Evaluating the Effect of a National Labour and Welfare Administration Reform (Nav-reform) on Employment, Social Insurance and Social Assistance

Arild Aakvik • Karin Monstad • Tor Helge Holmås

Evaluating the Effect of a National Labour and Welfare Administration Reform (NAV reform) on Employment, Social Insurance and Social Assistance

Arild Aakvik, Karin Monstad and Tor Helge Holm?s

Uni Rokkan Centre, Bergen, Norway

Abstract

A national labour and welfare administration reform (NAV reform) is evaluated on employment outcomes and use of social insurance and assistance schemes. The reform was implemented step-wise over time during 2006–2010 at different geographical locations. Thus, we are able to create treatment and control groups to facilitate causal interpretations of the empirical results. We use a data set on individuals where we match local implementation of the reform to individual data on social security, employment, and socioeconomic characteristics. The labour and welfare administration was until 2006 fragmented, involving three large public agencies (employment, social assistance, and social insurance services), with limited coordination and collaborations. The reform merged these institutions into a new organization called NAV, after the "one door for all" principle ("one-stop-centres"). The idea was to help users get faster back to work after a period of sickness or unemployment through coordinated help by generalist case worker competence. We find that this was not the case for the implementation year and the following years, and that the reform failed to reap the intended benefits of the merger process, at least in the short term. This could be due to the fact that most users still demand specialized case worker competence, inadequate planning of IT infrastructure, large training requirements of staff, increased number of users during the 2008 financial crisis, and the buildup of new specialized teams in 2008 that took competent workers away from the day-to-day operation of the new NAV organization. Large bureaucratic organizational reforms can take many years to implement after a turbulent implementation face, and can have short term costs for users. The turbulence of the reform was not fully anticipated even though 25 offices in different municipalities and city districts participated in a pilot project 6

months before the country-wide implementation process started.

1   Introduction

The social insurance system is under pressure in many countries mainly from an aging population and an increased number of people leaving the labour market. Old age pension is the largest social insurance scheme, but more than 10 % of the potential labour force[1] are now on disability pension in Norway. Absence through sickness represents the third major type of financial transfers from the social insurance system to individuals. There is also a worrying connection between long-term absence through sickness and progression to a disability pension where people permanently leave the labour force. On a given working day, around 6.0 % of the work-force (130,000 persons) receive sickness benefits based on a sickness certificate from a general practitioner (GP) (NAV Statistics 2012). Unemployment insurance and social assistance are also an important part of the welfare state.

Pressure to increase the efficiency of the governmental sector has led to several reforms. The pension reform implemented in January 2010 uses financial incentives to get people to stay longer in the labour force, and the reform has to a large degree archived that goal (Dahl and Galaasen (2013)). However, the Norwegian government has failed on several occasions to change the sickness insurance system by introducing financial incentives for employers and employees, as has happened in the Netherlands (Everhardt and Jong (2011)) and Sweden (Johansson and Palme (1996)).

The labour and welfare administration was until 2006 divided into three main public agencies whose coordination and collaboration was limited. Especially multi-service users or clients were struggling in a system where many case workers had specialized tasks. Around 15 % of all users needed to apply for help from several different offices; Fevang, Røed, Westlie, and Zhang (2004), Christensen and Lægreid (2010). In 2001, it was decided to reform the labour and welfare administration by merging three large public agencies: the Norwegian Public Employment Services (PES), the National Insurance Services (NIS), and the municipality based Social Welfare Services (SWS). The new entity is called NAV.

---

[1]Background: The population in Norway in the age group 18-66 was 3.2 million in 2012. The total population is 5 million. The number of employed persons was 2,8 million in 2012, of which 1,9 million are full time employed. There are 1.37 million employed males, of which 1,17 million are full time employed. For females, there are 1,23 million employed, of which 730,000 are full time employed.

The three main goals of the NAV reform are a) to get more people into work and reduce the number of persons on welfare schemes like sickness, social assistance and unemployment; b) to create a more efficient administrative apparatus, and c) to make the administration more service-oriented, (Askim, Christensen, Fimreite, and Lægreid (2010)). We will look closer into employment outcomes and the use of welfare schemes in our analysis for groups of the population that has a high probability of having contact with the welfare system. These groups include long-term unemployed, persons having health problems, and young school drop-outs.[2]

NAV was implemented step-wise over time at different geographical locations. The implementation period lasted from October 2006 to April 2011. Norway has 435 municipalities, and 19 counties including Oslo. The four large cities (Oslo, Bergen, Trondheim, and Stavanger) are in addition divided into urban administrative district with their own labour and welfare administration. 25 municipalities/districts implemented the reform in October 2006 as a pilot project, where one municipality was selected more or less randomly from each of the 19 counties. In addition, 6 urban administrative districts from the two largest cities (Oslo and Bergen) were also selected at the same time. The last municipality implemented the reform in April 2011, but most of the 457 new offices were established by the end of 2009, where the main roll out period was from 2007 to 2009. The randomisation of municipalities was not perfect in a statistical sense (no coin or other randomisation device was used), but it was important that the selection of municipalities should be geographical stratified and implemented progressively over time. This was also the case for most urban administrative district. For instance, Oslo has 15 urban districts, the first one implemented the reform in 2006 (Sagene/Nordre Aker) and the last one at the end of 2010 (Nordstrand). Thus, at each period from 2006 to 2010 we have individuals belonging to the diversified system of three organizations (comparison group) and persons belonging to the new merged office system (treatment group) which is now called NAV.

We use this step-wise implementation of the reform to estimate causal effects on employment and use of social insurance and benefits within the difference-in-difference framework (Heckman and Honore (1990), Lazear (2000), Bhuller, Havnes, Leuven, and Mogstad (2013)). We have population data from 2003 to 2011 with information on income, employment, use of different social insurance (unemployment benefits, disability benefits, sickness benefits, old age pension, rehabilitation benefits) and social assistance benefits. Thus we observe all people before, during and after the reform with

---

[2]Schreiner (2012), Schreiner and Markussen (2012) and Fevang, Markussen, and Røed (2013) have done similar studies focusing on duration of spells before, during and after the NAV reform. Their analysis has a different research design and study population.

a follow up period of at least one year. We sample different groups of individuals prior to the reform in 2007, and follow individuals over time in the two different systems until pratically all municipalities had implemented the reform in 2010. We use the panel data design since the fixed-effects model is robust compared to many other empirical models when it comes to selection issues into the "treatment" and "control" group. We evaluate the reform on employment[3] and the use of social insurance and social assistance after periods of sickness or unemployment prior to the start of the implementation process.

The idea of the NAV reform was to facilitate labour-market integration, e.g., help users get faster back to work after a period of sickness or unemployment through coordinated help by generalist case worker competence, based on an integrated service model that is speading in Europe, e.g. Minas (2014). We find that this not the case for the national implementation period from late 2006 to early 2010. We find that the NAV reform failed to reap the intended benefits of the merger process in terms of increased employment and reduced use of social insurance schemes. This could be due to the fact that most users still demand specialized case worker competence, inadequate planning of IT infrastructure, and large training requirements of staff. Large bureaucratic organizational reforms can take many years to implement after a turbulent implementation face, and can have short term costs for users. The turbulence of the reform was not fully anticipated even though 25 offices in different municipalities and city districts participated in a pilot project 6 months before the country-wide implementation process started.

The paper is organized as follows. In Section 2, institutional background on the reform. Section 3 presents the data used in the empirical section. Section 4 presents the empirical set-up. Section 5 presents the main estimation results from the model. Section 6 concludes the paper.

2 Background to the reform

The labour and welfare administration was until 2006 divided into three main public agencies whose coordination and collaboration was limited. Especially multi-service users or clients were supposedly struggling in a system where many case workers had specialized tasks. Around 15 % of all users needed to apply for help from several different offices; Fevang, Røed, Westlie, and Zhang (2004). In 2001, it was decided to reform the labour and welfare administration by merging three large public agencies:

---

[3]We base our measure of employment on actual work income. Spells of employment from register data are inacurate and does not include self-employed persons.

the Norwegian Public Employment Services (PES), the National Insurance Services (NIS), and the municipality based Social Welfare Services (SWS). The new entity is called NAV ("new employment and welfare administration"), involving 16,000 employees and 1/3 of the national government budget.

The NAV reform is a continuation of earlier reforms in the early 1990s focusing on faster return to work for persons on social insurance schemes, in particular for persons with long term health problems and disabilities. This was particular the case for persons on vocational rehabilitation (VR). Many persons could stay on for many years on VR, and in 1994 the responsibility for and the administration of VR clients related to the return to work was transferred from the National Insurance Services (NIS) to the Public Employment Services (PES), and a distinction between medical and vocational rehabilitation was made. The administrative reform in 1994 has never been evaluated, but Markussen and Røed (2014) estimate the impact of vocational rehabilitation on short- and long-term labour market outcomes for persons that came under the risk of being referred to VR from 1996 through 2005. They find that a strategy focusing on rapid placement in the regular labour market is superior to alternative strategies giving higher priority to vocational training or sheltered employment. Strategies prioritising subsidized regular education also tend to be relatively successful in terms of final outcomes, but at the cost of protracted periods of benefit dependency first. They suggest more use of subsidized employment in the regular labour market, and a more restricted use of training courses and referrals to sheltered firms.

Aakvik, Heckman, and Vytlacil (2005) find on average no positive employment effects of vocational rehabilitation for clients that ended their rehabilitation effort prior to the reform in 1994, but Aakvik (2001) find a positive effect for clients with a low estimated probability of employment. A comparison of studies before and after the reform in 1994 shows that more focus on active re-employment strategies by the employment service seems to have had a positive effect compared to passive receipt of social insurance without focus on re-employment as was the case for many clients prior to the reform in 1994.

The generous Nordic welfare states hinges on high employment rates, and a continuous concern about high disability and sickness rates has spurred other administrative changes such as the one in 2000 where the requirements for permanent disability benefits were made clearer and more demanding, see for instance Ekhaugen (2006). Other administrative reforms include the "Faster return to work" (FRW) scheme that focus on reducing waiting times for treatment, and thus the total length of sick leave. The main idea behind the FRW scheme is that long waiting times for hospital treatment lead to unnecessarily long periods of sick leave and postponed return to work after

4

illness or injury. The reform has been evaluated by Aakvik, Holmås, and Kjerstad (2012). They find that the average waiting period for treatment or consultation for FRW patients (treatment group) is 12–15 days shorter than for people on sick leave on the regular waiting list (comparison group). This reduction is only partially transformed into a reduction in the total length of sick leave. On average, the reduction is approximately eight days. There is a significant difference between surgical and non-surgical patients, where surgical patients benefit the most from the reform in terms of significantly faster return to work. There is find no effect of the reform for non-surgical patients.

