UiT THE ARCTIC UNIVERSITY OF NORWAY

School of Business and Economics

Absenteeism in Norway

How does the North differ from the rest? A panel data study of the period from 2002 to 2011

—

Thomas Nyrud
Master thesis in Economics May 2014

Acknowledgements

I would first and foremost like to thank my supervisor Mikko Moilanen for his excellent guidance and encouragement throughout the entire writing process.

I would also like to thank Bente Ødegaard and Oddmund Klæboe at NAV Troms for providing valuable input at our meetings, and for providing data for analysis.

A thank you also goes to my family, friends and girlfriend for their encouragement and support in this period. A special thank you goes to my girlfriend for proof-reading and commenting on my drafts, and for being patience with me.

Abstract

A panel data study of sick leave behavior in the Norwegian population is conducted on a municipality level for the period from 2002 to 2011. Data on sickness insurance usage is fitted against a selection of variables describing different regional characteristics that have been linked with absenteeism in the earlier literature. A basic model for sick leave usage is constructed, and further divided into two slightly different submodels where one is used to analyze sick leave behavior in the 43 northernmost municipalities of Norway, which are located in the counties Troms and Finnmark. The other submodel is used to look at the remaining 385 municipalities. The purpose is to explore why the northern municipalities have an overall higher level of sickness insurance usage than the rest of the country. The performed analysis further looks at men and women separately, as there exists significant gender differences in sick leave usage. The results reveal that especially unemployment, bankruptcies and downsizing are important factors on a national level, but lose explanatory value when zooming in on the north. The average education level is lower among the northern population, and the estimation results strongly suggest that this might be an important part of the answer when asking why sick leave is higher in this part of the country. Further, municipalities with a higher share of female workers between 16 and 25 years of age are found to have lower total sick leave when looking at the 385 municipalities outside of Troms and Finnmark. When shifting focus to the north however, a higher share of this age group is found to predict *higher* total sick leave usage.

Key words: Sick leave usage, absenteeism, social insurance, regional differences, health

Table of Contents

1 Introduction....4

2 Background and previous research....8

2.1 Rules and regulations....8

2.2 Gender differences....9

2.3 Troms and Finnmark....10

3 The individual and absence behavior....14

4 A model for sick leave usage....20

4.1 Dependent variables....20

4.2 Explanatory variables....22

4.2.1 Demographic variables....22

4.2.2 Social variables....26

4.2.3 Labor market variables....28

5 Data and methodology....31

5.1 Variable definitions and descriptive statistics....32

5.1.1 Dependent variables....33

5.1.2 Demographic variables....35

5.1.3 Social variables....39

5.1.4 Labor market variables....41

5.2 Methodology....44

6 Estimation results....48

6.1 Demographic effects....52

6.2 Social effects....53

6.3 Labor market effects....54

7 Discussion....55

Appendix A: Overview of variables used as regression input....66

Appendix B: Hausman-Taylor regression output....67

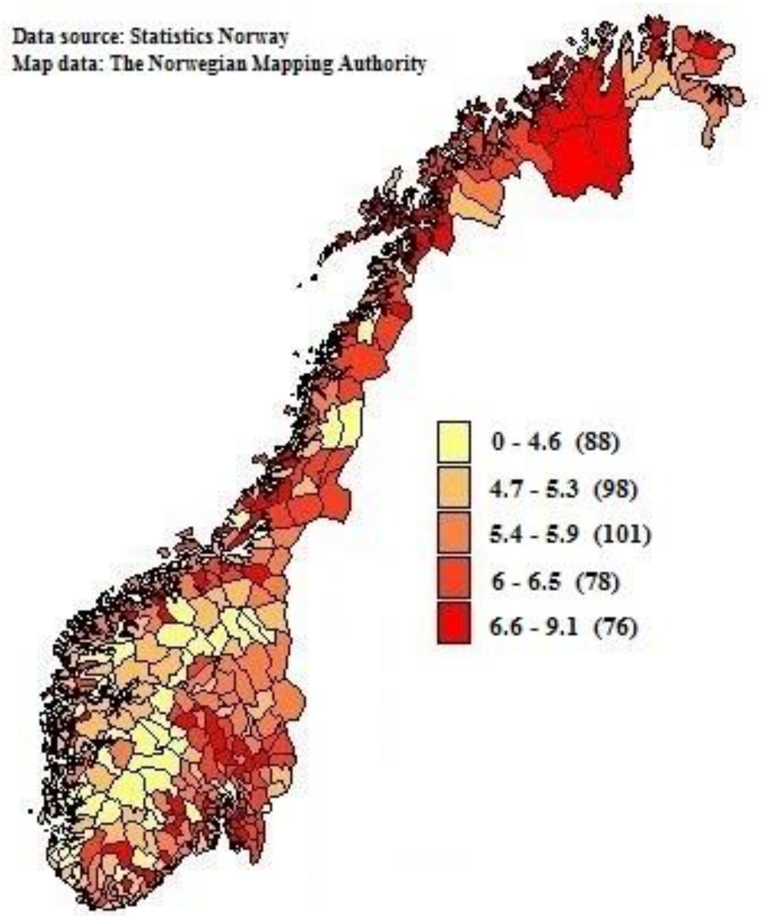
1 Introduction

Scientific research over the past decades has shown time and time again that there exists no definite relationship between health and sickness insurance usage. One study shows that through the years from 1996 to 2003 total sickness insurance usage in Norway increased substantially, but without any significant changes in the general public health taking place (Ihlebaek, Brage & Eriksen 2007). In another study Markussen, Røed, Røgeberg & Gaure (2009) examine determinants of sickness absence behavior and find that when the initial right to full income compensation has passed and the insurance receiver is transferred to other less generous benefit schemes, the probability of returning to working life increases exponentially. It is apparent that beside purely health related factors, the sick leave decision is highly influenced by other factors as well.

For the national economy, the costs surrounding sick leave are substantial. There are the direct costs of insurance payments, but also more indirect costs connected with forgone labor supply. Total payments of social insurance benefits by the Norwegian Labor and Welfare Service (NAV) in 2012 were at 344 billion Norwegian kroner (Bjørnstad, 2013). This sum constituted 16% of GDP for mainland-Norway the same year, which was an increase from 14% of GDP in 2007, but a slight decrease from 17% in 2003. Bjørnstad (2013) further finds that the total expenses of social insurance payments increased by 11% over the period from 2003 to 2012, but when adjusting for the growth in number of employed he finds an actual *decrease* of 4%.

The map of Norway below shows the sick leave percentage, which depicts the share of lost work days due to sick leave, on a municipality level in the fourth quarter of 2013.

Figure 1.1: The sick leave percentage on a municipality level in the fourth quarter of 2013



Physician-certified sick leave for workers, by municipality and sick leave percentage.
Time: Fourth quarter of 2013

The map reveals a clear overall darkening of the shades when moving further north in the country. When ranking the municipalities from highest to lowest sick leave, it turns out that 19 out of the 60 municipalities with the most sick leave usage this quarter were located in the northernmost counties of Troms and Finnmark. This constitutes approximately one-third of the highest values, which is impressive when considering that these two counties only hold one-tenth of the total municipalities in Norway as of 2013[1]. For the fourth quarter of 2013 the sick leave percentage was at 5.4% on a national level, while it was at 6.3% in Troms and 6.6% in Finnmark.

The present paper is focusing on sickness insurance usage in the north versus the rest of Norway. The north of Norway is experiencing a rapidly increasing demand for competent workers in both the public and private sector (Sparebank 1 Economic Survey for Northern-Norway, 2009). This demand is only expected to grow further in the coming years.

According to the quoted survey, there are three possible ways of increasing the labor supply. The first is to draw workers from other parts of the country, or from abroad, by active marketing of the northern regions as an attractive employer. The second is to increase labor force recruitment by securing that young people finish their education and settle down in the region afterwards. The third way is a more effective exploitation of the existing labor reserves, by *reducing sick leave* and disability benefits usage, and by keeping people in the labor force for a longer time before retirement. For the northern municipalities, all of these points will be important areas of study in the coming years. More in-depth information on the structure of sick leave in Northern-Norway might help with tapping into this source of unused labor supply, and it is therefore an important area to explore.

As seen, sick leave usage is significantly much higher in the north of Norway than in the rest of the country, both among male and female workers. By the second quarter of 2009 the difference in sick leave days taken between Northern-Norway[2] and the rest of the country was approximately equivalent to 1500 man-years (Sparebank 1 Economic Survey for Northern-Norway, 2009).

The Nordic Council of Ministers (2007) performed a phone-survey among 1000 randomly chosen individuals from each of the five Nordic countries; Norway, Sweden, Denmark, Iceland and Finland. The survey first asked the respondents about nine commonly used reasons for utilizing sick leave, where they were asked to rate how acceptable they found these reasons on a scale from 1 to 10. Next they were asked to suggest what they considered to be an acceptable length of the sickness spell in connection with 6 out of the originally 9 reasons. The results on Norway reveal that harassment in the workplace, difficulties connected with divorce/break-ups and having relatives that need support and care are among the most accepted reasons for absenteeism in Norwegian workplaces. On the other end of the scale, lack of sleep and feeling ill after drinking the night before are found to be least accepted. When reviewing the results, the survey-takers note that Norwegian employees are found to overall have a conservative view on sick leave usage.

The results on Norway are further divided by regions, and Northern-Norway[3] was found to have the highest level of accept on 5 out of the 9 specified reasons (out of 6 regions in total). Reasons for sick leave more accepted among northern-Norwegians included harassment in the workplace, dissatisfaction with conditions in the workplace and feeling ill after drinking (although this last one is found to have in general low accept also in the north). When reviewing the results on accepted length, they found the highest accepted average length in Northern-Norway on 4 out of 6 surveyed reasons. These included stress in the workplace and difficulties connected with divorce/break-ups.

It appears by these findings that the general attitude towards sick leave usage might be more open in the north than in the rest of the country. These types of effects, which might be defined as a general attitude or «culture», will influence sick leave behavior to a large extent. They are however difficult to quantify, and thus difficult to measure by conventional analytical models like the one presented in this paper. An attempt to capture some of these effects is made nonetheless, and as we will see, the results reveal some rather interesting findings.

The present paper will perform a panel data analysis of sick leave usage on a municipality level for the period from 2002 to 2011. A basic model will be constructed, and further subdivided into two almost identical models where one will be used on the municipalities in Troms and Finnmark, and the other on the municipalities in the rest of Norway. Regression analysis will then be performed in an attempt to define the relationships between sick leave and a selection of possible explanatory factors. The results for the north and for the rest of Norway will be compared in a search for differences that might explain why Troms and Finnmark experience a higher degree of sick leave.

The rest of the paper will be structured as follows. Section 2.1 presents the rules and regulations surrounding sick leave usage in Norway. Section 2.2 then explores gender differences in sick leave usage. Section 2.3 gives a more in-depth presentation of the sick leave percentage in Troms and Finnmark. Some earlier findings on average sickness spell length in the north is also presented, and an overview of how the labor market structure looks in the north compared to on a national level is given. Chapter 3 gives a brief introduction to some of the most common individual level theories surrounding sick leave behavior. Chapter 4 constructs the basic model for sick leave usage to be used in the regression analysis, and links the different variables used to existing literature. Chapter 5 presents the data and how the variables are constructed. It is also shown how the basic model is divided into two sub-models for the analysis, one for the north and one for the rest of the country. An overview of the chosen methodology is also included in this chapter.

Chapter 6 presents the estimation results for both models. Chapter 7 discusses the estimation results more in depth and gives some concluding remarks.

2 Background and previous research

This chapter will give an overview of sickness insurance usage in Norway. The existing legal framework and how it is governed in practice is covered in section 2.1. Section 2.2 gives some insight into the gender differences observed in sick leave patterns based on previous research. Lastly, section 2.3 zooms in on the northernmost counties of Troms and Finnmark and explores the sick leave percentage, average length of sickness spells and differences in labor market structure between the north and the rest of Norway.

2.1 Rules and regulations

Public social insurance in Norway is regulated through the National Insurance Act (Lov om Folketrygd – Folketrygdloven). All facts presented in this section can be referred back to this act, if not stated otherwise. The purpose of the act is described as follows on the webpages of the Norwegian foundation Lovdata[4] (translated from Norwegian):

"…to provide economic security by securing income and compensating for special expenses connected with unemployment, pregnancy and childbirth, single care for children, sickness and injury, disability, old age and death. The National Insurance should further contribute to an equalization of income throughout the individuals' life span and between groups of individuals…"

Most of the administrative work surrounding the National Insurance Act is as of 2006 performed by the Norwegian Labor and Welfare Service (NAV – Arbeids- og Velferdsetaten).

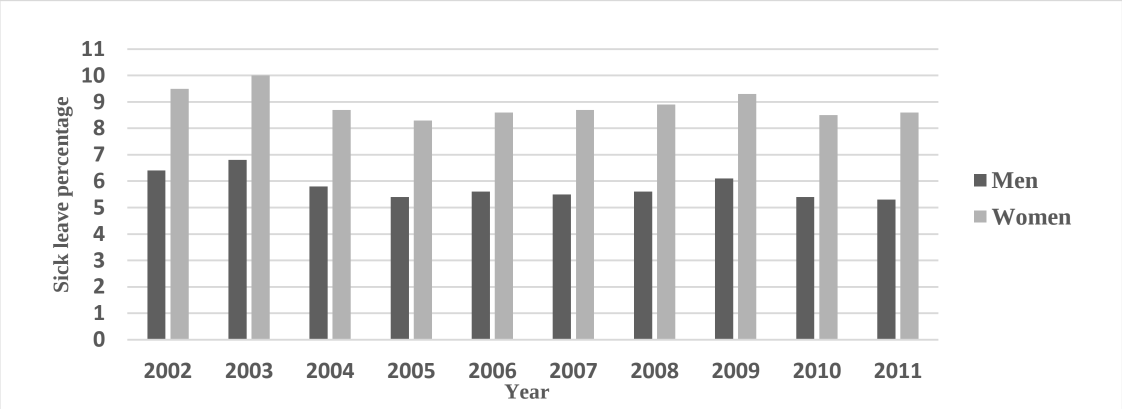
For a worker to earn the right to sickness insurance, he or she needs to have been employed for at least four weeks in the immediate time before the sickness spell. The Norwegian sick pay scheme covers 100% of regular earnings from the first day of sick leave. For the first sixteen days of absence the employer provides coverage, after which the government takes over and provides full salary coverage for up to one year. After one year, those who are still not able to return to work are transferred to some form of rehabilitation program such as work assessment allowance (AAP – Arbeidsavklaringspenger) which normally covers approximately 66% of original income. While on such programs the insurance receiver is obliged to actively attempt returning to work-life.

Sickness spells lasting more than three days must be certified by a physician, but if the employer participates in the IA-agreement[5] certification is not necessary until after the eighth day. The possibility of using self-certified sick leave is only granted after the worker has been employed a minimum of two months.

2.2 Gender differences

The regressions performed in the present paper looks at sick leave among men and women separately. It is therefore natural to examine more closely why exactly this might be a good idea. Graph 2.1 depicts the sick leave percentage for each of the years from 2002 to 2011 for both genders. Men had an average sick leave percentage of 5.79% over the period, while the female average was 8.91%. The percentage for women is significantly higher in each of the observed years, and this gender difference has come to be a well-established fact in the literature.

Graph 2.2.1: The sick leave percentage for the whole of Norway. 2002-2011. Divided by gender.



Notes: All used data obtained from Statistics Norway

The use of sick leave is higher among women than men, both when looking at married, separated/divorced and single workers, with or without children (Markussen, Røed, Røgeberg & Gaure, 2009).

The higher level of sick leave usage among women between 21-39 years of age might be partially explained by pregnancy-related sick leave (Hauge & Kann, 2007). However, Markussen, Røed, Røgeberg and Gaure (2009) study Norwegian register data and find that women have higher entry rates to certified sick leave than men, even when controlling for absence connected with pregnancies. They find that for minor diseases, the entry rates are 45-68 percent higher for women, and for major diseases they are 26-43 percent higher.

Some commonly used explanations for the gender differences are that women have a lower threshold for seeking out medical expertise, and that there is a majority of women in lower-paid jobs where sick leave is naturally higher (NOU 2000:27). Hauge & Kann (2007) further list inconvenient working hours and physically straining work tasks in female-dominated occupations as possible explanations. The theories on occupational differences between the genders are however to a large extent disproved by Mastekaassa (2000), who compares men and women in the same jobs at the same workplace and finds that the higher sickness absence among women is not explained by less healthy work environments.

The significant difference between the genders cannot be ignored, and the further analyses performed here will try to look at men and women separately wherever such separation is possible.

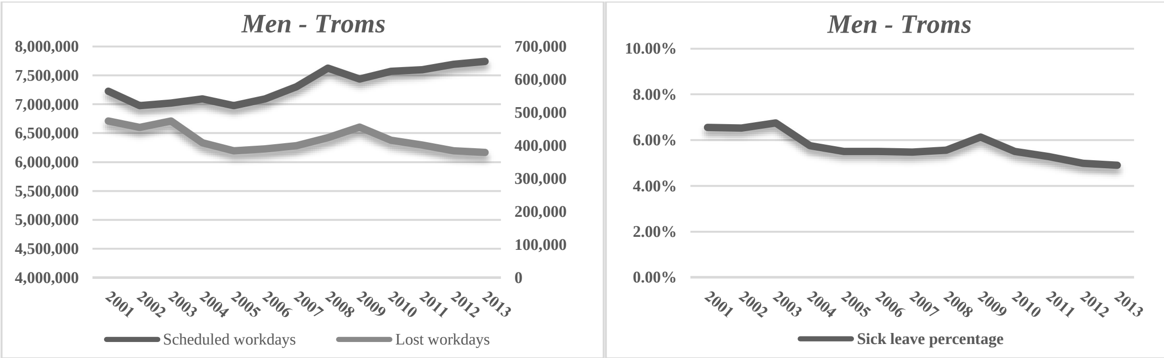
2.3 Troms and Finnmark

The main aim of this thesis is to study why the northern municipalities have an overall higher level of sickness insurance usage than the rest of the country. Dutrieux & Sjöholm (2003) conducts a regional study of sick leave in Sweden and reveals significantly higher levels of usage in the country's northern regions than in the southern and middle regions, even after controlling for explanatory factors such as demographic differences and labor market structure. Some sick leave statistics for the two northernmost counties are therefore presented in this subsection.

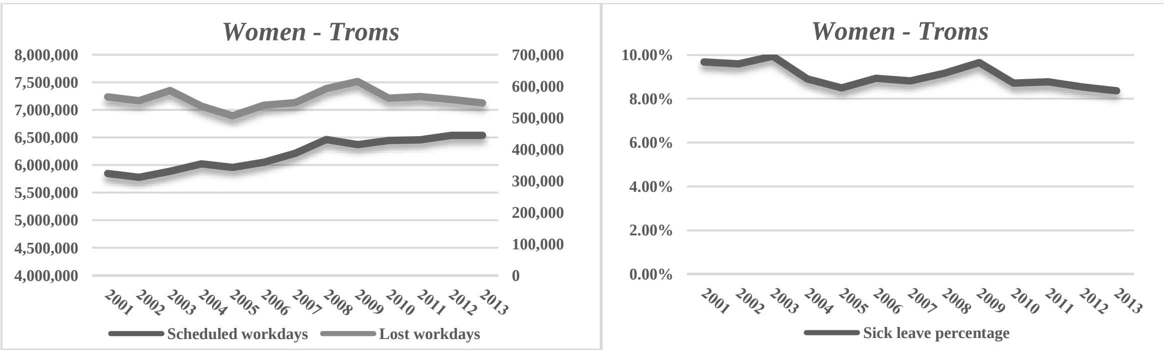
A decomposition of the sick leave percentage

The sick leave percentage is defined as a fraction with lost number of work days in a given period in the counter, and the total scheduled number of work days in the same period in the denominator. Graphs 2.2 - 2.5 present the number of lost work days, scheduled work days and the sick leave percentage separate for Troms and Finnmark, and for men and women. The graphs to the left have number of lost work days in the right hand y-axis and number of scheduled work days in the left hand y-axis. The graphs to the left show the sick leave percentage. All graphs are based on data retrieved from NAV. The years of 2001, 2012 and 2013 have also been included due to available data, and we thus obtain a broader picture.

