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

**Register-based census 2010 and census test 2006
in Austria**

Note by the Statistic Austria

INTRODUCTION

1.   In the year 2010 Austria is replacing the traditional census with a completely register-based census. The legal basis is the register-based census law of 16 March 2006. This law also stipulates that a census test will be carried out in October 2006.

2.   The census test with the reference date of 31st October 2006 covers the complete federal territory as the register-based census of 2010 will. The only difference to the regular census in 2010 is the fact that no legal consequences follow. The data collection and the processing of the data collected in the census test almost completed. We will provide a report on the results of the census test to the Federal Government in April. For the census test we tested all following described procedures and are able to present results and first evaluations at the presentation in May.

GE.08-

I.   THE CENSUS TOPICS

3.   The census topics are:
   a) Population census: people with residence in Austria (PC)
   b) Housing census: buildings and dwellings (HC)
   c) Census of enterprises and their local units of employment (CE)

II.   THE PRINCIPLES OF THE NEW REGISTER-BASED CENSUS

4.   Due to the fact that the registers were not connected with each other in the past and that the data has been collected independently, they often contain different values for the same variables of the same person. Therefore, it is not practicable to collect data only from one register and to trust it. To get data of a satisfactory quality, the principle of redundancy has been used: Data on sex, date of birth, nationality, place of residence, to mention just a few examples, is collected from as many registers as possible. Inconsistent data has to be checked and converted to plausible values, if necessary, also with the help of the registers or – in fewer cases – of the people concerned.

5.   To make matters even more complicated, the data protection commission did not allow the data to be matched by a commonly used personal identification number which is already included in the several registers (like the social security PIN). Instead, a tool has to be used which was created for e-government, called branch-specific personal identification number "Official Statistics" (bPIN OS, see below).

6.   The following points outline the principles of the new concept:

   a) Eight "base registers" are used (mentioned below). A decision was made about which register to use as the base register for each variable (if more than one register contains this variable). Many variables are also collected from seven "comparison registers" (mentioned below), which are used as proof for the confirmation of the values in the base registers (principle of redundancy).

   b) The data will be collected without the registered names and without the social security PIN, which both have to be replaced by the branch-specific personal identification number (bPIN OS, see below) in the delivery of the data to Statistics Austria for data protection reasons.

   c) The data from the registers will be matched using the bPIN OS and afterwards checked for consistency and adjusted according to plausibility rules.

   d) Remaining gaps in the registers have to be closed by statistically sound estimation procedures.

   e) Not the full range of the data available in the base and comparison registers is collected for the census. Only those variables are gathered which have already been used in previous censuses.

f) As some variables are not available in any register, it is not possible to collect them using a register-based census, e.g. colloquial language or some characteristics of commuting like the duration of the journey or the means of transport. The biggest disadvantage of the register-based census is the fact that the variable "occupation" is not included in large registers and is therefore not collected in the census, although it is a core variable of censuses in the UN census recommendations. This variable will therefore not be available at local level in Austria, but only at regional level of NUTS II ("Länder") using a different source like the labour force survey.

g) A lot of work has been done and is ongoing to improve the quality of many registers to fulfil the requirements of the census. In some cases even further variables have been added to a register. The most important example is the address of the place of work which has been added to the social security records to provide annual commuting data.

h) In order to check inconsistencies in the data Statistics Austria is entitled to get in contact with the respective register authority.

i) If further investigation is necessary where there are severe inconsistencies or there is the suspicion that a person has already left Austria in spite of being included in the central population register (CPR), the names and addresses of people have to be delivered from the register to Statistics Austria. According to the response or non-response to a letter addressed to the people concerned, the data can be clarified or a decision has to be taken on whether to include the people in the census or not. → residence analysis

j) In cases where Statistics Austria intends to exclude a person from being counted in the municipality in which they are registered in the central population register as having their main residence[1], this municipality has to be informed of the intended non-inclusion to enable it to apply for re-inclusion with proof of the existence of the main residence.

III.   THE REGISTERS

6.   For each variable it is regulated by law which register is the base one and which registers are the comparison ones.
Data owners are obliged to help Statistics Austria to complete the data. The data transmission is electronic.

