STATISTICS IN TRANSITION new series, December 2016
Vol. 17, No. 4, pp. 631–658

SHIFT IN METHODOLOGY AND POPULATION CENSUS QUALITY

Elżbieta Gołata[1]

ABSTRACT

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. We use 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.

Key words: population census quality, register-based census, coverage errors.

1. Introduction

A population census is not only the oldest investigation, best-known, well-formed in terms of methodology, but also an investigation which is widely regarded as the most reliable source of data. As the methods of conducting censuses, especially methods of data collecting, have changed incredibly over the last decades, it is important to address the issue of quality assessment of the population census under the shift in methodology. The purpose of this paper is to discuss the quality of information derived from the 2011 population census in Poland. Special attention is given to a comparison of census data accuracy in view of the conventional versus register-based approach.

[1] Poznań University of Economics. E-mail: elzbieta.golata@ue.poznan.pl.

Modern technologies, their development and application in all spheres of social and economic life influenced also the population census methodology. Huge changes and modifications can be seen at every stage of the census procedure, and, in principle, they are observed in all countries around the world (UN 2012). Some countries have opted for a fundamental change in methodology understood as a new source of data, while the others introduced only some innovations in the technology of data collecting and processing. One of the main reasons for these changes is the need of saving, but also improving census data quality (Longva, Thomsen, Severeide 1998, UN 2010a, 2012, 2013, CSO 2012). As noted by P. Valente (2010), it is essential to count the population, but census taking is costly, and a growing number of people are reluctant to participate. Similar opinions are also presented in the CSO report (2012), which emphasized that organization of the census turned out to be very expensive and laborious. For that reason, Poland decided to give up the traditional census in favour of a mixed method, in the 2011 round. National Statistical Institutes (NSIs) in many countries face growing challenges and difficulties: pressure to make greater use of information available elsewhere, lower public cooperation and participation, changing user demand and the need to control or reduce costs (UN 2011). These issues were discussed during various meetings and seminars and were ideally recognized by T. Holt in his presentation: *The Official Statistics Olympic Challenge: Wider, Deeper, Quicker, Better, Cheaper* (2007)[2].

According to the UN Recommendations for the 2010 censuses of population and housing (UN 2006), there are three primary approaches[3] to conducting a census, based on the method of data collection. These are: a) the conventional method of universal enumeration based on field operations at a given moment, b) the method of using registers and other administrative sources, and c) a combination of registers and other administrative sources and surveys. In the 2010 round, majority (56%) of the countries applied the traditional method, but the percentage using the register-based approach doubled amounting to 14.5%. However, if the complex method were also taken into account, this proportion would rise from 20% in the 2000 round up to 40% in the 2010 round (UN 2012).

Register-based censuses have already been conducted since the 70s (Statistics Denmark 1995, Statistics Finland 2004, Statistics Netherland 2004). Invaluable in this respect is the experience of the Nordic countries (Statistics Finland 2004, UN 2007, UN 2011). But the 2010 round brought a methodological shift in the way of conducting censuses in many countries (UN 2010a, 2012). The register-based approach and the mixed method greatly expanded. They were applied also by countries of a much more numerous population than Finland, Sweden, Denmark or Norway, and in countries with little experience in the use of administrative data in official statistics, like Poland and Germany. Poland is one of those countries

[2] This is the title of professor D. Tim Holt's invited presentation during a special seminar in celebration of the 100th meeting of Committee on National Statistics of The National Academies.

[3] In addition, two methods might be reported: the so-called "rolling census" carried in France and the traditional enumeration with yearly updates in the United States.

which decided to abandon the traditional method and to turn to the "mixed" one. Public administration registers and information systems were used as the census-data source, but data on different topics was also collected directly from population in a large-scale sample survey.