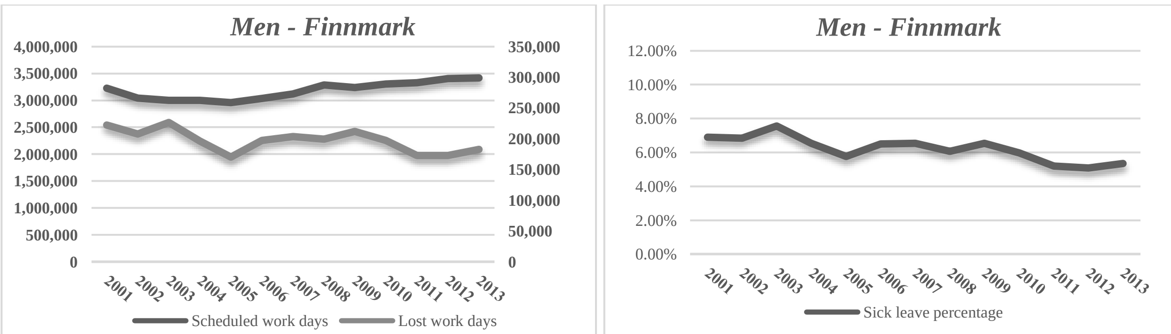
Graph 2.3.1: Lost work days, scheduled work days and the sick leave percentage. For Troms 2001-2013. Male workers.



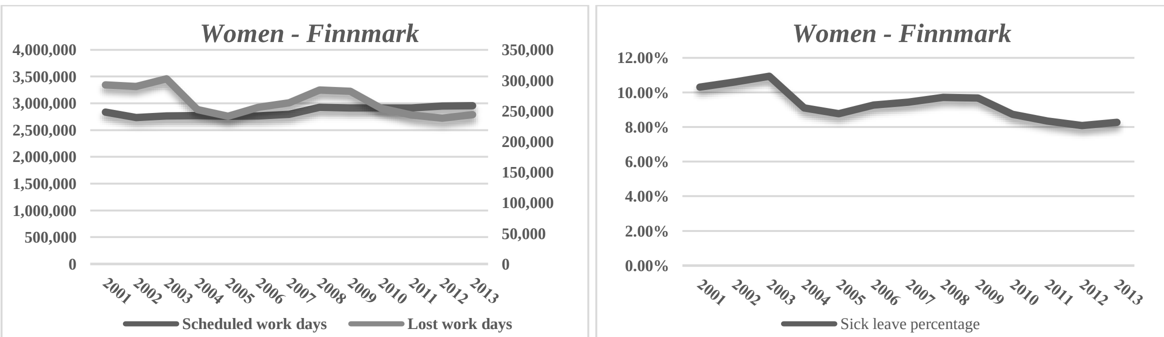
Graph 2.3.2: Lost work days, scheduled work days and the sick leave percentage. For Troms 2001-2013. Female workers.



Graph 2.3.3: Lost work days, scheduled work days and the sick leave percentage. For Finnmark 2001-2013. Male workers.



Graph 2.3.4: Lost work days, scheduled work days and the sick leave percentage. For Finnmark 2001-2013. Female workers.



The sick leave percentage shows a decreasing trend in both counties and for both genders over the period. In Troms, this can be attributed to a significant increase in the scheduled number of work days between 2001 and 2013, while the number of lost work days has been decreasing for men and remaining relatively stable for women. In Finnmark the increase in scheduled work days has been more modest however, with the change for women being miniscule over the period. The northernmost county has nonetheless experienced a decrease in the sick leave percentage due to a decreasing trend in number of lost work days for both genders.

The graphs further reveal that the level difference in the sick leave percentage between the genders can be explained both by women having a lower number of scheduled work days than men *and* by them having a higher number of lost work days.

Average length of sickness spells

Helde, Kristoffersen, Lysø & Thune (2010) analyzed a data set for 2008 on the rate of occurrence of different diagnoses under physician-certified sick leave, and how these differ between the Norwegian counties. An inspection of the included data tables in their paper reveals that the average length of sick leave in six major diagnostic groups is lower in Troms and Finnmark than in the country as a whole.

In their paper they present the broad categories of “musculoskeletal disorders” and “mental disorders” as the two most recurring types of diagnoses. They further subdivide the former into “shoulder syndrome”, “back syndrome with pain radiation” and “back syndrome without pain radiation”. The latter is subdivided into “depressive disorder”, “mental imbalance situational” and “laxity/weariness”. With data retrieved from NAV, they list average length, median length and number of cases for each of these diagnoses separated by county. A quick study of their tables makes for some interesting findings. For “depressive disorders”, Finnmark has the countries lowest average length at 76 days. The nationwide average for this diagnosis is 88 days. The same goes for “mental imbalance situational”. Here, Finnmark has the lowest average length at 33 days, while the country average is 45 days. Thirdly, for “laxity/weariness” Finnmark and Oslo share the lowest average length, at 45 days, with a country average of 55 days. Thus, the data reveal that Finnmark has the lowest average length of sick leave for all three subdivisions of “mental disorders”.

In addition, Troms has the lowest average on "back syndrome with pain radiation" with 62 days, while Finnmark comes a close second with 63 days.

Numbers for all diagnoses have been reproduced in table 1, together with a ranking comparing Troms and Finnmark to the other counties in Norway.

Table 2.3.1: Average length of sick leave by diagnosis and ranking of average length compared to other counties. Cases starting in 2008. For Troms, Finnmark and Norway.

	Norway	Troms		Finnmark	
Diagnosis	**Average length**	**Average length**	**National Ranking**	**Average length**	**National Ranking**
Shoulder syndrome	67 days	58 days	Second lowest	63 days	Fourth lowest
Back syndrome with pain radiation	72 days	62 days	Lowest	63 days	Second lowest
Back syndrome without pain radiation	34 days	33 days	Seventh lowest	31 days	Third lowest
Depressive disorder	88 days	86 days	Seventh lowest	76 days	Lowest
Mental imbalance situational	45 days	40 days	Third lowest	33 days	Lowest
Laxity/weariness	55 days	51 days	Seventh lowest	45 days	Lowest

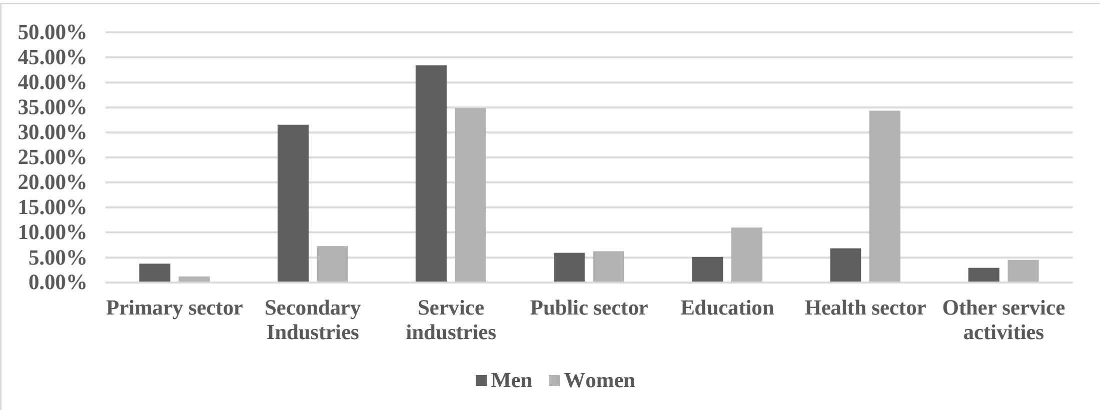
****All data retrieved from Helde, Kristoffersen, Lysø & Thune (2010). Based on statistics from the Norwegian Social Security Admin. (NAV)***

As seen, the statement that average length of sick leave in Troms and Finnmark is relatively low, is well justified.

Labor market structure

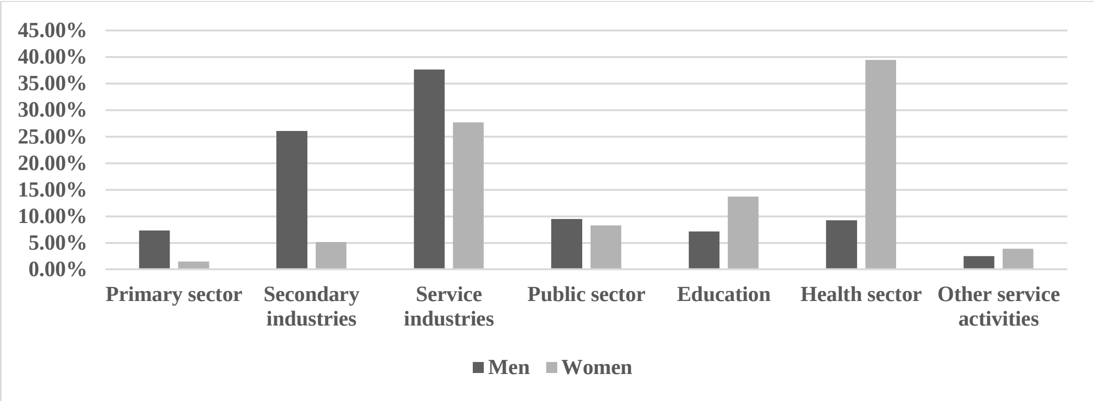
Both the physical and mental strains connected with the average working day will vary greatly between workplaces, industries and sectors. It is therefore natural to assume that there might be some variation in sick leave usage as well. Graph 2.2 shows how large a percentage of the total Norwegian labor force is employed across a broad categorization of workplaces.

Graph 2.3.5: Employment across different sectors and industries per 2011. Whole of Norway. Separate for men and women.



As the graph shows, the largest employers among men are the secondary industries and the service industries, while for women they are the service industries and the health sector. Graph 2.3 shows the same distribution for the northernmost counties of Troms & Finnmark.

Graph 2.3.6: Employment across different sectors and industries per 2011. Troms & Finnmark. Separate for men and women.



Also here the secondary and service industries are the largest for men while the service industries and the health sector are the largest for women. However, what are more interesting are the differences between the north and the rest of Norway. Both the primary sector and the public sector are significantly larger in the north, and the same goes for employment in the health sector among women. By contrast, relatively less people appear to be employed in the service and secondary industries in Troms and Finnmark than elsewhere in Norway.

The findings here create the foundation for how labor market structure is used in the further analysis. The primary sector, public sector, health sector and the service industries are chosen as further focus points that might contribute to understanding what sets the northernmost municipalities apart from the rest of Norway when it comes to sickness insurance usage.

3 The individual and absence behavior

The basic decision between utilizing sick leave or not is made on an individual level, and the construct of this decision making process can be rather complex. This chapter will give a quick run-through of some of the most commonly brought up theories regarding individual sick leave behavior. Hopefully, this chapter will also clarify how some of the underlying factors affecting the decision making process work, and lay a theoretical foundation for the model on sick leave usage to be constructed in the next chapter.

The neoclassical theory of individual labor supply[6]

The economic theory on absenteeism starts with the basic labor market model, looking at the mechanisms of labor supply. Here, work-participation, or the lack thereof, is a supply decision made exclusively by the worker. He or she will supply labor services if granted some form of payment. A *reservation wage* is defined, which is the minimum sum the worker must be paid to prefer working over not working. He will have a utility function that is increasing in both consumption and leisure, so the indifference curves of the worker will be negatively sloped and convex. Given these conditions, the worker will attempt to maximize utility by finding the optimal tradeoff between labor and leisure. For higher wages, the worker will be willing to trade more leisure for labor. In sum, this means that aggregate labor supply will be an increasing function of the offered wage.

On the other side we have labor demand. A profit-maximizing employer will hire workers up to a point where the marginal value of the work provided equals the marginal cost of employing the worker, which equals the wage. The higher the wage, the more expensive the labor becomes, and labor demand decreases.

The equilibrium wage level **w*** will then lie at the intersection of the two curves. Workers with a reservation wage below **w*** will realize a positive surplus from working, while those with a lower reservation wage will choose not to work.

The existence of social benefits will affect the labor market outcomes through two channels. First, they will increase the reservation wage of the workers, which induces more of them not to supply labor at a given market wage. Secondly, maintaining them requires increased taxation (payroll taxes), which will affect labor market outcomes through lower net wages.

In conclusion, when there are benefits connected with not working, the worker requires higher compensation for his or her time, in the form of a higher reservation wage. The higher the benefits, the higher the reservation wage. So, all forms of payment connected with not working will reduce the number of people that choose to work, given a constant market wage. In Norway, the sick leave insurance equals the market wage for the first year, which then according to basic labor market theory should raise the reservation wage substantially and increase the usage of sick leave.

In the context of economic incentives and their effect on the labor supply decision, it is also natural to discuss the concept of moral hazard.

Moral hazard

When the replacement rate is higher, it reduces the individuals' costs of being absent from work, and his or hers incentive to minimize sick leave therefore diminishes. This influence of economic incentives on absenteeism is often referred to as moral hazard (see for example Kahn & Rehnberg, 2009). A general definition of moral hazard is that the individual will have a tendency to take higher risks when the potential costs are borne partly or entirely by someone else. If the individual must bear the costs themselves, he or she will tend to be more risk-averse. Transferred to the context of social benefits and sick leave, this means that an individual might change his or her absence behavior under different degrees of exposure to the costs of not working. For example, under the protection of a well-functioning sick-pay scheme the individuals' economic loss from not working will be smaller, or even nonexistent as in the case of Norway. For a given health status, he or she will then be more inclined towards taking sick-leave than a person with an identical health status, but under a less generous sick-pay scheme.

This theoretical framework can at best just approximate reality, and will not hold true for extreme cases. The Norwegian insurance system provides full coverage for up to one year, but obviously not all Norwegian workers are utilizing this. Similarly, countries without functioning sick-pay schemes still experience some level of sick leave in its population. The affected group is therefore presumed to be those standing at the margin between working and not working.

Kahn & Rehnberg (2009) studied the individual behavior of workers in regards with utilization of sick leave under different levels of perceived job security. They found that especially short-term sick leave tends to increase with a higher perceived job security.

The social gradient and status theories

In his famous book *Status Syndrome,* Marmot (2004) argues the existence of a social gradient when it comes to health, where a higher social position facilitates better overall health in the individual. He uses examples from studies on work position (Marmot, Shipley

& Rose, 1984, Marmot & Shipley, 1996, and van Rossum, Shipley, Van de Mheen, Grobbee & Marmot, 2000), promotions (Marmot et al., 1991), fame (Redelmeier & Singh, 2001) and education (Erikson, 2001) in supporting his claim. It is pointed out that it is the relative position in the hierarchy that makes the difference, and that the surrounding social construction determines the effects of a given social rank. In an article published the same year where he summarizes key points of his book, he explains the construct of status and relative position in the following way:

"…your status is related to two fundamental human needs: to have control over you own life and to be a full social participant with all that implies about being a recognized member of society…the key to the status syndrome lies in the brain. It is stress arising from the inability to control our lives, to turn to others when we lose control or to participate fully in all that society has to offer…being part of a socially fractured community adds the insult of low social participation to the injury of low control over life circumstances."

Marmots reasoning is that being able to actively take part in society and feeling in control of one`s own life are basic elements influencing the general health and well-being of the individual. Transferred to the present context, meaning what might influence his or her sick leave behavior.

The theories of Marmot have also found scientific support in a Norwegian study by Markussen, Røed, Røgeberg & Gaure (2009) where they observe that "*…the gradient prevails whether we measure status by family background, own educational attainment, occupation, wealth or pay."*.

Central to how the social gradient affects health is the presence of stress, and a multitude of stress-theories across various disciplines have been developed over the years.

Stress-theories

When looking indirectly at sick leave behavior through health determinants, one must also mention stress-theories. Two often mentioned theories in this context are the person-environment fit theory (Edwards & Rothbard, 1999) and the effort-reward imbalance theory (Marmot et al., 1991), both rather self-explanatory. The former describes a mismatch between the worker and the work-environment which then creates stress, and the latter an imbalance between the work provided and what is perceived as a fair reward or payment for said work.

Bankruptcies, restructuring and downsizing of companies often create a more stressful working environment. It increases the pressure on the individual worker and gives him/her a diminished sense of control over his/her working situation. This leads to increased mental and physical strain and general dissatisfaction, which thus increases sick leave. This connection between a demanding work environment, lost sense of control and sick leave usage, is known in the literature as the Demand-Control hypothesis (See for example Røed & Fevang 2005).

Exposure to these forms of stress over time might facilitate health problems such as cardiovascular disease, anxiety and depression.

Social norms

Social norm-hypotheses in general state that humans are influenced by the actions, attitudes and behavior of people they regularly interact with through various social circles, and that they do not want to deviate from the norms and behavior of these groups (Kostøl, 2010). The group in question can be the workplace environment, the neighborhood, close family, friends or a variety of other social circles, depending on who is most influential in the individuals' life. Identifying the effects of such hypotheses can be challenging because the individual might also affect the group through own behavior, and there might be various external factors influencing both the group and the individual (Manski, 1993). It is also difficult to define whether the effects follow from the social interactions themselves, or from the flow of information they cause.

Lindbeck, Palme & Persson (2008) ask the question of whether group influence exists in sickness absence behavior, and if so, how large it might be. They use four strategies for estimating how social interaction within neighborhoods affects sick leave behavior. First, they exploit the difference in sick leave usage between public and private sector employees, and look at neighborhoods represented more or less by one of these groups to see if there are any intergroup influences. Next, they look at individuals moving from one neighborhood to another to see if there is any conformity in sick leave behavior. Third they look at immigrants to Sweden to see if there is a tendency to adjust to the behavior where they settle down. Lastly, they investigate the effect of network interaction in the neighborhood and on the workplace. They find that all four strategies "…*unambiguously indicate that such interaction effects do in fact exist.".*

Economic cycles

Empirical research on labor market conditions and absenteeism has shown that sick leave is pro-cyclical, meaning that it increases when the economy is doing well and unemployment is low, and likewise decreases in economic downturns when unemployment is rising (see for example Kahn, Gerdtham & Jansson (2004)). The three most frequently referenced hypotheses supporting this are the disciplinary hypothesis, the labor force composition hypothesis and the demand hypothesis.

