

Microdata for research and analysis

Siena Group on Social Statistics Helsinki 2005

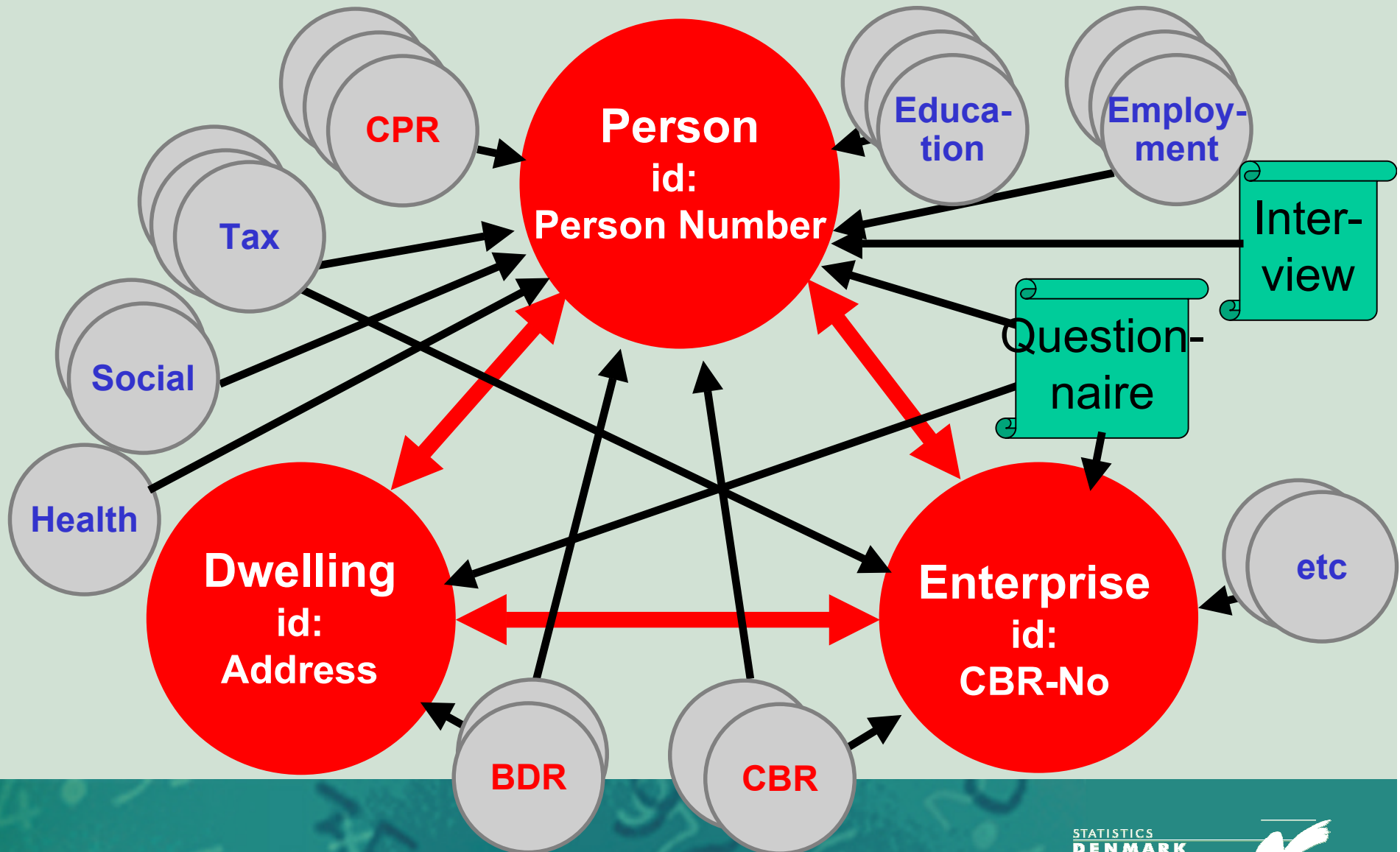
Presented by Flemming Petersson

What I will talk about

- A goldmine of data
- But there are problems ...
 - Social statistics / Person number
 - Workplace- and business statistics