This study attempts to provide quality assessment of data from the 2011 population census in Poland. There are a number of methods to evaluate censuses including: post enumeration surveys[4], demographic analysis, interpenetrating studies used in conjunction with a current census, record checks and a comparison of census data with results of existing household surveys (Fosu 2001, UN 2009, 2010). Evaluation methods differ with respect to the type of error to be measured (coverage and content error), technical sophistication, data requirements and quality of results (Baldrige, Brown, Jones, Keane 1985). Baldridge et al. (1985) presented a typology of such methods and distinguished methods based on a single source of data and methods based on comparison of data from two or more sources (matching and non-matching studies).

According to UN survey (UN 2013b), demographic analysis was the method used by the greatest number of countries (76 per cent) for the measurement of either undercoverage or overcoverage. Additionally, differences in the methods used by countries conducting different types of census were observed. For a traditional census, a larger proportion of countries used a census coverage survey and demographic analysis. More countries, which adopted the register-based approach, used comparisons with aggregate administrative datasets and comparisons with existing surveys. Also, the majority of countries that implemented the combined method, conducted the comparison with unit level administrative datasets, an analysis of questionnaire return rates and demographic analysis. In all countries, regardless of differences in adopted census methodology, comparisons with existing surveys were the highest or second highest reported method.

The census coverage survey is usually considered as one of the best methods to assess the accuracy of census estimates, particularly in terms of coverage (Baldrige et al., 1985, Kordos, 2007, 2012, Gołata, 2012). Although the coverage survey was conducted in Poland shortly[5] after 2011 census, its results have not been published yet and there is no information on their use in the estimation. Additionally, results of the coverage analysis have not been presented by CSO, nor detailed description of the methodology applied. And CCS data are unavailable for researchers for the purposes of scientific research. However, at the end of 2015 preparations were undertaken by CSO to make individual unidentifiable data from 2011 census available for scientists.

[4] Both surveys: Post Enumeration Survey (PES) and Census Coverage Survey (CCS) are non-demographic methods of post census adjustments for population estimates. Therefore, if there was no explicit reference to a particular study, these terms might be used interchangeably.

[5] The census coverage survey was conducted from 1 to 11 July 2011, while the census data was collected from 1 April until 30 June 2011, as of March 31, hours. 24:00.

In this study we use multiple sources of data and non-matching methods, in particular: demographic analysis based on the previous census, vital statistics, birth and death registers, population register (PESEL) and the comparison with other existing sources. The choice of these methods is a result of the availability of data. It is natural to question quality of data from all sources to assess their reliability. But the process of the comparison analysis included in the combined method of taking censuses and creating the 'gold record' makes it possible to work with quality assessment in a new way (Wallgren & Wallgren, 2013). Polish Population census in 2011 was a combination of registers, other administrative sources and surveys that were subjected to a thorough quality assessment. This allowed for a thesis that the mixed census, as a combination of multiple data sources and as a part of a register-based statistical system, provides estimates that are not of a worse quality than those from a traditional census.

The assessment of the census quality under the new approach involves many methodological and practical issues. Some of them are discussed in the paper. The first group of problems relates to change in the methods in statistical research adopted for conducting the census. In the next section questions on assessing the quality of the census are discussed under essential aims and objectives, to provide reliable information on the population by age and sex in a detailed territorial division. Important issues involving methodological questions, data sources and types of errors are indicated. However, the study is limited to assessing the quality of the census due to coverage errors. The results of an empirical study are presented in the fourth section. Considerations are closed with conclusions and some final remarks.

2. Shift in statistics and population census quality

Currently, we are witnessing a change in the way of conducting statistical surveys (Baffour, King, Valente, 2013, Zhang 2012, UN 2011). Q. A and B. Bakker (2000) and P. van der Laan (2000) define changes in Statistics Netherlands as a process of reorganization of social statistics. To take one example, they present Sociaal Statistisch Bestand (SSB), a micro-data base obtained as a result of record linkage and statistical integration of different administrative records. Among most important reasons for these changes one may indicate: an increase in demand for information, the pressure to improve the efficiency of statistical process to make savings in costs and staff resources, demands to reduce the burden placed on the respondents to statistical surveys, but also the development of computing, data collection methods, data editing and integration and in estimation methodology (Wallgren & Wallgren, 2013, UN 2011). These expectations are often contradictory and force statisticians to consider alternatives to the traditional survey approach. The most natural is to see if usable data already exist elsewhere and may be used for statistical purposes. These data are rarely direct substitutes for those collected via statistical surveys,

but there are many possibilities like a combination of variables from multiple sources to obtain satisfactory results (UN 2011).

