	35.4	35.9	24.0	21.3
Without	30.6	37.4	41.4	42.3	31.8	42.1
With – without	0.0	-9.1	-6.0	-6.6	-7.8	-20.8
	[-0.9 , 0.6]	[-10.6, -7.4]	[-8.0 , -3.7]	[-8.2 , -3.8]	[-9.5 , -5.9]	[-24.4,-17.3]
Number of participants						
With	39992	37515	21795	12751	54623	10908
Without	39926	36404	20735	13993	58004	8455
With – without	-21	1156	1104	-1263	-3484	2458

Note: The confidence interval on the estimated differences is reported in brackets [ ] and at a 98 percent level. Spell duration is measured in months. Around two percent of the spells with program effect, and one percent of the spells without, have no final destination.

EDU and WS have nearly the same effect on the disability probability, reducing it by around 10 percentage points. AMO also has a negative effect on the disability probability, reducing it by 6.7 percentage points. The final two programs, WTO and WTP, both increase the disability by 2.1 and 6.1 percentage points respectively. Table 14 does not report any program effects on duration. Since spell duration is affected by the penultimate program, this duration effect would not be a valid estimate of the contribution of each program.

Another interesting result in table 14 is the final outcome probability without program effects. This indicates how participants are selected into the different programs

based on their pre-program outcome probability. Participants with EDU as their last program have an employment probability of 44.8 percent without program effects. This is around 10 percentage points more than all other participants. A more surprising result is that WS participants have nearly the same employment probability as participants in the two other work training programs. One might have suspected that these participants would have characteristics that induce high employability, since the employer has a higher degree of codetermination. However, the caseworkers have been instructed to only offer WS to those with particular difficulties in reentering the labor market. In addition, the WS employer might focus on the participant's health status in order to minimize the expected future sickness payments. This is reflected by the disability probability, which is higher for both WTO and WTP participants than for WS participants.

6.1. Comparing the results of this paper with existing literature

Comparing these results with Aakvik et al. (2005), wee see that one of their main findings, that persons with characteristics indicating a low employment probability are the ones with the highest treatment effect, matches the results of this paper. However, one major difference between their results and the ones in this paper is that they report an average treatment effect of close to zero. This difference may be explained by the way the outcome variable is measured. Their outcome is the employment probability around four and a half year after the subjects apply for programs, thereby discarding all spells with a longer duration ending in employment later on. Put differently, their treatment effect combines the positive effect of increased employment probability with the cost of increased spell duration. In the data presented in this paper, around 10 percent of spells containing programs, last more than four and a half year. As a final comparison, a new simulation is conducted, but this time only for women and with a simulation window of four and a half year. In the simulation with program effects, 41 percent ended in employment, while the corresponding share without program effects ended on 37 percent (not reported). This leaves us with a program effect of 4 percentage points, which is considerably less than the effect of 8.4 percent presented in table 12.

Frölich et al. (2004) find that educational rehabilitation (i.e. EDU) performs worst when it comes to re-employment, which is quite the opposite of the result in this paper. Again this difference may be explained by the fact that their program effect includes both

the on-program and the after–program effect. Their data window allows them to see less than four years after the first person enters their data, indicating that the cost of longer spells may have large impact on their estimated program effect.

7. The instruments

Table 15 reports how the hazard rates for entering into the five programs are affected by changes in the instruments, i.e. α_k from equation (1). All effects are reported at a change of one standard deviation of the instruments. High values of *Share of new programs*, i.e. months with a relative high number of new program slots provided by local firms, increase the hazard rates into WTO, WTP and WS as expected. The hazard rate into AMO on the other hand decreases. One explanation for this pattern is that in times with a low number of available training spots in firms, participants are instead directed into AMO. As mentioned earlier, it may be easier to find a larger classroom than to find new firms willing to hire program candidates. Next, the *Share of new AMO courses* has a strong positive effect on the hazard rate into AMO, while it has no significant effect on any of the other programs. Finally, *work pressure on the caseworker* only affects EDU participation, not surprisingly in a negative way.

