UiT
THE ARCTIC UNIVERSITY OF NORWAY

School of Business and Economics

# Labor Market Programs in Norway

*Do labor market programs improve the job prospects for individuals with reduced working ability? An evaluation.*

**Espen Hauan**

*Master Thesis in Economics*
*May 2015*

### *Foreword*

This master thesis marks the end of my master's degree in economics at The School of Business and Economics at University of Tromsø. It has been a ride of great learning and challenges. I would like to thank the staff at The School of Business and Economics for all the great lectures and seminars and also for being easily available for discussions and guidance. I would also like to thank my classmates for making these five years enjoyable. This master thesis was written with great support from NAV, and I would like to thank Oddmund Klæboe, Hans-Jostein Melbøe and Bente Ødegaard at NAV Troms and also Helene Ytteborg at Arbieids- og velferdsdirektoratet. All of whom contributed with great advice and encouragement. Lastly, I would like to thank my supervisor Mikko Moilanen for his excellent guidance. I am very thankful for the enthusiasm he has shown regarding this project.

### *Abstract*

An econometric framework, anchored in the literature, is used to evaluate the effect of Norwegian labor market programs. In Norway, the social benefits can roughly be divided into: unemployment benefits, sickness benefits, work assessment allowance (WAA) and disability pension. Receivers of WAA are individuals that has been declared with reduced working ability because of physical/mental/social issues that makes them unable to work regularly. This thesis will look at how labor market programs are used and how it affects the job prospects for those individuals that have reduced working ability and are receiving WAA. As this is an observational study, the problem of selection bias arises. This is dealt with by using a selection model that allows for analyzing both the selection process into programs and also the employment outcomes after programs. As only one outcome is observed (work or no work), the model estimates the unobserved missing outcome, allowing for the estimation of various treatment effects. Over all the programs and the usage of programs appear successful. The treatment effect on the treated are positive at 5.8 percentage points. The other treatment effects suggests that programs are directed towards people who benefit from program participation rather than towards those who not benefit, in terms of labor market outcome. There is also shown that there is considerably heterogeneity in the response to treatment over different groups. As the former is based on observable characteristics, there is also shown heterogeneity in response to treatment based on unobservable characteristics. Lastly, sub analyses show that, among the most frequently used programs, training (education) appears as the best program, while assistance is in between and work practice performs the worst.



*Keywords: Labor market programs, vocational rehabilitation, program evaluation, discrete choice models, selection bias, treatment effect*

# Contents

# Tables



# Figures

# 1 Introduction

General labor market policies and regulations are important tools for well-functioning labor markets. Such policies are set out to increase employment and provide better social assistance and are thus of great interest for researchers, policy makers as well as the general public. How countries shapes and conduct their labor market policies varies with respect to development, wealth, ideology and so forth. Norway has a quite regulated labor market as does many other advanced economies. Norway also has extensive social policies for people who are struggling in the labor market. Particularly, the benefit arrangements are generous, highlighted by the fact that Norway has the highest level of spending on sickness and disability benefits, as a share of GDP, among OECD countries (OECD 2013). In the broad scope of what is labor market policy, one has labor market programs (sometimes called active labor market policies or active labor market measures). These are economic-political measures that are supposed to be a tool against unemployment, disability and inactivity. Labor market programs allows unemployed to participate and the goal is to increase job prospects for those who participates. In other words these are government services that are supposed to help and enable individuals that are unemployed to find and get work. Within the term labor market programs, there is a variety of programs that are supposed to fit different people with different needs of assistance. What kind of programs that is within this definition varies, as some are overlapping with pure social help. In general, labor market programs include such programs as work practice, training, education and general help and advice. An unemployed that participates on one of these programs are supposed to increase his/hers employability such that after the program duration is over, the person will be closer to regular employment.

Labor market programs (LMP) are usually managed by a country's labor and welfare administration. In Norway, this is NAV which is a government agency that administers the labor and welfare on behalf of the government. In Norway, LMPs are mainly targeted at those with reduced working ability, which by definition is some sort of issue (physical/mental/social) that, at the time of status, prevents the individual from maintaining his/hers full working ability. In September 2014, approx. 84 percent of the individuals on LMPs were also registered with reduced working ability. Since most of the LMPs are aimed towards those with reduced working ability, this thesis will rule out regularly unemployed and focus entirely on those with reduced working ability. More specifically, this thesis will look at those who receives work assessment allowance (WAA), which is a social benefit/entitlement. Most of those who are registered with reduced working ability are entitled to WAA.

The social benefits in Norway can roughly be divided into: unemployment benefits, sickness benefits, WAA and disability pension. WAA is a benefit for people who, because of some physical/mental/social issues, have a reduced ability to work and for this reason are (temporarily) out of the labor force. NAV pays the benefit and also provides a process where the ultimate goal is that the individual obtains regular work. This process might involve one or several LMPs. This thesis will look at those who started receiving WAA from March 2010 up to and including December 2011. The reason for this is that there was a reform in NAV that entered to force March 2010. This involved combining several benefits into one witch became WAA. This makes it particularly interesting to evaluate the effect LMP has on this relatively new group of receivers, as there has not been much research on the topic after the implementation of WAA.

Although the welfare gains and costs regarding labor market policies goes beyond budget chapters, some of the numbers from the state budget can help put things in perspective. In the Norwegian state budget for 2015 there is suggested to spend 7.8 billion NOK on LMP. This is important by itself, but maybe more importantly is the goal one hope to achieve which is to get people to work rather than receive social benefits. As mentioned, the main social benefits in Norway are unemployment benefits (12.4 billion), sickness benefits (39.6 billion), work assessment allowance (34.1 billion) and disability pension (78.1 billion). Combined this is 164.2 billion NOK, which is approximately 13.7 percent of the Norwegian state budget. In comparison this is almost twice as much as the total grant suggested to the Ministry of Education (85.8 billion). There are significant economic gains to achieve if LMPs can contribute to help people from social benefits and into employment.

This topic raises several questions of interest. What effect do LMPs have on the job prospects for those individuals that participate? Which individuals are selected to program participation? Is the selection process effective from a welfare perspective? Which groups benefits from LMP? Which of the specific programs work best? To shed some light on such questions, this thesis will use Norwegian data and construct a fitting econometric model. The topic of interest is how program participation affects employment outcome. This is studied by running a system of equations where program participation acts as a treatment. The goal is to measure the effect of the treatment (program participation) on the outcome (employment) in comparison to the outcome for the non-treated. The treated will make up a “treatment group” and the non-treated a “control group”. The individuals are not randomly selected to program participation, which means that they are not randomly assigned to the treatment group and the control group. This causes the selection problem which is a statistical challenge but also very interesting to deal with and study. Since the goal is to measure the treatment effect, the optimal would be to measure the employment outcome after program participation for an individual, in comparison to the employment outcome if that same individual did not participate on a program. Since this is not possible (comparing the same person in two different outcomes), the goal is to compare comparable people from the treatment and the control group. This kind of treatment effect framework is well anchored in the literature as a fitting method for these kinds of evaluations (Heckman et al. 1999).

# 2 Institutional settings

The following chapter will give an overview of the institutional settings. The goal is to get an understanding of how the Norwegian welfare system actually works. This involves a discussion about the manager of the Norwegian welfare system (NAV) and relevant social benefits, statuses and definitions.

## 2.1 NAV

NAV is the Norwegian labor and welfare administration. It is a government institution or more precisely an agency which is governed by the Ministry of Labor. Through its mandates, NAV manages all the social benefits including pensions on behalf of the Norwegian Government and thus the public. NAV’s main goals are to make sure that the labor markets function properly and be of social and financial assistance for people who needs help. There is also a stated goal for NAV to get more people in employment and activity and less on social benefits. NAV has offices all over the country such that case workers and administrators are near the people receiving services. NAV is now a brand name and is written in capital letters, but was originally short, in Norwegian, for new labor and welfare administration.

## 2.2 Definitions

### 2.2.1 *Labor market programs (LMP) and vocational rehabilitation (VR)*

There are many different terms and expressions for different measures and categories and it often has different touches across languages. This section will try to clarify some of these expressions related to NAV. By definition labor market programs (LMP) is a wide basket of work-related programs aimed towards anyone that is unemployed, be it because of health or lack of job availability. The programs are different in purpose and nature and NAV assesses which programs are appropriate for different individuals. The term vocational rehabilitation (VR) or vocational rehabilitation programs is often used to specify the programs that are aimed towards those with reduced working ability. In NAV though, several of the same training programs is applied to both those with reduced working ability and those without. The terms LMP and VR will be used somewhat interchangeably, but it means more or less the same as VR is a subset of LMP, and in Norway makes up most of it.

### 2.2.2 *Program participation*

Program participants (participating on an LMP/VR) must not be confused with receivers of unemployment/sickness benefits or work assessment allowance, as individuals can receive benefits without participating on a program and vice versa. Program participation and receiving social benefits are hence independent of each other, but the absence of regular employment or ability to work is one of the criterions for both. Say someone is registered as unemployed/sick/reduced working ability. Then, if entitled and after some casework, one starts receiving some sort of social benefit. Also, during this time, that is after being registered as unemployed/sick/reduced working ability, a case worker might assign the individual to a program, such that the individual is on a labor market program, hence become a program participant. An individual might start on a program before or after starting receiving social benefits, but always after being registered as unemployed/sick/reduced working ability. Therefore program participation and receiving social benefits is independent of another (although very often coinciding). In order to receive social benefits, the individual must be entitled to it and there is often some case work and applications while figuring out what the person is entitled to. For program participation on the other hand, the NAV case worker may assign the person to a program after his/hers own discretion.

### *2.2.3 Reduced working ability*

Reduced working ability is a condition or a status of a person who is unable to work because of his/hers health. Reduced working ability could be due to for instance physical, mental and/or social issues. The definition of reduced working ability requires that the reduction of the ability to work must be of such a nature that it prevents the person to keep or get regular income producing work. It is also defined that the reduction in the work ability shall be caused by illness, injury or disability. A person is declared with reduced working ability by NAV often based on a physician's assessment. This is a so called assessment of working ability and it must be underwent for the person to be declared with reduced working ability. Reduced working ability in its self does not include any benefits but is only a status.

### 2.2.4 *Work assessment allowance (WAA)*

If one has been registered with reduced working ability one most likely has the right to work assessment allowance (WAA), which is a social benefit given to people with reduced working ability. To be eligible for WAA the working ability must be reduced by at least a half, which means that reduced working ability in fact is one of the conditions for WAA. How much WAA a person can receive is based on his/hers previous income. WAA makes up 66% percent of the previous income except for a fixed maximum and minimal amount. As the word suggests work assessment allowance is an allowance persons receive while their work ability is assessed and eventually clarified, where the goal (in most cases) is regular employment. The maximum time of this benefit is four years, although extensions can be made under special circumstances (and as revealed by the data, is often granted). This process is often aided by some program and/or medical treatment. The process revolving reduced working ability and WAA is summarized in table 1, where some person's process is followed.

*Table 1. The process regarding WAA.*

| | | |
|---|---|---|
| Person is declared with reduced working ability based on an assessment of the working ability. (Often after being on a sick leave). → | Person already have a case worker assigned. Maybe some efforts are made to get the person back to work (for instance LMP). The person might apply for WAA. → | Person is declared to have his/hers working ability reduced by at least half, which entitles him/her to WAA. → |
| Person is now a receiver of WAA, and some further sort of plan is made to fit the persons need. → | Now the process revolves around assessing and eventually clarifying the person's (in)ability to work. → | Some sort of measure might be suggested. This could be medical treatment and/or some sort of LMP. → |
| The person has now received one or several measures/treatments, or maybe none. NAV treats each person individually. → | The maximum time span for receiving WAA is four years. At some point during this time or at the end, the work ability/status is clarified. → | The person ends up in regular employment, unemployment, out of the labor force or disability pension (or maybe other statuses). * |

### 2.3 The current landscape

In table 2, some numbers that are of interest is listed and measured as a percentage of the working age population. The working age population is the whole population in the age span 18 – 67 years old and must not be confused with the labor force which is the employed and the job seeking unemployed. The regularly unemployed and those on disability pensions are included in the table to get some more perspective on the magnitude of the amount of disabled workers.

Notice that the LMP's in Norway are mainly targeted on people with reduced working ability, as 55 346 of the 65 654 program participants have been registered with reduced working ability. The rest of the LMP participants are mainly long term unemployed, young unemployed and immigrants. As mentioned before, reduced working ability is in practice a criterion for eligibility to WAA. This is reflected by the numbers as they show that there were 203 355 people registered with reduced working ability and 151 796 recipients of WAA in October 2014. Further these number means that about one quarter of those with reduced working ability were on some form of an LMP in September 2014, as 55 346 of those where participating on some LMP.

*Table 2. Unemployed and disabled workers as percentage of working age population.*

| Status | Number of persons | Percent of working age population* |
|---|---|---|
| Unemployed (regular jobseeker) | 72 114 | 2.2% |
| Disability pension | 310 000 | 9.4% |
| Reduced working ability | 203 355 | 6.1% |
| Work assessment allowance | 151 796 | 4.6% |
| On LMP (program participants) | 65 654 | 2.0% |
| Reduced working ability, on LMP | 55 346 | 1.7% |

*Working age population (age 18-67) in Norway is approximately 3.3 million.

Source: NAV. Numbers are from September and October 2014.

### 2.4 The programs

The different programs within LMP vary in scope and contents. The idea is that different programs are to help different people with different needs of assistance. The most commonly used in NAV towards those with reduced working ability and in the dataset at hand are roughly categorized as: work practice, training (education), assistance (”following up”), wage subsidies, clarification and facilitated work. First, facilitated work is not of interest in this evaluation and will be dropped from the dataset as explained later. This is because facilitated work are aimed at individuals that requires special attention and are not expected to have regular jobs again, so it is really a tool of social assistance rather than a labor market measure. The other five, however, are commonly used programs that are aiming to help people find and get jobs. Table 3 gives a summary of how programs are used in the data at hand (this means for receivers of WAA).

Facilitated work, which is a commonly used program, is left out when the percentages are calculated in table 3. The first two columns are from the dataset at hand. Notice that people can participate on one or several programs during a spell of WAA, although only one at a time. Column one considers the main program of the individual during the spell at WAA i.e. the longest lasting.  The second column is the total usage of programs independent of individuals. One unit is one month of program participation. One can see that training becomes larger in this column because those programs are often more lengthy than the others, thus more months is counted. One can also see that clarification becomes much smaller, as those programs are often much shorter in term of months. Different views give somewhat different results but roughly speaking; work practice and training are the most used programs while assistance is third. Clarification and wage subsidies are the least used programs except other special programs here categorized under “other”.

*Table 3. Usage of programs towards receivers of WAA (in percentage of total program usage excluding facilitated work)*

| | Longest lasting program for each individual. | Total usage of program (in terms of months) independent of individual. | As reported by NAV in the month of September 2014. |
|---|---|---|---|
| Work practice | 32% | 32% | 33% |
| Training | 31% | 39% | 28% |
| Assistance | 16% | 17% | 22% |
| Clarification | 14% | 6% | 6% |
| Wage subsidies | 4% | 5% | 10% |
| Other | 3% | 1% | 1% |

The categories here have several subcategories as defined by NAV. Work practice is divided into work practice in ordinary environment and work practice in protected environment. These are actually quite different programs as work practice in protected environment is aimed at individuals with substantially more need of assistance. The data does not allow for dividing work practice into these two different categories, which is a weakness that will be discussed later. Training consists mostly of some sort of ordinary education at private or public schools/universities. The other significant subcategory within training is so called labor market training. These are shorter programs, up to ten months, were participants receive both theoretical and practical work related training. Training thus separates a bit from other programs, as it has the element of classroom schooling. Next, assistance or “following up” is mostly assisted work where the participant is working with some sort of assistance. This happens in companies that have special agreements with NAV and are obligated to follow up and assist the participants. Clarification is short programs that are meant to help the participant clarify what he/she could do. The duration is usually around a month and it involves an evaluation by NAV in cooperation with the participant on where to go next, so it is not really interesting as a labor market measure. It includes such things as exploration of possibilities for customization at current/previous work place or other work places. Often the clarification is followed by other measures. Wage subsidies are roughly divided into a limited and an unlimited time period. Wage subsidies with an unlimited time period are usually aimed at individuals that are not expected to work regularly again.