---

[1] The definition of "main residence" is similar to the UN definition of the "usual residence" in the UN Recommendations on Statistics of International Migration.

A.   THE BASE REGISTERS

7.   For the register-based census Statistics Austria receives data for the reference date from eight base registers containing the variables, as required by law. Statistics Austria is the owner of four base registers.
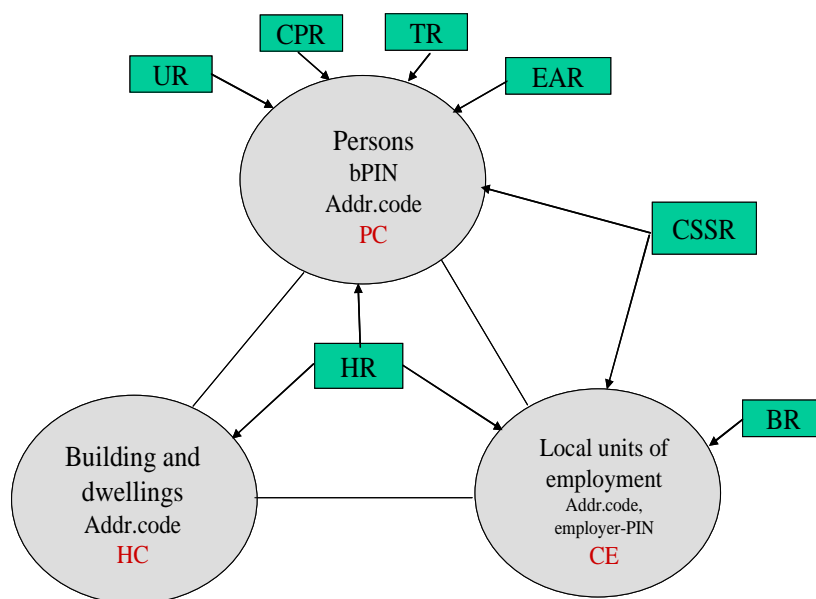
Statistics Austria's registers:

a) Business register of enterprises and their local units (BR)
b) Housing register of buildings and dwellings (HR)
c) Register of educational attainment (EAR)
d) Register of enrolled pupils & students (PSR)

Other registers:

e) Central population register (CPR)
f) Central social security register (CSSR)
g) Tax register (TR)
h) Unemployment register (UR)

Figure 1

The data from these registers contains all the variables required for the register-based census.

B.   THE COMPARISON REGISTERS

8.   To keep the quality of the census results as high as possible, the base registers will be compared with the comparison registers (principle of redundancy).

Seven registers used only for comparison:

a) Child allowance register (CAR)
b) Central foreigner register (CFR) - available as of 2008
c) Registers of public servants of the federal state and the Länder (RSP)
d) Register of car owners (COR)
e) Register of social welfare recipients (SWR)
f) Conscription Register (CR)
g) Register of alternative civilian service (ACSR)

9.   Further comparison registers are the base registers for each other, because for many variables the base registers are used for comparisons with other base registers too.
The comparison-only registers contain primarily basic demographic data e.g. main residence, nationality, sex and information about employment e.g. full-time or part-time. On the other hand, they provide very special information, like place (local unit) and branch (NACE) of work, or about people doing military or alternative civilian service.

IV.   MATCHING AND ADJUSTING BASIC AND COMPARATIVE DATA

A.   MATCHING OF PERSONS

10.   The matching is conducted by a "branch-specific personal identification number for official statistics" (bPIN OS). The complete governmental administration is divided by law into several branches like "social security", "taxes", "health", "official statistics" and so on. Each branch has its own branch-specific PIN. These PINs, which should serve for the data protected communication between public authorities within e-government, are derived by the Austrian Data Protection Commission (DPC) from the central population register by the following procedure.

11.   The owner of the register has to ask the DPC for this PIN for each person by sending the name, date of birth, birth place and address. The bPIN OS as well as the other bPINs are derived from the PIN in the central population register (CPR) using a special and very complex algorithm known only by the DPC. The bPIN OS is given to other register owners than Statistics Austria only in an encrypted (ciphered) form. The owner of the register has to send the data together with the encrypted bPIN OS and an accompanying number of their own to Statistics Austria. The latter number serves to identify the respective record in case of further inquiries by Statistics Austria.

