Stockholms universitet

The Impact of Work Organisation on the Employee's Future on the Labour Market

Martina Aksberg

Masteruppsats i matematisk statistik
Master Thesis in Mathematical Statistics

Masteruppsats 2012:7
Matematisk statistik
September 2012

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Matematisk statistik
Stockholms universitet
Masteruppsats **2012:7**,
http://www.math.su.se/matstat

The Impact of Work Organisation on the Employee's Future on the Labour Market

Martina Aksberg*

september 2012

Sammanfattning

In this paper we explore the relationship between work organisation and the employee's future on the labour market. Based on the results from a survey performed in 1998, the work organisation at a sample of Swedish workplaces is observed. The roughly 230000 employees are followed over the years 1999-2008 with respect to their positions on the labour market. Logistic regression and generalized estimating equations are used to find possible connections between work organisation and the probability of having a job, being employed by the same firm, employed by another firm, unemployed, on sick leave, disability pensioner or having no or a very low income. By the aid of multiple linear regression the corresponding analysis is performed with income development in focus. The main conclusion that may be drawn from our study is that work organisation does have an impact on the employee's future on the labour market.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige. E-post: ma.aks@hotmail.com. Handledare: Tom Britton.

Preface

This paper constitutes a Master thesis in Mathematical Statistics. It has been performed at Statistics Sweden (SCB), at the department of National Accounts.

First of all I would like to give my sincerest thanks to PhD in Economics Hans-Olof Hagén, Senior Advisor at Statistics Sweden, for guiding me through this project. Thank you for all your time, help and good advice and for always being such an encouraging and welcoming person.

I would also like to thank my supervisor at Stockholm University, Tom Britton, Professor in Mathematical Statistics, for useful discussions and comments.

Contents

1 Introduction

Work organisation is a broad concept, describing how the work at a company is organised and managed. How and by whom the work tasks should be performed, to which extent the individual employees have an impact on their daily work and work situation in general, the possibilities of personal development and learning for the employees, the company's way of handling fluctuating need for labour, the hierarchical levels in the company, among other things, are all determined by the work organisation. It is obvious that the work organisation is of great importance in the daily life of the employees. In this thesis, however, we want to investigate the impact of working conditions in a broader perspective. Does the work organisation have an influence not only on the work situation and well-being of the employees today, but also on the future life of the individuals?

Our goal with this study is to examine whether it is possible to find connections between the work organisation of the firm in which someone is employed, and his or her future position on the labour market. Does the fact that one has worked under certain kinds of working conditions increase the probability of still being in work in the years immediately after? Furthermore, if that is the case, we are also interested in investigating the subgroups of those who do and do not have a job. That is, what characterises the work organisation of companies in which people tend to stay, and is work organisation one of the factors that seem to lead people into finding new jobs? Among those who do not work any longer, possible connections between working conditions and the probability of unemployment, sick leave, disability pension or having no or a very low income are of interest. Finally, we aim to answer the question whether the work organisation has an impact on the future wage development of the employees.

We start this paper by introducing the data upon which our study is based. This section includes both general information on the raw data and more specific information on the datasets and variables created for our purposes. Particularly, a more precise definition of how the concept work organisation is quantified and measured is given in this chapter. After this we further describe the goal of our study and how it can be achieved with the data at hand. An overview of the statistical models used follows immediately after. We continue by presenting and analysing our results and, thereafter, a discussion on some aspects of statistical models follows. Finally we summarize our results and try to draw some general conclusions from our study.

2 Description of data

Our analyses are mainly based on two sources of data – the FLEX 2-survey and the LISA database, which are described briefly in the two subsequent sections. Thereafter we proceed by introducing the datasets and variables used in our study, and show how they are created.

2.1 The FLEX 2-survey

FLEX 2 is an extensive survey, performed in 1998 by Swedish National Board for Technical and Industrial Development (NUTEK), which aimed to make researches into the characteristics of Swedish workplaces. It covers a wide range of topics such as financial responsibility, work organisation, training of employees, wage setting processes, strategies of development and profitability, markets in which the companies are operating, customers, service and product development, co-operation, personnel and use of information technology. A large number of analyses based on the results have been made, mainly published in *Enterprises in Transition – Learning Strategies for Increased Competitiveness* (Institutet för tillväxtpolitiska studier [ITPS], 2001), that focuses on describing the learning strategies in Swedish companies and on finding connections between learning strategies and profit and productivity.

The survey was not directed to enterprises but to workplaces. A workplace is defined as "a geographically limited place where a permanent activity is carried on" whereas "an enterprise can consist of several workplaces at different places" (ITPS, 2001, pp. 92). The workplaces included in the sample group were chosen by stratified random sampling from Statistics Sweden's central workplace and company database covering the first quarter of 1998. Strata were formed with respect to business sector and number of employees and in total the gross sample consisted of 5681 workplaces with at least five employees, which was estimated to be representative for about 80000 Swedish workplaces in trade and industry. First telephone interviews were performed and thereafter a postal questionnaire with additional questions was sent out. More detailed information on the data collection process may be found in (ITPS, 2001, pp. 92-100) and the entire questionnaires are presented in (ITPS, 2001, pp. 101-119).

2.2 The LISA Database

The LISA database is a longitudinal database of the Swedish population. It includes yearly information as from 1990 on all individuals, nationally registered in Sweden, of at least 16 years of age in each year respectively. LISA is an acronym for Swedish *Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier*, with the English translation *Integrated database for labour market research*. The database is built upon a large number of population registers and includes hundreds of variables, principally describing the demographics, education and sources of income of the individuals. A complete description (in Swedish) of the database and the variables included is found in (Statistiska centralbyrån [SCB], 2011:4).

2.3 Our Datasets

As previously mentioned, the goal of our study is to draw conclusions about whether or not and possibly in which way the work organisation of the company at which someone is employed has an impact on his or her future situation on the labour market. To explore this, we need information on how at least a number of workplaces were organised at some certain point of time and data that makes it possible to follow the employees of these workplaces over the years. The FLEX 2 survey provides the former kind of information, while the LISA database includes data that enables us to draw conclusions about the employees' later situations on the labour market as well as other relevant information about them.

To be able to explore the work organisation of a certain workplace, with the results from the FLEX 2 survey as a basis, it is needed that all questions relevant for describing the work organisation have been answered. Imposing this requirement leaves us with a sample of 1311 workplaces, representing 997 different firms. We need to keep in mind that this systematic selection of workplaces limits the opportunity of generalizing the results. On the other hand, although we cannot regard our sample as representative, there is no evident reason to believe that the response group should differ widely from the non-response group with respect to the questions of interest for our study. It is possible and probable that certain types of work organisations are under- or overrepresented in the sample, but intuitively it is unlikely that a certain type of organisation would affect the employees of the non-responding firms in a completely different way than the employees of the responding ones. Therefore, in our point of view, we should be able to regard our results as a good indication of the situation on the labour market in general, although such a generalization of course needs to be made with caution and should be seen as highly approximate.

To start with, the employees of the 1311 workplaces that answered the questions relevant for our study are sorted out. That is, we create a list of all individuals who worked at these workplaces in 1998. This is done by a matching of both the workplace number (*CfarNr* in the LISA database) and the corporate identification number (*PeOrgNr* in the LISA database) of the companies that answered all relevant questions of the FLEX 2 survey, with the LISA database of 1998. Since we are only interested in drawing conclusions about people who are likely to having been employed by, and with that affected by the work organisation of, the FLEX 2 companies during at least fairly long periods of time, we additionally require that the individual should have had a declared income from work of at least 70000 SEK in 1998 to be included in the dataset. It may be discussed which income limit is appropriate to use, especially since wages differ widely among people, which for a given level of income makes the time spent at work different for different people. However, in our opinion, this amount is large enough to exclude most of those who worked only occasionally or during very short periods of time. Worth mentioning is that the average income in the resulting group was 245393 SEK and that the 25 percent quantile was 184800 SEK which indicates that the majority of the employees in the sample are far from this income limit and presumably have had work as their main source of income during the year.

The above procedure yields a list of in total 229387 persons who were employed at the workplaces that answered the survey in 1998. This fixed group of individuals is followed over the years 1999-2008. As time goes by, several people die or move abroad permanently and are therefore no longer nationally registered in Sweden (and thereby no longer included in the LISA database, nor relevant for our purposes). This makes the sample size decrease over the years. Furthermore, we wish to base our analyses only on people who are expected to take part on the Swedish labour market and who are entitled to public allowances, and not on those who are outside the labour market due to age. Therefore, individuals are excluded from our study the year they reach the age of 65, regardless of whether they work or not. The number of individuals left in the sample in each of the years is presented in Table 1.

Table 1: Number of individuals left in the sample in each of the years 1999-2008

Year	Number of observations
1999	227 457
2000	225 414
2001	223 289
2002	220 790
2003	217 847
2004	214 054
2005	209 656
2006	205 159
2007	200 323
2008	194 965

2.4 Variable Description

In the following section we present the response and explanatory variables included in our analyses and describe how they are created.

Response variables

For each year, 1999-2008, a division into the following eight different categories of occupation is made.

1. Employed within the same firm.
2. Employed within another firm.
3. Unemployed.
4. On sick leave.
5. Disability pensioner.
6. Student.
7. Other, low income.
8. Other, higher income.

The data used for the categorisation is collected from the LISA database. A specification of the variables upon which each category is based is found in table A1 in the appendix.

To start with, the individuals employed by the same firm as in 1998 are sorted out and placed in category one. This is done by a matching of the corporate identification numbers, CID. Changes of the CID from year to year, and mergers or expansions of firms, are accounted for by the use of the FAD database (acronym for "Företagens och Arbetsställenas Dynamik"), which we choose to not describe any further in this report. This means that in each year respectively the individuals who are registered at any of the corporate identification numbers that may be traced back to the company of 1998, are categorised as still employed by the same firm.

