



AALBORG UNIVERSITY
DENMARK

Aalborg Universitet

Documentation on the DISKO4-IDA Merge and the Creation of the Panel Dataset DISKO2-DISKO4

Timmermans, Bram

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Timmermans, B. (2008). Documentation on the DISKO4-IDA Merge and the Creation of the Panel Dataset DISKO2-DISKO4. (pp. 1-30). Department of Business Studies, AAU: Institut for Erhvervsstudier, Aalborg Universitet.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- ? Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- ? You may not further distribute the material or use it for any profit-making activity or commercial gain
- ? You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Working Paper Series

Department of Business Studies

No. 24, 2008

Documentation on the DISKO4-IDA Merge and the
Creation of the Panel Dataset DISKO2-DISKO4

By

Bram Timmermans

Documentation on the DISKO4-IDA Merge and the Creation of the Panel Dataset DISKO2-DISKO4

Bram Timmermans

IKE Group

Department of Business Studies

Aalborg University

Fibigerstræde 4

DK-9220 Aalborg Ø

E-mail: bram@business.aau.dk

December 2008

Abstract:

This document is a description of the merging process of the IDA-DISKO data, based on the latest DISKO survey (DISKO4), located at the Statistics Denmark. It provides a guide for researchers on how the different datasets are created and which variables are present in the datasets as well as the way the different files should be used. Two types of datasets have been created using the different datasets in the Danish Integrated Database for Labour Market Research (IDA). First, there is a database using all the observations from DISKO4 and IDA for the years 2000-2004. Second, a panel dataset for all those firms present in DISKO2 and DISKO4 including all the variables from IDA covering the period 1997-2004.

ISBN 9788791646287

1 Introduction

The purpose of this document is to describe the merging process and content of two newly created datasets. Starting point of these datasets is the latest DISKO survey, called DISKO4, conducted in 2006 by Statistics Denmark on behalf of four research groups (IKE, CARMA, CIP and CCWS) at Aalborg University. This survey is a follow up of previous surveys in the so-called DISKO project, which have been conducted in 1996 (DISKO1), 2001 (DISKO2), and 2004 (DISKO3), that focusses on organizational and technological change in Danish firms.

Since the firms that participated can be linked to Danish government register data (IDA) it is possible to link both firm and employee information to these DISKO4 firms. As a result important research questions can be addressed, including labour market research, intra-organizational research, industrial dynamics, innovation based analysis. This directly illustrates the strength of DISKO. The two databases that will be constructed using the DISKO4 survey are:

- A merge between DISKO4 and IDA based on all respondents in the DISKO4 questionnaire survey
- A panel-dataset selecting those firms that participated in both DISKO2 and DISKO4 and merging these firm together with IDA.

Reichstein and Vinding (2003)¹ have written a document on an earlier merging process between DISKO1 and DISKO2. This merging process will be different since they only merged the data together based on the largest plant for each firm. In this new merging process all plants will be included in order to provide the researcher more flexibility in using the dataset. Partly due to this flexibility and comprehensibility of the datasets is it undesirable to create two large datasets. Instead I have chosen to create smaller datasets divided between firm and person based information for each year, which can be merged together easily according to the researcher's needs.

This document is structured as follows. Section 2 will present the different raw datasets used in the merging process. In addition the variables that will be used are described and a list will be provided of questions that are similar in both DISKO2 and DISKO4. Section 3 discusses the merging process, provides a graphical representation of this merge and gives an overview of the number of observations.

2. Construction of the Datasets

¹ Reichstein, T. and A. L. Vinding (2003). Documentation of the IDA-DISKO Database. Aalborg, Department of Business Studies - IKE Group. Aalborg University.

The above-mentioned two datasets are constructed by using multiple but compatible databases. These databases are:

- DISKO2/PIE - A questionnaire survey concerned with technological and organizational change in the time period 1998-2000.
- DISKO4 - A questionnaire survey concerned with technological and organizational change in the time period 2003-2005.
- IDA - Integrate Database of Labour Market Research divided in:
 - FIDA - key dataset to connect the plant identification number (LBNR) to the firm identification number (JURNR).
 - IDAPERSON - Personal information on the individuals.
 - IDAANSAT - Employee information on the individuals based on their primary workplace.
 - IDAARBSTED - Plant level information.
- FIRMAGF - Firm level accounting data (only available since 1999 and for this reason not used in the panel dataset).
- REGNSKAB - Firm level accounting data.

2.1 DISKO2

DISKO2, also referred to as PIE, is a survey data which got carried out in the winter of 2001 as part of the LOKE-Project². The survey was send out to all firms in the private sector with 25 or more employees, supplemented with a stratified proportional sample to firms with 20-25 employees. Those firms that did not respond on the postal questionnaire were contacted and asked for a telephone interview. The respondents are primarily high-level executives complemented with a survey among the employees in the same firms.

This survey asks for changes regarding the period 1998-2000 on issues such as major organizational change, change in firm's management structure, development in employment conditions, change in the character of work, employee training, innovation, co-operation and the introduction of new technologies. In addition questions where asked on the state of things in 2001 and the impact of the above-mentioned changes.

During DISKO2 two samples were carried out. One sample targeting management and another one targeting employees, which respectively resulted in 2,007 and 473 returned questionnaires. In addition the results of a large ICT survey were added into this DISKO2 dataset resulting in a total of 2,738 observations. The questionnaires filled in by the management team are those that are in line with DISKO4 and for this reason the panel sample will be constructed using DISKO4 and these 2,007 observations.

²See for a short introduction <http://www.business.aau.dk/pie/> More information can also be found in a document prepared by Reichstein and Vinding (2003)

2.2 DISKO4³

DISKO4 is the latest DISKO survey conducted in 2006 by Statistics Denmark on behalf of four research groups (IKE, CARMA, CIP and CCWS) at Aalborg University. Just like the previous DISKO survey will DISKO4 take its point of departure in firms that participated in earlier DISKO survey augmented with firms in those industries that can be found in the firm statistics of 2004.