12.   Only Statistics Austria is able to decrypt (decipher) its "own" PINs, which were dedicated for the use for statistical purposes, namely the bPIN OS, and to use these PINs as a common matching variable.

13.   In practice, Statistics Austria has the problem that the names and dates of birth in the registers are sometimes inaccurate. "Data-twins" and other people who cannot be found by the procedure mentioned and therefore cannot be matched pose problems. For solving this problem Statistics Austria uses deterministic record linkage methods.

B.   MATCHING OF ADDRESSES

14.   As the diagram shows, the housing register of buildings and dwellings is very important as it is principally able to connect all addresses in the registers by a numerical address code. The housing register is linked with the central population register and with the business register. These registers contain the same addresses (numerical codes) for buildings but not always for door numbers. The register of buildings and dwellings is highly reliable concerning the buildings. As far as dwellings are concerned, the matching of dwellings with people registered in the central population register is less successful due to many missing door numbers. In general, the matching of the central population register, the housing register and the business register is much more reliable for buildings than for dwellings. This may pose problems for the identification of households and families.

C.   LINKING OF PERSONS TO THE BUSINESS REGISTER

15.   Each record of an employee in the central social security register includes a PIN for the employer. This employer-PIN can be used to match the records of employees with the business register, where it is used as an identification variable.

V.   THE POPULATION CENSUS

A.   THE POPULATION

16.   The base information for the population comes from the central population register, e.g. main residence, date of birth etc.

17.   There are some special rules for people who change their main residence in the central population register around the reference date of the census. Firstly, immigrants are counted as being mainly resident in Austria only if they stay for at least 90 days. Otherwise, they are considered as mere visitors who are not part of the population of Austria. The same happens with emigrants the other way round: If people deregister to leave the country and register again within 90 days in Austria, they are considered as being absent only temporarily and therefore as having lived without a break in Austria (in the same or previous municipality). This rule is consistent with the definitions given in the UN Recommendations on Statistics of International Migration.

18.   Secondly, another rule should reduce so-called "census tourism": In the past, due to pressure by mayors, a personal attachment to the municipality of their secondary residence or to the residence where they grew up, a lot of people changed their main residence shortly before the reference date of the census and returned to their previous residence shortly afterwards (not in reality, but in their register record). For this reason, there is a special procedure if they migrate

back to the former municipality after the reference date, within 180 days of the first migration. They have to be counted in the former municipality in the census (but not in the central population register), although they were registered in the other municipality on the reference date.

19.   In order to be able to count people who are recorded in the central population register also in the census population it is necessary to check if there is information about them in the other registers. For that purpose the central social security register is the most important comparison register. An employment or pension entry can be found in this register for the vast majority of the population. If people become unemployed or are receiving social welfare benefits, an entry is found in the unemployment register and the register of social welfare recipients.

20.   If an entry can only be found in the central population register without any entries in other registers, a special procedure called "residency analysis" is used which is a signs-of-life analysis. Statistics Austria is authorised to ask the registrar for the names and addresses of the people concerned, to whom a letter is addressed that they are obliged to answer. If there is no reaction to the letter and there are no other signs of existence, the person is supposed to be deleted from the census records. As already mentioned in the introductory section on principles, in this case the municipality in which the people are registered as having their main residence[2] has to be informed of the intended non-inclusion to enable it to apply for re-inclusion with proof of the existence of the main residence. Otherwise, or if there is a negative decision by Statistics Austria the records of the people concerned have to be deleted.

B.   HOUSEHOLDS AND FAMILIES

21.   The lack of many flat door numbers in the housing register poses some problems for the household and family statistics, as already mentioned in section 0. It is planned to use not only the information on door numbers from the two registers mainly containing that information, i.e. the central population register and the housing register, but also information from other registers with addresses. Additional sources could be used, e.g. information about living in the same household from the child allowance register.

22.   It is partially possible to deduce the relationship between the members of a family from the register data to construct the family statistically and to build the variable "family status" which means the position of a person within the family (parent, child). There is basic data available from the social security register for people who were sometimes out of the labour force and therefore were not self-insured but covered by a family member's national health insurance. For these people their relationship to the family member is registered. One of the comparison registers is the child allowance register with information about the parent–child relationship for children up to 18 or, if studying, up to 27 years of age. The remaining data gap has to be closed by estimates.

