

Working Paper No. 6
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)

Invited paper

**THE ROLE OF BUSINESS REGISTER (BR) IN SOCIAL STATISTICS - PRINCIPLES,
PRACTICES AND FUTURE DEVELOPMENTS**

Submitted by Magne Bråthen, Svein Gåsemeyr, Kjell Lorentzen, Helge Næsheim, and
Harald Utne, Statistics Norway

I. INTRODUCTION

1. The paper aims to discuss some principles of the role and functioning of the BR in a comprehensive and integrated statistical system. Practical aspects of the use of BR in social statistics are presented using Norwegian examples.
2. The most important units of administrative procedures and official statistics are natural person, establishment, and dwelling (the dwelling is related to property/building/address). The infrastructure of basic registers of these three units is developed in most European countries and represents an important tool in developing efficient administrative procedures.
3. In an increasing number of countries, the basic administrative registers assign official and unique ID numbers for person (PIN), business unit (BIN), and dwelling/address (DIN) to be used for administrative and statistical purposes. Interesting opportunities emerge when basic registers can be linked at unit level. Examples of this could be the use of the official and unique address (DIN) in the register of natural persons and the business register. In Norway there exists a rather rare link between natural persons and business units through the Social Security Authorities' Employer - Employee Register. Through this register, the unit of job is identified by PIN and BIN represented by the link between the registers of persons and businesses.
4. The use of BR in register-based social statistics is related to the unit of job. In a Swedish paper presented at the third Work Session on Administrative Records, Geneva 1999 [1], the register-based job file was described as a basic register of the Swedish statistical system.

II. THE ADMINISTRATIVE AND STATISTICAL BUSINESS REGISTER - THE CASE OF NORWAY

5. The maintenance of a statistical BR should be based on administrative sources on business units. In Norway a good and very comprehensive cooperation between the administrative and statistical business registers has developed over decades. The Norwegian administrative legal unit register and the cooperation with the statistical BR is a model for other countries.

III. THE ADMINISTRATIVE LEGAL UNIT REGISTER (LUR)

6. The LUR was established in 1995. The official name is the Central Co-ordinating Register for Legal Entities. The LUR is a basic register that is made up of mainly legal units from state government registers. These are units such as employers, value-added units, corporations, foundations and the enterprises of the statistical BR. All other governmental administrative registers are obliged to use these same units with the registered information. A legal unit is identified by use of name, address and identification number. Most units of the LUR are registered in several administrative registers and much work has been done to ensure that each unit is registered only once in the LUR.
7. As most legal unit registers in the public sector were already linked to the BR since the 1970s, Statistics Norway (SN), through the Division for Business Registers, was heavily involved in creating the linkage between the different registers and the LUR. The work of SN in developing the new register to be covered by the LUR was organised as a paid commission.
8. One interesting quality of the Norwegian system is, under the legislation, a bank account has to be identified either by a PIN (person ID) or by the ID assigned by the LUR. The legislation related to the LUR ensures that a new business unit is registered in the LUR very early on in the process of establishment and that the LUR should obtain a very complete coverage of all kind of business units. The LUR covers all sectors of the economy including all units in agriculture and the governmental sector. Some of the variables of LUR are available to the public, e.g. name, address, ID number, NACE code and institutional sector.

9. The Division for Business Registers (DBR) in Statistics Norway is responsible for important functions in the operation of LUR. The statistical unit of establishment is registered in LUR as a sub-unit used by the Employer-Employee Register. Thus the DBR is responsible for profiling the unit of enterprise (or legal unit) into units of establishment. Also the governmental sector is profiled into legal units and establishments. In addition, intermediate units such as governmental agencies or institutions are being profiled by DBR. The idea is to create units as similar as possible to units of the private sector. Also DBR is responsible for the activity code classification, based on NACE, of all the units and for the institutional sector classification of the enterprises. The result of this cooperation is that, in principle, all the units and their NACE codes and institutional sector codes are exactly the same in the LUR and BR.

10. The legal unit (the enterprise) and the functional unit (the establishment) have their own separate ID numbers - BIN enterprise (BINen) and BIN establishment (BINes). BINen is a variable connected to the establishments so that we can identify to which establishment a legal unit belongs.

IV. THE STATISTICAL BR

11. When the statistical BR was established in 1965 only a few industries were covered. Business censuses and surveys were important sources for updating. Year by year, an increasing number of administrative sources such as employer and VAT registers were used to contract new business units from administrative sources. From 1995 the BR and LUR are integrated and the result is a full coverage of the total population of units.