Eventually, the questionnaire has been sent out to a total of 4,136 danish firm that were selected on a number of criteria. The first criteria was to include those firms that participated in previous DISKO surveys. In total 1,552 firms were identified as still being operational. The second criteria was to include all the firms with more than 100 employees and finally a unbiased selection of firms in the size category 20-49 and 50-99 employees. Table 1 provides an overview of the distribution of these 4,136 firms based on size and industry.

Table 1: The distribution of the sample based on industry and size.

Industry	Size Groups				Total
	< 20 employees	20-49 employees	50-99 employees	Over 100 employees	
Manufacturing	37	342	415	576	1370
Construction	58	217	125	78	478
Trade and Repair	79	401	294	266	104
Hotel and Restaurants	6	45	29	23	103
Transport	24	107	97	124	352
Financial services	1	21	33	80	135
Business services	24	197	161	201	583
Culture and Sports		24	32	19	75
Total	229	1354	1186	1367	4136

Although the DISKO surveys have a general requirement to include firms larger than 20 employees there are 229 firm that are smaller. These are firms that participated in the previous DISKO surveys and during the years became smaller. Despite this fact a decision has been made to include them in the large sample.

Statistics Denmark received 1,781 questionnaires from these 4,136 firms 1,781, 6 of these questionnaire turned out to have been answered already so in total there are 1.775 unique questionnaires resulting in a response rate of 42,9 percent. Please note that in the recent DISKO4 database only 1,770 questionnaires are included, the reason is unknown to the

³ For a more thorough description of the DISKO4 there is a method report available in Danish produced by Statistics Denmark

author. The distribution of the answers based on the same categorization as in Table 1 are presented in Table 2. The questionnaire itself can be found in Appendix B.

Table 2: The distribution of the answered questionnaire based on industry and size.

Industry	Size Groups				Total
	< 20 employees	20-49 employees	50-99 employees	Over 100 employees	
Manufacturing	14	190	177	210	591
Construction	26	97	53	38	214
Trade and Repair	34	169	116	98	417
Hotel and Restaurants	2	16	14	5	37
Transport	12	45	38	48	143
Financial services		9	17	43	69
Business services	10	91	711	97	269
Culture and Sports		12	10	8	30
Total	98	629	496	547	1770

2.3 Panel DISKO2-DISKO4⁴

When aligning the 1,770 answered questionnaires from DISKO4 with the 2,007 comparable observations from DISKO2 there are 791 firms that participated in both surveys. For these firms it may be interesting to emphasize the question which are more or less similar in both surveys. These questions are:

- Question 8 in DISKO2 and Question 2 in DISKO4 concerning the organization of work.
- Question 5 in DISKO2 and Question 3 in DISKO4 concerning organizational change.
- Question 6 in DISKO2 and Question 4 in DISKO4 concerning the objectives of organizational change.
- Question 35 in DISKO2 and Question 8 in DISKO4 concerning introduction of product innovation
- Question 36 in DISKO2 and Question 9 in DISKO4 concerning market location of the new product innovation
- Question 41 in DISKO2 and Question 15 in DISKO4 concerning pressure of competitiveness
- Question 42 in DISKO2 and Question 16 in DISKO4 concerning development of closer relationships with external actors
- Question 25 in DISKO2 and Question 19 in DISKO4 concerning the change in the character of work
- Question 26 in DISKO2 and Question 20 in DISKO4 concerning management efforts in continuous skill development of employees

⁴ DISKO3 was a different type of questionnaire and would not have any comparability with other DISKO surveys

- Question 27 in DISKO2 and Question 21 in DISKO4 concerning the number of employees that participated in courses
- Question 28 in DISKO2 and Question 22 in DISKO4 concerning the importance for a firm's competitiveness to develop employees' skills.

2.4 IDA

IDA makes a distinction between two types of register data, on the one hand person and employee information on individuals legally residing in Denmark and on the other hand plant information and firm accounting data on Danish firms. The variables obtained in these datasets are described below for each sub-dataset.⁵

2.4.1 FIDA

DISKO4, just as the other DISKO surveys, is firm level data while IDA, with the exception of the accounting data, is plant level data. FIDA is used to identify which plant belongs to which firm. In the case of 2003 and 2004 there is a special plant code (DSKODE) which needs to be transformed in order to connect a plant identification number to it.⁶ In earlier versions of FIDA the plant identification number (LBNR) is provided by Statistics Denmark. Keep in mind that in both situation there might be some inaccuracy, especially regarding the secondary workers. Statistics Denmark has indicated that they are not able to accurately determine a persons secondary workplace. An additional feature of FIDA makes it possible to connect individuals to the plant that have this plant as a secondary workplace. These variables are included in both the person and firm based datasets. Variables extracted and created from this FIDA are:

- The type of job the individual possess in the plant⁷
- Year
- Unit code for the plant (DSKODE starting with a letter are fictional plants, they have no fixed address. Plant information is not available for these plant codes)
- Personal number
- Plant identification number
- Firm identification number (CVR number)

⁵ Appendix A lists these variables including the coding. For a more thorough description visit the website of Statistics Denmark www.dst.dk

⁶ In FIDA2003 and FIDA2004 a DSKODE is used to identify which individuals work together in the same unit. This unit is, according to DST, a plant. In order to find the plants identification number (LBNR) that are connected to each firm identification number (JURNR) there is a need to identify which DSKODE corresponds to which LBNR. In order to do this the following steps have been undertaken. (1) In IDAANSAT one finds the person number (PNR) and the LBNR to which each individual is connected. Keep in mind that IDAANSAT only gives employee information for the primary workplace. (2) In FIDA three variables are important, i.e. PNR, PSTILL and DSKODE. PSTILL indicates whether or not it is the primary workplace of the individual. So, for each DSKODE a individual will be chosen that has this DSKODE as the primary workplace. (3) For FIDA it is known which DSKODE is connected to a primary worker and in IDAANSAT it is known which LBNR is connected to this primary worker. This allows to connect a LBNR to a DSKODE. Everybody who has the same DSKODE will thus have the same LBNR.

