



**Economic and Social
Council**

Distr.
GENERAL

ECE/CES/AC.6/2008/7
3 March 2008

Original: ENGLISH

ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Joint UNECE/Eurostat Meeting on Population and Housing Censuses

Eleventh Meeting
Geneva, 13-15 May 2008
Item 2 of the provisional agenda

CENSUS QUALITY ASSURANCE AND EVALUATION

Quality control in Finnish Censuses in 1970-2000

Note by Statistics Finland

Summary

This document presents briefly the history of the register-based population censuses conducted in Finland from 1970 to 2000. The document focuses on methods and measures taken to control the quality of census operations, and on the evaluation of the quality of census results.

I. PERIOD 1970-1980

1. In Finland the use of register-based data in censuses began already in 1970. At first the names of persons were pre-filled on questionnaires on the basis of place of residence. Personal ID numbers and addresses were also pre-filled. Demographic data from registers and income data from taxation were combined with the census data. Population was, however, counted on the basis of the questionnaire data. From the pre-filling Statistics Finland identified the major sources of errors in the central population register, namely over-coverage, errors in place of residence and errors in household-dwelling units.

2. The personal data collected in the census were used to establish a statistical register of persons at Statistics Finland, which was updated with the population change data arriving at the central population register. This population register was maintained from 1971 to 1975. In the context of the 1970 census Statistics Finland also established a statistical Register of Completed Education and Degrees, which has since then been the source of data on degrees and qualifications for all statistics and sample surveys.
3. In addition, map co-ordinates were defined for all buildings already in 1970. These co-ordinates may be regarded as the basis for geographic information.
4. In the 1975 census persons' names were pre-filled as lists by permanent place of residence. Personal ID numbers and addresses were pre-filled as well. The census data were complemented with the demographic data drawn from registers, the income data from taxation and the data on degrees and qualifications from the Register of Completed Education and Degrees. Population was counted on the basis of questionnaire data.
5. Data on a person's place of work were extracted from the Register of Enterprises and Establishments.
6. In connection with the 1975 census an extensive comparison was made between the register of persons and the central population register. In the period 1970-1975 some 30,000 population change notifications had been entered into the register of persons for which a counterpart person could not be identified (these persons had moved, women who had given birth, had died, or were otherwise not in the register). The comparison revealed at least a part of the under-coverage of the population census. On the basis of comparisons done in 1975 it was decided that the quality of the central population register was good enough for all demographic statistics, and since then the production of demographic statistics (and the population census) has been based on the central population register. After the 1975 census Statistics Finland discontinued its own register of persons.
7. The compilation of family statistics based on data from the central population register began in 1978.

II. PERIOD 1980-1985

8. In the 1980 population census the population, demographic data on the population as well as data on families and household-dwelling units were extracted from the central population register and only the data on the activities of the population and their places of work were collected with questionnaires. The physical data on buildings and dwellings were also collected with questionnaires. Data on dwellings and buildings made up the basic data of the Register of Buildings and Dwellings, which is why they were collected by virtue of a separate act and the owner of the questionnaire data was the Population Register Centre operating under the Ministry of the Interior. They were thus officially not census questionnaires. The Tax Administration's data on buildings and dwellings were also pre-filled on questionnaires, as were data on municipalities in the municipalities which had already established their own building registers.

9. The industry of a person's place of work and its location, etc., were extracted from the Register of Enterprises and Establishments.
10. Population was counted on the basis of the central population register.
11. The assessment of reliability started in 1980 when 20 different municipalities were identified from the census data and comparison data for the residents of these municipalities were produced from the registers. Data were compared as municipal statistics and by person. Statistics were then produced for the selected 20 municipalities on the basis of both questionnaires and registers. This comparison mostly concerned the definition of main activity, but even then it could be seen that surprisingly similar municipality statistics could be produced on the basis of registers as with censuses.
12. In the 1985 population and dwelling census direct data collection by questionnaires concerned only the definition of main activity, place of work and occupation. All demographic data, family data, data on education and income as well as all data on dwellings, housing and buildings were already at that time extracted from registers.
13. The 1985 census collection of data on place of work and occupation had one interesting special feature. Data on place of work and occupation collected in the preceding census in 1980 were pre-filled on personal questionnaires. If the place of work and/or occupation had not changed, the questionnaire forms did not need to be processed. For more than one-half of the employed persons the data had not changed, which meant that the processing costs fell significantly.
14. In addition, the 1985 census included fewer call-backs than usual. Instead of call-backs, register data were used for persons who did not respond or returned an incompletely filled-in questionnaire. For some 137,000 employees, or over 3 per cent of the population, the missing data could be taken from registers. This significantly improved the coverage of the census and the quality of data.
15. The questionnaire data of all the persons included in the 1985 census were compared with register data on them. The comparison concerned mainly the data on main activity. The data were compared at the level of the whole country, by region and municipality. The differences were surprisingly small in these comparisons and logical explanations were found for them.