Table 15.
Estimated effect of instruments on the hazard rates into treatment (α)

	WTO	WTP	AMO	EDU	WS
Share of new program slots in WTO WTP and WS*	0.050	0.086	-0.020	0.007	0.038
	0.004	*0.007*	*0.008*	*0.004*	*0.009*
Share of new AMO courses*	0.000	-0.006	0.186	0.004	-0.005
	0.004	*0.007*	*0.005*	*0.004*	*0.009*
Work pressure on the caseworker*	-0.010	-0.031	0.082	-0.131	0.061
	0.029	*0.045*	*0.047*	*0.026*	*0.063*

* all estimates are evaluated at one standard deviation of the instrument.

Exclusion restrictions are identifying restrictions, so they can not be tested directly (van den Berg (2007)). Even so, it is possible to investigate whether the employment probability differs depending on the realization of the three instruments. This is done in appendix A2. A brief summary will be presented in this sub-section. The main test in A2 is to see whether any correlation exists between the instrument and the employment probability. However, since the instruments affect the participation probability, and

participation affects employment, the test is conducted on program participants only. The main test result indicates that the instruments have no effect on the employment probability for those who have participated in at least one program. That is, neither the instrument value in the first month nor the average over the first twelve months has any impact on the employment probability. However, the critique presented in Van den Berg (2007), that potential participants may have knowledge of (and hence base their job search efforts on) future realizations of the instrument, is not considered in this test.

8. Concluding remarks

The main findings of this paper can be summarized as follows: First, VR program participation increases the employment probability of an average VR client by 8.4 percentage points. However, the employment effect varies substantially between the different programs. Classroom programs (AMO or EDU) increase the employment probability by 11.7 and 15.4 percentage points. Among the work training programs I find the same pattern as in previous studies on ordinary labor market programs (see for instance Gerfin et al. (2005)), i.e. the closer the program resemble a real job, the larger the effect is. Wage subsidy (WS) in particular has a large effect, increasing the employment probability by 30.6 percentage points. However, as previously discussed, this estimate may be biased upwards if subsidy is provided to firms who would have hired the VR candidate anyway. Even if this criticism is valid, WS is still more effective than the other work training programs. This is probably due to more relevant training and a stronger signaling effect (i.e. an employer has been willing to pay half the salary). The other program that provides work training in ordinary firms, WTO, also has a positive effect, increasing the employment probability by 5.8 percentage points. WTP, on the other hand, has no employment effect.

Second, VR programs reduce the disability probability by 2.7 percentage points. Again the different programs have different effects. Re-educating participants to enable them to find a new profession (AMO and EDU) or providing work training in a competitive environment helps VR clients avoid permanent exclusion from the labor market. The two other work training programs (WTO and WTP) show a positive effect, increasing the disability probability by 2.2 and 6 percentage points respectively. This can be explained by the dual objectives of these programs, i.e. they provide participants with

relevant work training as well as detecting their work capacity. In a welfare state, where one of the main objects is to provide income security for all inhabitants, this positive effect may not be a bad thing. In fact, the results show that the increase in disability comes at the expense of the last outcome, namely temporary withdrawal from the labor market. The cost of the increased disability probability depends on the extent to which this group manages to return to the labor market.

Third, all these desirable effects come with a cost of increased VR duration. On average, a VR spell lasts 7.4 month longer due to the VR programs. More than 50 percent of this increase is caused by EDU. WS, on the other hand, has a small negative effect.