⁷ 0 is secondary workplace

2.4.2 IDAPERSON

IDAPERSON contains personal information on all individual that are connected to the firms. The variables extracted from this dataset are:

- The age of the individuals
- The gender of the individuals
- The highest obtained general education
- The highest degree of education of the individuals
- The end date of the highest education
- The education the individual is currently following
- Citizenship of the individual
- The individuals A-income during leave of absence due to illness
- The individuals B-income during leave of absence due to illness
- The degree of unemployment during the year
- The work experience of the individual before 1979
- The work experience of the individual from 1980
- changes in the identify of the plant
- The title of the individual as a municipality employee
- Period of unemployment of the individual
- Net salaries of the individual
- Gross income of the individual
- Taxable income
- Profit for firm owner
- Aggregation of degree of unemployment from 1980
- GEO-kode 1 for the location of residence
- GEO-kode 2 for the location of residence
- Municipality code where the individual lives
- Municipality code where the individual is obtaining an education

2.4.3 IDAANSAT

IDAANSAT contains employee information. This information is only present for the primary workplace of the individual. Whenever the DISKO firm is not the individuals primary workplace information is not present in the dataset. The variables extracted from IDAANSAT are:

- The year the individual was hired to work for the plant
- The position of the individual in the plant
- Association to the primary workplace
- Indicator whether the individual is employer or employee
- Does the individual leave one workplace for another the following year
- Has the individual left one work place for another the past year
- Primary work title of the individual

- The type of job the individual possesses
- The hourly wage rate for the individual
- The quality of the valuation of the individual's hourly wage rate
- Distance between plant location and location of residence

2.4.4 IDAARBSTED

IDAARBSTED contains plant-level information. A firm might consist out of several plants for the participated firms all plants are included. The variables extracted from IDAARBSTED are:

- Number of employees in the third week of November (full time equivalent)
- Number of employees in the third week of November (head count)
- Number of employees that have been working there throughout the entire year (head count)
- Sideline occupations
- Identity of workplace the year before
- Identify of the workplace the year after
- Municipality code of the workplace's location
- Industry code of the plant
- Identification of the type of ownership

2.4.5 FIRMAGF (*Accounting Statistics 1*)

FIRMAGF contains general firm information including accounting statistics. This dataset offers more information than REGNSKAB. Since the same data is not present before 2000 this information will not be valuable for the panel dataset. The variables extracted from this dataset are:

- Industry code of the entire firm
- Number of employees in the entire firm (full time equivalent)
- Number of employees in the entire firm (headcount)
- Fixed assets
- Equity
- Value added growth
- Exports
- Turnover
- Purchases
- Regular Profit
- Profit after tax
- Gross Profit
- Balance Sum
- Wages pensions and other social security expenses
- Method for calculating turnover
- Municipality code

- Function code

2.4.6 REGNSKAB (*Accounting Statistics 2*)

REGNSKAB is available for the years that covered by the panel dataset. Although it does not contain so many variables as in FIRMAGF it is easier to compare the variables. The variables in this dataset are:

- Yearly Profit
- Number of employees in the entire firm (full time equivalent)
- Fixed Assets
- Total Assets
- Equity
- Current Assets
- Turnover
- Liabilities

3 The merging process

This Section will discuss the merging process of the two datasets. Each of these sets have a folder including the smaller sub datasets that need to be merged together and the characteristics of each datasets. An example of a merging process is provided.

3.1 The DISKO4-IDA merge

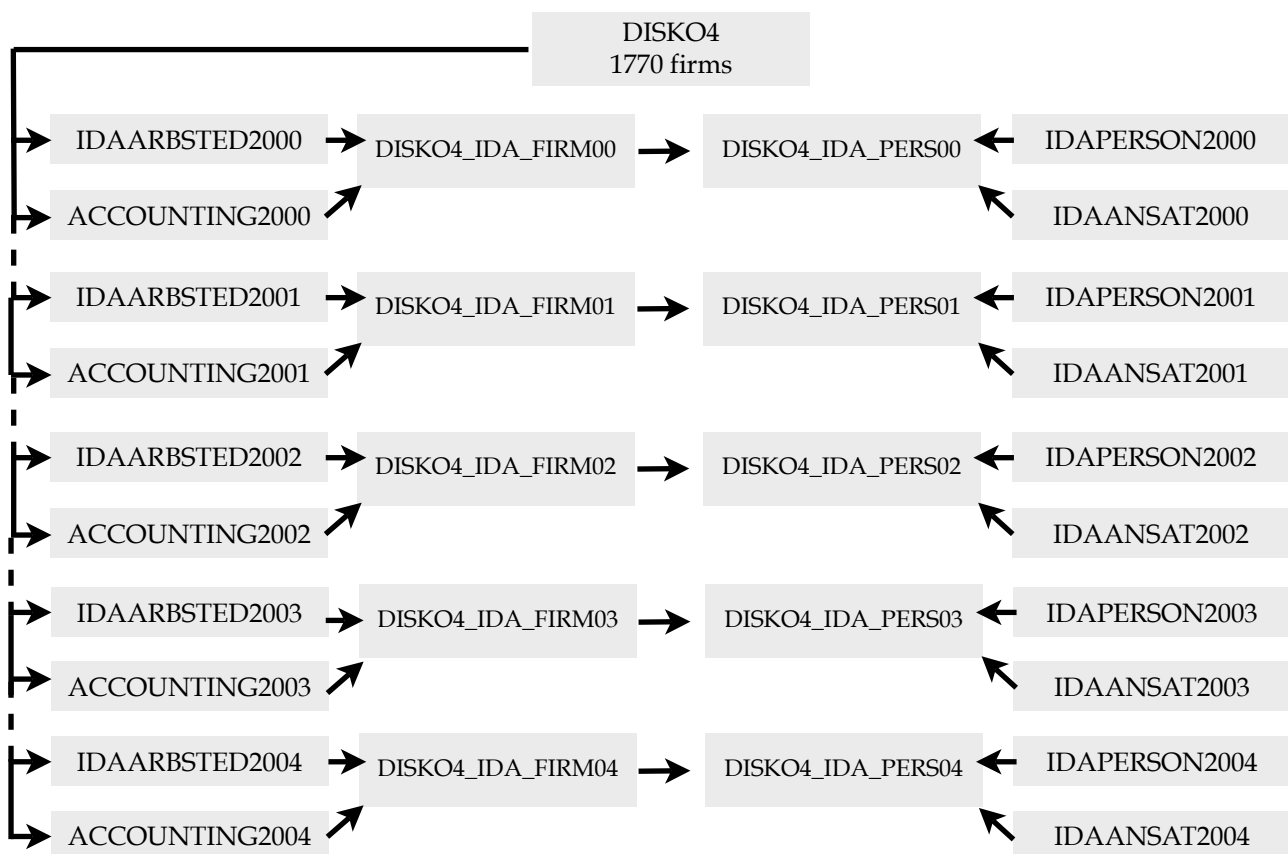
In the explanation of DISKO4 it was already made clear that there are two types of question asked. One being questions regarding the time period 2003-2005 and other targeting the situation at the time the questionnaire was handed out, i.e. the year 2006. For that reason all these years have to be included in the merging process. Due to a time lag in the available data in the current year there is only data available until 2004. Currently DISKO4 will be merged with IDA data from 2000, 2001, 2002, 2003 and 2004 since this are the years covered by the survey in addition there is information available in the period up to the period the questionnaires covers. Of course can the researcher connect other years if there is a need to do so and access to the data is possible. Since several employee information is provided to Statistics Denmark in November of any given year can one treat the information of a specific year as the start position of the year after.

