UNITED NATIONS **E**

Economic and Social Council

Distr.
General

ECE/CES/2006/28
30 March 2006

Original: ENGLISH

ECONOMIC COMMISSION FOR EUROPE **STATISTICAL COMMISSION**

CONFERENCE OF EUROPEAN STATISTICIANS

Fifty-fourth plenary session
Paris, 13-15 June 2006
Item 6 of the provisional agenda

SEMINAR ON POPULATION AND HOUSING CENSUSES
SESSION II

The Dutch Virtual Census of 2001: a register-based approach combined with survey information[1]

Submitted by Statistics Netherlands

I. ABSTRACT

1. Data from many different sources were combined to produce the Dutch Census tables of 2001. Since the last Census based on a complete enumeration was held in 1971, the willingness of the population to participate has fallen sharply. Statistics Netherlands found an alternative in the Virtual Census, using available registers and surveys. The table results are not only comparable with the earlier Dutch Censuses but also with those of the other countries in the 2000 Census Round.

2. For the 2001 Census, more detailed information is required than was the case for earlier Census Rounds. The acquired experience in dealing with data of various administrative registers for statistical use enabled Statistics Netherlands to develop a Social Statistical Database (SSD), which contains coherent and detailed demographic and socio-economic statistical information on

[1] This paper has been prepared at the invitation of the secretariat. The views expressed in this paper are those of the author and do not necessarily reflect the policies of Statistics Netherlands. This paper can be considered as an updated and shortened version of [8].

persons and households. The Population Register forms the backbone of the SSD. Sample surveys are still needed for information that is not available from registers.

3. To achieve overall numerical consistency across the Census tables set of 2001, the methodologists at Statistics Netherlands developed a new estimation method that ensures numerically consistent table sets if the data are obtained from different data sources. The method is called repeated weighting, and is based on the repeated application of the regression method to eliminate numerical inconsistencies among table estimates from different sources.

II. INTRODUCTION

4. In 2003, data were combined to produce the Dutch 2001 Census tables. In the Netherlands, this was done using data that Statistics Netherlands already had available rather than by interviewing inhabitants in a complete enumeration. This way, the Dutch taxpayer received a much lower census bill. The costs for a traditional census would be about three hundred million Euros, while the costs using this method are 'only' about three million. The estimate includes the costs for all preparatory work such as developing a new methodology and accompanying software. The costs of the registers are not included, but the analyses of the results are. Registers are not kept up-to-date for censuses but for other purposes. Saving money on census costs is only possible in countries that have sufficient register information. As an example, we can compare the costs of the Dutch Virtual Census of 2001 with the costs of the traditional Census that was held in Canada. In Canada, the census costs amounted to approximately 450 million Euros. Canada has about 31.6 million inhabitants, twice as many as the Netherlands. A virtual census would be impossible in Canada because of the lack of sufficient register data.

5. The 2001 Census relates to forty extensive tables. Twenty-eight are about the Netherlands as a whole, nine are at the COROP level (NUTS 3) and three at municipal level (NUTS 5). The forty tables fall into a number of groups. Eight tables concern housing, two tables concern commuting and the other thirty tables are demographic tables, relating to occupation, level of education and economic activity. Additionally, demographic, housing and labour figures are compiled at sub-city district level for ten large cities that participate in Urban Audit II [9].

6. Except the financial aspect, other important differences exist between a traditional census and the virtual census conducted in the Netherlands. In spite of the mandatory character of a traditional census, a certain part of the population will not participate (unit non-response) and the part that does participate will not answer certain questions (item non-response). Correcting non-response by weighting and imputation techniques is worth trying. A well-known problem with traditional censuses is that participation is limited and selective. Traditional correction methods fall short of the need to be able to publish reliable results. The last traditional Census in the Netherlands (in 1971) met with much privacy objections against the collection of integral information about the population living in the Netherlands. This increased the non-response problem and the expectation was that non-response would be even higher if another traditional census were held in the Netherlands [1]. There are almost no objections to a virtual census and the non-response problem only plays a role in the surveys of which the data are used. If non-response can be corrected in a survey, it will certainly be possible to correct for the selectivity of that survey in the census where it is used.