In particular, the above changes apply to a population census which is the oldest survey for counting people and recording their characteristics. Censuses are normally carried out once every decade for the whole population residing in the country. For centuries, the census was the most common form of examination the entire population (Bethlehem 2009). Census data constitute one of the most important source of information relating to demographic and socio-economic characteristics, because it provides a broad overview of a country's population to the lowest level of geographical division.

Although it dates back to works presented by John Graunt (1620-1674) and William Petty (1620 – 1683), the idea of studying some representatives instead of the entire population became more popular at the beginning of the XX century. The representative method has been developed by the works of Jerzy Neyman, Karl Pearson and Sir Ronald Fisher. Currently, sample surveys are best known and most commonly used method of conducting statistical surveys. However, due to the growing financial restrictions and an increasing number of data, nowadays a survey would not be carried out automatically. First, one would rather look at registers, administrative records and other existing sources (including the Internet and Big data (Ruggles, 2014)) to learn what information is available. Different data sources, like parish records, or other administrative records were also used in past. The research conducted by J. Graunt is widely known. In *Natural and Political Observations Made upon the Bills of Mortality* (1662), he used the mortality rolls in London to construct first life tables. But this was rare. At present, administrative records are easily available, although they are created for different purposes, there are many possibilities of exploring them for statistics. A new approach to obtain information for statistics is observed (Zhang, 2013, Al, Bakker 2000).

This change applies to the nature of statistics, understood as the whole process of obtaining information which is the basis for further research and analysis. And it is not just the data collecting, but the whole process of statistical survey (Wallgren and Wallgren, 2007, 2014). The increasing use of administrative data for statistical purposes is called by Wallgren and Wallgren (2013) the transition to the register-based statistics production system. This means a shift from a system based on address lists and interviews to the one in which sample surveys become 'register-based'. Registers are not only helpful in updating frames, improving sample scheme and survey design, but are also used in estimation process, provide auxiliary data for estimation or serve for evaluation purposes (UN 2011, Baffour, King, Valente 2013, Zhang 2011, 2012). The administration register data may be combined with other data sources as well as it can be used to improve other surveys in the system.

The shift in the process of statistical research, and particularly in the way of conducting population census, enlarged also the palette of evaluation topics, as many different criteria for a successful census might be listed (UN 2013a). Of

course, at first one would mention the two types of census errors: coverage and content errors (Baldrige et al., 1985). But the 2010 round of population censuses showed that the change in data sources induced a wider use of modern technologies and new methods. Instead of classical enumeration, data were extracted from administrative records, conventional field operations were replaced with Internet transfers. The use of multiple sources of data for the census induced developments in the use of data imputation, record linkage, calibration, estimation using auxiliary variables from external sources. Implementation of each of these projects can be considered as one of the evaluation criteria: improving quality of the registers, accuracy of the estimates, cost reduction, use of modern ICT in data collecting and dissemination. In view of the new methodology, including new data sources, quality assessment might be considered in a structural way. Berka et al. (2012) proposed a three stage approach to derive quality of raw, combined and imputed data in three hyperdimensions (Documentation HD_D, Pre-processing HD_P and External Source HD_E) to satisfy such requirements as transparency, accuracy and feasibility (UN 2012b). Discussing the change in census methodology, some authors indicate the need for conceptualization and measurement of the statistical accuracy in register statistics, which would enable application of rigorous statistical concepts such as bias, variance, efficiency and consistency, as in the case of survey sampling (Zhang 2011).