The NAV reform is a continuation of the focus on high employment for all groups in society as a foundation for a generous welfare state that characterizes the scandinavian countries. The unemployment rate in Norway is very low, but the number of persons on disability pension, rehabilitation benefits and sickness benefits are high compared to other countries, OECD (2010).

3 Implemention of the NAV reform

The NAV reform was implemented step-wise over time at different geographical locations as shown in Figure 1. Norway has 435 municipalities, and 19 counties including

Figure 1: Implementation of the NAV reform at the municipality level by quarter-year (N=457).
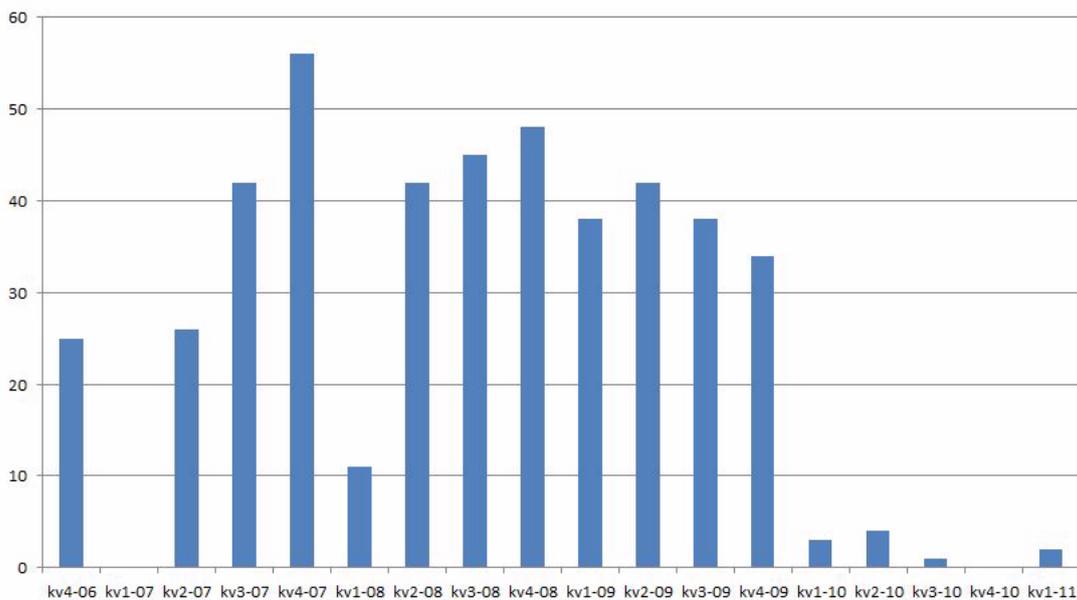
Oslo. The four large cities (Oslo, Bergen, Trondheim, and Stavanger) are in addition divided into urban administrative district with their own labour and welfare administration. 25 municipalities and city districts implemented the reform in October 2006 as a pilot project, where one municipality were selected more or less randomly from each of the 19 counties. In addition, 6 city administrative districts from the two largest cities (Oslo and Bergen) were also selected at the same time. The last municipality implemented the reform in April 2011 (Kristiansand), but most of the 457 new offices were established by the end of 2009, where the main roll out period was from 2007 to 2009, as can be seen from the figure. Municipalities finance social assistance, whereas social insurance is covered by the central government. In addition to NAV offices in each municipality/city distritc, NAV has also established county offices with particular tasks and administrative functions.

The randomisation of municipalities was not perfect in a statistical sense (no coin or other randomisation device was used), but it was important that the selection of municipalities should be geographical stratified and implemented progressively over time. This was also the case for most urban administrative district. For instance, Oslo has 15 city districts, the first one implemented the reform in 2006 (Sagene/Nordre Aker) and the last one at the end of 2010 (Nordstrand). Thus, at each period from 2006 to 2010 we have individuals belonging to the diversified system of three organizations (comparison group) and persons belonging to the new merged office system (treatment group) which is now called NAV.

3.1 The labour market

Labour market conditions have varied in the period from 2004 until today. The number of unemployed persons decreased steadily until the financial crises in 2008. The Norwegian stock exchange was at it lowest in November 2008 with an index value of 162. Three months earlier the stock exchange index was 375. Thus, there was a substantial drop in the value of companies listed at the stock exchange of almost 60 percent during the turmoil in fall 2008. The economy started to recover in late 2009. Two years after its lowest value, the stock exchange was back at its pre-crisis value in November 2010.

The number of unemployed persons varied a lot during our observation period. Figure 2 shows the number of unemployed male persons in the relevant period.[4] We can see that the number of registered unemployed persons doubled in 2008 from around 20000 to 40000 (male) individuals. The number of registered unemployed persons vary

---

[4] The figure for females looks very similar.

6

Figure 2: Average number of unemployed male persons, 2005–2013. Unadjusted monthly data, N=428 municipalities. Source: Statistics Norway.
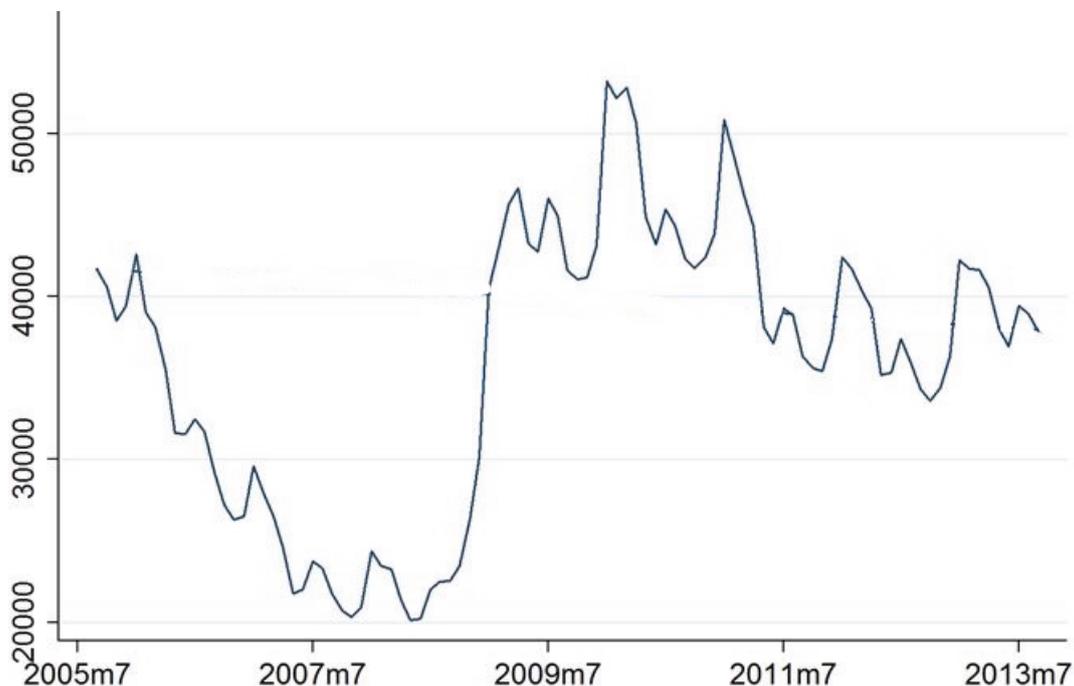

by season, but the general trend is decreasing from the autumn of 2009. We have not seasonal adjusted the data as is common in official statistics.

The NAV reform was implemented stage wise (see Figure 1). Figure 3 shows the implementation date and how this date varied a lot in terms of the general unemployment rate. Figure 3 shows the number of unemployed male persons in the relevant evaluation period and the timing of the implementation of the NAV reform at the municipality level. The early municipalities implemented the reform in a period with very low unemployment before the credit crises in the autumn of 2008 compared to the unemployment today. The unemployment rate increased during 2009 and reached its maximum early 2010. The credit crises falls just in the middle of the NAV implementation reform period from 2006 to 2010.

Figure 4 shows the unemployment rate for two groups of municipalities depending on the implementation period. The blue line in Figure 4 shows the unemployment rate for municipalities that implemented the reform in 2007 (second, third and fourth quarter). The red line shows the unemployment rate for municipalities that implemented the reform in 2009. The period up to and including the first quarter of 2007 is the pre-treatment period for both municipalities, and the period from and including 2010 is the post-treatment period for both groups. In 2008 we have both a treatment

Figure 3: Average number of unemployed male persons and the stage wise implementation of the NAV reform. Unadjusted monthly data, N=428 municipalities. Source: Statistics Norway (SSB), and Norwegian Social Science Data Services (NSD).
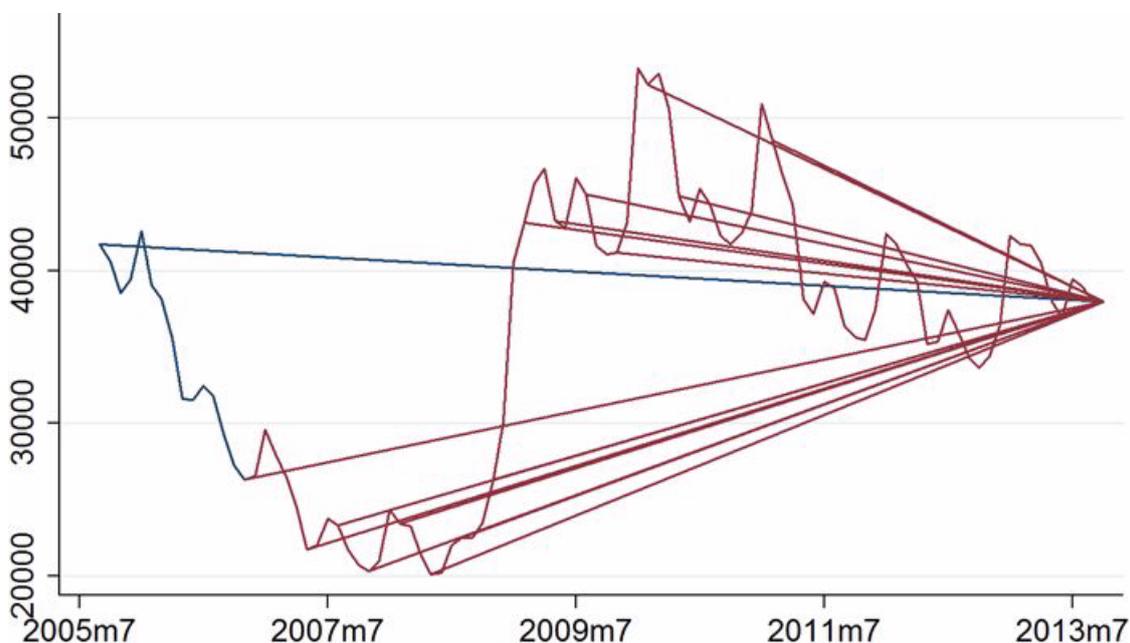
group (those that implemented the reform in 2007) and a comparison group (those that implemented the reform in 2009). The two horisontal lines are based on the average implementation period for the two groups of municipaliteis shown in the figure.