---

[2] The definition of "main residence" is similar to the UN definition of the "usual residence" in the UN Recommendations on Statistics of International Migration.

23.   The necessary estimation procedures may be based not only on the register data but also on data from the labour force survey which is a very large sample survey.

C.   HINTS ON SOME VARIABLES OF THE POPULATION CENSUS

24.   Information about marital status should come from the central population register as the corresponding base register. However, this register collects the marital status only for new registrations or registrations changed since July 2006. Only partial or even rudimentary information exists in the miscellaneous comparison registers.

25.   Statistics Austria obtains the base information for the employed and pensioners from the social security register and the tax register. To analyse other kinds of livelihood Statistics Austria uses the registers of military and alternative civilian service, the unemployment register, the register of social welfare recipients and information from the central social security register on people who are out of the labour force but covered by a family member's national health insurance. Information on pupils and students concerning their enrolment, the type of school or university they attend, the field of studies and the location of the school or university is provided by the register of enrolled pupils and students.

26.   The register of educational attainment is the base register for the highest completed educational level.

27.   There is no data available in the registers for the some variables although these variables have always been collected in Austrian censuses:

  a) Occupation
  b) Mode of transport for commuters
  c) Duration of the daily commute
  d) Language(s) most commonly spoken
  e) Religion

28.   Most of them are only optional according to the UN Recommendations on Population and Housing Censuses, with occupation as the only exception which is a core variable in censuses. The easiest way to collect these variables is to use a sample survey, the data is matched with the census data. However, this data will not be available at local level because the largest sample survey in Austria is the labour force survey which is only representative for the NUTS II-regions ("Länder").

29.   Another possible way of collecting these variables is offered by the register-based census law: It authorises the government to enact a regulation which stipulates the need for a special survey on the variables "religion" and "language(s) most commonly spoken". However, such a regulation has not passed the Council of Ministers yet, presumably due to the costs of such a survey.

VI.   THE HOUSING CENSUS

30.   The base register for the housing variables is the housing register of buildings and dwellings. No comparison register is available. There are almost no problems analysing the different variables for buildings. However, in urban areas with apartment buildings, there are the previously mentioned problems of the partial lack of door numbers (see 0). It is necessary to use estimation methods for this group of variables too. Again, an important source for sound estimations is the labour force survey.

VII.   THE BUSINESS CENSUS

31.   The base register for the variables of enterprises and their local units of employment is the business register. It contains aggregated numbers of employed people for the whole enterprise or institution as well as for the local units of employment.

32.   For comparisons in the public sector Statistics Austria is able to use the registers of public servants of the federal state and the "Länder".

33.   Statistics Austria has more or less complete data on enterprises (concerning the number and structure of employees), but less information on local units of employment. Therefore, enterprises with only one local unit of employment are well registered, but those with more than one local unit are partly incomplete. The main data source for the latter is the structural business survey which covers NACE sections C-K for enterprises with 20 or more employed people (NACE C-F) or above a specified turnover (NACE G-K). Only limited information is available for local units of enterprises below these thresholds and for those which belong to other economic branches (NACE sections L to O).

34.   Statistics Austria will receive better information for those local units of employment from 2008 onwards. Each employer is obliged to submit the address of the place of work of each employee together with their social security record and their tax record to Statistics Austria once a year. This information will be used to provide individual commuting data and will also be incorporated into the business register in terms of the number of employed people in the local units.