The requirement for falling into the second category is being registered at another firm with a declared yearly income above a certain level. We choose an income limit corresponding to 70000 SEK in 1998, approximately corrected for the inflation by a yearly increase of 2%, i.e.

$$L_i = 70000 \times 1.02^i,$$

where $L_i =$ income limit year i, $i = 1, 2, \dots 10$ corresponding to the years $1999 - 2008$. The rest of the categories are tested for in subsequent order, so that each person takes part only in one category. The unemployed are defined as those with unemployment benefits making up at least one third of the yearly total income (i.e. one third of the sum of unemployment benefits and declared income). Included in the fourth category are those with at least 61 days of sick leave during the year, and being categorised as disability pensioner only requires the presence of disability pension. We define students as those taking part of student benefits and with a total income (i.e. the sum of student benefits and declared income) on at least the maximum level of student benefits the current year. The individuals not defined as any of the above are included in category seven if their declared yearly income is below L_i and in eight if it is higher than L_i. Finally, we define the category non job as those who do not belong to neither of the categories one or two.

To be able to investigate whether there is a connection between wage development and work organisation we calculate the percentage income change since 1998 for each year 1999-2008 and individual as

$$Income\_percent\_x\_1998 = \frac{Income_x - Income_{1998}}{Income_{1998}} * 100,$$

where x denotes the years 1999-2008. This is based on the variable *DekLon* from the LISA database.

Explanatory variables

In our analyses we distinguish between two types of explanatory variables – index variables based on the results from the FLEX2-survey and register-based variables describing the individuals. The first category is of main interest while the other category should be seen as a collection of variables to be controlled for.

In short, the register-based variables chosen are: gender, age-group, level of education, industry in 1998, regional group and Swedish or foreign background. These characteristics are known to generally influence people's chances on the labour market. The variable regional group specifies whether the individual lives in an area close to Stockholm, another large town, a small town et cetera. In total it has six categories. With industry we mean the line of business in which the company operates. Twelve rough categories are used, for example energy, trade, manufacturing and transport. Having in mind that our focus in this report does not lie on analysing this variable, we choose not to specify the categories used since each one of them is based on a large number of different very specific smaller lines of business, which would make such an exposition very extensive.

While gender, ethnic background and industry of the company in which the individual worked in 1998 are of course fixed variables, the regional group in which the individual lives, the level of education and the age-group to which the individual belongs may differ from year to year. A specification of the variables, their possible values (categories) and the LISA variables used to create them is found in Table 2 below.

Table 2: Specification of the register-based explanatory variables.

Variable	Categories	Variable from the LISA database
Gender	Male / Female	Kon
Age	-35 / 36-49 / 50-	Alder
Education level	Compulsory school / upper secondary school / post secondary education less than 3 years / university 3 years or more	Sun2000niva_old
Industry in 1998	12 different industries	SNI* – svensk näringsgrensindelning
Foreign background	Individual born outside Sweden or individual born in Sweden with both parents born abroad / individual born in Sweden with at least one parent born in Sweden	UtlSvBakgAlt
Regional group	6 different possibilities	Kommun

**The variable is collected from a Swedish company database and not the LISA database. For further information on SNI (in Swedish), see (Statistiska centralbyrån [SCB], n.d.).*

Dummy variables are used for the coding of the different categories. That is, e.g. the variable *education level* does actually consist of four different variables, of which for each person and year one takes the value one and the other three take the value zero (depending on his or her level of education).

The work organisation of the workplaces at which the individuals were employed in 1998 is measured by three index variables - numerical flexibility, decentralisation and individual learning. These concepts are well-known in organisational and economic theory and are described theoretically in e.g. (Atkinson, 1984, pp. 28-31). Our goal is to catch the meaning of these variables with the data at hand. An intuitive interpretation of numerical flexibility is that it aims to describe to which extent the company adjusts the number of employees according to the current need for labour. Decentralisation gives an indication of the level of responsibility, decision-making possibilities and co-operation among the employees. As the name suggests, individual learning describes the employees' learning and professional development in daily work.

Table 3 presents the questions upon which the indices are built. For exact formulations of the answer choices we refer to (ITPS, 2001, pp. 101-119), where the complete questionnaires may be found in their original layout. The indices are, with some minor exception, based on the same questions and data as the ones used in (Statistiska centralbyrån [SCB], 2011, pp. 235-272).

Table 3: Questions included in the indices. T denotes that the question was asked in the telephone interview and Q that it was included in the postal questionnaire.

Numerical flexibility	T22: Is the everyday/normal work in direct production organised so that a) The employee alternates between a number of different working tasks/operations b) The employee is continuously assigned new working tasks Q19: Did you employ personnel on a temporary basis in 1997? Q20: If you employed personnel on a temporary basis in 1997, what was the proportion of this type of personnel? Q22: Did you adjust working hours to business cycles in 1997? By adjusting to business cycles we mean adjusting working hours to peaks and slumps in the market. Q24: Did you use the services of other enterprises in order to pursue your core business in 1997? If so, state which type of personnel that was used and for what reasons. Q25: If you used the services of other enterprises to pursue your core business in 1997, how large was this input compared to your own labour force?
Decentralisation	T13: Is the whole workplace organised so that people with different professional functions carry out work together? a) In production of services and goods

	b) In planning of the work c) In follow-ups of the results and quality control d) In selection of production technology e) In service and product development T17: Which of the personnel normally carry out the following tasks in direct production? a) Daily planning of one's own work b) Weekly planning of one's own work c) Quality control d) Follow-up of results e) Planning and training related to work f) Introduction/training of employees g) Personnel administration h) Service and product development i) Selection of production technology j) Maintenance of technical equipment, e.g. machines and computers Q21: What proportion of the personnel had the following forms of working hours in 1997? a) Fixed working hours b) Flexible working hours c) Free disposition of working hours
Individual learning	T20: Does the everyday work contain elements of organised skills development? T22: Is the everyday/normal work in direct production organised so that: c) The supervisor continuously makes higher demands in respect of existing working tasks. d) The employee himself further develops existing working tasks e) The employee himself develops new working tasks T23: What proportion of the employees in direct production participated in training/courses which were wholly or partly paid by the employer in 1997?

The answers are assigned numerical values and thereafter the indices are calculated as weighted sums of each question. The questions included in the numerical flexibility and individual learning indices are used in a direct way. That is, a positive answer or a relatively large percentage increases the value of the index. In the decentralisation index we aim to catch the degree of team work by question T13, but also by noting whether any of the tasks mentioned in T17 are performed by work teams. The employee's possibility of planning his or her own work and participating in the quality control and follow-up of results is also captured by T17, but in this case if the tasks a)-d) are reported to be performed by the individual employee. Also for decentralisation, a positive answer increases the value of the index. A precise description on how the indices are calculated is found in the appendix.

3 Problem Formulation

The aim of our study is to investigate whether the three index variables, describing the work organisation of the workplaces in 1998, have an impact on the employees' probability of having a job, still being employed by the same firm, being employed by another firm, being unemployed, on sick leave, early retired or having no or a very low income in the years 1999-2008. Furthermore, we are not only interested in whether or not the work organisation influences the future lives of the employees but also in which direction, and how the impact of the organisational variables changes with time. That is, do numerical flexibility, decentralisation and individual learning respectively increase or decrease the probability of being part of each of the categories mentioned above and does this possible impact of the work organisation change over the years 1999-2008? Finally, we want to make the corresponding analysis of the impact of work organisation on the income development.

4 Statistical Models

In this chapter we describe the statistical models used to analyse the data. We focus on intuitive descriptions that aim to simplify the understanding of the results later presented. Discussion of, and motivation for, our choices of methods and models is left to chapter 6, *Discussion of the Statistical models*. Theoretical details are either referred to or presented in the appendix.

To sum up, the statistical models used in our study are multiple logistic regression, ordinary least squares (OLS) multiple regression and generalized estimating equations. The former two are used to make analyses of the impact of the organisational variables on each year separately. The generalized estimating equations approach makes it possible to perform longitudinal analyses in which all years are studied simultaneously.

4.1 Logistic Regression

Our first approach is to perform logistic regressions on each year, 1999-2008, separately. For each response variable and year, models of the type

$$log\left(\frac{p}{1-p}\right) = \alpha + \beta_1 x_1 + \cdots + \beta_n x_n + \beta_{NF} x_{NF} + \beta_D x_D + \beta_{IL} x_{IL} + \varepsilon \quad \mathbf{(1)}$$

are estimated, where α is the intercept, $\beta_{1,\ldots}\beta_n$ denote the parameter estimates of the register-based dummy variables described in the chapter *Explanatory variables*, β_{NF} denotes the parameter estimate of numerical flexibility, β_D the parameter estimate of decentralisation, $\beta_{IL}$the parameter estimate of individual learning and ε is a random error. The probability of belonging to a certain group is denoted by *p,* and $log\left(\frac{p}{1-p}\right)$ is the natural logarithm of the odds. A negative parameter estimate indicates that the variable diminishes the log odds, and thus indirectly the probability of being part of the group under investigation, while a positive parameter estimate has the opposite effect.

Linear combinations of the explanatory variables are avoided by excluding one category-dummy of each register-based explanatory variable from the model, or equally including it in the intercept. P-values based on Wald test statistics are used to assess the significance of the explanatory variables. A more thorough, and theoretical, description of the logistic regression model may be found in e.g. (Agresti 2002, chapters 5-6).