# 3 Literature review

The development within the field of labor market evaluations and within econometrics has led to what  Kluve (2006) characterizes as “third generation” evaluation studies, and argues that this now is more or less standard. This is a microeconometric framework that involves some sort of treatment effect model with some method to adjust for selection bias, as will be discussed more throughout. This thesis will look at those with reduced working ability, but most of the literature actually concerns the effect programs has on regularly unemployed and not necessarily on those with reduced working ability. Kluve (2006) states that “…several countries also have specific active labor market programs for the disabled (workers), but very few evaluations of these measures exists.” Although many of the program evaluations are aimed at a slightly different policy then what will be evaluated in this thesis, the methodology will be similar which makes such papers interesting to review. Further it would make sense to look at some evaluation studies done in Scandinavia and particularly in Norway since this is the actual objective of the thesis, and also because Scandinavian social arrangements are generally more generous than other countries. Based on the literature, the general results from evaluation studies are ambiguous and seems to range around no, little and some positive effect on the desired outcome. The outcome is often employment but could also be for instance earnings. Evaluation studies that are done in Norway also show mixed effects.

## 3.1 Methods in the literature

### 3.1.1 *Observational vs. experimental studies*

Labor market program evaluations are either based on experimental data or observational data. An experimental study allows the researcher to control the statistical sample and randomly assign the treatment, thus creating a randomized treatment and control group. This is regarded as the most robust form of evaluation (Heckman et al. 1999; Kluve 2006). This is because it, to a large extent, eliminates the selection problem. By having control of the study the researcher knows that the ones who receive the treatment (program participation) and the ones who do not are randomly assigned. An observational study, also called a non-experimental study, allows the researcher only to observe the data. This means that the treatment group and control group are not randomized. Regarding observational studies as this thesis, there may in fact be several reasons that some people are selected for program participation and some are not.

### 3.1.2 *Selection problem*

In an experimental study the individuals who receive treatment is randomized by the researcher. This means that their personal characteristics, such as demographics, education, work experience, personal motivation etc. should not be the cause of their selection or non-selection to a program. The researcher also knows that the settings and surroundings do not affect individual's selection or non-selection to a program.

When dealing with observational data the selection bias is a known problem. The problem is that we expect "… differences in unobservables that are related to program participation" (Blundell & Costa Dias 2000). That is to say that there are characteristics of the individuals, that the researcher cannot observe, that affect their potential program participation and hence straight forward regressions would produce biased estimates.

In observational studies the researcher only has the ability to observe the data. A group of unemployed individuals at the start of the observation period makes up a sample. Some of these individuals participate in programs and some do not. If one runs an evaluation one would find some results regarding the program participants and some results regarding the non-participants. Results from such an evaluation would be influenced by selection bias and would not give any insight to the real effect of program participation. The selection process to the program would cause biased estimates. The two main reasons for the selection problems are in the literature known as self-selection and administrative selection (Aakvik 2001; Frölich et al. 2004). Self-selection is because of the influence the individuals themselves have on their potential program participation. For example, it is expected that more "motivated" individuals are more likely to get themselves selected to programs. Administrative selection regards the social case workers. The case worker does not randomly assign individuals to programs but does this on the basis of his/hers own discretion, which again is influenced by the policy and procedure of the administration. The problem of self-selection is expected to lean towards more employable people being selected to programs, while the administrative selection is more unclear. However, if the case worker's performance are measured by the employment outcomes of their clients, this could also lean towards more employable people being selected for programs (Aakvik et al. 2005).

Most of the evaluations are observational studies as experimental studies are more costly and difficult to conduct. There have been more experimental studies in recent years though. In their meta-analysis, Card et al. (2010) surveys 97 studies where nearly 10% of the studies were experimental studies. This allows them to compare results between the two types of studies after controlling for the nature of the programs. They find that there are not significant differences in the results of observational studies in comparison to those of experimental studies. They argue that this shows the robustness of the econometrical methods used to adjust for selection bias in observational studies.

### 3.1.3 *Adjusting for selection problems*

There are different methods of trying to overcome the selection problem, and there is not one "fix all" method. Which methods researchers apply, depends on several things but most importantly on the data available. Some of the most common methods being used are: instrumental variables, selection models, difference-in-differences and matching methods (Heckman et al. 1999; Blundell & Costa Dias 2000). Instrumental variables (IV) and selection models are best fitted for cross sectional data. Difference-in-differences needs data over time to work and can provide more robust estimates when dealing with panel data, which is cross sectional data over time. Matching methods can be used for both types of data, but the method is very dependent on good and detailed data to provide consistent and unbiased estimates. The different kinds of methods are supposed to adjust for the selection bias by comparing comparable people. Matching methods does this on the basis of the observable data, thus the heavy reliance on detailed data. The other methods work to compare comparable people on the basis of the unobserved characteristics. In the following sections, the papers that are reviewed are mostly based on matching methods or selection models. A method that is quite new and becoming more popular is the so called timing-of-events approach (Abbring & van den Berg 2003). This method is seen as very reliable, but it requires much computational power and it will not be much discussed here.

## 3.2 Results from the literature

### 3.2.1 *Meta-analyses*

Card et al. (2010) does a broad meta-analysis on LMPs in Europe. A meta-analysis is in short a statistical analysis of other studies. They use 97 studies conducted between 1995 and 2007 which contains 199 program impacts. They find that evaluation studies done over longer time periods seems to be more favorable than short term evaluations with regards to positive employment outcomes. The results show that several programs have a positive effect beyond the first year and have significant positive impact after 2 and 3 years. This is in a sense good news and maybe not unexpected as the goal of many programs is to increase the human capital of the participant, which should increase his/hers employability. In the comparison of the programs, the results showed that subsidized job programs had the worst outcome while training and job search assistance had the best. Overall the effects of programs range around no to some effect.

Heckman et al. (1999) reviews several microeconomic evaluations done in the U.S and Europe. The results vary from some to no positive effect. Government employment and training programs showed some effect, especially when applied to low-skilled workers. Further it is found that the different programs have very different effects on different groups with regards to demography and skill-level. Further, they report that youth specific programs have no impact or even negative impact on earnings in the U.S. The results on such programs are better in Europe, but mixed between some and little effect. The general conclusion though is that LMPs in the U.S. and Europe have modest effects at best when it comes to earnings.

Kluve (2006) also does a meta-analysis. The data set is constructed from available microeconometric evaluation studies carried out across many European countries. The data consists of 95 evaluation studies, which here gives 137 observations, as one observation corresponds to the evaluation of a particular training program. In the overall sample he finds that 75 studies (55%) found a positive effect while 62 (45%) did not. The programs that are evaluated are mainly applied on regularly or long term unemployed, which excludes those with reduced working ability. Furthermore, the results from the study are somewhat ambiguous as it there is no clear evidence that all LMPs are effective at increasing the likelihood of individuals getting employed. It is found that basic Service and Sanctions emerge as a promising measure. This is a category that consists of basic job search assistance such as counseling, monitoring and assistance, and with corresponding sanctions (on the social benefits) in case of noncompliance. Job training in the form of education, work practice etc. shows mixed to modestly positive effects.

### 3.2.2 *Scandinavian evaluations*

Frölich et al. (2004) does a study of how vocational rehabilitation affects the labor market outcome in Sweden. They use a matching estimator in their evaluation of the treatments. Further, they analyze a sample of long term sick people, who because of their sickness have got their working ability reduced. The sickness arrangement in Sweden is quite generous as it pays 80% of the previous income with a max ceiling, and the period is unlimited. Their data is provided by the Swedish labor administration. Data on long term sick individuals were randomly selected from 67 different local labor administration offices during a three year span in the 90's. After doing some shaving on the data set, they were left with 6287 sickness cases (individuals) were 3087 had participated on some sort of program. Their background data on the individuals are very rich including many socioeconomic variables as well as detailed medical history and status. Frölich et al. (2004) actually found that several rehabilitative programs decreases the chances for re-employment and reintegration compared to no programs. According to their study job training performed better than the other programs although no evidence is found that it performs better than no programs. Further, passive and educational programs performed the worst according to their study. So none of the programs showed any positive effect on employment and several actually showed negative effect. As this study only measures the effect of programs, which is shown to be negative, the costs of these measures can be added to the economic outcome from the programs. According to this study those programs are failures, at least from a labor market perspective. The negative effects are mostly explained by the lock-in effects that occur, which will be more discussed.

There are several differences across countries on how labor market policy is conducted. As mentioned, the Norwegian policy is constructed to aim the programs mainly on those with reduced working ability, often referred to as vocational rehabilitation (VR) programs (approx. 84% of LMP participants was registered with reduced working ability in September 2014). This is obviously not by coincidence, and is probably a strategic priority as Norway have had low unemployment compared to other European countries over several years, but have had high rates of sickness and disability absence. In Norway, spending on disability and sickness benefits amounts to around 5% of GDP which is the highest level of all OECD countries

(OECD 2013). Evaluation studies done in Norway reflect the fact that programs are mainly aimed at people with reduced working ability.

Arild Aakvik has done several studies on the effects of programs, especially the effect it has on those with reduced working ability. Aakvik (2001) uses a matching estimator to evaluate the effect of the Norwegian vocational rehabilitation (VR, i.e. the part of LMP that are directed towards those with reduced working ability). His data consists of 4416 people who received VR-benefits, where 2908 of those participated on a program (the different program types are not considered separately but considered in one category). The rest of the individuals make up the control group. VR-benefits are the precursor of WAA, such that the study is comparable to today's arrangement. Further, the data set consists of people who became registered as VR-clients during the year 1989 and he follows those people until 1993. The results suggest that programs has an overall significant positive effect. He finds that the average training effect is 6.3 percentage points, which means that the treatment (some program) increases the job probability with 6.3 percentage points. Further he finds that the different programs have a larger effect on the individuals with relatively lower job prospects to begin with. This implies that the programs should be aimed at those with relatively lower job prospects in the first place, but the results suggests that it may be the other way around. In the literature this is known as "cream-skimming" (Bassi 1984), and it implies that the case workers are choosing the ones with the best initial job prospects for program participation (will be discussed more thoroughly). Hence, the results from Aakvik (2001) may be promising but leave much room for improvement.

Aakvik et al. (2005) study people with reduced working ability, who receives (the precursor to) work assessment allowance. The data consists of 1924 randomly selected females who applied for training programs in 1989 and these females are followed through 1993. The different forms of programs are collapsed into a single category because of data limitations. Other variables include income, age, years of education and more. Further they use a selection model framework that matches the individuals on the unobservables rather than on the observables. A latent variable model is used to capture the unobservable characteristics. The results from Aakvik et al. (2005) suggests that program effects are negative, but that it becomes better for individuals with characteristics that predict lower job prospects in the first place. Again the "cream-skimming" problematic is discussed, as results points towards the presence of this. Aakvik et al. (2005) suggests that guidance and evaluations for NAV's case workers should be more thorough and be based on research rather than rule of thumb. The study suggest that employment gains would be achieved if the selection of program participants is more aimed towards those who are less employable.

Of the two Norwegian studies that have been reviewed, one showed promising results and the other did not. Further in Norway, Westlie (2008) also considers those with reduced working ability and he finds that both classroom training (public or by NAV) and wage subsidies gave positive effects on employment outcome. He also finds that program participation reduces the probability for disability pension. On the more negative side he finds that program participation increases the time of receiving social benefits due to the lock-in effect. Røed & Raaum (2006) analyses all regularly unemployed that received benefits and found positive transition rates to employment after labor market programs, but when adjusting for opportunity costs of the more lengthy time as unemployed the net effect of programs was close to zero. Dahl & Lorentzen (2005) found no effect from work programs, but some from effect from training programs. Their study compromises of all the beneficiaries as well, including regularly unemployed. The results are thus ambiguous and there is hard to argue that there is clear cut evidence in regards to Norwegian LMPs.

### 3.2.3 *Lock-in effect*

The lock-in effect is an important driver for bad results in the evaluation literature of LMPs. When people are unemployed and have reduced working ability they are receiving social benefits which are usually substantially less than what they had in a regular job. So during the time with social benefits people at least have monetary incentives to apply for jobs if they are able to. When they participate on a program it is expected that eventual job searching decreases, as people become "locked in" to a program. If the goal is regular employment such measures as subsidized jobs, which would include job practice and wage subsidies, are locking the participants to the program and reducing their job searching effort (van Ours 2004). Also education which normally is a lengthy program is expected to have considerable "lock-in" effects, as is found in Norway by Westlie (2008). Frölich et al. (2004) argues that their negative results are explained by the fact that the individuals are decreasing their job searching during programs. It is clear that if an individual is declared to a program, the job search intensity is expected to decrease, as the individual are committed to a program. Lock-in effects are often discussed in the literature as the main downside of applying labor market programs.

### 3.2.4 *Cream skimming*

In evaluation studies the presence of cream skimming is sometimes discussed as a potential unfortunate outcome. In the Norwegian case, cream skimming would be related to NAV selecting the most employable people for program participation. The evaluation of performances within NAV (and also within other country's labor administrations) are often based on the outcome, which is how many people get employed after the programs. If this is the case, the NAV case worker who assigns people to programs, have incentives to assign the individuals that are more likely to get regular work. This is certainly a risk if case workers performances are measured by the employment rates of their clients. This may lead to assigning the most "employable" individuals to programs rather than those who might need it the most and benefit the most. The opposite of cream skimming is bottom fishing and it implies that the least employable are assigned for programs. However, those least employable are not necessarily always the most "treatable" (gain the most from program participation).

In Norway, signs of cream skimming is found in evaluation studies as discussed and it is also found that those least employable are most treatable (Aakvik 2001; Aakvik et al. 2005). In the U.S., Heckman et al. (2002) study the performance standards and behavior of the case worker explicitly. Their results suggests that there is little effect from cream skimming. However in Sweden, Skedinger & Widerstedt (2007) finds that there seems to be cream skimming when considering the selection process to an assisted work program.

# 4 Method

In this observational study only one regime (participation or nonparticipation) and one outcome (work or no work) is observed for each individual, which means that for one individual one can only observe the outcome after program participation or nonparticipation. It is not possible to observe the outcome after both regimes. When selecting the framework in this case, there are three things that become crucial. First, the evaluation only allows for one regime and one outcome to be observed. Second, the data is observational rather than experimental. Third, although the data has a time dimension and is collected over time, it is treated as cross-sectional data in the analysis. Given the characteristics of the evaluation and data, some of the proposed methods are propensity score matching or a selection model (Heckman et al. 1999; Blundell & Costa Dias 2000). A selection model is the method of choice as the propensity score method relies on very detailed data to produce reliable estimates.

### 4.1 The model

The model is a selection model with an endogenous switching equation. This means that a system of three equations is estimated where one determines the probability of the outcome with the treatment, the other determines the probability of the outcome without the treatment and the third determines the probability of the treatment (the switching equation). The selection, or the determination of the probability of the treatment, happens inside the model, meaning that the switch is endogenous. Such a switching equation or regression that estimates the probability between two regimes was termed by Quandt (1972). Further, similar methods are proposed by Heckman (1979) and its application to microeconomic evaluations are proposed. The latter became known as Heckman's selection model and has become a known tool to cope with selection bias in evaluations (Blundell & Costa Dias 2000; Dutoit 2007). Heckman's selection model have been generalized and customized in many ways. This thesis will rely on a command written in Stata for such questions, a switching probit model. The command and related theory is described in Lokshin & Sajaia (2011). This method is also similar to the one used by Aakvik et al. (2005). It is a system of three equations consisting of one equation for the selection and two for the outcome, where one is for the treated case and the other for the non-treated. Using this framework, one can estimate a latent variable that strives to describe an individual's unobservable characteristics. This can be interpreted as an individual's propensity to work. If this latent variable or propensity is precisely estimated, one can match individuals and isolate the effect of the treatment, that is, one can compare comparable people such that the isolated effect of program participation can be measured directly. This allows for estimating treatment effect on the treated, treatment effect on the untreated and treatment effect on a randomly picked individual. Furthermore, all the dependent variables are binary variables; they can only take the shape of zero or one. Since this is the case, a probit model that uses maximum likelihood estimation for the equation system is the method of estimation. The model setup is as follows.