VIII.   DATA PREPARATION

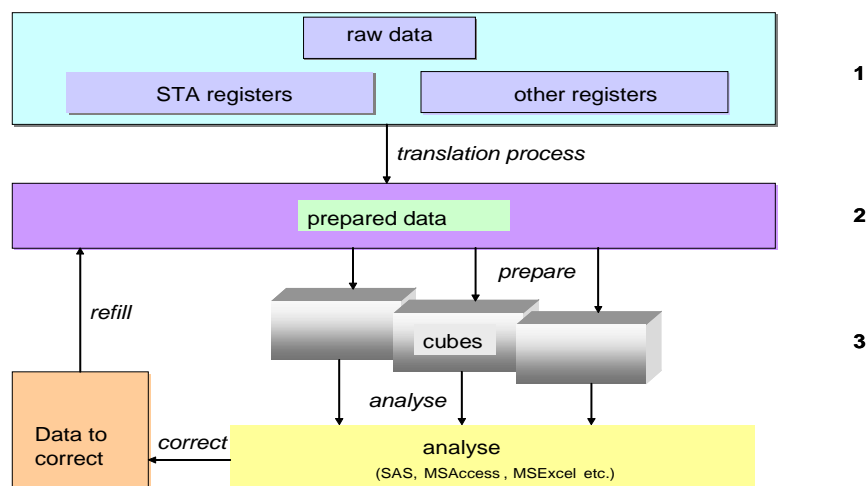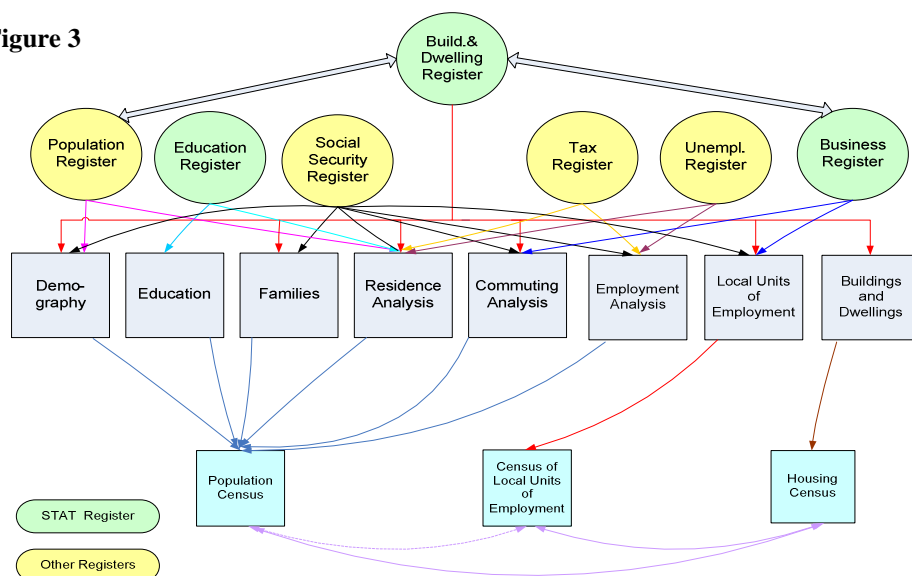35.   Following scheme shows the way of preparing the data:

Figure 2

36.   First the raw data is stored in the so-called staging area (1), in a next step the data is prepared (2), which means linked on personal level and checked for plausibility during a translation process. Then OLAP Cubes are built for analysis. With OLAP technology data can be displayed in a multidimensional way by using measures and dimensions. A dimension, for example, is the geography dimension, which has as top level Austria, then nine Länder (federal provinces), further down municipalities to the base level of each dwelling. Using OLAP cubes one can view the data on the required level of aggregation for each dimension. On micro data level further checks of plausibility are carried out and corrected in a loop back process, rebuilding the cubes with corrected data. This ensures that every analyst uses the latest version of data.

A.   ANALYSIS-CUBES

37.   There are cubes for all areas of analysis. The following figure shows all basic registers as well as the cubes for analysis for register based census test 2006:

Figure 3

### B.	MULTIPLE ATTRIBUTES

38.	Attributes with origin from more than one register are called multiple attributes. For each of these multiple attributes there is one analyst responsible, who defines the rules for plausibility and valid values per person.

39.	At the moment the following multiple attributes are used in the test register based census:
 a)	Marital status
 b)	Date of birth
 c)	Country of birth
 d)	Sex
 e)	Nationality
 f)	Main residence

### C.	EXAMPLE SEX:

40.	To illustrate the process of analysis the example sex is described step by step. Information on sex is supplied from seven different registers, the base registers shown in figure 2 as well as three comparison registers.

 a)	Central population register (CPR)
 b)	Central social security register (CSR)
 c)	Tax register (TR)
 d)	Unemployment register (UR)
 e)	Register of car owners (COR)
 f)	Register of social welfare recipients (SWR)
 g)	Registers of public servants of the federal state and the Länder (RSP)

41.	The area of analysis "demography" is responsible for defining the valid value. During the translation process the rules for deciding the valid value are determined.