To get an impression of the adequacy of the models we use pseudo R^2. Several such measures have been proposed, a choice of which are discussed in (Menard, 2000). We choose to rely on the one proposed by Nagelkerke (Menard, 2000, pp. 20),

$$R^2_{pseudo} = \frac{1 - \left(\frac{L_0}{L_M}\right)^{2/n}}{1 - {L_0}^{2/n}},$$

where L_0 denotes the likelihood of a model containing only the intercept, L_M the likelihood of the model with all explanatory variables and *n* is the sample size. It is important to note that the interpretation of pseudo R^2 measures is not equivalent to the interpretation of the coefficient of determination, R^2, in OLS regression. R^2_{pseudo} does not represent the proportion of variability explained by the model. Instead we should consider it "a measure of strength between the dependent variable and the total set of predictors" (Menard, 2000, pp. 17). In similarity with the ordinary R^2 measure, R^2_{pseudo} ranges from zero to one, with higher values indicating better model fit.

4.2 Multiple Linear Regression

To analyse the yearly connections between the organisational variables and income development, multiple linear regression is used. The model is of the type,

$$Income\_percent\_x\_1998 = \alpha + \beta_1 x_1 + \cdots + \beta_n x_n + \beta_{NF} x_{NF} + \beta_D x_D + \beta_{IL} x_{IL} + \varepsilon \; (\mathbf{2}),$$

where $Income\_percent\_x\_1998$ denotes the percentage income increase or decrease of the individual since 1998, and the remaining notation corresponds completely to the one of the logistic regression model. In this case the parameter estimates represent the direct impact on the percentage income development of the variable. The coefficient of determination, R^2, is used to determine the proportion of variability in data that is accounted for by the model.

A model assumption, to make correct inferences based on the linear regression model, is that the residuals, ε, should be independent and normally distributed with an expected value of zero and constant variance. In our case, however, the normality assumption needs not to be checked for. Considering the large number of observations, the central limit theorem assures that the p-values, based on the t-test statistics, are asymptotically correct, given that the other assumptions hold. More detailed information on the multiple linear regression model may be found in e.g. (Lindgren, 1993, pp. 508-512) or (Sundberg, 2009, ch. 3).

4.3 Generalized Estimating Equations

Finally we want to make use of the fact that we have got longitudinal (panel) data, and analyse the datasets from all years in the same model. A method called generalized estimating equations, GEE, is used. Although the mathematical theory behind differs, the practical aspects of the method (such as the structure of the model and the interpretation of the results) are almost identical to ordinary multiple logistic regression. The advantage of GEE is that the fact that the same persons are observed over time is taken into account. For example, it is reasonable to believe that having been on sick leave in one year generally increases the probability of being on sick leave also in some of the other years. GEE allows this kind of correlation within subjects (individuals), while the logistic regression model has an underlying assumption of independence between all observations, which is very likely to not be fulfilled when the datasets from all years are used in the same analysis. Furthermore, the model allows attrition, i.e. in our case that the individuals leave the sample the year they reach the age of 65 or in case of death or moving abroad.

An overview of our longitudinal GEE model follows below.

$$log\left(\frac{p}{1-p}\right) = \alpha + \beta_1 x_1 + \cdots \beta_n x_n + \beta_{NF} x_{NF} + \beta_D x_D + \beta_{IL} x_{IL}$$

$$+\beta_{1999} x_{1999} + \cdots + \beta_{2007} x_{2007} + \beta_{NF}\beta_{1999} x_{NF} x_{1999} + \cdots \beta_{NF}\beta_{2007} x_{NF} x_{2007}$$

$$+\beta_D\beta_{1999} x_D x_{1999} + \cdots \beta_D\beta_{2007} x_D x_{2007} + \beta_{IL}\beta_{1999} x_{IL} x_{1999} + \cdots \beta_{IL}\beta_{2007} x_{IL} x_{2007} + \varepsilon \quad \mathbf{(3)}$$

The interpretation of the elements on the first row, and the interpretation of ε, correspond completely to the interpretation of the elements in the ordinary multiple logistic regression models (equation (**1**)). For natural reasons the number of people who are still working in the same organisation as in 1998 decreases with time, which means that the share of individuals in the other categories of occupation is smaller in the years immediately after 1998, compared to later. This, in combination with the fact that the state of the market may differ from year to year, motivates the inclusion of year dummy variables. To avoid linear combinations of the time dummies, one year needs to be excluded from the model. Our choice is 2008. We denote the parameter estimates β_{1999}, ... β_{2007}. For the same reason, we also exclude one category-dummy from each of the register based explanatory variables. The impact of these variables is instead caught in the intercept.

The organisational variables describe the work organisation of the company in which the employee was employed in 1998. Since people quit their jobs and reorganisations may take place, we do believe that the importance of these variables differs over time, and we wish to study these possible differences. Therefore, in addition to the main effects β_{NF}, β_D and β_{IL}, interaction terms between the three organisational variables and the year dummies are introduced, with parameter estimates denoted by $\beta_{NF}\beta_{1999}, \ldots, \beta_{IL}\beta_{2007}$.

P-values based on z-test statistics are used to determine whether the explanatory variables are significant or not. An exposition on some important theoretical aspects of the model that should be mentioned, and references for further reading is found in the appendix.

5 Results

5.1 The Probability of Having a Job

Our first analysis aims to find possible differences between individuals having and not having a job with respect to the organisational variables of the companies in which they were employed in 1998. For each of the years 1999-2008, logistic regressions are performed as described in equation (1), with not being employed in the same or in another organisation as response variable (defined as those who do not belong to neither category 1 nor 2 according to the definitions in chapter 2.4) and explanatory variables as previously specified.

Thereafter a longitudinal study is performed, with the same explanatory variables as in the former analysis, but with an additional time effect and interaction effect between time and the organisational variables included, as described in equation (3). The parameter estimates of the organisational variables and in the latter case, additionally the interaction effect between the organisational variables and the year dummies, are presented in Table 4 and Table 5 respectively.

Throughout the report, we let * denote statistical significance on the 10 percent level, ** on the 5 percent level and *** on the 1 percent level. The reader should note that this is not the most commonly used system for denoting statistical significance. As previously mentioned, the explanatory variables describing the gender, age, education level, industry, ethnic background and regional group are not of main interest, but are rather included in the models to be controlled for. Therefore we do not to present the parameter estimates of these variables in our report.

Table 4: Parameter estimates of numerical flexibility, β_{NF}, decentralisation, β_D, and individual learning, β_{IL}, from logistic regressions performed on each year separately, as described in equation (1), with *non job* as response variable.

Year	Numerical flexibility	Decentralisation	Individual learning	R^2_{pseudo}
1999	0.4305***	-0.0488***	-0.1979***	0.0484
2000	0.2078***	-0.0700***	-0.2199***	0.0601
2001	0.0431	-0.0538***	-0.1255***	0.0641
2002	0.1043***	-0.0524***	-0.1759***	0.0674
2003	0.1282***	-0.0680***	-0.1326***	0.0773
2004	0.1106***	-0.0690***	-0.1389***	0.0820
2005	0.0848***	-0.0690***	-0.0772***	0.0873
2006	0.0617**	-0.0647***	-0.0713***	0.0931
2007	0.0389	-0.0612***	-0.0831***	0.0963
2008	0.00818	-0.0619***	-0.0709***	0.0981

The overall trend is that the impact of numerical flexibility diminishes over the years, whereas the coefficient of decentralisation lies on a relatively stable level. The level of individual learning at the companies in which the individuals were employed in 1998 loses importance as time goes by, although there is still a non-negligible connection between the occupation of the individuals and the index in 2008. We note that the coefficients of determination, R^2_{pseudo} , are low. This is further commented in chapter 6, *Discussion of the Statistical Models*.

Table 5: Parameter estimates of the main effects of numerical flexibility, β_{NF}, decentralisation, β_D, individual learning, β_{IL}, and interaction effects between the three variables and the year dummies, $\beta_{NF}\beta_{1999}$, … $\beta_{IL}\beta_{2007}$, from the GEE analysis, as described in equation (3), with *non job* as response variable.

Numerical flexibility	0.0345		
Decentralisation	-0.0072		
Individual learning	-0.0704***		
Year	**Numerical flexibility*year**	**Decentralisation*year**	**Individual learning*year**
1999	0.5465***	-0.1130***	-0.3273***
2000	0.2393***	-0.0364***	-0.3204***
2001	0.0879**	-0.0341***	-0.1982***
2002	0.0708**	-0.0215***	-0.1419***
2003	0.0839***	-0.0390***	-0.0078
2004	0.0714**	-0.0419***	-0.0220
2005	0.0526**	-0.0329***	0.0341
2006	0.0505**	-0.0243***	0.0339*
2007	-0.0037	-0.0093**	0.0027

The overall impression is that the results from the longitudinal study are similar to the results from the separate logistic regressions. After having controlled for gender, age, education level, industry, foreign background and regional group, numerical flexibility significantly increases the probability of being out of work until 2006, and also in this case the effect of the variable is decreasing over the years. As in the former analyses, a high level of decentralisation decreases the probability of not having a job in each of the years. The effect is generally relatively small and stable, except for in 1999 when the parameter estimate is much larger, which was not the case in the former analysis. Since the main effect of individual learning is significant, we draw the conclusion that this variable decreases the probability of being out of work during all of the years. In addition to this, the significant interaction effects between time and individual learning show that the level of individual learning at the company of 1998 has a larger impact during the years 1999-2002, compared to later. This more or less corresponds to the results where logistic regressions were performed on each year separately. We emphasize the fact that since industry is included in the model, these results are valid within each line of business.

The results from both types of analysis above show that there is a positive connection between having worked in a company with a high degree of numerical flexibility and being out of work in each of the following eight years (with one exception). This most likely reflects the fact that the more numerically flexible a firm is, the more temporary personnel it is likely to use, which by definition means that a relatively large share of the employees will lose their jobs within a certain period of time. While some of them continue working in other firms, others do not succeed in finding new jobs, which explains the positive sign of the parameter.

