

Kordos, Jan

Article

THE CHALLENGES OF THE POPULATION CENSUS ROUND OF 2020. OUTLINE OF THE METHODS OF QUALITY ASSESSMENT OF POPULATION CENSUS DATA

Statistics in Transition New Series

Provided in Cooperation with:

Polish Statistical Association

Suggested Citation: Kordos, Jan (2017) : THE CHALLENGES OF THE POPULATION CENSUS ROUND OF 2020. OUTLINE OF THE METHODS OF QUALITY ASSESSMENT OF POPULATION CENSUS DATA, Statistics in Transition New Series, ISSN 2450-0291, Exeley, New York, NY, Vol. 18, Iss. 1, pp. 115-138,
<https://doi.org/10.21307/stattrans-2016-060>

This Version is available at:

<https://hdl.handle.net/10419/207848>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

STATISTICS IN TRANSITION new series, March 2017
Vol. 18, No. 1, pp. 115–138

THE CHALLENGES OF THE POPULATION CENSUS ROUND OF 2020. OUTLINE OF THE METHODS OF QUALITY ASSESSMENT OF POPULATION CENSUS DATA¹

Jan Kordos²

ABSTRACT

The author begins with the challenges of the population census round of 2020 and gives an outline of the methods of quality assessment of population census data. Next, the synthesis of two Eurostat documents (2007, 2009) and UN Statistics Division (2010) monograph relating to census data quality assessment are considered. Special attention is paid to three methods of census quality assessment: (i) demographic method, (ii) post-enumeration survey, and (iii) comparison with existing household surveys. Census quality assessment methods are discussed in the context of the international recommendations and experience of selected countries. Some Polish experience in these fields is also discussed, and some suggestions to the 2021 Polish Census of Population and Housing preparations are given.

Key words: census of population, demographic analysis, quality assessment, post enumeration survey, household survey.

1. Introduction

Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing is published³. The main objectives of this Recommendations are:

- (a) to provide guidance and assistance to countries in the planning and execution of their population and housing censuses; and
- (b) to facilitate and improve the comparability of a census at the UN regional level through the identification of a core set of census topics and the harmonization of concepts, definitions and classifications.

¹ This is an updated and extended version of the author's paper presented at the Plenary Session of the Committee on Demographic Studies, Polish Academy of Sciences, on 15 November 2010.

² Warsaw Management University.

³ <http://unstats.un.org/unsd/demographic/sources/census/census3.htm>.

The Recommendations are also expected to be used as the general framework for the European Union programme for the 2021 round of population and housing censuses, from which data are to be provided to Eurostat by Member States under the provisions of Regulation (EC) 763/2008.

The author would like to focus his attention on the quality of census of population and housing presented in different documents. His first selection are two Eurostat publications: *Handbook on Data Quality Assessment Methods and Tools* (Eurostat, 2007) and *Handbook on Quality Reports* (Eurostat, 2009), which relate to the quality assessment in statistics, and particularly to the quality assessment of censuses of population. However, his special concern is the publication of the UN Statistics Division on *Post-enumeration Surveys* (2010), which is extremely important for population census results evaluation.

These Eurostat publications underline that the production of high quality statistics depends on the assessment of data quality (Eurostat, 2007, p.5): *“Without a systematic assessment of data quality, the statistical office will risk to lose control of the various statistical processes such as data collection, editing or weighting. Doing without data quality assessment would result in assuming that the processes can not be further improved and that problems will always be detected without systematic analysis. At the same time, data quality assessment is a precondition for informing the users about the possible uses of the data, or which results could be published with or without a warning. Certainly, without good approaches for data quality assessment statistical institutes are working in the blind and can make no justified claim of being professional and of delivering quality in the first place. Assessing data quality is therefore one of the core aspects of a statistical institute’s work”*.

The quality of population and housing census data is very important for many reasons, building public trust and understanding in the national statistical system. The purpose of census evaluation is to provide users with a level of confidence when utilizing the data, and to explain errors in the census results. It is therefore important to choose an appropriate way of sending out these messages to the right group of people.

2. Errors in population censuses and methods of their assessment

It is universally accepted that a population census is not perfect and that errors can and do occur at all stages of the census operation. Errors in the census results are classified into two general categories:

- coverage errors, and
- content errors.

Coverage errors are the errors that arise due to omissions or duplications of persons or housing units in the census enumeration.

Content errors are errors that arise in the incorrect reporting or recording of the characteristics of persons, households and housing units enumerated in the census.

Many countries have recognized the need to evaluate the overall quality of their census results and have employed various methods for evaluating census coverage as well as certain types of content error. Numerous methods are available to estimate the coverage and content error of censuses. These include:

- simple techniques of quality assurance such as internal consistency checks;
- comparisons of results with other data sources including previous censuses, current household surveys and/or administrative records are also useful techniques⁴.

However, for evaluating census data there are two very important methods:

- *demographic analysis* and
- *post enumeration surveys*.

Nevertheless, countries have frequently encountered problems in implementing evaluation methods, and especially in designing and implementation post enumeration surveys (UN Statistics Division, 2010). For this reasons, the author decided, after presenting basic methods of census quality evaluation and some experience of selected countries in this field, to focus his attention on *post enumeration survey designing and implementation*.

2.1. Demographic analysis of census results

By undertaking a demographic analysis, results from a census may be compared with data from other demographic systems such as vital registration of births and deaths including net migration if such data are available. The cohort component method of demographic analysis uses data from successive censuses as well as life-table survival rates, age-specific fertility rates and estimated levels of international migration between censuses. The population is projected forward to the reference date of the second census based on estimated levels and age schedules of fertility, mortality and migration, and the expected population is compared with enumerated population in the second census.

Another method of analysis involves comparing age distributions of successive censuses. This method is widely used because it requires little data. Yet another method in use is the cohort survival regression method, which uses population counts by age from two censuses and deaths by age during the inter-census period to estimate the coverage rate.

⁴ See: UN Economic Commission for Europe (2011, 2015).

Record checks

In this type of analysis, census records are matched with a sample of records from the vital registration or other identification systems, where the relevant respondents to the census questionnaire are traced to the time synchronized with the census. Such sources include previous censuses, birth registrations, school enrolment registries, voter registration list, health and social security records, immigration registers, national or citizen registration cards, etc. Both coverage and content errors could be measured through such comparisons. For coverage evaluation purposes, the following pre-conditions are necessary:

- (i) a large proportion of the census target population should be covered in the record system;
- (ii) the census and record system should be independent from each other; and
- (iii) there should be sufficient information in the records so that accurate matching is possible.