7. The Virtual Census in the Netherlands was off to a later start than in other countries where a traditional Census was conducted. It did not make sense to begin the 2001 Census Project until all sources were available; some registers were available relatively late. Nevertheless, the Netherlands was quicker with the compilation of the forty census tables than most of the other countries that participated in the 2000 Census Round. In fact, the Netherlands was one of the first to send the complete set of forty tables to Eurostat, which coordinated the contributions of all European Union (EU) member states, accession countries and European Free Trade Association (EFTA) member states. The Netherlands had the advantage that the incoming census forms did not need to be checked and corrected. However, it must be noted that for some variables only sample information is available, which implies that it was impossible to meet the level of detail required in some Dutch tables.

8. An interesting option for the future is to use small area estimation techniques to estimate the cell values that could not be estimated adequately. A theoretical framework for small area estimation can be found in [6]. The ONS studied the application of this technique in the context of its Neighbourhood Statistics Programme. This is a major initiative to bring together and make widely available statistics on a small area level. In each case of implementation of indirect small area estimates particular attention was paid to model specification. Some experimental synthetic estimates were published in the United Kingdom. Currently, the Netherlands is conducting small area estimation studies on its Labour Force Survey. If these experiments turn out to be successful, possibly the techniques of repeated weighting and small area estimation can be combined in the 2010 Census Round.

9. Currently, the advantages of the virtual census in cost and non-response problems amply make up for the loss of some detail compared to a traditional census. Moreover, not all required information will consistently be available for the users in traditional censuses. This is because traditional correction methods such as weighting and imputation sometimes do not correct for limited and selective participation. This means no reliable results can be published for some of the cells in the set of tables. One may wonder why simply applying mass imputation (filling in valid values for all missing scores) is not considered to overcome these problems. An important advantage of mass imputation is that once the records are imputed, any user will be able to reproduce results when using the same imputed file. However, mass imputation is not a viable strategy for raising survey outcomes to population totals. There are not enough degrees of freedom to sustain a sufficiently rich imputation model accounting for all significant data patterns between sample and register variables. Only if the interest is in totals of subsets of the population defined by the explanatory variables in the model does the imputation approach lead to (almost) design-unbiased and hence reliable estimates (at least if the variances are reasonably small) [5].

10. The Nordic countries (Denmark, Finland, Iceland, Norway and Sweden) have currently more variables available in registers than the Netherlands. So the problem of insufficient detail in the outcome does not play a major role there. Moreover, some Nordic countries conducted a (limited) enumeration for variables missing in the registers. Most of the other countries are in a similar position as the Netherlands where some variables relevant for the census can be found in registers, while other variables are available on a sample basis only. That’s why much interest exists in the Dutch approach to combine registers and surveys and to use modern statistical techniques and software to compile the tables. It is, of course, crucial that statistical bureaus be

able to make use of registers that are relevant for the census. For Statistics Netherlands, this possibility was strengthened in the new statistical law that came into force in the beginning of 2004. Nevertheless, in the years to come Statistics Netherlands will have to establish good contacts with register holders. Timely deliveries with relevant variables for Statistics Netherlands are essential for statistical production.

11. The reason why Statistics Netherlands has compiled the set of tables is a gentlemen's agreement. In 1991 the Census Act was rescinded, officially cancelling Statistics Netherlands' obligation to hold a census once every ten years [1]. There is no European obligation to supply census data of the 2000 Census Round, but it is inconceivable that the Netherlands would not compile census data for the international organisations just like all other European countries do. Eurostat has a coordinating role in collecting harmonised data on the EU and a duty to make international comparisons of the outcome.

12. It takes several years before all countries participating in the 2000 Census Round have sent their final set of tables to Eurostat. Therefore, Statistics Netherlands took the initiative to compare the 2001 results of a limited number of European countries. The results of the Dutch 2001 Census were also compared to earlier Dutch Censuses. Such work has been carried out in the past as well. The data compiled on 1981 and 1991 were much less detailed than the set of tables of the 2001 Census. The 1991 Dutch Census was largely based on a register count of the population in combination with the Labour Force Survey 1991 and the Housing Demand Survey 1989/1990. Contrary to 1981 and 1991, Statistics Netherlands has published census information for 2001 on the municipal level.