> ***The disciplinary hypothesis*** was launched by Shapiro and Stiglitz (1984) and takes the assumption that the existence of unemployment in the labor market has a disciplinary effect on workers, thereby reducing their inclination towards shirking when the employer is not watching. This will be a dynamic relationship. When the unemployment rate is low, people will shirk more. When it is high, people will shirk less. Under the assumption that sick leave is used as a shirking mechanism and that a significant part of aggregate sick leave can be explained by shirking, this points towards a negative relationship between the sick leave rate and the unemployment rate.
>
> ***The labor force composition hypothesis*** states that people with poorer health, defined as "marginal workers", will to a larger extent find employment in economic upturns when labor demand increases (See for example Ose, Jensberg, Eidsmo, Suandsund & Dyrstad (2006)). They lower the collective health of the labor force and increase aggregate sick leave usage. Conversely, when times are worse and unemployment increases, these workers will be the first to lose their job.
>
> ***The demand hypothesis*** relates to the psychosocial and physical working conditions during different stages of the economic cycle. In an economic upturn the pace and physical strains at the work place might increase, leading to increased risk of injuries, getting "burned out" and other health-related problems. In a recession the pressure on the individual employee lessens, thereby reducing the need for sick leave (Askildsen, Bratberg & Nilsen (2005)).

Askildsen, Bratberg & Nilsen (2005) tested the relationship between unemployment and sick leave across a random sample of the Norwegian work force, and found it to be negative. When they restricted the sample to only those workers who were present over the whole period, thus excluding changes in labor force composition as an explanatory variable, the negative relationship became even more significant. The results show that stable workers do in fact change their absence behavior depending on economic cycles. The results further indicate that procyclical fluctuations in the use of sick leave are mainly caused by stable workers and the disciplinary hypothesis, and not by the labor force composition.

4 A model for sick leave usage

This chapter will construct a basic model for analyzing differences in sick leave usage across municipalities. Each variable chosen as input is explained in connection with earlier findings, related theories and, where it is appropriate, some practical examples.

4.1 Dependent variables

Two different dependent variables will be used in the model in an attempt to capture different aspects of sick leave behavior. They are the *sick leave percentage* and the *percentage of sick workers*. Mainly three types of variables have been used earlier in the literature where regression analysis has been performed. These former measurements have been expressed as:

(1) *Total Sickness Insurance Usage (TSIU)/Sick-days per worker/Sick-pay costs per worker* (Dutriex & Sjöholm, 2003, Olsson, 2004, Bragstad, Regbo & Sagsveen, 2006, Olsson, 2006, Osterkamp & Röhn, 2007, Krogsgård, 2009)
(2) *Sick-listings per worker* (Dutriex & Sjöholm, 2003, Bragstad, Regbo & Sagsveen, 2006, Krogsgård, 2009)
(3) *Average absence length* (Bragstad, Regbo & Sagsveen, 2006)

Where (1) measures the total amount of sick leave usage, (2) measures the frequency of sick leave cases and (3) measures the average length of a sickness spell.

The sick leave percentage

The sick leave percentage is the measurement for sick leave usage that has been referred to so far in the present paper. It measures the share of planned work days that gets lost due to reported sickness absence. Out of the other used measures in the literature it lies closes to TSIU and the first group. However, in the sick leave percentage the number of lost work days is measured against the number of planned work days, and not against number of workers.

In recent years, the sick leave percentage has become the most commonly used measurement when discussing sick leave in Norway. It is frequently seen reported by the media and in publications and articles by NAV (See for example Helde, Kristoffersen & Lysø 2011, Brage, Nossen & Thune 2013 and Kann, Thune & Galaasen 2013).

Percentage of sick workers

This sick leave measurement has not been covered so far. It shows the number of workers with one or more cases of registered sick leave in a given period, as a share of total number of workers. It does not measure sick leave directly but rather the amount of workers utilizing it, and hence can not be put into any of the three mentioned categories. It has not been found used in any previously performed regression analyses, but is often found reported in the media.

The sick leave percentage can be seen as describing the depth of sickness insurance usage in the population, while the share of sick workers is more of a width measurement. The former is regulated through measures such as graded sick leave, where the possibility of combining sickness with work is made more convenient. The total amount of sick leave can thus be minimized. The latter however requires measures where the decision between using sick leave or not is more directly influenced. The target group there would thus be those who have more of a choice between using it or not, so called marginal insurance users. Logically, the type of sick leave to target would then be mainly short-term spells because longer sickness spells often involve more serious illnesses where working is not an option in any case.

Using both of these variables in the model gives a more detailed view of sick leave usage, and thus allows for a more in-depth discussion of the results.

4.2 Explanatory variables

With a base in earlier research, this section will list a selection of factors that might influence how sickness insurance usage changes over time, and how it differs between municipalities. The chosen explanatory variables for the model have been broadly divided into the following three groups:

- Demographic factors
- Social factors
- Labor market factors

The creation of these particular groups is based on how similar analyses have been structured in the past (See for example Dutriex & Sjöholm, 2003 and Bragstad, Regbo & Sagsveen, 2006). Each variable is discussed in the framework of its categorization.

4.2.1 Demographic variables

When the population composition changes over time, sick leave usage is affected. The overall effect of demographic changes may however not always be obvious, as there will be multiple relationships pulling in different directions at the same time. Hauge & Kann (2007) looked at the years 2002-2005 and show that the sum of demographic changes over time had a relatively miniscule effect on how sick leave usage changed.

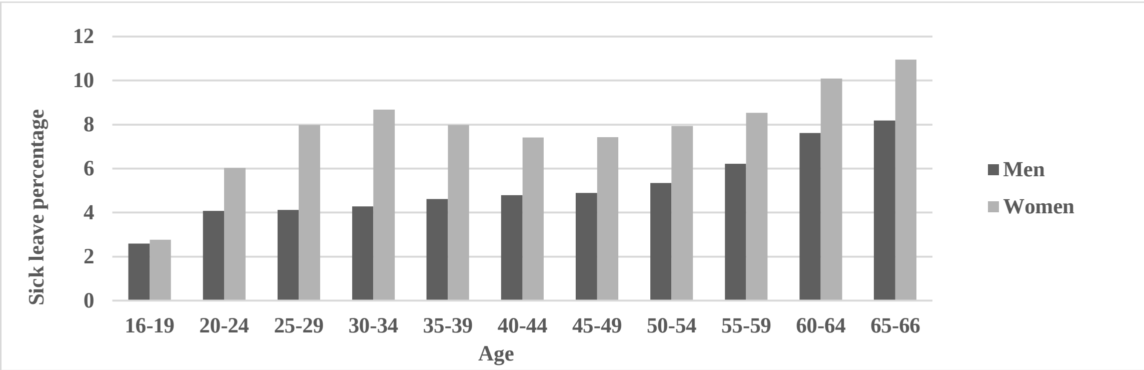
The demographic variables used in the model are *Age*, *Immigrants* and *Municipality size*, each discussed in turn.

Age

Figure 4.1 shows the sick leave percentage for different age groups over the period 2002-2011, divided by gender. We see a general rising trend with age for both sexes. However, Hauge & Kann (2007) state that the labor force participation among older people is expected to increase over time due to improved health and a higher average level of education. Thus a decreasing use of sick leave might also be expected. Further, Berge (2010) shows that the oldest age group has experienced a significant decrease in total sick leave throughout the 2000`s, which is partially explained by a gradual reduction in number of long term absences.

The graph also shows a relatively high rate of sick leave for women between 21-39 years of age, which as mentioned in the discussion on gender differences might be partially attributed to pregnancy-related sickness absence.

Graph 4.2.1: Self-certified and physician-certified sick leave in Norway averaged over the period 2002-2011. Divided by gender and age groups



Notes: All data obtained from Statistics Norway

Markussen, Røed, Røgeberg & Gaure (2009) finds that the probability of entering into sick leave *declines* sharply with age up to around 45 years, where it then stabilizes for minor diseases and rises for major diseases. They mention a few possible explanations. Older employees might have had more time to find a satisfactory job match. Also, there might be some natural selection in the labor force composition with age, meaning that those with poorer health and higher absence propensity eventually fall out, leaving the remaining workers with a higher average health level. Lastly, young workers might be bearers of a less strict norm set and lower thresholds for utilizing sick leave.

In sum, although the level differences between the age groups are rather clear over the period, the relationship between sick leave usage and age is rather ambiguous.

Immigrants

Per first of January 2014, first- and second-generation immigrants in Norway consisted of 759 185 individuals from 221 different countries and autonomous regions, constituting 14.9% of the Norwegian population[7].

Dutriex & Sjöholm (2003) perform a cross-sectional analysis of Swedish municipalities for the year 2000, and find a higher share of immigrants to have a significant negative effect on their dependent variable *Total Sickness Insurance Usage*. They do not explore these results further, only note that ”*results from individual-level studies suggest that the effect* *should be the contrary...*". When they run a further regression on only the three big city regions in their sample, they find TSIU to be positively correlated with the share of foreign born.

Dahl, Hansen & Olsen (2010) present four groups of factors that briefly summarize the existing literature on health differences between natives and immigrants. They are: (1) Psycho-social factors, mainly stress-related due to major life changes; (2) Socio-cultural factors, such as lifestyle-differences between countries and selective bias among individuals that migrate; (3) Social exclusion, e.g. a weaker labor market position; (4) Norms, in the form of differing attitudes and understandings of what constitutes acceptable absence. Some of these factors indicate higher sick leave among immigrants while other lower, so no definite sign can be put on the aggregate effect of health differences.

The same article studies a panel data set for 1992-2003 on long-term sick leave among immigrants and ethnic Norwegians, and finds there to be a higher level of sickness insurance usage among immigrants from Asia (both men and women) and Africa (only men) than among ethnic Norwegians. However, among immigrants from North-America and Oceania they found the level to be lower. They further found immigrants from other nordic countries, Western- and Eastern-Europe, and all second-generation immigrants in Norway to have a level of sickness insurance usage indifferent from that of ethnic Norwegians.

In sum, the expected sign of this variable is rather ambiguous and no definite relationship between share of immigrants and use of sick leave is assumed.

Municipality size

Krogsgård (2009) creates a variable defined as the *natural logarithm of regional population* and uses it as input in a cross-sectional regression on averaged out sick leave data from 2003-2007. He finds the variable to have a significant negative relationship with sick leave cases per worker, but gets no significant results when running it against total sickness insurance usage.

A higher prevalence of sick leave in more rural regions has been found repeatedly in Swedish studies (SOU 2000:121 & SOU 2002:5). This is attributed to weaker labor markets and a higher occurrence of people with work impairments. Healthier, more active people can more easily relocate to areas with better labor market prospects.

The following practical example depicts how municipality size and sick leave usage are linked. All presented numbers in this section were retrieved from the online databases of Statistics Norway.

In 2002 the lowest observed sick leave percentage for men was at 1.33% in Kvitsøy, while the highest was at 11.4% in Lavangen. For women the lowest was at 4.05% in Bjerkreim, while the highest was at 14.63% in Berg.

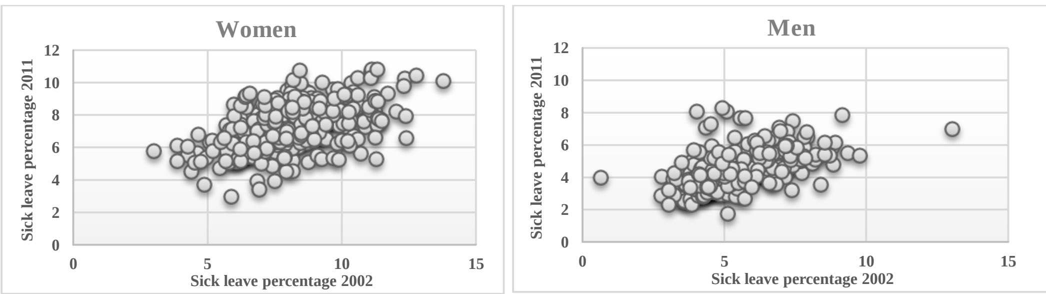
In 2011 the lowest observed sick leave percentage for men was at 2.33% in Tydal, while the highest was at 9.63% in Lavangen. For women the lowest observed was at 3.73% in Hornindal, while the highest was at 12.8% in Ballangen.

Per first of January 2002 the population sizes in these municipalities were as follows: Kvitsøy and Lavangen had 529 and 1131 inhabitants respectively. Bjerkreim had 2483 inhabitants, while Berg had 1094 inhabitants. Tydal had 935 inhabitants, while Ballangen had 2731 inhabitants. The average population size over all municipalities per 2002 was 10 485.

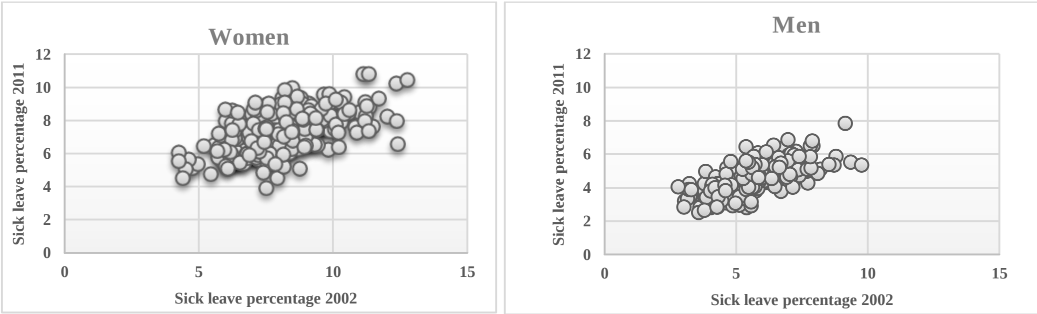
So, the extreme values for 2002 and 2011 were all from relatively small municipalities. It turns out that there exists a clear relationship between observation spread and municipality size when looking at sick leave. The below scatter plots illustrate this effect. The first two plots show the sick leave percentage for women and men separately over all municipalities. The next four show the same plots, but the observations have been limited to include only those with more than 2000 and more than 5000 inhabitants. This excludes the extreme value municipalities mentioned above and other municipalities of similar sizes. The plots clearly show a gradually more concentrated bundle of observations as smaller municipalities are excluded.

Scatter plot 4.2.1: Sick leave percentage for 2002 and 2011. Separated by gender and municipality size.

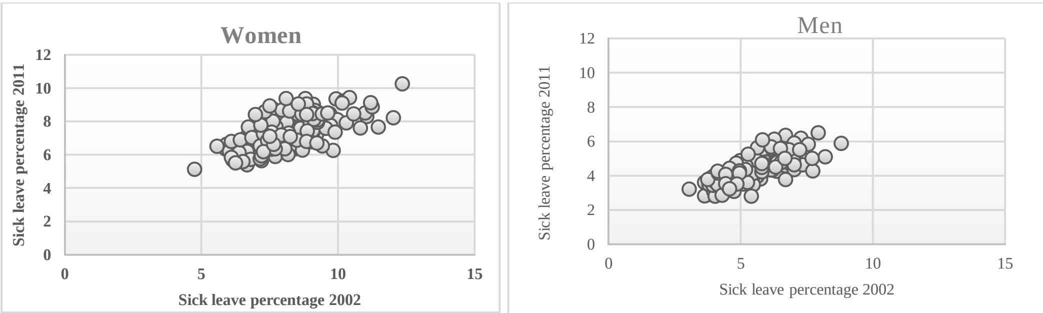
All municipalities (N=428)



Municipalities with more than 2000 inhabitants (N=337)



Municipalities with more than 5000 inhabitants (N=189)



While this example says nothing about how the municipality size affects the propensity for sick leave usage, it does indicate that the two are strongly connected.

4.2.2 Social variables

Three variables have been chosen under this category. They are *level of education*, *participation in elections* and *disability benefits usage*.

Level of education

A higher average level of education is presumed to be negatively correlated with sick leave rates, based on findings by Dutriex & Sjöholm (2003), Olsson (2004), Olsson (2007) and Markussen, Røed, Røgeberg & Gaure (2009) to name a few.

As covered in chapter 3, studies suggest a strong social gradient in health and the use of sick leave (e.g. Marmot 2004). A higher education is one of the factors connected with a higher social status, and thus an improved overall health.

Markussen, Røed, Røgeberg & Gaure (2009) further find that the type of education is of lesser importance, it is the level that matters. As a further indirect effect they formulate a connection between average level of education and social norms, stating that colleagues with higher education and thus lower sick leave will raise the threshold among all workers for claiming sickness benefits.

A STAMI-report from 2008 by Foss & Skyberg finds the highest prevalence of long-term sickness absence among unskilled and low education workers. They explain that even though they do not have enough information on the connections between work environment, health problems and exclusion from working life, it may be assumed that exclusion most often affects unskilled workers, and that work environments with negative health effects primarily affect those with less choices in the labor market. It is considered important to utilize preventive work environment measures in professions and workplaces with a relatively high share of unskilled labor, tight time schedules and low levels of codetermination in deciding work processes.

Participation in elections

Participation in elections is included as an indicator for social participation in the population. Marmot (2004b) stresses the importance of social participation and a feeling of being in charge of one's own life when it comes to personal well-being and health. The level of electoral participation is chosen as a proxy for the general level of this type of social involvement in a municipality. The variable was also included by Krogsgård (2009), but no significant results were obtained in the regressions.

Disability benefits usage

Payments of disability benefits in a municipality are in themselves a health measurement on equal grounds as sickness insurance usage. Bragstad, Regbo & Sagsveen (2006) state that the share of people on disability benefits in a given period points toward how the level of sick leave has been in the municipality in previous periods. They further state that high disability benefits payments combined with high levels of sick leave might imply that the overall health in the municipality is poor. High sick leave usage often leads to disability benefits usage, so to reverse this relationship and rather capture the effects of disability benefits usage on sick leave, they lag the variable by one year. As estimation results they find a positive connection between disability benefits usage and sick days per worker for women. They obtain no significant results for men. When they exclude the time factor and only look at inter-municipality differences, they get a highly significant positive relationship for both genders, and across all three dependent variables (sick days per worker, sick leave cases per worker and average absence length)

In accordance with the method chosen by Bragstad, Regbo & Sagsveen (2006), the variable on disability benefits usage is used in lagged form in the model. The lag time is set to one year. The variable is therefore defined as *disability benefits usage preceding year* in the remaining part of the paper.

4.2.3 Labor market variables

Labor market factors included are *unemployment*, *bankruptcies*, *labor market structure* and *small companies*. Gender separated data were obtained on *unemployment* and *labor market structure*.

Unemployment

This factor has already been partially covered in chapter 3, in the context of sick leave and economic cycles. As shown there, Askildsen, Bratberg & Nilsen (2005) found a significant negative relationship between unemployment and sick leave. Further, Bragstad, Regbo & Sagsveen (2006) analyze a panel data set for 1993-2004 on Norway, and also find the relationship to be negative for both genders when looking at changes over time. When looking at differences between regions they find a negative relationship for women but get no significant results for men.

Dutriex & Sjöholm (2003) found that municipalities with high levels of sick leave often had a small labor market and a consistently high unemployment rate. They raise the question of whether cyclical and structural unemployment might have partially opposite effects on sick leave behavior. A permanently higher supply than demand for labor in the municipality points toward a higher propensity towards sickness insurance usage.

Bankruptcies

Workers, who know that they are at high risk of losing their job, or are already in their period of notice, might utilize sickness insurance as a strategic measure (Nossen 2010b). They will then receive sickness benefits equal to 100% of regular income for up to one year, as an alternative to spending a period on work assessment allowance or unemployment benefits. This behavior might also be profitable for the firm itself if they are in a situation where they have excess labor but find it difficult to lay off people due to employment protection regulations.

Bratsberg, Fevang & Røed (2010) look specifically at companies going out of business, and find that bankruptcies increase the probability of ending up on disability benefits in the next six years with 123% for men and 50% for women, compared to employees in companies that experienced a less than 10% decrease, or an increase, over the same period.