For content evaluation purposes, the record system should contain some relevant items covered in the census such as age, sex, education, income, etc. It is important to ensure that the definitions of items are the same.

2.2. Post Enumeration Surveys

According to the Conference of European Statisticians (2007): *Principles and Recommendations for Population and Housing Censuses*, the post enumeration survey (PES) is a complete enumeration of a representative sample of a census population followed by matching each individual enumerated in the PES with information from the census enumeration. The results of the comparison are mainly used to measure coverage and content error in the context of the census. Some countries only confine the PES to evaluating coverage error (ONS, 2005). Coverage error refers to housing units and people missed in the census or those erroneously included. On the other hand, content errors evaluate the response quality of selected items. In general, an evaluation of the magnitude and direction of errors in a census is necessary in order to present to users the extent of reliability and accuracy of some characteristics reported. The evaluation, therefore, allows for a better interpretation of census results by presenting limitations to users by quantitatively evaluating the accuracy of census results with respect to coverage or/and quality of responses to questions on selected variables. For some countries the results of the PES are used to adjust census results if, for instance, there is evidence of major coverage errors (Hogan, 1992, 2001; Whitford & Banda, 2001).

2.3. Comparison with existing household surveys

Theoretically, any probability sample of households or individuals can be used to evaluate coverage and content errors in a census if they have some identical items with the same concepts and definitions. However, the post enumeration survey discussed above is specifically designed and most ideal to do so. In the absence of a post enumeration survey, other households survey results can be used to evaluate census results provided the principles of independence from the census and closeness to the census date are upheld. In addition, there should be sufficient identical information to perform accurate matching. For content evaluation, it is essential that several of the same data items are collected.

In Poland, comparisons with existing household sample surveys, such as the Household Budget Survey, the Labour Force Survey, and the Statistics of Income and Living Conditions survey (EU-SILC), to evaluate coverage and content errors in a census have not been used. Some indirect attempts have been tried for the census of agriculture (Bartosińska, 2006). The author thinks that the household sample surveys conducted in Poland may fulfil most of the above prerequisites, and be used to evaluate coverage and content errors. However, some research in this field is required.

2.4. Interpenetration studies on current census

This method involves drawing subsamples, selected in an identical manner, from the census frame, with each subsample capable of providing a valid estimate of the population parameter (UN Statistics Division, 2010). The assignment of census personnel (enumerators, coders, data entry staff, etc.) is also done randomly. This method helps to provide an appraisal of the quality of the census information, as the interpenetrating subsamples can, for example, be used to secure information on content error. In censuses and surveys, nonsampling errors, for instance, arise from differential interviewer bias, different methods of eliciting information, etc. After the subsamples have been enumerated by different groups of interviewers and processed by different teams of workers at the tabulation stage, comparison of estimates based on the subsamples provides a broad check on the quality of the census results. The results, from such studies, could be useful in improving the operations of future censuses and large-scale sample surveys. Until now, the interpenetrating studies have not been used in Poland (Kordos, 1987, 1988, 2007).

3. Some experience of other countries in census quality assessments

The author would like to refer to some experience in the census quality assessments of the U.S. Census Bureau, the Statistics Canada, the UK Office for National Statistics (ONS), and the Nordic Countries experience in register-based statistics and census of population, and comment of some Polish experience in this field.

3.1. Experience of the U.S. Census Bureau in census quality assessment⁵

The U.S. Census Bureau has used post-enumeration surveys with dual system estimation to measure coverage in the Decennial Censuses of Population and Housing since 1980. This approach involves case-by-case matching of persons in an independent survey with persons in the census to determine who was missed or counted in error. The post-enumeration survey-based coverage measurement program associated with the 1980 Census was called the Post-Enumeration Program (PEP); in the 1990 Census it was called the Post-Enumeration Survey (PES); in the Census 2000 it was called the Accuracy and Coverage Evaluation (A.C.E.); and for the 2010 Census it is called Census Coverage Measurement (CCM).

*Coverage Measurement in the 2010 Census*⁶. The primary goal of the 2010 CCM program was to measure coverage error in the 2010 Census such that this information could be used to improve the coverage of future censuses. As a result, the scope of coverage measurement was broader and the emphasis was different than it had been in the past. Specifically, the 2010 CCM goals were to:

- 1) produce measures of coverage error, including its components of omissions and erroneous enumerations;
- 2) produce these measures of coverage error not only for demographic groups and geographic areas, but also for key census operations; and
- 3) continue to provide measures of net coverage error.

National Research Council (2010), *Envisioning the 2020 Census*⁷: Planning for the 2020 census is already beginning. This book from the National Research Council examines several aspects of census planning, including questionnaire design, address updating, non-response follow-up, coverage follow-up, de-duplication of housing units and residents, editing and imputation procedures, and several other census operations. The book recommends that the Census Bureau overhaul its approach to research and development. The report urges the Bureau to set cost and quality goals for the 2020 and future censuses, improving efficiency by taking advantage of new technologies.

3.2. Experience of the Statistics Canada in Census Quality Assessment⁸

Conducted every five years, the Canadian Census of Population is a major undertaking whose planning and implementation spans a period of over eight years. Statistics Canada generally works on two and even three censuses at any given point in time. Before the final results of one census are out, planning and systems development are already well under way for the next one.

⁵ https://www.census.gov/coverage_measurement/post-enumeration_surveys/.

⁶ <http://www.nap.edu/catalog/12524/coverage-measurement-in-the-2010-census>.

⁷ <http://www.nap.edu/catalog/12865/envisioning-the-2020-census>.

⁸ <http://unstats.un.org/unsd/dnss/docs-nqaf/Canada-12-539-x2009001-eng.pdf>.

Consultations on the 2011 Census content were held in 2007 with a broad range of users, including key federal government departments, provinces and territories, local authorities, libraries, academia, the private sector, special interest groups and the general public. The Internet has become the primary vehicle for consultation material. Written submissions, dedicated meetings, conferences and working groups also continue to be highly effective ways to engage with users.

Statistics Canada with the 2011 Census continues to address both internal and external pressures to change its collection, processing and dissemination strategies. Using a number of pilot tests and phasing changes over the 2001, 2006 and 2011 Census cycles, the Agency moves from a decentralized, manually intensive collection and data entry operation to a more centralized and automated approach. This in particular addresses key concerns regarding confidentiality and security of personal census data. At the same time, a more proactive census communications and dissemination strategy has been adopted which has led to substantial increases to the amount of media coverage, to interest by a large number of censuses.

