

Working Paper No. 2
ENGLISH ONLY

**STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE**

**STATISTICAL OFFICE OF THE
EUROPEAN COMMUNITIES (EUROSTAT)**

CONFERENCE OF EUROPEAN STATISTICIANS

Joint UNECE-EUROSTAT Work Session on Registers and
Administrative Records for Social and Demographic Statistics
(Geneva, 9-11 December 2002)

Supporting paper

USE OF BUSINESS REGISTERS IN SOCIAL STATISTICS IN FINLAND

Submitted by Kaija Ruotsalainen, Statistics Finland¹

¹ The views expressed in this paper are those of the author and do not necessarily reflect the views of Statistics Finland.

I. INTRODUCTION

1. The history of producing and maintaining a business register goes back half a century in Finland. The first enterprise census was drawn in 1952, the next one followed ten years later, and its regular maintenance was started at the beginning of the 1970s. In its present form the business register was set up at the beginning of the 1980s and it was also exploited in a population census in the same decade. When the register-based census system was being created, the business register became one of the most important basic registers for it. The production of annual employment statistics became possible because annual updating was also introduced in the business register towards the end of the 1980s. Today, the Business Register and employment statistics work closely together to the benefit of both sets of statistics.

II. THE BUSINESS REGISTER

2. The business register is a basic statistical register. Its purpose is the maintenance of exhaustive data on enterprises, and on their establishments and activity. Since 1999 the scope of the register has also extended to the public sector, i.e. central and local government units.

3. The basic units in the business register are the legal unit, its local units and establishments. Each unit is given a unique identifier. The identifier of a legal unit is its Business Identity Code (Business ID). In Finland, all enterprises and corporations are registered in a common enterprise information system, the Business Information System (BIS). The BIS is shared by the tax authorities and the National Board of Patents and Registration, where they are assigned their Business ID codes. Identifiers for the other basic units are given at Statistics Finland.

4. To a large extent the business register is maintained and updated today by utilising data from diverse administrative registers. A prerequisite for being able to do this is that the different registers use the same, unique identifiers. In Finland, unique identifiers are used almost without exception in administrative registers, so the preconditions for their use as statistical sources are well met.

5. The most important administrative data sources are:

- Diverse data registers of the Tax Administration, e.g. customer database, business taxation data files, VAT and PAYE data (monthly), data on annual wages (annual PAYE register);
- Data of the Bank of Finland;
- Data of the National Board of Customs;
- Data of the National Board of Patents and Registration;
- Register of Buildings and Dwellings;
- Data of the State Treasury;
- Data of the Local Government Pension Institution;
- Data of Finland Post Ltd.

6. The data obtained from administrative registers are utilised in two ways as source data for the business register. The data may be drawn as such from the source registers or they can be used to estimate data that are not available directly from the registers. Examples of direct exploitation would be information about the establishment of new enterprises or about changes in information concerning existing enterprises, which are received quarterly from the customer database of the Tax Administration. The Tax Administration's data on annual wages is used to estimate the number of employees in enterprises, while the Central Pension Security Institute's data on persons contributing to self-employed persons' insurance schemes are used for estimating the number of self-employed persons.

7. Besides administrative data, data obtained from diverse inquiries are also utilised in the updating of the business register. Inquiries are sent to all multiple establishment enterprises, a sample of single establishment enterprises, new enterprises having started operations, and to central government agencies. The so-called quality inquiry additionally yields information on approximately 5,000–10,000 enterprises. The main purpose for the business register's own inquiries is to maintain establishment level data on business structures and changes, because these are not available from any administrative register.

III. USE OF THE BUSINESS REGISTER IN POPULATION CENSUSES AND IN EMPLOYMENT STATISTICS

8. Information on the industry of employed persons is an essential population census data item. There are two commonly used ways to collect this information. The first one is to establish the industry of an employed person's workplace by asking about it and then code the industry according to the answer. Another way is establish the name and address of a person's workplace, whereafter it can be identified from e.g. the business register and an employed person can then be assigned the industry of his or her workplace. In population censuses, as well as in other inquiry-based surveys or interviews, the industry is usually defined for a person on the basis of his or her own description and statement. Exploitation of the business register in population censuses was started in Finland in 1975 and 1980, in other words 10 to 15 years before the register-based population census system was established.