Temporary contracts typically run over short periods, so within a couple of years most employees who are affected by this have already quitted, which explains the higher values of the parameter estimates in the years 1999-2000. In spite of this, the variable still has some importance until 2006. One possible reason could be the fact that people with short time contracts are likely to have gotten new short time contracts in other companies before falling into one of the non job categories. We may also suspect that numerical flexibility is more common in certain kinds of professions, and since it is more likely to change job within your former profession than trying a completely new one, the variable may still be fairly relevant for the year under investigation.

The relatively stable effect of decentralisation is more surprising. Having had a job in a, to a large extent decentralised firm in 1998 significantly increases the chance of having a job in each of the following years. A credible explanation is that having worked in a decentralised company has had a large influence on the individual. Someone who is used to planning one's own work, who has been responsible of the quality control and follow-up of results and who is an experienced team worker, is generally more attractive on the labour market than others even after having changed work more than once or having been in one of the non job categories. It is a positive experience that you either have or do not have. Another possible explanation is that the stable impact is a consequence of the fact that people who get new jobs tend to stay within the same field of work, and that the level of decentralisation is likely to be similar at these companies. However, the result is a clear indication that employees of highly decentralised firms tend to be in work to a larger extent than others in the future. Our conclusion is thus that working at a decentralised company is positive for the individual.

Individual learning decreases the probability of being out of work. As time goes by, the probability of having changed work at least once increases which in turn increases the probability of having worked at a company with a different level of individual learning. This should be one reason for the weaker impact of individual learning over the years. Another reason could be the fact that what the employees learned in 1998 is not relevant for their work anymore, so that the variable loses importance also for the group of employees who are still working in the same company. Since the variable still has some impact in 2008 there is reason to believe that for a relatively large group, the knowledge and experience that resulted from the high level of individual learning at the company was of a type that not only increased the employer´s willingness to keep the employee, but also other employers' willingness to employ him or her.

The above analyses shed some light on the general trends, but also raise further questions of interest. We may for example believe that numerical flexibility does not have an impact on the whole non job group, but mainly on the risk of unemployment or the risk of falling into the group of people with a very low yearly income. It is also reasonable to think that having worked in a numerically flexible firm should increase the probability of working at another company and decrease the probability of continuing working at the same company as in 1998. One may suspect that the effect of the variables describing the degree of decentralisation and individual learning should be more stable for the group of people still working in the same company, in comparison to the other groups, since the variable is likely to still be highly relevant for this group. To further investigate these, among other, questions of interest we proceed by performing logistic regressions and longitudinal studies on these subgroups.

5.2 Two Categories of Work

To find the characteristics of the two groups of people with employments, logistic regressions and longitudinal analyses are performed, with employed within the same or employed within another organisation as response variables, respectively. Table 6 and 7 present the results of the analyses of the group of people still employed within the same firm and in table 8 and 9 the corresponding results for those who are employed within other firms are found. Thereafter follows an analysis of the results in all four tables. As previously mentioned, the parameter estimates for the register-based dummy variables describing the individuals, are not presented.

Table 6: Parameter estimates of numerical flexibility, β_{NF}, decentralisation, β_{D}, and individual learning, β_{IL}, from logistic regressions performed on each year separately, as described in equation (1), with *employed within the same firm* as response variable.

Year	Numerical flexibility	Decentralisation	Individual learning	R^2_{pseudo}
1999	-0.5281***	0.1186***	0.3493***	0.0521
2000	-0.3072***	0.0590***	0.1099***	0.0476
2001	-0.2213***	0.0289***	0.1091***	0.0529
2002	-0.1331***	0.0448***	0.1436***	0.0804
2003	-0.1189***	0.0719***	0.1417***	0.1008
2004	-0.0977***	0.0656***	0.1380***	0.1011
2005	-0.1142***	0.0621***	0.1407***	0.1035
2006	-0.0463**	0.0540***	0.1169***	0.1058
2007	-0.1096***	0.0638***	0.1355***	0.1166
2008	-0.1358***	0.0636***	0.1426***	0.1177

It is clear that numerical flexibility decreases the probability of staying in the same company the most in 1999-2001, and thereafter levels away. The effect of decentralisation is relatively constant, and positive, except for a slightly larger value in the first year. Individual learning increases the probability of staying in the same firm in a rather constant way over time, except for in 1999 when the variable had a larger impact.

Table 7: Parameter estimates of the main effects of numerical flexibility, β_{NF}, decentralisation, β_{D}, individual learning, β_{IL}, and interaction effects between the three variables and the year dummies, $\beta_{NF}\beta_{1999}$, … $\beta_{IL}\beta_{2007}$, from the GEE analysis, as described in equation (3), with *employed within the same firm* as response variable.

Numerical flexibility	-0.2594***
Decentralisation	0.0589***
Individual learning	0.2252***

Year	Numerical flexibility*year	Decentralisation*year	Individual learning*year
1999	-0.6174***	0.0643***	0.5373***
2000	-0.3302***	-0.0403***	0.2355***
2001	-0.1921***	-0.0663***	0.1905***
2002	-0.0395**	-0.0312***	0.0594***
2003	0.0082	0.0065*	-0.0035
2004	0.0317**	0.0028	0.0117
2005	0.0153	0.0023	0.0212*
2006	0.0719***	-0.0110***	0.0010
2007	0.0161*	-0.0018	-0.0037

The significant main effects show that throughout the period, numerical flexibility decreases and individual learning increases the probability of staying within the same organisation as in

1998. Besides this, the significant interaction effects imply that both variables have a larger, but diminishing, effect in the years 1999-2002, compared to the later years, which was also the case in the former analysis. Decentralisation increases the probability of staying within the organisation the most in 1999, and is on a rather stable level as from 2003. In 2000-2002 the impact is much weaker, which is seen by taking both the main and the interaction effect into account (the sum of these variables). Whereas numerical flexibility, and to some extent decentralisation, follow the same pattern as in the former analysis, we have found a diminishing effect of individual learning that was not evident before.

Table 8: Parameter estimates of numerical flexibility, β_{NF}, decentralisation, β_{D}, and individual learning, β_{IL}, from logistic regressions performed on each year separately, as described in equation (1), with *employed within another firm* as response variable.

Year	**Numerical flexibility**	**Decentralisation**	**Individual learning**	R^2_{pseudo}
1999	0.5275***	-0.1444***	-0.3931***	0.0735
2000	0.3090***	-0.0457***	-0.0382*	0.0753
2001	0.2675***	-0.0118**	-0.0742***	0.0889
2002	0.1254***	-0.0299***	-0.0857***	0.1131
2003	0.0871***	-0.0496***	-0.0951***	0.1273
2004	0.0656***	-0.0389***	-0.0798***	0.1287
2005	0.0953***	-0.0328***	-0.1144***	0.1356
2006	0.0301	-0.0258***	-0.0894***	0.1422
2007	0.1085***	-0.0384***	-0.1037***	0.1567
2008	0.1482***	-0.0363***	-0.1149***	0.1621

Numerical flexibility increases the probability of being employed in another firm, especially during the years 1999-2001. Decentralisation makes it less likely to be working at another company although the effect seems to be relatively small as from the year 2000. Except for one non-significant result, individual learning significantly decreases the probability of being employed in another firm, particularly in 1999.

Table 9: Parameter estimates of the main effects of numerical flexibility, β_{NF}, decentralisation, β_{D}, individual learning, β_{IL}, and interaction effects between the three variables and the year dummies, $\beta_{NF}\beta_{1999}$, … $\beta_{IL}\beta_{2007}$, from the GEE analysis, as described in equation (3), with *employed within another firm* as response variable.

Numerical flexibility	0.2007***
Decentralisation	-0.0687***
Individual learning	-0.1892***

Year	**Numerical flexibility*year**	**Decentralisation*year**	**Individual learning*year**
1999	0.5380***	-0.1298***	-0.5544***
2000	0.2934***	0.0291***	-0.1692***
2001	0.2064***	0.0640***	-0.1593***
2002	-0.0025	0.0273***	-0.0012
2003	-0.0591***	0.0000	0.0180
2004	-0.0799***	0.0082*	0.0043
2005	-0.0511***	0.0104***	-0.0422***
2006	-0.1065***	0.0202***	-0.0117
2007	-0.0321***	0.0024	0.0043

The positive main effects show that there is a connection between having a new job and numerical flexibility and individual learning respectively, during the whole period. Both variables are particularly important until 2001, and thereafter stay on a more stable level. Decentralisation generally has a weak negative impact, except for in 2001 when the sum of the main and interaction effect was nearly zero. We draw the conclusion that numerical flexibility and decentralisation generally behave in the same way as when separate logistic regressions were performed, while the diminishing impact of individual learning only could be detected in this later analysis.

In comparison with the general results in the previous section, that aimed to find the differences between those having and not having a job, we see that the tendency of diminishing impact of numerical flexibility as well as the relatively constant impact of decentralisation remain. Individual learning on the other hand, that lost importance with time in the former analysis stays on a relatively constant level when logistic regressions are performed for each year separately, but has the same diminishing tendency in the two longitudinal studies. We note that the parameter estimates of the analyses of the group of people still employed by the same firm are close to the parameter estimates of the group of people employed by other firms, but with opposite sign. This is hardly surprising since the vast majority of those who do not work in the same firm work in another firm, and what is positive for the probability of staying within the firm is on the other hand negative for the probability of leaving it.

Our interpretation of the diminishing effect of numerical flexibility corresponds to the one of the analysis of the whole non job group, which in short was that numerical flexibility by definition means that some employees will lose their jobs within a short period of time. As expected, the signs differ in the analyses of the two groups. Numerical flexibility decreases the probability of staying at one´s job, but also increases the probability of getting a new one. The latter should probably be seen as a consequence of the fact that those who were employed in highly numerically flexible firms in 1998 have been forced to find new jobs to a larger extent than others, rather than a positive outcome of numerical flexibility. Worth mentioning is that in the analyses of both groups the variable remains significant during the whole period (with one exception), which differs from the result of the whole non job group where the variable only had a significant influence until 2006.