In view of these realities, Statistics Canada has a long tradition of providing guidance in its survey designs by consolidating its experience and conclusions about what constitute “best practices” into a set of Quality Guidelines. The first edition of Quality Guidelines appeared in 1985. Revised editions were released in 1987, 1998 and 2003. In keeping with the need to keep the guidelines evergreen, the present document, i.e. Statistics Canada (2009), has been significantly updated from the previous edition to reflect further advances in survey methodology over the past six years.

Data quality assessment provides an evaluation of the overall quality of census data. The results are used to inform users of the reliability of the data, to make improvements for the next census and, in the case of two coverage studies, to adjust the official population estimates. Quality assessment activities take place throughout the census process, beginning prior to data collection and ending after dissemination.

3.3. British One Number Census project⁹

The author would like to present a synthesis of the British One Number Census project as an example of correctly designed and implemented undertaking. Census 2001 results were the first to represent the entire population. This was achieved through a new strategy known as the ‘*One Number Census*’ (ONC). One of the key elements of the ONC was an independent follow-up survey. The Census Coverage Survey (CCS), as it is known, involved face to face interviews with a sample of 320,000 households from every local authority in the UK. In the past, the total population given by the census was the raw count, reflecting a response rate of 98 per cent. But by combining the results of the census and the

⁹ (ONS, 2005); One Number Census: Evaluation Report, Census 2001 Review and Evaluation, More available at: <http://www.ons.gov.uk/ons/guide-method/census/census-2001/index.html>.

CCS, it was possible in 2001 to estimate the total resident population - the 'one number' - to a high level of precision, plus or minus 0.2 per cent.

The Census Coverage Survey (CCS) was specifically designed to enable census population counts to be adjusted for under-enumeration at the national, local and small area level. It consisted of a completely independent and intensive face-to-face survey of a sample of over 16,000 postcodes containing 320,000 households drawn from all local authorities in England and Wales. The sample design took into account the uneven distribution of under-enumeration across the country by stratifying by a '*Hard to Count*' index based upon characteristics likely to be associated with under-enumeration, such as the number of multi-occupied addresses.

The CCS was operationally independent from the census enumeration exercise. The CCS sample postcodes were kept confidential, CCS interviewers did not have any sight of the address lists produced in carrying out the census, nor the census forms returned in the area in which they were interviewing. The interviewers focused on making as many calls as necessary to achieve an interview, and the timing of these calls was varied to maximise the probability of making contact.

The CCS in England and Wales achieved a response from 91 per cent of the households identified by interviewers. This is a high response rate for such a large-scale voluntary survey when compared to other the U.K. national surveys. The survey succeeded in meeting its objective of identifying households and persons that had been missed by the 2001 census.

For the ONC process to produce unbiased estimates of the population it is necessary for the census and Census Coverage Survey to be as independent of each other as possible. Practical arrangements were put in place to achieve this with census and CCS operations being kept entirely separate on the ground. If the two attempts at enumerating the same population are independent, it is possible to not only estimate those missed by either the census or CCS but to also estimate those missed by both - the dual system approach.

Through this approach, independence of the process was achieved. However, there is an additional component of dependence which needs to be taken into account. This is dependence caused by the fact that those people who are difficult to count in a census are also difficult to count in a post-enumeration survey such as the CCS. This was expected and a methodology was developed to identify those areas where dependency was marked and to adjust for that dependence. This added an additional 230,000 to the ONC population estimates for England and Wales as a whole.

3.4. The Nordic Countries Register-based statistics and Census of Population

The Nordic countries have a long tradition in using administrative registers in the production of official statistics. One of their common experiences is that the use of administrative records in censuses is the last step in a process that begins

with producing statistics on different subject areas, depending on the type of registers available. By producing statistics on population or employment based on administrative data, they have learned about the influence the actual registers have on the quality of the statistics. After a period of testing and improvement, they realized that the quality of the administrative data was compatible with the quality recommended for censuses, and decided to also use the registers for census purposes. The time it takes from the establishment of an administrative register to possessing the quality data needed for census purposes may differ from one subject area to another. Nevertheless, the process is similar from subject area to subject area and from country to country. This also seems to be the case for establishing a statistical system based on register information.

In recent years, an increasing number of countries in the UNECE region have been considering the possibility of producing statistics based on administrative registers. Therefore, the National Statistical Institutes (NSIs) of the Nordic countries (Denmark, Finland, Iceland, Norway and Sweden) decided to share their experience and knowledge with the international statistical community, by producing comprehensive documentation of their best practices. The UN Economic Commission for Europe¹⁰ published this document in 2007: The objective of this volume is to give strategic and planning officers in the NSIs and understanding of what register-based statistics are, covering also the necessary technical and administrative capacity, and the possible applications of these methods to produce official statistics. The emphasis of the publication is on the use of administrative registers to produce demographic and social statistics. In publishing the present volume, the United Nations Economic Commission for Europe (UNECE) would like to support the Nordic countries in sharing their experience in this field with the international statistical community at large. The volume represents a valuable tool for all NSIs (both within the UNECE region and outside it) that are planning to produce official statistics based on administrative registers. It also supported the implementation of the 2010 round of population and housing censuses, and preparation of the 2020 round of population and housing censuses.

It is necessary to stress that developing a register-based data system has been a step-by-step process in all Nordic countries over a rather long period. Statistical registers have been established in several areas and by 2011 all Nordic countries, at least according to national plans, will have a totally register-based population and housing census system. And even if we call it a census system, the same data sources are used in the corresponding subject matter statistics.

¹⁰ Register-based statistics in the Nordic countries - Review of best practices with focus on population and social statistic: available at:
http://www.unece.org/fileadmin/DAM/stats/publications/Register_based_statistics_in_Nordic_countries.pdf.

A register-based statistical system should never be regarded as completed once and for all. As new user needs arise and new administrative registers are established, new information should be integrated in the system.

What is a register-based census?

At one end of the line we have traditional censuses collecting data by the use of enumerators and questionnaires, using no register information at all. At the other end we have the totally register-based census. Some countries use mixed mode data collection with a combination of data from registers and questionnaires (as a total count or a sample survey). However, even countries conducting mainly traditional censuses may use register information to some extent, for instance an address list for persons or households for mailing out census forms.

For some countries, such as in the Nordic region, the goal has been to develop fully register-based censuses. For other countries, the goal may be to use registers to a certain extent, even if it is not possible or desirable to conduct a fully register-based census. And for some countries, a register-based census may not be an option at all.