We can see from Figure 4 that the reform had a negative effect in early 2008, where the blue line (treatment group) is above the red line (comparison group). The negative effect is larger in the beginning of 2008 before the unemployment rate started to increase.[5] We can see from the figure that the new welfare system (NAV) did not tackle the financial crisis as did the comparison municipalities. From the figure we see that the blue line is above the red line in the spring of 2009. From then on both groups have implemented the reform (post-reform period) and the difference between the two groups are only minor in terms of unemployment rates when we compare municipalities that implemented the reform in 2007 to those that implemented the reform in 2009.

Figure 7 in the appendix shows the unemployment rate for municipalities that implemented the reform in the second quarter of 2007 (blue line) and those that implemented the reform in the second quarter of 2009 (red line). The early adopters

---

[5]The total effect is the cumulative difference between the two lines over the relevant evaluation period where we have both a treatment and comparison group.

Figure 4: Average unemployment rate for two groups of municipalities; those that implemented the reform in 2007 (blue line) and those that implemented the reform in 2009. Unadjusted monthly data. Source: Statistics Norway (SSB) and Norwegian Social Science Data Services (NSD).

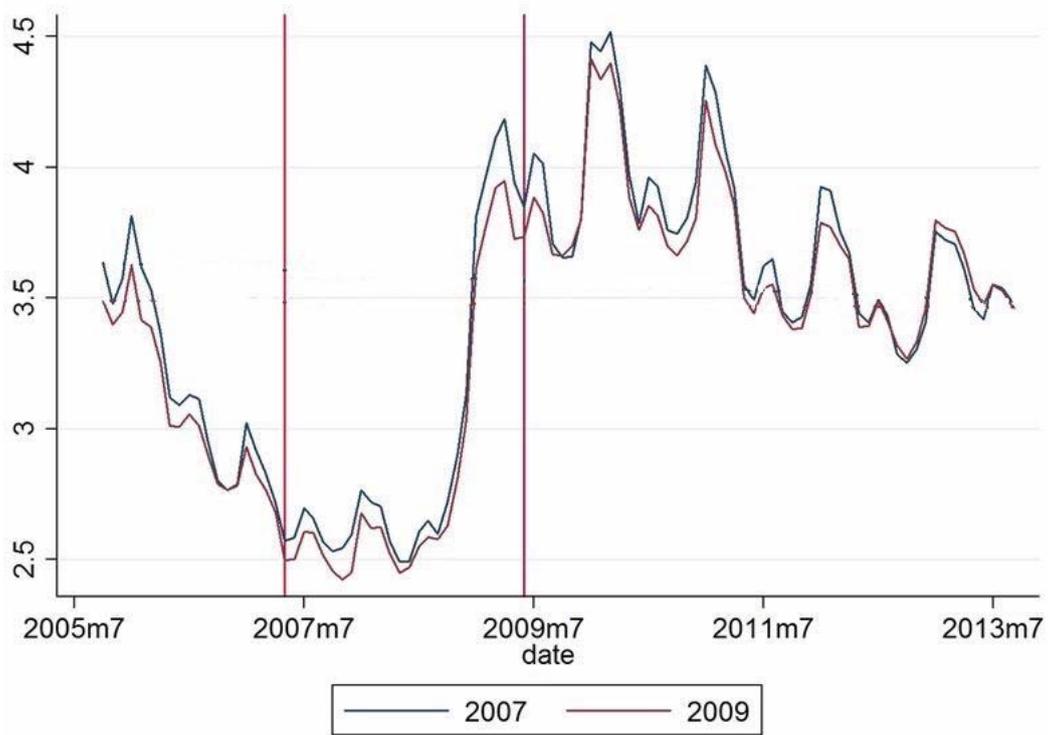have a slightly higher unemployment rate prior to the reform. In the beginning of 2009 the gap begins to widen, and the municipalities that implemented the reform in 2007 continues to have higher unemployment rates compared to those that implemented the reform in the second quarter of 2009. The figure indicates that the effect of the reform is negative for these municipalities.

Figure 8 in the appendix shows the unemployment rate for municipalities that implemented the reform in the fourth quarter of 2007 (blue line) and those that implemented the reform in the fourth quarter of 2009 (red line). Those municipalities that implemented the reform late (i.e. fourth quarter 2009) had higher unemployment rates compared to those that implemented the reform in the fourth quarter of 2007, but the changes over time is similar. The fixed-effect regression approach evaluates the reform in terms of changes over time (diff-in-diff-approach). An important assumption for the model is that the treatment and comparison groups experience the same kind of development in unemployment rates prior to the reform. Figure 8 indicates that this is actually the case. Thus, even if the level of unemployment is different for the two groups, we can still rely on the diff-in-diff-approach as a method of estimating the effect of the NAV reform.

Figure 4 indicates a small negative effect on unemployment rates of the refom, as do the other figures. An unconditional measure based on all municipalities indicate a negative effect of the NAV reform of around 0.05 percentage points on unemployment rates (for instance from 3.40 percent unemployment to 3.45 percent unemployment), which translates to around 1 percent more unemployed persons due to the NAV reform. When doing the same unconditional analysis[6] on all municipalities, we find similar effects. We also find that the smaller municipalities had a smaller negative effect compared to the larger municipalities on the unemployment rate at the municipality level, where the difference is around 0.02 percentage points.

When we include municipality and year fixed effects, as well as municipality-specific time trends, e.g. as in Wolfers (2006), we find that the effect drops to 0.03 percentage points increased unemployment. A more dynamic specification, inspired by equation (2) in Wolfers (2006), shows that the effect improves somewhat over time and becomes positive (i.e. reduces the unemployment rate) first after 2,5 years.

We have so far shown figures for unemployed persons only. The number of persons on long-term sickness, rehabilitation, and disability pension also varies over time and between municipalities. We can see that the number of long-term person on the sickness benefit scheme increased somewhat more in NAV municipalities compared

---

[6]We include dummy variables for time (i.e. fixed-effects for running month) and municipalities, but include no other control variables.

to the municipalities in the comparison groups, i.e. municipalities that have not yet implemented the reform. These numbers also shows a small negative effect, but the figures are much messier, and are not reported here.

Before we move on to a more rigorous analysis of different groups of users of social insurance and social assistant schemes in Norway, we sketch some of the reforms that have taken place in other European countries, tackling the fragmentation problem of social security systems.

Results from reforms in other countries

In the recent decade, several western European countries have reformed their welfare services to improve cooperation and coordination between different institutions. This reorientation has been characterized as "a paradigm shift in welfare state objectives from income protection to labour-market integration" (Van Berkel and Borghi (2008)). In an overview article, Champion and Bonoli (2011) categorize coordination initiatives along two dimensions: how many types of services and benefits the reform encompasses ("inclusiveness") and to what extent the governance structure is changed, with mergers as an extreme outcome ("intensity"). Champion and Bonoli (2011) characterize the NAV reform as "one of the most radical coordination initiatives in Europe". This is because of the scope of the reform (1/3 of budget, encompasses a very wide range of services and benefits, which implies that 50% of the population are in contact with NAV each year (Hansen, Lundberg, and Syltevik (2013), chapter 1) and because the reform involved a merger of two large state agencies followed by mandatory partnerships between central government and each municipality. Along the two dimensions mentioned above, the NAV reform resembles most closely the establishment of Jobcentre Plus in the UK and the SUWI reform in the Netherlands, both implemented in 2002. In France and Denmark, intensive but less encompassing reforms have been set into force (Champion and Bonoli (2011)).[7]

The UK Jobcentre Plus reform, like the NAV reform, was implemented gradually, as it took several years to roll out the reform to all districts of the country. Karagiannaki (2007) characterizes it a natural experiment, and study effects in the first 2 years of the implementation period. Data are at the district (as opposed to office) level, and the independent variable is therefore the percentage of offices within a district offering the integrated Jobcentre Plus services per quarter of year. Notably, there is no correction

---

[7]The UK has, alongside The Netherlands and countries such as Australia and the United States, been at the forefront of implementing "active" welfare state reform, whereas Germany, for example, is often viewed as a "laggard" of active welfare reform" (Christensen, Knuth, Lægreid, and Wiggan (2009).

for time trend in the analysis. The reform is estimated to have a positive effect on job entry outcomes (Karagiannaki (2007)).

The NAV reform is mainly a structural reform. The substance of the welfare services remained little changed in the implementation period, with two exceptions: "kvalifiseringsprogrammet" and "arbeidsavklaringspengar". These benefits were implemented in 2008 and May 2010, respectively. Notably, they were implemented on a national basis, i.e., more or less independent of each municipality's reform status at that time. In contrast, the Hartz reforms in Germany brought about an important change in the substance of the social security system, as unemployment and social assistance were integrated into one new benefit.

The Hartz reforms in Germany were implemented in the years 2003-2005. It is considered the most far-reaching reform endeavour in the German welfare state (Jacobi and Kluve (2006)). The broad aims of the reform were to improve employment services and policy measures, activate the unemployed, and deregulate parts of the labour market. An important cornerstone of the Hartz reforms was to improve the effectiveness and efficiency of the labour market services. This was done by modernising the organizational structure of the local employment services along the lines of New Public Management. The organizational change should facilitate better co-ordination of institutional arrangements, especially between active and passive policy measures. The Hartz reform converted the former employment offices into custumer-oriented one-stop centres where every job-seeker is assigned to a fixed caseworker, which is somewhat contrary to the generalist case worker competence strategy in the NAV reform.