III.1 Use Of The Business Register In Questionnaire-Based Censuses

9. Identification of workplaces was started in Finnish population censuses of 1975. The census questionnaire included a question about the name and address, as well as industry, of a person's workplace, and the answers were used to assign to employed persons an establishment code from a list that had been compiled for the census. The lists had been drawn up from the data in the business register at that time. The register then only covered employer enterprises in the private sector, and not even all of their establishments exhaustively. Not all public sector (central and local government) establishments were included in the register, either. The list of establishments was continuously supplemented during the enumeration work so that whenever a new establishment was found in the questionnaire data it was immediately added to the list and the enumerators subsequently used the new establishment code given to it. Industries, exact locations and legal forms were later defined for these new establishments, and these were then assigned to those working in them with the help of the establishment code. This means that as far back as 1975, all persons working at a given establishment received the same industry.

10. Only if a person's workplace was at home, mobile or outdoors, e.g. agricultural workers, construction site workers, professional drivers, market sellers, family day-carers, etc., was his or her industry defined directly from the questionnaire data.

11. After the census the list of establishments was, therefore, quite exhaustive, because the establishment of every person who worked at a fixed establishment had been added to it. The list in fact became over-exhaustive, because some establishments were entered into it twice or, in a few cases, more than twice. This is because employees could refer to the same employer by names that differed from each other so much that it was difficult to recognise that they meant the same establishment and consequently the establishment was added to the list once more.

12. The same method was used in the 1980 and 1985 population censuses. The population of enterprises and establishments was again drawn from the business register, but it was supplemented with establishment data that were missing from the business register but that had been obtained with the previous census, such as data on public sector establishments. Local governments checked the correctness of the list and added to it the missing establishments, removed establishments that had closed down and corrected addresses on the list. The extent of these corrections to the list varied between 20 and 40 per cent by municipality.

III.2 The Business Register In The Register-Based Population Census System

13. The use of administrative data in population censuses was gradually introduced in Finland starting with the population census of 1970. In 1987, data concerning the workplace of a person were produced from registers for the first time and the 1990 population census was totally register-based. At the same time, annual production of almost all population census data items was started.

14. The population for employment statistics is the entire population of the country, and the scope of the statistics extends to data on the population's demographic characteristics, education, main activity, employment, unemployment, industry, employer sector, workplace location, income, family relationships, etc. The data sources for employment statistics comprise over 30 administrative registers or datafiles. In addition, an annual inquiry asking for information about employees' establishments is conducted among multiple establishment enterprises and local government's operating units.

15. The business register is exploited in the production process of employment statistics at two different stages: in the collection of data, and as a data source in the actual production of the statistics. Information about a person's employment relationship and employer is obtained from employment pension insurance datafiles. These registers contain enterprise level data concerning the private sector, meaning that their data reveal only a person's employer, and not the establishment at which he or she works. Where an enterprise with only one establishment is concerned, there is no problem about a person's establishment as the only establishment of the enterprise is assigned as his or her workplace. The situation becomes more complicated concerning enterprises with multiple establishments.

16. An annual enterprise inquiry is conducted in order to ascertain establishments for persons who work for multiple establishment enterprises. The inquiry is conducted in collaboration with the business register. The inquiry sample is drawn from the business register and it covers all enterprises that employ more than 10 persons and have two or more establishments. The size of the annual sample is approximately 4,500 enterprises.

17. The inquiry is comprised of two parts: the inquiry of the business register and the inquiry of employment statistics concerning employees. The business register inquiry asks enterprises for information about the sizes, locations and industries of their establishments for the purpose of updating the data in the register. Employment statistics ask for information about the persons who work in each establishment. Enterprises are sent lists of employees that have been prefilled with data obtained with the previous year's inquiry and are asked to update these. Enterprises are requested to check the information and confirm that the division into establishments agrees with the current division they have reported to the business register.