We see that decentralisation does not only increase the probability of having a job, but it also increases the probability of staying at one´s job and decreases the probability of being employed by a new firm. It is not unrealistic to draw the conclusion that decentralisation is appreciated by the employees and that the level of decentralisation has an influence on the willingness to start looking for new jobs.

Individual learning makes it more likely to stay within the firm and hence less likely to be employed in a new one. Relevant knowledge increases the employer's willingness to keep the employee but probably also the employee's willingness to stay within the firm.

5.3 Four Categories of Non Job

To further analyse the non job group, logistic regressions and longitudinal analyses, as defined in equations (**1**) and (**3**), are performed with the unemployed, those on sick leave, the disability pensioners and others with low income as response variables respectively. The parameter estimates of the organisational variables are presented in tables 10-16.

Table 10: Parameter estimates of numerical flexibility, β_{NF}, decentralisation, β_{D}, and individual learning, β_{IL}, from logistic regressions performed on each year separately, as described in equation (1), with *unemployed* as response variable.

Year	Numerical flexibility	Decentralisation	Individual learning	R^2_{pseudo}
1999	0.6803***	-0.0574***	-0.2482***	0.0415
2000	0.3730***	0.00582	-0.5923***	0.0564
2001	0.1867***	0.00139	-0.5173***	0.0528
2002	0.2091***	-0.0432***	-0.3641***	0.0389
2003	0.1460***	-0.0979***	-0.2198***	0.0470
2004	0.2573***	-0.0956***	-0.3024***	0.0468
2005	0.2097***	-0.0588***	-0.2132***	0.0431
2006	0.2108***	-0.0654***	-0.0664	0.0431
2007	0.1980***	-0.0293*	-0.1409**	0.0415
2008	0.2827***	-0.0664***	-0.1387**	0.0464

From table 10 we draw the conclusion that all three indices do have an influence on the probability of unemployment. As expected, the risk is larger for those who worked in firms with a high level of numerical flexibility in 1998, especially in the two subsequent years. Both decentralisation and individual learning generally decrease the probability of unemployment.

Table 11: Parameter estimates of the main effects of numerical flexibility, β_{NF}, decentralisation, β_{D}, individual learning, β_{IL}, and interaction effects between the three variables and the year dummies, $\beta_{NF}\beta_{1999}$, … $\beta_{IL}\beta_{2007}$, from the GEE analysis, as described in equation (3), with *unemployed* as response variable.

Numerical flexibility	0.2248***		
Decentralisation	-0.0289*		
Individual learning	-0.1567***		
Year	**Numerical flexibility*year**	**Decentralisation*year**	**Individual learning*year**
1999	0.6978***	-0.0263	-0.3877***
2000	0.2647***	0.0796***	-0.6189***
2001	0.0233	0.0725***	-0.4776***
2002	0.0330	-0.0217	-0.2974***
2003	-0.0808	-0.0877***	-0.0063
2004	0.0341	-0.0652***	-0.0850
2005	0.0000	-0.0229	-0.0419
2006	-0.0009	-0.0167	0.0420
2007	-0.0529	0.0300**	-0.0179

As can be seen in table 11, the positive main effect and the positively significant interaction effects in 1999-2000 confirm the conclusions about numerical flexibility that were drawn from the separate logistic regressions. Individual learning decreases the probability of unemployment throughout the period, with a peak in the four earliest years. Quite contrary to what was found before, decentralisation increases the probability of unemployment in 2000-2001, and significantly decreases this probability only in a few of the later years. Since the results differ in the two analyses and the parameter estimates are rather small in comparison with the other two organisational variables, we choose not to draw any conclusions about the impact of decentralisation.

Table 12: Parameter estimates of numerical flexibility, β_{NF}, decentralisation, β_D, and individual learning, β_{IL}, from logistic regressions performed on each year separately, as described in equation (1), with *on sick leave* as response variable.

Year	Numerical flexibility	Decentralisation	Individual learning	R^2_{pseudo}
1999	0.1923*	-0.1548***	-0.0234	0.0798
2000	0.0608	-0.1324***	-0.0789	0.0776
2001	-0.0687	-0.1390***	-0.0795	0.0652
2002	-0.00221	-0.1410***	-0.1240**	0.0569
2003	0.1010	-0.1550***	-0.0766	0.0518
2004	-0.0529	-0.1470***	-0.1505**	0.0451
2005	-0.1404*	-0.1462***	-0.0896	0.0399
2006	-0.1444*	-0.1274***	-0.1445**	0.0373
2007	-0.1272	-0.1330***	-0.1949**	0.0375
2008	-0.0497	-0.1229***	-0.1716**	0.0333

Table 12 shows that we cannot establish any statistical connection (on the 5% level) between the level of numerical flexibility and being on sick leave. Neither is there any strong evidence that the degree of individual learning would have an impact on the probability of getting sick, at least not in the short run. What we do see, on the other hand, is that decentralisation clearly decreases the risk of being on sick leave.

Table 13: Parameter estimates of the main effects of numerical flexibility, β_{NF}, decentralisation, β_D, individual learning, β_{IL}, and interaction effects between the three variables and the year dummies, $\beta_{NF}\beta_{1999}$, … $\beta_{IL}\beta_{2007}$, from the GEE analysis, as described in equation (3), with *on sick leave* as response variable.

Numerical flexibility	-0.0353
Decentralisation	-0.1295***
Individual learning	-0.1313*

Year	Numerical flexibility*year	Decentralisation*year	Individual learning*year
1999	0.2301*	-0.0231	0.0785
2000	0.0741	-0.0095	0.0378
2001	-0.0152	-0.0084	0.0153
2002	0.0273	-0.0134	0.0050
2003	0.1183	-0.0210	0.0939
2004	-0.0165	-0.0197	0.0142
2005	-0.1027	-0.0161	0.0749
2006	-0.1099	-0.0000	0.0229
2007	-0.0524	-0.0037	-0.0582

In table 13 we see that the conclusions that may be drawn from the longitudinal study correspond to what was seen in the former analysis. Decentralisation diminishes the probability of being on sick leave on a constant level in all of the years under investigation. In this analysis, individual learning does not have an impact on the five percent level in any of the years. On the other hand, the main effect is significant on the ten percent level. This fact, in combination with the occasionally significant impact seen in the separate logistic regressions, makes us draw the conclusion that individual learning is likely to influence (diminish) the probability of being on sick leave to at least some extent.

Table 14: Parameter estimates of numerical flexibility, β_{NF}, decentralisation, β_D, and individual learning, β_{IL}, from logistic regressions performed on each year separately, as described in equation (1), with *disability pensioner* as response variable.

Year	Numerical flexibility	Decentralisation	Individual learning	R^2_{pseudo}
1999	0.1441	-0.0743	-0.0247	0.1263
2000	0.0512	-0.1653***	0.4377***	0.1363
2001	-0.1018	-0.1665***	0.2573***	0.1523
2002	-0.1305*	-0.1220***	0.1402**	0.1573
2003	-0.0158	-0.1013***	0.0342	0.1584
2004	0.00492	-0.0882***	0.0120	0.1629
2005	0.00616	-0.1034***	0.0571	0.1615
2006	-0.00317	-0.1018***	0.0637	0.1604
2007	-0.0491	-0.1014***	0.0671	0.1576
2008	-0.1069**	-0.0943***	0.0709*	0.1557

In table 14 we see that the probability of living on disability pension is affected by the level of decentralisation at the company of 1998. Once again, having worked in a decentralised firm seems to be positive for the individual, this time by decreasing the risk of becoming a disability pensioner.

Table 15: Parameter estimates of the main effects of numerical flexibility, β_{NF}, decentralisation, β_D, individual learning, β_{IL}, and interaction effects between the three variables and the year dummies, $\beta_{NF}\beta_{1999}$, … $\beta_{IL}\beta_{2007}$, from the GEE analysis, as described in equation (3), with *disability pensioner* as response variable.

Numerical flexibility	-0.1022**
Decentralisation	-0.0831***
Individual learning	0.0678*

Year	Numerical flexibility*year	Decentralisation*year	Individual learning*year
1999	0.1220	0.0419	-0.1809
2000	0.0956	-0.0525	0.4202***
2001	-0.0177	-0.0775***	0.2553***
2002	-0.0149	-0.0384**	0.1175*
2003	0.0896	-0.0149	0.0089
2004	0.0844*	-0.0041	-0.0148
2005	0.0760**	-0.0162**	0.0038
2006	0.0701**	-0.0156*	-0.0036
2007	0.0358	-0.0074	0.0028

From table 15 we may draw the conclusion that the results differ slightly from the results of the separate logistic regression analyses. In both cases, decentralisation clearly decreases the probability of early retirement. The GEE analysis however, shows that also numerical flexibility has a negative effect. Since this is quite contrary to what was found in the former analyses, we choose not to draw any conclusions about the impact of numerical flexibility on the probability of disability pension.

Table 16: Parameter estimates of numerical flexibility, β_{NF}, decentralisation, β_{D}, and individual learning, β_{IL}, from logistic regressions performed on each year separately, as described in equation (1), with *other, low income* as response variable.

