Analyseinstitut for Forskning

Job Mobility – LFS versus Registers

Data Sources on Human Capital Mobility – the Danish case

Ebbe Krogh Graversen

Notat 2002/7 fra
Analyseinstitut for Forskning

The Danish Institute for Studies in
Research and Research Policy
Finlandsgade 4
DK-8200 Aarhus N
www.afsk.au.dk

Job Mobility – LFS versus Registers

Data Sources on Human Capital Mobility – the Danish case

By

Ebbe Krogh Graversen

The Danish Institute for Studies in Research and Research Policy

Paper prepared for the 1st ENMOB-meeting 25th-26th October 2001, Leuven, Belgium.
Updated January 2002. Final version October 2002.

Table of Contents

1. Danish Domestic Mobility

Everybody in Denmark has, in principle, a central personal register number, CPR. The number is uniquely connected to each person and identifies the person in all areas of the public register system, i.e. health, income and tax, subsidies and benefits, etc. local as well as national. The CPR system uses a ten-cipher code to represent each individual. The first six ciphers gives the date, month and year (without century) of birth, the next three ciphers are random while the last tenth cipher is calculated using a complex algorithm on the first nine ciphers. The last cipher also reveals the gender of the person, unequal for men, equal for women. Use of the complex algorithm backwards can also reveal whether a given CPR-number is a valid one.

Foreign immigrants are all given a permanent or temporary CPR-number when they approach the Danish public system for the first time. Emigrants keep their original number and get it back when they eventually return.

The existence of such a unique identification system open up for a wide spectre of matching and merging of data sources. Law prohibits this, unless the Register supervision authorities give a special permission. This is seldom done. Instead, the Statistics Denmark has a general permission to collect data, merge it and sell access to anonymous versions of the datasets. However, every time they create new dynamic databases, they also need permissions from the Register supervision authorities.

The following note concerns the part of collected data in Denmark, which can be used to analyse human work place mobility. Especially, the data sources, which may have similarities with data sources in other countries.

1.1 The Danish LFS

The Labour Force Survey, LFS, in Denmark is collected by Statistics Denmark and follows the definitions set up and used by Eurostat in their annual publication on Labour Force Survey results. Besides this, several extensions are attached or can be attached to the data using the unique personal register mentioned above. For example can the survey information be enriched by a merge with the register data collected from other sources, see Section 1.2 below. Similarly, a full time series history can be attached by merge of the LFS data from a given year and register data from previous (or future) years.

In Denmark, the LFS has been collected annually since 1984, since 1994 as a continuously survey. Before 1994, the sample unit was families, after 1993 it has been persons. The survey sample is collected and made representative by use of data from

registers on the CPR, population, unemployment, labour marked classifications, education classification and telephone modules.

The LFS contains information on the labour market attachment of the 15 to 66 year old Danish population. The population is split into people inside and outside the labour force. The people inside are further split into employed and unemployed persons. The attachment into a specific subgroup is determined by the status in the interview week. The LFS data is collected each quarterly. The sample is split into 13 equally sized groups, who are interviewed, one group each week during the quarter. Hence, information is collected for all weeks in a year. Aggregated quarterly or annual results are published by Statistics Denmark.

The aggregated results are corrected for sample errors that origins from the selection of and drop out in the survey sample. The results from the remaining sample are weighted by a factor, which in average is 333 (1st quarter 2000) over all the strata's used, i.e. a stratified weighting procedure, in order to make the results representative for the Danish population. In an attempt to increase the number of observations among unemployed person, one-third of the sample (ca. 5000) is sampled among persons who were unemployed in the previous quarter. The over sampling is corrected again in the aggregated measures. The average drop out percent is close to 30.

The Danish LFS cover a quarterly sample on 15600 persons and is collected primarily by telephone interviews. Persons not answering the phone get a questionnaire and are, if not responding on this, reminded once. Persons are asked about the previous weeks activities. The data collection is done from a rotating panel where one-third of the sample is replaced each quarter. Another one-third is interviewed again in the next quarter and again a year later. The retrospective information in the sample can be easily validated through the fact that persons are represented three times; both quarterly and annual human capital mobility can be analysed. Only the annual retrospective information is demanded by the regulation for the European LFS, see Eurostat.