Wallgren & Wallgren (2013) discuss quality assessment for register-based statistical systems as a process consisting of two parts, each of which has two levels. The first one is to analyse the source itself. It includes a discussion of metadata regarding the analysed source to determine its relevance, and an analysis of microdata from the source to determine its accuracy. The second part is a comparison analysis of the source with its base register and with other sources in the system containing similar variables. Systematic comparisons between surveys and registers in the system give new knowledge of quality in different surveys, and also give new possibilities to redesign surveys to improve their quality.

By 2011, population censuses in Poland were carried out using traditional methods involving census enumerators visiting all inhabited units and noting down information obtained from respondents on census forms (available in hard copy). The 2011 Polish Census of Population and Housing (NSP 2011) was the first census conducted since Poland's accession to the European Union, and it took place in the period from 1 April to 30 June 2011 (as of 31 March 2011, at 00.00). The census was conducted by applying the mixed method with the use of administrative records (full survey - short form), supplemented by information from Internet self-enumeration. Additionally, a sample survey (long form) was carried out on approximately 20% of randomly selected dwellings. Data collected from administrative registers and sample survey formed the so-called golden record. This record was the result of integration of information from all data sources in the environment of Operational Micro-Database. Further processing

allowed for creation of Analytical Micro-Database which was used as the basis for census estimates.

As described in Berka et al. (2012) and Wallgren & Wallgren (2013), during the construction of the golden record, detailed studies and comparative analysis were carried out. In preparation for the 2011 census metadata about 300 various administrative registers were collected and analysed. All variables in those systems were rated with regards to the possibility of obtaining information on population, housing and buildings, in line with the recommendations and classifications of the United Nations Economic Commission for Europe (UNECE) and Eurostat (UN 2006). In preparation for the 2011 census Central Statistical Office (CSO) examined many administrative records and conducted a large-scale research of their conformity for the census as concerns concepts, definitions and classifications (*The report on the work of sub-group for the use* ..., 2007, *Memo from the current state of research* ..., 2007 Dziubiński, 2008 Kobus, Smolka, Nowakowska, 2009, *List of concepts and definitions* ..., 2007, Golata 2009). As a result of a detailed analysis, 28 registers were selected. Among them, as a priority, the following systems should be mentioned: Common Electronic System of Population Register (PESEL), Social Security System (ZUS), the Health Insurance System, Land and Buildings, Register of Territorial Division of the Country, data from the State Fund for Rehabilitation of Persons with Disabilities. Information collected from administration sources, which was properly structured and divided into strata, was also used in creating the frame for the census sample survey.

Social assessment of the new census methodology and attention of the scientific community are diverse: some opinions give full recognition and others are negative. The traditional census was perceived by the public, local government, and also by many scientists, as an indisputable source of 'certain' and unquestionable information (Barwiński, 2014, *Raport* … 2011). However, there are also clear assessments indicating that the previous arrangements were not ideal because of coverage errors (Sakson 2002, Śleszyński 2004, 2005) and due to the fact that the data was not collected directly from the respondents (Paradysz 2002). J. Paradysz (2010) underlines the need for critical evaluation of previous censuses and suggests the usage of all available data sources to improve census estimates. On the other hand, lack of comprehensive information on such topics as families and households, as well as unavailability of data for a detailed territorial division (information that was available only from the sample survey) is often considered as a disadvantage of the mixed census (Gołata 2013).

Coverage errors refer to either an undercount or overcount of units owing to omissions, duplication or erroneous inclusion. In the traditional census, an undercount was a typical situation (Paradysz 2002, 2010). Operational guidelines for conducting Post Enumeration Surveys (UN 2010) illustrate the use of various procedures with results obtained for selected countries, in majority, undercounts. As for the register-based census, Lenk (UN 2012b) underlines the importance of detecting inactive records in the population register and to eliminate them to avoid

overcoverage. Statistics Austria, for example, used the residence analysis, which allowed ensuring that only individuals with a pre-defined number of "signs of life" were counted in the census. All the individuals covered only by the population register, but not by other administrative source, were asked in a written form to confirm their main place of residence. Finally, approximately 0.5 percent of the initial population was not counted[6]. Unfortunately, CSO did not provide any information on applying similar procedure to avoid overcoverage in 2011 census in Poland.