III. METHOD OF COMPILING

13. The current virtual census relates to 2001. The backbone of the census is the central Population Register (PR), which is the combination of all municipal population registers. PR data of 1 January 2001 were used as the basis for the set of tables. The set of tables focuses on frequency counts and not on quantitative information. Different variables, such as occupation and level of education, were obtained from the Labour Force Survey (LFS). The variable job size was obtained from the large Survey on Employment and Earnings (SEE). To obtain sufficient records, information on persons from the LFS 2000 and the LFS 2001 was combined. For the housing tables, we used PR data of 1 January 2001, the Housing Register 2001 and the Survey on Housing Conditions (SHC) 2000.

14. Some variables of the PR and SSD datasets are available on an integral basis. Examples are age, sex, marital and employment status. Survey variables are only available for a part of the population. Examples are the highest level of education attained (LFS) and whether someone rents or owns the property they live in (SHC). We guaranteed consistency among the tables by using the technique of repeated weighting. The method of repeated weighting has been described extensively in [3] and [4]. It generates a new set of weights for each estimated table and is based on the repeated application of the regression estimator. The results of five simulation studies testing various aspects of repeated weighting can be found in [2]. When using repeated

weighting, the weights of the records in the microdata are adapted in such a way that a new table estimate is consistent with all earlier table estimates.

15. To apply the technique of repeated weighting, we used the latest version of the software package VRD developed by Statistics Netherlands. The letters VRD stand for Vullen (Filling) Reference Database and the aim of the application is to fill and manage the reference database. The main functions of VRD are the estimating of tables via repeated weighting, the addition of these tables to the reference database, and the withdrawal of aggregates from the reference database. Under the condition of small, independent samples, the variances of the table values can also be estimated. The estimating of the tables does not occur in VRD itself, but takes place in Bascula 4.0 automatically without the VRD-user seeing this explicitly. Estimating the tables and the variances can be done in the batch or interactively.

16. To be able to estimate every table as accurately as possible, each estimate is based on the largest possible number of records. Tables that contain register variables only are counted from the registers. Tables that contain at least one variable from a survey are estimated from the largest possible combination of registers and surveys.

17. The figures of the 2001 Census relate to persons living in the Netherlands on 1 January 2001 (counting unit persons). The persons who were living in the Netherlands at the beginning of that day according to the PR were ‘counted’ in the Virtual Census. Most of the Dutch population lives in private households, the remainder being part of institutional households. The number of employees in the tables relates to the end of the year 2000 for which 22 December 2000 was used as reference date to fix the number of jobs of employees in the Netherlands. It was impossible to have a reference day in 2001 for the number of employees since the SSD datasets 2001 were not available in time to use in the 2001 Census. The SSD data used registers’ information on the jobs of employees. If an employee holds several jobs at the same time, he or she can appear several times in the employee register. The features of the main job are used in the set of tables. The main job of an employee has been defined as the job with the highest gross wage for the social insurances.

18. The 2001 Census was compiled partly on the basis of sample data. Therefore, margins of inaccuracy have to be taken into account for some results of the 2001 Census. Because of the reliability of the results, rules of thumb are being applied for cell values that are based on a sample from the census population. The exact margins of inaccuracy cannot be given because of the complex design of the sample surveys used for the Census. The rules of thumb for records of observations from the LFS run as follows:

(a) Table cells based on less than 10 persons are always suppressed.

(b) Table cells based on 25 or more persons are always published.

(c) Table cells based on 10–24 persons are only published if they form a part of a breakdown (by age or sex), in which no cells based on less than 10 persons occur, and at least 50 percent of the cells in the breakdown have more than 25 persons. The threshold of 25 persons corresponds to an estimated relative inaccuracy of at most 20 percent (i.e. the estimated margins amount to 40 percent at most).

19. The rules of thumb for records from the SHC are of the same form. However, somewhat higher threshold values are applied because the sample size of the SHC is somewhat more limited than the one of the LFS. For table cells with households or dwellings as counting unit, analogous rules of thumb are applied for the Dutch Census of 2001.