Year	**Numerical flexibility**	**Decentralisation**	**Individual learning**	R^2_{pseudo}
1999	0.0723	-0.0896***	-0.1672**	0.0562
2000	0.0415	-0.0331*	-0.0703	0.0451
2001	0.0742	0.00870	0.0108	0.0443
2002	0.0565	0.0729***	-0.1099**	0.0518
2003	0.1309***	0.0721***	-0.0609	0.0565
2004	0.0533	0.0283**	-0.00517	0.0551
2005	0.0612	0.00696	0.00813	0.0498
2006	0.0657	0.00606	-0.0691*	0.0424
2007	0.0148	-0.0129	-0.0808**	0.0419
2008	0.0466	-0.00725	-0.1013***	0.0401

In table 16 we see that there is no clear connection between any of the organisational variables and being part of the group of people with low incomes. Since this group is likely to include people with widely spread reasons for not having an income, the result is not very surprising. Presumably the group mainly includes people who are receiving support from the social security, people who are financially supported by the members of their families and to some extent people living on saved money.

Table 17: Parameter estimates of the main effects of numerical flexibility, β_{NF}, decentralisation, β_D, individual learning, β_{IL}, and interaction effects between the three variables and the year dummies, $\beta_{NF}\beta_{1999}$, … $\beta_{IL}\beta_{2007}$, from the GEE analysis, as described in equation (3) with *other, low income* as response variable.

Numerical flexibility	0.0383		
Decentralisation	0.0350***		
Individual learning	-0.0178		
Year	**Numerical flexibility*year**	**Decentralisation*year**	**Individual learning*year**
1999	0.1701	-0.2372***	-0.3336***
2000	0.0060	-0.1290***	-0.0947
2001	0.0461	-0.0672***	-0.0329
2002	0.0144	0.0262**	-0.1113**
2003	0.0684	0.0334***	-0.0213
2004	0.0114	-0.0007	0.0414
2005	0.0277	-0.0099	0.0893***
2006	0.0360	-0.0073	0.0362
2007	-0.0179	-0.0118*	0.0195

The results from the longitudinal study, presented in table 17, show that numerical flexibility does not influence the probability of being part of this group. Neither is there any clear evidence that individual learning is important, although it has a significant impact in some of the years. Decentralisation on the other hand has a decreasing effect until 2001, and thereafter weakly increases the probability of being part of this group. This effect could not be seen in the former analysis although there is a tendency of a decreasing effect in the first years and an increasing effect in some of the later years.

From the above analyses we draw the conclusion that numerical flexibility does have an influence on the probability of unemployment, but not on any of the other non job categories. For the same reasons as mentioned in earlier chapters, the impact of the variable is largest in the years immediately after 1998.

Individual learning mainly has an impact on the probability of unemployment. It decreases the probability of getting unemployed to a relatively large extent. On one hand this is likely to reflect the fact that employers would not invest in the learning and the education of the employee if the intention was not to keep the employee in the company for a longer period. On the other hand, participating in organised skills development activities and in the development of working tasks continuously puts the employee in a situation where he or she has a knowledge specific for the work that he or she performs, which may be hard to replace. Also, the increased knowledge of the employee is likely to increase his or her chances of getting employed by other companies.

Decentralisation clearly decreases the probability of being on sick leave and of living on disability pension. That is, we have detected an association between health and having worked in a highly decentralised company. Being given the possibility to plan one's own work, taking part in the quality control and follow-up of results, working in teams and having an impact on one's own working schedule seems to decrease the risk of having to leave the labour market due to illness.

5.4 Interaction Effects Taken into Account

Our next step is to perform the corresponding analyses with the two-factor interaction effects between the organisational variables taken into account. With this approach we aim to investigate whether there is a connection between the variables describing the work organisation. That is, does the fact that a firm has a certain level of e.g. decentralisation affect to which extent and in which way numerical flexibility or individual learning influences the response variables? An overview of the model is found below.

$$log\left(\frac{p}{1-p}\right) = \alpha + \beta_1 x_1 + \cdots + \beta_n x_n + \beta_{NF} x_{NF} + \beta_D x_D + \beta_{IL} x_{IL} + \beta_{NF}\beta_D x_{NF} x_D + \beta_{NF}\beta_{IL} x_{NF} x_{IL} + \beta_D \beta_{IL} x_D x_{IL} + \varepsilon \quad (\mathbf{4})$$

Except for the interaction effects $\beta_{NF}\beta_D$, $\beta_{NF}\beta_{IL}$ and $\beta_D\beta_{IL}$, the model is identical with the one described in equation (**1**).

We start by performing logistic regressions, for each of the years 1999-2008 separately, with non job as response variable. The results are presented in table 18.

Table 18: Parameter estimates of numerical flexibility, β_{NF}, decentralisation, β_D, individual learning, β_{IL}, and the two-factor interaction terms between the three variables, from logistic regressions performed on each year separately, as described in equation (4), with *non job* as response variable.

Year	Numerical flexibility	Decentralisation	Individual learning	Num*Dec	Num*Ind	Dec*Ind	R^2_{pseudo}
1999	0.6576***	0.0818**	-0.3814***	-0.1992***	0.3064**	-0.00067	0.0487
2000	0.1085	-0.0291	-0.3765***	-0.0495	0.3026***	-0.0135	0.0602
2001	0.0932	-0.0692**	-0.1392	-0.00462	-0.0608	0.0253	0.0641
2002	-0.1565*	-0.1022***	-0.3135***	0.0604**	0.1774*	0.0144	0.0675
2003	-0.0952	-0.1234***	-0.1893**	0.0784***	0.0683	0.00658	0.0774
2004	0.0702	-0.1089***	-0.0340	0.0692**	-0.1589*	-0.00553	0.0820
2005	0.0968	-0.1235***	0.0395	0.0709**	-0.2400	0.0131	0.0874
2006	0.00642	-0.1219***	-0.0633	0.0578**	-0.1075***	0.0276	0.0932
2007	-0.0141	-0.0818***	0.0202	0.0575*	-0.1021	-0.0217	0.0963
2008	-0.1690*	-0.0875***	-0.0236	0.0749**	0.0190	-0.0303	0.0982

As in the former analyses of the non job group, having worked in a company that has a high level of decentralisation generally increases the probability of having a job later in life. Individual learning does also significantly affect this probability in the same direction, but only in the first five years. Having worked in a company that had both these characteristics at the same time does not decrease or increase this probability any further apart from the impact of each of the main effects separately, i.e. there is no significant interaction effect in any of the years.

Having had a job that was numerically flexible increases the probability of not having a job in 2002-2008, but to which extent depends on the level of decentralisation (i.e. the interaction effect between the two variables is significant). In other words, a given level of numerical flexibility has a larger impact for those who worked in highly decentralised companies compared to those who worked in less decentralised ones. In 1999-2000 the interaction effect between numerical flexibility and individual learning is significant, and we draw the analogous conclusion that the level of individual learning decides to which extent numerical flexibility influences the probability of not having a job.

Next we study the two groups of working people separately. Logistic regressions are performed, as described in equation (**4**), with employed within the same organisation and employed within another organisation as response variables respectively. The parameter estimates of the organisational variables are found in table 19 and 20.

Table 19: Parameter estimates of numerical flexibility, β_{NF}, decentralisation, β_{D}, individual learning, β_{IL}, and the two-factor interaction terms between the three variables, from logistic regressions performed on each year separately, as described in equation (4), with *employed within the same organisation* as response variable.

Year	Numerical flexibility	Decentralisation	Individual learning	Num*Dec	Num*Ind	Dec*Ind	R^2_{pseudo}
1999	-0.4170***	0.00298	0.8403***	0.1920***	-0.7612***	-0.0117	0.0529
2000	-0.6118***	0.0166	0.3119***	0.1577***	-0.0641	-0.0818***	0.0480
2001	-0.5500***	0.00480	-0.1821***	0.0129	0.4136***	0.0199	0.0532
2002	-0.1105*	0.1509***	-0.2289***	-0.1874***	0.5641***	0.0186	0.0812
2003	-0.2272***	0.1974***	-0.3245***	-0.1980***	0.7836***	0.000574	0.1020
2004	-0.0188	0.2529***	-0.2860***	-0.2752***	0.7734***	-0.0157	0.1028
2005	0.0393	0.2738***	-0.2259***	-0.2944***	0.7345***	-0.0319*	0.1053
2006	-0.3302***	0.2176***	-0.2064***	-0.1462***	0.8815***	-0.0964***	0.1074
2007	-0.1678***	0.2277***	-0.1776***	-0.1952***	0.7193***	-0.0539***	0.1178
2008	-0.1210*	0.2090***	-0.2007***	-0.2063***	0.6497***	-0.0191	0.1187

As expected, when numerical flexibility is significant it decreases the probability of being part of the group of people still working in the same organisation. Individual learning increases the probability of being part of this group in the first two years, and decreases it thereafter. Decentralisation as a main effect is positively significant as from 2002. In 1999-2000 the negative impact of numerical flexibility is smaller in highly decentralised firms, compared to less decentralised ones (the interaction effect is positive).

The interaction effect between numerical flexibility and individual learning is generally positively significant. This means that the negative impact of having worked in a numerically flexible firm is reduced if the firm at the same time had a high degree of individual learning.

Table 20: Parameter estimates of numerical flexibility, β_{NF}, decentralisation, β_{D}, individual learning, β_{IL}, and the two-factor interaction terms between the three variables, from logistic regressions performed on each year separately, as described in equation (4), with *employed within another organisation* as response variable.