The change of employment services into one-stop centres was evaluated by WZB and infas (2005) using a conditional difference-in-difference-analysis, see section 4. They used the fact that the new centres have been introduced at different points in time. Ten employment offices that had been modernized into one-stop-centres were matched to ten agencies that had not been changed at that time. The results were evaluated in terms of integration into regular employment showed no significant effects. However, the observation period was maximum nine months, and the short observation period could explain lack of significance according to Jacobi and Kluve (2006). WZB and infas (2005) finds that the effects are more positive in former East Germany, where labour markets are poorer, and that the change into one-stop-centres work better for men than women. Eichhorst and Kaiser (2006) discussed the Hartz and Agenda packages, and whether the reforms are overrated or just need some time to show the desired effects.

Minas (2014) synthesizes the work on one-stop shops as a strategy to combat fragmentation of social security systems in Western European countries. Unemployed

people loosely attached to the labour market, and people receiving social assistance benefits often have multiple problems simultaneously which implies contact with several different public agency. The article focuses on integrated service models that provide benefits or social assistance to unemployed persons. The analysis shows that integrated services might not fulfill the promises of more seamless services and that users might actually experience stricter work conditionality. This conclusion is based on governmental reports, academic research, and interviews with experts in the field.

4 Empirical model and identification of causal effects

For each person i we have two potential outcomes (Y_{0i}, Y_{1i}) corresponding to the potential outcomes in the untreated and treated states respectively. Let $D_i = 1$ denote the receipt of treatment and $D_i = 0$ denote non-receipt of treatment. In our case treatment indicates that a person lives in a municipality where the NAV reform has been implemented, and non-treatment means that the NAV reform has still not been implemented. All municipalities will eventually implement the reform. This happened in 2011. Between 2006 and 2011 we had a dual system where some municipalities had implemented the reform, indicated by $D_{it} = 1$, and some not, indicated by $D_{it} = 0$. Let Y_{it} be the measured outcome variable so that

$$Y_{it} = D_{it}Y_{1,it} + (1 - D_{it})Y_{0,it}.$$

This is the classical model of potential outcomes: see Neyman (1923), Fisher (1935), Roy (1951), Cox (1958), Quandt (1972), Rubin (1978), and Heckman and Honore (1990). The model can be used to estimate structural econometric models. The model has two potential outcome states of which only one is observed for each individual at a given point in time t. $Y_{1,it}$ is the outcome of an individual under the NAV reform in period t, and $Y_{0,it}$ is the outcome of an individual under the previous system in period t.

The comparison of expected outcome conditional on the NAV reform can be linked to the following effect measure:

$$
\begin{aligned}
E(Y_{it}|D_{it} = 1) - E(Y_{it}|D_{it} = 0) &= E(Y_{1,it}|D_{it} = 1) - E(Y_{0,it}|D_{it} = 1) \\
&+ E(Y_{0,it}|D_{it} = 1) - E(Y_{0,it}|D_{it} = 0),
\end{aligned}
$$

where $D_{it} = 0, 1$ indicates whether an individual resides in a municipality that has

implemented the NAV reform ($D_{it} = 1$) or not ($D_{it} = 0$) in period t. $E(Y_{1,it}|D_{it} = 0)$ is the expected outcome of participants had they not participated, $E(Y_{0,it}|D_{it} = 0)$ is the expected outcome of non-participants, $E(Y_{1,it}|D_{it} = 1)$ is the expected outcome of participants, and $E(Y_{0,it}|D_{it} = 1)$ is the expected outcome of non-participants had they participated.

The term $E(Y_{1,it}|D_{it} = 1) - E(Y_{0,it}|D_{1,it} = 1)$ show the outcome-effect solely linked to the NAV reform and is the average causal effect of NAV reform for those who participated. This is a conditional measure and we need identifying assumptions to be able to estimate the effect of the NAV reform, since we cannot observe an individual both in a reform municipality and pre-reform municipality at the same time.

The Difference-In-Difference model (DID)

The model that will be used to analyse our problem whether the NAV reform increases or decreases employment will be the differences-in-differences (DID) model. Known from the field of effect evaluation, the DID-estimator is taking advantage of a comparison group and a treatment group, and observing them over time before and after a policy change. Our data structure is illustrated in Figure 5 below. The DID model is

Figure 5: Implementation of the reform.
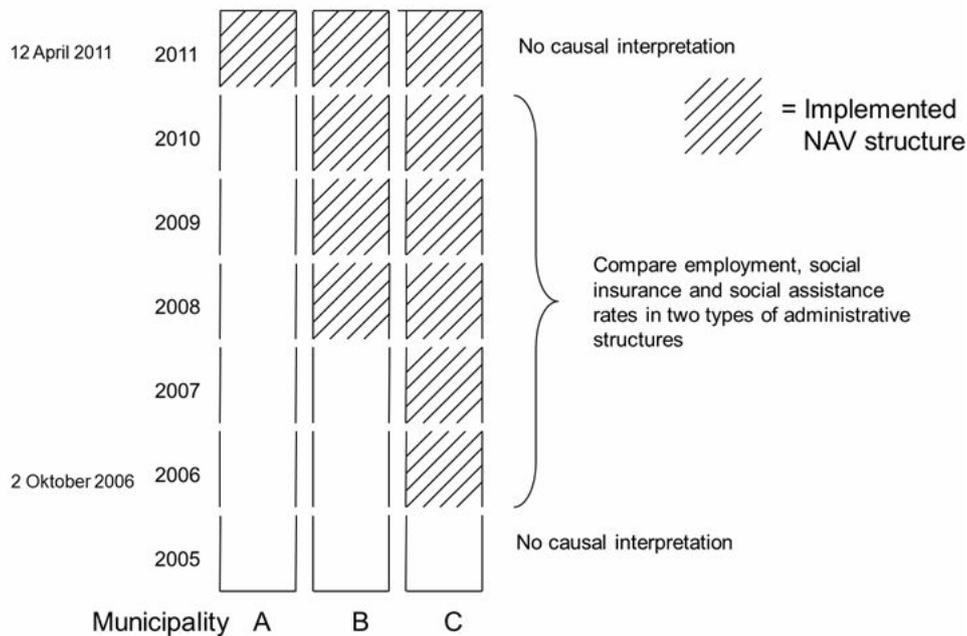
one of several methods for analysing and evaluating social programs. The evaluation

in itself seeks to estimate the difference in an outcome between the "treated" and "untreated" in the population, after a treatment has occurred. The DID-estimator hence makes use of either two distinct time-periods (before and after) or an average over the observations before and after the exogenous incident (Heckman, Lalonde, and Smith (1999), Verbeek (2008), Angrist and Pischke (2008)). We seek to estimate:

$$E(Y_{1,t+1}|X, D = 1) - E(Y_{0,t+1}|X, D = 1) \tag{1}$$

$E(Y_{1,t+1}|X, D = 1)$ is the expected outcome for the individuals or group that received the treatment if they in fact were given the treatment. $E(Y_{0,t+1}|X, D = 1)$ is the expected outcome for the individuals or group that did not receive a treatment if they in fact were given the treatment. How can we estimate the population average outcome for individuals who were not given treatment had they been given treatment ($D = 1$)? This counterfactual outcome is represented by $E(Y_{0,t+1}|X, D = 1)$.

We cannot observe a group or a person in two different states at the same time. What we seek to estimate is a measure of the expected outcome of those not treated had they in fact received the treatment. We cannot just compare the after-treatment outcomes of the two groups because they might have some other characteristics rendering them not comparable from the outset Heckman, Lalonde, and Smith (1999). Hence we would most likely find an estimation-bias in that the differences might be due to other factors than the treatment which we seek to evaluate, if we simply compare the two groups of individuals after the exogenous incident.

We can more easily estimate $E(Y_{1,t+1}|X, D = 1)$. This equation is the estimated outcome of the treated in a state of being treated. This can be estimated by using the sample average for this group after the treatment, either at a specific point in time, or to take the average over the time-periods in which we have data from after the treatment.

Implementation of the Difference-in-Differences Model

The DID-model assumes access to longitudinal data on both the treated and the comparison group. It also assumes that the estimated differences between individuals over time in the no-treatment outcome do not systematically change. What we mean by this is, if there were no treatment, we should not expect any systematic differences in the outcome-variables over time between the two groups. In fact in this case there would not even be two groups, because there has not been any exogenous incident dividing the individuals into groups. In the case of an exogenous incident, the groups may have different qualities from the outset rendering the outcomes different, but this

can be controlled for by using individual fixed-effects. Some individuals might for instance have a much higher initial level of outcome than the others, which would ultimately result in different levels of outcome also over time. The lack of systematic change in the no-treatment outcome does not mean that the individuals have to be homogenous from the beginning. In fact, we use these individual fixed-effects to control for heterogeneity amongst individuals.

We start off by observing a dependent variable Y_{it} for the entire sample and we propose that the exogenous incident, or the treatment, might have an impact on this outcome. We observe $i = 1, 2, \ldots, N$ individuals over $t = 1, 2, \ldots, T$ time-periods. We are assuming that this exogenous incident takes place sometime between period t and t+1. We introduce a fixed-effects model given by:

$$Y_{it} = \alpha_i + \gamma_t + \theta D_{it} + \epsilon_{it} \tag{2}$$

where α_i indicates the fixed-effects attributed to each individual in the sample, γ_t indicates the fixed-effects attributed to each time-period, and ϵ_{it} is an error-term which is identically and independently distributed and is assumed to take the expected value 0 and constant variance σ_e^2. We assume that the binary variable D takes the value $D_{i,t-1} = 0$ before a specific treatment for each individual while for some individuals it takes the value $D_{i,t+1} = 1$ after these specific individuals where given the treatment.

Estimating deviations from means is algebraically the same as treating the individual effects as parameters to be estimated (Angrist and Pischke (2008)). We may then eliminate these individual fixed-effects by first calculating the individual averages over time:

$$\bar{Y}_i = \alpha_i + \bar{\gamma} + \theta \bar{D}_i + \bar{\epsilon}_i, \tag{3}$$

and then subtracting this from equation (2):

$$Y_{it} - \bar{Y}_i = (\alpha_i - \alpha_i) + (\gamma_t - \bar{\gamma}) + \theta(D_{i,t} - \bar{D}_i) + (\epsilon_{it} - \bar{\epsilon}_i). \tag{4}$$

Here we have eliminated the individual fixed-effects which make us able to make the comparisons needed to evaluate the effect of the reform. This approach is often termed the within transformation. We may also eliminate the fixed effects by first-differencing the model (2). This essentially means that we subtract the time period from before the treatment from the period after:

$$Y_{i,t+1} - Y_{i,t-1} = (\alpha_i - \alpha_i) + (\gamma_{t+1} - \gamma_{t-1}) + \theta(D_{i,t+1} - D_{i,t-1}) + (\epsilon_{i,t+1} - \epsilon_{i,t-1}) \tag{5}$$

or:

$$\Delta Y_{i,t+1} = \Delta \gamma_{t+1} + \theta \Delta D_{i,t+1} + \Delta \epsilon_{i,t+1} \tag{6}$$

For consistency we must assume that $E(\Delta D_{i,t+1} \Delta \epsilon_{i,t+1}) = 0$. This condition do allow correlation between for instance $D_{i,t+1}$ and $\epsilon_{i,t-1}$ in levels, but we have to assume that they are uncorrelated in changes from one period to another. The assumption is therefore not as strong as the strict exogeneity condition: $E(D_{i,t} e_{i,s}) = 0$ for all s, t. We can estimate consistently the effect of θ with OLS from a regression of $\Delta Y_{i,t+1}$ on $\Delta D_{i,t+1}$.