Furthermore, the survey includes direct questions on previous work attachment. Hence, not only job-to-job mobility can be analysed, but also out of job as well as into job mobility as illustrated in Table 1. Here, the outflow mobility rates can be calculated in two different ways as illustrated in column three and four in Table 1:

- The 'out of job' mobility rate can only be calculated in one way as the share of employees in year t who are not employed in year t+1 divided by the total stock of employees in year t (column three in Table 1)
- The 'job-to-job' outflow mobility rate can be calculated in two ways as the share of employees employed in both years but at another employment place in year t+1 divide by

- the total stock of employees in year t (column three in Table 1)
- the total stock of employees employed in both years (column four in Table 1)

Similarly, the inflow mobility rate can be calculated in two ways:

- The 'into job' mobility rate can only be calculated in one way as the share of employees in year t+1 who are not employed in the previous year t divided by the total stock of employees in year t+1 (column seven in Table 1)
- The 'job-to-job' inflow mobility rate can be calculated in two ways as the share of employees employed in both years but at another employment place in year t divide by
  - the total stock of employees in year t+1 (column seven in Table 1)
  - the total stock of employees employed in both years (column six in Table 1)

Only the job-to-job mobility rates with employment stock of employees employed both years produce equal inflow and outflow mobility rates, see column four and six. Mobility rates calculated this way are significantly higher than the job-to-job mobility rates presented with the total employment stock used as divisor, see column four and seven.

Table 1: Inflow and outflow job mobility rates from Danish LFS data, example from 4th quarter 1998 to 4th quarter 1999. Weighted to population figures. Base 1999

State of employment	Persons (1999)	Outflow		Stock or flow type	Inflow	
	Millions	----- % -----			----- % -----	
Employed in 1998	2712	**100.0**		Stock in 1998		
Outgoing employees	289	10.7		'Out of job' mobility		
Employed in both years	2422		**100.0**	Common stock	**100.0**	
➢ Employed same place	2136	78.8	88.2	Stable workers	88.2	79.5
➢ Changed employment	286	10.5	11.8	'Job to job' mobility	11.8	10.6
Employed in 1999	2688			Stock in 1999		**100.0**
Employees coming from						
➢ Unemployment	43			'Into job' mobility		1.6
➢ Schooling	184			'Into job' mobility		6.8
➢ Other reasons for no job	38			'Into job' mobility		1.4

Source: Statistics Denmark: Statistiske Efterretninger. Arbejdsmarked. 2000:14.

The information in Table 1 is constructed from a sample of retrospective questions to respondents when they participate in the LFS. They are asked about their main activity a year before. This results in information on a collection of previous employment states

which is used to construct Table 1. Hence, the table is not constructed on basis of the information in the rotating sample procedure, i.e. what they are doing at the time of the interview in two consecutive years merged by an identification code. Apparently, this procedure where the respondents own memory decide whether they have had a job shift reduces the mobility numbers considerably. An exercise on shift in job place based on register data almost doubles the mobility rates; see Section 1.2 below. Hence, memory may not be the best information source on these matters. However, using the actual information in two consecutive years reduces the sample size so much that only macro numbers have any reliability.

1.1.1 The exact information on mobility issues in the Danish LFS

Although the questions concerning the actual job situation is very detailed, the retrospective questions on the previous job are relative vague in the Danish LFS. The retrospective part only concerns the main employment a year ago. The time span is not specified except from the phrase ‘a year ago’. If the respondent was employed at the interview time, he is asked whether it was the ‘same employment as a year ago’ or ‘another employment a year ago’ or if not employed a year ago, what else he was doing.

This means that the retrospective question leaves a lot of possibilities for misinterpretations. A new employment may happen inside the same establishment by promotions, i.e. an employee can shift function inside the organisation, or may not happen even though the respondent shifts job since the employment may be considered as the same, i.e. a brick layer do the same at two different employers. This means that the LFS based mobility rates may overestimate as well as underestimate the actual job mobility rates depending on which definition that is used.

1.2 Register Data in Denmark

Danish authorities collect data on many aspects of the economy and peoples life. The CPR number mentioned above uniquely identifies all registered persons. Similarly, a central business register, CVR, number identifies all firms and work places. All the registers are approved by the Register supervision authorities. If not, the data collection in the registers is illegal. Authorities having several registers may not merge these unless the Register supervision authorities approve it, which is very often difficult.

However, Statistics Denmark collects copies of all the different registers. Parallel with this, Statistics Denmark also collects their own data on a long list of subjects. Statistics Denmark has by law permission to collect these data and the respondents are obliged to answer.