Year	Numerical flexibility	Decentralisation	Individual learning	Num*Dec	Num*Ind	Dec*Ind	R^2_{pseudo}
1999	0.1593*	-0.0513*	-0.9976***	-0.1437***	0.9755***	-0.00171	0.0745
2000	0.7199***	0.000261	-0.2091***	-0.1741***	-0.0362	0.0964***	0.0753
2001	0.6359***	0.0367*	0.3083***	-0.0147	-0.4603***	-0.0486***	0.0889
2002	0.2033***	-0.1266***	0.4390***	0.1991***	-0.7382***	-0.0402**	0.1139
2003	0.2919***	-0.1693***	0.4809***	0.2017***	-0.9262***	-0.0105	0.1285
2004	-0.0516	-0.2340***	0.3402***	0.2891***	-0.7599***	0.0146	0.1300
2005	-0.1372**	-0.2435***	0.2198***	0.3034***	-0.6472***	0.0230	0.1368
2006	0.3191***	-0.1702***	0.2536***	0.1363***	-0.8562***	0.0817***	0.1430
2007	0.1409**	-0.2026***	0.1622***	0.1887***	-0.6627***	0.0620***	0.1572
2008	0.1768***	-0.1743***	0.2081***	0.1859***	-0.6434***	0.0287	0.1626

Numerical flexibility increases the probability of having a new job. As from 2002, decentralisation decreases this probability whereas individual learning decreases it in 1999-2000 and thereafter increases it. The interaction effect between numerical flexibility and decentralisation is generally significant. In the first two years it decreases the probability of being employed by a new company, and later it increases this probability. Our interpretation for the first two years is that decentralisation decreases the probability of being employed by a new company, and the more numerically flexible the firm was, the larger is this reducing effect. The interaction effect between numerical flexibility and individual learning is generally negatively significant. That is, the more numerically flexible the company of 1998 was, the less important is the positive effect of individual learning.

Although the interaction effect between decentralisation and individual learning is significant in some cases, the parameter estimates are relatively small and the symbols shift in an apparently random manner from year to year. Therefore we draw the conclusion that this variable is not very important for the probability of having a new job.

Table 21 contains the corresponding results of logistic regressions performed with unemployed as response variable.

Table 21: Parameter estimates of numerical flexibility, β_{NF}, decentralisation, β_{D}, individual learning, β_{IL}, and the two-factor interaction terms between the three variables, from logistic regressions performed on each year separately, as described in equation (4), with *unemployment* as response variable.

Year	Numerical flexibility	Decentralisation	Individual learning	Num*Dec	Num*Ind	Dec*Ind	R^2_{pseudo}
1999	0.8312***	0.1645**	-0.2402	-0.1874**	0.4186**	-0.1384**	0.0416
2000	0.3678**	0.1725***	-0.8675***	-0.1861***	0.7040***	-0.0756	0.0560
2001	-0.1714	0.0764*	-0.7708***	-0.0366	0.7216***	-0.0868*	0.0518
2002	-0.4168**	-0.1083**	-0.4137***	0.1877***	0.3586**	-0.0867*	0.0383
2003	-0.4832***	-0.2205***	-0.2037	0.2516***	0.1513	-0.0587	0.0471
2004	-0.1253	-0.2106***	-0.1179	0.2325***	-0.1548	-0.0500	0.0471
2005	0.1624	-0.1208***	0.1205	0.1523***	-0.3957**	-0.0485	0.0433
2006	0.0109	-0.1323***	-0.2307	0.0515	0.1153	0.0459	0.0431
2007	0.1019	-0.0219	-0.0377	0.0478	-0.00380	-0.0516	0.0415
2008	-0.1827	-0.2076***	-0.2120	0.1965***	0.0548	0.0170	0.0466

In the first two years, having worked in a company that was numerically flexible increases the probability of unemployment. Decentralisation has the same effect. High values of both these variables reduce this impact. As from 2002, decentralisation shifts sign and instead decreases the probability of unemployment, whereas the interaction effect between numerical flexibility and decentralisation instead increases it.

Individual learning decreases the probability of unemployment in some of the earlier years, but at companies with much numerical flexibility this positive effect is smaller (i.e. there is a positive interaction effect).

Next logistic regressions are performed with disability pensioner as response variable. The parameter estimates of the organisational variables are found in table 22.

Table 22: Parameter estimates of numerical flexibility, β_{NF}, decentralisation, β_{D}, individual learning, β_{IL}, and the two-factor interaction terms between the three variables, from logistic regressions performed on each year separately, as described in equation (3), with *disability pensioner* as response variable.

Year	Numerical flexibility	Decentralisation	Individual learning	Num*Dec	Num*Ind	Dec*Ind	R^2_{pseudo}
1999	1.1603	0.6297***	-0.4582	-0.9046***	1.5715*	-0.2205	0.1313
2000	0.1969	0.1176	0.4487	-0.2438	0.5421	-0.1720	0.1368
2001	-0.2110	-0.0304	-0.0371	-0.1691	0.6562**	-0.0446	0.1526
2002	-0.2642	-0.0602	-0.0465	-0.0717	0.4049*	-0.0248	0.1574
2003	0.0703	0.0312	-0.0185	-0.1365**	0.3076	-0.0650	0.1586
2004	0.2434	0.0149	0.0931	-0.1117*	0.0184	-0.0459	0.1630
2005	0.2631*	-0.0466	0.0484	-0.1089**	-0.0261	0.0145	0.1616
2006	0.1932	-0.0610	0.00665	-0.0966*	0.0193	0.0254	0.1604
2007	0.3567**	-0.0283	0.0634	-0.1568***	-0.0952	0.0339	0.1578
2008	0.3730**	-0.0243	0.1125	-0.1594***	-0.1931	0.0403	0.1559

The results for the group of disability pensioners are rather unclear. There is no clear connection between the main effects and living on disability pension. What we do see on the other hand is that the interaction effect between numerical flexibility and decentralisation significantly decreases the risk of early retirement in many of the years. This means that those who worked in companies with both these characteristics in 1998 are less represented in this group of people, at least in some of the years.

For those on sick leave, no significant interaction effects were found. We see that decentralisation reduces the risk of being sick, which was also the case in the analysis without interaction effects. Since nothing new is found, we choose not to show this table. Nor do we show the results of the logistic regressions with *other, low income* as response variable. This is due to the fact that no clear conclusions may be drawn from these analyses.

Comparisons of the R^2_{pseudo} between the models with and without interaction effects show that there are only minor differences in model fit. The values of the Akaike information criterion, AIC, (not presented) give the same indication, which is hardly surprising since both measures are likelihood-based. This, in combination with the fact that the results are rather unclear and varying, leads us into the decision to not proceed with GEE analyses with interaction effects between the organisational variables included in the models. Nor do we make the corresponding analyses of the income development.

The Pearson correlation matrix, presented in table 23, reveals that the organisational variables are relatively uncorrelated, which may be seen as a further indication that the models without interaction effects are to be preferred.

Table 23: Pearson correlation between the organisational variables.

	Decentralisation	Numerical flexibility	Individual learning
Decentralisation	1	0.07522***	0.16540***
Numerical flexibility	0.07522***	1	0.12777***
Individual learning	0.16540***	0.12777***	1

Furthermore, keeping in mind that one should always strive for statistical models as simple as possible, we make the decision to draw our main conclusions from the models without interaction effects. Nevertheless, the results presented in this section should be seen as interesting indications of possible connections between the organisational variables.

5.5 Income Development

In this section we want to investigate whether the work organisation has an impact on the development of the future wages of the employees. At a first thought, one may prefer to only include those who are part of the two job categories in this analysis, but this would not reflect what we are interested to examine. With our analyses we aim to catch the future real income development of a randomly chosen person, given the work organisation of his or her workplace. We do not wish to analyse the income development given that he or she will have a job in the future, since such an assumption is far from realistic. Thus we choose to perform our analyses on the whole group of people included in the sample in each year, regardless of whether they work or not. We need to keep in mind that according to the definitions given in the section *Explanatory variables*, some work income is possible also for those who are included in the non job categories, which means that the income change is not automatically $-100\left(= \frac{\boldsymbol{Income_x - Income_{1998}}}{\boldsymbol{Income_{1998}}} * \mathbf{100} = \frac{\boldsymbol{0 - Income_{1998}}}{\boldsymbol{Income_{1998}}} * \mathbf{100}\right)$ for everyone in this category.

As previously mentioned, the response variable chosen for this analysis is the percentage income change from 1998 to each year respectively, *Income_percent_x_1998*. That is, the definition of the response variable changes from year to year. Multiple linear regression, as described in equation **(2)** is used to make the analyses. The results are presented in table 24. Plots of the residuals against predicted values (performed but not presented) do not reveal any structure, which indicates that the residuals are indeed independent with constant variance, and as previously mentioned, thanks to the central limit theorem, the normality assumption needs not to be checked for. Thus the model assumptions may be considered fulfilled.

Table 24: Parameter estimates of numerical flexibility, β_{NF}, decentralisation, β_{D}, and individual learning, β_{IL}, from logistic regressions performed on each year separately, as described in equation (2), with *Income_percent_x_1998* as response variable.

Year	Numerical flexibility	Decentralisation	Individual learning	R^2
1999	-0.87056***	0.40831***	0.61842**	0.0296
2000	-2.08238***	0.29807***	1.16914***	0.0572
2001	-1.37818***	0.54980***	0.51414	0.0751
2002	-1.66293***	0.30669**	0.85818**	0.0817
2003	-1.68651***	0.18169	1.80304***	0.0853
2004	-1.81454***	0.30266**	2.32260***	0.0888
2005	-0.61676	0.59482***	0.97553*	0.0989
2006	-0.29640	0.36310**	1.35713**	0.1075
2007	0.26833	0.30584	0.69338	0.1181
2008	-0.76967	0.14478	1.17867	0.1244

It is evident that generally, numerical flexibility decreases and decentralisation and individual learning increase the work income of the employees in the earlier years. Since the response variable changes from year to year, it is not possible to decide whether the impact of the organisational variables is decreasing. The fact that the coefficients of determination increase with time is most likely a consequence of the fact that as time goes by, the income differences become larger and clearer. The income change from one year to another is likely to be affected by random differences to a larger extent than the income change over a ten year period.

Since the response variable changes from year to year, a longitudinal study of the income development based on GEE, is not appropriate.

6 Discussion of the Statistical Models

In this chapter we bring some statistical issues that affect our study up to discussion.