IV. KEY RESULTS OF THE 2001 CENSUS IN THE NETHERLANDS

A. Population by sex, age and type of household

20. At the start of 2001 a total of 16.0 million people were living in the Netherlands, 7.9 million male and 8.1 million female. In the age categories 0-14 and 15-74 year there were more men than women, but in the category 75 and older there were almost twice as many women as there were men. Most people live in private households. More than 200 thousand people lived in institutional households, such as health care institutions and institutions for retired and elderly people. About 36 percent of this group was male and 64 percent female. Of the people in institutional households, 57 percent was over 75. This group was dominated by women. More information about the population by sex, type of household and age group can be found in Table 1.

Table 1. Population by sex, type of household and age group

Sex and type of household	***All ages***	***Age in years***		
		0-14	***15-74***	***75+***
Total population	**15,985,538**	**2,977,283**	**12,036,171**	**972,084**
Male	**7,909,052**	**1,522,811**	**6,047,425**	**338,816**
Female	**8,076,486**	**1,454,472**	**5,988,746**	**633,268**
		0-14	***15-74***	***75+***
Population in private households	**15,766,606**	**2,970,545**	**11,947,996**	**848,065**
Male	**7,829,914**	**1,518,611**	**5,998,189**	**313,114**
Female	**7,936,692**	**1,451,934**	**5,949,807**	**534,951**
		0-14	***15-74***	***75+***
Population in institutional households	**218,932**	**6,738**	**88,175**	**124,019**
Male	**79,138**	**4,200**	**49,236**	**25,702**
Female	**139,794**	**2,538**	**38,939**	**98,317**

B. Population by economic activity

21. At the start of 2001, just under half of the people living in the Netherlands belonged to the economically active population (labour force). The working labour force included 7.4 million people: 6.8 million were employees and 0.6 million self-employed. The unemployed labour force comprised almost 200 thousand people. In the organisational set-up of the census, employees, the self-employed and unemployed are mutually exclusive categories. Self-employed people who also work a number of hours a week for pay are counted as employees. Someone in the working labour force cannot be unemployed at the same time. The number of unemployed is estimated on the basis of sample information.

22. Of the economically active population 58 percent was male, while of the economically inactive population, 58 percent was female. The economically inactive include attendants at educational institutions, retired people and people engaged in family duties. The number of housewives is more than 18 times the number of househusbands. More information about the population by economic activity and sex can be found in Table 2.
Table 2.

Table 2. Population by economic activity and sex

Economic activity	*Total*	*Male*	*Female*
Economic active population	**7,586,914**	**4,388,239**	**3,198,675**
Working	**7,394,777**	**4,287,967**	**3,106,810**
Employed	**6,786,511**	**3,883,813**	**2,902,698**
Self-employed	**608,266**	**404,154**	**204,112**
Unemployed	**192,137**	**100,272**	**91,865**
Economic inactive population	**8,398,624**	**3,520,813**	**4,877,811**
15-74	**4,449,257**	**1,659,186**	**2,790,071**
Attendant at educational institutions	**640,446**	**342,934**	**297,512**
Retired	**1,355,940**	**620,493**	**735,447**
Engaged in family duties	**1,270,420**	**65,821**	**1,204,599**
Other economically inactive	**1,182,451**	**629,938**	**552,513**

V. THE 2001 CENSUS COMPARED TO EARLIER DUTCH CENSUSES

23. The first Census in the Netherlands was held in 1795 for the purpose of establishing voting constituencies. At that time, the united provinces of the Netherlands were still a republic and the borders were different from the current borders. After Napoleon, the Netherlands became a kingdom and once every ten years a census was held. The first Census in the Kingdom of the Netherlands was held in 1829. Before Statistics Netherlands was established, another six Censuses were held in 1839, 1849, 1859, 1869, 1879 and 1889 under the responsibility of the Ministry of the Interior. In 1899, Statistics Netherlands was established and was put directly in charge of the eighth Census. In the 20th century six more traditional Censuses were carried out in 1909, 1920, 1930, 1947, 1960 and 1971. The three most recent Censuses (1981, 1991 and 2001) were not based on a complete enumeration but on registers and surveys available to Statistics Netherlands.