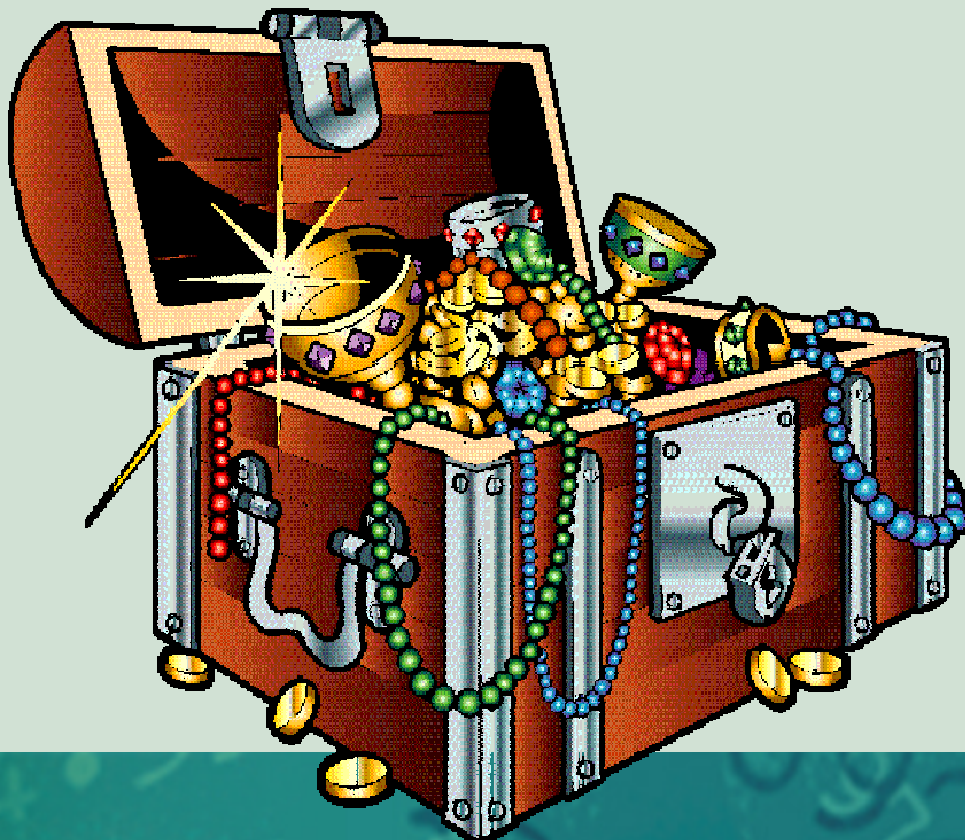


The Statistical Information System



A gold mine!

- High data quality
- Comprise entire population
- Longitudinal studies – cover several years
- Combining & linking through a set of keys



Linkage – process

- Registers linked by various keys:
 - Persons – Families - Dwellings
 - Workplaces
 - Enterprises
- Possibilities of combining data are infinite
- Illustrated by many research projects in:
 - economy, sociology, demography, epidemiology etc.
 - Access to Register of Medicinal Products