The estimate we get for θ is the sample average of $Y_{i,t+1} - Y_{i,t-1}$ for the treated minus the sample average of $Y_{i,t+1} - Y_{i,t-1}$ for the untreated. Define $\Delta \bar{Y}_{(t+1)}^{treated}$ as the average for the treated where $D_{i,t+1} = 1$ and $\Delta \bar{Y}_{(t+1)}^{untreated}$ as the average for the untreated where $D_{i,t+1} = 0$. Then by performing an OLS regression we get:

$$
\begin{aligned}
\hat{\theta} &= \Delta \bar{Y}_{(t+1)}^{treated} - \Delta \bar{Y}_{(t+1)}^{untreated} \\
&= (\bar{Y}_{(t+1)}^{treated} - \bar{Y}_{(t+1)}^{untreated}) - (\bar{Y}_{(t-1)}^{treated} - \bar{Y}_{(t-1)}^{untreated}).
\end{aligned}
$$

The difference-in-difference estimator in given by $\hat{\theta}$. This is then first the difference between the treated and untreated averages and secondly the difference over the two time-periods. Figure 2: Illustration of research design and notation. The figure can

Figure 6: Research design and notation.
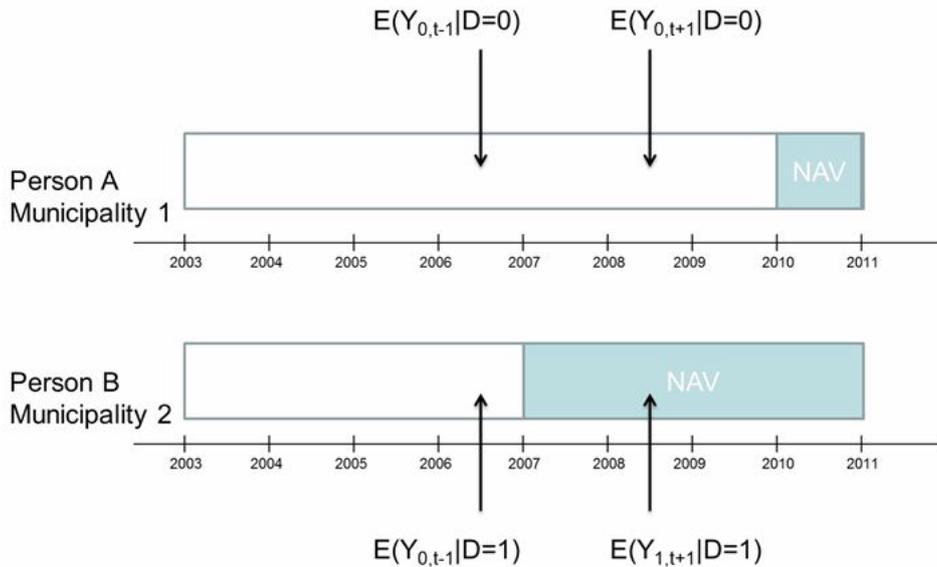
illustrate the DID-estimatore. Using first-differencing for illustration, we have that

$\bar{Y}_{(t+1)}^{treated}$ will be an estimate on $E(Y_{1,t+1}|D = 1)$, and $\bar{Y}_{(t+1)}^{untreated}$ will be an estimate for $E(Y_{0,t+1}|D = 1)$. $\bar{Y}_{(t-1)}^{treated}$ is an estimate for $E(Y_{1,t-1}|D = 1)$ and $\bar{Y}_{(t-1)}^{untreated}$ is an estimate for $E(Y_{0,t-1}|D = 0)$. We then see that:

$$\hat{\theta} = (\bar{Y}_{(t+1)}^{treated} - \bar{Y}_{(t+1)}^{untreated}) - (\bar{Y}_{(t-1)}^{treated} - \bar{Y}_{(t-1)}^{untreated})$$

is an estimate of

$$\hat{\theta} = \left[E(Y_{1,t+1}|D = 1) - E(Y_{0,t+1}|D = 0) - E(Y_{1,t-1}|D = 1) - E(Y_{0,t-1}|D = 0) \right].$$

5   Econometric specification

We are interested in testing whether the NAV reform has an effect on employment outcomes and the use of social insurance or social assistance. Define NAV as a dummy variable equal to one if the person lives in a municipality that has implemented the NAV reform and equal to zero if the person lives in a municipality where the old system is still in place. We start by using a linear probability model for panel data with the following specification

$$y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \delta_1 \, \mathrm{NAV}_{it} + \boldsymbol{\gamma}_I + \boldsymbol{\gamma}_T + \boldsymbol{\gamma}_M + \epsilon_{it}, \tag{7}$$

where $\mathbf{x}_{it}$ includes background variables such as educational level (eight categories), marital status (three categories: unmarried/married/divorced), and number of children (under the age of 11/18). $\boldsymbol{\gamma}_I$ is a set of dummy variables for individual (fixed-effects), $\boldsymbol{\gamma}_T$ is a set of year dummies (2003-2010), and $\boldsymbol{\gamma}_M$ is a set of dummy variables for municipalities of which there are 435. The outcome variable y_{it} consists of three alternative variables of which all are dummy variables: Empl $\in \{0, 1\}$ indicates whether a person is employed at a given time, SI $\in \{0, 1\}$ indicate whether a person receives some kind of social insurance (unemployment benefit, sickness benefit, etc) at a given time, and SA $\in \{0, 1\}$ indicates whether a person receives social assistance. In the case of a dichotomous outcome variable we have $\mathrm{E}(y_{it}) = \mathrm{Pr}(y_{it})$ and NAV_{it} is a dummy variable in the regressions. δ_1 in equation (7) measures the effect of the NAV reform in percentage points.

Step-wise implementation

We also allow for the fact that the effect of the NAV reform can change over time by including interaction terms between the NAV dummy and time dummies after the

implementation year:

$$
\begin{aligned}
y_{it} =\ & \mathbf{x}'_{it}\boldsymbol{\beta} + \delta_1\,\mathrm{NAV}_{it} \\
& +\ \delta_2\,\mathrm{NAV}_{it}\cdot 2\mathrm{Year} + \delta_3\,\mathrm{NAV}_{it}\cdot 3\mathrm{Year} + \delta_4\,\mathrm{NAV}_{it}\cdot 4\mathrm{Year} \\
& +\ \boldsymbol{\gamma}_I + \boldsymbol{\gamma}_T + \boldsymbol{\gamma}_M + \epsilon_{it},
\end{aligned} \tag{8}
$$

The first municipalities that implemented the NAV reform did this at the end of 2006, see Figure 1. We define 2008 as the second year of NAV for those municipalities that implemented the reform in 2006 and 2007. The parameter δ_1 will give the effect of the NAV reform the first year, δ_2 will give the effect the second year, δ_3 after two year, and δ_4 will give the effect the fourth year. Most municipalities had implemented the reform prior to 2010 but some (four) municipalities implemented the reform in 2010, so this year can also be used since we still have some municipalities defining the comparison group in our data.

We also test for learning effects, i.e. if the order of implementation of the reform at the county level matters. We define the variable FP_F as the order of implementation at the county level, and use the following specification

$$
y_{it} =\ \mathbf{x}'_{it}\boldsymbol{\beta} + \delta_1\,\mathrm{NAV}_{it} + \delta_2\,\mathrm{FP}_\mathrm{F} + \boldsymbol{\gamma}_I + \boldsymbol{\gamma}_T + \boldsymbol{\gamma}_M + \epsilon_{it}. \tag{9}
$$

The idea is that municipalities could get feedback from early adopters at the county level, and learn from their implementation process, since each county has their own administration overseeing the implementation process.

We implement the linear probability model in equation (7) as a panel data fixed effect model.

Index model

We also estimate a fixed-effect-logit-models. We start with the following index specification

$$
\begin{aligned}
& y^*_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \delta\,\mathrm{NAV}_{it} + \boldsymbol{\gamma}_D + \boldsymbol{\gamma}_I + \boldsymbol{\gamma}_C + \epsilon_{it} \\
& y_{it} = 1 \text{ if } y^*_i \geq 0 \\
& y_{it} = 0 \text{ if } y^*_i < 0
\end{aligned} \tag{10}
$$

where

$$
\mathrm{E}(y_i|\mathbf{x}_i) = \Pr(y_i = 1|\mathbf{x}_i) = \Pr(\mathbf{x}'_i\boldsymbol{\beta} + \epsilon_i \geq 0) = \mathrm{F}(\mathbf{x}'_i\boldsymbol{\beta}).
$$

We use a logit specification of $F(\cdot)$, and estimate the model using the "xtlogit" model in Stata. We specify a conditional fixed effects model. For details see for instance Arellano (2003), Skrondal and Rabe-Hesketh (2004), Pendergast, Gange, Newton, Lindstrom, Palta, and Fisher (1996).

6 Data set

We start out by using several different data set in the analysis of the effect of the NAV reform. First, we look at three different groups of users of NAV: a) Long-term unemployed (unemployed for more than 6 months), b) School drop-outs (person not finishing high school), c) Long-term sickness absence (length of sick leave is greater than 6 months).

We sample our data in three different ways, where we include person full-filling our inclusion criteria. a) Sample persons up to the beginning of the implementation period of the NAV reform in October 2006, which is our main specification. b) Sample persons up to the NAV reform was implemented in the particular municipality in which the person live. c) Sample person up to and including 2009. Data sampling procedure in c) (and b)) can create endogeneity bias since inclusion of persons in our data set is done also after the implementation of the NAV reform. For instance, if the NAV reform has a positive effect, then this could affect the number of persons in our sample and thus bias the estimated effects. Suppose we sample persons prior the reform and include persons whose length of sick leave are more than 6 months, and look at employment status for instance one year after. Now suppose the NAV reform has a positive effect in the sense of reducing length of sick leave. If a person was sick for more than 6 months prior to the reform, there is a possibility that the same person (or similar persons) would only be sick for instance 5 months and 30 days in the case if NAV had already been implemented and had a positive effect on sickness outcomes. In that case, the person would not be included in the sample, and by that bias the estimated effect on for instance employment status. We use specification b) and c) as a sensitivity analysis. We also do sensitivity analysis on issues relating to persons moving between municipalities.