24. Originally, the censuses had two aims. First, they were meant to correct errors in the municipal population registers. Second, they were used to obtain extra information about the socio-economic phenomena in the country. Since the Netherlands conducts a register-based census, the first aim no longer exists. Also, the quality of the central Population Register (PR), which unites all municipality population registers, has improved considerably over time. This is because the incentive for municipalities to keep their population registers up-to-date is the allocation of central government funds among municipalities, which is generally based on the population size according to the local registers. Another reason is that it is extremely difficult to live in Dutch society without being included in the PR. So both municipalities and citizens have enough incentives to keep the PR of good quality. Recent actions in Rotterdam to improve the quality of the municipal population register for some old quarters prove this statement. The

second aim is still valid and many census results are published in a historical or international context. Currently, census data are popular for comparisons among countries.

25. Table 3 presents some key results of the Dutch Censuses in the period 1829-2001. The ageing of the Dutch population is worth noting, especially in the post-war period.

Table 3. Population by age group in the period 1829-2001

Census		*All ages*	*Age group*		
Number	*Year*		*0-19*	*20-64*	*65+*
		× 1,000	*in % of the total population*		
1	*1829*	2,613.3	44	50	5
2	*1839*	2,860.6	45	50	5
3	*1849*	3,056.9	43	53	5
4	*1859*	3,309.1	42	53	5
5	*1869*	3,579.5	43	52	6
6	*1879*	4,012.7	44	50	5
7	*1889*	4,511.4	45	49	6
8	*1899*	5,104.1	44	50	6
9	*1909*	5,858.2	44	50	6
10	*1920*	6,865.3	42	52	6
11	*1930*	7,935.6	40	54	6
12	*1947*	9,625.5	38	55	7
13	*1960*	11,462.0	39	53	9
14	*1971*	13,060.1	36	54	10
15	*1981*	14,216.9	31	57	12
16	*1991*	15,070.0	25	62	13
17	*2001*	15,985.5	24	62	14

VI. THE DUTCH 2001 CENSUS COMPARED TO OTHER COUNTRIES

26. More than fifty countries in the United Nations Economic Commission for Europe region participated in the 2000 Census Round. Many countries chose a day in 2001 as their reference day, although they chose many different days. As it takes a long time before all countries finish the tables required by the international organisations, the Netherlands took the initiative to make some simple comparisons among nine European countries that were relatively quick in compiling the set of tables for Eurostat and that were willing to join the comparison analyses. The calculations in this paper are the author's own and are based on the set of standard tables produced from census data for Eurostat by nine different countries. It is expected that there are definitional differences among the countries which will affect comparisons. Also, the statistics produced by the author do not necessarily reflect the way countries usually choose to present their data. Furthermore, they may not be considered in some cases to be the definitive source of information on these topics. The nine countries are the Netherlands (NL), Norway (NO), Sweden (SE), Finland (FI), Estonia (EE), Switzerland (CH), Slovenia (SI), Greece (GR) and the United Kingdom (UK). The nine countries differ in size, but all except the United Kingdom have a fairly small number of inhabitants compared to France and Germany.

27. The nine countries are members of the European Union (EU) or the European Free Trade Association (EFTA). The Netherlands joined the European Community at the start in 1958, the United Kingdom joined in 1973 and Greece in 1981. The European Community became the European Union in 1995 when Sweden and Finland joined. Estonia and Slovenia joined the EU in 2004. Norway and Switzerland are EFTA members and work closely together with the EU countries. Norway is also a member of the European Economic Area (EEA). The EEA agreement came into force on 1 January 1994. EEA countries are the EU 15, Norway, Iceland and Liechtenstein. Switzerland did not join the EEA, but works together with the EU countries on a bilateral basis. Statistics is one of the issues on which the EEA countries work together. The aim of the statistical co-operation in the EEA is to build a European Statistical System that gives a coherent and comparable description of the economic, social and environmental developments in the EEA countries.

28. The nine countries that are compared have different census reference dates: 31 March 2000 (Estonia), 5 December 2000 (Switzerland), 1 January 2001 (The Netherlands, Sweden and Finland), 18 March 2001 (Greece), 29 April 2001 (United Kingdom), 3 November 2001 (Norway) and 31 March 2002 (Slovenia).