Fourth, the programs have different effects for different participants based on observable characteristics. In particular, program effects differ considerably according to pre-VR status. The main pattern is that the main target group of these programs, i.e. the long-term ill, is the one that stands to gain the most by participating in terms of increased employment probability and reduced disability probability. Other individual characteristics indicating a weak connection to the labor market (such as little previous work experience) also coincide with higher program effect. This result corresponds well with previous findings in the literature (see for instance Aakvik et al. (2005)). It is important to note, however, that the model does not estimate the effect of an early entry into the VR regime. It may be desirable to transfer persons into the VR regime soon after they are cured, even though they experience less effect of the programs. One last thing worth noting is that participants older than 44 years, who face one of the highest disability probabilities, are also the ones whose participation results in the highest reduction in the disability hazard.

References

Aakvik, A. (2001), 'Bounding a Matching Estimator: The Case of a Norwegian Training Program', Oxford Bulletin of Economics and Statistics Vol: 63, No. 1: 115-143.

Aakvik, A. (2003), 'Estimating the Employment Effects of Education for Disabled Workers in Norway', Empirical Economics Vol: 28 No. 1: 515-533.

Aakvik, A. and S. A. Dahl (2006), 'Transitions to Employment from Labour Market Enterprises in Norway', International Journal of Social Welfare Vol: 15, No. 2: 121-130.

Aakvik, A., J. J. Heckman and E. J. Vytlacil (2005), 'Estimating Treatment Effects for Discrete Outcomes When Responses to Treatment Vary: An Application to Norwegian Vocational Rehabilitation Programs', Journal of Econometrics Vol: 125, No. 1-2: 15-51.

Abbring, J. H. and G. J. Van den Berg (2003), 'The Nonparametric Identification of Treatment Effects in Duration Models', Econometrica Vol: 71, No. 5: 1491-1517.

Brinch, C. N. (2007), 'Nonparametric Identification of the Mixed Hazards Model with Time-Varying Covariates', Econometric Theory Vol: 23, No. 2: 349-354.

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 Vol: 64, No. 4: 655-682.

Frölich, M., A. Heshmati and M. Lechner (2004), 'A Microeconometric Evaluation of Rehabilitation of Long-Term Sickness in Sweden', Journal of Applied Econometrics Vol: 19, No. 3: 375-396.

Gaure, S., K. Røed and T. Zhang (2007), 'Time and Causality: A Monte Carlo Assessment of the Timing-of-Events Approach', Journal of Econometrics Vol: 141, No. 2: 1159-1195.

Gerfin, M., M. Lechner and H. Steiger (2005), 'Does Subsidised Temporary Employment Get the Unemployed Back to Work? An Econometric Analysis of Two Different Schemes', Labour Economics Vol: 12, No. 6: 807-835.

Heckman, J. and B. Singer (1984), 'A Method for Minimizing the Impact of Distributional Assumptions in Econometric-Models for Duration Data', Econometrica Vol: 52, No. 2: 271-320.

Heckman, J. J., R. J. Lalonde, and J. A. Smith (1999), 'The Economics and Econometrics of Active Labor Market Programs', in Handbook of Labor Economics, Chapter 31: Elsevier, 1865-2097.

Kluve, J. (2006), 'The Effectiveness of European Active Labor Market Policy', IZA Discussion paper No. 2018.

Lechner, M. and S. Wiehler (2007), 'Does the Order and Timing of Active Labor Market Programs Matter?' IZA Discussion paper No. 3092.

Lindsay, B. G. (1983), 'The Geometry of Mixture Likelihoods - a General-Theory', Annals of Statistics Vol: 11, No. 1: 86-94.

Martin, J. P. (1998), 'What Works among Active Labour Market Policies: Evidence from OECD Countries' Experiences', OECD Labour Market and Social Policy Occasional Papers, No. 35.

McCall, B. P. (1994), 'Identifying State Dependence in Duration Models', American Statistical Association 1994 Vol: Proceedings of the Business and Economics Section, No.: 14 - 17.

Meyer, B. D. (1990), 'Unemployment-Insurance and Unemployment Spells', Econometrica Vol: 58, No. 4: 757-782.