Rege, Telle & Votruba (2005) look at the connection between company downsizing and disability entry rates and find that the likelihood of entering into disability is significantly larger among workers in a company that is downsizing. Interestingly, they find the relationship to be nonlinear, with higher entry rates when the company downsizes between 65-95%, than if the company completely goes out of existence.

Labor market structure

The effect of industry structure on sick leave has been explored in a multitude of studies. Bragstad, Regbo & Sagsveen (2006) examine official sickness absence statistics and find that sick leave varies both across sectors and industries. In their analysis they find that an increasing share of municipal employees has a significant negative effect on average number of sick days per worker and on average absence length. Dutriex & Sjöholm (2003) however measure the same variable and find it to have a *positive* effect on TSIU. The Swedish report also measures the effect of state employees and here finds a negative effect.

In another report by Dutriex & Sjöholm from 2003 (RFV Analyserar 2003:12) they inform that one important aspect of municipalities with higher sick leave is that the labor market structure differs from that observed elsewhere. Rural municipalities in northern regions, which they discover often have above average sickness rates, also tend to have a lower share of workers working in industry and a higher share in education, healthcare and social services than the country average.

Olsson (2004) uses various labor-market structure variables as input in a panel data regression with TSIU as a dependent variable, and gets several significant results: share of municipality employed has a positive effect; share of privately employed has a negative effect; share working in trade has a negative effect; share working in hotels and restaurants has a positive effect, and share that works in transport has a positive effect.

Markussen, Røed, Røgeberg & Gaure (2009) find entry rates into sick leave to vary with as much as 30% across major industries. Their analysis finds highest sickness rates in the manufacturing, teaching, and health care sectors, while the lowest are found in the oil industry, retailing and R&D.

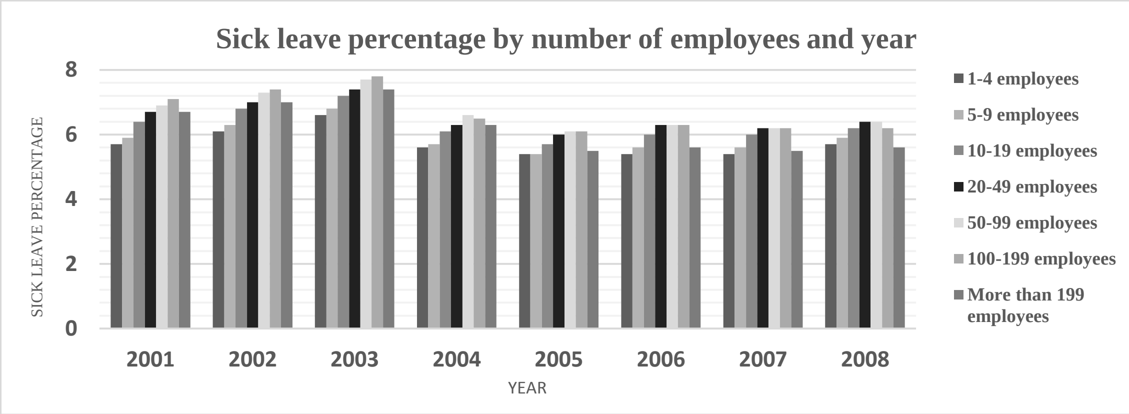
Helde,Kristoffersen & Lysø (2011) look at Norwegian data from 2009 on number of sick leave cases in different sectors. They find health- and social services and administrative and support service activities to have the highest share of cases per worker, while real estate and technical activities has the lowest. As a general explanation for occupational differences in sick leave they mention that the composition of diagnoses will differ between professions, due to different degrees and types of physical and mental strain. Further, certain diagnoses will be more/less inhibiting in certain jobs.

Small companies

A study by Barmby & Stephan (2000) runs a panel data analysis on German individual and firm data and finds absence rates to increase with firm size. As possible psychosocial explanations they theorize that bigger companies have a lower level of social control, and thus more shirking-related sick leave. The employees will feel less influential in the work place, possibly have a more distant relationship with superiors and a lower sense of responsibility towards colleagues. These factors might encourage marginal sick leave. The study further looks at the labor demand side of the equation and shows that larger firms might be more able to diversify risk from absence than smaller firms. Therefore they are capable of withstanding a higher equilibrium level of sick leave.

Figure 4.2 shows the aggregate sick leave percentage for Norway, divided by company size over the years from 2001 to 2008. A general increase in sick leave with company size is seen up to a certain point, were it then flattens out and starts to decrease. A possible explanation for this is that in the biggest companies the prevention and follow-up systems for sickness absence might be more developed. Also, the work force might be more subdivided into smaller work environments, which then takes on the characteristics of smaller companies.

Figure 4.2.2: Sick leave percentage by number of employees and year. Whole of Norway, 2001-2008.



In sum one might assume that the smallest companies should overall experience lower sick leave usage than larger companies, and this will thus be examined in the further analysis by including the prevalence of companies with 1-9 employees as an explanatory variable in the model.

5 Data and methodology

This chapter will define on a more technical level how the chosen variables are constructed and how they will be used in the regression analyses, together with some basic descriptive statistics. Also covered is the methodology surrounding the use of panel data, why this regression form is preferable and how the basis for the performed regression analysis is constructed.

Based on the information presented here, the general model for sick leave usage presented in chapter 5 will be divided into two slightly different models to be used further. One of them is used when performing regressions on the 43 northern municipalities of Troms and Finnmark, while the other is used in regressions on the remaining 385 municipalities. For definitional simplicity the latter model will in the rest of the paper be referred to as the country-wide model, even though this is not one hundred percent correct seeing that it excludes the municipalities of Troms and Finnmark. The former will simply be referred to as the Troms & Finnmark model or the northern model.

The data used is based on the municipal boundaries as of first of January 2013. This gives as mentioned a total of 43 municipalities, or individuals, in the northern model and 385 municipalities in the country-wide model.

5.1 Variable definitions and descriptive statistics

A general explanation of the structure of each variable is shown in table 5.1. Some of these, like *age* and *labor market structure*, are further subdivided in the analysis. Some of the variables are also separated by gender. Details on this are covered in subchapters 5.1.1 - 5.1.4.

Table 5.1.1: Description of dependent and independent variables used in the regression models.

Variable name	**Description**
The sick leave percentage	Number of work days lost due to reported sick leave as a share of total number of scheduled work days
The share of sick workers	Number of workers with one or more cases of registered sick leave as a share of total number of workers
Demographic factors	
Age	Number of people in a given age group as a share of total population between 16-66 years of age
Immigrants	Number of people with two foreign-born parents and four foreign-born grandparents as a share of total population. Asylum seekers and people on short-term stays are not included
Municipality size	The population size per first of January 2002
Social factors	
Level of education	Share of population over the age of 25 that has completed a high school education or higher
Participation in elections	Share of population that exercised their right to vote at parliamentary elections.
Disability benefits usage in preceding year	Share of population between 18 and 67 years of age that received disability insurance payments in the preceding year
Labor market factors	
Unemployment	Share of population between 15 and 74 years of age that is registered as unemployed
Company bankruptcies	Number of bankruptcies in a given year as a share of total number of registered companies at the beginning of the year
Labor market structure	Number of workers in a given sector or industry as a share of total number of workers
Small companies	Number of companies with 1-9 employees as a share of total number of companies

The following sections will present some more in-depth details and descriptive statistics on each variable. The estimated statistics on variance include overall, between and within variance, which describe total variation, cross-sectional variation and variation over time respectively. The cross-sectional statistics are created by averaging out the data over time for each separate municipality, while the within-variation is found by averaging out the data over all municipalities in each separate year.

5.1.1 Dependent variables

The sick leave percentage

As described in table 5.1.1, the sick leave percentage over a period depicts the total number of work days lost due to reported sick leave as a fraction of total number of scheduled work days.

In the country-wide model, percentages on all of the 385 municipalities over the period from 2002 to 2011 have been obtained from Statistics Norway[8]. These data are rounded to one decimal point, and are therefore not optimal. Their accuracy is however considered more than sufficient to give a satisfactory picture of the causal relationships under study.

Further, data on the sick leave percentage for the 43 municipalities in Troms and Finnmark, with accuracy down to the 14th decimal point, have been obtained from NAV Troms. Due to different methods of measurement[9] the data from NAV are not completely identical to those supplied by Statistics Norway. However, no loss of homogeneity in the input takes place. They are used in separate regression models and are therefore defined as two separate dependent variables.

The share of sick workers

The data on share of sick workers have been obtained from Statistics Norway for all 428 municipalities in Norway. As with the sick leave percentage they only include one decimal point and are therefore not optimal, but sufficient, in terms of accuracy.

Some descriptive statistics on the dependent variables as they will be divided in the regressions are presented in table 5.1.2 for the country-wide model and table 5.1.3 for the northern model.

Table 5.1.2: Descriptive statistics on dependent variables used in the country-wide model.

Variable		*Mean*	*Std. Dev.*	*Min*	*Max*	*Observations**
Male workers						
Sick leave percentage	overall	5.11%	1.21%	0.57%	11.35%	N = 3850
	between		0.94%	2.03%	8.82%	n = 385
	within		0.77%	2.34%	9.09%	T = 10
Share of sick workers	overall	4.54%	1.09%	0.53%	9.43%	N = 3850
	between		0.89%	1.88%	7.82%	n = 385
	within		0.64%	0.93%	7.61%	T = 10
Female workers						
Sick leave percentage	overall	7.69%	1.49%	2.83%	15.33%	N = 3850
	between		1.16%	4.73%	11.44%	n = 385
	within		0.94%	3.84%	13.53%	T = 10
Share of sick workers	overall	7.20%	1.41%	1.58%	13.83%	N = 3850
	between		1.14%	4.32%	10.78%	n = 385
	within		0.84%	3.19%	12.12%	T = 10

****N is the total number of data points, n is the number of individuals and T is the number of time periods.***

Table 5.1.3: Descriptive statistics on dependent variables used in the northern model.

Variable		*Mean*	*Std. Dev.*	*Min*	*Max*	*Observations**
Male workers						
Sick leave percentage	overall	6.39%	1.46%	2.78%	12.27%	N = 430
	between		1.03%	3.71%	8.59%	n = 43
	within		1.05%	2.66%	11.69%	T = 10
Share of sick workers	overall	5.40%	1.29%	2.20%	10.88%	N = 430
	between		0.91%	3.16%	7.44%	n = 43
	within		0.92%	2.07%	9.30%	T = 10
Female workers						
Sick leave percentage	overall	8.30%	2.08%	3.24%	14.57%	N = 430
	between		1.14%	5.94%	10.59%	n = 43
	within		1.75%	3.42%	14.12%	T = 10
Share of sick workers	overall	7.80%	1.63%	3.03%	14.63%	N = 430
	between		1.15%	5.33%	9.96%	n = 43
	within		1.16%	3.66%	13.65%	T = 10

****N is the total number of data points, n is the number of individuals and T is the number of time periods.***

We see that the estimated means are higher in the north than in the rest of the country for all variables, and in general higher among women. For the country-wide model there is more between-variation than within-variation in the variables, while in the north the within-variation is highest. This follows logically from the latter having fewer, more homogenous municipalities. The standard deviations of the variables used on the north are seen to be in general higher than on the variables used on the rest of Norway, both between, within and overall.

5.1.2 Demographic variables

Age

Two separate variables are used for capturing age effects. The number of people between 16-25 years of age and the number of people between 55-66 years of age, both as a share of total population between 16-66 years of age. The age variables are also divided by gender in the regressions, where the variables then are defined as the number of men (women) in a specific age group as a share of total number of men (women) in the population between 16-66 years of age.

Immigrants

As seen in table 5.1.1, this variable is defined as the number of first and second generation immigrants as a share of total population. It is also divided by gender in the regressions, where it is then defined as number of male (female) first and second generation immigrants as a share of total male (female) population.

Municipality size

The population size per first of January 2002, which is the beginning of the measured period. The variable is used as a time-invariant, since the population size is not expected to have changed considerably over the period. Changes in population size over the sampled time period are mostly relatively small and slow, and their effect on sick leave behavior is considered miniscule. Any significant effects caused by this variable are rather found in level differences between municipalities.

In creating the variable, the municipalities in the country-wide model are divided into seven categories based on size of population.

1) More than 40 000 inhabitants
2) Between 20 000 and 40 000 inhabitants
3) Between 10 000 and 20 000 inhabitants
4) Between 5000 and 10 000 inhabitants
5) Between 2000 and 5000 inhabitants
6) Between 1000 and 2000 inhabitants
7) Less than 1000 inhabitants

The division is based on Statistics Norway`s database for standard classifications[10]. The category «more than 50 000 inhabitants» has been changed to «more than 40 000 inhabitants» to get a larger sample size in this group. Further, the category «less than 2000 inhabitants» has been divided into «1000-2000 inhabitants» and «less than 1000 inhabitants».

For the regressions on Troms and Finnmark the categories listed above had to be altered due to a limited number of municipalities in several of the groups[11]. Category groups 1, 2, 3 and 4 were instead combined to one group, as were groups 6 and 7. The occurrence of municipalities in the different categories, separated by model, is as seen in tables 5.1.4 and 5.1.5.

Table 5.1.4: Frequency of municipalities in different size categories. All municipalities except those located in Troms and Finnmark.

Number of inhabitants	*Frequency*	*% of total*
More than 40 000	14	3.64
Between 20 000 and 40 000	27	7.01
Between 10 000 and 20 000	55	14.29
Between 5000 and 10 000	84	21.82
Between 2000 and 5000	131	34.03
Between 1000 and 2000	56	14.55
Less than 1000	18	4.68
Total:	**385**	**100.00**

Table 5.1.5: Frequency of municipalities in different size categories. Municipalities in Troms and Finnmark.

Number of inhabitants	*Frequency*	*% of total*
More than 5000	9	20.93
Between 2000 and 5000	17	39.53
Less than 2000	17	39.54
Total:	**43**	**100.00**

Categories 2 and 3 in the regressions for Troms and Finnmark, and 2-7 in the regressions for the rest of Norway are set as indicator variables, both with its respective category 1 as a reference point. The estimated results then depict if and how sick leave usage in the biggest municipalities differs from sick leave usage in municipalities of smaller sizes. A taste of what the results might look like can however be given already now, by looking at the mean sick leave percentage in the different categories. This is presented in tables 5.1.6 and 5.1.7.

Table 5.1.6: The sick leave percentage for each size category in the country-wide model per 2002. Divided by gender.

The sick leave percentage	*Mean*	*Std.dev*	*Min*	*Max*	*Observations*
Male workers					
Municipalities w/ more than 40 000 inhabitants	4.69%	1.06%	2.83%	7.70%	195
Municipalities w/ 20 000-40 000 inhabitants	4.98%	0.88%	3.20%	7.93%	364
Municipalities w/ 10 000-20 000 inhabitants	5.26%	1.10%	2.80%	9.55%	741
Municipalities w/ 5000-10 000 inhabitants	5.26%	1.06%	2.80%	10.18%	1157
Municipalities w/ 2000-5000 inhabitants	5.16%	1.31%	2.18%	11.35%	1924
Municipalities w/ 1000-2000 inhabitants	5.31%	1.54%	2.30%	13.03%	936
Municipalities w/ less than 1000 inhabitants	4.45%	1.79%	0.57%	12.28%	247
Female workers					
Municipalities w/ more than 40 000 inhabitants	7.65%	1.10%	5.53%	10.48%	195
Municipalities w/ 20 000-40 000 inhabitants	8.10%	1.02%	5.75%	11.35%	364
Municipalities w/ 10 000-20 000 inhabitants	8.11%	1.29%	4.70%	13.23%	741
Municipalities w/ 5000-10 000 inhabitants	7.96%	1.25%	4.63%	12.28%	1157
Municipalities w/ 2000-5000 inhabitants	7.65%	1.62%	3.35%	15.33%	1924
Municipalities w/ 1000-2000 inhabitants	7.66%	1.85%	3.40%	14.98%	936
Municipalities w/ less than 1000 inhabitants	7.13%	2.12%	2.83%	15.23%	247

Table 5.1.7: The sick leave percentage for each size category in the Northern model. 2002-2011. Divided by gender.

The sick leave percentage	*Mean*	*Std.dev*	*Min*	*Max*	*Observations*
Male workers					
Municipalities w/ more than 5000 inhabitants	5.83%	1.18%	3.74%	9.54%	117
Municipalities w/ 2000-5000 inhabitants	6.28%	1.37%	3.24%	9.99%	221
Municipalities w/ less than 2000 inhabitants	6.65%	1.67%	2.78%	13.16%	221
Female workers					
Municipalities w/ more than 5000 inhabitants	9.18%	1.24%	6.57%	12.77%	117
Municipalities w/ 2000-5000 inhabitants	9.13%	1.43%	4.07%	13.08%	221
Municipalities w/ less than 2000 inhabitants	9.03%	1.91%	4.14%	14.57%	221

Table 5.1.6 shows that on a national level, the lowest average sick leave percentage for both women and men can in fact be found in the *smallest* municipalities. The same holds for female workers in the north. However, the standard deviations are found to be relatively large, so it is not possible to give any definite ranking of sick leave usage in the size groups based on this. That there is an apparent trend can however not be denied.

Tables 5.1.8 and 5.1.9 present descriptive statistics on all the demographic variables chosen, divided by model. *Municipality size* does not change over the time period and therefore has no estimates on within-variation. All statistics is on percentage form except *municipality size*, which shows the number of inhabitants.

Table 5.1.8: Descriptive statistics on demographic variables in the country-wide model

Variable		***Mean***	***Std. Dev.***	***Min***	***Max***	***Observations****
Male workers						
Age 16-25	overall	19.38%	1.97%	12.75%	28.90%	N = 3850
	between		1.73%	15.12%	25.47%	n = 385
	within		0.94%	14.46%	23.68%	T = 10
Age 55-66	overall	22.03%	3.30%	12.27%	36.98%	N = 3850
	between		2.68%	14.51%	30.96%	n = 385
	within		1.92%	13.36%	30.85%	T = 10
Immigrants	overall	4.90%	3.34%	0.00%	29.60%	N = 3850
	between		2.86%	0.05%	25.06%	n = 385
	within		1.73%	-4.57%	21.53%	T = 10
Female workers						
Age 16-25	overall	19.07%	2.17%	7.08%	31.12%	N = 3850
	between		1.95%	14.06%	28.90%	n = 385
	within		0.96%	11.68%	25.01%	T = 10
Age 55-66	overall	22.22%	3.14%	11.76%	36.83%	N = 3850
	between		2.62%	14.38%	30.26%	n = 385
	within		1.72%	12.92%	29.60%	T = 10
Immigrants	overall	5.16%	2.91%	0.00%	27.20%	N = 3850
	between		2.57%	0.42%	22.89%	n = 385
	within		0.14%	-0.37%	14.36%	T = 10
Non gender-specific						
Municipality size	Overall	11 071	31 281	232	508 726	N = 3850

N is the total number of data points, n is the number of individuals and T is the number of time periods.