18. The inquiry concerning business registers is returned to the business register and the establishment-specific employee forms are returned to employment statistics. The business register and employment statistics work in constant cooperation at the data verification stage. Verifications are made to ensure that establishments correspond with each other, numbers of employees are comparable with each other, etc. Enterprises sometimes only report information about their changed location or new establishment to employment statistics, and from here the information is then forwarded to the business register. Thus the cooperation benefits both parties and harmonisation is achieved between two large statistical systems.

19. As already mentioned, data concerning a person's employment relationship are obtained from employment pension insurance registers. In the public sector these registers also contain information about a person's establishment, not just on his or her employer. Combining data on employment relationships in the private sector with the establishment-specific data on employees collected with the inquiry also yields establishment data for persons working for private sector enterprises. Persons working for single establishment enterprises in the private sector are assigned the only establishment of the enterprise as their establishment.

20. Once a person defined by employment statistics as employed has been assigned the Business ID and establishment codes of his or her workplace, the codes are used to draw data concerning that person's workplace, such as industry, employer sector, enterprise size and location (municipality, address, coordinates), from the business register.

21. From the point of view of production of regional statistical data, it is very important that this link at the establishment level exists between a person and his or her workplace. Employment statistics is the only system that is capable of producing annual data on the whole population, employed persons and workplaces at the small area level. Therefore, it is of utmost importance that the quality of this link from an employed person to an establishment be maintained to as high a degree as possible. Additionally, the link also facilitates the production of structural enterprise statistics that incorporate data on the demographic characteristics, education, turnover, etc., of the employees of enterprises.

III.3 Advantages And Problems In The Use Of The Business Register In The Current Population Census System

22. In register-based statistics the data on, for example, the industry and workplace location of a person are obtained from the business register. This means that all persons working at the same establishment are assigned exactly the same industry and workplace location. In questionnaire-based censuses, the data concerning the industry and location of a person's workplace were determined by the person's own reporting. This meant that there could be discrepancies in these data even among persons working at the same establishment, for the reporting of industry can be difficult in the case of a multiple industry enterprise, for example. Moreover, when the business register is used the data on industries and locations are verified and updated regularly, meaning that the data obtained from it are more reliable and uniform than data collected with questionnaires.

23. From the point of view of regional statistics, register-based statistics production does have its problems. The following are examples of problem situations that may arise in the production of data by area, and especially those on workplaces. As already pointed out, the most vital prerequisites to the production of register-based statistics are identifying codes. If they do not exist, names can in certain cases be used to identify a unit and combine it with another register. Because of this, an error in a specific data item generally affects a larger number of persons in register-based statistics than in questionnaire-based statistics. For example, if the information about the location of an establishment is wrong in the business register, all those working at that establishment are assigned erroneous location data. If the same information is asked with a questionnaire, perhaps only a small number of respondents give erroneous data about the location of their workplace and the error would be a random one. Thus, the aspect that on the one hand is an advantage with a view to the uniformity of data may, in certain cases, also become a disadvantage.

24. Large national enterprises with regional organisations create problems because they may have just a few establishments in the largest towns in different parts of the country, but their employees may work in different parts of the area. Such cases are often connected with occupations in which persons do not necessarily work at the same location all the time but their work may be of a mobile nature, such as e.g. cleaning. Another example of this type of problem would be enterprises practising labour rental activities, in which case the enterprise operates in one location but the persons it hires out to other enterprises may work in different locations within the radius of several hundreds of kilometres. Yet, these persons have employment relationships with the labour rental enterprise.

25. Register-based statistics cannot be produced on persons performing mobile work in the way they can be produced if data are collected with questionnaires. Persons doing so-called mobile work, e.g. transport or cleaning work, are assigned as their workplace locations the establishments from which their work is supervised. In some cases this may result in concentrations of industries in certain localities. An example in Finland would be the attachment of the personnel of a certain large shipping company to its head office, which is located on the Åland Islands.