What is required in order to call a census "register-based"? The main requirement must be the existence of a population register and a dwelling register. After all, resident persons and housing units are the basic units in a population and housing census. The link between persons and their dwellings is equally important, giving the household unit. These requirements constitute the cornerstones for developing a register-based census.

Register-based population and housing censuses system

The population and housing census provides the best example of the use of administrative records in statistical production. The Finnish system is used here as an example. The base registers cover the statistical units relevant to a census: *persons* resident in the country, the *buildings and dwellings* in the country as well as all *enterprises and their establishments* (business register). All statistical units can be linked to one another by means of the identification systems: *persons* can be linked to *families* and *households*, to the *dwelling* and *building* in which they live, and to the *employer* for whom they are working. Similarly, all units can be located on maps using geographical coordinates, because all buildings have been provided by coordinates.

3.5. Some Polish experience in census quality assessment

In Poland, demographic analyses of censuses evaluation have been widely used. There are a number of publications devoted to evaluation of census quality results using demographic methods (Gołata, 2009, 2010, 2012, 2014, 2016; Nowak, 2002, 2008; Paradysz, 2010, Strzelecki et al., 2002). The author will refer to some of them later, and begin with post enumeration survey (PES).

In Poland, first PES on sampling basis was applied in 1978 for the 1978 Population Census (Zasepa, 1993), in 1988 for the 1988 Population Census

(Nowak, 1998), the Micro-census 1995 and the 2002 General Population and Housing Census (GUS, 1996; Szablowski et al., 1996).

In the 2002 Population and Housing Census the PES was conducted three weeks after the main census. A primary sampling unit was Census Enumeration Area (CEA). Out of all 177,591 CEAs for PES 903 CEAs were selected using stratified sampling designs by region with proportional allocation. Altogether 60,029 dwellings were selected and 27 census items were checked.

In the 2011 Population and Housing Census Sampling frame for the PES were sampling units selected for the sampling census, i.e. 2,744 thousand. For PES covered dwellings which had done self-enumeration, dwellings done by census enumerators or CATI, and not covered from different reasons. The PES enumeration was carried out from 1 April to 30th June. Additional condition was added that a dwelling should be contacted by telephone (50, 5%) . . . As by now no PES results have been officially published.

The author would like to refer here shortly to three papers: Gołata (2012, 2016), and Paradysz (2010), where some interesting aspects of census quality are discussed.

Gołata (2012) discusses the quality of population census data; in particular, the quality assessment of census data refers to the Polish experience. The author aims to propose an integrated approach to censuses, based on administrative records and sample surveys. She has stressed that despite a long tradition, as well as a well-developed research methodology, censuses do not provide 'perfect' results. First of all, the census as a comprehensive investigation may be burdened with different kinds of non-random errors. Taking into account all the different methods for conducting censuses, including the traditional method, using data from administrative records with sample surveys and the mixed method, different sources of error are analyzed.

The paper also attempts to define the census population in the light of international standards concerning usual residents, permanent residents and actual residents, with the resulting consequences. Sources of errors, methods of their identification and elimination are discussed. Particular attention is paid to coverage errors, which are illustrated by examples from the Polish censuses. The results of previous studies assessing the quality of administrative records, which were conducted in preparation for 2011 Population Census, are taken into account.

Evaluation of population censuses is traditionally carried out based on the results of post-census surveys. Polish experience in this field is rather limited. The discussion also includes harmonization problems, particularly the danger of divergent results and estimates, in terms of several data sources. In this regard, the British One Number Census project is mentioned (ONS, 2005). The idea of the project was to reduce the population underestimate and provide outputs which would be consistent in cross-section of territorial units at the regional and national level.

Golata (2016): This is the best updated article on quality assessment of the last two population censuses in Poland conducted in 2002 and 2011. The article refers to the shift in methods to conduct a population census: from a conventional enumeration through a sample survey and a mixed approach to administrative data, as a new standard in statistics. The paper compares two Polish censuses of 2002 and 2011. It is aimed at quality assessment in the case of both: the traditional method (2002 census) and the combined approach (2011 census).

The quality of census data is discussed with essential aims and objectives to provide reliable information on the population age and sex structure in detailed territorial division. Therefore, quality assessment is provided for the whole country and at regional level. First of all, coverage errors are considered. She uses multiple sources of data and non-matching methods, in particular: demographic analysis based on previous censuses, vital statistics and a comparison with other existing sources. Different cross-sections according to sex, age and place of residence are considered. In each separate domain adequacy and divergence assessments are accompanied by an attempt to provide substantive explanations.

Paradysz (2010) has stressed that Poland prepares the new population census 2011 on the basis of a Virtual Census, using the administrative registers and combining them with surveys. This approach to modern census needs, however, modern indirect estimators. The paper also deals with nonresponse problems and other biases due to new approach. In the Polish 2011 census calibration and imputation methods are used. It also means indirect estimation on a different level of statistical aggregation of population. In conclusions he stressed that those systematic errors in population censuses are much higher than previously assumed. It concerns not only mobile persons (20 - 29 age group), but also much older persons. He concludes that post enumeration surveys had no important meaning since none of them did not show the most important, evident systematic errors, but no references are mentioned. He suggests that for next census administrative registers, and particularly a population register, should be used as well as indirect estimation. The author agrees with Paradysz's conclusions, and would like to add that previously conducted post enumeration surveys in Poland were not correctly prepared and implemented if we take into account recommendations for the post enumeration surveys (UN Statistics Division, 2010). Some parts of these recommendations are presented later.

4. Overview of Post Enumeration Surveys

As it has been stressed earlier the primary objective of a census evaluation programme is to determine the sources and magnitude of coverage error and content error (for some selected variables). For many developing countries the post enumeration survey (PES) has become a reasonable independent evaluation programme. This is partly because other independent sources of data with

relevant, comprehensive and reliable information are not available (Hogan, 1992; Hogan and Wolter, 1988, 1999; UN Statistics Division, 2010).

In conclusion, it should be pointed out that for many developing countries basic data, to facilitate census evaluation, are lacking or insufficient. For example, to undertake a demographic analysis there is a need for very reliable data to make it possible to calculate the demographic components of the population, namely fertility, mortality and migration. In some developing countries these data are not available. In addition, many developing countries do not have comprehensive vital registration systems; therefore, sophisticated demographic analysis to evaluate the census may not be feasible.

