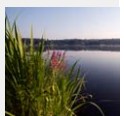


Issues and plans for the disclosure control of the Swedish Census 2011

Ingegerd Jansson, Statistics Sweden

ESSnet Workshop on Statistical Disclosure Control of
Census data

Luxemburg 19-20 April 2012





Outline

- A register-based census
- Important issues for disclosure control
- Current situation and feasible methods





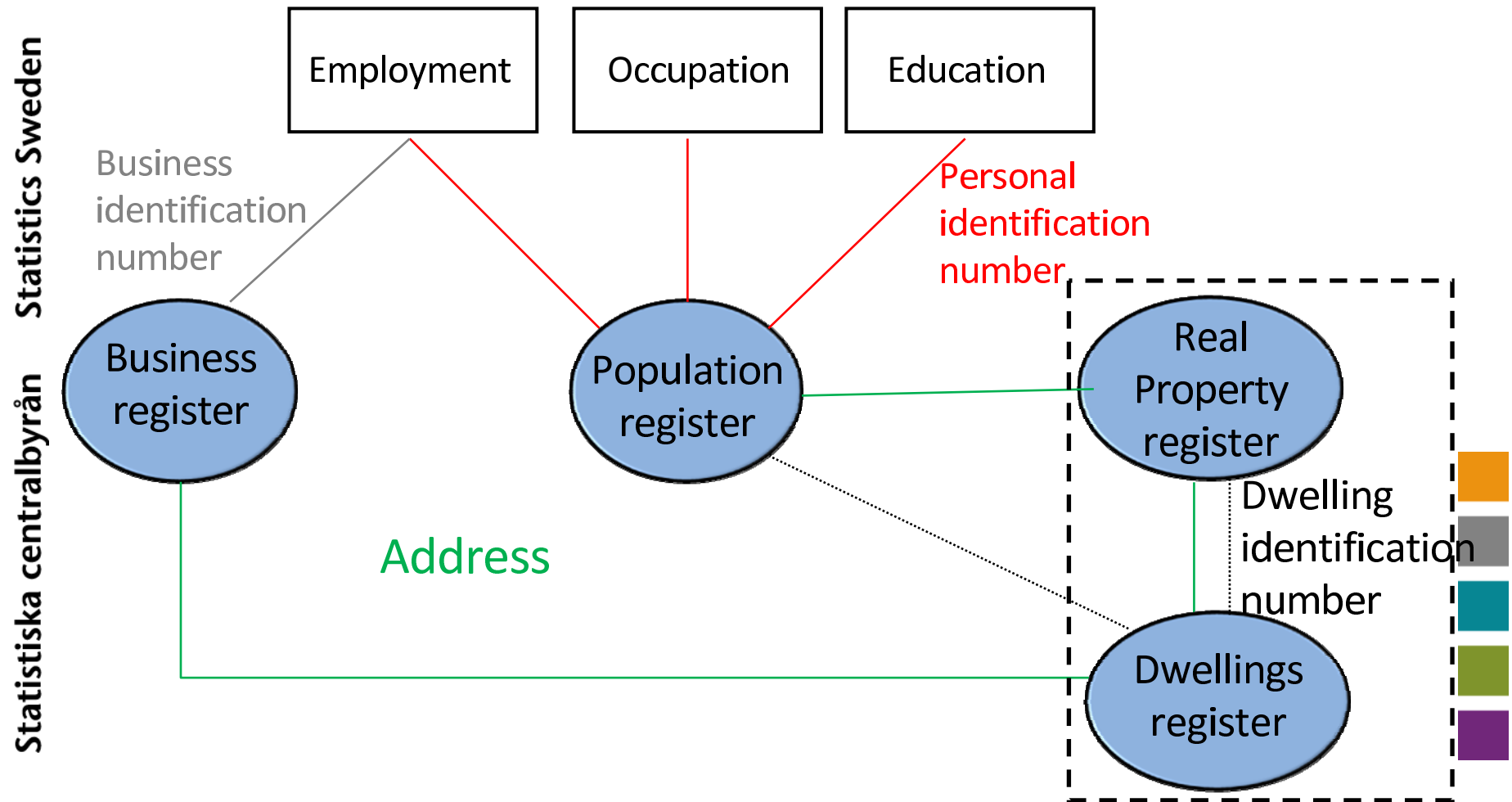
A brief history of census in Sweden

- 1960–1990 A traditional census every fifth year
The Swedish parliament decided that henceforth censuses will be carried out based solely on registers. Discussions on personal integrity delayed a law and regulation.
- 1995 The census was cancelled
- 2006 Law to form a register of dwellings passed
- 2007 Regulation regarding the register of dwellings adopted
- 2008 The collection of data for the register of dwellings started in January, finished in September 2010. Population registration on dwelling numbers 2010-2011
- 2011 December 31 is the reference date





A system of statistical registers





Legal requirements

- National legislation applies (EU regulation No 763/2008)
- All official statistics are confidential without exception and may only be published if it can be guaranteed that a disclosure will not cause harm or damage to an individual (Offentlighets- och sekretesslag 2009:400)
- Balance between the statistical notion of disclosure risk and the "soft" assessment of harm





Scope

- specific disclosure control solution for the Census 2011
- general solution that may apply to all types of official statistics on households and housing





Strategic issues

- Where are the greatest risks? small values, detailed geography, attribute or identity disclosure, group disclosure, risk of differentiation, or anything else?
 - involve management
 - consensus within the agency
 - systematic way of working to assess risk scenarios in statistical products, using Census 2011 as a case study





Strategic issues

- What is most important to keep in the published data?
 - reduce risk, keep utility
 - important users, what are their requirements





Publishing

- Publication in the Swedish Statistical Databases (SSD), aiming at 30 June 2013
- SSD does not currently have a function for disclosure control. Protect tables before uploading or as they are requested?
- National publication still being discussed





Current work

- Census SDC project, initial phase autumn 2011
 - literature review, important considerations
- Census SDC project phase two starting in May
 - strategy and criteria for choosing a method
 - tests on real data
 - make a choice of method(s)
 - outline for an IT-solution





Which method will we choose?

- Tables will be sparse
 - 290 municipalities, lowest level of geography
 - 9 482 855 persons registered in the TPR as of December 31, 2011
 - 4 508 373 dwellings in 2010 (apartments and one family houses)
- There will be missing data
 - 95,6 % coverage – 384 000 persons without dwelling number
 - In addition a presently unknown number of apartments without a dwelling number
 - Evaluation of households (running) and buildings (planned)





Which method will we choose?

- General solution
 - Pre-tabular method to be preferred?
 - Targeted record swapping, ABS cell perturbation method, other method?
- Specific census solution
 - Targeted record swapping one option, other method?
 - Post-tabular method has been discussed: random rounding of small values + additional rules (average cell size, threshold value, etc.), other method?