(1) $$T_i = 1 \quad if \quad \gamma Z_i + \mu_i > 0$$

$$T_i = 0 \quad if \quad \gamma Z_i + \mu_i \leq 0$$

(1) is the selection equation. Here, $T_i$ determines if person $i$ participate on a program, it is equal to one if the person participate and zero otherwise. There are *n* individuals in the sample so $i = 1, 2, \dots, n$. $\gamma$ is a vector of parameters, $Z_i$ is a vector of variables that explains potential program participation and $\mu$ is the error terms also interpreted as the unobserved factors that explain potential program participation. Together this determines whether the individual will participate on a program or not.

$$(2) \qquad y_{1i}^* = \beta_1 X_{1i} + \epsilon_{1i} \quad y_{1i} = I(y_{1i}^* > 0)$$

$$(3) \qquad y_{0i}^* = \beta_0 X_{0i} + \epsilon_{0i} \quad y_{0i} = I(y_{0i}^* > 0)$$

(2) and (3) are the two outcome equations. Here, $y_{1i}^*$ and $y_{0i}^*$ are the latent variables, that is the unobserved propensity for working. $X_{1i}$ and $X_{0i}$ are vectors of observed variables that explains individual $i$'s propensity for working. $\beta_1$ and $\beta_0$ are vectors of parameters and $\epsilon$ are the error terms. $y_{1i}$ and $y_{0i}$ are the observed outcome, which is realized by the unobserved latent variable.

From the model, the individual's potential program participation is determined by (1). Further, from (2) and (3), the latent variables $y_{1i}^*$ and $y_{0i}^*$ are the unobserved determinants of the outcome. *I* is just the indicator function $\{0, 1\}$ so that the outcome ($y_{1i}$ and $y_{0i}$) is binary. $y_{1i}$ and $y_{0i}$ is observed and take the form of one if the outcome is work and zero otherwise. $y_1$ is the outcome conditional on the treatment, that is for those who participate on program(s). $y_0$ is the outcome conditional on non-treatment, that is for those who do not participate on program(s). Hence, the observed outcome is:

$$y_i = \begin{cases} y_{1i} & if\ T_i = 1 \\ y_{0i} & if\ T_i = 0 \end{cases}$$

The empirical specification for equation (1) is that $Z_i$ is a vector of personal characteristics and administrative statuses that influences the probability for individual $i$'s potential program participation. As has been mentioned before, both self-selection and administrative selection are expected to influence potential program participation. The dataset at hand allows this evaluation to implement a bit of both elements in the selection equation, although mostly personal background information. This personal background information is first and foremost expected to influence the degree of self-selection (but also administrative selection), while administrative statuses are expected to influence mainly the administrative selection. $\gamma$ and $\mu_i$

are parameters and error terms (the unobserved factors) respectively, and they are found by estimating the model. Together, equation (1) determines the individual's probability for program participation.

The empirical specification of equation (2) is that $X_{1i}$ is a vector of characteristics that is expected to influence individual $i$'s probability of working, conditioned on, that individual $i$ has participated on program(s). Through estimation one finds the vector of parameters and error terms; $\beta_1$ and $\epsilon_{1i}$ respectively. The dataset allows for much background information, such as work background and education, in the vector of characteristics. By doing the estimation one finds $y_{1i}^*$, the propensity for individual $i$ working based on the information available. The outcome $y_{1i}$ is potentially realized through the estimated latent variable. The specification of equation (3) is exactly the same as for (2), but conditioned on non-treatment, rather than treatment.

(4)
$$\Omega = \begin{pmatrix} 1 & \rho_0 & \rho_1 \\ & 1 & \rho_{10} \\ & & 1 \end{pmatrix}$$

(4) is the correlation matrix of the error terms; $\mu, \epsilon_0$ and $\epsilon_1$. The underlying assumptions of the model are that the error terms are jointly normally distributed with a mean-zero vector. Also it is assumed that the correlation between $\epsilon_0$ and $\epsilon_1$ is equal to one ($\rho_{10} = 1$ ) as those error terms cannot be observed simultaneously due to the fact that $y_{0i}$ and $y_{1i}$ cannot be observed simultaneously. $\rho_0$ and $\rho_1$ are the correlation between $\epsilon_0, \mu$ and $\epsilon_1, \mu$ respectively.

The simultaneous system of equations (1)-(3), constrained by the underlying assumptions, is then estimated with a log likelihood function. The log likelihood function is specified as follows.

$$\log(\mathfrak{I}) = \sum_{T_i \neq 0, y_i \neq 0} \log\{\Phi_2(X_{1i}\beta_1, Z_i\gamma, \rho_1)\} + \sum_{T_i \neq 0, y_i = 0} \log\{\Phi_2(-X_{1i}\beta_1, Z_i\gamma, -\rho_1)\}$$
$$+ \sum_{T_i = 0, y_i \neq 0} \log\{\Phi_2(X_{0i}\beta_0, -Z_i\gamma, -\rho_0)\} + \sum_{T_i = 0, y_i = 0} \log\{\Phi_2(-X_{0i}\beta_0, -Z_i\gamma, \rho_0)\}$$

Here, $\Phi_2$ is the cumulative function of a bivariate normal distribution. The model is a probit model which explains the presence of the cumulative distribution function. Further, the maximum log likelihood estimates of the parameters are computed over the four possible outcomes which are: treatment/work, treatment/no work, no treatment/work and no treatment/no work. The log likelihood function will not be dealt with in more depth as this will be estimated by a computer, but the smoothness and rate of convergence of the model will be of interest in the estimation.

### 4.2 Estimated parameters

The computed model will give insight to which of and how the independent variables affects the dependent variables. The three dependent (and also binary) variables in the model are program participation, work conditional on program participation and work conditional on nonparticipation. In other words the estimation will show how the vector $Z_i$ will affect potential program participation and how the vectors $X_{1i}$ and $X_{0i}$ will affect the probability of work. From the computed model it is also possible to get numerical estimates that shed light on the main question; what is the effect of program participation on the individual's job prospects? The model will enable the estimation of the treatment effect on the treated (TT), treatment effect on the untreated (TU), treatment effect (TE) and marginal treatment effect (MTE).

$$(5) \qquad TT(x) = \Pr(y_1 = 1|T = 1, X = x) - \Pr(y_0 = 1|T = 1, X = x) = \frac{\Phi_2(X_1\beta_1, Z\gamma, \rho_1) - \Phi_2(X_0\beta_0, Z\gamma, \rho_0)}{F(Z\gamma)}$$

(5) is the treatment effect on the treated (TT). F is a cumulative function of the univariate normal distribution. The treatment effect on the treated for some given characteristics $x$, are the difference between the predicted probability of a program participant getting work, and the predicted probability of that individual getting work given that he/she had not participated on a program. It is thus the expected value of $(y_1 - y_0)$ given treatment and some characteristics $x$. The interpretation of this is that a positive value of TT means that the probability of getting work (for some individual) is higher with program participation than without, and vice versa if it is negative. The TT is calculated for all program participants, but the individual specific estimate is not necessarily very informative so for analyzing purposes, means and distributions are more interesting (Aakvik et al. 2005).

(6) $$TU(x) = \Pr(y_1 = 1|T = 0, X = x) - \Pr(y_0 = 1|T = 0, X = x) = \frac{\Phi_2(X_1\beta_1, -Z\gamma, -\rho_1) - \Phi_2(X_0\beta_0, -Z\gamma, -\rho_0)}{F(-Z\gamma)}$$

Equation (6) is the treatment effect on the untreated. This estimates the effect the treatment would have had on the untreated if they had been treated. That is to say that it estimates the effect program participation would have had on the nonparticipants had they been selected to program(s). It has the same attributes as TT, as it is estimated for some individual with characteristics $x$, and it is best interpreted as means or distributions.

(7) $$TE(x) = \Pr(y_1 = 1|\, X = x) - \Pr(y_0 = 1|\, X = x) = F(X_1\beta_1) - F(X_0\beta_0)$$

Equation (7) is the treatment effect and it is the estimated treatment effect on a randomly selected individual with some given characteristics $x$. In other words it is just the expected value of $(y_1 - y_0)$ given $x$. One can interpret the treatment effect as the effect from program participation when averaging over both the treated and the untreated group. So if hypothetically, there was no selection bias, that is, if the program participants and nonparticipants were selected completely at random, then the treatment effect, the treatment effect on the treated and the treatment effect on the untreated would all be equal.

As mentioned the TT, TU and TE are best interpreted as means or distributions. For this reason, emphasis will be put on presenting the mean of TT($x$) for the whole treated group when presenting the results. The same goes for TU and TE, but for the whole untreated group and the whole group respectively. Another interesting feature is the consideration of estimated parameter means for subgroups. That is the estimated treatment effects for groups that share some characteristics, for instance old age. By looking at these means it is possible to compare treatment over different subgroups. For instance one could consider the average treatment effect on the treated for the subgroup $n_k$:

(8) $$ATT(n_k) = \frac{1}{n_k}\sum_{i=1}^{n_k} TT(x_i)$$

Here, $n_k$ could be for instance female, so that $ATT(n_k)$ is the average treatment effect on all the females who participated on programs. Such a parameter is interesting and could be compared to that of the whole group (in this case the TT mean of the treated group) to see if there are any differences for this particular group. It could shed some light on which groups benefit most and which groups benefits the least from treatments. Also it could be interesting to compare to other relevant subgroups (in this case males). Equation (8) here, also works for consideration of TU and TE for different subgroups. TT is just swapped with TU or TE.

The general interpretation of the estimated parameters would be that, if the TT and the TU were similar, there would be little selection bias. The ones who received treatment and the ones who did not would have similar effects from program participation. If the TT were greater than the TU, this would indicate that the treated group benefits more from program participation then the untreated group. In other words this would mean that the selection process is biased towards those who benefit more from program participation.

The last estimated parameter of interest is the marginal treatment effect (MTE), introduced by Björklund & Moffitt (1987) and developed by Heckman & Vytlacil (1999, 2005).

$$\begin{aligned} (9) \qquad MTE(x,\bar{\mu}) &= \Pr(y_1 = 1|X = x, \mu = \bar{\mu}) - \Pr(y_0 = 1|X = x, \mu = \bar{\mu}) \\ &= F\left(\frac{X_1\beta_1 + \rho_1\bar{\mu}}{\sqrt{1-\rho_1^2}}\right) - F\left(\frac{X_0\beta_0 + \rho_0\bar{\mu}}{\sqrt{1-\rho_0^2}}\right) \end{aligned}$$

The marginal treatment effect is the expected treatment effect on employment outcome conditional on observed characteristics $x$ and conditional on unobserved characteristics $\mu$. It follows from the model that $\mu$ can be interpreted as the unobserved characteristics that influence the decision for program participation (the selection process). As seen in (1), program participation are determined by both observed ($Z$) and unobserved ($\mu$) characteristics. The estimated average MTE is the treatment effect on some individual, given $x$, that are on the margin of program participation and nonparticipation given the unobservables. It may be more interesting though to evaluate the MTE for different values of $\mu$ as it ranges from those who are least likely to participate on program(s) to those who are most likely. For small values of the normalized $\mu$, the MTE are the estimated treatment effect for individuals whose unobservables makes them less likely to participate on programs. For large values of $\mu$, the MTE is the treatment effect on those who are more likely to participate on programs.[1] This distribution gives insight to the selection process and is an important feature of the evaluation.

The model should be able to give insight to certain mechanics. First, how the individual's observed characteristics ($Z$) affects their potential selection to programs ($\gamma$). Secondly, how the individual's observed characteristics ($X_1$) affects their probability for work if program participation ($\beta_1$). And third, how the individual's observed characteristics ($X_0$) affects the probability for work in the case of nonparticipation ($\beta_0$). Furthermore, considering estimated parameters from the outcome equations ($\beta_1$ and $\beta_0$) along with the consideration of treatment effects for subgroups can shed light on how variables affect the outcome differently in general and over different subgroups. The most important feature is the TT, as its goal is to estimate the program participant's job probability compared to those same participant's job probability if they had not participated, even though both outcomes cannot be observed.

# 5. Data

## 5.1 Data source

The data is delivered by NAV which is the source itself for the employment and social benefit data. NAV also has information on the background of the individuals including such things as age, gender, education and work background. The dataset consists of all the people that started receiving work assessment allowance between and including March 2010 and December 2011. This was 91 013 people and they were all who started receiving WAA during this time stretch. This means that those who were receivers during this time stretch but started before is not included. The observation period is up to and including November 2014. Beneficiaries are (initially) entitled to WAA up to four years which means that many are still receiving as their entitlement can stretch to 2015 if they started in 2011. This creates some problem in the data which will be discussed.

[1] The fact that an increasing $\mu$ is related to higher probability for participation follows from the model setup (equation (1)). This could have been reversed, depending on the model setup. The statistical program at hand will provide the reverse curve of what is described here.

## 5.2 Adjustments and weaknesses in the data

Missing data, necessary refinements and information making individuals irrelevant to the questions of interest led to rounds of substantial downward revisions of the sample. At the point of the main analysis the data consisted of 32 924 individuals. The reason for the reduction and which observations are left out will be discussed as this will impact the results. The individuals that are left out from the main analysis are either because of missing data on critical variables (which is believed to be random), the need to limit the time span or individuals that are not of interest for the questions at hand.

### 5.2.1 *Special cases*

There are some individuals that are considered to not be of interest for the evaluation of LMPs. One of the variables contains the goal or the objective for each individual at the time they start receiving WAA. Some of these individuals had the objective “permanent facilitation”, which in reality means that these are individuals that are not expected to have regular jobs again, but rather receive benefits as part of their livelihood. This could be for instance “permanent wage subsidies” and hence they are left out from the data. Individuals that are on full age pension at the end of the observation period are also dropped. Also those who are born before 1952 are dropped from the data. These are people who were 60 years or older in 2011, when they latest started WAA. It is considered that those are people who are close to retirement and the WAA is in those cases considered more of a waiting room for age pension. Further, individuals that are emigrated or passed away are dropped from the data. Together 10 336 individuals were left out due to the above reasons.

### 5.2.2 *Missing data*

Missing values on critical variables is another reason for the downward revision. The data contains information such as age, gender, municipality of residence, education, work background and objective of WAA. These variables are essential to the switching equation and the explanation of the selection. The goal is to compare comparable people and those variables play an important role in just that. Education and work background are the variables with most missing values and this is because NAV do not have complete data on this, as individuals fills out this information by themselves. It is a weakness in the data that individuals are left out because of missing values. An alternative is to give the missing data some value, for instance “unknown”. This has been tested and the results were similar. The literature is highly concerned with the quality of the data in evaluation studies (Heckman et al. 1999), and it is thus assessed that dropping individuals that have missing data is the most consistent solution. Frölich et al. (2004) for instance also decided to: "ignore individuals with missing values on important variables". It remains a weakness though that many individuals are dropped, but the alternative; to run the model with vast amount of missing data, seems undesirable. The missing data that is because of administrative routines and possible slippages in the NAV system is not expected to be systematic. Some data is though expected to be missing because of incomplete self-reporting from the unemployed. This is on the other hand not necessarily completely random, as some are maybe more likely or more willing to fill out self-reporting schemes. However, there seems not be any good solution to this, so the model will have to be estimated with the data available. For this reason 25 620 individuals were left out for the analysis.

#### 5.2.3 *Conditioning on exit from WAA*

Ideally the data should have been over a longer time span and further back in time. Since the data is so close to the present day it was decided to cut of the data at a certain time. The initial data consisted of all the people who started receiving WAA from March 2010 up to and including December 2011. The reason for this is that there was a reform regarding social benefits that was initiated March 1st 2010, and that it would be interesting to evaluate the arrangement after the reform and also because of practical reasons. Since the WAA is initially up to four years for beneficiaries, those who started in 2011 are still entitled into 2015, whereas the data is up to November 2014. This means that many individuals in the data were still receiving WAA at the last observation and the effect of program participation or nonparticipation may not yet be observed.

The data for the main analysis was thus conditioned on exit from WAA no later than June 2014. This involves leaving out 22 133 individuals from the analysis. The advantage of this is that it ensures that one can observe the effect of program participation or nonparticipation at least 6 months after finishing WAA, and also after finishing all programs. If this was not done the data would contain individuals that are still on programs. The disadvantage though is that this does something with the sample that disassembles its similarity to the population, which in turn can cause biased estimates. Those who are left out may be long term recipients of social welfare and may be far away from regular work. On the other hand those might be individuals that are on long term programs such as education and may be close to regular employment when programs are finished. It was clear from the data that, among those who were dropped; a high proportion were program participants and a very low proportion were regularly employed at the last observation (obviously). The choice of limiting the time span was made with the evaluation question in mind. Since the goal is to evaluate the programs this should be done after participants and nonparticipants have participated or not and their labor market outcome is observed.