26. Small and casual employers that are missing from the business register also constitute a problem area. Examples of this would be builders and decorators hired by private households, the data on whose industry, etc., will remain deficient.

IV. EXPLOITATION OF THE BUSINESS REGISTER IN THE LABOUR FORCE SURVEY

27. The labour force survey is a panel survey in which the monthly sample is divided into five rotating groups. Each group comprises approximately 2,400 persons, so the total population consists of 12,000 persons. Continuous survey week was introduced in the Finnish labour force survey at the beginning of the year 2000.

28. The labour force survey is implemented as computer-assisted telephone interviews. The interviewers record the obtained information as such and only code the information concerning occupation during the interview. Besides the information on occupation, the data that need coding are those on the industry and employer sector of a person. The panel design is exploited in the collection and processing of the data. During rounds 2 to 5 the interviewer checks whether the data on workplace and occupation have remained unchanged from the previous interview. The data from the previous interview have been prefilled on the respondent's questionnaire (screen display). If the occupation and workplace are unchanged the information need not be saved and recoded but is copied from the previous interview round. Only completely new or changed data are coded.

29. The business register inquiry base, as well as the classification database, are utilised in the coding. An establishment corresponding with a person's workplace is sought from the business register against the workplace name and address given by the person in plain language. If a matching establishment is found in the business register, the person is assigned an industry in accordance with the business register's data on establishments. If the workplace given by the person cannot be identified in the business register, industry is coded with the help of the classification database from where the industry code is obtained against a description of the industry in plain language.

30. Data on the employer sector are obtained from the business register. However, these data are always checked because the classification used in the labour force survey differs from that used in the business register.

31. The new system of coding will be in use as from the beginning of 2003. Then, if an enterprise and an establishment corresponding to the information given by the interviewee are found in the business register, the enterprise and establishment codes are also saved in the labour force survey database.

V. CONCLUSION

32. The business register has been exploited in population censuses for a couple of decades now. A register-based system for producing employment statistics was built towards the end of the 1980s and this also facilitated the introduction of annual data updating in the business register. Thus, it also became possible to start producing the conventional population census data annually. Utilisation of the business register as a source of data on the industry, workplace location, etc., of employed persons has reduced data collection significantly in comparison to questionnaire-based censuses. Moreover, the data obtained from the business register are more unified than those obtained with questionnaires. In questionnaire censuses, even persons working at the same establishment could give the information concerning their workplace in very different ways. The labour force survey has also started to exploit the data in the business register to determine the industry and employer sector of a person. Thus, use of the business register in different statistical systems increases coherence and comparability between them.

33. Although exploitation of the business register poses certain problems from the perspective of regional statistics, for example, the benefit gained from its utilisation far outweighs these problems. The production of annual regional employment statistics would be impossible without a well functioning and continuously updated business register.

Sources

Kiiski, Salme (2002). Labour Force Survey, monthly interviews (memory). Only in Finnish.

Leinonen, Tuija (2000). Administrative and Statistical Units in the Finnish Business Register. A paper prepared for the 14th International Roundtable on Business Survey Frames, Auckland, 29 October – 3 November.

Myrskylä, Pekka and Kaija Ruotsalainen (1997). Annual System of Small Area Statistics Based on Administrative Records and Registers. Contributed Papers for the 51st Session of the ISI, Istanbul. Book 1, 511-512.

Ruotsalainen, Kaija (2001). Annual System of Small Area Statistics Based on Administrative Records and Registers - the Possibilities and the Problems. A paper prepared for the forty-ninth plenary session of the Conference of European Statisticians, Geneva, 11-13 June 2001.

Viitaharju, Tuula (2002). Managing Administrative Sources, the Business Register as a Link between Economic and Social Statistics. A paper for the 16th International Roundtable on Business Survey Frames, Lisbon, 21-25 October 2002.