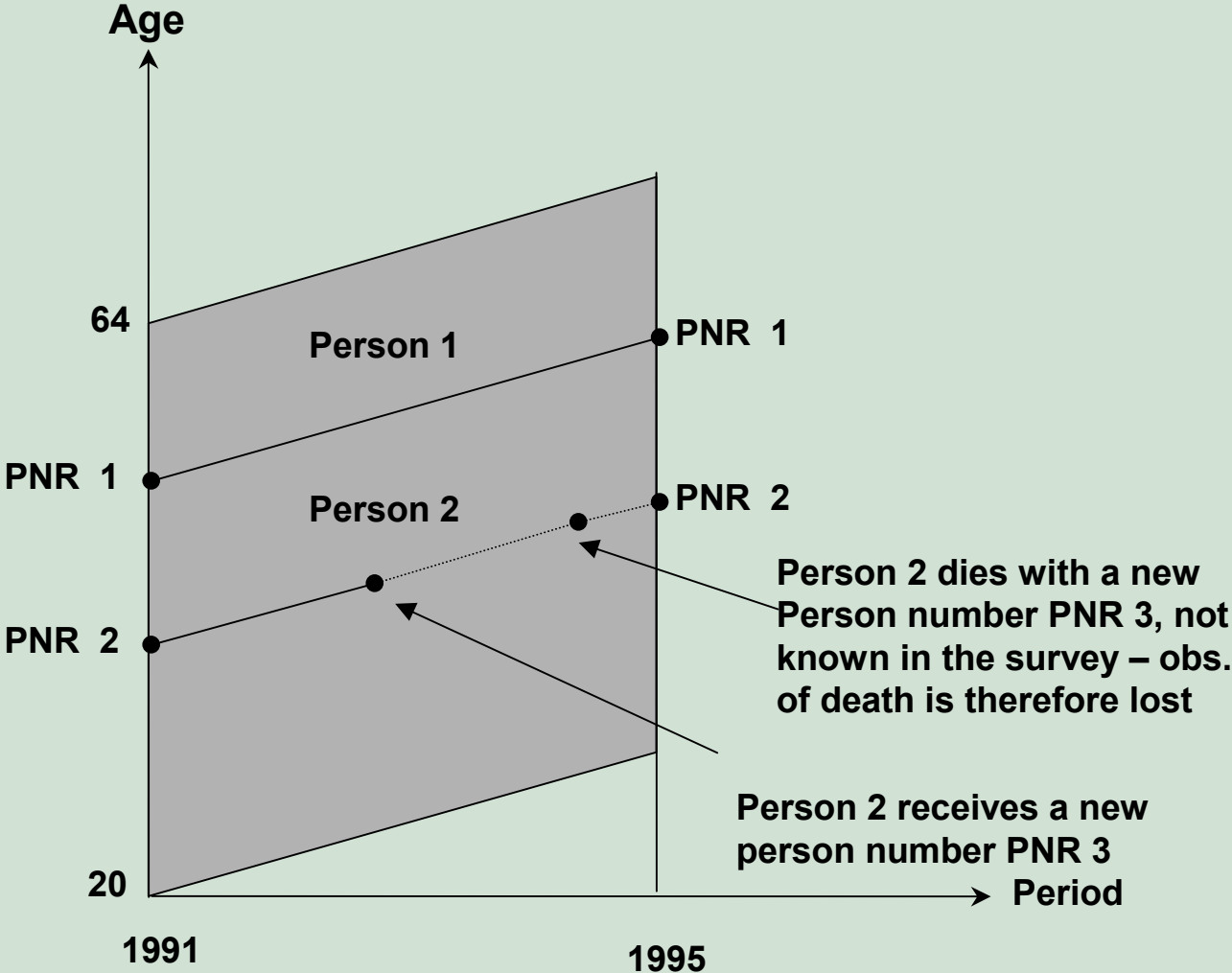
But there are problems

- Researchers prefer longitudinal data
- Unique identification is important
- We assume that person number is unambiguous and stable key, but:
- Person number is not a unique identifier
 - Some people change person number

Mortality and occupational survey

- A cohort of persons is observed for five years 1991–1995 (age 20-64)
- Measures differences in mortality
- Observed in terms of emigration and death
- Only person number assigned at the beginning of the period included
- If one change Person number and subsequently dies the event is recorded under the new Person number:

Change in person number - example



Change in person number since 1968

Number of changes in person number	Persons	Percent
0	7.187.615	98,65
1	92.876	1,27
2	4.875	0,07
3	400	0,01
4+	100	0,00



Danish / immigrants with address in DK

1.1. 2003 with change in person number

Danish/ immigrants	Persons with change in person number	Percent	Total population	Percent
Danish	46.586	82,1	4.952.818	92,0
Immigrants	10.152	17,9	430.689	8.0



Immigrants with change in person number

Country of origin	Immigrants with change in person number	Percent	All Immigrants	Percent
More developed country	2.816	27,7	182.001	42,3
Less developed country	7.223	71,2	246.412	57,2
Unknown	113	1,1	2.276	0,5

Why do people change person number

- Error in the registration process
 - Wrong date of birth and wrong sex
- Unknown age when entering the country
- Change of sex
- The person number will be changed
 - When the person make an application to the local community for some reason



What is the solution ?

Replace PNR with unique identifier

- We keep up with all changes in Person number
- Unique identification has been implemented in The Demographic database, that contains detailed information on the demographic events from 1980 and onwards:
 - Change of address
 - migration
 - marital status etc. and
 - change in person number



Employees and workplaces - a great potential for researchers

The Labour Market Research Data Base - IDA:

- Contains a large number of person, family and workplace variables
- Connects employees with workplaces fra 1980 - 2002
- Set of rules adopted for determining when a workplace change identity:
 - Ownership
 - Industry
 - Labour force
 - Address



Business data - a problem in research projects

How do we connect workplaces with enterprises?

- We use a key called FIDA
- FIDA covers only 1995-2001
- Before 1995 it is not possible to link workplaces with enterprises
 - Correct key require manual checks / data editing because of complex ownership structure fx:
 - Employment data in IDA may be recorded under one entity in the group, while:
 - Enterprise Statistics (turnover etc.) recorded under different entity of the same group, therefore:
- FIDA is not an absolutely correct key



Business data

- a problem in research projects

The enterprise statistics

- Deals with sale, export, industry, full time employed etc.
- Subject to many changes throughout the 1990s
 - Before 1999 several numbering system existed
 - New method in 1999 with the establishment of Central business Register: CBR number
- No business demography
 - No method to determine the identity of an enterprise over time
 - If an enterprise change ownership a new CBR number is created
 - The enterprise will appear as a discontinued entity and
 - A new entity is created
 - We are working on methods for business demography
 - Will create whole new options for researchers



Business data

- a problem in research projects

The account statistics

- Deals with detailed account items
- The quality of the statistic methods was improved substantially in 1995
 - Combine information from Tax authorities and questionnaires
- Not all sectors are covered 1995-1998
- From 1999 all sectors are covered except agriculture



Business data

- a problem in research projects

Conclusion:

- There exist more than one version of both accounting and enterprise statistics
- Limits the possibilities for consistent business statistics data over a longer period
- In research project it is often a problem that solid data only exist for short periods

The end