It should, however, be emphasized that the PES can only generate reliable and accurate results if the sample is well and efficiently designed, its implementation is of high standard, the matching exercise is carefully done, and the analysis of results and estimation are correctly executed (UN Statistics Division, 2010). Basic principles of planning and implementing of a PES, using UN Statistics Division extensively, are presented below¹¹.

4.1. Planning of a Post Enumeration Survey

“The planning of a PES should be preceded by a clear and unambiguous statement of objectives of the evaluation. Planning for a PES should, to the extent possible, be synchronized with the planning for the census. It should start early and adequate resources should be devoted to it as part of the overall census programme. The success of the PES depends mainly on the availability of qualified human and other adequate resources and it has to be independent from the census operations. In some cases, insufficient resources are at the disposal of PES planners to support its thorough conduct. Without adequate resources, the quality of the PES results would seriously be compromised. It is therefore necessary for the organization responsible for the conduct of the PES to develop a plausible survey plan with adequate budgetary and manpower requirements clearly spelled out.

Cost

The cost is a determining factor as to whether a PES should be undertaken or not. There is need to have adequate financial allotment to ensure availability of qualified enumerators and supervisors, competent matching clerks, qualified data processors, adequate training for all involved and effective operational and quality assurance in the whole PES process. The sample size in turn will depend on whether only national estimates are required. In this case the sample size can be relatively smaller compared to an overall sample size aimed at getting reliable results for many different domains. The latter will require independent estimates,

¹¹ Adapted from: UN Statistics Division (2010), Post Enumeration Surveys, Operational guidelines, Technical Report, New York, April 2010. Available at: <http://unstats.un.org/unsd/demographic/standmeth/handbooks>.

which can only be reliable if the sample size is reasonably large, implying that adequate sample sizes are obtained for each domain with its specified reliability levels. Common domains include rural/urban, regional provincial or other sub-regional domains. Large samples, for example, will demand the recruitment of a large pool of enumerators, supervisors, data entry clerks, etc.

Publicity

In order to encourage active participation in the PES by as many respondents as possible, it is advisable to plan for and mount publicity campaigns. Advance publicity is necessary because it prepares potential respondents for the PES by soliciting their cooperation. In this way response rates may be increased. Different approaches to publicity may be adopted depending on prevailing circumstances in different parts of a country.

4.2. Implementation of a Post Enumeration Survey

A number of factors contribute to errors in executing a PES, among them: use of faulty maps defining enumeration areas and unclear addresses especially in rural areas, poor publicity, shortage of transport facilities and limited communication during the data collection exercise, poor planning for data collection and data processing activities coupled with resource constraints.

Pilot test. A comprehensive test of all PES procedures is advisable. This can be a dress rehearsal of the actual PES just as the pilot census is a dress rehearsal of the census. The pilot test can cover some selected administrative divisions. The aim is to test the adequacy of the entire PES plan and the PES organization. The PES pilot test should preferably be undertaken in conditions similar to the actual enumeration. This implies that the Pilot PES should immediately follow the census pilot test. The purpose of the pilot PES is to prepare for the main PES; however, while it is not a source of usable substantive data, it provides lessons pertaining to the operational aspects of data collection that can be implemented in the current census. Ideally, the pilot should be taken a year before the actual PES just as the pilot census is taken one year before the planned census

Matching of records between a PES and the census is one the main features of the evaluation exercise. It should, however, be stated that it is one of the complex and challenging undertaking in a PES programme. It has to be done well for the PES results to be useful. The results of a pilot contribute to the establishment of matching rules, reconciliation procedures, and logistical flow of documents between the PES and the census. It may be possible to make broad estimates of precision and accuracy on the pilot PES results, such as sampling errors and certain bias components of the total mean square error.

Data collection. The method commonly used in data collection with respect to the PES is the personal interview method. The method entails enumerators going to households, in selected EAs, and interviewing the respondents, thereby collecting information by asking questions from the PES questionnaire. The main

advantage of this approach is that the enumerator has the potential to ask probing questions. This is, in most cases, necessary in a PES. Additionally, enumerators are in an interactive mode with respondents such that they can explain to respondents the objectives of the PES when asked.

Questionnaire design. Questionnaires for the PES should be designed based on the final census questionnaire in order to facilitate an objective evaluation of the census. The PES questionnaire plays a central role in the survey process in which information is transferred from the respondents to the survey analysts. It is the vehicle through which the information needs of users are expressed in operational terms as well as the main basis of input into the data-processing system. It may be worth noting that if the questionnaire is to be used for recording responses by enumerators in the field, it should be sturdy enough to survive handling. It is also advisable that the questionnaire should be designed to facilitate the collection of accurate information.

A good questionnaire should have the following qualities:

- (a) Enable the collection of accurate data to meet the needs of potential data users in a timely manner;
- (b) Facilitate the work of data collection, data processing and tabulations;
- (c) Ensure economy in data collection avoiding the collection of non-essential information;
- (d) Permit comprehensive and meaningful analysis and purposeful utilization of collected data.

Selection and training of enumerators. As earlier stated, enumerators are at the interface with the respondents. Their work is critical to the success of the PES field work. In general, an enumerator should be able to effectively communicate with respondents and should have qualities needed to collect accurate information from respondents in a timely manner. An enumerator for the PES must have an adequate level of education and should be able to record information honestly. It is important that the selected enumerators should follow instructions and use definitions and concepts as provided in the enumerators' manual.

The selected enumerators should be thoroughly trained before being assigned to do field work. It should be noted that the main objective of the training programme is to enhance uniformity and minimize measurement error, in interviewing procedures of the PES. Qualified instructors, who are well versed in the objectives of a PES, should be responsible for training. It is advisable that the trainers should be part of the PES planning and implementing team. In addition to the following lectures, trainees should take turns in explaining to others the various items in the questionnaire. In addition, practical sessions should be arranged both in the classroom environment and actual field situation. The training programme should result in decision by the PES director of which trainees may require additional training and whether any of them are entirely unsuited for the assignment.

The role of supervisors. It is recognized that training is a prerequisite of effective and successful PES field work. Notwithstanding the foregoing, training without proper supervision may not yield accurate results. In reality the success of PES fieldwork demands dedicated and effective supervision by supervisors that are supposed to be more experienced and better qualified than the enumerators. Like the enumerators, supervisors should undergo extensive training in all aspects of the PES. It should be underscored that a supervisor is an important link between the PES planners or management and the enumerator. The supervisor is supposed to organize work for enumerators, by determining field assignments. They review completed work and maintain a high level of commitment by enumerators to the PES programme. In order to achieve the above it is suggested that about five enumerators, at the maximum, should be assigned to one supervisor. A supervisor can play a central role in making follow-up of non-respondents. Non-response is a common phenomenon during a PES field work, just like in any other surveys. Because supervisors are supposed to have better qualifications and more experience than enumerators, there are therefore in the best position to contact the non-respondents and try to collect the requisite information.