12. For several years the DBR has worked to secure a sufficient quality of register as the source of demographic statistics of business units. Such statistics are being published for the first time this year.

V. THE UNIT OF ESTABLISHMENT TO BE REPORTED IN ADMINISTRATIVE DATA SYSTEMS

13. Usually administrative data systems are based on the unit of enterprise (or legal unit). In cases where the units of enterprise and establishment are the same unit, there are no problems in using such sources for statistical purposes. In cases where an enterprise consists of two or more establishment units, there could be problems in creating statistics by industry or geography.

14. Most enterprises consist of only one establishment. But the large enterprises are usually split into many establishments of different industries or geographic areas. The large enterprises represent a large share of turnover, value added and employment. So, for register-based statistics focusing on activity code or geography, it is important that the establishment unit be specified in administrative records.

15. One of the success stories of Norwegian register-based statistics is the specification of the unit of establishment in the Social Security data system of employers and employees - the EE-register. Firstly, the DBR profiles the enterprise units comprising more than one establishment, and registers the establishment units. Secondly, the employer is informed by the LUR about the establishment units and the assigned ID number (BINes) for each establishment unit. Thirdly, the employer is supposed to use the assigned BINes in reporting employees to the data system.

16. For many administrative purposes, it is irrelevant which establishments within the enterprise are being used for the reporting of employees. Therefore, it is SN that has to verify that the employer is using the correct BINes in the reports. Methods to check that the reporting of employees is related to the proper establishments are in operation.

VI. THE STATISTICAL UNIT OF JOB

17. The unit of job is a transaction in the SNA (System of National Accounts). An employed person is not. The unit of job has a key role in integrating economic and social statistics. The variables of the statistical unit of job should be grouped in 4 types: variables that identify the job, variables that are related to the unit of job, person and establishment.

18. The main variables of a job file by type are:

Variables that identify the unit of job

- PIN of the employed person
- BIN of the establishment if the job
- Date when the job started

Variables related to the employed person

- Age
- Sex
- Educational attainment
- Residence address

Variables related to the work place of the job (establishment)

- Address, locality
- Size group
- Economic activity
- BINen (and through this; institutional sector, ..)
- Demographic variables of the establishment

Variables related to the unit of job

- Main or secondary job
- Status in employment
- Hours paid for and actually worked
- Wage sum
- Occupation
- Socio-economic group
- The date when the job terminated.

VII. THE EXTENDED JOB FILE - A BASIC REGISTER?

19. The Swedish job file that is defined as a basic register includes units other than jobs. First, the unit of job should be extended to the unit of labour force, i.e. two units registered in Employment Service are included, spells of unemployment and periods in a labour market measured. Second, other time use activities should be registered in the file, for example persons active in an educational programme and in household work. For most persons, the listed time use activities are related to an income source, labour income, unemployment benefit, scholarship and student loan, and calculated value of household work. So the extended job file should cover available income sources. The variables of the extended job file would then be closer to the variables of the Population Census and the Eurostat list of harmonised core variables.

20. The reason for using the term 'basic register' of the extended job file is the important role of the job file in developing an integrated statistical system. A Norwegian Working Group is studying the Swedish report on strategies for developing statistics based on administrative sources to see what could be learned. Statistics Norway will probably not use the term 'basic register', but will continue to follow the traditional definition for a basic register.

VIII. LABOUR MARKET STATISTICS

VIII.1 Register based employment statistics (RES)

21. Statistics Norway produces employment statistics based on data from different registers. When defining employment, the main register is the EE-register run by the Social Security Authorities. The employers have to send information on employees starting and quitting during the last month. The employees are identified by their personal ID number (PIN) and the employer has to use the BINes for the establishment in which the person works. The information about the employee and the establishment is represented by their ID numbers in this EE-register. Concerning the job, the employer gives data on occupation, contractual working hours per week and date for start and stop. To cover self-employment, SN use some other administrative sources, mainly from the tax authorities.

22. To collect information on variables connected to the employee and the establishment, SN uses the ID numbers in the EE-register (PIN and BINes) to link to other registers. SN is also able to collect more information on variables connected to the job. As a linking number, we then use the combination of PIN and BIBes and, in addition, the date when the job started. A job is not defined as precisely as this in all registers, so this linking is not straightforward.