The NAV reform is not an individual treatment but an organizational reform at the municipality or city district level. Thus it is very important to **cluster** at the municipality level. Our model is non-nested since individuals can move between municipalities from one period to another. We use the clustering option for non-nested models when estimation the standard error in our models. Clustering has a large effect on standard error, since the reform status vary between only 435 municipality or city

districts, while we have several hundred t

   ***

7   Results

The main results from our analysis are reported in Table 1. The effects are generally small and negative, and for some of our specifications we do not find a significant effect of the NAV reform. The NAV reform has not increased the number of employed persons is the main conclusion. Table 1 do not report all our specifications, but the main result is robust to different samples and specifications. Our first set of models

Table 1: Estimating the effect of the NAV reform on employment. Numbers are in percentage points except for the logit model.

	Uempl (1)	Uempl (2)	Sick leave (3)	Sick leave (4)	Drop outs (5)
NAV	-0.92(-2.08)	-1.12(-6.84)	-1.26(-0.87)	-0.72(-0.05)	-0.45(-0.64)
NAV (logit)	-4.31(-3.12)	-5.10(-3.47)	-3.21(-1.37)	-2.82(-0.87)	-0.88(-0.06)
NAV	-1.06(-3.65)	-1.16(-3.96)	-1.25(-0.88)	-0.87(-0.72)	-0.95(-2.42)
NAV FP	0.21(5.04)	0.19(5.79)	0.25(0.87)	0.01(1.86)	0.08(4.24)
NAV 1st year	-1.04(-1.96)	-1.23(-2.19)	-1.43(0.97)	–	–
NAV 2nd year	-0.71(-1.47)	-1.05(-1.59)	-1.42(0.98)	–	–
NAV 3rd year	-0.21(-0.84)	0.44(0.71)	0.49(0.45)	–	–
NAV 4th year	0.15(0.56)	0.85(1.35)	0.36(0.36)	–	–
Time dummies	Yes	Yes	Yes	Yes	Yes
Municip	Yes	Yes	Yes	Yes	Yes
X	Yes	Yes	Yes	Yes	Yes

Note: Numbers are presented in **percentage points**.
t statistics in parentheses. Outcome variable: Employment. Individual fixed-effects dummies are included but are not reported. Other background variables (X) are also included, i.e. dummies for educational level, marital status, children, etc, but are not reported.
Standard error are clustered at the municipality or city district level.
NAV FP = Order of implementation at county level (0-12).
Column 1: Sample up to and including 2006 (long term unemployed). Column 2: Sample up to and including 2009 (long term unemployed). Column 3: Sample up to and including 2006 (long term sick). Column 4: Sample up to and including 2009 (long term sick). Column 5: Sample up to and including 2006 (school drop outs).

reported in the first row in the table is based on equation (7). In this specification, we estimate the effect of the NAV reform using a dummy variable for the reform. The numbers are reported in percentage points. We find that the reform for different groups of potential NAV uses decreases the probability of employment by around one percentage points. This is similar to unconditional measures we find as indicated for

21

instance in Figure 4. The negative effect is significantly different from zero for the unemployment sample, and different versions of this sample. However, after clustering at the municipality level, the effect is estimated to be negative but is not statistically different from zero for the sample of persons on long-term sick leave and school drop-outs, as reported in column 3-5 in Table 1.

We also estimate equation (7) for different groups of users. Table 3 in the appendix shows the results where we analyse the effect of NAV for males and females separately, as well as for small and large municipalities. We find only minor differences for males and females. However, we find a more negative effect of the NAV reform for larger municipalities compared to smaller municipalities, where we have defined large and small based on a cut off level of 5000 inhabitants. Andreassen, Drange, Thune, and Monkerud (2007) concludes that the implementation of the reform has been easier for smaller municipalities. Table 4 shows the results for the logit specifications for the long term unemployed sample, where we show the results divided by gender and also age where we devide the sample in two at age 45 (columns 4 and 5).

The logit specification based on equation (10) confirms the results from the linear probability model. Two things should be commented on regarding this model. First, the estimated regression coefficient is not interpreted as the coefficients in the linear probability model, where the effect is in percentage points. There is no clear cut way to change these logit coefficients to marginal (percentage points) coefficients. Thus, the magnitude of the effects are not comparable to the other models, but the sign and significant level can be interpreted straightforward. Thus, we can conclude from the logit specifications that the effect is negative also in these models. Second, the logit model for panel data (xtlogit) only use the sample for which persons change their employment status. Thus, in these models we do not uses persons that either are unemployed for the entire period, or employed for the entire period, meaning that the sample used in the logit model are always smaller compared to the sample in the other models.

Our third model specification is based on equation (9), confer rows 3-4 in Table 1. Here we define a variable where we test for learning effects, i.e. if the order of implementation of the reform at the county level matters. We find a positive learning effect at the county level. Thus, those municipalities that implemented the reform late compared to others in their county have slightly better effect compared the those that implemented the reform early. Again, these effects are significant only for the sample of long term unemployed.

Our last specification is based on equation 8, see rows 5-8 in Table 1. In this specification we allow for the fact that the effect of the NAV reform can change over

time by including interaction terms between the NAV dummy and dummies for number of years after the implementation year. We define 2008 as the second year of NAV for those municipalities that implemented the reform in 2006 and 2007. These regressions indicates that the effect of the NAV reform improves slightly over time. We find a significant negative effect on employment the first year. The effect is less negative for the second year, and tends to become zero for the last year. However, we find very few significant effects for this specification.

We also estimate the models using social insurance and social assistance as outcome variables, both as a linear probability model and conditional logit model. For both these two outcome variables we do not find indications that the reform has had any positive effects, and the results confirm our conclusion from Table 1 that the point estimate are unfavourable from a welfare state point of view. Persons living in municipalities that have implemented the reform at a given point in time have higher use of social insurance and social assistance compared to persons living in municipalities that have not yet implemented the reform. The results for social insurance are reported in Table 2 below. In the specification where we estimate the effect of the NAV reform on the use of social assistance defined as an indicator variable and where NAV is a dummy variable as in equation (7) we find that the use of social assistance is increased by almost 1 percentage points due to the NAV reform, and thus the reform is not successful in decreasing the use of social assistance. This result is robust to different data samples and econometrics specifications. The logit specification based on equation (10) confirms the results from the linear probability model. The specification based on equation (9) indicates some positive learning effects, i.e. if the order of implementation of the reform at the county level matters. Municipalities that implemented the effect late at the county level has a slightly better effect compared the those that implemented the effect early.

Our last specification reported in Table 2 is based on equation 8. In this specification we allow for the fact that the effect of the NAV reform can change over time by including interaction terms between the NAV dummy and dummies for number of years after the implementation year. We find very few significant effects, and we cannot conclude the the NAV reform have had any favourable effects on social assistance.

23

Table 2: Estimating the effect of the NAV reform on social assistance. Numbers are in percentage points.

	Uempl (1)	Uempl (2)	Sick leave (3)	Sick leave (4)	Drop outs (5)
NAV	0.76(1.98)	0.84(3.84)	0.87(0.76)	0.98(0.45)	0.96(1.24)
NAV (logit)	2.23(2.32)	3.82(2.61)	2.71(1.69)	2.38(1.12)	1.88(1.52)
NAV	1.02(2.37)	1.23(2.42)	0.95(1.18)	0.87(0.72)	1.03(1.62)
NAV FP	-0.18(-2.13)	-0.19(2.37)	-0.21(-1.31)	-0.08(-1.64)	-0.18(-1.76)
NAV 1st year	0.98(1.97)	0.99(2.15)	1.03(1.18)	1.07(1.49)	-
NAV 2nd year	0.54(1.38)	0.76(1.64)	0.98(0.87)	0.76(1.11)	-
NAV 3rd year	0.21(-0.82)	-0.23(-0.87)	-0.37(-0.92)	0.07(1.03)	-
NAV 4th year	-0.09(-0.76)	-0.39(-1.16)	-0.31(-1.51)	-0.17(-1.15)	-
Time dummies	Yes	Yes	Yes	Yes	Yes
Municip	Yes	Yes	Yes	Yes	Yes
X	Yes	Yes	Yes	Yes	Yes

Note: Numbers are presented in **percentage points** except for the logit specification.

t statistics in parentheses. Outcome variable: Employment. Individual fixed-effects dummies are included but are not reported. Other background variables (X) are also included, i.e. dummies for educational level, marital status, children, etc, but are not reported.

Standard error are clustered at the municipality or city district level.

NAV FP = Order of implementation at county level (0-12).

Column 1: Sample up to and including 2006 (long term unemployed). Column 2: Sample up to and including 2009 (long term unemployed). Column 3: Sample up to and including 2006 (long term sick). Column 4: Sample up to and including 2009 (long term sick). Column 5: Sample up to and including 2006 (school drop outs).

8   Conclusion

A national labour and welfare administration reform (NAV) is evaluated on employment outcomes and use of social insurance and social assistance. We focused on groups that have a high risk of needing the services from NAV. The reform was implemented step-wise over time at different geographical locations. Thus, we are able to create treatment and control groups to facilitate causal interpretations of the empirical results. We use a unique data set on individuals where we match local implementation of the reform to individual data on social security, employment, and socioeconomic characteristics. The labour and welfare administration was until 2006 fragmented, involving three large public agencies (employment, social assistance, and social insurance services), with limited coordination and collaborations. The reform merged these institutions into a new organization called NAV, after the "one door for all" principle. The idea was to help users to get faster back to work after a period of sickness or unemployment through coordinated help by generalist case worker competence.

We find that the reform failed to reap the intended benefits of the merger process. We find that the NAV reform decreased the individual employment propensity by around 1 percentage points. However, there seems to be indications that the effect of NAV improves slightly over time.

The lack of positive effects of the NAV reform could be related to the fact that most users still demand specialized case worker competence and to inadequate planning of IT infrastructure. Large bureaucratic reforms can take many years to implement after a turbulent implementation face, and can have large short term costs for users, as indicated by this specific reform. The turbulence of the reform was not anticipated even though 25 municipalities and city districts participated in a pilot project 6 months before the country-wide implementation process started.