Field data collection. The following are some of the socio-demographic variables included in a census questionnaire and repeated in the PES questionnaire for matching content error. The listed variables below are relatively easy to measure and are considered important demographic and social variables worth measuring response error if any. They include: a) age, b) sex, c) relationship to head of household or reference person, d) marital status, e) education level, f) type of housing unit.

4.3. Independence

The Dual System Estimation methodology is based on the assumption that the PES is an independent collection from the Census. The method is modelled on the technique of capture-recapture commonly used to estimate the population of wildlife. The methodology assumes a closed population. The assumption being that the population remains unchanged during the period of the study. Independence, therefore, requires that the PES must not be influenced by what took place in the census. For example, in theory, the frame should be independent from that of the census, the planning of a PES must be done by an independent group of people; in addition the PES implementation should be independent from that of the census. This implies that different enumerators and supervisors are supposed to be selected and deployed for the PES from the census. However, what is maintained is operational independence from the census at every stage of the PES such as enumeration, data processing and administering the survey. In order to maximize the independence between the two exercises, in some countries, the enumeration in the census is through the self-completed questionnaire, while the PES is conducted face-to-face and the frame for PES is independent from the census enumeration areas. In addition, the estimation procedure is based on the case-by-case matching of two different and independent sources describing the same event.

4.4. Control of non-sampling error

In all surveys there are bound to be non-sampling errors, the focus should be to minimize them. Both sampling and non-sampling errors should be controlled and reduced to a level at which their presence does not compromise the usefulness of the PES results. Non-sampling errors are particularly harmful when they are non-random, because they introduce bias in PES estimates. Such bias is complicated or difficult to measure. The best way to control non-sampling error is to follow the right procedures in all PES activities including planning, sample design, data collection, processing and analysing results. Emphasis is laid on careful and intensive training of field staff.

The following are some of the factors which contribute to nonsampling errors in PES:

- (i) Vague objectives resulting in inadequate and/or inconsistent specifications with respect to objectives;
- (ii) Duplication or omission of units in the PES due to imprecise definition of the boundaries of the EAs;
- (iii) Inappropriate methods of interviewing, observation using ambiguous questionnaires, definitions or enumerator or supervisors instructions;
- (iv) Lack of trained and experienced field interviewers including lack of good-quality field supervisors;
- (v) Incomplete identification particulars of sampling units or faulty methods of interviewing;
- (vi) Errors occurring in data processing;
- (vii) PES respondent does not know census occupants or remembers incorrectly;
- (viii) Adequate information for movers is inadequate or not available;
- (ix) Other identifying information is poorly remembered for out-movers”.

4.5. Challenges of Post Enumeration Surveys

There are a number of problems which constrain some countries from carrying out post enumeration surveys. UN Statistics Division (2010) briefly presents below some of the key problems and suggested solutions:

- (i) “In some developing countries there is lack of technical personnel with requisite skills and experience in survey methodology and in designing and implementing the whole PES process. These challenges are in areas such as sample design, implementation, matching and estimation. However, it can be argued that if a country can design and conduct good and efficient household surveys, the chances of conducting a good post-enumeration survey are high. There is therefore need for countries to develop and maintain capacity in sample survey methodology and implementation;
- (ii) Lack of financial resources after investing so much in a census. It is against this background that it is advisable that, with respect to financial

resources for the PES, the planning should be an integral part of the overall census Programme;

- (iii) It is not possible to maintain the theoretical independence required between the census and the post enumeration survey. The commonly adopted strategy is to devise practical or operational approaches of maintaining independence. For example, planning and management of a post enumeration survey has to be undertaken by personnel that is separate from census personnel. In practice, independence is maintained by putting into place seemingly independent field procedures that are implemented to try and improve the enumeration in the post enumeration survey compared to the census count. These include:
 - (a) the use of enumerators and supervisors who are better qualified than those used in the census or the best staff used in the census;
 - (b) assigning the post enumeration survey staff to areas in which they did not work during the census;
 - (c) using a questionnaire which asks more detailed and probing questions on selected characteristics;
 - (d) census results for designated areas should not be known by staff who are assigned to those areas;
 - (e) all census materials from the selected PES areas should be collected before PES enumerators go into the field;
- (iv) The design of the survey, matching and estimation procedures may be perceived to be complex. These problems can be solved or mitigated by having good sample survey methodologists and analysts including employing qualified and well-trained enumerators and matching clerks and supervisors;
- (v) The post enumeration interview can be demanding. It usually incorporates questions to determine if a respondent should really be counted at the residence in question. In addition, the post enumeration interview takes place after the census interview, at which point the respondent may feel overburdened and not be as forthcoming with accurate information.
- (vi) In some countries, census planners feel that it is enough to put in place good quality assurance procedures at various stages of a census. The truth is that a census is a massive operation such that despite the assurance procedures put in place, error is bound to creep in. It is against this background that an evaluation of census results is still necessary despite the quality assurance procedures having been implemented in a census;
- (vii) Some countries are ambivalent about conducting a post enumeration survey after the fatigue resulting from the census operation. A census being a difficult and taxing operation, which saps the energy of those associated with it, discourages some national statistical/census offices to conduct a PES. Others feel that the exposure of discrepancies between the census and post enumeration survey results to users would be

detrimental to the reputation of the census or statistical organization. With respect to fatigue this is the reason why (UNSD, 2010) advocates that the PES should be as independent, to the extent possible, from the census activities. So that a different team would be responsible for the PES.

- (viii) In the case of the ambivalence of having evaluation figures for the census, it should be recognised that it is common in all credible statistical studies to have quantitative measures of error. PES evaluation results, therefore, would enhance confidence among informed *users* of census data, contrary to the negative view stated above. The evaluation would not diminish the importance of the census as long as users understand the limitations of the data and errors do not affect the major uses of the data.”

It should further be emphasized that an evaluation of a census is necessary for a number of reasons, among them is the fact that the population census is the most extensive and expensive data collection exercise for many countries. In addition, censuses have in recent years become complex. With vast amounts of resources spent, there is usually considerable pressure on census takers to ensure that census results are accurate to facilitate informed decision making at national and other domain levels. In addition, because of the massive nature of the census operation, it is inevitable that some inaccuracies such as errors of coverage and content/responses are unavoidable.