Main type of activity according to the Census and Registers

	Men			Women		
	Registers	1985 Census	Difference %	Registers	1985 Census	Difference %
Labour force	1'272'532	1'262'641	0.8	1'179'719	1'153'350	2.3
Employed	1'198'532	1'186'833	1.0	1'113'524	1'089'988	2.1
Wage and salary earners	1'015'546	991'736	6.2	975'337	964'400	1.1
Own-account workers	182'986	195'097	-6.2	138'187	125'588	10
Unemployed	74'066	75'772	-2.3	66'195	63'332	4.5
Outside the labour force	618'867	628'824	-1.6	888'065	914'434	-2.9

16. The register situation has improved since then, and more comprehensive data on employment relationships are now available from various registers than in 1985.

17. Employment data are among the key census data and their reliability is especially important. Data on employment is primarily obtained from employment pension insurance records, as employment pension insurance is obligatory for all employees (wage and salary earners and own-account workers). If an employer neglects to pay the pension insurance, data describing the existence of an employment relationship may be obtained from taxation records when the employer reports the withheld advance tax. The third route for obtaining data on employment relationships is the own report of the person liable to pay taxes. Employed persons report annually the income they have received from their employers.

18. Some systematic differences can be seen in register and questionnaire data. The number of the employed is roughly the same, but according to register data, some male own-account workers are classified as wage and salary earners and some women engaged in domestic work are classified as own-account workers. Many own-account workers who are engaged in paid work in addition to entrepreneurship consider themselves as own-account workers, even if the income from paid work makes up the major part of their livelihood. Similarly, some of the women engaged in farm work do not report it, in which case they are classified as housewives in the census.

19. Some of the students who also work forget to report their limited employment on the questionnaire, but on the basis of register data they will be classified as employees. In such situations the questionnaire reduces the number of employees and increases the number of students. In all these cases the registers produce a more correct result than questionnaire data collections, as registers do not overlook limited employment and reveal the financially most significant employment relationship. The biggest advantage of register data is indeed that at least in Finland they cover nearly all gainful work.

20. The 1985 census data were also compared with data from the Labour Force Survey for the same period. The results are very similar to those of the comparison with registers. The Labour Force Survey produces slightly more employees, as it covers some work which is not insured or reported in taxation. The very low lower limit of work, i.e. one hour per week is enough to fulfil the criteria of employment, also contributes to this. Students end up as employees in the Labour Force Survey interviews more easily than in questionnaire censuses, but some own-account workers remain as own-account workers even though their income from paid work is higher than their entrepreneurial income.

21. In the Labour Force Survey industry was defined according to the respondent's report and in the questionnaire census according to the Register of Enterprises and Establishments. This causes systematic errors in the Labour Force Survey, e.g. the builders of a factory producing houses get construction as their industry, whereas the correct industry would be manufacturing. Similarly, e.g. the mechanical engineering personnel at the railways may be classified under transport, whereas the correct industry would be metal industry. The same applies to vehicle drivers in trade and manufacturing; in interviews they often end up as transport personnel.

III. THE USE OF THE REGISTER OF ENTERPRISES AND ESTABLISHMENTS IN CENSUSES

22. When the employed are linked to the Register of Enterprises and Establishments, the people working at an establishment will automatically be assigned with the real industry of that establishment. Defining industry is difficult in a questionnaire data collection and the risk of errors is high. The defining of data on place of work with the help of the Register of Enterprises and Establishments began already in the 1975 census and the same procedure was applied in the 1980 and 1985 censuses. This procedure has been very helpful, when industrial data were harmonised between censuses and economic statistics.

IV. QUALITY CONTROL IN THE FIRST REGISTER-BASED CENSUS IN 1990

23. An extensive quality survey was conducted in connection with the 1990 register-based census.

- (a) At first two per cent of the real estate in the country were selected.
- (b) Then building data were collected from the owners of these real estate (23,000 buildings) and all the dwellings in these buildings were counted (45,000 dwellings).
- (c) Thirdly, all the persons living in these dwellings were interviewed (96,000 persons).