23. Concerning variables of the workplace/establishment, we link to the statistical BR. The quality of establishment variables depends on:

- to what degree the employer is registered with all his establishments;
- to what degree the employer connects the employees to the correct establishment;
- the coding quality of the variables in the statistical BR;
- timeliness for updating the variables in the statistical BR.

24. The responsibility for updating units and values of variables and to check quality in the statistical BR is extended over different divisions in Statistics Norway. Those divisions producing structural statistics have the responsibility concerning issues 1, 3 and 4 for their respective industry codes. The Division for Employment Statistics is the main responsible unit for issue 2. The Division for Business Registers has an overall responsibility for the register, standards and definitions and most of the connection to external governmental bodies responsible for or using LUR.

25. There are several advantages to using the statistical BR in this way:

- We reduce the response burden for the employers. There is no specific employment survey of the employers in Norway;
- We improve quality: both the economic and social statistics have a common interest in improving the same statistical BR. Resources used for improving this register in one division are a benefit for all other divisions as well;
- We obtain more consistent data between different statistics;
- We produce statistics using less resources. The employee statistics do not have to maintain a separate business register or to code economic activity and other establishment/enterprise variables.

26. A main problem when producing short-term employment statistics in this way is that the updating of the industry code is not very timely, since it is connected to revisions in structural statistics.

VIII.2 Developing an employer-employee linked database

27. Traditionally, longitudinal analysis has not been the primary focus of administrative registers. Registers focus on the provision of an accurate snapshot of the employee or business population at a point in time. The increase in comprehensive national data systems, both national and international, has led to a number of studies concerning job and labour force mobility, and other studies of the dynamic aspects of the

labour market. To meet the demand for standardised data in this field, Statistics Norway has, since 1997, been working on constructing a longitudinal database of linked employer-employee data from 1992 on. The project has been influenced by work done in other countries such as Sweden and Denmark, who have already constructed similar kinds of databases.

28. The information/data is collected from different administrative registers, the EE-register being the most important since it provides a link between natural persons and establishments. As already outlined, the statistical BR is the data source for data on establishments. In respect of quality on longitudinal data, it is important that there is a high quality of data on birth and death of an establishment.

VIII.3 The Labour Force Survey (LFS)

29. The Norwegian LFS is a continuous survey covering 8000 persons each month. The sample of persons is determined before the interviews start by using a statistical version of the administrative Population Register for extracting the sample. This way, we have the PIN code on each person in the sample.

30. When coding workplace variables, we make use of the BR. The sample of persons in the LFS is linked on a micro level (using PIN) to the register-based employment statistics (RES) before the results from the interviews are finished. When coding, the coding staff have information on the screen concerning the name and address of the establishment in both LFS and RES. In RES this information comes from BR. If the establishment is considered to be the same in the two data sources, the coder presses a button for accept and the data from RES/BR are transmitted to the LFS file. In LFS the respondents sometimes mention the name of the enterprise instead of the establishment. This is also then usually considered to be the same job and the data from RES is used.

31. If there is no job found in RES or RES indicates a quite different job, the coders seek the name of the establishment, given in the questionnaire, in the BR. If the establishment is found, they put in the BINEs and establishment/enterprise variables are transmitted.

IX. Wage statistics annual and quarterly

32. The annual survey of wages is based on a sample of enterprises. The wage statistics are annual and cover all industries except agriculture, hunting, forestry and fishing (sections A and B of the Standard Industrial Classification), apart from fish hatcheries and fish farm which are covered from 2002. The population covers all enterprises in Statistics Norway's Central Register of Establishments within the relevant industries according to the Standard Industrial Classification (Norwegian version of the NACE rev. 1). Each enterprise covers one or more establishments grouped by industrial category. The data for wage statistics are obtained for each establishment at the person level via forms or electronic media from the units covered by the sample. Information is obtained on wages, bonuses and commissions, variable additional allowances, overtime, occupation and working hours of the individual employee in the establishment. The wage statistics makes use of several variables related to the establishment as they are defined in Statistics Norway's Central Register of Establishments and Enterprises.

33. Furthermore Statistics Norway produces a quarterly wage index which covers fewer industries and is based on a more limited sample. Apart from this it covers all the variables also covered in the annual survey except identification of occupation.