5. Strengths and weaknesses of evaluation methods

Methods based on a single source of data provide less insight into the magnitude and types of errors in the census data than the methods based on comparison of two or more sources of data. Examples are age and sex distribution analyses, which provide a general impression of the quality of the census results, but provide little insight on relative contributions to coverage and content error. The advantage is that such methods do not require additional data to be collected for evaluation purposes and, in general, there is no need for sophisticated matching operations. Such methods can, however, complement other methods of evaluation such as post enumeration surveys.

Demographic analysis has the advantage that no additional data is needed to be collected to perform this analysis. Information is already available, therefore, it is less costly and where the national statistical/census office has demographers there may be no need for additional staff to carry out the technical analysis. The limitation of single data source methods is that they provide less insight into the type and magnitude of errors present in the sources of data.

Results of well-designed and implemented interpenetrating subsamples can give good insights into different contributions of component errors to total error. This type of evaluation helps in the identification of operational stages that contribute to census error. However, such studies are relatively costly involving many field staff, intensive training and close supervision.

The PES is an independent evaluation method of a census. It demands adequate financial, human and other resources. A successful PES calls for a good sample design and survey implementation. Mention should be made that the matching exercise can be somewhat complex. The major advantage of matching over non-matching studies (an analysis that does not require matching censuses records with another source) centres on their ability to provide separate estimates of coverage and content error. On the other hand, the non-matching studies, because they review census results at the aggregate rather than unit level, for example, housing units, households or persons, provide only estimates of net census error. The characteristics that can be evaluated from matching studies are much more than those for non-matching studies, which are usually limited to age and sex distributions. Matching evaluation methods, however, require high level technical skills, managerial and financial resources.

In summary, since the demographic analyses are, in general, undertaken irrespective of a PES being conducted, the critical decision as to whether or not to conduct a PES lies in the quality and variety of demographic data available. Accurate data on fertility, mortality and migration levels and trends are needed.

Demographic analysis often depends on previous census data which may also be flawed. In such situations, therefore, the PES approach, though relatively complex, may be the only reliable way of evaluating census error. This, however, does not preclude complementing the PES with demographic analysis approaches in situations where required data are available.

6. Concluding remarks

The author has outlined different methods of quality assessment of population census data, using mainly UN Statistics Division (2010) and National Research Council (2009) recommendations. He has added also UN (2015) Recommendations for the 2020 Censuses of Population and Housing, accepted by the Conference of European Statisticians. He had in mind the Polish experience in this field and preparation of the Polish 2021 Census of Population and Housing. Some of the methods presented here were used in the last Polish population censuses, such as *demographic analysis* and *post enumeration survey*, *small area estimation methods* and *imputation* (Gołata, 2012, 2016; Paradysz, 2010), but final solutions are not yet known. Nevertheless, he would like to draw attention to the following aspects connected with census quality assessment:

1. ***Different modes of data collection*** were used in the 2011 Polish Census of Population, such as: traditional census using electronic questionnaire and a terminal of hand-held type, CAPI - Computer Assisted Personal Interview, Self-registration by Internet, CAII - Computer Assisted Internet Interview and CATI - Computer Assisted Telephone Interview. Each of these modes is the source of some kind of errors. They should be verified by the 2021 Polish Pilot Census.

2. **Basic data are to be collected on sampling basis.** On average, 20 percent of dwelling units have been selected in each powiat (county), but sampling fractions were different in each county: the smaller numbers of dwelling units in the county, the higher sampling fraction were selected. Sampling allocation in different strata should be verified for the next census.
3. **Demographic analysis** has the advantage of no additional data needed to be collected to perform this analysis. Information is already available, therefore, it is less costly and where the national statistical office has demographers, as it is in Poland, there is no need for additional staff to carry out the technical analysis. The limitation of single data source methods is that they provide less insight into the type and magnitude of errors present in the different sources of data. Demographic analysis often depends on previous census data which may also be defective. In such situations, therefore, the PES approach, though relatively complex, may be the only reliable way of evaluating census error. In summary, since the demographic analyses are, in general, undertaken irrespective of a PES being conducted, the critical decision as to whether or not to conduct a PES lies in the quality and variety of demographic data available. The GUS should give priority to research on improving demographic analysis in the four areas: (i) improving the measurement of undocumented and documented immigrants, (ii) development of sub-national geographic estimates, (iii) assessment of the uncertainty of estimates from demographic analysis, and (iv) refining methods for combining estimates from demographic analysis and post enumeration survey data.
4. **A post-enumeration survey** is worth conducting if it is carefully planned and functioning within operational and statistical constraints. Cooperation of the different kind of experts involved in preparation, implementation, processing and publication of a population census is very important for the quality of census results. The dual system estimation methodology, which is key to the PES philosophy, assumes *independence* between *the census* and *the PES*. However, it should be noted that conducting a PES is demanding in terms of planning, sampling design, data collection and supervision, matching of PES and census results. The prerequisites for a successful PES are having adequate resources; qualified enumerators and supervisors; good survey statisticians and analysts; and efficient and careful implementation of all the activities related to the survey. **If these conditions are not fulfilled, it is better to stop conducting PES.**
5. **Application of small area estimation methods.** To increase the precision of estimates obtained on sampling basis, small area estimation methods may be applied if good additional data are available. Different kinds of registers are to be used in the census if the qualities of these registers are verified. How the registers are to be used is still not clear. It seems that in Poland we have some experience in this field but regarding the last population census, it is necessary to show how effective the methods of small area estimation were for obtaining data from sample surveys and available statistical sources.

6. ***Permanent data quality improvement.*** The author thinks that a PES proposed in the census is not efficient and will produce the same kind of data as post enumeration surveys in years 1978, 1988 and 2002 (Kordos, 2007). There were problems with independent matching and double system of estimation. He suggests conducting post enumeration survey after the main census using methods presented in (UN Statistics Division, 2010) adjusted to the Polish conditions or stop conducting PES.
7. ***Preparation for the Polish 2021 Census of Population and Housing:*** Preparation for the 2021 Polish Census and Housing has been started, but first of all, the assessment of the quality of the last census should be completed. All obtained results of analysis should be published, giving opportunity of other experts for contributions in this field.