However, results from the model were similar even if the sample is conditioned on exit from WAA or not. Variables shows similar signs and explanatory power, and the treatment parameters are also similar. It remains a weakness though that the timespan is short, especially in regards to that it does not allow for evaluating more long term effects.

## 5.3 Variables

### 5.3.1 *Explanatory variables*

Among the explanatory variables that are used are the background variables: age, gender, relationship status, country of birth, size of municipality, work background and education. This is in line with what is suggested and used in the literature for evaluation studies. Blundell & Costa Dias (2000) for instance suggests all the above except work background. Frölich et al. (2004) uses all the above, including health information, while for instance Aakvik et al. (2005) uses several of the same including education and work background. Furthermore, there is also a variable regarding the objective (or goal) for the unemployed at start of WAA which will be discussed, and along with the background variables makes up the explanatory variables. All the explanatory variables are created as dummy variables.

The variables of choice are selected because they are believed to influence both the individual’s potential program selection and the individual’s prospects for regular work. Age is expected to influence the selection equation negatively, that is to say, the probability for program participation is expected to be decreasing with higher age. Younger persons are expected to have a higher probability for program participation and this is because there is a clearly expressed strategy from NAV to prioritize young people for program participation. Gender, relationship status and possible immigrants are interesting background variables that could affect both selection and outcome. The size of the municipality of residence is added to see whether there are any patterns between urban and less urban places of residence. This could have an effect on both the outcome and maybe more so to the selection as there are probably differences among NAV offices with regards to capacity, resources and access to potential program slots. Work background and education are important background variables as they are expected to explain both outcome and selection. It is expected that those with less “skilled” work backgrounds have lower job prospects than those who have more “skilled” work background. Furthermore it is expected that those with higher education have better job prospects than those with lower education.

Age is divided into dummy variables, rather than being a continuous variable to allow for comparison between age groups, and also to comply with restrictions from the privacy authorities regarding the project. More precisely age is divided into five categories: 18-24, 25-31, 32-41, 42-51 and 52-60 years old. These are the individuals’ respective age at the end of the year 2011, and the year is picked because everybody has started receiving WAA at the end of 2011. The age goes up to 60, as those older than that were ignored, as discussed. The youngest persons in the data were 18 years old at the end of 2011. Male and female are dummy variables. Relationship status is also divided into dummy variables and is categorized as: married, unmarried and divorced/separated. Initially there was also a status in the data for those who were widowed. Since this was a tiny share, this status was incorporated under unmarried. A dummy variable for individuals born in another country than Norway is also included. The size of the municipality of residence is represented as three dummy variables. It is a categorized as small (<10 000 residents), medium (10 000 – 50 000 residents) and large (>50 000 residents). Further both work background and education are categorized according to Statistics Norway’s standards and then collapsed into fewer categories. Statistics Norway has nine main categories for occupational classification; those are here collapsed into five categories that make up dummy variables. The five categories are as follows: (1) manager/executive or academician; (2) office and administration; (3) sales and service or care; (4) craft, construction, transportation, fishing or farming; (5) cleaners, assistants etc. Further, Statistics Norway has seven main categories for classifying education. Here this is collapsed into four categories and they are as follows: (1) university or college degree; (2) three years of secondary school and/or technical vocational school; (3) one or two years of secondary school; (4) primary school.

The objective at start of WAA is also included as an explanatory variable, although this is not as obvious as the other background variables. This variable allows for the creation of three dummy variables and they are as follows: “retain work”, “obtain work” and “increase participation”. This classification of the unemployed or objective for the unemployed is set by

NAV in consultation with the unemployed. This status or objective is set at start of WAA (although updated along the way, but here only objective at start is used), and is expected to capture some of the health status and capabilities of the individuals. Those who receive WAA has already been through an assessment of their working ability and it has been established to be reduced. Also, along with this, there is established an objective or goal for the individual and this is what is used here. The objective that is set reflects the capabilities and health of the individual and what is considered a realistic objective for the individual to achieve. NAV's guidelines for consideration of an individual's objective mainly includes user's wishes, health conditions and capabilities. Therefore this variable is considered as a proxy variable for health and motivation in the analysis.

The three objectives are actually quite descriptive (ranking not by coincidence):

- Objective retain work are people who cannot work at the time because of their reduced working ability but have jobs and has an ultimate goal of returning to it.
- Objective obtain work are people who do not have jobs and the goal is to find and get one (the clearly biggest group).
- Objective increase participation are set for people who are struggling more than the other two groups. In fact the goal in the first place might be to increase social participation and comfort rather than work participation. An objective or goal of regular work is not clearly stated (as with the other two).

The employment rates of the groups reflects the above ranking. The category retain work shows far better employment rates as well as far lower program participation rates than the other two and thus underlining that these individuals probably has more temporary struggles. The category increase participation have the clearly lowest employment rates and the highest disability pension rate at the end. The variable objective at start of WAA is thus considered as a proxy variable for health and capabilities. This is not detailed health information but it is considered to capture some information in a broad sense, as it has been discussed that there are significant differences between the groups. The dataset does not include any explicit information on health as it is sensitive data and is thus difficult and costly to obtain. Objective at start of WAA will serve as a proxy variable for health and is expected to capture some of the characteristics that the other background variables miss.

The variables that are picked as explanatory variables are picked because of its expected relevance to the outcome and selection. It is also chosen among the data available. The amount of background information is good and should make for a decent analysis. Although there are not any explicit information on health background which is something that would have been valuable for the analysis.

### 5.3.2 *Instrumental variable*

To make the estimates more robust, an instrumental variable is introduced to the selection equation. Adding an instrument should make the estimates more robust to alternative functional assumptions (Lokshin & Glinskaya 2009). Furthermore the use of instruments is suggested in particular for the model used here as it helps identify the model (Lokshin & Sajaia 2011). The desired instrumental variable should be able to affect the selection, but not directly affect the job outcome. The instrumental variable of choice is the share of nonparticipants in the municipality. It is calculated as the share of nonparticipants in the individual's municipality (for interpretation purposes it is created such that it numerically ranges from 0 to100). It thus become larger when a smaller share of the sample is participating on programs, in that individual's municipality. It can be seen as a measure for the degree of rationing and when it increases less people participate on programs and it thus decreases some person's probability for program participation. This is inspired by and is more or less the same as used by Aakvik et al. (2005). It is expected to affect (decrease) the probability of program participation and is not expected to affect the probability for regular work. This is because the degree of rationing is not expected to relate to the individual, it is more a result of external factors.

### 5.3.3 *Employment outcome*

There are two outcome variables which both are dummy variables, employment and program participation. In the model, employment is defined only as regular work and not partly work and partly disability pension for instance. Such arrangements means that people are still receiving benefits and is not considered regular work here. The most common outcomes for the sample are listed in table 4, to give a thorough overview of this. These are the statuses November 2014. There are also several other outcomes as registered by NAV which are not included in the table. For instance, regularly unemployed (3%) did not even "make the list".

*Table 4. The observed final outcome of the sample (percentages and frequency of the sample, 32 924 individuals).*

| Outcome (status November 2014) | Frequency | Percent |
|---|---|---|
| Regularly employed | 12 717 | 39% |
| Disability pension | 5727 | 17% |
| Reduced working ability (which means that they probably is back on receiving WAA) | 4709 | 14% |
| Not registered in NAV's registers (which means that they may be temporarily withdrawn from the labor force) | 3610 | 11% |
| Partly disabled/partly working (which means that they are receiving graded disability pension) | 1871 | 6% |
| Partly reduced working ability/partly working (which means that they are probably receiving graded WAA) | 1782 | 5% |

5.3.4 *Program participation and control group*

Program participation is also a binary variable, such that individuals are either participating on programs or not. Program participation considers any program independent of length, so the presence of some program during some time over the observation period is considered program participation. From the way the selection variable is presented, it follows that those who participate on a program, independent of program and length, make up the treatment group. In that regard the remains makes up the control group. It was decided that the remaining individuals in the dataset that do not participate on programs, are all part of the control group. This means that the treated are compared to all the remaining individuals and that the whole group of individuals are included in the main analysis (this is after the omission of individuals as discussed in chapter 5.2, leaving 32 924 for the main analysis).

The data allows for separation between the different programs into broad categories. The programs that will be looked at in more depth are training, work practice and assistance. As was previously discussed these are the most used programs. Clarification will not be looked as NAV does not see this as an active labor market measure in that sense. Clarification is, as mentioned, short duration measures that are supposed to help clarify the possibilities and capabilities of the participant. Even if the programs are divided into categories, those categories are relatively broad in the sense that the categories of the dataset have several subcategories. Training consists mainly of ordinary public or private education, but also so called labor market training. Labor market training is a combination of practical and theoretical training at a work place. The data does not allow for the separation of these two. Furthermore, work practice is divided between work practice in ordinary environment and work practice in protected environment and these are quite different in terms of nature and who they are aimed at. Work practice in protected environment is, as the name implies, aimed at people who needs a protected environment and may be quite a long way from regular work, while work practice in ordinary environment are at regular work places. The inability to separate these two in the dataset is a weakness as those programs ideally should have been considered separately. Although, many of those who participated on work practice in protected environment are expected to have been omitted when those with status permanent facilitation were left out, as discussed previously. Lastly, assistance is divided into several subcategories but those subcategories are quite similar in nature such that considering the program as one seems reasonable.

It is not unusual for individual's to participate on more and maybe several different programs. This may be situations where individuals start with clarification and then go onwards to participate on another program. The method of choice when assigning *one* program to each individual in the analysis was to go for the longest lasting program per individual in terms of months. In the data, individuals were registered for each month they participated on a program. This was collapsed on the individuals by the most frequent program, such that there is one program (that which lasted the longest) per individual and the individuals are counted once. This was considered the best alternative as it captures the "main program" for each individual.

### 5.4 Descriptive statistics

Thus the main analysis is carried out with 32 924 individuals. Table 5 presents some descriptive statistics on the main variables and some other interesting variables (may not sum up to one because of rounding errors). The data are conditional on exit from WAA no later than June 2014 as discussed.

*Table 5. Descriptive statistics. Variables of interest (means).*

| | LMP participants | Nonparticipants |
|---|---|---|
| Number of persons | 18 339 | 14 585 |
| Employment November 2014 | 0.36 | 0.41 |
| Employment 6 months after exit WAA | 0.32 | 0.39 |
| *Objective at start of WAA* | | |
| Retain work | 0.11 | 0.35 |
| Obtain work | 0.86 | 0.60 |
| Increase participation | 0.03 | 0.05 |
| WAA duration | 666 days | 478 days |
| Program duration | 13 months | |
| Female | 0.50 | 0.58 |
| Male | 0.50 | 0.42 |
| Year of birth | 1973 | 1968 |
| Age 52 – 60 | 0.13 | 0.28 |
| Age 42 – 51 | 0.24 | 0.29 |
| Age 32 – 41 | 0.29 | 0.25 |
| Age 25 – 31 | 0.22 | 0.12 |
| Age 18 – 24 | 0.13 | 0.06 |
| Born in foreign country | 0.13 | 0.12 |
| *Relationship status* | | |
| Unmarried | 0.54 | 0.40 |
| Married | 0.29 | 0.41 |
| Divorced/separated | 0.17 | 0.19 |
| *Place of residence* | | |
| Small municipality (<10 000) | 0.25 | 0.27 |
| Medium municip. (10 000 – 50 000) | 0.43 | 0.39 |
| Large municipality (>50 000) | 0.32 | 0.34 |
| *Work Background* | | |
| Manager or academician | 0.08 | 0.13 |
| Office and administration | 0.18 | 0.24 |
| Sales, service or care. | 0.34 | 0.31 |
| Craft/construction, transportation etc. | 0.29 | 0.23 |
| Cleaners, assistants etc. | 0.11 | 0.08 |
| *Education* | | |
| Primary school | 0.13 | 0.14 |
| 1 or 2 years of secondary school | 0.36 | 0.29 |
| 3 years of secondary school | 0.30 | 0.29 |
| University/college degree | 0.21 | 0.28 |
| Share of nonparticipants in municip. | 0.45 | 0.47 |

About 56 percent of the sample participated on a program some time during their spell on WAA, so it is clear that programs are actively used towards people with reduced working ability. It is also a well-balanced proportion in regards to the analysis. From the descriptive statistics it is clear that the nonparticipants has better employment rates both six months after exit from WAA and also at the end of the observation period (November 2014), and quite considerably with 7 and 5 percentage points respectively. Emphasis throughout this thesis will be on the employment outcome at the end of the observation period so notice is taken of the higher employment rate for nonparticipants of 5 percentage points.

Not surprisingly, program participants have on average longer WAA spells. On average program participants stays about 6 months longer on WAA. For the program participants, the average time spent on programs over the whole observation period is 13 months. This is regardless of participating on several programs and spells of nonparticipation in between.

As younger people are a priority for program participation there are no surprises that the share of young program participants is substantially higher than the share of young nonparticipants. Further a substantially higher share of the unmarried participates on programs while the married does not. Among the nonparticipants there is a clearly higher share of females at 58% than males, but for some reason the gender shares are equal among the participants. Drawing from this it is clear that there are more females in the dataset, but males are for some reason as represented as females among the program participants. Work background shows some pattern of lower participation rates for the more skilled background and vice versa. Looking at education shows less of a pattern. There is a higher rate of one or two years of secondary school among the program participants than among the nonparticipants. There is also a lower rate of university degree among the participants than the nonparticipants. The descriptive statistics give some insight to the data that are being dealt with, but does not provide any causality. At least it does not show any surprises when moving forward.

# 6 Results

This chapter will present the results from the analyses. In this thesis four models are considered. The first is the main analysis which is the most emphasized. This is the model that contains the whole sample (32 924 individuals), and any program is considered treatment. The other three considers the three programs training, work practice and assistance as single treatments leaving out those individuals who not participates on the specified program of interest. A consideration of the different programs are carried out since it is interesting to evaluate which programs works for whom.

Weight will be given to the explanation and interpretation of both the selection and the outcome equations. Further, estimated parameters of interest will be presented and suggested interpretations will be discussed. Focus will be on the treatment effects from the main analysis, both for subgroups and the whole sample. The estimated parameters from the analyses of the different programs will also be presented and discussed.

## 6.1 Main analysis

### 6.1.1 *Selection to programs*

In table 6 the coefficients of the selection equation from the estimated model are presented. These are the probabilities of being selected for program participation and how the variables affects this. Age affects the probability for program participation negatively, as expected. In table 6, the oldest age group (52 – 60 years) are the reference group and compared to that, all other groups have significantly higher probability for program participation. It also shows that there is an increasing probability for program participation the younger one gets. The youngest group have the highest coefficient and then it decreases in order with the age dummies. The results are in line with NAV’s priorities of young people for program participation. The age dummy variables are all statistically significant at a 1 percent significance level.

According to the model, males are more likely to be selected for programs than females. The reason for this is unclear. One suggestion might be that females are, more often than males, single parents and sole providers. Those who are sole providers for their children, have entitlements to WAA even if they not necessarily meet all the regular criterions, this is stated in NAV’s guidelines. This may cause more females to receive WAA than males. From the descriptive statistics it was clear that females are more represented among the nonparticipants, while there were an even share of males and females participating on programs. For there to be an even share of males and females that participates on programs there must such that males are more likely to be selected to programs. Another thing that might relate is that males are highly represented (87%) among those with work background from craft/construction etc. which is a group that is relatively likely to participate on programs.

Furthermore, those who are unmarried are more likely to participate on programs than those who are married. This may be linked with age as those who are younger are more likely to be unmarried. On the other hand unmarried and divorced are not significantly different, and those who are divorced are not expected to be young. Both the objectives retain work and increase participation, decreases the probability of program participation significantly in comparison to obtain work. This means that the main group (obtain work) and what is considered the second best health status are most likely participants. With regards to retain work this is not surprising, as NAV does not prioritize those who have jobs and have the best health status. On the other hand is it more surprising that the objective increase participation gives lower probability for program selection than obtain work, as one might expect those who needs the most assistance as more likely program participants.

Those who are born in a foreign country are significantly more likely to participate on a program than the Norwegian born. In their stated priorities NAV does not say anything about prioritizing immigrants among those with reduced working ability, but they do have a policy of prioritizing immigrants among the regularly unemployed. The results here suggests that this also applies to reduced working ability, to some degree, although not stated as a goal. The presence of a large municipality decreases the probability of program selection in comparison to small and medium sized municipalities. The reason for this is not clear, but may have to do with available resources in the respective municipalities.