Prentice, R. L. and L. A. Gloeckler (1978), 'Regression-Analysis of Grouped Survival Data with Application to Breast-Cancer Data', Biometrics Vol: 34, No. 1: 57-67.

Røed, K. and O. Raaum (2006), 'Do Labour Market Programmes Speed up the Return to Work?' Oxford Bulletin of Economics and Statistics Vol: 68, No. 5: 541-568.

Røed, K. and L. Westlie (2007), 'Unemployment Insurance in Welfare States: Soft Constraints and Mild Sanctions', Essay 1 in this dissertation.

Raaum, O. and H. Torp (2002), 'Labour Market Training in Norway - Effect on Earnings', Labour Economics Vol: 9, No. 2: 207-247.

Van den Berg, G. J. (2007), 'An Economic Analysis of Exclusion Restrictions for Instrumental Variable Estimation', IZA Discussion Paper No. 2585.

Appendix

A1. Selected estimates

In this appendix, some of the other explanatory variables are presented and discussed.

A1.1 The effect of previous social security and labor market history

Table A1 presents some selected parameter estimates from the model. Persons with short-term illnesses have higher hazard rates into AMO, EDU and WS and a lower hazard rate into WTO than the reference group (long-term illnesses). They also have a higher hazard rate into employment and a lower hazard rate into disability and temporary withdrawal. People arriving from unemployment share much of the same pattern. However, the unemployed experience an even higher hazard rate into employment and temporary withdrawal, and a lower participation hazard into EDU. In contrast to the interaction terms in the treatment effect, I have included a dummy for persons combining sickness and unemployment (i.e. those who become ill while searching for employment). This group resembles the long-term ill in most of the transitions, relative to the other groups. This group is assumed to experience the same treatment effect as the long-term ill.

Persons with high levels of previous labor market income follow much of the same pattern as those with short-term illnesses, i.e. high hazard rates into employment, AMO, EDU and WS, and low hazard rates into WTO and WTP. Previous income has, however, no impact on the transition rate into disability. This may be explained by the fact that previous income is an important factor when calculating the pension.

The following rows report the effect of some selected combinations of age and work experience, i.e. previous work experience for persons aged 25-29 and 35-39. The younger ones have higher employment- and lower disability-hazards than the older ones. In addition, the employment hazard and the participation hazard into AMO, EDU and WS increases with work experience conditional on age. The effect of work experience between age groups is quite similar however. Finally, the age effect dominates the work experience. For instance, comparing 25-29 and some work experience with 35-39 and most work experience, it becomes apparent that the younger group has higher employment hazards and lower disability hazards.

A higher level of previous education results in higher hazards into employment and EDU and lower hazards into AMO. At the high school level, the vocational track leads to work training while general studies leads to EDU, all other things being equal.

A1.2 Family background

Having a spouse increases the hazard rate into disability, although having a spouse in the labor market also increases the employment rate. This employment rate increases even more the higher the spouse's income is. Having a spouse that is not receiving a disability pension reduces the hazard rate into temporary withdrawal. Being a parent, on the other hand, has a clearly negative effect on the disability rate, both for males and females. This effect diminishes as the children grow older. The employment hazard for mothers is strongly affected by the age of the child, whereas males' employment hazard is only affected by having children or not. In fact, females with small children have low hazard rates towards all eight destinations (compared to both females without kids and with older kids) except for temporary withdrawal. Males with young children have relative high hazard rates into employment and WS.