As can be seen in Figure 1 is the data divided in two groups, i.e. a firm and plant level dataset on the left (Filename: DISKO4_IDA_FIRMxx), and a person based dataset on the right (Filename: DISKO4_IDA_PERSxx), for each year. Firm and plant based data are the variables obtained from accounting data (FIRMAGF and REGNSKAB) and IDAARBSTED for the firms, including all the plants, that participated in DISKO4. In this dataset all information is based on plant level. This means that firm level data is provided for each individual plant. Whenever a researcher only wants firm level data one should omit plant level data and make sure there are no duplicates in the dataset. The person based dataset

contains the variables obtained from IDAPERSON and IDAANSAT. The firm and plant level data is used to make sure that the individuals in this dataset are those connected to the firms in that participated in DISKO4. In total there are seven⁸ datasets that can be merged according to the need of the researcher. The DISKO4 survey dataset is also included separately. The variables obtained from the raw datasets and their coding in the dataset are described in Appendix A.⁹ Some of these variables are not the same or simply not present for all years because:

- there is a recoding, e.g. change in industry coding between the years;
- variables giving information on the situation in 2005 (information that is not yet available); or
- the variable is not present in the raw dataset of that year.

Figure 1: Graphical presentation of the formation of the DISKO4-IDA



One glimpse at Table 3 already shows that the number of firms is not similar with the 1,770 returned questionnaires of the DISKO4 survey. Some firms are apparently not registered in IDA. The new dataset will thus have information on 1,634 firms in 2000, 1,686 firms in 2001, 1,735 firms in 2002, 1,767 firms in 2003 and 1,762 firms in 2004. Each of these firms have one or more plants. The total number of plants in the dataset is indicated in Table 3. Keep in mind that there are plants whose identification number (LBNR) is not

⁸ These 7 datasets and their filename are listed in Appendix C

⁹ A more thorough description can be found on the website of Statistics Denmark www.dst.dk

known. These plants are grouped together under one plant identification number, i.e. 0000000000. Within this group there are also the earlier mentioned fictional workplaces (DSKODE starting with a letter). If the researchers interest is to focus on only one plant, e.g. the largest one, one should be aware not to use this group but a plant with an actual identification number. The last column of Table 3 shows the number of individuals that are connected to these firms in respectively 2000, 2001, 2002, 2003, and 2004.

Table 3: Number of observations in firms plants and employees based on DISKO4 and IDA

Year	Firms	Plants	Employees
2000	1634	5448	283524
2001	1686	6024	314969
2002	1735	6283	329172
2003	1767	6356	339540
2004	1762	6383	338056

As noted earlier the researcher can merge any combination of these seven datasets according to their needs. Box 1 provides a description of possible merges using the SAS procedure.

Box 1. The different merging procedures

Merging DISKO4 with firm data

```
data a;
set DISKO4_IDA_FIRMxx (keep=[...]);
run;
```

```
data b;
set DISKO4 (keep=[...]);
run;
```

```
data c;
merge a b;
by jurnr;                                (sort the databases accordingly)
run;
```

Merging DISKO4 with person data

```
data a;
set DISKO4_IDA_PERSxx (keep=[...]);
run;
```

```
data b;
set DISKO4 (keep=[...]);
run;
```

```
data c;
merge a b;
by jurnr;                                (sort the databases accordingly)
run;
```

... *Box 1. continued*

Merging DISKO4 with person data and firm data

```
data a;  
set DISKO4_IDA_PERSxx (keep=[...]);  
run;
```

```
data b;  
set DISKO4_IDA_FIRMxx (keep=[...]);  
run;
```

```
data c;  
merge a b;  
by jurnr lbnr;  
run;
```

(the correct plant information should be connected to the person that work for this plant, this is the reason why the merge should be on both JURNR and LBNR)

```
data d;  
set DISKO4 (keep=[...]);  
run;
```

```
data e;  
merge c d;  
by jurnr;  
run;
```