Table 5.1.9: Descriptive statistics on demographic variables in the northern model

Variable		***Mean***	***Std. Dev.***	***Min***	***Max***	***Observations****
Male workers						
Age 16-25	overall	18.44%	2.48%	10.61%	25.48%	N = 430
	between		2.12%	12.68%	23.09%	n = 43
	within		1.33%	15.12%	22.48%	T = 10
Age 55-66	overall	23.89%	4.10%	13.45%	35.57%	N = 430
	between		3.40%	17.20%	31.89%	n = 43
	within		2.35%	15.17%	30.11%	T = 10
Immigrants	overall	4.03%	2.40%	0.60%	13.50%	N = 430
	between		2.05%	0.99%	10.37%	n = 43
	within		1.28%	0.66%	11.48%	T = 10
Female workers						
Age 16-25	overall	18.46%	1.83%	13.67%	25.17%	N = 430
	between		1.42%	15.72%	22.00%	n = 43
	within		1.18%	14.36%	22.68%	T = 10
Age 55-66	overall	23.59%	3.84%	14.30%	36.05%	N = 430
	between		3.22%	16.73%	30.21%	n = 43
	within		2.15%	16.74%	29.72%	T = 10
Immigrants	overall	5.57%	3.02%	1.10%	19.20%	N = 430
	between		2.80%	2.50%	16.70%	n = 43
	within		1.20%	2.32%	10.45%	T = 10
Non gender-specific						
Municipality size	overall	5239	9509	963	60 086	N = 430

N is the total number of data points, n is the number of individuals and T is the number of time periods.

Average municipality size in the northern municipalities, defined by size of population, is found to be less than half of the average for the rest of the country. Further, in comparison to the rest of the country, Troms and Finnmark appear to have a lower average share of younger people and a higher average share of older people for both genders. Lastly, the north is found to have a lower share of male immigrants, but a higher share of female immigrants than the country average.

5.1.3 Social variables

Level of education

This variable is defined as the share of total population over the age of 25 that has completed a high school education or higher. Unfortunately, no gender specific data were available.

Participation in elections

Shows the percentage of total population that exercised their right to vote. Data from the Parliamentary elections in 2001, 2005, 2009 and 2013 are used. The electoral participation on a municipality level for these four years is averaged out to create a time-invariant variable.

Disability benefits usage in preceding year

The numbers are calculated as the number of people receiving disability benefits as a share of total population between 18 and 67 years of age[12] [13]. When used in the regressions the variable is input on lagged form, e.g. when put against the sick leave percentage in time *t*, it will be used on the form *t-1*.

As with the sick leave percentage, these data have been collected both from Statistics Norway and from NAV, where the former are used in the country-wide model and the latter in the northern model. They differ in the same way as specified in section 5.1.1. They are further divided by gender in the northern model, but not in the country-wide model.

Descriptive statistics for all the social variables are included in tables 5.1.10 and 5.1.11. As seen, disability benefits usage preceding year is divided by gender in the northern model but not in the country-wide model. Further, as with municipality size, participation in elections is a time-invariant variable and has no statistics on within-variation.

Table 5.1.10: Descriptive statistics on social variables in the country-wide model

Variable		*Mean*	*Std. Dev.*	*Min*	*Max*	*Observations**
Non gender-specific						
Level of education	overall	69.20%	7.25%	32.00%	88.00%	N = 3850
	between		6.97%	37.50%	87.50%	n = 385
	within		2.04%	62.60%	77.90%	T = 10
Disability benefits usage preceding year	overall	10.94%	2.92%	3.88%	23.15%	N = 3850
	between		2.85%	4.43%	21.27%	n = 385
	within		0.62%	6.97%	15.04%	T = 10
Participation in elections	overall	76.32%	3.08%	66.08%	86.53%	N = 3850

***N is the total number of data points, n is the number of individuals and T is the number of time periods.**

Table 5.1.11: Descriptive statistics on social variables in the northern model

Variable		*Mean*	*Std. Dev.*	*Min*	*Max*	*Observations**
Male workers						
Disability benefits usage preceding year	overall	13.02%	3.27%	6.50%	23.07%	N = 430
	between		3.13%	6.86%	20.53%	n = 43
	within		1.03%	9.51%	16.48%	T = 10
Female workers						
Disability benefits usage preceding year	overall	16.38%	3.39%	8.93%	24.49%	N = 430
	between		3.31%	9.28%	22.46%	n = 43
	within		0.88%	13.39%	20.10%	T = 10
Non gender-specific						
Level of education	overall	60.90%	8.10%	43.00%	78.00%	N = 430
	between		7.90%	46.60%	76.20%	n = 43
	within		2.13%	55.81%	65.81%	T = 10
Participation in elections	overall	72.13%	2.82%	65.88%	79.38%	N = 430

***N is the total number of data points, n is the number of individuals and T is the number of time periods.**

As we see, the average share of people with a completed high school education is 8.3 percentage points lower in the northern municipalities than in the rest of the country. Further, there is lower average electoral participation in the north. Also, the share of both men and women in the north utilizing disability benefits is higher than the share of the total population utilizing it elsewhere in Norway.

5.1.4 Labor market variables

Unemployment

Share of population between 15 and 74 years of age that is registered as unemployed. The variable is separated for men and women in the regressions.

Bankruptcies

Number of company bankruptcies in a given year as a share of total number of registered companies at the beginning of the year.

Labor market structure

Defined as number of workers employed in a given sector or industry as a share of total number of workers. Four areas of the labor market will be given special attention.

The primary sector - Agriculture, forestry and fishing.

The public sector - Public administration, military and social insurance.

The health care sector - Health and social services.

The service industry - Wholesale and retail trade, accommodation and food service activities, transportation and storage, financial services.

As covered in chapter 2, the health sector and the service industry are the two biggest labor markets on a national scale and naturally also have the biggest shares of total sick leave usage. For this reason they are included as explanatory variables. The primary sector is further included due to being a significant labor provider in the northern municipalities, while the public sector is included due to being significantly bigger in the north than the rest of the country, and it is further a frequently used variable in the literature.

Small companies

The effects of company size are reflected through the share of registered companies that have between one and nine employees.

Descriptive statistics on all the labor market variables are shown in tables 5.1.12 and 5.1.13.

Table 5.1.12: Descriptive statistics on labor market variables in the country-wide model

Variable		*Mean*	*Std. Dev.*	*Min*	*Max*	*Observations**
Male workers						
Unemployment	overall	2.70%	1.33%	0.10%	10.50%	N = 3850
	between		0.93%	0.78%	8.72%	n = 385
	within		0.95%	-0.62%	8.64%	T = 10
Employees in primary sector	overall	11.28%	9.23%	0.18%	57.69%	N = 3850
	between		9.12%	0.29%	52.04%	n = 385
	within		1.49%	-0.21%	21.26%	T = 10
Employees in public sector	overall	5.70%	3.27%	0.73%	40.99%	N = 3850
	between		3.13%	2.09%	30.32%	n = 385
	within		0.98%	-3.62%	16.37%	T = 10
Employees in health sector	overall	5.40%	2.09%	0.00%	17.36%	N = 3850
	between		1.98%	1.42%	14.97%	n = 385
	within		0.66%	1.90%	9.03%	T = 10
Employees in service industry	overall	31.99%	9.38%	9.68%	63.35%	N = 3850
	between		9.23%	13.29%	61.91%	n = 385
	within		1.71%	22.46%	45.79%	T = 10
Female workers						
Unemployment	overall	2.36%	1.16%	0.00%	9.90%	N = 3850
	between		0.84%	0.61%	6.32%	n = 385
	within		0.80%	-0.36%	6.95%	T = 10
Employees in primary sector	overall	4.14%	3.58%	0.00%	23.65%	N = 3850
	between		3.46%	0.14%	20.04%	n = 385
	within		0.93%	-1.90%	11.44%	T = 10
Employees in public sector	overall	6.56%	2.80%	2.13%	35.37%	N = 3850
	between		2.39%	2.68%	29.59%	n = 385
	within		1.47%	-1.62%	35.47%	T = 10
Employees in health sector	overall	35.74%	5.30%	10.66%	55.78%	N = 3850
	between		4.94%	23.58%	53.54%	n = 385
	within		1.93%	13.50%	43.50%	T = 10
Employees in service industry	overall	29.43%	6.21%	13.46%	49.10%	N = 3850
	between		6.00%	15.06%	48.07%	n = 385
	within		1.61%	21.50%	37.22%	T = 10
Non gender-specific						
Bankruptcies	overall	0.64%	0.53%	0.00%	4.95%	N = 3850
	between		0.33%	0.00%	1.78%	n = 385
	within		0.42%	-0.82%	4.18%	T = 10
Small companies	overall	27.15%	5.00%	14.47%	46.58%	N = 3850
	between		4.69%	15.61%	40.86%	n = 385
	within		1.74%	16.25%	35.38%	T = 10

N is the total number of data points, n is the number of individuals and T is the number of time periods.

Table 5.1.13: Descriptive statistics on labor market variables in the northern model

Variable		*Mean*	*Std. Dev.*	*Min*	*Max*	*Observations**
Male workers						
Unemployment	overall	5.15%	2.58%	1.00%	15.00%	N = 430
	between		2.09%	1.71%	11.31%	n = 43
	within		1.55%	1.29%	10.93%	T = 10
Employees in primary sector	overall	15.70%	8.63%	2.67%	37.60%	N = 430
	between		8.53%	3.44%	33.03%	n = 43
	within		1.81%	10.13%	24.44%	T = 10
Employees in public sector	overall	10.76%	8.14%	2.08%	49.44%	N = 430
	between		8.07%	4.32%	44.63%	n = 43
	within		1.56%	6.08%	17.88%	T = 10
Employees in health sector	overall	8.13%	2.60%	3.36%	17.25%	N = 430
	between		2.32%	4.47%	15.00%	n = 43
	within		1.20%	2.89%	12.50%	T = 10
Employees in service industry	overall	29.41%	6.54%	14.39%	47.65%	N = 430
	between		6.19%	17.50%	45.63%	n = 43
	within		2.28%	22.94%	37.83%	T = 10
Female workers						
Unemployment	overall	3.22%	1.87%	0.50%	11.80%	N = 430
	between		1.46%	1.45%	6.78%	n = 43
	within		1.19%	0.41%	9.52%	T = 10
Employees in primary sector	overall	3.37%	2.35%	0.13%	14.02%	N = 430
	between		2.05%	0.62%	7.82%	n = 43
	within		1.18%	-0.06%	11.08%	T = 10
Employees in public sector	overall	9.93%	4.30%	2.11%	25.03%	N = 430
	between		3.97%	5.21%	23.53%	n = 43
	within		1.75%	4.50%	18.60%	T = 10
Employees in health sector	overall	39.93%	6.67%	20.31%	60.91%	N = 430
	between		6.21%	29.27%	56.34%	n = 43
	within		2.58%	30.55%	48.16%	T = 10
Employees in service industry	overall	23.25%	5.57%	8.89%	36.14%	N = 430
	between		5.28%	13.85%	32.87%	n = 43
	within		1.95%	18.15%	28.72%	T = 10
Non gender-specific						
Bankruptcies	overall	0.79%	0.71%	0.00%	4.17%	N = 430
	between		0.31%	0.33%	1.50%	n = 43
	within		0.64%	0.63%	4.03%	T = 10
Small companies	overall	33.47%	5.35%	16.05%	47.50%	N = 430
	between		4.78%	20.42%	41.10%	n = 43
	within		2.50%	24.94%	41.02%	T = 10

N is the total number of data points, n is the number of individuals and T is the number of time periods.

The tables show a much higher unemployment rate for men in the north than in the rest of the country. It is also found to be higher for women. Employment in the primary sector is overall higher among men, and higher for both genders in the north. The share of employment in the public sector is found to be more than twice as high among men in the north compared to elsewhere in the country. For women this share is also nearly twice as big. Further, employment in the health sector is found to be more than six times bigger among women than among men on a national level, while it is more than four times as big in the northern municipalities. Employment in the service industry is found to be bigger in the north for men, but smaller for women.

Lastly, the share of bankruptcies is slightly higher in the northern municipalities, but insignificantly so due to a large standard deviation. The prevalence of companies with 1-9 employees also appears to be higher in the north.

5.2 Methodology

The inputs used in the regressions are on the form of a panel data set. Panel data is a combination of cross-sectional and time-series data, and describes a number of individual units over several sequential points in time. Regression analysis on this form of data can therefore capture variation both over units and over time. The panel data set used in the present paper is defined as a short panel, since it examines a large amount of individual units, or municipalities, over a relatively small time period. The benefits of using panel data include but are not limited to[14]:

(1) Panel data is able to control for unit-invariant or time-invariant variables affecting the analysis, whereas this cannot be done with cross-sectional or time series data. In short, it assumes that the individuals, firms or regions under study are heterogeneous, something which most often holds true.
(2) The data is more informative than simple time series data. Time series data are often filled with multicollinearity, but the added cross-sectional aspect of panel data gives more variation and information on the variables.
(3) Likewise, it has informational advantages over cross-sectional data since it does not simply look at the differences between variables at a point in time, but also measures how this difference changes over time both between and within units. Repeated cross-sections can also reveal how variables change over time, but unlike panel data they cannot capture the adjustements between one period and the next.
(4) Effects that are simply not measurable in cross-sectional or time series data can be measured by using panel data.

When running regressions on panel data, a choice has to be made between using a fixed-effects and a random effects model. Both assume the existence of individual differences in the municipalities, and that these are captured by the intercept parameter. The fixed-effects model however assumes that the intercepts of each individual are fixed, while the random-effects model defines them as consisting of two parts: the population average intercept plus a random individual specific effect, or error component. The error component is assumed to have a mean of zero, be uncorrelated across individuals and have a constant variance. For the present analysis the random-effects model has been chosen, due to inclusion of the important time-invariant explanatory variables *municipality size* and *participation in elections*. The fixed-effects model does not allow for such time-invariant effects[15].

After choosing a model, the natural next step is to test for heterogeneity between the municipalities to see if there actually are any random effects present, or if they share a common intercept. This is done by testing the null hypothesis that the total variance of the error components is equal to zero against the alternative hypothesis that it is greater than zero. The Lagrange Multiplier test[16] is used for this purpose. The test calculates the χ^2, or chi-square, random variable for each regression, where χ^2 is defined as the sum of squares of k standard normal random variables. Simply put, a larger value of χ^2 signifies a larger total variance around zero in the k random variables, which in this case are the individual random error components. Low values of χ^2 signify that there is little or no heterogeneity present. The results for each model is as reported in table 5.2.1.

Table 5.2.1: Estimation results for the Lagrange-Multiplier test. Chi-square and p-values.

	χ^2	**p-value**
Troms & Finnmark model		
Sick leave percentage men	109.25	0.000
Sick leave percentage women	27.07	0.000
Share of sick workers men	53.47	0.000
Share of sick workers women	34.85	0.000
Country-wide model		
Sick leave percentage men	2083.58	0.000
Sick leave percentage women	3002.47	0.000
Share of sick workers men	1625.76	0.000
Share of sick workers women	2448.02	0.000

As seen, the null hypothesis that the error component variance is equal to zero is thoroughly rejected for all reasonable probability values. The existence of individual specific random effects in the municipalities has therefore been established, and the random-effects model is deemed appropriate.

The random-effects model makes the rather strong assumption that the error components are uncorrelated with the explanatory variables. This might however not be the case. One trait of the error components is that they will reflect any significant explanatory variables not included in the model when running regressions. If these non-included factors are correlated with any of the used regressors, then their effects will be attributed to them, creating inconsistent results (Carter Hill, Griffiths & Lim, 2012). The fixed-effects model does however not make this assumption, and is therefore commonly used in the literature when comparing regions (see for example Bragstad, Regbo & Sagaveen, 2006).

However, when the analysis includes time-invariant variables, like the present analysis does, then it is not possible to use fixed-effects without dropping them. Some papers when faced with this dilemma have chosen the solution of simply using random-effects, and then specifying that the results might be suboptimal (Osterkamp & Röhn (2007)). Another solution is to use an instrumental variables estimator known as the Hausman-Taylor estimator[17]. The estimator provides coefficients with the consistency of the fixed-effects model without having to drop the time-invariant variables. When using this estimator one has to differentiate between which of the regressors are influenced by correlation and which are not. The influenced regressors are defined as endogenous while those not influenced are defined as exogenous.

Exogenous variables are per definition not influenced directly by other factors, and it is therefore natural to assume that all the demographic variables in the present models are exogenously given. The same goes for the variables on *labor market structure* and *small companies*. Further, *disability benefits usage preceding year* must be exogenous because the values of other factors in time *t* can not influence the value of disability benefits usage in time *t-1*. The remaining four factors are thus perceived as endogenous. They are: *level of education*, *participation in elections*, *unemployment* and *bankruptcies*. *Participation in elections* is time-invariant while the remaining three are time-variant.

We should inspect whether the time-varying variables contain enough within-panel variation to serve as their own instruments. These results have already been presented in section 5.1 but are repeated here in tables 5.2.2 and 5.2.3 for convenience.

Table 5.2.2: Descriptive statistics on endogenous variables in the country-wide model

Variable		*Mean*	*Std. Dev.*	*Min*	*Max*	*Observations**
Male workers						
Unemployment	Overall	2.70%	1.33%	0.10%	10.50%	N = 3850
	Between		0.93%	0.78%	8.72%	n = 385
	Within		0.95%	-0.62%	8.64%	T = 10
Female workers						
Unemployment	Overall	2.36%	1.16%	0.00%	9.90%	N = 3850
	Between		0.84%	0.61%	6.32%	n = 385
	Within		0.80%	-0.36%	6.95%	T = 10
Non gender-specific						
Level of education	Overall	69.20%	7.25%	32.00%	88.00%	N = 3850
	Between		6.97%	37.50%	87.50%	n = 385
	Within		2.04%	62.60%	77.90%	T = 10
Bankruptcies	Overall	0.64%	0.53%	0.00%	4.95%	N = 3850
	Between		0.33%	0.00%	1.78%	n = 385
	Within		0.42%	-0.83%	4.18%	T = 10

Table 5.2.3: Descriptive statistics on endogenous variables in the northern model

Variable		*Mean*	*Std. Dev.*	*Min*	*Max*	*Observations**
Male workers						
Unemployment	Overall	5.15%	2.58%	1.00%	15.00%	N = 430
	Between		2.09%	1.71%	11.31%	n = 43
	Within		1.55%	1.29%	10.93%	T = 10
Female workers						
Unemployment	Overall	3.22%	1.87%	0.50%	11.80%	N = 430
	Between		1.46%	1.45%	6.78%	n = 43
	Within		1.19%	0.41%	9.52%	T = 10
Non gender-specific						
Level of education	Overall	60.90%	8.10%	43.00%	78.00%	N = 430
	Between		7.90%	46.60%	76.20%	n = 43
	Within		2.13%	55.81%	65.81%	T = 10
Bankruptcies	Overall	0.79%	0.71%	0.00%	4.17%	N = 430
	Between		0.31%	0.33%	1.50%	n = 43
	Within		0.64%	-0.63%	4.03%	T = 10

The within-panel variation is deemed sufficient, although *level of education* might be a rather weak instrument.

The next assumption is that the exogenous variables are sufficient as instruments for the time-invariant variable *participation in elections*. For this purpose we must check the correlation matrix between these variables as presented below.