Table A.1

The effect of some selected estimates on transitions into

	Employment	Disability	Temporary withdrawal	WTO	WTP	AMO	EDU	WS
Short-term illness	0.59	-0.52	-0.29	-0.07	-0.04	0.38	0.31	0.28
	0.02	*0.04*	*0.04*	*0.06*	*0.06*	*0.06*	*0.06*	*0.06*
Long-term illness	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
Unemployed with sickness history	0.21	-0.31	0.10	-0.07	0.44	0.18	0.01	0.11
	0.02	*0.02*	*0.02*	*0.01*	*0.02*	*0.02*	*0.01*	*0.03*
Previously unemployed	1.72	-0.36	1.73	-0.31	1.59	0.19	-0.55	0.78
	0.05	*0.07*	*0.05*	*0.09*	*0.09*	*0.09*	*0.09*	*0.09*
Previous income*	0.08	0.00	0.03	-0.03	-0.15	0.02	0.08	0.04
	0.01	*0.01*	*0.01*	*0.01*	*0.01*	*0.01*	*0.00*	*0.01*
Expected disability pens*	-0.09	0.10	-0.18	-0.04	0.00	0.03	0.00	-0.10
	0.01	*0.02*	*0.02*	*0.01*	*0.02*	*0.02*	*0.01*	*0.03*
Only compulsory school	-0.16	0.15	0.29	0.05	0.37	-0.10	-0.61	-0.05
	0.02	*0.03*	*0.03*	*0.02*	*0.03*	*0.03*	*0.02*	*0.03*
High school, general studies (one or two years)	-0.19	0.04	0.11	0.07	0.27	0.01	-0.33	-0.05
	0.03	*0.04*	*0.04*	*0.02*	*0.04*	*0.04*	*0.02*	*0.05*
High school, general studies (three years)	-0.05	-0.01	-0.10	-0.13	-0.12	-0.17	0.14	-0.20
	0.03	*0.05*	*0.04*	*0.03*	*0.05*	*0.04*	*0.02*	*0.06*
High school, vocational	-0.13	0.09	0.12	0.10	0.35	0.03	-0.33	-0.01

Table A.1

The effect of some selected estimates on transitions into

	Employment	Disability	Temporary withdrawal	WTO	WTP	AMO	EDU	WS
track (one or two years)								
	0.02	*0.03*	*0.03*	*0.01*	*0.03*	*0.02*	*0.01*	*0.03*
High school, vocational track (three years)	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
Higher education (three years or less)	0.14	-0.02	-0.22	-0.13	-0.40	-0.26	0.32	-0.04
	0.03	*0.05*	*0.05*	*0.03*	*0.06*	*0.04*	*0.02*	*0.06*
Higher education (more than three years)	0.31	-0.01	-0.10	0.06	-0.52	-0.48	0.22	0.23
	0.03	*0.05*	*0.06*	*0.03*	*0.07*	*0.05*	*0.02*	*0.06*
Age25-29, no work exp.	-0.61	-0.29	0.90	0.11	0.79	-0.12	-0.14	-0.53
	0.06	*0.08*	*0.07*	*0.04*	*0.06*	*0.06*	*0.03*	*0.09*
Age 25-29, some work exp.	0.13	-0.45	0.24	-0.05	0.02	-0.09	0.15	-0.19
	0.04	*0.07*	*0.05*	*0.03*	*0.05*	*0.04*	*0.02*	*0.06*
Age 25-29, most work exp.	0.20	-0.56	-0.12	0.04	-0.19	0.00	0.19	-0.01
	0.04	*0.07*	*0.05*	*0.03*	*0.05*	*0.04*	*0.02*	*0.05*
Age 30-34, more work exp.	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
Age 35-39,no work exp.	-1.07	0.37	0.97	0.04	0.80	-0.26	-0.74	-0.25
	0.09	*0.09*	*0.08*	*0.06*	*0.08*	*0.09*	*0.06*	*0.12*
Age 35-39, some work exp.	-0.51	0.33	0.60	-0.03	0.21	-0.10	-0.35	-0.28
	0.05	*0.07*	*0.06*	*0.04*	*0.06*	*0.05*	*0.03*	*0.08*
Age 35-39, more work exp.	-0.23	0.33	0.24	0.01	0.18	-0.06	-0.26	-0.07
	0.04	*0.05*	*0.05*	*0.03*	*0.06*	*0.04*	*0.03*	*0.06*
Age 35-39, most work exp.	-0.05	0.30	-0.14	0.10	0.13	-0.04	-0.15	-0.04
	0.03	*0.05*	*0.05*	*0.03*	*0.05*	*0.04*	*0.02*	*0.05*
Age 45-49, no work exp.	-1.47	1.12	1.48	-0.04	0.86	-0.27	-1.50	-0.89
	0.15	*0.12*	*0.11*	*0.09*	*0.12*	*0.13*	*0.12*	*0.26*
Age 45-49, some work exp.	-0.65	1.08	0.75	0.04	0.38	-0.34	-0.81	-0.28
	0.06	*0.06*	*0.07*	*0.04*	*0.06*	*0.05*	*0.03*	*0.08*
Age 45-49, more work exp.	-0.24	1.02	0.36	0.17	0.28	-0.25	-0.69	-0.12
	0.05	*0.06*	*0.07*	*0.04*	*0.07*	*0.06*	*0.03*	*0.08*
Age 45-49, most work exp.	-0.16	1.14	0.01	0.14	0.43	-0.24	-0.72	0.00
	0.04	*0.05*	*0.06*	*0.03*	*0.05*	*0.04*	*0.03*	*0.05*
Single	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
Spouse's labor market income*	0.04	0.00	0.01	0.00	-0.05	0.01	0.01	0.01
	0.01	*0.01*	*0.00*	*0.01*	*0.01*	*0.01*	*0.01*	*0.01*
Spouse is working	0.26	0.24	-0.11	0.10	-0.21	-0.07	-0.01	0.18
	0.02	*0.02*	*0.02*	*0.01*	*0.02*	*0.02*	*0.01*	*0.03*
Spouse is receiving a disability pension	-0.01	0.34	-0.09	-0.02	0.02	-0.28	-0.33	0.02
	0.04	*0.04*	*0.05*	*0.04*	*0.05*	*0.05*	*0.04*	*0.07*
Spouse is not working	0.04	0.14	-0.10	0.06	0.11	-0.04	-0.06	0.10