When it comes to education, those with one to three years of secondary school are the most likely participants. Both these education groups (one or two years and full three years) have significantly higher probability of program participation than those who have a university/college degree and those who have primary school only. It is not surprising that those with university degrees are less likely as participants as those have better expected job prospects in the first place. Although it may be surprising that those who have the lowest level of education in this data are less likely participants than those who have secondary school education. Further, the influence of the work background on potential program selection is more as expected. Those with a work background as a manager/executive or academician are the least likely program participants. A work background of office and administration increases the probability of program selection in comparison to manager or academician at a 5% percent significance level. The other work backgrounds (sales, service or care; craft, construction, transportation, fishing and farming; cleaners, assistants etc.) also increases the probability for selection in comparison to manager or academician. Those coefficients are all significant at a 1 percent significance level. Lastly the instrumental variable, share of nonparticipants, is significant at a 1 percent level and has the expected sign as a higher share on nonparticipants in the municipality decreases the probability for program selection.

*Table 6. Estimated selection equation. Main analysis.*

| Variable | Coefficient | t-value |
|---|---|---|
| Age (reference: 52 – 60 years) | | |
| 42 – 51 years | 0.302** | 13.83 |
| 32 – 41 years | 0.459** | 20.34 |
| 25 – 31 years | 0.651** | 24.13 |
| 18 – 24 years | 0.685** | 20.79 |
| Male | 0.095** | 5.55 |
| Relationship status (reference: unmarried) | | |
| Married | -0.074** | -3.96 |
| Divorced | 0.026 | 1.18 |
| Objective (reference: obtain work) | | |
| Retain work | -0.801** | -42.85 |
| Increase participation | -0.418** | -11.73 |
| Foreign born | 0.078** | 3.42 |
| Municipality size (reference: small) | | |
| Medium sized municipality | -0.003 | -0.14 |
| Large sized municipality | -0.048* | -2.43 |
| Education (reference: university degree) | | |
| Primary school | 0.017 | 0.61 |
| 1 or 2 years of secondary school | 0.082** | 3.70 |
| 3 years of secondary school | 0.064** | 2.98 |
| Work background (reference: manager/academician) | | |
| Office and administration | 0.059* | 2.15 |
| Sales, service or care | 0.122** | 4.36 |
| Craft/constr., transportation, farming or fishing | 0.185** | 6.30 |
| Cleaners, assistants etc. | 0.200** | 5.76 |
| Share of nonparticipants | -0.025** | -25.13 |
| Constant | 0.905** | 15.17 |

*Significant at a 5 percent level, **significant at a 1 percent level.

### 6.1.2 *Employment outcome*

Table 7 shows the results from the outcome equations. Coefficients that are significant and have a positive sign increases the job probability with the presence of the related variable. The two columns that are presented gives the results for the treated group and the untreated group, given the selection. Comparisons between the outcomes of the treated group and the untreated will give an indication of how the treatment work. However, the estimated treatment effects will be most important in that regards, but looking at these numbers will give an indication on where the results are going.

Age shows significance for both the program participants and the nonparticipants. According to the results all other age groups shows higher probabilities for regular work in both the treated and the untreated state in comparison to the oldest age group. Among the program participants, the age groups 25 – 31 and 32 – 41 shows the highest job probability in comparison to the oldest group. Among the nonparticipants, those same age groups and also the youngest group shows the highest job probability in comparison to the oldest group. The youngest age group, 18 – 24 shows higher probability than those aged 42 – 51 among nonparticipants, but that relation is reversed among participants. Something that is certain is that the oldest group, 52 – 60, have the worst job prospects in comparison to the other age groups in both states. All the age dummy variables is significant at a 1 percent level.

Among the program participants, gender does not significantly affect the job outcome. Among the nonparticipants however, males have a significantly lower job probability indicating that females benefits less from program participation than males. In other words, the results indicate that females does better without programs than males. This is also reflected in the employment rates which are 44% for nonparticipating females and 37% for nonparticipating males. Furthermore, relationship status shows similar signs in both states. The presence of married increases the job probability in both states in comparison to unmarried, while the relationship status of divorced does not show any significance in comparison to unmarried, in either state.

*Table 7. Estimated employment outcome equations. Main analysis.*

| | Program participants | | Nonparticipants | |
|---|---|---|---|---|
| Variable | Coefficient | t-value | Coefficient | t-value |
| Age (reference: 52 – 60 years) | | | | |
| 42 – 51 years | 0.621** | 16.32 | 0.577** | 16.13 |
| 32 – 41 years | 0.821** | 20.09 | 0.691** | 15.65 |
| 25 – 31 years | 0.768** | 15.75 | 0.645** | 10.79 |
| 18 – 24 years | 0.591** | 10.80 | 0.680** | 9.45 |
| Male | 0.035 | 1.53 | -0.099** | -3.75 |
| Relationship status (reference: unmarried) | | | | |
| Married | 0.201** | 7.88 | 0.115** | 4.21 |
| Divorced | 0.036 | 1.17 | 0.018 | 0.54 |
| Objective (reference: obtain work) | | | | |
| Retain work | -0.001 | -0.01 | 0.404** | 8.63 |
| Increase participation | -0.255** | -4.30 | -0.098 | -1.70 |
| Foreign born | -0.178** | -5.87 | -0.132** | -3.77 |
| Municipality size (reference: small) | | | | |
| Medium sized municipality | 0.064** | 2.60 | -0.002 | -0.06 |
| Large sized municipality | -0.019 | -0.74 | -0.091** | -3.18 |
| Education (reference: university degree) | | | | |
| Elementary school | -0.497** | -12.87 | -0.452** | -10.95 |
| 1 or 2 years of secondary school | -0.262** | -8.64 | -0.306** | -9.37 |
| 3 years of secondary school | -0.124** | -4.25 | -0.137** | -4.40 |
| Work background (reference: manager/academician) | | | | |
| Office and administration | -0.123** | -3.07 | -0.157** | -4.28 |
| Sales, service or care | -0.128** | -3.20 | -0.260** | -6.61 |
| Craft/constr. etc. | 0.060 | 1.44 | -0.168** | -3.92 |
| Cleaners, assistants etc. | -0.211** | -4.33 | -0.377** | -7.10 |
| Constant | -0.879** | -8.92 | -0.549** | -7.49 |
| Log-likelihood | -40 729.26 | | | |

*Significant at a 5 percent level, **significant at a 1 percent level.

The health/motivation proxy, objective at start WAA, produces quite different results for the two different states. Among the program participants, objective increase participation (the poorest health status) stands out as it significantly decreases the job probability in comparison to the other two (objective obtain work is reference while retain work is insignificant). For the nonparticipants on the other hand, objective retain work (those who have jobs) now increases the job probability in comparison to the other groups, which is not surprising. Here objective increase participation is only significant at a 10 percent level rather than at a 1 percent level, but is still negative. This indicates that those who have the best proxy (retain work) benefits less from programs than those with the second best (obtain work). Further, those who are foreign born have significantly lower job probabilities in both states (the treated and the non-treated).

In comparison to a small municipality, a medium sized municipality increases the job probability in the treated state but not in the treated state. A large municipality has the reverse effect, it increases the job probability in the non-treated state but does not affect the outcome in the treated state, in comparison to a small municipality. The level of education has the expected signs, meaning that those with highest level of education have the best job prospects regardless of program participation and that it is decreasing with lower levels of education. In comparison to a university/college degree, the education level does not show much difference in affecting the job outcome over the two states. All the education level dummy variables are statistically significant at a 1 percent level.

When evaluating the coefficients related to work background, most coefficients shows significance and the same signs over the two states. It is as expected that in comparison to manager or academician the other work backgrounds decreases the job probability. The “least skilled” work background (cleaners, assistants etc.) decreases the job probability the most in comparison to manager or academician in both states. The work background that stands out from the other results are craft/construction etc. that actually shows an effect on the employment outcome in the treated state not significantly different from manager or academician, but in the untreated state it significantly decreases the employment outcome. This may be an indication that those with this work background benefits more from program participation than those with the other work backgrounds.

### 6.2 Program analyses

Since it is interesting to evaluate the different programs separately, this section will provide the results from the sub analyses of the different programs. The way the analyses is done is that the treatment group now only consists of those program participants on the program of interest. This is done for the programs training, work practice and assistance. The control group will remain the same, so the same switching probit estimation will be used, but now the treatment group are exclusively those who participated on the program of interest while the others are left out. Three different estimations are then ran, one for each program of interest. The average length of the different programs are 16.5 months (training), 12.5 months (work practice) and 13.5 months (assistance).

This may not be an ideal form of analysis. The preferred method might be to run a four-state model with multiple treatments. This would evaluate the four states; non-treatment, training, work practice and assistance together, where the last three are separate treatments. Although the method that will be used should give some insight to the selection process to the different programs and provide interpretable treatment effects. It is still the case that comparable individuals are compared on the observables and the unobservables when interpreting the treatment effects, even if these models are less fitted than that of the main analysis. Table 8 provides the estimated coefficients of the selection equations for the three programs of evaluation. The treatment groups for the different analyses consists of 6215 (training), 5045 (work practice) and 2808 (assistance) individuals.

The results from the selection equations from the three different models are provided in table 8. It shows that the estimated coefficients have similar signs as those in the main analysis. First, it is clear that age has similar effects on selection as was the results in the main analysis. Compared to the oldest age group, probability for program participation is always increasing with lower age groups (with one exception for training). Further, the selection to training perhaps shows the age factor more clearly than the others as there are larges differences in the probability of selection between age groups. This is not really unexpected as training first and foremost consists of ordinary education and it makes more sense to apply this to younger people than to older people. Gender, relationship status, objective and foreign born shows more or less the same results over the different treatments as before (except that divorced people are more likely and foreign born people are less likely to be elected for training).

For the training program a medium sized municipality increases the probability for selection while in the work practice program a small sized municipality increases the probability for selection. Further, those with the lowest level of education are the least likely to participate on training programs and assistance programs in comparison to other education levels. For work practice however, those with primary school and those with one or two years of secondary school are more likely participants than the others. These results indicates that work practice is more used towards people with only primary school education than the other two programs.

*Table 8. Estimated selection equations. Program analyses.*

| | Training (20 800 obs) | | Work practice (19 630 obs) | | Assistance (17 393 obs) | |
|---|---|---|---|---|---|---|
| Variable | Coefficient | t-value | Coefficient | t-value | Coefficient | t-value |
| Age (reference: 52 – 60 years) | | | | | | |
| 42 – 51 years | 0.624** | 18.33 | 0.239** | 7.54 | 0.283** | 7.73 |
| 32 – 41 years | 0.897** | 26.23 | 0.418** | 12.85 | 0.343** | 9.02 |
| 25 – 31 years | 1.175** | 30.27 | 0.601** | 15.75 | 0.484** | 10.82 |
| 18 – 24 years | 0.990** | 20.76 | 0.862** | 19.65 | 0.570** | 10.66 |
| Male | 0.051* | 2.13 | 0.063** | 2.59 | 0.124** | 4.41 |
| Relationship status (reference: unmarried) | | | | | | |
| Married | 0.038 | 1.49 | -0.155** | -5.77 | -0.186** | -5.93 |
| Divorced | 0.090** | 2.90 | -0.038 | -1.18 | -0.042 | -1.16 |
| Objective (reference: obtain work) | | | | | | |
| Retain work | -0.978** | -34.62 | -0.707** | -25.73 | -0.701** | -21.54 |
| Increase participation | -0.549** | -10.62 | -0.251** | -5.29 | -0.428** | -6.97 |
| Foreign born | 0.026 | 0.85 | 0.112** | 3.47 | 0.105** | 2.83 |
| Municipality size (reference: small) | | | | | | |
| Medium sized municipality | 0.082** | 3.13 | -0.145** | -5.59 | 0.049 | 1.53 |
| Large sized municipality | 0.038 | 1.40 | -0.250** | -9.09 | 0.000 | 0.01 |
| Education (reference: university/college degree) | | | | | | |
| Primary school | -0.244** | -6.14 | 0.143** | 3.68 | -0.126** | -2.74 |
| 1 or 2 years of secondary school | -0.024 | -0.81 | 0.179** | 5.56 | 0.002 | 0.05 |
| 3 years of secondary school | 0.049 | 1.70 | 0.061 | 1.93 | -0.036 | -1.03 |
| Work background (reference: manager/academician) | | | | | | |
| Office and administration | 0.017 | 0.44 | 0.058 | 1.39 | 0.087* | 1.96 |
| Sales, service or care | 0.124** | 3.26 | 0.176** | 4.24 | 0.082 | 1.78 |
| Craft/constr., transportation, farming or fishing | 0.271** | 6.76 | 0.205** | 4.74 | 0.097* | 2.01 |
| Cleaners, assistants etc. | 0.059 | 1.21 | 0.297** | 5.99 | 0.189** | 3.34 |
| Share of nonparticipants | -0.022** | -15.86 | -0.022** | -16.16 | -0.023** | -14.06 |
| Constant | -0.247** | -2.95 | 0.062 | 0.75 | -0.123 | -1.25 |

With regards to work background the results are similar as in the main analysis. For training however, cleaners, assistants etc. is now not more likely program participants than the reference, manager or academician. Indicating that those with the less skilled work background, cleaners, assistants etc. are not prioritized for training programs. The other thing that differs from the main analysis is that for assistance programs, those with work background from sales, service or care are not more likely participants than work background manager or academician.

The estimated employment equations for the different programs will not be reproduced here, but are found in the appendix. Now the various treatment effects will be discussed for both the main analysis and the specific program analyses.

## 6.3 Treatment effects

### 6.3.1 *Treatment effects from main analysis*

The following treatment effects are means. As mentioned the treatment effects are estimated for either the treated group (TT), the untreated group (TU) or the whole group (TE), but are not necessarily interesting or informative at an individual level.

From the main analysis the mean treatment effect on the treated, that is the estimated equation (5), are found to be:

$$TT(x) = \Pr(y_1 = 1|T = 1, X = x) - \Pr(y_0 = 1|T = 1, X = x) = 0.058$$

This means that, according to the model, the program participants do benefit from program participation if the goal is regular employment. The treatment effect on the treated are estimated to be 0.058 or 5.8 percentage points. This means that program participation increased the probability for regular employment by 5.8 percentage points for those who participated on programs. This is to say that, on average, the program participants had 5.8 percentage points higher probability for regular employment in comparison to if they had not participated. Differently put, the isolated effect from program participation on the outcome is estimated to be 5.8 percentage points, where the outcome is regular work.

Next is the treatment effect on the untreated (if they had been treated), that is the estimated equation (6), are found to be:

$$TU(x) = \Pr(y_1 = 1|T = 0, X = x) - \Pr(y_0 = 1|T = 0, X = x) = -0.137$$

This means that if the untreated had been treated, they would have been worse of in regards to regular employment. According to the model the nonparticipants would have had 13.7 percentage points lower probability of regular work had they participated in a program. This means that the nonparticipants have better job prospects without program participation than with.

The last parameter that are reasonably interpreted as a mean is the treatment effect. This is the estimated treatment effect for the whole group if everyone had participated on a program. The treatment effect are found from estimating equation (7):

$$TE(x) = \Pr(y_1 = 1 | \, X = x) - \Pr(y_0 = 1 | \, X = x) = -0.029$$

The mean treatment effect are found to be -0.029 or -2.9 percentage points. This means that for the whole group, both the treated and the untreated, the average effect of program participation is -2.9 percentage points in comparison to nonparticipation, on employment. This means that on average program participation decreases the probability for regular employment by 2.9 percentage points, if it had been applied to everyone.

If the goal is to improve the job prospects for those it is applied to, the results suggests that LMP programs are succeeding at its task. Those who participated on programs did benefit from it in terms of employment. Furthermore, these results seems rather encouraging for NAV as a social and economic caretaker. These results suggest that the labor market programs are directed towards people who gain from it such that the overall labor market effect is positive. Since the TT is positive this means that the participants benefits. On the other hand, the TU is negative which means that if the nonparticipants had been selected for programs, they would have been worse off in terms of employment prospects. By breaking the treatment effects into subgroups the reasons for the favorable selection can become clearer.