All the databases at Statistics Denmark are in theory accessible through Statistics Denmark. This means that it is possible to buy results in tables, graphs etc. Similarly, is it possible to buy access to their researcher environment, where it is possible to work on the original data. The data is always anonymised and the access is under supervision of the research office at Statistics Denmark. Physically, the access has to happen at the research offices of Statistics Denmark. The access is expensive, slow and bureaucratic.

Statistics Denmark has for labour market researchers made the Integrated Database for Labour Market research, the IDA database. This database is longitudinal, covers the entire population for now two decades, 1980-1999. It has unique employer-employee links using the CPR and CVR numbers, so both persons and firms can be followed over time. IDA contains information on 5.3 million persons (the population) and approximately 230,000 firms (all firms with employees). Persons can be characterised by information on their work place or work places can be characterised by information on their employees. IDA contains more than 250 variables. Unfortunately, no economic information on the employers is included in IDA. Neither does IDA contain information on reasons for an employee-employer split, i.e. mobility. IDA contains factual information and no behavioural information.

Although IDA only contains annual data, it contains the most usually used information on the persons, their family, their jobs and income, and employer relations. It is possible to buy access to the database for special research projects. However, if more than the usual information is needed, for example migration information, this has to be merged on the database as a special service performed by Statistics Denmark. This is possible but expensive and time consuming.

From a human mobility point of view, the registers contain at least in theory information on all paid work, when it is done and where. In practice, the registers have problem separating the periods individuals work different places, i.e. whether it is simultaneous or sequential. In theory, all aspects of the employer-employee relations can be analysed. In practice, this is very complex to do so, and the data is not necessarily better than survey data. The major advantage of register data is that they cover the entire population.

In the mobility analysis, a person shift work place if the employer CVR number changes. However, this also happens if the firm shifts owner, branch or address. In IDA, a person is not changing work place, even though the CVR number changes, if the owner and branch is the same, if the owner and labour force is the same, or if the labour force and the address or branch is the same. This correction secures the correct measuring of the actual mobility rates in the database by exclusion of the mentioned pseudo changes.

A severe obstacle in the use of register based data for mobility studies are law enforced through the monopoly status regarding information gathering and merging of Statistics Denmark. Even though researchers in the recent years have got easier access to data, it is still bureaucratic and costly. There seems not to be political will to find finances that could make the access easier for the research environment even though this might be able to help the same politicians by policy relevant analysis' in the longer run.

2.1.1 Mobility information in the register data

Given access to register data Danish mobility rates from 1994-95 is given in Table 2. The table is set up like Table 1 for comparable reasons; see Table 1 for explanations on the calculations of mobility rates in each column.

Forward and backward looking mobility variables are constructed in the registers through a merge of individual and workplace information. If the workplace id changes between year t and t+1 the employee is employed in both periods and have a job change. If the workplace id is the same, the employee is stable, i.e. does not change job. If no work place id exists in one of the periods but not in the other, either an into-job or an out-of-job employment change happens. This information is then merged back on the individual specific data giving the information presented in Table 2. The rates are corrected for mergers, take-over, new owner, split of work place etc. following a row of corrective rules. For example, if 30 percent or more of the workers are found again with a new common work place id, it is not treated as workplace mobility, since it is more likely cased by one of the reasons mentioned above. Hence, a work place can theoretically be split into three parts without giving any workplace mobility in the registers used, i.e. the IDA database.

Since we only has access to a sample containing all highly educated and 1% of the remaining population, the common stock of employees vary between 1995 and 1996. However, the difference in the theoretically identical job-to-job mobility rate based on employees employed both years is small in magnitude.

Of more interest is that the stock of employees is 5-10 percent smaller in the registers compared to the LFS in Table 1. Similar differences are observed for the outgoing and less clear for the incoming flow of employees as well as the stock of stable workers. The opposite is the case concerning mobile workers employed both years. According to the registers this number is more than 20 (1996) or 30 (1995) percent higher than the LFS stock of job-to-job mobile workers. This seems to be the major reason for the 50 percent higher mobility rates found in register data compared to LFS data. Again the conclusion may be that there exists a lack of memory or understanding among the

respondents of the questions in the LFS questionnaire. At least, there exists a difference so clear that it influences any comparison significantly.