There are a number of non-demographic methods of post census adjustments for population estimates: Post Enumeration Survey (PES), Coverage Surveys and the Reverse Record Check, Dual System Estimation (DSE) or Residents Temporarily Overseas (RTOs), used in many countries, e.g. in Australia, Canada, Japan, New Zealand, United Kingdom or USA (Newell and Smallwood 2010). Dual system estimation is one of the methods that can be used to generate population estimates from census data (Plewis et al. 2011, Brown et al. 2006). This method is based on the assumption of independence between census and census coverage survey data (different data collection methods, different personnel and a different address frame). The method applies matching procedures and detail analysis ensuring that individuals counted by both surveys are correctly allocated by age and sex within each of CCS estimation areas. Then, a two-dimensional distribution table is constructed to allow comparisons of the estimates. Having data from the population register and CCS, it seems possible to apply a similar approach (Tab. 1) to assess coverage.

Table 1. Basic Table for Dual System Estimation

Source of data	Observation	Census Coverage Survey		
		Observed	Not observed	Total
PESEL - Population Register	Observed	n_{11}	n_{12}	$n_{1.}$
	Not observed	n_{21}	n_{22}	$n_{2.}$
	Total	$n_{.1}$	$n_{.2}$	$n_{..}$

Source: Based on Plewis et al. (2011), Brown et al. (2006)

It is important to put attention to n_{12}, the number of persons observed in the population register, but not in the CCS. Differences in this dimension seem to be possible and significant, as Poland is a country of intensive migration, and people who migrate (irrespective of the length of their stay abroad) are included in the register unless they notify the authorities about leaving the country. Another real problem concerns the people who went abroad and died there. If nobody informs the Polish office of a death, the person may "live in the register" even up to 200 years. As a result, the oldest man of the world lives in Poland, according to the

[6] In Austria, due to the results of the test census, about 80 percent of the non-counted individuals were removed from the residence registers by local municipalities (UN 2012b).

register (Kuc 2014). Another important inconsistency is the number n_{21} , representing people observed in the CCS but not in the PESEL. This number may refer to all foreigners staying in Poland (even for more than 12 months) who do not have legal resident status. This status is associated with the registration for permanent residence, which requires submission of a document confirming the right of permanent residence[7]. In the case of a failure to meet formal requirements for permanent residence, even immigrants living in Poland for over 12 months were not counted as residents in the census.

Under independence assumption ($\theta = \frac{n_{11}n_{22}}{n_{21}n_{12}} = 1$) within a given age-sex group in the analysed territorial unit, we can estimate the census coverage and CCS coverage as well as the unobserved number of persons n_{22}. However, if the independence assumption is not valid then DSE will be biased: when $\theta > 1$, DSE has negative bias and when $\theta < 1$, DSE has positive bias (Brown et al. 2006). The assumptions of homogeneity and independence are very strong, and for several reasons they might not be met (Plewis, Simpson and Williamson 2011). Differences in probabilities of responding to CCS and of 'being included' in the register are possible, particularly by age and sex group, and for different territorial units. Some people may have no chance of being included in the register (for example due to legal regulations or because their propensity to respond to government enquires equals to zero). The matching process might be invalid for different reasons, e.g. migration, due to postponing the registration of newborn children, or inertness of the register in updating the reported changes.

However, not ignoring the importance of CCS in assessing the quality of the census, as there is no access to the data, this study provides census evaluation only in terms of demographic analysis in comparison to previous censuses, administrative data and other existing surveys, mirror statistics. Demographic analysis was carried out for the population of the whole country by sex and age, and with regard to certain aspects of territorial division.

3. Evaluation of 2011 population census in Poland

There is a considerable difficulty in identifying references for assessing accuracy of the estimates of the 2011 population census in Poland. The population register may serve as one of them. Another may be the census sample survey, which was conducted on a random sample of 20% of dwellings on the national scale. These two studies were the primary sources of census data, so they could hardly be considered as reference in assessing the census accuracy.