### 6.3.2 *Treatment effects by groups*

By using equation (8), the treatment effect on the treated and the treatment effect are considered for different subgroups. The results are reproduced in table 9. The left column provides the TT while the right column provides the TE. Emphasis will be on discussing the TT. The average treatment effect on the treated is 5.81 percentage points, as mentioned. The different groups can thus be compared with this and to each other to see the effect program participation has on different groups.

There are substantial differences between age groups in regards to the TT. Those who are between 25 and 41 years of age are the ones who benefits most while the two oldest groups benefit a bit less. What stands out is the negative TT on the age group 18 – 24. This means that the youngest groups are the one of the age groups that benefits the least from program participation. The reason for these results are unclear and are certainly not encouraging in regards to NAV's goals of prioritizing the young disabled workers. According to NAV though, such young aged individuals are rarely entitled to and granted WAA. Those very young people who do get granted WAA often have serious struggles, and may be far away from regular work. It could be that this group, which at first glance seemed more employable given that they are young, may not be very employable after all. Even though it still do not provide any explanation of why the programs have a negative effect.

Males benefit more from program participation than woman. The reason for this is not clear but it may be related to the high share of men with a work background from craft/construction etc. which will be shown to be the work background that benefits the most from program participation. With regards to relationships status, those who are married stands out with a higher TT. Further, there are substantial differences in treatment effects over the different objectives. Those who have the objective obtain work benefits most from program participation in this group. The group that had objective increase participation shows some positive treatment effect on the treated, while the group that had objective retain work have negative effect from program participation. The latter is not really surprising and is believed to be because this group already have jobs to return to, and it is thus expected that program participation will just lock in the participant and thus stall potential return to regular work. The sizes of the individual's municipality of residence and country of birth does not show much deviation as groups from the average treatment effect on the treated.

*Table 9. Treatment effects by groups. Main analysis.*

| | Treatment effect on the treated (TT) | | Treatment effect (TE) | |
|---|---|---|---|---|
| Variable | Estimate | Std. error | Estimate | Std. error |
| 52 – 60 years | 4.00 | 0.077 | -5.36 | 0.087 |
| 42 – 51 years | 5.66 | 0.084 | -4.68 | 0.093 |
| 32 – 41 years | 8.38 | 0.076 | -0.70 | 0.088 |
| 25 – 31 years | 7.39 | 0.077 | 0.96 | 0.092 |
| 18 – 24 years | -0.49 | 0.084 | -5.62 | 0.093 |
| Male | 8.59 | 0.054 | 1.11 | 0.056 |
| Female | 2.99 | 0.051 | -6.35 | 0.056 |
| Unmarried | 5.19 | 0.058 | -2.10 | 0.062 |
| Married | 7.67 | 0.077 | -3.16 | 0.081 |
| Divorced | 4.60 | 0.089 | -4.25 | 0.102 |
| Objective obtain work | 6.83 | 0.041 | 0.66 | 0.033 |
| Objective retain work | -1.60 | 0.110 | -14.79 | 0.062 |
| Objective increase participation | 3.57 | 0.184 | -4.56 | 0.121 |
| Foreign born | 5.32 | 0.104 | -2.14 | 0.114 |
| Small sized municipality | 4.79 | 0.085 | -4.01 | 0.087 |
| Medium sized municipality | 6.10 | 0.067 | -2.30 | 0.070 |
| Large sized municipality | 6.22 | 0.070 | -2.62 | 0.076 |
| *Education* | | | | |
| Primary school | 4.02 | 0.099 | -2.63 | 0.094 |
| 1 or 2 years of secondary school | 7.00 | 0.071 | -0.26 | 0.071 |
| 3 years of secondary school | 6.09 | 0.079 | -2.61 | 0.081 |
| University degree | 4.48 | 0.088 | -6.76 | 0.095 |
| *Work background* | | | | |
| Manager or academician | 2.71 | 0.119 | -8.68 | 0.134 |
| Office and administration | 2.84 | 0.081 | -7.32 | 0.091 |
| Sales, service or care | 4.43 | 0.064 | -3.40 | 0.065 |
| Craft/constr., transportation, farming | 10.36 | 0.069 | 2.91 | 0.072 |
| Cleaners, assistants etc. | 5.37 | 0.104 | -0.83 | 0.107 |
| Total | 5.81 | 0.042 | -2.85 | 0.045 |

Estimates are presented as percentage points and the presented standard errors are adjusted accordingly.

In this dataset education level is divided into four categories, and the results here indicates that it is those with the highest and those with the lowest level of education that benefits the least from program participation, although the differences between education levels are not that large. It may not be very surprising that those with the highest level of education here (university/college degree), benefits less than those with secondary school, as there is expected to be some lock in effect for people with better job prospects initially. It may be disappointing though that those with the lowest education level (primary school), are the ones who benefits the least from program participation when grouping into education levels. It is not clear why this is, employment rates are also similar (and relatively low) in both the treated and the untreated state. It may serve as a critique to the programs and administration that these group with relatively low job prospects does not improve this through program participation.

With regards to work background, manager or academician and office and administration shows the lowest gains from programs. This does not appear as very surprising as those are relatively skilled work backgrounds, and they are thus expected to have decent job prospects initially. Sales, service or care and cleaners, assistants etc. show similar treatment effects as the average treatment effect on the treated. The group that stands out with regards to work background and really also stands out in the whole sample are those with work background craft/construction, transportation, fishing or farming. This group shows a treatment effect on the treated of 10.36 percentage points. This means that this group are estimated to have over 10 percentage points higher job probability if they participate on programs in comparison to if they not participate. The reason for this high gain in this group is unclear, but one suggestion might be that those with this work background are more prone to physical injuries, both in terms of accidents but maybe mostly wear and tear. It may be that physical injuries are the reason for the reduction in working ability while the motivation and capabilities for regular work are still present. In this regard a program might serve as a retraining for a job that are not as demanding physically, and thus the program serves as a substantial help for the participant in getting readjusted for regular work.

### 6.3.3 *Treatment effects from program analyses*

The treatment effects from the different programs gave very different results. According to the model training shows by far the best effect as a treatment effect on the treated. The results from training were:

$$TT_{training}(x) = 0.151$$

$$TU_{training}(x) = -0.076$$

$$TE_{training}(x) = -0.008$$

Accordingly, the results from the model when having work practice as the only treatment were:

$$TT_{work\ practice}(x) = 0.009$$

$$TU_{work\ practice}(x) = -0.122$$

$$TE_{work\ practice}(x) = -0.088$$

Lastly, when running the model with assistance as the only treatment the results were:

$$TT_{assistance}(x) = 0.053$$

$$TU_{assistance}(x) = -0.031$$

$$TE_{assistance}(x) = -0.018$$

The treatment effect on the treated which are the most interesting parameters varies substantially over the different programs. Training shows a TT of 15.1 percentage points versus 5.3 percentage points for assistance. Worst of the estimated TT's is work practice with 0.9 percentage points. Training really stands out as superior to the other programs when applied to the "right" individuals. Work practice on the other hand gives little effect even if it seems to be directed towards individuals that are more treatable. The difference between the results advocates training as the best program when the goal is regular work. The weak results of work practice may be explained by the fact that in the dataset work practice contains both work practice in protected and ordinary environment. Those who are in the protected environment is expected to have weaker job prospects and the goal of regular work may not be as clear. Assistance shows results similar to that of the main analysis except that it does not show as negative effect on the untreated had it been applied to those.

All the three programs have a higher TT than TU indicating that the programs are directed against people who gain more relative to nonparticipants, such that the selection process is favorable from a labor market standpoint. Especially, training shows a large gap between TT and TU indicating that training is very effective as long as applied to the “right” people.

### 6.3.4 *Treatment effects from program analyses by groups*

In table 10 the treatment effects from the three program analyses are broken up into subgroups as was done in table 9 for the main analysis. The three columns in table 10 reproduces the treatment effects on the treated by subgroups from the three different models where training, work practice and assistance were specified as the single treatment.

Roughly speaking, the results from the different program analyses are similar, in terms of comparisons and relations within the models, to that from the main analysis. When considering the age groups, the youngest show the worst effect from all the programs. Both the programs training and work practice gives higher treatment effects with higher age, which contradicts that of the main analysis. The oldest age group has the highest treatment effect from both these programs, which may relate to that these groups have relatively low job prospects initially and those who participates on programs accomplishes a high gain from it. Males have higher treatment effect from all these three programs than females. Those with objective retain work have relatively low gain from training, while those with objective increase participation have relatively low gain from work practice. A possible interpretation of this is that objective retain work, individuals who are closer to employment, are locked into lengthy training programs. For those with objective increase participation, work practice may not have a good enough design to help and make those individuals more employable.

With regards to education level, results are similar as the main analysis in terms of comparisons within the model. What may be surprising is that work practice is better for those with a university/college degree than the others. For both training and assistance, those with secondary school level of education gain the most from program participation. Those with work background from craft/construction etc. still have the highest effect from all the programs. When considering assistance, the program works far worse for the work backgrounds manager/academician and sales/service or care, than for the other work backgrounds.

*Table 10. Treatment effects by groups. Program analyses.*

| | TT for training | | TT for work practice | | TT for assistance | |
|---|---|---|---|---|---|---|
| Variable | Estimate | Std. error | Estimate | Std. error | Estimate | Std. error |
| 52 – 60 years | 23.81 | 0.344 | 5.33 | 0.083 | 6.35 | 0.155 |
| 42 – 51 years | 18.53 | 0.186 | 2.11 | 0.089 | 7.61 | 0.163 |
| 32 – 41 years | 16.02 | 0.148 | 1.65 | 0.078 | 9.29 | 0.167 |
| 25 – 31 years | 13.29 | 0.155 | -0.66 | 0.077 | 1.37 | 0.178 |
| 18 – 24 years | 4.31 | 0.196 | -2.59 | 0.070 | -2.67 | 0.196 |
| Male | 18.99 | 0.131 | 2.41 | 0.062 | 7.47 | 0.145 |
| Female | 11.46 | 0.123 | -0.51 | 0.065 | 3.10 | 0.148 |
| Unmarried | 12.26 | 0.124 | 0.01 | 0.059 | 3.01 | 0.143 |
| Married | 20.45 | 0.167 | 3.18 | 0.095 | 8.24 | 0.180 |
| Divorced | 14.80 | 0.211 | 0.92 | 0.114 | 8.50 | 0.195 |
| Objective obtain work | 15.64 | 0.103 | 1.28 | 0.051 | 5.43 | 0.120 |
| Objective retain work | 9.05 | 0.401 | -0.58 | 0.183 | 4.11 | 0.336 |
| Objective increase participation | 14.22 | 0.660 | -1.83 | 0.211 | 6.81 | 0.532 |
| Foreign born | 12.21 | 0.259 | 1.21 | 0.126 | 7.51 | 0.291 |
| Small sized municipality | 15.40 | 0.225 | 0.70 | 0.093 | 2.57 | 0.221 |
| Medium sized municipality | 15.92 | 0.156 | 1.00 | 0.079 | 5.14 | 0.167 |
| Large sized municipality | 13.90 | 0.160 | 1.14 | 0.087 | 7.43 | 0.174 |
| *Education* | | | | | | |
| Primary school | 14.86 | 0.362 | 1.28 | 0.110 | 4.01 | 0.285 |
| 1 or 2 years of secondary school | 17.00 | 0.172 | 0.82 | 0.073 | 6.11 | 0.195 |
| 3 years of secondary school | 15.14 | 0.172 | 0.55 | 0.103 | 5.15 | 0.214 |
| University degree | 12.42 | 0.191 | 1.60 | 0.129 | 5.08 | 0.210 |
| *Work background* | | | | | | |
| Manager or academician | 12.40 | 0.308 | 0.50 | 0.213 | 1.88 | 0.328 |
| Office and administration | 11.34 | 0.200 | -0.27 | 0.127 | 6.81 | 0.207 |
| Sales, service or care | 12.61 | 0.151 | 0.06 | 0.075 | 2.89 | 0.191 |
| Craft/constr. etc. | 21.27 | 0.155 | 3.11 | 0.083 | 7.66 | 0.202 |
| Cleaners, assistants etc. | 14.11 | 0.298 | 0.28 | 0.103 | 6.97 | 0.302 |
| Total | 15.11 | 0.101 | 0.95 | 0.050 | 5.34 | 0.111 |

Estimates are presented as percentage points and the presented standard errors are adjusted accordingly.

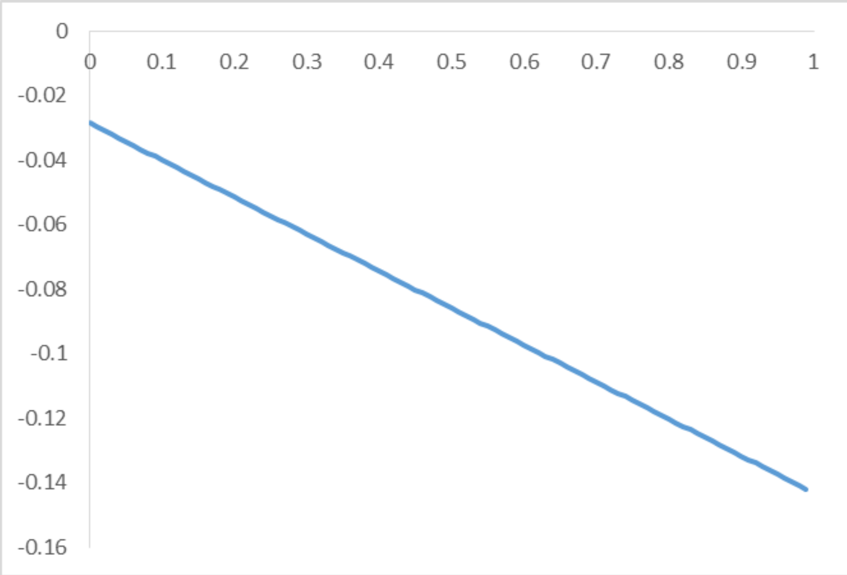
6.3.5 *Marginal treatment effect*

Different programs has been assessed, but now the results from the main analysis is again the focus. From the main analysis, the marginal treatment effect (MTE) will now be considered. How the selection process works have been discussed with respect to observable characteristics but not in terms of the unobservable. The advantage of the MTE is that it helps interpreting the selection process to programs in terms of unobserved characteristics. The MTE is the expected effect of treatment conditional on observed characteistics $X$ and conditional on $\mu$, the unobservables that affects potential program participation (Heckman & Vytlacil 2005).

Figure 1 plots the MTE against different normalized values of the unobserved characteristics $\mu$, given $x$. As was discussed, $\mu$ is the unobserved characteristics of the individuals that affects how likely they are to participate on programs. Here it ranges from 0.01 to 1 and the small values of the unobservables are related to those who are the most likely program participants, while the larger values are related to those who are the least likely participants. An interpretation of $\mu$ is that it might be something like personal motivation for participating on programs, which obviously is not observable from the data. The low values are consistent with higher probability for program participation, maybe because the individual are more motivated and gets himself/herself self-selected. High values of $\mu$, which means less likely program participation, may be because the individual are unmotivated for program participation. The fact that the curve is not flat confirms the presence of heterogeneity in the responses to program impacts. In other words, programs affects the individuals differently in terms of their employment outcome.

The marginal treatment effect are downward sloping, or decreasing in $\mu$, meaning that those who are the most likely program participants, based on the unobservables, have higher effects from treatment. The individuals whose unobservables make them the least likely program participants (large $\mu$) are those who are worst off from program participation. If one interprets $\mu$ as personal motivation, as before, those more motivated are more likely program participants and vice versa.

*Figure 1. Marginal treatment effect.*




# 7 Discussion

This chapter will provide a discussion of the results. Comparisons are made to other studies and the selection process to programs are discussed more thoroughly as it is highly relevant for administrators and policy makers.

## 7.1 Comparison to other results

This section will provide a short comparison of the results from this analysis compared to other comparable evaluations. Since this is a consideration of Norwegian LMPs and its effect on employment outcomes for people with reduced working ability, evaluations of similar programs in Norway are used to compare.