References

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

AAKVIK, A., J. J. HECKMAN, AND E. J. VYTLACIL (2005): "Estimating treatment effects for discrete outcomes when responses to treatment vary: an application to Norwegian vocational rehabilitation programs," *Journal of Econometrics*, 125(1), 15–51.

AAKVIK, A., T. H. HOLMÅS, AND E. KJERSTAD (2012): "Hospital Capacity, Waiting Times and Sick Leave Duration–An Empirical Analysis of a Norwegian Health and Labour Policy Reform," Working paper (10) 2012, Deparment of Economics, University of Bergen.

ANDREASSEN, T. A., I. DRANGE, T. THUNE, AND L. C. MONKERUD (2007): "På vei mot en integrert velferdsforvaltning? Erfaringer fra pilotprosessen i den nye arbeids- og velferdsforvaltningen.," *AFI-rapport 4/2007*.

ANGRIST, J. D., AND J.-S. PISCHKE (2008): *Mostly harmless econometrics: An empiricist's companion.* Princeton University Press.

ARELLANO, M. (2003): *Panel data econometrics.* Oxford University Press.

ASKIM, J., T. CHRISTENSEN, A. L. FIMREITE, AND P. LÆGREID (2010): "How to assess administrative reform? Investigating the adoption and preliminary impacts of the Norwegian welfare administration reform," *Public Administration*, 88(1), 232–246.

BHULLER, M., T. HAVNES, E. LEUVEN, AND M. MOGSTAD (2013): "Broadband Internet: An Information Superhighway to Sex Crime?," *The Review of Economic Studies*, nn(n), nn–nn.

CHAMPION, C., AND G. BONOLI (2011): "Institutional fragmentation and coordination initiatives in western European welfare states," *Journal of European Social Policy*, 21(4), 323–334.

CHRISTENSEN, T., M. KNUTH, P. LÆGREID, AND J. WIGGAN (2009): "Reforms of welfare administration and policy–a comparison of complexity and hybridization: an introduction," *Intl Journal of Public Administration*, 32(12), 1001–1005.

CHRISTENSEN, T., AND P. LÆGREID (2010): *Reforming Norway's welfare administration.* Statskontoret Sweden.

COX, D. R. (1958): *Planning of experiments.* Wiley.

DAHL, E. S., AND A. M. GALAASEN (2013): "I pose og sekk? Hvorfor velge å kombinere jobb og pensjon," *Nav publication: Arbeid og velferd*, 3, 56–67.

EICHHORST, W., AND L. C. KAISER (2006): "The German labor market: still adjusting badly?," *IZA discussion paper*.

EKHAUGEN, T. (2006): "Utfall av yrkesrettet attføring i Norge 1994-2000," Report 6, Ragnar Frisch Centre for Economic Research.

EVERHARDT, T., AND P. JONG (2011): "Return to Work After Long Term Sickness: The Role of Employer Based Interventions," *De Economist*, 159(3), 361–380.

FEVANG, E., S. MARKUSSEN, AND K. RØED (2013): "Til, fra og mellom inntektssikringsordninger - før og etter NAV," *Frisch Report 1/2013*, 1.

FEVANG, E., K. RØED, L. WESTLIE, AND T. ZHANG (2004): "Veier inn i, rundt i, og ut av det norsketrygde- og sosialhjelpssystemet," *Frisch Report 6/2004*, 6.

FISHER, R. A. (1935): *The design of experiments.* Oliver & Boyd.

HANSEN, H.-T., K. G. LUNDBERG, AND L. J. SYLTEVIK (2013): *Nav - med brukeren i sentrum?* Universitetsforlaget.

HECKMAN, J. J., AND B. E. HONORE (1990): "The empirical content of the Roy model," *Econometrica*, pp. 1121–1149.

HECKMAN, J. J., R. J. LALONDE, AND J. A. SMITH (1999): "The economics and econometrics of active labor market programs," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3, chap. 31, pp. 1865–2097. Elsevier.

JACOBI, L., AND J. KLUVE (2006): "Before and after the Hartz reforms: The performance of active labour market policy in Germany," *IZA discussion paper*.

JOHANSSON, P., AND M. PALME (1996): "Do economic incentives affect work absence? Empirical evidence using Swedish micro data," *Journal of Public Economics*, 59(2), 195–218.

KARAGIANNAKI, E. (2007): "Exploring the effects of integrated benefit systems and active labour market policies: evidence from Jobcentre Plus in the UK," *Journal of Social Policy*, 36(2), 177.

LAZEAR, E. P. (2000): "Performance Pay and Productivity," *American Economic Review*, 90(5), 1346–1361.

MARKUSSEN, S., AND K. RØED (2014): "The Impacts of Vocational Rehabilitation," Discussion Paper 7892, IZA.

MINAS, R. (2014): "One-stop shops: Increasing employability and overcoming welfare state fragmentation?," *International Journal of Social Welfare*.

NEYMAN, J. (1923): "Statistical problems in agricultural experimentation," *Supplement to the Journal of the Royal Statistical Society*, 2(2), 107–180.

PENDERGAST, J. F., S. J. GANGE, M. A. NEWTON, M. J. LINDSTROM, M. PALTA, AND M. R. FISHER (1996): "A survey of methods for analyzing clustered binary response data," *International Statistical Review/Revue Internationale de Statistique*, pp. 89–118.

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

ROY, A. D. (1951): "Some thoughts on the distribution of earnings," *Oxford Economic Papers*, 3(2), 135–146.

RUBIN, D. B. (1978): "Bayesian inference for causal effects: The role of randomization," *The Annals of Statistics*, pp. 34–58.

SCHREINER, R., AND S. MARKUSSEN (2012): "NAV-reformen: "Færre i arbeid og lenger på trygd"," *Søkelys på arbeidslivet*, 1-2, 151–162.

SCHREINER, R. C. (2012): "NAV-reformen: Flere i arbeid - færre på trygd," *Frisch Report 1/2012*, 1.

SKRONDAL, A., AND S. RABE-HESKETH (2004): *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. CRC Press.

VAN BERKEL, R., AND V. BORGHI (2008): "Introduction: The governance of activation," *Social Policy and Society*, 7(3), 331–340.

VERBEEK, M. (2008): *A Guide to Modern Econometrics*. Wiley.

WOLFERS, J. (2006): "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results," *The American Economic Review*, 96(5), pp. 1802–1820.

WZB, AND INFAS (2005): *Evaluation der Massnahmen zur Umsetzung der Vorschlaege der Hartz-Kommission*. Research Report of the Federal Ministry for Labour and Social Affairs.

Sample of long term unemployed

Table 3: Estimating the effect of the NAV reform on employment for different groups. All/males/female/small/large.

	(1)	(2)	(3)	(4)	(5)
	Empl	Empl	Empl	Empl	Empl
2004.year	-0.100***	-0.0954***	-0.105***	-0.112***	-0.0991***
	(-63.52)	(-41.71)	(-48.22)	(-22.10)	(-59.62)
2005.year	-0.231***	-0.225***	-0.238***	-0.256***	-0.229***
	(-146.22)	(-98.13)	(-108.92)	(-50.61)	(-137.35)
2006.year	-0.329***	-0.328***	-0.330***	-0.340***	-0.328***
	(-202.52)	(-139.40)	(-147.40)	(-66.40)	(-191.41)
2007.year	-0.154***	-0.129***	-0.178***	-0.194***	-0.150***
	(-93.32)	(-54.08)	(-78.21)	(-35.57)	(-86.82)
2008.year	-0.0387***	-0.0230***	-0.0544***	-0.0936***	-0.0342***
	(-21.26)	(-8.78)	(-21.59)	(-14.55)	(-17.99)
2009.year	-0.00800***	-0.00773**	-0.00937***	-0.0557***	-0.00472**
	(-3.65)	(-2.45)	(-3.08)	(-7.29)	(-2.06)
2010.year	-0.00314	-0.00685**	-0.000972	-0.0461***	-0.000320
	(-1.38)	(-2.09)	(-0.31)	(-6.01)	(-0.13)
NAV	-0.0112***	-0.00930***	-0.0130***	0.00771	-0.0115***
	(-6.84)	(-3.93)	(-5.76)	(1.35)	(-6.71)
Constant	0.292***	0.295***	0.290***	0.282***	0.293***
	(163.57)	(120.27)	(109.56)	(50.90)	(155.58)
FE	Yes	Yes	Yes	Yes	Yes
Municip	Yes	Yes	Yes	Yes	Yes
X	Yes	Yes	Yes	Yes	Yes
NT	754097	367548	386549	71390	682707

t statistics in parentheses

Outcome variable: Employment. Background variables (X) are also included, i.e. dummies for educational level, marital status, children, etc, but these are not reported.

Column 1: All individuals. Column 2: Males. Column 3: Females.

Column 4: Small municipality. Column 5: Large municipality.

Sample of long term unemployed, where sampling is up to and including 2009.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Estimating the effect of the NAV reform on employment. Fixed-effects logit estimates. (34)

| | (1) | (2) | (3) | (4) | (5) |
	Empl	Empl	Empl	Empl	Empl
Empl(dummy)					
2004.year	-0.825***	-0.759***	-0.894***	-0.849***	-1.123***
	(-62.39)	(-40.60)	(-47.75)	(-45.12)	(-45.98)
2005.year	-2.219***	-2.061***	-2.395***	-2.266***	-2.928***
	(-140.52)	(-93.96)	(-104.82)	(-100.62)	(-92.86)
2006.year	-24.56	-23.80	-23.35	-25.11	-23.57
	(-0.04)	(-0.03)	(-0.05)	(-0.02)	(-0.05)
2007.year	-1.558***	-1.222***	-1.927***	-1.477***	-2.700***
	(-101.80)	(-58.03)	(-85.76)	(-67.75)	(-85.11)
2008.year	-0.606***	-0.374***	-0.857***	-0.433***	-1.837***
	(-36.80)	(-16.38)	(-35.89)	(-18.28)	(-55.80)
2009.year	-0.417***	-0.294***	-0.563***	-0.191***	-1.691***
	(-20.70)	(-10.55)	(-19.23)	(-6.61)	(-41.90)
2010.year	-0.418***	-0.310***	-0.554***	-0.170***	-1.726***
	(-19.76)	(-10.55)	(-18.02)	(-5.56)	(-40.93)
NAV (logit coeff)	-0.0494***	-0.0330	-0.0700***	-0.0472**	-0.0375
	(-3.25)	(-1.56)	(-3.19)	(-2.21)	(-1.20)
FE	Yes	Yes	Yes	Yes	Yes
Municip	Yes	Yes	Yes	Yes	Yes
X	Yes	Yes	Yes	Yes	Yes
NT	434881	215201	219680	221500	123013
R^2					
LL	-133478.7	-67454.3	-65441.6	-67026.7	-34377.2
AIC	266991.4	134942.7	130917.1	134087.5	68788.5

t statistics in parentheses

Outcome variable: Employment. Individual fixed-effects dummies are included but are not reported. Other background variables (X) are also included, i.e. dummies for educational level, marital status, children, etc, but are not reported. Logit estimates.