Table 2: Inflow and outflow job mobility rates from Danish IDA database, example from 1995 (outflow) and 1996 (inflow). Weighted population figures. Base 1995 and 1996 respectively

State of employment	Persons (1995)	Outflow		Stock or flow type	Inflow		Persons (1996)
	Millions	----- % -----			----- % -----		Millions
Employed in 1995	2472	**100.0**		Stock in 1995			
Outgoing employees	241	9.7		'Out of job' mobility			
Employed in both years	2231		**100.0**	Common stock	**100.0**		2269
➢ Employed same place	1851	74.9	83.0	Stable workers	84.6	76.0	1920
➢ Changed work place	380	15.4	17.0	'Job to job' mobility	15.4	13.8	349
Employed in 1996				Stock in 1996		**100.0**	2526
Employees coming from				'Into job' mobility		10.2	257

Source: Graversen 1999. Own calculations. Based on a sample of all highly educated and 1 percent of the remaining population aged 20-65. Hence, the population weighted common stock deviates between 1995 and 1996.

1.3 Electronic CVs and other new data sources

This is a new expanding field in Denmark. There seems not to be any standardization yet, but a kind of common sense seems to exist. Several work places hide their employee information, but a few in the private sector publish it on the Internet. In the public sector the picture is equal except in the research branch. Here it seems more usual to have CVs on the employees, especially the researchers. No common style seems to exist.

A quick search on the Danish part of the Internet reveals 28 Internet addresses with CVs. These CVs are usually from jobseekers using either union based Internet sites or private job matching firm Internet sites.

Regarding standardization work and structuring of data sources, a quick search on Human Resources Mark-up Language resulted in no hit while a similar search on XML

standards gave 29 hits. EDIFACT gave 7 hits. There seems to be a coordinated effort to implement standards in the field.

Also the publicly run national job centres in Denmark has a central Internet site with CVs and possibilities for job search etc. The site is relatively active compared with the Danish number of unemployment, 150.000 in the autumn 2001. The job centres site has 4157 jobs available, and 28500 active CVs. Yesterdays activity numbers were 146 new job opening, 74675 searches in the job bank, and 2185 searches in the CV bank.

2. International mobility

The lack of data in this field is a well-known problem. In Denmark, all movements of permanent character are in principle collected by Statistics Denmark. Through the CPR number this information can be merged with other relevant information, for example history of mobile persons. However, no information is collected on the temporary cross border mobile people.

Another problem is the lack of information on education, work experience, and other relevant aspects regarding the incoming people, the immigrants. What the abilities are among these people is not known in the registers or anywhere else at Statistics Denmark. There have been no censuses in Denmark since 1971. An attempt to correct the missing information among foreigners immigrating has been done in year 2000. Through a questionnaire, this snapshot tries to update information on these foreigners. Although it is planned to be performed on a regularly basis it is not decided yet. Similarly, information on Danish citizens updating their abilities abroad, are not updated in the Danish registers; neither in the period abroad nor when they returns.

Regarding researchers and other high-educated individuals, no centralised attempts are made to keep track on their cross border mobility. The Danish Research Academy has since the early 1990s kept track on the PhDs educated in the period. However, this is only a subsample of the relevant population for the studies.

Through special investigations of annual reports etc. from research environments it may be possible to dig out relevant cross border mobility information. We are not aware of any attempts to organise such an information gathering in practice.

The OECD reporting system on Migration, See OECD (2000, 2001) seems to be the best transnational information source at the moment although this information builds on national data sources as well.

References

OECD. 2001. *Innovative People. Mobility of skilled Personnel in National Innovation Systems.* OECD Proceedings.

OECD. 2001. *International Mobility of the Highly Skilled*. OECD Proceedings.

OECD. 2001. *Trends in International Migration. Continuous reporting system on Migration. Annual Report 2001.* SOPEMI, OECD.

OECD. 2000. *Trends in International Migration. Continuous reporting system on Migration. Annual Report 2000.* SOPEMI, OECD.

Graversen et al. 2002. *Knowledge transfer by labour mobility in the Nordic countries*. The Danish Institute for Studies in Research and Research Policy. Working Papers 2002/1. http://afsk.au.dk.

Graversen. 1999. *Formal Competencies in the Danish National Innovation System*. The Danish Institute for Studies in Research and Research Policy Report 1999/4. http://afsk.au.dk.

Statistics Denmark. 2001. http://www.dst.dk. In Danish.

Statistics Denmark. 2001. The LFS Questionnaire for telephone interviewing. In Danish.

Statistics Denmark. 2001. The LFS Questionnaire for paper surveying. In Danish.

Statistics Denmark. 2000. Statistiske Efterretninger. Arbejdsmarked. 2000:14. Danmarks Statistik. In Danish.

Various Danish Internet sites.