Aakvik (2001) and Aakvik et al. (2005) uses a matching method and a selection model respectively and both considers and evaluates programs independently of which program, and for groups of peoples with reduced working ability. In the first article the program effect is estimated to 6.3 percentage points (but considered sensitive to selection bias. In the second article, the program effect (a TT is used as in this thesis) is estimated to -11 percentage points. In both of the papers, cream skimming is found to be present as more employable individuals are more frequently selected to programs, even though they may gain less than those less employable. In this thesis the program effect (the TT) when considering all programs as one was found to be 5.8 percentage points which is roughly similar to that of Aakvik (2001). On the other hand there is little signs of cream skimming in these results, such that this contradicts the results of Aakvik. There is however differences in this dataset and those of Aakvik's. In particular, those papers are based on data from the late 80's and the early 90's. In this regard these results may be an indication of improvements in NAV's system of selecting program participants.

Dahl & Lorentzen (2005) found that training programs gave positive effects while employment programs did not, except for certain subgroups (they use earnings as the outcome). That result is in line with the program specific analyses in this thesis, where it was estimated that training is the best program with regards to employment outcome. Further, Røed & Raaum (2006) uses a timing-of-events approach and data on program participation in Norway from 1989 to 2002. They found a positive post-program effect but this is almost undone by the negative on-program effect (lock-in effect) although effects varies over different groups. They also found evidence that programs are best utilized when they are directed towards individuals with initially lower job prospects and thus counteract long term unemployment spells, while they lock in individuals with better initial job prospects. The results of program impact heterogeneity across groups are in line with the results in this thesis, but here there are not necessarily those with lower expected job prospects that are most treatable. Further, Røed & Raaum (2006) also finds that those with lower job prospects benefits the most. Which again is not as clear in this thesis' results, as for instance prime aged (32 – 41 years) individuals are the age group that gain most from program participation. However other results may be more in line Røed & Raaum, for instance can a work background as manager/academician and office/administration be seen as more skilled, and these groups also show less effect from program than the other work backgrounds.

Westlie (2008) also uses a timing of events approach to cope with selection problems and uses Norwegian data from 1994 to 2003 on individuals with reduced working ability. He considers multiple treatments so that five programs are evaluated. He finds an overall positive effect from program participation at 8.4 percentage points, which is not very different from 5.8 percentage points, as found in this thesis. When considering the different programs, Westlie finds that training gave the best results at 11.7 and 15.4 percentage points (he has data that allows for splitting up training into two subcategories). He also finds that work training gave the worst results of the programs, both results corresponding to what was found in this thesis. Westlie found the effect to be 5.8 percentage points for work practice in ordinary environment and no effect for work practice in protected environment. In this thesis the data does not allow for splitting up work training, as discussed, but it was found to be small at 0.9 percentage points (when running work practice as the only treatment).

### 7.2 Selection to programs

Based on the unobervables, the downward sloping MTE suggests a successful selection in terms of selecting the individuals that gains the most from program participation in terms of employment outcome. By looking back at the selection equation and then taking into account the results from the outcome equations and the TT on subgroups, it can be argued that the selection on the observed characteristics is favorable as well. Good selection in this context is selecting those who are more treatable, that is those who gain the most from treatment. Some of the characteristics that increased the probability for selection were: young age, male, objective obtain work, secondary school education and work backgrounds from craft/construction etc. and cleaner/assistants etc. and as shown; several of these groups are also advantageously to select for program participation in terms of labor market gains.

The three youngest age groups; 18 – 24 years, 25 – 31 years and 32 – 41 years are the most likely program participants in terms of age, and the probability for selection is increasing with lower age. The two age groups; 25 – 31 and 32 – 41 have particularly high treatment effects and are thus contributing to the good selection because they are also more likely as program participants. From a labor market perspective NAV is “right” in selecting these groups for treatment in comparison to other age groups. On the other hand, NAV is “wrong” in selecting the youngest age group (18 – 24) for program participation. This group shows negative employment effect from participating on programs, meaning that this group should not have been selected so frequently to programs. While keeping in mind here that this is discussed purely in a short term labor market perspective.

Males are more likely to participate on programs and also has a higher TT than woman, contributing to selecting the “right” individuals. There are differences in the selection probability and the treatment effects over the different objectives at start of WAA. Objective obtain work shows higher probability for selection and also higher treatment effect than the other objectives, such that this also contributes an advantageously selection in terms of labor market gains. Those who have secondary school level education are more likely to participate and also gains the most from programs, in comparison to the other levels of education. Those who had the lowest level of education were as discussed less likely to participate on programs but that group also shows the lowest gain from program participation indicating that this drives the results towards selecting the "right" individuals.

Lastly, the grouping the individuals in terms of work background, shows several interesting results. It was clear that work backgrounds from craft/construction etc., cleaners/assistants etc. and sales, service and care were more likely participants. These work backgrounds also showed higher treatment effects on the treated than the other two. This is particularly true for craft/construction that showed the second highest probability for selection among these groups and also a very high effect from program participation. The groups that have been mentioned mostly contributed to selecting the "right" individuals for program participation. The only exception were that the youngest age groups showed negative results from treatment, and should thus not be so frequently selected. One other result contributing to the selection of "wrong" individuals are that of relationship status. Those who are married showed lower probability for program participation in comparison to the other statuses, but showed higher gain from it.

### 7.2.1 *Cream skimming?*

The results from this evaluation does not have enough information to directly accept or reject cream skimming or bottom fishing (reverse cream skimming). A consideration of cream skimming is limited to the discretional interpretation of the results. There seems to be effects leaning both ways.

It was considered that the selection process to programs were efficient in terms of labor market gains. That is, the selection process seems to select people who benefits from programs, rather than those who do not benefit. If it was a fact that those who benefit the most from programs also were the ones who were farthest away from regular work, the results would have suggested bottom fishing. This is not necessarily the case, as the only thing that can be interpreted from the results are that people who gain more from participation than nonparticipation are more frequently selected. So, given the design and nature of the programs, individuals that are more treatable are more frequently participating on programs.

The results further indicate that some of the characteristics that are expected to make people less employable on average are related to a smaller effect from treatment. For instance, one might expect those with the lowest level of education (primary school) to have a larger effect from programs than those with higher levels of education. The results here though, indicate that when comparing the whole group in terms of their level of education, those with the lowest level are the ones who gain the least. One might argue the same way in regards to the oldest and the youngest age groups in comparison to those who are in between and in their prime age. The youngest and the oldest groups are the ones who benefits the least even though they may be less employable with regards to little work experience and old age respectively. Such that the consideration of the selection process of “right/wrong” individuals is not necessarily linked with cream skimming according to the model.

It is thus hard to say much about cream skimming in terms of the whole model. Although, some variables may be linked with lower expected job prospects and show increased probability for program selection. NAV has a youth guarantee, meaning that young people are prioritized for programs. This was clear from the main analysis as the youngest group had the highest probability for program participation and that the probability was decreasing with age. This means that the prime aged groups; 25 – 31 and 32 – 42 were more likely to participate than those at 52 – 60 years old. This may have an element of cream skimming as one would at least expect that unemployed aged 32 – 42 years have initially better job prospects than unemployed aged 52 – 60 years. When considering education levels, the highest and the lowest level of education shows the lowest probability of selection. The fact that the highest level (university/college degree) shows this is related to bottom fishing, but that the lowest level shows this, could be related to cream skimming. Especially, those with the lowest level of education have the lowest probability of participating on training programs (when looking at training as only treatment), and training programs main ingredient is education. In terms of work background the selection shows signs of bottom fishing rather than cream skimming as those with the more skilled professions are least likely to participate on programs.

# 8 Summary and conclusion

I have in this thesis evaluated the effect of labor market programs on unemployed with reduced working ability in Norway, with data delivered by NAV. The evaluation considers 32 924 individuals that started receiving work assessment allowance (WAA) from March 2010 until and including December 2011. WAA is a social benefit granted to individuals that have a proven reduced ability to work because of physical/mental/social health issues. A selection model was the method of choice to cope with the selection bias that arises from such observational studies. It was found that programs have a positive effect on those who participates and that the selection of program participants are successful in terms of short term labor market gains. The estimated treatment effect on the treated was 5.8 percentage points meaning that the estimated job probability for the program participants were 5.8 percentage points higher than it would have been had they not participated on a program. The treatment effect on the untreated and the treatment effect was found to be -13.7 and -2.9 percentage points respectively. This indicates that the programs would have decreased the job probabilities for the nonparticipants had they participated and also would have decreased the job probability on average had it been applied to everyone. A TT higher than TU and TE indicates that program participants are also those who, on average, gain more from it.

There was done sub analyses for the three most frequently used programs in the data which are training, work practice and assistance. It was found that training gave the best results with an estimated TT of 15.1 percentage points. Then followed assistance with 5.3 percentage points and worst was work practice with 0.9 percentage points. However, the data does not allow for splitting up work practice in ordinary and protected environment which would have been desirable as those measures are quite different in its purpose. Regardless, training performed best and this result is supported in other evaluation studies done in Norway (Dahl & Lorentzen 2005; Westlie 2008). Also it may be encouraging that training show good results even if the time span is short. One might expect training to have long term effects as well as short term, when considering that training first and foremost is some sort of education.

Prime age and particularly young age groups were more likely to participate on programs than older age groups. The three prime age groups (combined to 25 – 51 years) showed the highest effect from program participation while the oldest showed less and the youngest showed the worst and also negative. Males showed both higher probability for program participation and also a higher effect from it. In terms of education, those with secondary school education (1-3 years) and eventual technical school showed higher probability for being selected to programs than both those with the lowest level (primary school) and those with the highest level (university/college degree) of education. Those with secondary school education also showed the highest gain from programs. Further, when considering the sample in terms of work backgrounds the probabilities for program participation is more as expected. Manager or academician give the lowest probability for program participation and further it is increasing through office and administration; sales, service and care; craft, construction, transportation, fishing and farming and lastly with the highest probability; cleaners, assistants etc. Those with work background as manager or academician and office/administration gained the least from programs. A work background that really stood out was that of craft/construction etc. When considering those with this work background as a subgroup the TT was estimated to 10.4 percentage points which is more than any other of the “subgroups”. Combining subgroups for the creation of a “perfect program participant” in terms of how much higher the individual’s job probability is with program participation than without, would produce a married male in his 30’s with 1 or 2 years of secondary school and a work background from craft/construction. Such an individual is also very likely to be selected to program participation, as all those characteristics, except being married, increases the probability for selection.

A consideration of the marginal treatment effect (MTE) was provided. The distribution of the MTE over the unobserved characteristics showed that those who have unobserved characteristics that makes them more likely to participate on programs also gain more from it and vice versa. If the unobserved characteristics is interpreted as for instance personal motivation those who are more motivated increases their own probability of getting selected to programs and also has a better effect from the program on employment outcome. Further, the distribution of the MTE is not flat which confirms the presence of unobserved heterogeneity of program impacts on the individual’s employment outcome.

When considering potential cream skimming or bottom fishing the results are not conclusive. There are observable characteristics that works both ways. For instance, are prime age (25 – 41 years) related to a higher probability of program participation even though one might expect those to have decent initial job prospects. Also, those who have the lowest level of education are less likely to participate on programs in comparison to two of the other three education levels. These results relates to cream skimming. On the other hand, those who come from less skilled work backgrounds are more likely program participants, which relates to bottom fishing. However, the selection process is successful in selecting those who are more treatable and thus the results indicates that, based on the observables, those who are most treatable are not necessarily those with the poorest job prospects. So given the design and nature of the programs, the least employable may not necessarily gain the most from programs.

Further research at this topic will come and it should be that and previous research that lays the foundation for the design and the execution of labor market programs in Norway. This evaluation did not have access to health information which would have been preferred, and may be an improvement of future evaluations. This evaluation also had a short time span to work with after exit from WAA. This means that potential long term gains from program participation may be overlooked. Furthermore, this thesis and other economic evaluations often considers mainly the economic labor market outcome. Research on health and social effects should also be considered. NAV has been through reforms in the recent decade and many of the new implications and what follows from those should be of interest for researchers and policy makers as well as the public.

# References

Aakvik, A., 2001. Bounding a matching estimator: the case of a Norwegian training program. *Oxford Bulletin of Economics & Statistics*, 63(1), pp.115–143.

Aakvik, A., Heckman, J.J. & Vytlacil, E.J., 2005. Estimating treatment effects for discrete outcomes when responses to treatment vary: An application to Norwegian vocational rehabilitation programs. *Journal of Econometrics*, 125(1-2 SPEC. ISS.), pp.15–51.

Abbring, J.H. & van den Berg, G.J., 2003. The nonparametric identification of treatment. *Econometrica*, 71(5), pp.1491–1517.

Bassi, L.J., 1984. Estimating the effect of training programs with non-random selection. *The Review of Economis and Statistics*, 66(1), pp.36–43.

Björklund, A. & Moffitt, R., 1987. The estimation of wage gains and welfare gains in self-selection models. *The Review of Economics and Statistics*, 69(1), pp.42–49.

Blundell, R. & Costa Dias, M., 2000. Evaluation methods for non-experimental data. *Fiscal Studies*, 21(4), pp.427–468.

Card, D., Kluve, J. & Weber, A., 2010. Active labour market policy evaluations: a meta-analysis*. *The Economic Journal*, 120(548), pp.F452–F477.

Dahl, E. & Lorentzen, T., 2005. What works for whom? An analysis of active labour market programmes in Norway. *International Journal of Social Welfare*, 14(2), pp.86–98.

Dutoit, L., 2007. Heckman ' s selection model, endogenous and exogenous switching models, a survey. *SelectedWorks of Laure C Dutoit*, (May).

Frölich, M., Heshmati, A. & Lechner, M., 2004. A microeconometric evaluation of rehabilitation of long-term sickness in Sweden. *Journal of Applied Econometrics*, 19(3), pp.375–396.

Heckman, J.J., 1979. Sample selection bias as a specification error. *Econometrica*, 47(1), pp.153–161.

Heckman, J.J., Heinrich, C. & Smith, J., 2002. The performance of performance standards. *The Journal of Human Resources*, 17(4), pp.778–811.

Heckman, J.J., LaLonde, R.J. & Smith, J.A., 1999. The economics and econometrics of active labor market programs. In: Ashenfelter, O & Card, D. (eds.) Handbook of labor economics, Vol 3. Elsevier Masson SAS, pp. 1865-2097.

Heckman, J.J. & Vytlacil, E.J., 1999. Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences of the United States of America*, 96(8), pp.4730–4734.

Heckman, J.J. & Vytlacil, E.J., 2005. Structural equations, treatment effects, and econometric policy evaluation. *Econometrica*, 73(3), pp.669–738.

Kluve, J., 2006. The effectiveness of European active labor market policy. *RWI Discussion Papers*, (37).

Lokshin, M. & Glinskaya, E., 2009. The effect of male migration on employment patterns of women in Nepal. *World Bank Economic Review*, 23(3), pp.481–507.

Lokshin, M. & Sajaia, Z., 2011. Impact of interventions on discrete outcomes: maximum likelihood estimation of the binary choice models with binary endogenous regressors. *The Stata Journal*, 11(3), pp.368–385.

OECD, 2013. Mental health and work: Norway. *OECD Publishing.*

Van Ours, J.C., 2004. The locking-in effect of subsidized jobs. *Journal of Comparative Economics*, 32(1), pp.37–55.

Quandt, R.E., 1972. A new approach to estimating switching regressions. *Journal of the American Statistical Association*, 67(338), pp.306–310.

Røed, K. & Raaum, O., 2006. Do labour market programmes speed up the return to work? *Oxford Bulletin of Economics and Statistics*, 68(5), pp.541–568.

Skedinger, P. & Widerstedt, B., 2007. Cream skimming in employment programmes for the disabled? Evidence from Sweden. *International Journal of Manpower*, 28(8), pp.694–714.

Westlie, L., 2008. Norwegian vocational rehabilitation programs: Improving employability and preventing disability? *Memorandum // Department of Economics, University of Oslo*, (2008, 24).