Column 1: All individuals. Column 2: Males. Column 3: Females.

Column 4: Below age 45. Column 5: Above age 45.

Sample of long term unemployed, where sampling is up to and including 2009.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Estimating the effect of the NAV reform on employment. **Random-effects** estimates.

| | (1) | (2) | (3) | (4) | (5) |
	Empl	Empl	Empl	Empl	Empl
2004.year	-0.111***	-0.105***	-0.116***	-0.120***	-0.110***
	(-65.42)	(-43.24)	(-49.43)	(-22.47)	(-61.49)
2005.year	-0.251***	-0.242***	-0.258***	-0.270***	-0.248***
	(-148.04)	(-99.30)	(-110.38)	(-50.60)	(-139.25)
2006.year	-0.340***	-0.340***	-0.341***	-0.352***	-0.339***
	(-197.25)	(-136.32)	(-143.05)	(-65.16)	(-186.28)
2007.year	-0.177***	-0.152***	-0.202***	-0.214***	-0.174***
	(-99.84)	(-59.44)	(-82.17)	(-37.12)	(-93.07)
2008.year	-0.0622***	-0.0475***	-0.0772***	-0.115***	-0.0575***
	(-31.55)	(-16.79)	(-28.18)	(-16.92)	(-27.91)
2009.year	-0.0313***	-0.0314***	-0.0323***	-0.0809***	-0.0275***
	(-13.27)	(-9.26)	(-9.82)	(-9.96)	(-11.14)
2010.year	-0.0259***	-0.0306***	-0.0226***	-0.0699***	-0.0227***
	(-10.58)	(-8.68)	(-6.63)	(-8.58)	(-8.83)
NAV	-0.00477***	-0.00331	-0.00620**	0.00943	-0.00480***
	(-2.69)	(-1.30)	(-2.52)	(1.55)	(-2.59)
Constant	0.290***	0.296***	0.284***	0.286***	0.291***
	(153.09)	(113.60)	(100.68)	(49.03)	(145.19)
FE	Yes	Yes	Yes	Yes	Yes
Municip	Yes	Yes	Yes	Yes	Yes
X	Yes	Yes	Yes	Yes	Yes
NT	639936	314801	325135	62127	577809

t statistics in parentheses

Outcome variable: Employment. Background variables (X) are also included, i.e. dummies for educational level, marital status, children, etc, but these are not reported.

Column 1: All individuals. Column 2: Males. Column 3: Females.

Column 4: Small municipality. Column 5: Large municipality.

Sample of long term unemployed, where sampling is up to and including 2006.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Sample of long term sick

Table 6: Descriptive statistics (means and standard deviations) 2007-2010: Full sample (2)

|	All individuals		Nav=0 group		Nav=1 group	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Outcome variables:						
Empl(dummy)	0.168	0.369	0.170	0.375	0.165	0.352
SI(dummy)	0.719	0.450	0.712	0.453	0.735	0.441
SA(dummy)	0.0874	0.282	0.0892	0.285	0.0826	0.275
SI amount/1000	129.4	90.79	128.5	91.44	131.7	89.03
WorkIncome/1000	37.16	93.79	39.28	97.67	31.62	82.60
WorkIncome/1000/zeros	176.2	136.0	180.4	140.1	163.8	122.1
Wyrkinnt/1000	95.39	179.8	96.49	180.8	92.54	177.2
Education level:						
Edu_2	0.346	0.476	0.347	0.476	0.342	0.474
Edu_3	0.186	0.389	0.183	0.387	0.193	0.395
Edu_4	0.260	0.439	0.260	0.439	0.259	0.438
Edu_5	0.0329	0.178	0.0339	0.181	0.0303	0.172
Edu_6	0.138	0.345	0.136	0.343	0.144	0.351
Edu_7	0.0210	0.144	0.0213	0.144	0.0203	0.141
Other variables:						
Married	0.467	0.499	0.465	0.499	0.474	0.499
Divorced	0.224	0.417	0.228	0.419	0.215	0.411
Child0_18	0.877	1.122	0.881	1.127	0.868	1.108
Child0_6	0.304	0.663	0.306	0.664	0.300	0.659
Age	44.70	9.894	44.60	9.879	44.94	9.929

Note: Sample of long term sick, where sampling is up to and including 2006 (first year=2007).

Figure 7: Average unemployment rate for two groups of municipalities; those that implemented the reform in the second quarter of 2007 (blue line) and those that implemented the reform in the second quarter of 2009. Unadjusted monthly data. Source: Statistics Norway (SSB) and Norwegian Social Science Data Services (NSD).
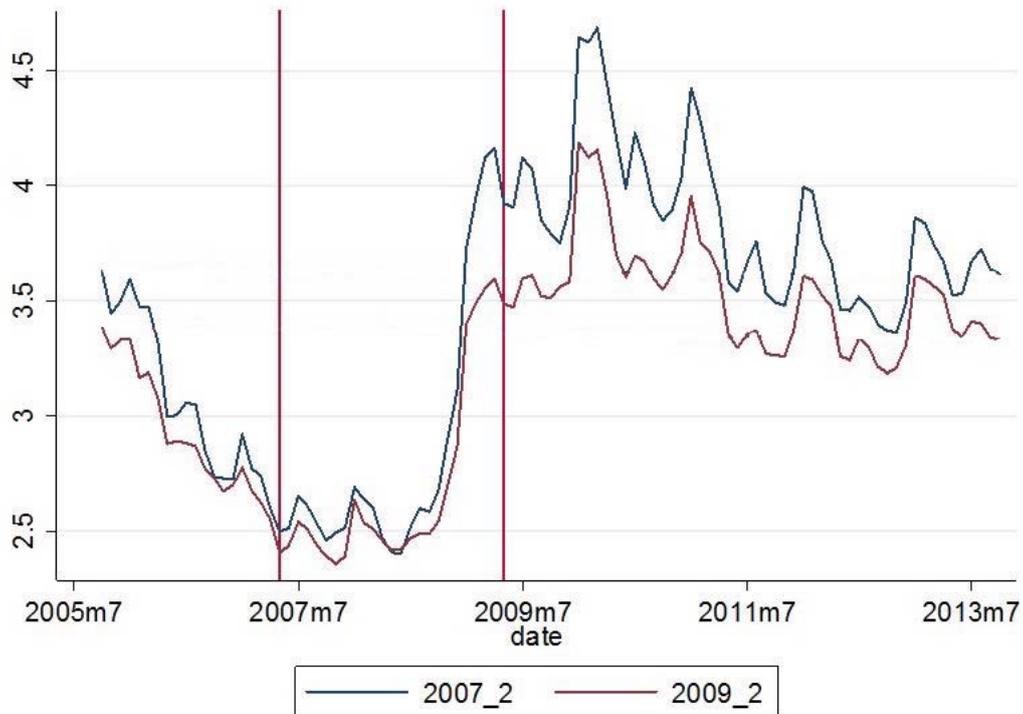
Note: The period in which we have a treatment group (the 2007 group) and a comparison group (the 2009) group is defined as the period between May 2007 and May 2009.

Figure 8: Average unemployment rate for two groups of municipalities; those that implemented the reform in the fourth quarter of 2007 (blue line) and those that implemented the reform in the fourth quarter of 2009. Unadjusted monthly data. Source: Statistics Norway (SSB) and Norwegian Social Science Data Services (NSD).
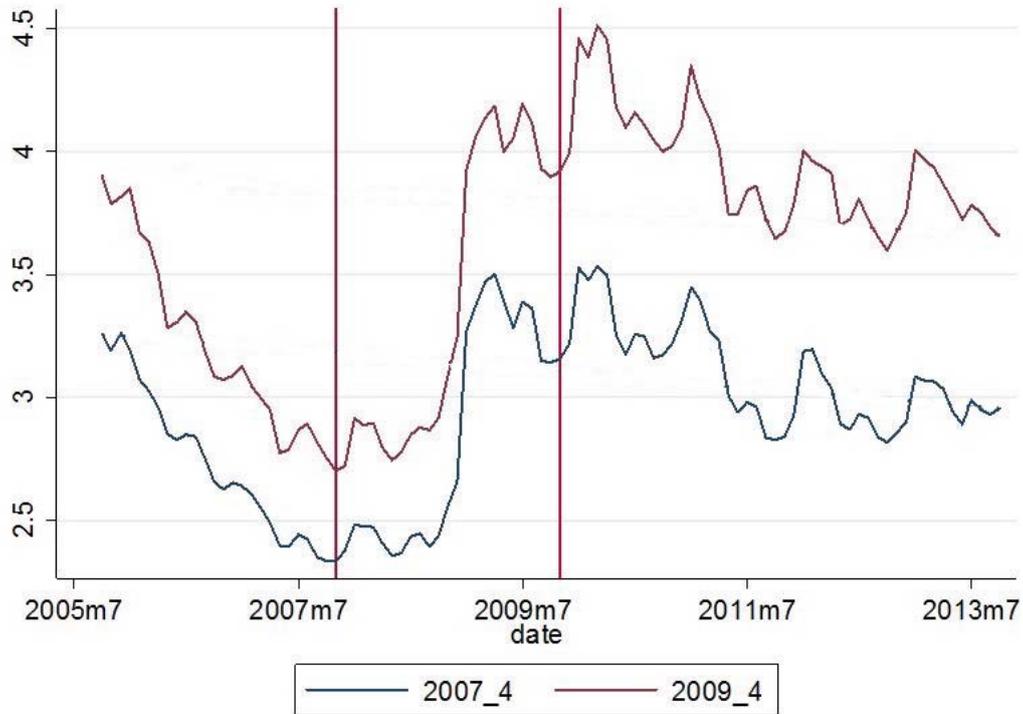
Note: The period in which we have a treatment group (the 2007 group) and a comparison group (the 2009) group is defined as the period between November 2007 and November 2009.