REFERENCES

- BARTOSIŃSKA, D., (2006). Attempts at Applying Small Area Estimation Methods in Agricultural Sample Surveys in Poland, *Statistics in Transition*, Vol. 7, No. 6, December 2006, pp. 1203–1218.
- DAUPHIN, M., CANAMUCIO, A., (1993). Design and implementation of post-enumeration survey: developing country example. Washington: International Statistical Programs Center, US Bureau of the Census.
- EUROSTAT, (2007). Handbook on Data Quality Assessment: Methods and Tools.
- EUROSTAT, (2009). Handbook on Quality Reports.
- GOŁATA, E., (2009). Economic activity in population census 2011 and administration resources. In: Gołata, E. (Ed.), *Methods and Sources of Obtaining Information in Public Statistics*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu. Poznań.
- GOŁATA, E., (2010). Indirect Estimation of Economic Activity for the Register-based Census. In: Gołata, E. (Ed.), *Measurement and Information in Economy*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań, pp. 85–104 (in Polish).
- GOŁATA, E., (2012). Population Census and Truth, *Studia Demograficzne*, No.1, 161, pp. 23–55 (in Polish).
- GOŁATA, E., (2014). Census quality of the new generation, *Wiadomości Statystyczne*, No. 5, pp. 26–53. (in Polish).
- GOŁATA, E., (2016). Shift in Methodology and Population Census Quality, *Statistics in Transition-new series*, December 2016, Vol. 17, No. 4, pp. 1–14.

- HOGAN, H., (1992). The 1990 Post-Enumeration Survey: An Overview, *The American Statistician*, Vol. 46, No. 4, pp. 261–269.
- HOGAN, H., (2003). The Accuracy and Coverage Evaluation: Theory and Design. *Survey Methodology*, 29 (2): 129–138.
- HOGAN, H., WOLTER, K. (1988). Measuring Accuracy in a Post-Enumeration Survey, *Survey Methodology*, Vol. 14, No. 1, pp. 99–116.
- KORDOS, J., (1987). Data Accuracy in Social Surveys (in Polish). BWS, GUS, t. 35, Warszawa.
- KORDOS, J., (1988). Quality of Statistical Data (in Polish), PWE, Warsaw , pp. 244.
- KORDOS, J., (2007). Some Aspects of Post-Enumeration Surveys in Poland, *Statistics in Transition-new series*, December 2007, Vol. 8, No. 3, pp. 563–576.
- NATIONAL RESEARCH COUNCIL, (2009). Coverage Measurement in the 2010 Census. Panel on Correlation Bias and Coverage Measurement in the 2010 Decennial Census, Robert M. Bell and Michael L. Cohen (Eds.). Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press: <http://www.nap.edu/catalog/12524/coverage-measurement-in-the-2010-census>.
- NATIONAL RESEARCH COUNCIL, (2010). Envisioning the 2020 Census, 2010. <http://www.nap.edu/catalog/12865/envisioning-the-2020-census>.
- NOWAK, L., (1998). Quality of Census Data, In: *Tendencies of Changes in Structure of Population, Households and Families in 1998-1995 GUS*, Warsaw, pp. 22–31 (in Polish).
- NOWAK, L., (2008). The 2011 Census of Population and Housing – Methodology and Topics, the Committee of Demographic Research of the Polish Academy of Sciences (in Polish). http://www.knd.pan.pl/images/stories/pliki/pdf/Nowak_28_luty_2008.pdf
- ONS, (2005). One Number Census: Evaluation Report, Census 2001 Review and Evaluation, <http://www.ons.gov.uk/ons/guide-method/census/census-2001/index.html>.
- PARADYSZ, J., (1989). On non-sampling errors in women's fertility survey in the 1970 National Population Census. In: *Problems of Statistical Surveys by Sampling*. GUS, Warszawa, BWS, Vol. 36, pp. 154–159 (in Polish).
- PARADYSZ, J., (2010). Necessity of Indirect Estimation in National Census. In: Gołata, E. (Ed.); *Measurement and Information in Economy*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu. Poznań, pp. 45 – 66 (in Polish).

- STATISTICS CANADA, (2009), Quality Guidelines, Fifth Edition – October 2009; <http://unstats.un.org/unsd/dnss/docs-nqaf/Canada-12-539-x2009001-eng.pdf>.
- STATISTICS FINLAND, (2004). Use of Registers and Administrative Data Sources for Statistical Purposes. Best Practices of Statistics Finland, Tilastokeskus, Statistikcentralen, Statistics Finland, Helsinki.
- SZABLOWSKI, J., WESOŁOWSKI, J., WIECZORKOWSKI, R., (1996). Index of Fitting as a Measure of Data Quality — on basis of the Post-enumeration Survey of Micro-census 1995, *Wiadomości Statystyczne*, No. 4. pp. 43–49 (in Polish).
- UN, (2007). UN Principles and Recommendations for Population and Housing Censuses:
http://unstats.un.org/unsd/demographic/sources/census/docs/P&R2_February%2012%202007.pdf
- UN, (2010). Post Enumeration Surveys Operational Guidelines. Technical Report. United Nations Statistics Division New York.
Available at: <http://unstats.un.org/unsd/demographic/standmeth/handbooks/>.
- UN, Economic Commission for Europe, (2011). Using Administrative and Secondary Sources for Official Statistic, A Handbook of Principles and Practices, New York and Geneva.
<http://www.bing.com/search?ei=UTF&pc=AV01&q=UN+Economic+Commission+for+Europe+%282011%29+%2CUsing+Administrative+and+Secondary+Sources+for+Official+Statistic%2C+A+Handbook+of+Principles+and+Practices%2C+New+York+and+Geneva.+&FROM=AVASDF>.
- UN, Economic Commission for Europe, (2015). Recommendations for the 2020 Censuses of Population and Housing, New York, Geneva.
<http://unstats.un.org/unsd/demographic/sources/census/census3.htm>.
- WHITFORD, D. C., BANDA, J., (2001). Post-enumeration Surveys: are they worth it? Symposium on Global Review of 2000 Round of Population and Housing Censuses: Mid-Decade Assessment and Future Prospects, UN Statistics Division, New York, 7-10 August 2001.
- WOLTMAN, H., ALBERTTI, N., MORIARITY, C., (1988). Sample Design for the 1990 Census Post-Enumeration Survey, Proceedings of the Survey Research Methods Section, American Statistical Association, pp. 529–533.
- ZASEPA, R., (1993). Use of Sampling Methods in Population Censuses in Poland, *Statistics in Transition*, Vol. 1, No. 1, pp. 69–78.
- ZHANG, L.-C, (2011). A Unit-Error Theory for Register-Based Household Statistics, „*Journal of Official Statistics*”, 27 (3), pp. 415–432.