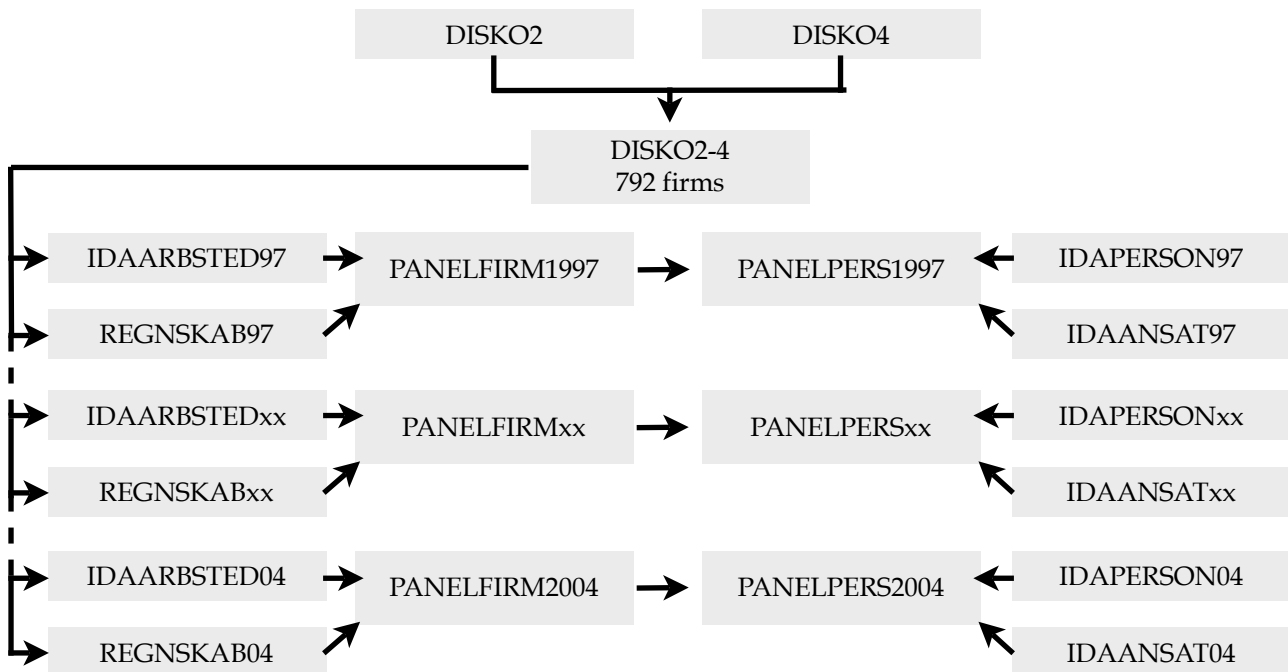
3.2 The Panel Dataset

As mentioned in Section 2 are there a total of 791 firms that participated in both DISKO2 and DISKO4. In the DISKO2 dataset there is, however, one firm that filled in two different questionnaires resulting in 792 observations. DISKO2 has just as DISKO4 two types of question, being the time period 1998-2000 and for the year the survey was conducted, 2001. The years obtained from IDA in the merger of the panel data contains information from years 1997, indicating the starting position, until 2004.

Figure 2 shows the merging process that has taken place to create the panel dataset. This panel dataset is, just as the large DISKO4-IDA dataset, divided in two groups i.e. a firm and plant level dataset (Filename: PANELFIRMxx) and a person based dataset (Filename: PANELPERSxx). The firm and plant level database is constructed using the plant information from IDAARBSTED for the years 1997-2004 only using the variables that are present in all the years. In addition accounting statistics from REGNSKAB since FIRMAGF only goes back to 1999 and thus cannot be used for all years. REGNSKAB is present for the entire time period of the panel.¹⁰

¹⁰ These datasets and their filename are listed in Appendix C

Figure 2: Graphical presentation of the formation of the panel dataset



For overview purposes the person and firm level databases are kept separately for each individual year allowing the researcher to obtain those variables needed for the longitudinal analyses. The variables present in the datasets are listed in Appendix A, including their coding. It is recommended to rename the variables of each yearly database when comparing them to another year. If one would for example like to compare the size of the firm in terms of employee head count between 1997 and 2000 one could rename the ANTNOV variable in ANTNOV97 and ANTNOV00 respectively. An example of this merging process is described in Box 2.

Box 2. The merging procedure for the panel dataset

```

data a;
set PANELFIRM1997 (keep=[...]); rename [...] = [...]97
run;

data b;
set PANELFIRM2000 (keep=[...]); rename [...] = [...]00
run;

data c;
merge a b;
by jurnr;
run;
  
```

Table 4 indicates the number of firms plants and employees that are to be found in the separate yearly databases. As indicated there is no firm information available on 31 firm in 1997, on 6 firms in 1998 and 1999, and 1 firm in 2002. Comparing these numbers with Table 3 one would notice the relative low number of plants. It shows from the data that this is due to the lower number of plants connected to the firms that answered the questionnaire. Be aware of the fact that in DISKO2 there is one firm with two questionnaires.

Table 4: Number of observations in firms plants and employees based in the different panel data

Year	Firms	Plants	Employees
1997	760	1500	99240
1998	785	1579	101249
1999	785	1672	104772
2000	791	1715	108325
2001	791	1691	111156
2002	790	1733	104856
2003	791	1736	100723
2004	791	1778	101220

APPENDIX A.

Coding and description of all the variables including the indication in which dataset they are available.