# Appendix

## *A.1 Estimated employment outcome equations for training**

| | | | |
|---|---|---|---|
| Switching probit model | Number of obs | = | 20800 |
| | Wald chi2(20) | = | 3177.84 |
| Log likelihood = -24003.444 | Prob > chi2 | = | 0.0000 |

(selection equation output omitted)

| | Coef. | Std. Err. | z | P>\|z\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| jobb_1 | | | | | | |
| alder51_42 | .477712 | .0934534 | 5.11 | 0.000 | .2945467 | .6608772 |
| alder41_32 | .5894606 | .1112933 | 5.30 | 0.000 | .3713296 | .8075915 |
| alder31_25 | .5398421 | .135618 | 3.98 | 0.000 | .2740356 | .8056485 |
| alder24_18 | .3268169 | .1331518 | 2.45 | 0.014 | .0658442 | .5877895 |
| mann | .0499765 | .0384009 | 1.30 | 0.193 | -.025288 | .1252409 |
| gift | .2936998 | .0404909 | 7.25 | 0.000 | .2143391 | .3730604 |
| skilt | .0495681 | .0515635 | 0.96 | 0.336 | -.0514945 | .1506306 |
| beholde_arb | -.1257268 | .124224 | -1.01 | 0.311 | -.3692013 | .1177477 |
| oke_deltak | -.2808834 | .1128011 | -2.49 | 0.013 | -.5019694 | -.0597973 |
| utenland | -.2517158 | .0497338 | -5.06 | 0.000 | -.3491922 | -.1542393 |
| medium_kom | .0826597 | .0437218 | 1.89 | 0.059 | -.0030334 | .1683527 |
| stor_kom | -.0561568 | .0443499 | -1.27 | 0.205 | -.143081 | .0307674 |
| grunnskole | -.4468759 | .0697081 | -6.41 | 0.000 | -.5835012 | -.3102505 |
| vgs_gk_vk1 | -.1908496 | .0483167 | -3.95 | 0.000 | -.2855486 | -.0961506 |
| vgs | -.083221 | .0464063 | -1.79 | 0.073 | -.1741756 | .0077336 |
| kontor | -.1210512 | .0664618 | -1.82 | 0.069 | -.2513141 | .0092116 |
| salg_serv | -.1043708 | .0662689 | -1.57 | 0.115 | -.2342555 | .0255139 |
| handv_anlegg_jord | .1171024 | .0722638 | 1.62 | 0.105 | -.0245321 | .258737 |
| hjelpe_renh | -.176029 | .0829455 | -2.12 | 0.034 | -.3385992 | -.0134587 |
| _cons | -.5991438 | .2578846 | -2.32 | 0.020 | -1.104588 | -.0936993 |
| jobb_0 | | | | | | |
| alder51_42 | .5874661 | .0395835 | 14.84 | 0.000 | .5098838 | .6650484 |
| alder41_32 | .7043597 | .0535683 | 13.15 | 0.000 | .5993677 | .8093517 |
| alder31_25 | .6636441 | .0761761 | 8.71 | 0.000 | .5143417 | .8129464 |
| alder24_18 | .7164423 | .0756225 | 9.47 | 0.000 | .5682249 | .8646596 |
| mann | -.0912093 | .0261006 | -3.49 | 0.000 | -.1423656 | -.0400531 |
| gift | .1064908 | .0273986 | 3.89 | 0.000 | .0527906 | .1601911 |
| skilt | .0165387 | .0333908 | 0.50 | 0.620 | -.0489061 | .0819834 |
| beholde_arb | .3678436 | .0510047 | 7.21 | 0.000 | .2678762 | .4678111 |
| oke_deltak | -.1171815 | .0606455 | -1.93 | 0.053 | -.2360445 | .0016815 |
| utenland | -.1264073 | .0349624 | -3.62 | 0.000 | -.1949323 | -.0578824 |
| medium_kom | .0031935 | .0288861 | 0.11 | 0.912 | -.0534222 | .0598092 |
| stor_kom | -.0926469 | .0290343 | -3.19 | 0.001 | -.149553 | -.0357408 |
| grunnskole | -.443237 | .0438004 | -10.12 | 0.000 | -.5290842 | -.3573898 |
| vgs_gk_vk1 | -.2981085 | .0328947 | -9.06 | 0.000 | -.362581 | -.233636 |
| vgs | -.1333815 | .0311318 | -4.28 | 0.000 | -.1943987 | -.0723643 |
| kontor | -.1531137 | .036895 | -4.15 | 0.000 | -.2254265 | -.0808009 |
| salg_serv | -.2548837 | .039607 | -6.44 | 0.000 | -.3325121 | -.1772553 |
| handv_anlegg_jord | -.1628277 | .0443892 | -3.67 | 0.000 | -.249829 | -.0758264 |
| hjelpe_renh | -.3597693 | .0526241 | -6.84 | 0.000 | -.4629105 | -.256628 |
| _cons | -.4710581 | .056913 | -8.28 | 0.000 | -.5826057 | -.3595106 |
| /athrho1 | .2358099 | .1502994 | | | -.0587716 | .5303914 |
| /athrho0 | -.1154451 | .1302615 | | | -.3707529 | .1398626 |
| rho1 | .2315341 | .1422422 | | | -.058704 | .4856802 |
| rho0 | -.114935 | .1285407 | | | -.3546501 | .1389577 |

LR test of indep. eqns. (rho1=rho0=0):chi2(2) = 3.27 Prob > chi2 = 0.1945

*see A.4 for an explanation of the variables

## *A.2 Estimated employment outcome equations for work practice*

```
Switching probit model                          Number of obs   =      19630
                                                Wald chi2(20)   =    2378.50
Log likelihood =  -21925.05                     Prob > chi2     =     0.0000
```

| | Coef. | Std. Err. | z | P>\|z\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|

(selection equation output omitted)

| | Coef. | Std. Err. | z | P>\|z\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| jobb_1 | | | | | | |
| alder51_42 | .461255 | .0769742 | 5.99 | 0.000 | .3103884 | .6121217 |
| alder41_32 | .5868971 | .0857199 | 6.85 | 0.000 | .4188892 | .7549051 |
| alder31_25 | .4917001 | .1024358 | 4.80 | 0.000 | .2909296 | .6924706 |
| alder24_18 | .478762 | .1239655 | 3.86 | 0.000 | .2357942 | .7217299 |
| mann | -.0126089 | .044361 | -0.28 | 0.776 | -.099555 | .0743371 |
| gift | .1402279 | .0542132 | 2.59 | 0.010 | .033972 | .2464837 |
| skilt | -.026079 | .0628768 | -0.41 | 0.678 | -.1493152 | .0971573 |
| beholde_arb | .1679107 | .1062366 | 1.58 | 0.114 | -.0403093 | .3761307 |
| oke_deltak | -.2812185 | .1040014 | -2.70 | 0.007 | -.4850574 | -.0773795 |
| utenland | -.1321443 | .0617042 | -2.14 | 0.032 | -.2530823 | -.0112063 |
| medium_kom | .0237855 | .0457868 | 0.52 | 0.603 | -.065955 | .113526 |
| stor_kom | -.0804063 | .0561234 | -1.43 | 0.152 | -.1904062 | .0295935 |
| grunnskole | -.4891363 | .0769096 | -6.36 | 0.000 | -.6398763 | -.3383963 |
| vgs_gk_vk1 | -.3148145 | .0639839 | -4.92 | 0.000 | -.4402206 | -.1894084 |
| vgs | -.1719218 | .0610337 | -2.82 | 0.005 | -.2915457 | -.0522979 |
| kontor | -.1401105 | .0850367 | -1.65 | 0.099 | -.3067794 | .0265584 |
| salg_serv | -.1537835 | .0851341 | -1.81 | 0.071 | -.3206432 | .0130761 |
| handv_anlegg_jord | -.01997 | .0886989 | -0.23 | 0.822 | -.1938166 | .1538766 |
| hjelpe_renh | -.2681468 | .1011632 | -2.65 | 0.008 | -.466423 | -.0698707 |
| _cons | -.6801191 | .2561341 | -2.66 | 0.008 | -1.182133 | -.1781056 |
| jobb_0 | | | | | | |
| alder51_42 | .5911424 | .033804 | 17.49 | 0.000 | .5248878 | .6573969 |
| alder41_32 | .7067556 | .0430826 | 16.40 | 0.000 | .6223152 | .791196 |
| alder31_25 | .665062 | .0600889 | 11.07 | 0.000 | .5472899 | .7828342 |
| alder24_18 | .672141 | .088549 | 7.59 | 0.000 | .4985882 | .8456937 |
| mann | -.093907 | .0260947 | -3.60 | 0.000 | -.1450516 | -.0427624 |
| gift | .1200872 | .0283616 | 4.23 | 0.000 | .0644995 | .1756749 |
| skilt | .0241174 | .0330698 | 0.73 | 0.466 | -.0406982 | .088933 |
| beholde_arb | .3761823 | .0420556 | 8.94 | 0.000 | .2937549 | .4586098 |
| oke_deltak | -.1218277 | .0541278 | -2.25 | 0.024 | -.2279162 | -.0157392 |
| utenland | -.1327125 | .035079 | -3.78 | 0.000 | -.201466 | -.063959 |
| medium_kom | .014466 | .0269567 | 0.54 | 0.592 | -.0383681 | .0673002 |
| stor_kom | -.0730772 | .0308662 | -2.37 | 0.018 | -.1335738 | -.0125806 |
| grunnskole | -.4621898 | .0412841 | -11.20 | 0.000 | -.5431051 | -.3812746 |
| vgs_gk_vk1 | -.3116451 | .0333228 | -9.35 | 0.000 | -.3769565 | -.2463336 |
| vgs | -.1341573 | .0310104 | -4.33 | 0.000 | -.1949365 | -.073378 |
| kontor | -.1554421 | .0368257 | -4.22 | 0.000 | -.2276192 | -.0832649 |
| salg_serv | -.2605963 | .0397432 | -6.56 | 0.000 | -.3384916 | -.182701 |
| handv_anlegg_jord | -.1644525 | .043169 | -3.81 | 0.000 | -.2490622 | -.0798429 |
| hjelpe_renh | -.3797926 | .0543071 | -6.99 | 0.000 | -.4862325 | -.2733526 |
| _cons | -.5048924 | .0647111 | -7.80 | 0.000 | -.6317238 | -.378061 |
| /athrho1 | .0221807 | .157406 | | | -.2863293 | .3306908 |
| /athrho0 | -.2039159 | .1558787 | | | -.5094325 | .1016007 |
| rho1 | .0221771 | .1573286 | | | -.2787528 | .3191413 |
| rho0 | -.2011357 | .1495725 | | | -.4695029 | .1012526 |

```
LR test of indep. eqns. (rho1=rho0=0):chi2(2) =     1.83  Prob > chi2 = 0.3997
```

## *A.3 Estimated employment outcome equations for assistance*

```
Switching probit model                              Number of obs   =      17393
                                                    Wald chi2(20)   =    1241.24
Log likelihood = -17840.141                         Prob > chi2     =     0.0000
```

| | Coef. | Std. Err. | z | P>\|z\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|

(selection equation output omitted)

| | Coef. | Std. Err. | z | P>\|z\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| jobb_1 | | | | | | |
| alder51_42 | .5692697 | .1034341 | 5.50 | 0.000 | .3665425 | .7719968 |
| alder41_32 | .7436615 | .1095321 | 6.79 | 0.000 | .5289826 | .9583405 |
| alder31_25 | .5103722 | .1275828 | 4.00 | 0.000 | .2603145 | .7604299 |
| alder24_18 | .4411816 | .1449686 | 3.04 | 0.002 | .1570484 | .7253148 |
| mann | .047448 | .0597248 | 0.79 | 0.427 | -.0696104 | .1645064 |
| gift | .2069631 | .0704062 | 2.94 | 0.003 | .0689694 | .3449567 |
| skilt | .1292479 | .0767221 | 1.68 | 0.092 | -.0211246 | .2796204 |
| beholde_arb | .2131473 | .1348634 | 1.58 | 0.114 | -.0511801 | .4774746 |
| oke_deltak | -.0838647 | .1633068 | -0.51 | 0.608 | -.4039402 | .2362109 |
| utenland | -.1262894 | .076947 | -1.64 | 0.101 | -.2771027 | .0245239 |
| medium_kom | .0985409 | .069672 | 1.41 | 0.157 | -.0380138 | .2350956 |
| stor_kom | .0712313 | .0699847 | 1.02 | 0.309 | -.0659362 | .2083987 |
| grunnskole | -.4463944 | .1019372 | -4.38 | 0.000 | -.6461877 | -.2466011 |
| vgs_gk_vk1 | -.1929469 | .073948 | -2.61 | 0.009 | -.3378822 | -.0480116 |
| vgs | -.0838742 | .0715806 | -1.17 | 0.241 | -.2241696 | .0564211 |
| kontor | .0250266 | .0975086 | 0.26 | 0.797 | -.1660868 | .2161399 |
| salg_serv | -.1204615 | .098808 | -1.22 | 0.223 | -.3141217 | .0731986 |
| handv_anlegg_jord | .0220476 | .102235 | 0.22 | 0.829 | -.1783294 | .2224246 |
| hjelpe_renh | -.1397662 | .1186862 | -1.18 | 0.239 | -.3723869 | .0928545 |
| _cons | -.8512716 | .3570252 | -2.38 | 0.017 | -1.551028 | -.151515 |
| jobb_0 | | | | | | |
| alder51_42 | .5907903 | .0356588 | 16.57 | 0.000 | .5209004 | .6606802 |
| alder41_32 | .7168585 | .0409662 | 17.50 | 0.000 | .6365663 | .7971508 |
| alder31_25 | .6815054 | .0559751 | 12.18 | 0.000 | .5717961 | .7912146 |
| alder24_18 | .7121779 | .0727487 | 9.79 | 0.000 | .569593 | .8547628 |
| mann | -.097715 | .0268486 | -3.64 | 0.000 | -.1503373 | -.0450928 |
| gift | .1208807 | .0291459 | 4.15 | 0.000 | .0637558 | .1780055 |
| skilt | .0243045 | .0331413 | 0.73 | 0.463 | -.0406514 | .0892603 |
| beholde_arb | .3694848 | .0432821 | 8.54 | 0.000 | .2846535 | .4543162 |
| oke_deltak | -.1127579 | .0592949 | -1.90 | 0.057 | -.2289738 | .003458 |
| utenland | -.1314573 | .0350574 | -3.75 | 0.000 | -.2001685 | -.062746 |
| medium_kom | .0003821 | .0289736 | 0.01 | 0.989 | -.056405 | .0571693 |
| stor_kom | -.0919372 | .0286971 | -3.20 | 0.001 | -.1481824 | -.035692 |
| grunnskole | -.4433134 | .0433462 | -10.23 | 0.000 | -.5282705 | -.3583564 |
| vgs_gk_vk1 | -.2979068 | .032908 | -9.05 | 0.000 | -.3624053 | -.2334083 |
| vgs | -.1277152 | .0312212 | -4.09 | 0.000 | -.1889076 | -.0665229 |
| kontor | -.1578227 | .0369713 | -4.27 | 0.000 | -.2302851 | -.0853603 |
| salg_serv | -.2540459 | .0392258 | -6.48 | 0.000 | -.3309271 | -.1771648 |
| handv_anlegg_jord | -.1559295 | .0422384 | -3.69 | 0.000 | -.2387152 | -.0731438 |
| hjelpe_renh | -.3699573 | .0531292 | -6.96 | 0.000 | -.4740886 | -.2658261 |
| _cons | -.491619 | .0631461 | -7.79 | 0.000 | -.615383 | -.367855 |
| /athrho1 | -.08808 | .1889908 | | | -.4584952 | .2823353 |
| /athrho0 | -.2238887 | .2096368 | | | -.6347693 | .1869918 |
| rho1 | -.0878529 | .1875322 | | | -.4288569 | .275065 |
| rho0 | -.2202214 | .1994699 | | | -.5613275 | .1848424 |

LR test of indep. eqns. (rho1=rho0=0):chi2(2) =     1.47  Prob > chi2 = 0.4805

### *A.4 Explanation of variables in A.1, A.2 and A.3*

| | |
|---|---|
| jobb_1 | Employment outcome conditional on treatment |
| jobb_0 | Employment outcome conditional on non-treatment |
| alder51_42 | Age 42 – 51 |
| alder41_32 | Age 32 – 41 |
| alder31_25 | Age 25 – 31 |
| alder24_18 | Age 18 – 24 |
| mann | Male |
| gift | Married |
| skilt | Divorced/separated |
| beholde_arb | Retain work |
| oke_deltak | Increase participation |
| utenland | Foreign born |
| medium_kom | Medium sized municipality |
| stor_kom | Large sized municipality |
| grunnskole | Primary school |
| vgs_gk_vk1 | 1 or 2 years of secondary school |
| vgs | 3 years of secondary school |
| kontor | Office and administration |
| salg_serv | Sales, service and care |
| handv_anlegg_jord | Craft, construction, transportation, fishing or farming |
| hjelpe_renh | Cleaners, assistants etc. |