24. In many cases a register-based system produces more correct data than surveys based on questionnaires or interviews. The register-based system produces a slightly higher number of employees, because some work performed while studying, for example, may be forgotten from questionnaires. Some people who are not working and are not registered as unemployed do, however, regard themselves as unemployed. Some students do not report their employment on the questionnaires and thus get counted as students.

25. Naturally the register of the unemployed also includes persons who are no longer factually seeking for work but only seek to receive unemployment benefits. On the other hand, young job seekers may become unemployed on the questionnaire, even if they are still continuing their studies. Though the term is not yet over, they may already regard themselves as unemployed.

26. The number of pensioners is roughly the same; only a small group of pensioners are still in employment relationships, in which case they end up as employees on the basis of register data. The mothers of own-account worker families, who have taken out an entrepreneur insurance, also become employees.

27. The below table illustrates the differences between the two surveys.

Main type of activity, 15-74 years old			
	Census 1990	Sample survey	
Total	3'751'200	3'751'800	-600
Employed	2'332'200	2'305'700	26'500
Unemployed	141'000	148'800	-7'800
Student	329'000	362'600	-33'600
Pensioner	757'000	766'200	-9'200
Conscript	28'000	24'000	4'000
Other	164'000	144'500	19'500

28. Some of the unemployed on the basis of questionnaires may be in a transition phase; they may be on leave from their last employment relationship but already report themselves as unemployed. The same applies for persons going into retirement. If the employment relationship continues as leave, the person concerned may already regard him/herself as a pensioner. Students may intentionally or unintentionally leave small-scale paid work unreported in the interview, but according to employment relationship data they are classified as employed (which is, indeed, correct).

29. Persons who have more than one job at the same time have trouble defining the most important employment relationship, e.g. a farmer who is in temporary paid work may regard him/herself as a farmer, even if his/her income from paid work was higher. Errors like this do not affect the number of the employed, but they do affect the data on industries.

Main type of activity, aged 15 to 74							
	CENSUS 1990						
Sample Survey	Total	Employed	Unemployed	Student	Pensioner	Conscript	Other
Total	3'751'200	2'332'200	141'000	329'000	757'000	28'000	164'000
Employed	2'305'300	2'172'600	34'500	34'000	24'600	4'000	35'600
Unemployed	148'800	23'900	97'100	5'400	2'300	1'200	18'900
Student	362'600	57'000	3'100	282'400	2'000	1'400	16'700
Pensioner	766'200	35'400	1'800	2'000	717'800	100	9'100
Conscript	24'000	1'100	600	200	100	21'200	800
Other	144'300	42'200	3'900	5'000	10'200	100	82'900

30. Population groups in transition phases, e.g. those transitioning from work into retirement, into unemployment, domestic work or studying, are always difficult. A leave period, for example, can confuse the definition of the transition date.

31. Similarly, transitions from studying into unemployment or other activities are difficult to define, because the ending of studies is not always clear cut. Self-directed studying is also difficult to define; some of these students are unemployed persons applying for entry into different educational institutions.

32. Since 1987 Statistics Finland have compiled the Register based Census annually. Every year census data have been combined with the data of Labour Force Survey at individual level. These combinations confirm that the quality remains at high level in Finnish register based statistics.

V. CHECKING FOR CHANGES

33. In the register-based method all data are produced annually and published at all area levels, small areas included. Statistics Finland has constructed a checking system which compares new data with the previous year's data by municipality. The system gives a notification when e.g. the number of jobs in a certain industry in a certain municipality changes illogically (by more than 10 per cent). In such cases an examination is launched into whether a new establishment has been opened or an old one closed down, or if there is an error in the location of the establishment. In most cases the error stems from the link between employees and establishments, which makes this checking necessary.

34. Sometimes missing data on persons can be completed with data from the year before. If data on occupation or establishment are missing, data from the previous year can often be used. The publication of data by municipal sub-areas and their assignment into municipalities often generates feedback, if users find the data illogical. It is very seldom, however, that errors turn up at the publication stage.

35. Longitudinal data files. Statistics Finland has compiled all register-based annual data into one longitudinal data file. With the help of this file the changes occurring to persons can be followed-up and it can be checked that the number of retiring persons matches that of beginning pensions, the number of newly unemployed matches that of new periods of unemployment, and the number of women transitioning to domestic work matches that of new maternity leave periods. The longitudinal data file also shows immigrants as new cases and emigrants as missing cases. The longitudinal file is, therefore, a good tool for examining e.g. the integration of immigrants into the country's labour market.