Table A.1

The effect of some selected estimates on transitions into

	Employment	Disability	Temporary withdrawal	WTO	WTP	AMO	EDU	WS
	0.03	*0.03*	*0.03*	*0.03*	*0.04*	*0.04*	*0.02*	*0.05*
Female (no children)	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*	*ref*
Female (1 child, 0-3 years)	*-0.38*	*-0.54*	*1.36*	*-0.15*	*-0.18*	*0.00*	*-0.13*	*-0.25*
	0.04	*0.08*	*0.04*	*0.03*	*0.06*	*0.05*	*0.03*	*0.10*
Female (1 child, 4-6 years)	*-0.22*	*-0.38*	*0.16*	*0.00*	*-0.15*	*0.00*	*-0.04*	*0.09*
	0.04	*0.07*	*0.05*	*0.03*	*0.06*	*0.05*	*0.03*	*0.08*
Female (1 child, 7-12 years)	*0.02*	*-0.29*	*-0.08*	*0.07*	*-0.14*	*0.15*	*0.12*	*0.06*
	0.03	*0.04*	*0.04*	*0.02*	*0.04*	*0.04*	*0.02*	*0.06*
Female (1 child, 13-16 years)	*0.23*	*-0.20*	*-0.13*	*0.11*	*-0.04*	*0.14*	*0.18*	*0.21*
	0.03	*0.04*	*0.04*	*0.03*	*0.04*	*0.04*	*0.02*	*0.06*
Male (no children)	0.10	-0.17	0.43	-0.14	0.35	-0.08	-0.19	0.61
	0.03	*0.03*	*0.03*	*0.02*	*0.03*	*0.02*	*0.01*	*0.04*
Male (1 child, 0-3 years)	0.32	-0.64	0.44	-0.11	0.18	0.07	-0.09	0.94
	0.04	*0.07*	*0.05*	*0.03*	*0.05*	*0.04*	*0.02*	*0.05*
Male (1 child, 4-6 years)	0.30	-0.43	0.60	-0.09	0.20	-0.03	-0.03	0.94
	0.04	*0.08*	*0.06*	*0.04*	*0.06*	*0.05*	*0.03*	*0.06*
Male (1 child, 7-12 years)	0.29	-0.40	0.51	-0.10	0.12	-0.09	-0.09	0.88
	0.04	*0.05*	*0.05*	*0.03*	*0.05*	*0.04*	*0.02*	*0.05*
Male (1 child, 13-16 years)	0.22	-0.36	0.54	-0.10	0.21	-0.11	-0.15	0.87
	0.04	*0.05*	*0.05*	*0.03*	*0.05*	*0.05*	*0.03*	*0.06*