A general remark is that the coefficients of determination, R^2_{pseudo} and R^2, are low, which indicates that the model fit is relatively poor. This is a natural consequence of the fact that the main reason for being part of a certain category is personal differences that are hard or impossible to measure and differences between individuals due to random chance. Intelligence, personal need for money, ambition, social skills, talent, health status, personality and having been in the right place at the right time are a few among a large number of characteristics that are very likely to affect people's positions on the labour market, but that cannot be taken into account with the data at hand. This fact implies that our models necessarily need to be clear simplifications of reality, and thereby, that we should expect the model fit to be relatively poor, especially in comparison with most statistical modelling of data from the natural sciences. The goal of our study is not to find the explanatory variables that best describe the work status of the individuals. Our ambition is to investigate whether or not the work organisation is one of the factors that influence the employees' future chances on the labour market. For this reason, the fact that the coefficients of determination are low should not be seen as a decisive problem, although larger values are of course to be preferred.

The fact that the main purpose of our study is not to find the best fitting model is also a crucial argument for our choice to not use any model selection procedure, such as stepwise regression, to make the selection of explanatory variables. One may argue that, because of the possibility of collinearity, this choice involves a risk of not getting significance in important variables. Fortunately, this seems not to be a problem in our case. Comparisons of the regression results of straightforward multiple regression and multiple regression where backward elimination is applied, on a sample of models, show that generally the significant variables in the former case are the same as the variables left in the model in the latter case (performed but not presented). Furthermore e.g. Flom and Cassell (2007, pp. 3) argue that "there is nothing intrinsic in multiple regression that requires only significant independent variables to be included". They point out that non-significant variables may be interesting to keep in the model if (in this case economic) theory suggests that they are important, partly since they may affect the parameter estimates of the other variables, and that in many cases the size of the parameter estimates are of greater interest than their statistical significance. In their opinion, substantive expert knowledge is an "excellent alternative that is often overlooked" when it comes to variable selection. This is applicable in our case, since the background variables are chosen with prior research in labour market economics in mind. However we want to point out that, in our analyses, the great majority of the register-based explanatory variables are in fact significant in most cases.

Another issue that should be mentioned is the one of multiple testing. By performing a large number of tests, the probability of getting significant results only by pure chance increases. A significance level of 5 percent means that one allows that in 5 percent of the tests performed we get significance only by chance. Thus we should have in mind that a few of our significant results may not be truly significant. However, since our study focuses on the general trends rather than on analysing the results in specific years, we are willing to accept this fact.

Furthermore we want to mention the weakness of the data upon which our study is based, apart from the problem with the representativeness that was already brought up to discussion. To start with, the explanatory variables of main interest, i.e. the organisational variables, are based on the results from a sample survey. This means that the results are highly dependent on the person who completed the questionnaire. Misinterpretation or misreading of the questions and choosing the wrong option by accident or by design are generally common problems. Secondly, our definitions of the response and some of the explanatory variables may be discussed. Although we do believe that the general trends that we have found in this study are good descriptions of reality, we need to remember that our results depend on our choice of categorisation of the occupational categories and our way of defining the explanatory variables.

Finally, it should be said that there are several other options for analysing our data. Survival analysis has been considered, which would have given us the opportunity to study how the time until falling into each one of the occupational categories depends on the organisational variables. That is, how does the intensity with which people become e.g. unemployed differ over time. Multivariate multiple logistic regression, with all categories of occupation in the same model, is another option that has been brought up to discussion. Both methods should be seen as interesting possibilities for future research.

7 Conclusions

Having worked in a numerically flexible firm does not only decrease the probability of still being employed by that very firm, but also the probability of having a job at all. Furthermore, and probably partly as a consequence of this, it also decreases the wage development. A positive connection between numerical flexibility and being employed by another firm is also found, although this most likely is due to the fact that people who are forced to quit their former jobs tend to find new ones to a larger extent than others.

Several positive outcomes of decentralisation are found. It clearly decreases the probability of being on sick leave or early retired later in life. Furthermore, the probability of having a job as well as the future income development is larger. The more decentralised the firm is, the longer the employees tend to stay within it.

Individual learning is positive for both the future wage as well as the probability of having a job in the future. It increases the probability of staying within the firm and decreases the probability of being employed in another one. Besides this, it diminishes the risk of unemployment.

In short, the most evident, but also most important, conclusion that may be drawn from our study is that the work organisation does have an influence on the future careers and lives of the employees.

Appendix

Table A1: Specification of LISA variables used for the categorisation into the different categories of occupation

Category	Variables from the LISA database
Employed within the same firm	PeOrgNr*
Employed within another firm	PeOrgNr*, DekLon
Unemployed	ArbLos, AmPol, DekLon
On sick leave	SjukSum_Ndag
Early retired	ForTid
Student	StudMed, DekLon
Other, low income	DekLon
Other, higher income	DekLon

**Changes of the corporate identification number from year to year and expansions of firms into multiple firms are accounted for by the use of the FAD database.*

Description of the Composition of the Work Organisation Indices

In this section the composition of the indices numerical flexibility, decentralisation and individual learning is described. For each variable, we start by presenting the numerical values assigned to the answers of the relevant questions. Thereafter, the formula for calculating the index is given.

Numerical flexibility

Question	Assigned the value one if the answer is:
T22a	Yes
T22b	Yes
Q19	Yes
Q20	More than 20 percent
Q22	Yes
Q24	Yes to at least one of the possible options
Q25	More than 5 percent

We define the variables

$$ROTATE = \frac{T22a + T22b}{2}$$

and

$$TEMPORARY = Q19 + Q20$$

The index variable is defined as

$$NUMERICAL\ FLEXIBILITY = \frac{1}{8}ROTATE + \frac{1}{8}Q22 + \frac{1}{2}TEMPORARY + \frac{1}{8}Q24 + \frac{1}{8}Q25$$

Decentralisation

Question	Value
T13	1 if the answer to at least one of the questions a)-e) is *Yes, normally*, and 0.5 if the answer to at least one of them is *Yes, in special cases*
T17 (individual planning)	1 if the answer to at least one of a) and b) is *Individual employee*
T17 (individual control)	1 if the answer to at least one of c) and d) is *Individual employee*
T17 (team work)	1 if the answer to at least one of the questions is *Work teams*
Q21	1 if the answer to a) is less than 90 percent

We define the variable

$$TEAM = \frac{T13 + T17(team\ work)}{2}$$

The index variable is defined as

$$DECENTRALISATION = TEAM + T17(individual\ planning) + T17(individual\ control) + Q21$$

Individual learning

Question	Assigned the value one if the answer is:
T20	Yes
T22	Yes to at least one of c)-e)
Q21	At least 51 percent

The index variable is defined as

$$INDIVIDUAL\ LEARNING = \frac{1}{2}T20 + \frac{1}{4}T22 + \frac{1}{4}Q21$$

Further Discussion on Generalized Estimating Equations

The generalized estimating equations approach was first introduced by (Liang & Zeger, 1986) as an extension of the generalized linear model for correlated data. We choose not to specify the equations or the theoretical aspects of the method in this report. The reader is instead referred to e.g. (Diggle, Heagerty, Liang & Zeger, 2002, pp. 146-160) where a thorough exposition of the theoretical details of the generalized estimating equations approach is found. However, in short, the parameter estimates are calculated by solving generalized estimating equations numerically by iteration. A so called working correlation matrix needs to be specified to describe the correlation between outcomes in the different years, to be used in the first iteration step. An overview of different possibilities and suggestions on which working correlation matrix to use in which situation may be found in e.g. Katz (2011, pp. 195). However, one can show that in generalized estimating equation analysis “the final parameter estimates are not generally dependent on the choice of working correlation matrix” (Katz, 2011, pp. 197), i.e. they are consistent. Katz (2011, pp. 197) points out that for this reason, for example the computational difficulty of the different working correlation matrices should be taken into account and that the so called exchangeable structure, which assumes that data has equal correlation within each cluster (individual) is a common choice even for data that is known to not fulfil this assumption. For simplicity, our choice in this study has thus been the exchangeable working correlation matrix.

References

Agresti, A., (2002), *Categorical Data Analysis*, John Wiley & Sons

Atkinson, J., *Manpower Strategies for Flexible Organizations,* Personnel Management, August 1984, pp. 28-31.

Diggle, P.J., Heagerty P., Liang K-Y, Zeger, S.L. (2002), *Analysis of Longitudinal Data*, Oxford University Press

Flom, P.L., Cassell, D.L. , (2007) *Stopping stepwise: Why stepwise and similar selection methods are bad, and what you should use*, SAS Institute

Institutet för tillväxtpolitiska studier, ITPS (2001), *Enterprises in Transition – Learning Strategies for Increased Competitiveness*

Katz, M. H., (2011), *Multivariable Analysis – A Practical Guide for Medical and Public Health Researchers*, Cambridge University Press

Liang, K.Y., Zeger, S.L., *Longitudinal Data Analysis Using Generalized Linear Models*, Biometrika, Vol. 73, No. 1 (Apr. 1986), pp. 13-22

Lindgren, B.W. , (1993), *Statistical Theory*, Chapman & Hall/CRC

Menard, S., *Coefficients of Determination for Multiple Logistic Regression Analysis*, The American Statistician, Vol. 54, No. 1 (Feb., 2000) pp. 17-24

Statistiska centralbyrån (SCB), Bakgrundsfakta, Arbetsmarknads- och utbildningsstatistik 2011:4 *Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier (LISA), 1990-2009*.

Statistiska centralbyrån (SCB), *Branschindelning: SNI 2007, SNI 2002, SNI 92* http://www.scb.se/Pages/List___257409.aspx

Statistiska centralbyrån (SCB) (2011), *Learning Organisations Matter*

Sundberg, R., (2009), *Lineära Statistiska Modeller*, Stockholm University