Table 5.2.4: Correlation between the endogenous time-invariant variable and the exogenous variables in the country-wide model and the northern model

Variable	***The country-wide model***	***The northern model***
	Participation in elections	*Participation in elections*
Male workers		
Age 55-66	-0.2135	0.0869
Age 16-25	0.1212	0.1698
Immigrants	0.1213	-0.0958
Employment primary sector	-0.2628	-0.2812
Employment public sector	0.0108	0.2485
Employment health sector	-0.0860	0.1279
Employment service industry	0.2644	-0.2175
Disability benefits preceding year	.	0.1865
Female workers		
Age 55-66	-0.2632	0.2023
Age 16-25	0.1091	0.0985
Immigrants	0.1301	-0.1546
Employment primary sector	-0.0905	-0.0619
Employment public sector	0.0451	0.1389
Employment health sector	-0.1186	0.4323
Employment service industry	0.1342	-0.3355
Disability benefits preceding year	.	0.3361
Non-gender specific		
Disability benefits preceding year	-0.4113	.
Small companies	0.0160	0.0708

The strength of the correlations differs between the models and across the variables, but the relationships are presumed strong enough for the continued estimation. The Hausman-Taylor estimates for the coefficients has thus been used and the results are as presented in the following chapter.

6 Estimation results

In the regression analyses two dependent variables are used, and further looked at for men and women separately. These four regressions are run twice, first on the country-wide model and afterwards on the model for the 43 northern municipalities of Troms and Finnmark. This leaves a total of eight sets of regression output over two different models to be examined in the present chapter.

The regressions performed on Troms and Finnmark and those performed on the rest of Norway are not completely identical in terms of variable definitions. The country-wide regressions use data on the sick leave percentage and on disability benefits usage that have been obtained from Statistics Norway, while the regressions on Troms and Finnmark use data on these variables obtained from NAV Troms. The details on how data from these two sources differ was covered in section 5.1.1. The main reason for utilizing NAVs data on the regressions for Troms and Finnmark is that they are more accurate due to a higher number of decimal points. The northern municipalities represent a much smaller sample size than the rest of Norway, and the higher prevalence of relatively small municipalities leads as shown in section 4.2.1 to a higher variance in sick leave usage. Higher accuracy in the inputs is therefore valuable for optimal estimation results. Ideally, NAVs data would have been obtained for the whole country and used in all the regressions, but unfortunately this was not feasible in the scope of this Master thesis.

The results of the eight regression analyses are separated for Troms & Finnmark and the rest of the country, and are as presented in tables 6.1 and 6.2 on the next pages. The results are also summarized in the following sub-chapters. The regressions are strongly balanced, meaning that all municipalities are observed in all time periods.

Table 6.1: Regression output with Hausman –Taylor estimation. Country-wide model. Separated for men and women, and for two dependent variables.*

	Country-wide model			
	The Sick Leave Percentage		*The Share of Sick Workers*	
	Men	Women	Men	Women
R-square: Overall	**0.3726**	**0.2284**	**0.3681**	**0.2365**
Between	**0.5044**	**0.3162**	**0.5437**	**0.3485**
Within	**0.1741**	**0.0946**	**0.0403**	**0.0371**
Demographic variables				
Age 16-25 years	0.00514 (-0.41)	-0.0523*** (-3.33)	-0.00496 (-0.44)	-0.0258+ (-1.75)
Age 55-66 years	-0.0111 (-1.08)	-0.0620*** (-4.55)	0.0136 (1.47)	-0.00685 (-0.54)
Immigrants	-0.0531*** (-5.52)	0.00261 (0.16)	-0.0456*** (-5.23)	0.00457 (0.31)
Municipalities w/ more than 40 000 inhabitants (reference category)				
Municipalities w/ 20 000-40 000 inhabitants	-0.000732 (-0.21)	0.00148 (0.30)	0.000129 (0.04)	0.00273 (0.57)
Municipalities w/ 10 000-20 000 inhabitants	0.00190 (0.58)	0.000811 (0.18)	0.00313 (1.00)	0.00308 (0.70)
Municipalities w/ 5000-10 000 inhabitants	0.000425 (0.13)	-0.00448 (-0.98)	0.00250 (0.80)	-0.00199 (-0.45)
Municipalities w/ 2000-5000 inhabitants	0.000826 (0.25)	-0.00767+ (-1.71)	0.00287 (0.92)	-0.00431 (-1.00)
Municipalities w/ 1000-2000 inhabitants	0.000619 (0.17)	-0.0110* (-2.27)	0.00252 (0.73)	-0.00716 (-1.54)
Municipalities w/ less than 1000 inhabitants	-0.0105* (-2.36)	-0.0194*** (-3.34)	-0.00499 (-1.19)	-0.00886 (-1.59)
Social variables				
Level of education	-0.120*** (-9.97)	-0.152*** (-10.05)	-0.0607*** (-5.54)	-0.0382** (-2.69)
Participation in elections	-0.00131* (-2.52)	-0.00202* (-2.42)	-0.00193*** (-4.01)	-0.00352*** (-4.45)
Disability benefits usage preceding year	-0.0522** (-3.03)	-0.0829*** (-3.50)	-0.0252 (-1.60)	-0.0179 (-0.81)
Labor market variables				
Unemployment	0.119*** (8.62)	-0.188*** (-8.17)	0.0779*** (6.26)	-0.225*** (-10.47)
Workers in the primary sector	-0.0446*** (-6.19)	-0.0264 (-1.65)	-0.0322*** (-4.84)	-0.0314* (-2.07)
Workers in the public sector	0.00882 (0.79)	-0.0245* (-2.04)	0.00281 (0.27)	0.00255 (0.23)
Workers in the health sector	0.0118 (0.74)	0.00856 (0.86)	0.0151 (1.03)	0.0211* (2.24)
Workers in the service industry	0.0147* (2.35)	0.0106 (1.03)	0.0197*** (3.40)	0.00689 (0.71)
Small companies	0.0194** (2.88)	0.0423*** (4.79)	0.0165** (2.68)	0.0285*** (3.43)
Bankruptcies	0.148*** (5.18)	0.246*** (6.86)	0.0932*** (3.62)	0.125*** (3.72)
Constant	0.235*** (6.14)	0.363*** (5.79)	0.225*** (6.31)	0.366*** (6.15)

***t-statistics in parentheses. Coefficients marked by significance level. Coefficients significant at 0.1%, 1%, 5% and 10% level marked by ***, **, * and + respectively.**

Table 7.2: Regression output with Hausman-Taylor estimation. Troms & Finnmark model. Separated for men and women, and for two dependent variables.*

	Troms & Finnmark model			
	The Sick Leave Percentage		*The Share of Sick Workers*	
	Men	**Women**	**Men**	**Women**
R-square: Overall	**0.3209**	**0.1500**	**0.2906**	**0.2925**
Between	**0.4936**	**0.2408**	**0.5455**	**0.6213**
Within	**0.1591**	**0.1053**	**0.0504**	**0.0059**
Demographic variables				
Age 16-25 years	0.0145 (0.35)	0.121* (2.18)	-0.00390 (-0.10)	0.0776 (1.51)
Age 55-66 years	-0.165*** (-5.08)	-0.104* (-2.25)	-0.0990** (-3.28)	0.0195 (0.46)
Immigrants	-0.0542 (-1.30)	-0.0289 (-0.52)	-0.0513 (-1.33)	0.00640 (0.13)
Municipalities w/ more than 5000 inhabitants (reference category)				
Municipalities w/ 2000-5000 inhabitants	0.00271 (0.48)	-0.00652 (-0.73)	0.00291 (0.56)	-0.00214 (-0.30)
Municipalities w/ less than 2000 inhabitants	-0.00218 (-0.27)	-0.0196+ (-1.77)	-0.00443 (-0.60)	-0.0103 (-1.12)
Social variables				
Level of education	-0.0699+ (-1.92)	-0.151*** (-3.34)	-0.0165 (-0.49)	-0.0437 (-1.09)
Participation in elections	0.00237+ (1.75)	0.00353+ (1.83)	0.00282* (2.26)	0.00241 (1.50)
Disability benefits usage preceding year	0.0636 (1.37)	-0.0200 (-0.31)	0.0595 (1.38)	-0.000335 -0.01
Labor market variables				
Unemployment	-0.0699 (-1.63)	-0.187* (-2.56)	-0.0642 (-1.60)	-0.152* (-2.21)
Workers in the primary sector	0.00815 (0.28)	-0.0712 (-1.31)	0.0308 (1.14)	-0.0573 (-1.13)
Workers in the public sector	-0.0587* (-2.36)	-0.0972* (-2.11)	-0.0612** (-2.67)	-0.0209 (-0.49)
Workers in the health sector	-0.0103 (-0.24)	0.0411 (-1.24)	-0.0460 (-1.18)	0.00388 (0.13)
Workers in the service industry	-0.0251 (-1.01)	-0.0594 (-1.53)	-0.0176 (-0.76)	-0.000189 (-0.01)
Small companies	0.0508* (2.39)	-0.00239 (-0.09)	0.0416* (2.10)	-0.0182 (-0.70)
Bankruptcies	0.135 (1.63)	0.0856 (0.83)	0.0351 (0.45)	0.109 (1.11)
Constant	-0.0352 (-0.36)	-0.00543 (-0.04)	-0.120 (-1.34)	-0.0709 (-0.65)

***t-statistics in parentheses. Coefficients marked by significance level. Coefficients significant at 0.1%, 1%, 5% and 10% level marked by ***, **, * and + respectively.**

As seen in the tables, there are some considerable level differences in overall R-square between the genders. It appears that the model has more explanatory power for male workers than female workers. Especially the result for the sick leave percentage among women in the north is much lower compared to men in the north, indicating that there are more uncaptured factors affecting the sick leave percentage for women than there are such uncaptured factors for men.

Further, overall R-square changes considerably between the country-wide regressions and the northern regressions. For the sick leave percentage it appears that the model loses explanatory power for both genders when changing perspective from a national level to only the north. This effect is much larger for women than for men.

In the following four sections the results for each explanatory variable is listed in turn. The estimations for both models are mentioned together for easier comparison.

6.1 Demographic effects

Age

The regressions return some very interesting results on the age factors. It appears that on a national level, both the youngest and the oldest among female workers contribute to a lower level of sick leave usage among women. In the north however, workers between 16 and 25 actually *pull up* total sick leave usage in the female workforce.

For men the age group between 55 and 66 also contribute to lower sick leave, but the effect is only prominent in the north. No effects are found in either direction from a high share of young male workers.

Age mostly affect the sick leave percentage for women, while the share of workers utilizing sick leave is also affected for men.

Immigrants

When looking at the whole country, the results reveal that more male immigrants in the population *reduce* both total sick leave usage among men and the number of male workers having one or more reported sickness episodes in a given year. However, no significant relationships are found in the northern municipalities. Further, the share of female immigrants in the population is not found to have any significant effects on sick leave among women.

Municipality size

The effects of municipality size on sick leave turn out to be limited, but interesting. The results for both the north and for the rest of the country state that the sick leave percentage is *lower* in the smallest municipalities than in the biggest municipalities. This effect is especially strong for female workers on a national level.

A pattern is also revealed in the results for the size groups. In general, the estimated coefficients seem to grow as the municipalities become smaller. This hints towards that the power of the estimated difference grows as the size difference between the municipalities and the control group grows. Spoken more plainly, the importance of municipality size when measuring differences in sick leave usage across municipalities *is bigger when the size gap between the compared municipalities is bigger*. This result should not at all be surprising. Consequently, since there is less of a size gap between the municipalities in the north, this explains why the effects of smaller municipalities is less apparent there than in the country-wide model.

6.2 Social effects

Level of education

According to the results, municipalities where more people have acquired a high school education or higher should also experience lower levels of sick leave. In the country-wide model this effect holds across both genders and both dependent variables.

In the north however the effect seems to be in general weaker. It disappears completely on the share of sick workers among both genders, meaning that it only affects how much sick leave that is taken in total, but it has no found influence on the amount of workers using it.

Participation in elections

All other things equal, a higher electoral participation is accompanied by less workers on sick leave and less sick days in total when looking at municipalities in the whole country excluding Troms and Finnmark. These results are the same for men and for women.

However, in the northern municipalities the situation is quite the opposite. Municipalities where more people use their right to vote seem to have *more* workers reporting in sick, and the total number of sick days across workers is *higher*.

Disability benefits usage

A rather surprising result is obtained on this variable. On a national level, it appears that more people on disability benefits in the preceding year *pulls down* the sick leave percentage in the present year. The effect holds for both women and men. No such effects are however found in the northern model.

6.3 Labor market effects

Unemployment

On a national level, a higher unemployment rate predicts a lower sick leave percentage and share of sick workers among women. For men the results are however completely opposite, predicting a *higher* sick leave percentage and share of sick workers. It appears that higher unemployment in the municipality causes less sick leave among women and more sick leave among men, all other things equal.

The same negative effect on sick leave among female workers is found in Troms and Finnmark. Interestingly, for men the effect however disappears completely in the north.

Bankruptcies

In the country-wide model, a high number of bankruptcies in a given year predicts both more sick days taken and more workers reporting sick that same year. This holds equal for both men and women, and is consistent with existing theory and former research.

However, in the north of Norway, company bankruptcies appear to have no effect on sick leave usage. It neither affects the amount of people using it or the total amount of sick days.

Small companies

Changes in the share of companies that have between 1 and 9 employees, thus defined as small companies, have a rather unexpected effect on sick leave behavior. In the country-wide model municipalities with relatively more small companies also experience *higher* sick leave usage and more sick workers. These results are highly significant and consistent across genders.

For men the same results are discovered in Troms and Finnmark, but no connection is here found for women.

Labor market structure

Municipalities were a larger portion of the male workers are employed in the primary sector seem to experience lower sick leave usage among men and fewer sick male workers. More women in the primary sector also leads to fewer sick workers among women in total. Surprisingly, even though the primary sector is bigger in the north, no significant effects of employment there are found for Troms and Finnmark.

Employment in the public sector appears to have a general decreasing effect on sick leave, which is unexpected. On a national level the result is only evident for female workers and the sick leave percentage. The same holds when zooming in on only the north, but here also men are affected by public employment. In fact, while sick leave among men is unaffected in the country-wide model, it becomes negatively affected for both dependent variables when shifting focus to the north.

Employment in the health sector only yields one significant result across all eight regressions. When it increases, a higher share of sick workers among women is predicted on a national level. No such effects are found in the north.

Finally, higher employment in the service industries predicts a larger total amount of sick days and more sick workers among men in the country-wide model. This effect however disappears when looking at the north. Total sick leave usage among women is not predicted to be significantly influenced by working in the service industry.

7 Discussion

With a base in existing literature, this paper has attempted to identify a selection of explanatory variables for differences in the use of sick leave across municipalities. Two usable measurements for sick leave usage were included, defined as the sick leave percentage and the share of sick workers. The analysis was further divided between men and women, since the genders are theorized to have differing patterns of sick leave and to be afflicted differently by outside factors. The results of the regressions were covered in chapter 6, and are further discussed in the following.

The finding that an increasing share of older workers decreases sick leave usage is consistent with earlier findings by Markussen, Røed, Røgeberg & Gaure (2009); Older workers will have had more time to find a job that matches their physical and psychological profile, something which also minimizes stress and dissatisfaction in the work situation. Further, some form of selection effect might exist, where those with poorer health to a large extent fall out of the work force before reaching the age of 55. On a national level, the negative effect is only evident for women. However, also male workers experience this effect when looking at the north. Based on the possible explanations provided by Markussen, Røed, Røgeberg & Gaure (2009), one might then hypothesize that males between 55 and 66 in Troms and Finnmark on average experience higher job satisfaction than their peers in other parts of the country, and therefore use sick leave less often. Also, the selection effect might be stronger among men in the north, something which is supported by the fact that both Troms and Finnmark overall have higher disability benefits usage than the rest of the country[18]. However, the findings in the present paper are far too limited for this to be more than speculations.

The results on age effects further show that an increasing share of women between 16 and 25 contribute to lower sick leave usage on a national level. As opposed to the findings on workers between 55 and 66, this result is more in accordance with the conventional belief that younger workers are healthier and therefore should have less use for sick leave.

What is interesting is that in Troms and Finnmark young workers have the opposite effect. Municipalities in the north were a large part of the female labor force is under 25 tend to experience *more* sick leave. The formerly mentioned theories on why sick leave can be lower among older workers, and thus higher among younger workers, could also be applied here. An additional explanation, also mentioned by Markussen, Røed, Røgeberg & Gaure (2009), is that «*…young workers are bearers of a new and less strict norm set, and hence have lower thresholds for claiming sickness benefits.*». A survey among 1278 company managers and 1044 workers in the north of Norway performed by the employment agency Proffice revealed that 7 out of 10 managers and 8 out of 10 employees were under the impression that some groups in the workplace utilize sick leave more than others[19]. Apparently, based on their personal experience both workers and managers considered workers under the age of 35 to be the most absent, closely followed by parents of young children and women. Although based on opinions and not facts, this hints towards supporting the presently estimated connection between a higher share of young women and higher sick leave usage.

Following the same train of thought as earlier one could hypothesize that young women in the north possibly have overall poorer health than women of the same age elsewhere in the country, but at the same time a higher threshold for withdrawing completely from working life. A less strict norm set and a more relaxed attitude towards what constitutes legitimate sickness absence is also a possible factor, albeit a rather strong statement based on these limited findings. An important point to take into account here is that the estimated regression results on the other dependent variable, *the share of workers utilizing sick leave*, is not affected by this age group when looking at the north. Only *the amount of sick days* *taken* is found to differ. This indicates that there are not more workers in this age group utilizing sick leave than in the rest of the labor force, but rather that the average length and/or frequency of sickness spells is higher. The findings by Helde, Kristoffersen, Lysø & Thune (2010) as presented in section 2.2 however indicate that the average length of sickness spells for most major diagnostic groups is *lower* in the north of Norway. A possibility then, although resting more on assumptions than on scientific research, is that women in the age group 16-25 in the north of Norway have a higher frequency of sickness spells than women in other age groups, and thus more total sick leave.

A larger share of *immigrants* in the population does not seem to have any effects on sick leave usage in the north, signifying that in the north this group to a large extent shares the same health predispositions and attitudes towards sick leave usage as the native population. When looking at the rest of the country however, it appears that male immigrants utilize sick leave *less* than their native coworkers. As discussed in section 4.2.1 no definite results in either direction where expected on this variable (Dahl, Hansen & Olsen, 2010). Dutriex & Sjöholm (2003) also find sick leave to be lower in regions with more immigrants, but offer no further explanation of their results.

The variable on *disability benefits usage preceding year* was included for two reasons: in an attempt to capture its intrinsic property as an indicator of the general health level, and as a proxy for the general mindset towards receiving social benefits in the population. However, the found *negative* effect on sick leave usage speaks against the basic idea of both these theories, which is that sick leave and disability benefits usage is positively correlated.

A third possible explanation is then that it might be attributed to a form of health-based selection in the work force, much like discussed in relation to the effect of older workers. If many people with high amounts of sick leave transfer to disability benefits in a given year, then the total health of the workforce in the following year will be better, and thus total sick leave usage will be lower. A similar idea is proposed by Bragstad, Regbo & Sagsveen (2006), but when running regressions they however find a positive correlation between sick leave and disability benefits.

Disability benefits usage is higher in the northern municipalities than in the rest of the country, but no connection with sick leave is found in the regressions. The apparent lack of connection with sick leave can be interpreted as suggesting that the higher level of disability benefits usage in the north is *not* because of poorer health. No signs of an «absenteeism culture» are found either. So why are there more people on disability benefits in the north than in the rest of the country? The answer might lie in the high prevalence of limited labor markets in many of the smaller northern municipalities. Dutrieux & Sjöholm (2003) looked at Swedish municipality differences in sick leave for the year 2000 and showed that regions with high levels of sick leave often had limited labor demand and a high unemployment rate. They explain that a small local labor market limits the spectrum of possible industries and sectors to work in, and the selection of jobs in each of these. Difficulties in finding a relevant job in the region leave some workers with a choice between moving and utilizing an available social benefits scheme. The latter will then often be perceived as the less costly option.