29. Table 4 presents the estimated costs of the Censuses in the 2000 Round, and the population and area of the nine countries. Estonia, Slovenia, Greece and the United Kingdom held traditional censuses; Switzerland used a combination of a traditional census and register information to produce the census tables. Norway relied largely on registers, but conducted a census for some missing housing variables. Sweden and Finland held entirely register-based censuses and the Netherlands performed a virtual census based on existing registers and surveys. The Census costs for Norway, Estonia, Switzerland, Slovenia, Greece and the United Kingdom include enumeration costs. In the Netherlands, Sweden and Finland such enumeration costs do not exist for their 2001 Censuses, so the costs presented in Table 4 for these three countries are rough indicators of the extra costs of producing census tables for the international organisations and of analysing and publishing the results. Table 4 shows that the costs per inhabitant in those countries that required completion of a census form for the census were much higher than the countries that did not have enumeration costs. In Table 4, the population densities among the nine countries can be compared. The Netherlands has the highest population density, followed by the United Kingdom and Switzerland. The population density in the Nordic countries (Norway, Sweden and Finland) and in Estonia is relatively low. Slovenia and Greece occupy a middle position.

Table 4. Comparison of nine countries according to the Census results in the 2000 Round

	NL	NO	SE	FI	EE	CH	SI	GR	UK
Cost of the Census (in millions of Euros)	**3.0**	**14.6**	**1.0**	**0.8**	**10.2**	**99.1**	**8.0**	**49.7**	**367.4**
Population (× 1,000,000)	**16.0**	**4.5**	**8.9**	**5.2**	**1.4**	**7.3**	**2.0**	**10.9**	**58.8**
Area (× 1,000 km^2)	**41.5**	**323.9**	**450.0**	**338.1**	**45.1**	**41.3**	**20.3**	**132.0**	**244.1**
Cost of the Census per inhabitant (in Euros)	**0.2**	**3.2**	**0.1**	**0.2**	**7.3**	**13.6**	**4.0**	**4.6**	**6.2**
Population density (persons per km^2)	**386**	**14**	**20**	**15**	**31**	**177**	**99**	**83**	**241**

VII. CONCLUSIONS

30. The virtual census has proved to be a successful concept in the Netherlands. It has many advantages compared to traditional censuses. The costs are now considerably lower. Nevertheless, data on the Netherlands have become available that could be compared to results of earlier Dutch censuses and to the results of other countries that took part in the 2000 Census Round. It was the third time that the Netherlands conducted a virtual census. However, the Dutch data that have been compiled for 1981 and 1991 were of a much more limited character than the set of tables of the 2001 Census. Moreover, they were largely based on a register count of the population in combination with the then existing surveys on the labour force and housing conditions.

31. The technique of repeated weighting has been used successfully to produce a consistent set of tables for the 2001 Census. Before compiling tables with this new technique, micro-integration of the different sources in the SSD remains important. In the micro-integration process, the data are checked and incorrect data are adapted. It is strongly believed that micro-integrated data will provide more reliable results, because they are based on a maximum amount of information. Also the coverage of subpopulations will be better, because when data are missing in one source, another source can be used. Another advantage of micro-integration and repeated weighting is that there is no reason for confusion among users of statistical information: there will be one figure on each socio-economic phenomenon, instead of several figures depending on which sources have been used.

32. It is possible to use the technique of repeated weighting in other countries as well. However, first it should be possible to use registers for statistical purposes. In most countries, not all census variables can be derived from register information. Additional surveying then remains a necessity, but a consistent set of census tables can be produced using the technique of repeated weighting.

33. At the end of 2003 the complete set of forty census tables for the Netherlands was sent to Eurostat. A book about the Dutch Virtual Census of 2001 was written afterwards [7]. This book provides a wide-ranging description of the socio-demographic and socio-economic state of the Netherlands based on the 2001 Census results. It discusses differences in size and composition among households, economic activity of households, individual activity status by region, age,

education level and branch of economic activity. There are separate chapters on the economic activities of young people and people of retirement age. The economic activities, levels of education and occupation of foreigners from various countries of origin are compared with each other and with the native Dutch population. Regional aspects are also examined, including commuting. The results of the 2001 Census are compared with the Census results of some other European countries and with earlier Dutch Censuses. Lastly, the Virtual Census methodology used is described in some detail.