[7] This documentation differs depending on the immigrant's home country and may involve complex procedures. For a person from a country outside the European Union such documentation includes a permit for a long-term residency in the European Union, the decision to grant the refugee status in the Polish Republic, the award of subsidiary protection or tolerated stay permit in Poland, among other things.

The population register was evaluated during preparations for the census (Józefowski and Rynarzewska-Pietrzak,2010, Paradysz, 2010, Roszka, 2013). Recognizing generally very positive results obtained, the population register was accepted as the basis for the census data system. The census survey was one stage sampling scheme with deep stratification and consisted of more than 2,744 thousand dwellings, out of nearly 13.5 million. Although for all census results precision tables were provided, the original weights had to be adjusted due to 13.7% of non-response. Nevertheless, the analysis of non-response has not been available yet. In turn, Census Coverage Survey did not meet the requirements of an independent survey carried out in a more precise way. It was conducted by CSO using the same frame. A sample of 80 thousand dwellings was drawn out of 2,744 thousand flats drawn earlier to the census sample survey. But the frame was restricted only to flats with at least one person with an assigned phone or mobile phone number, and the survey was performed by CATI. Additionally, it covered all dwellings that took part in self-enumeration by the Internet.

All the above reasons influenced the decision to discuss the census quality in terms of demographic analysis in comparison to several existing data sources including the previous census. Previous traditional censuses in Poland were evaluated mainly by demographers, who used the possibilities of demographic analysis based on other existing data sources. There is quite well documented evidence on coverage errors in Polish censuses (Jończy, 2010, Kordos, 2007, Paradysz, 2010, Sakson, 2002, Śleszyński, 2004, 2005, Zasępa 1993). As concerns the coverage assessment, it is common to have a net census undercount as the number of omissions usually exceeds the number of duplications. Among the biggest coverage errors, J. Paradysz (2002, 2010) indicated a shortage of up to 30% of women with the shortest duration of marriage (1988 Census), omission of 10% of the youngest infants up to 6 months (2002 Census), omission of the population with increased mobility (2002 Census), lack of the elderly aged over 90 (2002 Census).

In the coverage assessment, special attention was paid to population at the age of an increased risk of biased estimates. These age groups were defined on the basis of earlier studies (Paradysz, 2010) and an introductory analysis. Special consideration was also paid to the fact that Poland is a country of intensive emigration, and consequently to the population at the age of particularly intensive migration mobility (as in the classical Rogers and Castro model).

We started the evaluation with a simple survival analysis. This phase consisted of a comparison of the census population in 2002 and 2011 by sex and age, including relevant aging. Survival rates are a basic tool in this case. The compatibility of survival rates was examined at first for: a) census data from 2002 and 2011 (Fig. 1 - thicker lines denoted as SR Census: solid for the entire population, dashed for men and dotted for women), and then also for b) projection of 2002 census data for 2011 (Fig. 1 - thinner lines denoted as SR Projection). Since both censuses, in 2002 and 2011, were carried out in spring (with the critical moments of May 21, 2002 and March 31, 2011), a simplifying assumption that the single age groups overlap was adopted. This means that an additional ageing for 1 month was omitted, and we assumed that a person aged 1 completed

year (according to the 2002 census), at the critical moment of the 2011 census, was aged 10 completed years. The survival coefficients were obtained according to formula (1).

$$SR_{\frac{2002}{2011}}(s,x) = \frac{\mathrm{P}_{2002}(s,x)}{\mathrm{P}_{2011}(s,x+9)} \tag{1}$$

where:

$SR_{\frac{2002}{2011}}(s,x)$ – between 2002 and 2011 censuses survival rate by sex and age

P_t – census population: t=2002 or t=2011

s – sex: *T*- both sexes; *F*- females; *M*- males

x – age

Both estimates of survival rates: (a) based on census data and (b) on demographic projection (of 2002 census data for 2011) are very much in line. This similarity of estimates obtained by using different data and different methodology indicates compatibility of the data. However, some values of the survival rates are cause for concern. The obtained results indicate the existence of such single year age groups, for which survival rates between censuses (SR Census) take values greater than unity. A closer look at the values of the SR Census coefficients allowed us to note that higher values were assigned for women than for men, although similar tendencies were observed. For both sexes the same age groups focus special attention. These are: 9-13 years, 16-19 years, 30 years and 33-35 with survival rates (SR Census) exceeding one. In addition, for the age of 65 years we observe a temporary collapse of the survival rate. It drops for about 5.5% from 0.89 to 0.84, but in the age of 66 years, almost 4% increase in the value (up to 0.88) was observed.