The results on *municipality size* show that the smallest municipalities experience less sick leave than the biggest municipalities, keeping all other things constant. This is consistent with the descriptive statistics presented in section 5.1.2, where municipalities with less than 1000 inhabitants on a national level had the lowest average sick leave percentage of all the size groups for both men and women. Thus, the fact that Troms and Finnmark experience more sick leave than the rest of the country can *not* be attributed to the on average higher prevalence of small municipalities there. A possible explanation for these results might be found in the individual theories on social norms, and especially social control. Bovin & Wandall (1989) found that workers in small municipalities or rural areas have fewer sick leave days than workers living in larger municipalities and cities. They suggest that this might be due to social control, and that staying at home from work is harder in smaller societies. This might especially be true with regards to shirking-related absence.

The lower prevalence of sick leave usage in the smallest northern municipalities might also be due to the already mentioned higher level of disability benefits usage in many of them. As discussed above, when many of those with poor health drop out of the workforce we get a selection effect with regards to health, and the overall health of the population in the municipality will improve. This then drags down the total usage of sickness insurance.

However, defining the connection between municipality size and absence behavior is a rather difficult task, because it might not always be clear what it is we are comparing. It might be the labor market, health, medical practice, culture or general attitudes towards social insurance usage.

Obtained results on *level of education* supports earlier findings by Dutriex & Sjöholm (2003), Olsson (2004), Olsson (2007), Foss & Skyberg (2008) and Markussen, Røed, Røgeberg & Gaure (2009). A more educated population contributes to less total sick leave usage. In Troms and Finnmark however only the total amount of sick days decreases as the population gets more educated. The amount of people using sick leave is not affected. This indicates that when those with higher education get sick, they will have a lower frequency and/or average length of their sickness spells than the remaining labor force, but the share of educated workers *getting sick* will not differ from the share of workers getting sick in the remaining labor force. So in total the effect of education is stronger outside of the north, because not only the amount of sick days will decrease, but the workers basic decision between using it and not using it *at all* is affected.

A further important point is that the share of people with a completed high school education or higher is almost ten percentage points lower in the north than in the rest of the country (60.9% in the northern municipalities against 69.2% in the rest of the country). This indicates that part of the explanation for why sick leave usage is higher in the north might be found in the fact that the population has a lower average level of education. More precisely, it might be found in the underlying individual characteristics of those with a completed education beyond secondary school. As mentioned in chapter 3 on individual theories, Marmot (2004a) and the social gradient are essential in this context. Workers with a level of education *lower* than high school will have much more difficulties finding jobs that match their physical and psychological profile, or even finding steady work at all. They will consequently be less content, feel less in control of their own lives and due to a higher representation in lower-paid manual labor jobs, feel that they are less recognized for their individual skills and qualities. Further, an increasing share of the jobs being created in society today demands an education on high school level or above, or some form of craft certificate. The relative need for unskilled labor has on the other hand diminished severely over the last decades. Conversely, a higher education gives more freedom financially and a greater feeling of control in general. All these factors combined will affect the health of the individual worker and create a correlation with sick leave usage.

Moving on, a rather surprising result for Troms and Finnmark is that the social participation in the population, measured by the amount of people who use their right to vote, seems to increase as sick leave usage increases, which is completely opposite of the results on the remaining country. Marmot's theories suggest that more social participation should indicate a greater feeling of well-being and satisfaction with life in general, and thus be connected with better overall health. A clear explanation as to why this is apparently not the case in the north can unfortunately not be given.

The large gender differences in how sick leave is affected by the *unemployment rate* might become clearer when viewed in connection with a broader look on the variable *bankruptcies*.

Bankruptcies and downsizing might also serve as a proxy for fluctuations in the economy, which then links it with the level of unemployment. As a result, increasing unemployment will have two partially counteracting effects. It will reduce sick leave due to the disciplinary and composition hypotheses, but through a higher number of company bankruptcies it might also increase the number of long sick leave spells, as discussed when reviewing the model. So is it possible to distinguish between these different effects? A key element at least is the differing effects of economic downturns and economic *crises* on sick leave behavior. Nossen (2010a) finds a large increase in sick leave among men, and a moderate increase among women, in the period following the financial crisis that started in the fall of 2008. For the three first quarters of 2009 he finds sick leave to increase by 9.7% for men and 2.9% for women compared to the same period the year before. The increase among men is found to be mainly in long term sickness spells.

As shown in the tables on labor market structure, men are mostly employed in the private sector while women more commonly in the public sector. This might to a large degree explain why sick leave among men was overall more affected by the crisis than among women in Nossens findings.

Thus, a possible explanation for the gender differences in the present paper's findings on the unemployment variable is then that male workers will in general be more vulnerable to these types of economic crises, and are therefore more prone to a positive correlation between sick leave and unemployment. This line of thought does however not explain very well why women experience a negative correlation.

In the estimated regressions, bankruptcies are found to have an overall facilitating effect on high sick leave usage on a national level. This effect however disappears completely in the north. Considering the aforementioned suggestions by Nossen (2010b) on how sick leave is used as a strategic measure when the worker is faced with the risk of unemployment, a natural assumption is then that workers in the north are less prone to this form of illegitimate insurance usage, or at least that the total effect on sick leave by such use is insignificant compared to other factors.

The obtained results on *small companies* are rather unexpected considering that other findings have shown that larger companies should experience more sick leave, while small companies less (Barmby & Stephan, 2000). Individual level theories on working conditions in small versus large companies in general state that smaller companies should have a higher degree of social control, in the form of a closer relationship with supervisors and coworkers (Barmby & Stephan, 2000). One possible explanation might be poorer working environments in many smaller companies, which then facilitates sickness and injuries. A news story published by the Norwegian Labor Inspection Authorithy reveals that as much as half of all fatal accidents in the workplace in recent years have been in companies with less than 10 employees[20]. These are companies that often have less time and resources to spend on improving work environment factors. However, the present findings are much to limited to make any definite statements on this topic, but it would be an interesting area to look at in future research.

The results on public employment show in general a lower level of sick leave in municipalities with a larger public sector when looking at the north. The same results are obtained for women on a national level. This hints towards that the higher level of sick leave in northern Norway can not be explained by the relatively larger share of public employment found there. The results are supported by the findings of Dutriex & Sjöholm (2003), which show a negative correlation between the share of state employees and total sickness insurance usage. The same report however finds a positive correlation between municipality employed and sick leave. This positive relationship is also found by Olsson (2004).

The performed analysis in this Master thesis was to some extent limited by restricted data-access. In addition, it would have been more optimal to use individual level data when studying these forms of causal relationships where individual behavior is the point of interest. However, such data were not obtainable in the scope of this thesis. Further, changes in total sick leave usage is constructed by two parts; changes in average length and changes in average frequency. Numbers on these two would have given a more in-depth depiction of the relationship between sick leave usage and the explanatory factors. The used meaurements for sick leave instead gives a broader picture without such nuances. Optimally, the analysis would have included data on short-term and long-term absences, and the prevalence of different diagnoses across municipalities. This was unfortunately not possible, and the inclusion of these more detailed depictions of sick leave is left as a possible subject for future research.

References

Askildsen, J.E., Bratberg, E. & Nilsen, O.A. (2005). «Unemployment, labor force composition and sickness absence: a panel data study». *Health Economics*, 14, 1087-1101

Baltagi, B.H. (2008). «Econometric Analysis of Panel Data, fourth edition». West Sussex: John Wiley & Sons Ltd

Barmby, T.,Stephan, G. (2000). «Worker absenteeism: Why firm size may matter». *The Manchester School Vol 68 No.5. September 2000*

Berge, S. (2010). "Uendret sykefravær siden 2001". *Samfunnsspeilet* 24, Nr 2, 16-23

Bjørnstad, A. F. (2013). «Utbetalingene av trygdeytelser siste 10 år». *Arbeid og velferd,* nr. 3

Boeri, T. & van Ours, J. (2008). «The Economics of Imperfect Labor Markets». New Jersey: Princeton University Press.

Bovin, B. & Wandall, J. (1989) «Sygedage: fravær blant ansatte i amter og kommuner». Denmark: AKF Forlaget

Brage, S., Nossen, J. P. & Thune, O. (2013). «Flere blir sykmeldt for symptomer og plager». *Arbeid & velferd* nr. 2

Bragstad, T., Regbo, U. & Sagsveen, A. (2006). «Variasjon i sykefravær. En analyse av geografisk variasjon i forbruk av sykepenger i perioden 1993-2004». NAV Rapport 04/2006. Rikstrygdeverket

Bratsberg, B., Fevang, E., Røed, K. (2010). «Disability in the welfare state: An unemployment problem in disguise?». IZA discussion paper no. 4897

Carter Hill, R., Griffiths, W. E. & Lim, G. C. (2012). «Principles of Econometrics, fourth edition». Asia: John Wiley & Sons Ltd

Dahl, S.Å., Hansen, H.T., Olsen, K. (2010). «Sickness absence among immigrants in Norway, 1992-2003». *Acta Sociologica, March 2010, 53:35*-52.

Dutrieux, J., & Sjöholm, K.R. (2003). «Regionala skillnader i sjukskrivning – hur ser de ut och vad beror de på?». RFV analyserar 2003:12

Dutrieux, J., & Sjöholm, K.R. (2003). «Sjuka kommuner? Skillnader i Sjukfrånvaro mellan Sveriges Kommuner år 2000». RFV analyserar 2003:17

Edwards, J. R. & Rothbard, N. P. (1999). «Work and Family Stress and Well-Being: An Examination of Person-Environment Fit in the Work and Family Domains». *Organisational behavior and Human Decision Processes,* Vol 77, No 2, February, 85-129

Erikson, R. (2001). "Why do graduates live longer?". In Cradle to Grave: Life-course Change in Modern Sweden (eds J. O. Jonsson and C. Mills). Durham: Sociology Press.

Foss, L. & Skyberg, K. (2008). "Sykefravær i ulike bransjer: Utvikling i sykefravær of uførhet I lys av individuelle faktorer og forhold ved arbeidsplassen". STAMI-report, Vol.9, No.18

Hauge, L. & Kann, I.C. (2007). «Hva betyr alders- og kjønnssammensetningen for sykefraværet?». *Arbeid og velferd,* nr.1

Helde, I., Kristoffersen, P., Lysø, N., & Thune, O. (2010). "Sykefravær: Store fylkesvise forskjeller i varigheten innenfor samme diagnose." *Arbeid og Velferd* nr. 4, 38-55

Helde, I., Kristoffersen, P. & Lysø, N. (2011). "Kan næringsstruktur forklare fylkesvise forskjeller i sykefraværet?." *Arbeid og Velferd* nr. 3

Hsiao, C. (2003). "Analysis of panel data". Cambridge University Press, Cambridge.

Ihlebaek, C., Brage, S. & Eriksen, H.R. (2007) "Health complaints and sickness absence in Norway, 1996-2003". *Occup Med (Lond),* 57, 43-49

Kann, I. C., Thune, O. & Galaasen, A. M. (2013). “Gir lavere sykefravær færre på langtidsytelser?”. *Arbeid og velferd*, nr. 3

Khan, J., Gerdtham, U.G., Jansson, B. (2004). “Effects of macroeconomic trends on social security spending due to sickness and disability”. Am. J. Public Health, 94(11): 2004–2009

Khan, J. & Rehnberg, C. (2009). “Perceived job security and sickness absence: a study on moral hazard”. *The European Journal of Health Economics*, Vol. 10, Issue 4, 421-428

Kostøl, A. R. (2010). «Kunnskapsoppsummering av samfunnsvitenskapelig litteratur om årsaker til sykefravær og utstøting». Statistics Norway.

Krogsgård, A. (2009). «Regional Variance in Sickness Insurance Usage». Master’s thesis, Jönköping International

Lindbeck, A., Palme, M. & Persson, M. (2008). «Social Interaction and Sickness Absence». CESIFO Working paper No. 2215. Category 3: Social Protection

Manski, C. (1993) “Identification of Endogenous Social Effects: The Reflection Problem”. *The Review of Economic Studies* 60, 531-542.

Markussen, S. (2009) “How physicians can reduce sick leave – evidence from a natural experiment”. *Ragnar Frisch Centre for Economic Research, project 4132*

Markussen, S., Røed, K., Røgeberg, O.J., & Gaure, S. (2009). “The Anatomy of Absenteeism”. *Journal of Health Economics,* volume 30, issue 2, 277-292

Marmot, M. G., Shipley, M. J. & Rose, G. (1984). “Inequalities in death – specific explanations of a general pattern”. *Lancet*, 323, 1003-1006.

Marmot, M. G., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., Feeney, A. & Davey Smith, G. (1991). “Health inequalities among British civil servants: the Whitehall II study”. *Lancet*, 337, 1387-1393.

Marmot, M. G. & Shipley, M. J. (1996). “Do socioeconomic differences in mortality persist after retirement? 25 year follow up of civil servants from the first Whitehall study”. *British medical journal*, 313, 1177-1180.

Marmot, M. G. (2004a). “Status syndrome”. London: Bloomsbury.

Marmot, M. G. (2004b). “Status Syndrome”. *Significance,* volume 1, issue 4, 150-154

Mastekaasa, A. & Dale-Olsen, H. (2000). “Do Women or Men Have the Less Healthy Jobs? An Analysis of Gender Differences in Sickness Absence”. *European Sociological Review*, Vol.16 No.3, 267-286

Nordic Council of Ministers (2007). “Aksepterte årsaker til sykefravær – holdninger i de fem nordiske landene – resultater for Norge”.

Nossen, J.P. (2010a). «Økt legemeldt sykefravær etter finanskrisen: Flere langvarige sykefravær for menn». *Arbeid og velferd*, nr. 1

Nossen, J.P. (2010b). «Hvilken betydning har nedbemanning for sykefravær og andre helserelaterte ytelser?». *Arbeid og velferd*, nr. 4

NOU 2000:27: «Sykefravær og uførepensjonering»

Olsson, S. (2004). «Skillnader mellan kommunernas utgifter för förtidspension plus sjukpenning, 1993-2000».

Olsson, S. (2006). «Attityder och regionala skillnadar i sjukförsakringen». In Palmer, E. (ed.). SKA-projektet: «sjukförsakring, kulturer och attityder: fyra aktörers perspektiv». Stockholm: Forsäkringskassan.

Ose, S. O., Jensberg, H., Eidsmo, R., Suandsund, M. & Dyrstad, J.M. (2006). «Sykefravær: Kunnskapsstatus of problemstillinger». *SINTEF*, A325.

Osterkamp, R. & Röhn, O. (2007). «Being on Sick Leave: Possible Explanations for Differences of Sick-leave Days Across Countries». *CESIFO Economic Studies,* Vol.52, 1/2007, 97-114

Redelmeier, D. A. & Singh, S. M. (2001). "Survival in Academy Award-winning actors and actresses". *Annals of Internal Medicine,* 134, 955-962

Rege, M., Telle, K. & Votruba, M. (2005). "The Effect of Plant Downsizing on Disability Pension Utilization". Statistics Norway Discussion papers No. 435

Røed, K., Fevang, E. (2005). «Organisational Change, Absenteeism and Welfare Dependency». Working paper 2005:9, The Ragnar Frisch Center for Economic Research and HERO.

Shapiro, C. & Stiglitz, J.E. (1984) Equilibrium Unemployment as a Worker Discipline Device. *American Economic Review,* 74,433-444

SOU, Swedish Government Official Reports. (2000). Slutbetänkande av Sjukförsäkringsutredningen SOU (Final report form the sickness insurance Investigation) Stockholm, SOU, 2000: 121.

SOU, Swedish Government Official Reports. (2002). Handlingsplan för ökad hälsa i arbetslivet. Slutbetänkande.

SOU (Implementation plan to enhance health in working life. Final report). Stockholm, SOU, 2002: 5.

Sparebank 1 Nord-Norge (2009). "Economic Survey for Northern-Norway".

Van Rossum, C. T. M., Shipley, M. J., Van de Mheen, H., Grobbee, D. E. & Marmot, M. G. (2000). "Employment grade differences in cause specific mortality. A 25 year follow up of civil servants from the first Whitehall study". *Journal of Epidemiology and Community Health*, 54, 178-184.

Appendix A: Overview of variables used as regression input

Dependent variables	***Translation***
sfpnavm	The sick leave percentage for men in the northern model
sfpnavk	The sick leave percentage for women in the northern model
sfpm	The sick leave percentage for men in the country-wide model
sfpk	The sick leave percentage for women in the country-wide model
arbm2	The share of sick workers among men
arbk2	The share of sick workers among women
Demographic variables	
alder1625m	Share of men aged 16-25
alder1625k	Share of women aged 16-25
alder5566m	Share of men aged 55-66
alder5566k	Share of women aged 55-66
invm	Share of immigrants among men
invk	Share of immigrants among women
	Country-wide model:
_Istrkat_2	Municipalities w/ 20 000-40 000 inhabitants
_Istrkat_3	Municipalities w/ 10 000-20 000 inhabitants
_Istrkat_4	Municipalities w/ 5000-10 000 inhabitants
_Istrkat_5	Municipalities w/ 2000-5000 inhabitants
_Istrkat_6	Municipalities w/ 1000-2000 inhabitants
_Istrkat_7	Municipalities w/ less than 1000 inhabitants
	Northern model:
_strkat2_2	Municipalities w/ 2000-5000 inhabitants
_strkat2_3	Municipalities w/ less than 2000 inhabitants
Social variables	
utdvgs	Level of education
stv	Participation in elections
laguf	Disability benefits usage preceding year in the country-wide model
lagufnavm	Disability benefits usage preceding year for men in the northern model
lagufnavk	Disability benefits usage preceding year for women in the northern model
Labor market variables	
ledm	Unemployment among men
ledk	Unemployment among women
sysprim	Employment in the primary sector for men
sysprik	Employment in the primary sector for women
sysoffm	Employment in the public sector for men
sysoffk	Employment in the public sector for women
syshelm	Employment in the health sector for men
syshelk	Employment in the health sector for women
systjem	Employment in the service industries for men
systjek	Employment in the service industries for women
konk	Bankruptcies
bed1	Small companies