34. The PDF version of the book can be found at the Statistics Netherlands website, at page http://www.cbs.nl/NR/rdonlyres/D1716A60-0D13-4281-BED6-3607514888AD/0/b572001.pdf. An extra Chapter (number 15) is available at page http://www.cbs.nl/NR/rdonlyres/7A45A707-D4F6-4F23-92E5-130C5BC1A144/0/b572001hoofdstuk15.pdf with an overview of the used data sources, methods and definitions. Hard copies of the book were sent to all authors of the book, to the management of Statistics Netherlands and to several libraries. The book was also offered to the Prime Minister, the Minister of Economic Affairs and the Minister of Education, Cultural Affairs and Science of the Netherlands and to Director-Generals of statistical offices in several countries. In August 2004, the book was publicly released at an official presentation in the Statistics Netherlands' office in Voorburg. The research process and the main findings were then presented to an audience of academics, press representatives, government officials, as well as Statistics Netherlands' employees. Several articles were written in national and regional newspapers about the Dutch Virtual Census of 2001 and its results. Announcements and interviews appeared in several mailing lists, newsletters and journals.

35. The set of forty standard tables for the Netherlands (in Excel format) can be found at page http://www.cbs.nl/nl-NL/menu/themas/dossiers/volkstellingen/cijfers/incidenteel/maatwerk/2003-volkstelling-excel.htm and the table annotations at page http://www.cbs.nl/NR/rdonlyres/D8D55875-0630-492F-8125-BA71D7608009/0/tableannotationsNLcensus2001.pdf. Protected 1 percent samples of the microdata of the Dutch Censuses of 1960, 1971 and 2001 were in 2005 disseminated via the IPUMS (Integrated Public Use Microdata Series) project, see http://www.ipums.org/international. These micro datasets contain a number of demographic and economic variables and can also be analysed via the institute DANS (Data Archiving and Networked Services), see http://www.dans.knaw.nl/en/. Bona fide researchers who want to make more detailed studies on these three censuses can work on-site at the premises of Statistics Netherlands. More information about this last option can be obtained via Statistics Netherlands' Centre for Policy Research (http://www.cbs.nl/nl-NL/menu/informatie/beleid/centrum-voor-beleidsstatistiek/cvb/default.htm).

VIII. GENERAL RECOMMENDATIONS FOR FUTURE CENSUS ROUNDS

36. In this section, based on the experiences of the Dutch Virtual Census of 2001, some recommendations are made for future Census Rounds. It is very useful to have Census guidelines and a table program, but some errors in this program have to be corrected and some aspects have to be described in some more detail for the next Census Round. The guidelines have originally been written with the idea in mind that all countries conduct a traditional census, but more and more countries have chosen other options. In revising the guidelines for the 2010 Census Round, this aspect has been taken into account to a certain extent. It is a challenge to change the table program according to the remarks made below.

37. The number of different classifications for one variable (e.g. age) is sometimes too large in the table program. This implies that sometimes cells in tables cannot be estimated consistently, whereas less estimation problems would be occurred if the number of different classifications would have been reduced. This problem is much more severe if non-nested classifications of the same variable are used.

38. What to do with people who have different economic activities at the same time? In a traditional census one can ask the respondent about his or her main economic activity. In register-based censuses one has to find a criterion to choose one economic activity per person. If different countries use different criteria, the comparability of the results will be hampered. In the Netherlands we could not distinguish between employers and other employed people and therefore these two self-employed categories have been merged. People who were at the same time employed and self-employed were counted as employees. As we have no register of unemployed people, this group is the only part of the economically active population that has to be estimated. Therefore, sometimes the number of economically active people is an estimate, although we know exactly how many people are employee or self-employed. People who are at the same time both economically active and student are counted as economically active people. The population of retired people is not clearly defined. What to do with people who are partly retired? And what with people who live from their property instead of from their pension? It is necessary to make international priority rules on the variable economic activity to improve the comparability of the results in different countries.