*Evaluated as deviation from the mean. The unit is one standard deviation.
Note: The standard errors are reported in italic below the estimates.

A2 Testing the instruments

Table 7 reports the logit estimates on how the realizations of the three instruments, at the entry month into the VR regime, affect the log ratio into employment relative to non-employment. This estimation is only performed on persons who actually participate in a program, since the instrument may affect the participation decision and participation may affect the employment probability. The results are reported as log ratios, with the corresponding standard errors and t-values. We see that the value of the instruments at the inflow have no impact on the employment probability.

Table A2.
Effect of the instruments at spell start on the final outcome. Only spells containing programs.

	Odds Ratio	Standard errors	t-value
Share of new training slots	0.999	0.001	0.00
Share of new AMO slots	0.998	0.002	0.49
Work pressure on the caseworker	1.028	0.017	1.22
Number of observations	73,715		
F-test all instruments (P-value)	0.34		

Other covariates included are pre-VR state, age, previous work experience and labor market income, education, gender, business cycle condition at spell start and spell completion, number of children.

The main model estimated in this paper not only uses the instrument value at *entry*, but at *each month* the candidates are likely to enter a program. A more correct test-indicator could thus be the each candidate's mean value in the first twelve months of the spell. This test is also applied on participants only. In addition, we restrict the sample to spells with duration of at least 12 months. This duration condition is applied in order to be able to construct an average level of the instrument. By choosing 12 months, any seasonal variation in the inflow will be neglected. Table 8 reports the odds ratios from this estimation. We see that neither the share of new training slots nor work pressure on the caseworker has any significant effect on the odds ratio to employment. However, the t-value of the relative changes in new AMO slots is close to two. This does not necessarily mean that the exclusion restriction is violated. Table 15 showed how this instrument only affected the AMO hazard. Since the different programs have a different effect on the employment probability, this test picks up the differences in the allocation of programs. Making the same regression conditioned only on AMO participants reduces the t-value to 0.96 (not reported). This indicates that the correlation between the instrument and the employment ratio is due to people choosing different programs at different values of the instruments, and these different programs may have different effects on the employment ratio. Since the main model estimates the transition rate into each of the five different programs, this will not be a problem in the main model.

Table A3.
Effect on the final outcome of the average value of the instruments in the first year of the spell. Only spells containing programs and with a duration of at least 12 months.

	Effect on the odds ratio	Standard errors	t-value
Share of new training slots	0.996	0.004	-1.13
Share of new AMO slots	0.985	0.008	-1.92
Work pressure	0.949	0.120	-0.42
Number of observations	61,610		
F-test all instruments	0.078		

Other covariates included are pre-VR state, age, previous work experience and labor market income, education, gender, business cycle condition at spell start and spell completion, number of children.

Summing up, it seems that the exclusion restrictions are valid as the instruments have no association with labor market outcomes. Another concern however, as mentioned by van den berg (2007), is that information about future realizations of the instruments are common knowledge among the participants. If so, participants may base their job search efforts on these values and hence the exclusion restriction conditions are not fulfilled. In my case however, I use small time-varying *shocks*, making it difficult to foresee the future values. Based on that, I will assume that people are not behaving according to future values of the instruments.