Appendix B: Hausman-Taylor regression output

B.1 Country-wide model results

B.1.1 - The sick leave percentage for men

```
Hausman-Taylor estimation                      Number of obs      =      3850
Group variable: num                            Number of groups   =       385

                                               Obs per group: min =        10
                                                              avg =        10
                                                              max =        10

Random effects u_i ~ i.i.d.                    Wald chi2(19)      =   1002.14
                                               Prob > chi2        =    0.0000
```

sfpm	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
TVexogenous						
ald1625m	-.0051386	.0124292	-0.41	0.679	-.0294994	.0192222
ald5566m	-.0110597	.0102211	-1.08	0.279	-.0310928	.0089734
invm	-.0530543	.0096176	-5.52	0.000	-.0719044	-.0342042
laguf	-.05216	.0172034	-3.03	0.002	-.0858781	-.018442
sysprim	-.0446367	.0072068	-6.19	0.000	-.0587617	-.0305116
sysoffm	.0088249	.0111113	0.79	0.427	-.0129527	.0306026
syshelm	.0118477	.0160445	0.74	0.460	-.0195989	.0432943
systjem	.0147173	.0062684	2.35	0.019	.0024315	.0270031
bed1	.0194408	.0067591	2.88	0.004	.0061932	.0326885
TVendogenous						
utdvgs	-.119939	.0120326	-9.97	0.000	-.1435224	-.0963555
ledm	.1188442	.0137835	8.62	0.000	.091829	.1458593
konk	.1478629	.0285554	5.18	0.000	.0918954	.2038304
TIexogenous						
_Istrkat_2	-.000732	.0035624	-0.21	0.837	-.0077142	.0062502
_Istrkat_3	.0018993	.0032861	0.58	0.563	-.0045413	.00834
_Istrkat_4	.0004254	.0032641	0.13	0.896	-.0059722	.0068229
_Istrkat_5	.0008261	.0032877	0.25	0.802	-.0056177	.0072698
_Istrkat_6	.0006185	.0036219	0.17	0.864	-.0064804	.0077174
_Istrkat_7	-.0104841	.0044442	-2.36	0.018	-.0191946	-.0017736
TIendogenous						
stv	-.0013081	.0005183	-2.52	0.012	-.002324	-.0002922
_cons	.2352223	.0383076	6.14	0.000	.1601408	.3103038
sigma_u	.01059681					
sigma_e	.00720834					
rho	.68365667	(fraction of variance due to u_i)				

B.1.2 - The sick leave percentage for women

```
Hausman-Taylor estimation                       Number of obs      =      3850
Group variable: num                             Number of groups   =       385

                                                Obs per group: min =        10
                                                               avg =        10
                                                               max =        10

Random effects u_i ~ i.i.d.                     Wald chi2(19)      =    526.52
                                                Prob > chi2        =    0.0000
```

sfpk	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
TVexogenous						
ald1625k	-.0522736	.0157087	-3.33	0.001	-.0830621	-.0214851
ald5566k	-.0619866	.013622	-4.55	0.000	-.0886853	-.035288
invk	.002611	.015913	0.16	0.870	-.0285778	.0337999
laguf	-.082857	.023663	-3.50	0.000	-.1292357	-.0364782
sysprik	-.0264022	.0160353	-1.65	0.100	-.0578307	.0050263
sysoffk	-.0244956	.0120259	-2.04	0.042	-.048066	-.0009253
syshelk	.0085566	.0099968	0.86	0.392	-.0110368	.02815
systjek	.0106325	.0103384	1.03	0.304	-.0096304	.0308954
bed1	.0422684	.0088255	4.79	0.000	.0249707	.0595661
TVendogenous						
utdvgs	-.1522262	.0151462	-10.05	0.000	-.1819122	-.1225402
ledk	-.1877096	.0229818	-8.17	0.000	-.232753	-.1426661
konk	.2462974	.0358968	6.86	0.000	.1759408	.3166539
TIexogenous						
_Istrkat_2	.0014821	.0049949	0.30	0.767	-.0083077	.011272
_Istrkat_3	.0008107	.0045848	0.18	0.860	-.0081753	.0097967
_Istrkat_4	-.0044761	.0045461	-0.98	0.325	-.0133863	.0044341
_Istrkat_5	-.0076682	.0044858	-1.71	0.087	-.0164603	.0011239
_Istrkat_6	-.0109936	.0048499	-2.27	0.023	-.0204992	-.001488
_Istrkat_7	-.0194235	.0058116	-3.34	0.001	-.0308142	-.0080329
TIendogenous						
stv	-.0020173	.0008333	-2.42	0.015	-.0036506	-.000384
_cons	.362608	.0626104	5.79	0.000	.239894	.4853221
sigma_u	.01490907					
sigma_e	.009269					
rho	.72123363	(fraction of variance due to u_i)				

B.1.3 - The share of sick workers among men

```
Hausman-Taylor estimation                       Number of obs      =      3850
Group variable: num                             Number of groups   =       385

                                                Obs per group: min =        10
                                                               avg =        10
                                                               max =        10

Random effects u_i ~ i.i.d.                     Wald chi2(19)      =    364.88
                                                Prob > chi2        =    0.0000
```

arbm2	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
TVexogenous						
ald1625m	-.0049622	.0112482	-0.44	0.659	-.0270082	.0170838
ald5566m	.0136393	.0093079	1.47	0.143	-.0046038	.0318824
invm	-.0455807	.0087111	-5.23	0.000	-.0626542	-.0285072
laguf	-.0251735	.015686	-1.60	0.109	-.0559176	.0055706
sysprim	-.0321829	.0066485	-4.84	0.000	-.0452137	-.019152
sysoffm	.0028055	.0102521	0.27	0.784	-.0172883	.0228993
syshelm	.0150775	.0146392	1.03	0.303	-.0136148	.0437699
systjem	.0197269	.0057943	3.40	0.001	.0083703	.0310835
bed1	.0165234	.0061668	2.68	0.007	.0044368	.02861
TVendogenous						
utdvgs	-.0607319	.0109537	-5.54	0.000	-.0822007	-.0392631
ledm	.0778722	.0124332	6.26	0.000	.0535036	.1022408
konk	.0932403	.0257396	3.62	0.000	.0427916	.143689
TIexogenous						
_Istrkat_2	.0001294	.0034146	0.04	0.970	-.006563	.0068218
_Istrkat_3	.0031321	.0031461	1.00	0.319	-.0030342	.0092984
_Istrkat_4	.0025039	.0031215	0.80	0.422	-.0036141	.008622
_Istrkat_5	.002874	.0031354	0.92	0.359	-.0032712	.0090191
_Istrkat_6	.0025236	.0034434	0.73	0.464	-.0042253	.0092726
_Istrkat_7	-.00499	.0042107	-1.19	0.236	-.0132429	.0032628
TIendogenous						
stv	-.0019305	.000481	-4.01	0.000	-.0028731	-.0009878
_cons	.2247416	.035631	6.31	0.000	.1549062	.2945771
sigma_u	.01020886					
sigma_e	.00650879					
rho	.71099144	(fraction of variance due to u_i)				

B.1.4 - The share of sick workers among women

```
Hausman-Taylor estimation                      Number of obs      =      3850
Group variable: num                            Number of groups   =       385

                                               Obs per group: min =        10
                                                              avg =        10
                                                              max =        10

Random effects u_i ~ i.i.d.                    Wald chi2(19)      =    251.82
                                               Prob > chi2        =    0.0000
```

arbk2	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
TVexogenous						
ald1625k	-.0258333	.0147508	-1.75	0.080	-.0547443	.0030777
ald5566k	-.0068539	.0127929	-0.54	0.592	-.0319276	.0182197
invk	.004572	.0149713	0.31	0.760	-.0247714	.0339153
laguf	-.0179281	.0222316	-0.81	0.420	-.0615013	.025645
sysprik	-.0313522	.0151299	-2.07	0.038	-.0610063	-.0016981
sysoffk	.0025539	.0112998	0.23	0.821	-.0195933	.024701
syshelk	.0210508	.0094092	2.24	0.025	.0026091	.0394924
systjek	.0068895	.009738	0.71	0.479	-.0121967	.0259756
bed1	.0284774	.0083113	3.43	0.001	.0121874	.0447673
TVendogenous						
utdvgs	-.0382085	.0142238	-2.69	0.007	-.0660866	-.0103304
ledk	-.2254375	.0215367	-10.47	0.000	-.2676487	-.1832263
konk	.1250387	.0336261	3.72	0.000	.0591327	.1909446
TIexogenous						
_Istrkat_2	.0027287	.0048055	0.57	0.570	-.0066898	.0121473
_Istrkat_3	.0030842	.0044091	0.70	0.484	-.0055576	.0117259
_Istrkat_4	-.0019886	.0043706	-0.45	0.649	-.0105548	.0065776
_Istrkat_5	-.0043083	.0043095	-1.00	0.317	-.0127547	.0041381
_Istrkat_6	-.007162	.004656	-1.54	0.124	-.0162876	.0019636
_Istrkat_7	-.0088558	.0055731	-1.59	0.112	-.0197789	.0020673
TIendogenous						
stv	-.0035182	.0007914	-4.45	0.000	-.0050693	-.001967
_cons	.3661819	.0595071	6.15	0.000	.2495501	.4828137
sigma_u	.01436083					
sigma_e	.00868213					
rho	.73232965	(fraction of variance due to u_i)				

B.2 Troms & Finnmark results

B.2.1 - The sick leave percentage for men

```
Hausman-Taylor estimation                        Number of obs      =       430
Group variable: num                              Number of groups   =        43

                                                 Obs per group: min =        10
                                                                avg =        10
                                                                max =        10

Random effects u_i ~ i.i.d.                      Wald chi2(15)      =     93.35
                                                 Prob > chi2        =    0.0000
```

sfpnavm	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
TVexogenous						
ald1625m	.0145376	.0409599	0.35	0.723	-.0657422	.0948175
ald5566m	-.1651251	.0325006	-5.08	0.000	-.228825	-.1014252
invm	-.0542257	.0415922	-1.30	0.192	-.135745	.0272935
lagufnavm	.0635527	.0464132	1.37	0.171	-.0274156	.154521
sysprim	.0081491	.0289478	0.28	0.778	-.0485875	.0648857
sysoffm	-.0586817	.0248399	-2.36	0.018	-.1073669	-.0099965
syshelm	-.0102515	.0419821	-0.24	0.807	-.092535	.0720319
systjem	-.025102	.0249664	-1.01	0.315	-.0740352	.0238312
bed1	.0507624	.0212627	2.39	0.017	.0090883	.0924365
TVendogenous						
utdvgs	-.0699442	.0364191	-1.92	0.055	-.1413242	.0014358
ledm	-.069945	.0429204	-1.63	0.103	-.1540675	.0141775
konk	.134706	.0827723	1.63	0.104	-.0275248	.2969367
TIexogenous						
_Istrkat2_2	.0027123	.0056934	0.48	0.634	-.0084464	.0138711
_Istrkat2_3	-.0021756	.0081012	-0.27	0.788	-.0180537	.0137024
TIendogenous						
stv	.0023701	.0013562	1.75	0.081	-.0002879	.0050282
_cons	-.0352212	.0974293	-0.36	0.718	-.2261791	.1557366
sigma_u	.01132137					
sigma_e	.00998426					
rho	.56251246	(fraction of variance due to u_i)				

B.2.2 - The sick leave percentage for women

```
Hausman-Taylor estimation                    Number of obs      =       430
Group variable: num                          Number of groups   =        43

                                             Obs per group: min =        10
                                                            avg =        10
                                                            max =        10

Random effects u_i ~ i.i.d.                  Wald chi2(15)      =     60.50
                                             Prob > chi2        =    0.0000
```

sfpnavk	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
TVexogenous						
ald1625k	.1208724	.0553429	2.18	0.029	.0124022	.2293425
ald5566k	-.1043066	.0464004	-2.25	0.025	-.1952497	-.0133635
invk	-.028942	.0556727	-0.52	0.603	-.1380584	.0801744
lagufnavk	-.0199907	.0650472	-0.31	0.759	-.1474808	.1074995
sysprik	-.0711788	.0545361	-1.31	0.192	-.1780677	.03571
sysoffk	-.097192	.0460831	-2.11	0.035	-.1875132	-.0068709
syshelk	-.0411243	.0331082	-1.24	0.214	-.1060152	.0237666
systjek	-.0594187	.0388241	-1.53	0.126	-.1355125	.0166751
bed1	-.0023921	.0278189	-0.09	0.931	-.0569161	.0521318
TVendogenous						
utdvgs	-.150824	.0451633	-3.34	0.001	-.2393425	-.0623056
ledk	-.1866028	.073033	-2.56	0.011	-.3297449	-.0434606
konk	.0856405	.1036821	0.83	0.409	-.1175727	.2888537
TIexogenous						
_Istrkat2_2	-.0065171	.0088966	-0.73	0.464	-.0239541	.0109199
_Istrkat2_3	-.0196395	.0111058	-1.77	0.077	-.0414065	.0021275
TIendogenous						
stv	.0035265	.0019293	1.83	0.068	-.0002548	.0073077
_cons	-.0054262	.1326015	-0.04	0.967	-.2653204	.254468
sigma_u	.01873956					
sigma_e	.01262958					
rho	.68765742	(fraction of variance due to u_i)				

B.2.3 - The share of sick workers among men

```
Hausman-Taylor estimation                       Number of obs      =       430
Group variable: num                             Number of groups   =        43

                                                Obs per group: min =        10
                                                               avg =        10
                                                               max =        10

Random effects u_i ~ i.i.d.                     Wald chi2(15)      =     42.92
                                                Prob > chi2        =    0.0002
```

arbm2	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
TVexogenous						
ald1625m	-.0039011	.0381876	-0.10	0.919	-.0787474	.0709453
ald5566m	-.0990323	.0301799	-3.28	0.001	-.1581838	-.0398809
invm	-.0513132	.0386778	-1.33	0.185	-.1271203	.0244938
lagufnavm	.0594846	.0430677	1.38	0.167	-.0249264	.1438957
sysprim	.0307853	.0269692	1.14	0.254	-.0220732	.0836439
sysoffm	-.0612261	.0229136	-2.67	0.008	-.106136	-.0163162
syshelm	-.0459561	.0390715	-1.18	0.240	-.1225347	.0306226
systjem	-.0175936	.0232703	-0.76	0.450	-.0632026	.0280153
bed1	.0416464	.0198085	2.10	0.036	.0028225	.0804703
TVendogenous						
utdvgs	-.0165307	.0337981	-0.49	0.625	-.0827737	.0497122
ledm	-.0641703	.0401012	-1.60	0.110	-.1427671	.0144266
konk	.0350831	.077391	0.45	0.650	-.1166005	.1867668
TIexogenous						
_Istrkat2_2	.0029078	.0052094	0.56	0.577	-.0073024	.0131181
_Istrkat2_3	-.0044307	.007445	-0.60	0.552	-.0190226	.0101612
TIendogenous						
stv	.0028173	.0012477	2.26	0.024	.0003719	.0052627
_cons	-.1201052	.0897916	-1.34	0.181	-.2960935	.055883
sigma_u	.01029871					
sigma_e	.00933053					
rho	.54920367	(fraction of variance due to u_i)				

B.2.4 - The share of sick workers among women

```
Hausman-Taylor estimation                    Number of obs      =       430
Group variable: num                          Number of groups   =        43

                                             Obs per group: min =        10
                                                            avg =        10
                                                            max =        10

Random effects u_i ~ i.i.d.                  Wald chi2(15)      =     17.12
                                             Prob > chi2        =    0.3116
```

arbk2	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
TVexogenous						
ald1625k	.0776333	.0512525	1.51	0.130	-.0228198	.1780864
ald5566k	.0195177	.0419863	0.46	0.642	-.0627738	.1018093
invk	.0063954	.050427	0.13	0.899	-.0924397	.1052304
lagufnavk	-.0003354	.0586387	-0.01	0.995	-.1152652	.1145944
sysprik	-.0572792	.0508292	-1.13	0.260	-.1569026	.0423443
sysoffk	-.0208718	.0423889	-0.49	0.622	-.1039525	.0622089
syshelk	.0038842	.0309015	0.13	0.900	-.0566815	.06445
systjek	-.0001892	.0357525	-0.01	0.996	-.0702628	.0698844
bed1	-.0182119	.0258333	-0.70	0.481	-.0688443	.0324205
TVendogenous						
utdvgs	-.0436744	.0399021	-1.09	0.274	-.121881	.0345323
ledk	-.1518581	.0687653	-2.21	0.027	-.2866356	-.0170806
konk	.1088041	.097933	1.11	0.267	-.0831411	.3007493
TIexogenous						
_Istrkat2_2	-.0021374	.007242	-0.30	0.768	-.0163314	.0120567
_Istrkat2_3	-.0103477	.0092419	-1.12	0.263	-.0284615	.0077662
TIendogenous						
stv	.0024094	.0016048	1.50	0.133	-.0007359	.0055548
_cons	-.0708679	.1095813	-0.65	0.518	-.2856433	.1439076
sigma_u	.0148447					
sigma_e	.011944					
rho	.60702582	(fraction of variance due to u_i)				

Endnotes

[1] More specific, per first of January 2013, 43 out of a total of 428 municipalities are located in Troms and Finnmark

[2] Northern-Norway as defined here means the three counties Nordland, Troms and Finnmark. The analyses performed in the present paper however focus on only the counties Troms and Finnmark, and is referring to these two when talking about «the north», unless specified otherwise.

[3] See footnote 2

[4] Lovdata, www.lovdata.no. A private foundation established in 1981 by the Norwegian ministry of Justice and the law faculty in Oslo.

[5] Agreement on an Inclusive Labor Market (IA). An agreement between the government, trade union and employers' organization aimed at reducing sick leave in Norway through improving the working environment and preventing workers from dropping out.

[6] For a more in-depth discussion of the neoclassical theory of individual labor supply, see for example Boeri & van Ours (2008).

[7] Source: Statistics Norway. «Innvandrere og norskfødte med innvandrerforeldre, 1.januar 2014». Published 24.04.14. Last checked 03.05.14. (http://www.ssb.no/befolkning/statistikker/innvbef).

[8] Obtained data were orignially on a quarterly form. The used input data has been calculated by averaging out the percentages over the four quarters to obtain yearly data.

[9] There are two main differences: (1) The numbers from Statistics Norway include both self-certified and physician-certified sick leave, while the numbers from NAV only include physician-certified sick leave. (2) Statistics Norway divides reported sick leave between municipalities based on the insureds registered home adress, while NAV divides it by NAV office where the case is registered.

[10] Stabas -> (00) General -> «Classification of municipalities by population size 1998» (http://stabas.ssb.no/ItemsFrames.asp?ID=8104002&Language=en&VersionLevel=ClassVersion)

[11] In fact, Tromsø is the only municipality in these two counties that has more than 40 000 inhabitants, while Nesseby is the only one with less than 1000 inhabitants

[12] Folketrygdloven Del IV. Ytelser ved sykdom m.m. § 12-4

[13] More precisely, 11/12 of the 18 year olds and 1/12 of the 67 year olds are counted. A person can at the earliest start receiving disability benefits a month after his/her eighteenth birthday, and he/she can at the latest transfer to retirement pension a month after his/her sixtyseventh birthday. This is in accordance with the general rule (See: Folketrygdloven Del VII. Forvaltningsmessige bestemmelser. § 22-12) that social benefits paid on a monthly basis should have its first disbursement one month after the individual meets the conditions needed for having a right to the benefit.

[14] Based on Hsiao (2003), as reviewed by Baltagi (2012)

[15] For a more in-depth discussion of fixed-effects vs. random effects, see for example ch.15 of *Priciples of Econometrics* by Carter Hill, Griffiths & Lim (2012)

[16] For a thorough walkthrough of the Lagrange-Multiplier test, see ch.4, p.63-72, in *Econometric Analysis of Panel Data* by Baltagi (2008)

[17] See for example ch.7, p. 133-136, in Baltagi (2012) for more information on the Hausman-Taylor estimator

[18] As seen in the descriptive statistics, both the gender separated percentages in Troms and Finnmark are higher than the aggregated percentage for the rest of the country. In the north it is 13.02% for men and 16.38% for women, while in the rest of the country it is 10.94% for the genders combined.

[19] YR, (http://m.yr.no/nyheter/distrikt/nordland/1.5939595). Published 09.06.2008, Last checked 28.05.2014.

[20] Mehli, H. «Farligere på jobb i små bedrifter». (http://www.arbeidstilsynet.no/nyhet.html?tid=237992) Published 07.02.2013. Last checked 30.05.2014.