		2000-2002	2003	2004	Panel
FIDA					
AAR	Year of observation	x	x	x	x
DSKOD	Unit code for the plant (DSKOD starting with a letter are fictional workplaces, meaning no address etc, therefore also no plant information regarding these codes)		x	x	
JURNR	Firm identification number (CVR nr)	x	x	x	x
LBNR	Plant identification number	x	x	x	x
PNR	Person number	x	x	x	x
PSTILL	The type of job the individual possesses (0 means secondary workplace)	x	x	x	
REAL_ARBSTED	Fictional workplace or not (SEE DSKOD)		x	x	
DISKO1	Did the firm participate in DISKO1				x
DISKO2	Did the firm participate in DISKO2				x
DISKO3	Did the firm participate in DISKO3				x
DISKO4	Did the firm participate in DISKO4				x
PERSON					
ADAGP	The Individuals A-income during leave of absence due to illness	x	x	x	x
ALDER2	The age of the individuals	x	x	x	x
ALMFSP	The highest obtained general education	x	x	x	x
ARLEDGR	The degree of unemployment during the year	x	x	x	x
BDAGP	The Individuals B-income during leave of absence due to illness	x	x	x	x
BOPGEOK1	GEO-kode 1 for location of residence	x	x		
BOPGEOK2	GEO-kode 2 for location of residence	x	x		
BOPKOM	Municipality code for the location of residence	x	x	x	x
BRINDK2	Gross income	x	x	x	x
ERHVER	The work experience of the individual from 1980	x	x	x	x
ERHVER79	The work experience of the individual until 1979	x	x	x	x
HFAFGTP	End date of highest degree of education	x	x	x	x
HFFSP	The highest degree of education of the individual	x	x	x	x
IGFSP	Education the individual is currently following	x	x	x	x
KON2	Gender	x	x	x	x
LEDPERI	Period of unemployment for the individual	x	x		
LONIND	Wage	x	x	x	x
OVSKEVI	Profit of firm	x	x		
PSTILL2	The type of job the individual possesses (when secondary workplace this value has been changed in a 0)	x	x	x	x
SKPLIND2	Taxable income	x	x	x	x
STATKOD	Citizenship	x	x	x	x
SUMGRAD	Aggregation of degree of unemployment from 1980	x	x	x	x
UDDKOM	Municipality code for the location of education whenever this individual is currently following an education	x	x		

ANSAT					
ANSAAR	The year the employee was hired in the plant	x	x	x	x
ANSXFREM	Does the individual leave one plant for another the following year	x	x		
ANSXTILB	Has the individual left one plant for another the past year	x	x	x	x
DFKSTIL	Position of the individual	x	x		
KMAFST	Distance between workplace and residence	x	x		
PJOB	Jobtype	x	x	x	x
TILKNYT	Association to the primary workplace	x	x	x	x
TIMELON	Hourly wages	x	x	x	x
TLONKVAL	The quality of the valuation of the individuals hourly wage rate	x	x	x	x
TYPE	Employee or employer	x	x	x	x
ARBSTED					
AARSVRK	Number of employees in the plant(full time equivalent)	x	x	x	x
ANTAAR	Number of employees that have been working for the plant in the last year (head count)	x	x	x	x
ANTNOV	Number of employees in the plant(head count)	x	x	x	x
ANTNOVBI	Sideline occupations	x	x	x	x
ARBKOM	Municipality code for the location of the plant	x	x	x	x
BRANCHE1	Industry code for the plant	x	x		x
BRANCHE03	Industry code for the plant		x	x	
EJERKO	Type of ownership	x	x	x	x
IDFREM	Identity of the workplace forward in time		x		
IDTILB	Identity of the workplace back in time	x	x	x	x
FIRMAGF					
GF_AARE_1	Yearly Profit after tax	x	x	x	
GF_AARSV_1	Number of employees in the entire firm (full time equivalent)	x	x	x	
GF_AAT_1	Fixed assets	x	x	x	
GF_ANSATTE_1	Number of employees in the entire firm (head count)	x	x	x	
GF_AT_1	Balance sum	x	x	x	
GF_BAV_1	Gross profit	x	x	x	
GF_BRANCHE_03	Industry code according to the 2003 classification		x	x	x03-04
GF_BRANCHE_93	Industry code according to the 1993 classification	x			x 97-02
GF_EGUL_1	Equity	x	x	x	
GF_EKSP_2	Total Export	x	x	x	
GF_FUNK_KODE_1	Function code	x	x	x	
GF_KOB_1	Purchases	x	x	x	
GF_KOM_KODE_1	Municipality code for the location of residence	x	x	x	
GF_ILGAGMV_1	Wages, pension and other social security expenses	x	x	x	
GF_OMS_2	Turnover	x	x	x	
GF_OPR_OMS_1	Method for calculating the turnover	x	x	x	
GF_RFEP_1	Regular profit	x	x	x	
GF_VIRKFOD_1	Organizational form code	x	x	x	
GF_VTV_3	Value added growth	x	x	x	
REGNSKAB					
AARE	Yearly profit after tax	x	x	x	x
AARSV	Number of employees in the entire firm (full time equivalent)	x	x	x	x

AAT	Fixed assets	x	x	x	x
AT	Total assets	x	x	x	x
EGUL	Equity	x	x	x	x
OMAT	Current assets	x	x	x	x
OMS	Turnover	x	x	x	x
PAST	Liabilities	x	x	x	x

Appendix B: The DISKO4 Questionnaire

Questionnaire for the management
about organization, employment and development activities

1 Is the firm part of a concern/group?

Yes No If No, go to question 2

Is the firm headquarter for this concern/group?

Yes No

1a Is the firm parent company or subsidiary company?

Parent Company Go to question 2

Subsidiary Company

1b In what country has the parent company for the concern/group domicile?

Denmark

Other EU country

Country outside EU

Organization and management of the firm

2 Does the firm make use of some of the following ways of organizing the work?

(Please, check with an X how many employees are included)

No/ none	Less than 25%	25- 50%	Over 50%	Don't know
-------------	------------------	------------	-------------	---------------

Planned job rotation

Autonomous groups

Systems for collecting proposals from employees

Quality circles/groups (Formal delegation of quality control)

Delegation of responsibility

Interdisciplinary workgroups

Integration of functions (e.g. sales, production)

3 Has the firm carried through important organizational changes during the period 2003–2005?

Yes

No

Go to question 5

Don't
know

Go to question 5

4 Have the organizational changes primarily had as their objective to strengthen?

(Please, check each row)

High extent	Some extent	Small extent	Not at all	Don't know
----------------	----------------	-----------------	---------------	---------------

The effectiveness of the daily work

Co-operation and co-ordination across the organisation

The ability to adapt to more turbulent surroundings

The ability continuously to adapt new products/services

The ability continuously to strengthen and renew knowledge and know-how
 The ability to outsource activities
 Quality and customer service

5 Which mechanisms does the company use to improve the relationship between the daily operations and the company's development activities?