39. More attention has to be paid to the variable country of birth and less to the variable citizenship in the next Census Round. Different countries have namely different policies towards changing nationality. International more relevant comparisons can be made by distinguishing first and second generation foreigners on the basis of the variables country of birth of the person and his or her parents. This is done in the PAU (Population Activities Unit of the UN) tables and Chapter six of [7]. In the family nuclei it is useful to add extra categories for same sex (married or cohabiting) couples. This is a growing group that has to be taken into account in the next Census Rounds. For the variable family status the rest category ‘child with other family status’ is missing.

40. The NUTS classifications change over time. Therefore, it is crucial to provide the countries with the table lay-outs including the classifications of the Census year. The list of country codes has to be improved and used consistently in all tables. Now only a general unknown category is included. This implies that for countries not in the list all the records are counted as unknowns, which implies that the totals per continent become incorrect. What to do with no longer existing countries not in the list? This is especially relevant for the variable country of birth if the original country was split.

41. For the NACE codes it is useful to add a separate category unknown. Now we have attributed the unknowns proportionally to the other categories. Also for the variable occupation a category unknown should be introduced. For 2001 we have included the unknowns in the total, but that implies that the sum of the occupation categories is often smaller than the total. Finally, for the ISCED a category unknown is included in the table, but here we merged the unknowns with the categories no education at all and level of education pre-primary as we could not

distinguish among these three categories.

42. When we take the lessons learnt on the importance of nested (age) classifications, the priority rules for economic activities, the variable country of birth, and on the NUTS and other classifications into account for the table program, the comparability of the results of the different countries will be improved enormously.

IX. REFERENCES

[1] Corbey, P., 1994. Exit the population Census. Netherlands Official Statistics, Volume 9, summer 1994, pp. 41-44.
[2] Duin, C. van and V. Snijders, 2003. Simulation studies of repeated weighting. Discussion paper 03008, Statistics Netherlands, Voorburg / Heerlen. http://www.cbs.nl/NR/rdonlyres/203C85C6-7075-47A0-97BA-A3B748D393FE/0//Discussionpaper03008.pdf.
[3] Houbiers, M., 2004. Towards a social statistical database and unified estimates at Statistics Netherlands. Journal of Official Statistics, Volume 20, No. 1, pp. 55-75.
[4] Houbiers, M., P. Knottnerus, A.H. Kroese, R.H. Renssen and V. Snijders, 2003. Estimating consistent table sets: position paper on repeated weighting. Discussion paper 03005, Statistics Netherlands, Voorburg / Heerlen. http://www.cbs.nl/NR/rdonlyres/6C31D31C-831F-41E5-8A94-7F321297ADB8/0/discussionpaper03005.pdf.
[5] Kroese, A.H. and R. H. Renssen, 2000. New applications of old weighting techniques, constructing a consistent set of estimates based on data from different sources. ICES II, Proceedings of the second international conference on establishment surveys, survey methods for businesses, farms, and institutions, invited papers, June 17-21, 2000, Buffalo, New York, American Statistical Association, Alexandria, Virginia, United States, pp. 831-840.
[6] Rao, J.N.K., 2003. Small area estimation. Wiley, New York, United States.
[7] Schulte Nordholt, E., M. Hartgers and R. Gircour (Eds.), 2004. The Dutch Virtual Census of 2001, Analysis and Methodology, Statistics Netherlands, Voorburg / Heerlen, July, 2004. http://www.cbs.nl/NR/rdonlyres/D1716A60-0D13-4281-BED6-3607514888AD/0/b572001.pdf.
[8] Schulte Nordholt, E., 2005. The Dutch virtual Census 2001: A new approach by combining different sources. Statistical Journal of the United Nations Economic Commission for Europe, Volume 22, Number 1, 2005, pp. 25-37.
[9] Statistics Netherlands, 2003. Urban Audit II, the implementation in the Netherlands. Report, BPA no. 2192-03-SAV/II, Statistics Netherlands, Voorburg. http://www.cbs.nl/NR/rdonlyres/8C6E4C9D-4338-4E32-848B-8D43B9B3242D/0/urbanauditIINetherlands.pdf

* * * * *