X. Annual wage statistics from The Register of End of the Year Certificates

34. This register is obtained from the Norwegian Tax Administration. Among other uses for this register one is the annual publishing of wage sums. In this register both persons and establishments are identified so that it is possible link information from the Central Register of Establishments and thereby

utilize information from this source. Information concerning age and sex for individuals is derived from the identification number of the individual. However information on occupation, length of work and working hours are not accessible in the finalized dataset used for publishing figures on wage sums.

XI. STATISTICS ON INCOME FROM WORK

35. From the Tax Agency, Statistics Norway receives a micro file on wage sums per job identified by PIN and BINen. To connect the job to the establishment, this file is linked to RES where, for most jobs, we have information regarding the establishment to which the job is connected.

36. The Tax Agency also gives micro data on self-employed identified by PIN. For many of the self-employed there is a link between PIN and BINen in the LUR/BR.

XII. POPULATION AND HOUSING CENSUS

37. The 2001 Population and Housing Census in Norway is partly traditional and partly register-based. The population part of the Census is based on administrative records and the housing part on questionnaires. Basic registers of persons and establishments have existed for a number of years. As a part of the census project a basic register of dwellings is being established (integrated into the Register of Ground Properties, Addresses and Buildings). This means that future censuses will be totally register-based.

38. In the 2001 Census, the labour market variables are for the first time based on registers only. This means that all variables related to the unit of job are based on the BR. In the 1990 Census, the labour market variables were based on a combination of registers and information from questionnaires. The BR was used both in the register data and in the coding of industry. Also in the 1980 Censuses, the BR was used in the coding of industry.

39. In a traditional census where a business register is not available, other methods must be used. The questionnaires provide information on all work places. This means that the census micro-file implicitly specifies a business register. A measure for coding of industry is to make a list of all establishments of the municipality and use these lists for editing and coding.

40. The 2001 Census uses the same data sources as the annual labour market statistics. This means that the requirement for the BR in general is the same as stated in the chapter on labour market statistics. In previous censuses there have been several variables describing commuting or journey to work: Distance, frequency and mode of transport. In a register-based census, the only variables available will be place of residence and place of work. This means that the information on journey to work will be less comprehensive in a register-based census than in a traditional census.

41. Furthermore, in previous censuses employed persons have been classified by place of reporting for work. For some persons this will not be the same as the address of the establishment. Some persons have been classified as reporting for work in different places during the census week, for instance construction workers. In a register-based census, all employed persons will be classified by place of work, that is, the address of the establishment. It is not possible to distinguish between the construction workers and their colleagues working in the office, since they all belong to the same establishment (according to the definition of establishment). Information on occupation may be used to identify persons who work in different places, but this information is not yet available from registers.

42. The BR has until recently only held information on the location of working places on the municipality level, but no information on exact location. This means that SN has not been able to produce statistics on place of work for small geographical areas (below the municipality level). As a part of the of the census project, numerical address have been assigned for most working places in the BR. The Division of Business Registers in SN has been in charge of this project. This will make it possible to link the BR to

the Register of Ground Properties, Address and Buildings and thereby transfer detailed geographical information on working places (geographical coordinates, local statistical units, urban settlements).

43. The Census file comprises exact location for both place of residence and place of work (this data will also be available on a current basis). By using GIS technologies, it will be possible to compute the length of journey to work, that is, the distance between the place of residence and the place of work. Such a method has not been developed as a part of the Census project, but the necessary data have been established.

44. In Norway the plan is to produce a Census micro file annually and to include new data sources as they improve. The strength of the Census micro file is twofold. First, the file is a source for small area statistics. Secondly, the Census file is a linked and integrated micro file of most of the statistical domains of social statistics: demography, education, labour, transfer and income. Variables of importance for developing integrated statistics are main time use activities, main source of livelihood, and main social status.

45. A traditional population census specifies all work places. This means that the census micro file implicitly specifies a business register. A measure to obtain the coding of the industry of a work place close to the BR is to make a list of all establishments of the municipality and the local labour market and use these lists for census editing and coding.

46. The Population and Housing Census is based on administrative records and statistical sample surveys. Basic registers of person, dwelling, and establishment are needed. In principle, the data sources are sufficient to also include a Business Census. If the BR could serve as a source for demographic statistics of the unit of establishment and enterprise, the practice should be to create a Population, Housing and Business Census.

XIII. OTHER HOUSEHOLD SURVEYS

47. The use of BR in other household surveys is as described for the labour force survey.