(Please check each row)	High extent	Some extent	Small extent	Not at all	Do not know
Coordination through use of rules and standards					
Cross-functional groups (combined across departments/areas)					
Systematic Continuous Improvement or similar program					
A link is ensured through daily management & leadership					
A link is ensured through recruitment, selection, training and development of personnel					
Work forms that support coordination between departments (for ex., projects, meetings, and information channels)					
Integrated product development					
Linking is ensured through strategy development and implementation of strategy					
By assigning people to coordinating roles					
Systems and processes that support knowledge sharing between development activities and daily operations tasks					
A culture that supports organizational learning and long term development					
A link is ensured by establishment of new business units with their own operating budgets and resources					

6 a How has the company prioritized the past year's improvement efforts?

(Please check each row)	Very high priority	High priority	Moderate priority	Low priority	Very low priority
Product/Service development					
Market development					
Technologic development (for example, product, service or process technology).					
Organizational development					
Business process development (for example, order process, supply chain process, or approach to product/service development)					

6b How has the company prioritized the coming year's improvement efforts?

(Please check each row)	Very high priority	High priority	Moderate priority	Low priority	Very low priority
Product/Service development					
Market development					
Technologic development (for example, product, service or process technology).					
Organizational development					
Business process development (for example, order process, supply chain process, or approach to product/service development)					

7 How would you rate the results of the company's improvement efforts in 2003-2005?

High results	Some results	Poor results	No results	Do not know
--------------	--------------	--------------	------------	-------------

Developments and outsourcing from the firm

8 Has the firm introduced new products/services during 2003–2005, when excluding minor improvements of existing products?

Yes, one	Yes, more than one	No	Go to question 10	Don't know	Go to question 10
----------	--------------------	----	-------------------	------------	-------------------

9 Are similar products/services found?

	(Please, check each row)		
	Yes	No	Don't know
On the Danish market			
On the world market			

10 Has the company outsourced during the course of 2003–2005?

Yes	Skip to question 12	No
-----	---------------------	----

11 Has the company considered outsourcing?

Yes	Skip to question 15	No	Skip to question 15
-----	---------------------	----	---------------------

12 To which countries has the company outsourced the following activities?

	Denmark	Other countries with high labour costs	Countries within Europe with low labour costs	Countries outside of Europe with low labour costs	No outsourcing of this type	This activity is not relevant for the company
Physical Production						
IT tasks						
Research & Development						
Other Activities						

13 To what degree have the following factors influenced the decision to outsource?

	To a high degree	To some degree	To a small degree	Not at all	Do not know	This activity is not relevant for the company
Cost reductions						
Strengthen and renewal of knowledge and know how						
Flexibility of production						
Quality and Customer Service						
Risk balancing						
Proximity to market						

14 Have one or more of the following conditions led to difficulties with outsourcing to countries with low labour costs?

	Yes	No	Do not know
Administrative barriers			

Lack of qualified employees in the partner company
 Lacking documentation and specifications of own products
 Uncertainty regarding quality and standards
 Cultural and language difficulties
 High establishment costs
 Difficulties with finding a suitable partner
 Attitudes/reactions in own company
 Delivery time
 Poor product development of the supplier

15 To which extent has the firm experienced competition from other firms during 2003–2005?

(Please, check each row) High extent Some extent Small extent Not at all Don't know/
 not relevant
 On the Danish market
 On the world market

16 To which extent has the firm co-operated with the following actors during 2003–2005?

(Please, check each row) High Some extent Small extent Not at all Don't know/
 extent not relevant
 Danish customers
 Foreign customers
 Danish subcontractors
 Foreign subcontractors
 Universities, institutions of higher education etc.
 Consultants

Employee use and participation in the firm

17 To what extent does the firm use the following possibilities to ensure that the personnel resources are in accordance with the needs of the firm?

(Please, check each row) High Some Small Not at all Don't
 extent extent extent all know
 By recruitment
 By dismissal
 By moving personnel between different job functions
 By regulation of working hours (overtime, flexitime)
 By changing the intensity of work
 By temporary employment
 By substitute
 By part time work

18 Are the employee representatives or the employees affected involved in decisions concerning:

Yes, employee Yes, employees No
 representative affected
 Recruitment
 Dismissals
 Internal organization developments
 Education and training
 Implementing new technology
 Developing new products/services

Qualification needs and competence development

19 Has the character of work changed during 2003–2005 along the lines of:

(Please, check each row)

	Employees with higher education	Employees with vocational training (Skilled)	Other employees
--	------------------------------------	--	--------------------

- Increased independence and responsibility
- Increased technical – professional demands
- Increased knowledge contents
- Increased interdisciplinary cooperation
- Demands to increase productivity
- Don't know
- Not relevant

20 How great importance do the following conditions have for management's efforts to ensure that the employees continuously develop their skills?

(Please, check each row)

	Great	Some	Small	None	Don't know
--	-------	------	-------	------	---------------

- Learning by doing
- Giving time for sparring with management/other employees
- Planned job rotation
- Organizing the work in teams
- Prompting co-operation and networking across divisions and groups
- Standard courses/educational schemes (e.g. vocational schools and AMU-centres)
- Educational activities tailored to the needs of the firm
- Long term educational planning

21 How many employees have participated in internal or external courses during 2003–2005?

	None	Less than 25%	25–50%	Over 50%
--	------	---------------	--------	----------

- Employees with higher education
- Employees with vocational training (Skilled)
- Other employees

22 How important is it for the competitiveness of the firm that the employees continuously develop their skills?

Great	Some	None	Don't know
-------	------	------	------------

Recruitment

23a Has the firm employees with advanced education (MA. and higher)?

Yes	Go to question 24	No
-----	-------------------	----

23b Does the firm plan to employ candidates with advanced education (MA. and higher)?

Yes	No
-----	----

24 Which expectations have the firm when it hires or consider hiring of a candidate with advanced education in preference to other types of employees?