42.	Example for Rules:

R1:	Same content in all sources
R2:	If there are only data from CPR, this is the valid value
R3:	If CPR_SEX is unknown, CSR_SEX is the valid value
R4:	If CPR_SEX and CSR_SEX are equal, this is valid (even if it is not consistent with other sources)

43.   For further analysis information on sex is shown in detail on a personal level in the analysis cube „Demography".

cube demography										
PID	sex_CPR	sex_CSR	sex_TR	sex_UR	sex_COR	sex_SWR	sex_RSP	sex_valid	sex_rule	sex_change
124586	1	1	1	0	0	0	1	1	1	
124587	1	0	0	0	0	0	0	1	2	
124588	9	2	2	0	0	2	0	2	3	
124589	2	2	1	1	0	1	0	2	4	

PID = Personal identification number

Change = Change field

44.   During the process of analysis there can appear cases where these rules cannot or should not be applied. Therefore the analyst has the possibility to change the multiple attribute via loop back process (Field change).

D.   CENTRAL ATTRIBUTES

45.   Apart from multiple attributes there is the concept of central attributes – meaning variables that are used in more than one cube of analysis. Similar to the multiple attributes there is always one area of analysis responsible for changes. Only in the cube belonging to this area the central attribute can be changed. A multiple attribute can be a central attribute as well, e.g. sex, which is used in all cubes containing persons. The value fixed in the demography cube is then used in the other personal cubes.

46.   If at least one value of a central attribute is changed, all analytic cubes containing this attribute have to be rebuilt to guarantee consistency.
Examples for central attributes:

   a) quality of residence
   b) geographical dimension
   c) sex

47.   The prepared data contains information of the building and dwelling register (Figure 2) for the geographical dimension. All cubes (for example the cube for demography, the cube for local units of employment, the cube for families and the cube for housing census) use the same geographical dimension.

E.   RESIDENCE ANALYSIS

48.   A main result of the register based census is the number of Austrian citizens and the living population of Austria, determined by an analysis of residence deals with this question.

49.   To be part of the Austrian population a person has to appear both in the central population register and one or more comparison registers. Otherwise it is a case of clearing and further examination. For this analysis the main information lies in the fact that the person appears in another register at all, so mainly flags are used in this cube.

cube residence analysis							
PID	flag_CPR	flag_CSR	flag_TR	flag_UR	flag_COR	flag_SWR	flag_RSP	resid. valid
789101	1	1	1	0	0	0	1	1
789102	1	0	0	0	1	0	0	1
789103	1	1	1	0	0	1	0	1
789104	0	1	1	1	0	1	0	1

50.   Within the residence analysis cube all persons - also those not (uniquely) linked during translation process – are represented. With different record linkage procedures these persons are linked later in the process.

F.   REGISTER CHECKS

51.   Using the procedure as stated above, single registers as well as consistencies between registers can be analysed and evaluated. This should lead to a relatively small amount of clearing cases. Those cases will be cleared, leading to deleting or approving a main residence, by obtaining further information from registers ("sign of life") or questioning all people relevant.

IX.   DISSEMINATION OF RESULTS

52.   Statistics Austria has to publish the census data within a year of the last data transmission from the registers.

X.   SPECIAL FEATURES OF THE 2006 CENSUS TEST

53.   The register-based census test of 2006 is a full register-based census. However, it has no financial consequences for the municipalities and no consequences for the number of votes necessary to win a seat in parliamentary elections. Statistics Austria is checking the results of this test by comparing them with a sample survey with the same reference date. The questionnaire only contains items designed to check the quality of the registers.

54.   The Modalities of the sample survey:
   a) A surface area sample of 25,000 people = three per thousand of the total population, in 100 sample areas
   b) The people concerned were obliged to answer the questionnaire

XI. FUTURE WORK

a) Evaluation report on the census test to the federal government by April 2008
b) Amendment of the concept by summer 2008
c) Required modifications of regulations from autumn 2008 onwards
d) Preparation of the 2010 register-based census
e) Implementation of a register-based census with the reference date 31th October 2010

*****