	High extent	Some extent	Small extent	Not at all	Don't know
- We expect that such a candidate ...					
Has better methodological and analytical skills to problem solving etc.					
Are better to systematize existing processes					
Is better to deliver innovative schemes					
Are more able to learn new competences and adjust to changing working conditions					
Has better IT skills					
Has higher level of relevant professional knowledge					
Is better to absorb new knowledge from universities, institutions of higher education etc.					
Is better to establish relevant contacts of cooperation with Universities, Institutions of higher education etc.					
Is better to identify and conceptualize possibilities and problems					

25 Does the firm experience one or more of the following barriers for employing a candidate with advanced education?:

(Please, check each row)	Yes	No	Don't know
Lack of information on their competences			
Do not have sufficient tasks to this type of employees			
Lack of candidates with sufficient qualifications			
Salary level is too high for this group of employees			
Their approach to practical problems is too theoretical			
Internal resistance in the organization from other employees			

26a How do the company typically post vacancies?

(Please, check each row)	Always	Most of the time	Seldom	Never	Don't know
Through the public employment service	52	223	399	315	11
Advertisement in newspapers, professional journals etc.	118	541	282	53	6
Through job bases at the Internet	235	472	176	106	11
Through the 'mouth to mouth method' via the employed.	160	453	332	42	13
Through the 'mouth to mouth method' via connections in the sector.	86	310	452	132	19
Through direct contact to former employed	24	155	584	215	22
Through direct contact to applicants on waiting list / uninvited job application	41	284	533	124	18

26b How important are these methods of posting vacancies?

(Please, check each row)	Very important	Important	Not that important	Not at all important	Don't know
Through the public employment service					
Advertisement in newspapers, professional journals etc.					
Through job bases at the Internet					

Through the 'mouth to mouth method' via the employed.

Through the 'mouth to mouth method' via connections in the sector.

Through direct contact to former employed

Through direct contact to applicants on waiting list /uninvited job application

27 How significant is it for the job chances if an applicant has got a positive recommendation?

(Please, check each row)

	Decisive significance	Great significance	Some, significance	No, significance	Don't, know
From a former employer, written					
From a former employer, orally					
From others in the sector					
From own employed					
From person at the public employment service					
From educational institution					

28 There can be a number of reasons for using informal contacts ('mouth to mouth method') when hiring new staff. How significant are the factors mentioned below?

(Please, check each row)

	Decisive significance	Great significance	Some, significance	No, significance	Don't, know
It secures applicants that quickly can adapt socially at the company					
It secures qualified applicants					
It is a quick method					
It is a cheap method					

29 How many applicants do company typically for a free vacancy?

(Please, only one cross)

Typically only on hand-picked applicant

Below 5 applicants

5-9 applicants

10 - 49 applicants

Above 50

Don't know

30 Do your company typically have a formal job interview when hiring new employees.

Yes

No

31 Do your company demand documentation for a clean record of convictions when hiring new employees?

Yes, always

Yes, sometimes

No

32 How would you judge the job possibilities for these groups at your company?

(Please, check each row)

	Very good possibilities	Good possibilities	Limited possibilities	Very limited possibilities	Don't know
--	----------------------------	-----------------------	--------------------------	-------------------------------	------------

Unemployed with at long record of unemployment
Unemployed with ethnic (non-Danish) background
Unemployed above 50 years

33 Do your company have a formulated policy about giving good job possibilities to ...

(Please, check each row)

yes

No

Don't know

...applicants with ethnic (non-Danish) background
...older applicants

34 What are the most important risks with employing unemployed above 50 years?

(set more marks)

No risks
Lack of professional skills
Lack of work motivation
Lack of willingness to change
Lack of ability to change
Bad health
The chemistry in relation to other employees
The chemistry in relation to superiors
The chemistry in relation to the customers
Other things

35a What are the most important risks with employing unemployed with ethnic (not-Danish) background?

No risks
Lack of language skills
Lack of professional skills
Lack of work motivation
Lack of willingness to change
Lack of ability to change
The chemistry in relation to other employees
The chemistry in relation to superiors
The chemistry in relation to the customers
Other things

35b What are the most important risks with employing unemployed with a long record of unemployment?

No risks
Lack of professional skills
Lack of work motivation
Lack of willingness to change
Lack of ability to change
Bad health
The chemistry in relation to other employees

The chemistry in relation to Superiors
 The chemistry in relation to the customers
 Other things

Job training (Unemployed in public subsidised employment)

36 Job training i.e. employment of unemployed subsidised by the public, influence the job possibilities of the unemployed. How significant are the following factor?

(Please, check each row)	Decisive signifi- cance	Great signifi- cance	Some signifi- cance	Llimited or no signifi- cance	Don't know
Job training allow the unemployed to show their professional skills					
Job training allow the unemployed to show that the "chemistry" suits the rest of the company					
Job training gives the unemployed new qualifications that are necessary for employment					
Job training gives unemployed the job motivation back					

37 Has the company within the last year had unemployed in job training?

Yes	No	Go to question 39	Don't know	Go to question 39
------------	----	-------------------	-------------------	-------------------

38 Following your judgment. Had these vacancies been established on normal conditions if the arrangement with subsidised employment did not exist?

Yes, definitely	Yes, probably	No, probably not	No, definitely not
-----------------	---------------	------------------	--------------------

Appendix C: Names of the databases:

The DISKO4 IDA merge

Firm and Plants based	Person Based	Survey
DISKO4_IDA_FIRM02	DISKO4_IDA_PERS02	DISKO4 (1770 obs)
DISKO4_IDA_FIRM03	DISKO4_IDA_PERS03	
DISKO4_IDA_FIRM04	DISKO4_IDA_PERS04	

The Panel dataset

Firm and Plant based	Person based	Survey
PANELFIRM1997	PANELPERS1997	PANELDISKO2 (792 obs)
PANELFIRM1998	PANELPERS1998	PANELDISKO4 (792 obs)
PANELFIRM1999	PANELPERS1999	PANELDISKO2_4(792 obs)
PANELFIRM2000	PANELPERS2000	
PANELFIRM2001	PANELPERS2001	
PANELFIRM2002	PANELPERS2002	
PANELFIRM2003	PANELPERS2003	
PANELFIRM2004	PANELPERS2004	