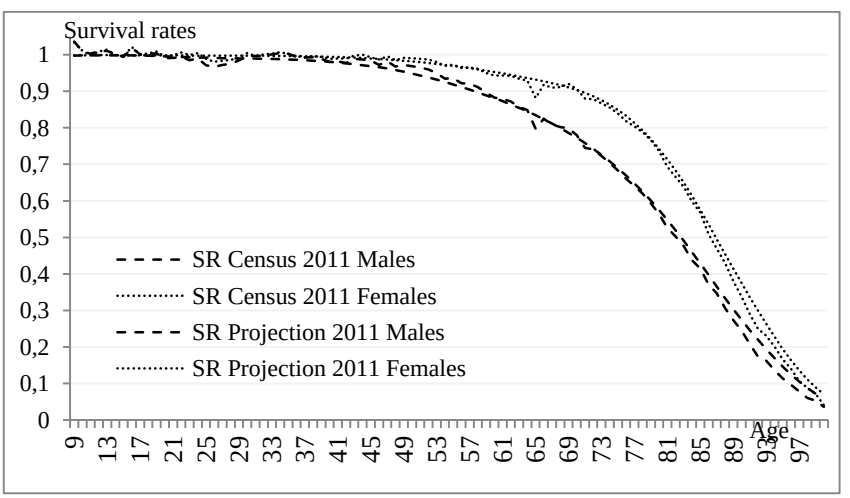

Figure 1. Survival rates for single year intervals between 2002 and 2011 census, Poland

Source: Estimates based on 2002 and 2011 Polish Population Census data and life tables, http://demografia.stat.gov.pl/bazademografia (Accessed 10 May 2014).

As between the censuses the survival rates (SR Census) relate to the population of people living at the time of the previous census, they should not exceed unity. Situations when this condition is not met suggest poor quality of the data or intense impact of migration. First, we analysed the impact of migration. The inclusion of migration in the analysis required relevant data that were available only for 5-year age intervals. Estimation for five-year intervals required an appropriate amendment in the formula. Summing the population for 5-year age intervals allowed the smoothing of the survival rates function according to age. The inclusion of migration allowed the elimination of unexpected values of survival rates (Fig. 2 and Tab. 2). The explanation of questionable survival rates was possible as the migration balance indicated a higher number of emigrants than immigrants. Nevertheless, Poland, which is historically a country of emigration, in recent years has served also as the host country for immigrants. However, both numbers of emigrants and immigrants differ significantly. The number of emigrants (according to 2011 census) amounted to over 2 million and was more than 50 times higher than the number of immigrants, which in 2011 census was estimated at approx. 40 thousand.

Statistics relating to migration raises legitimate uncertainty. However, they do not undermine the observed relationship (Kicinger and Koryś 2011, Fihel, Kaczmarczyk, Okólski 2006). Furthermore, recent migration has shifted to be more fluctuating with changing destination due to studies, employment or family reasons, indefinite period of stay and changes of the country of residence. This applies to the Polish migration after accession to the UE in 2004 in particular. Researchers are exploring these new phenomena under the concept of transnationalism (Borket and Penninx 2011).

Table 2. Survival rates for five year intervals between 2002 and 2011 census with and without migration, Poland

Age	Survival Rates					
	Without migration			With migration		
	Total	Males	Females	Total	Males	Females
10-14	**1.006**	**1.006**	**1.006**	0.974	0.974	0.974
15-19	**1.005**	**1.004**	**1.006**	0.979	0.979	0.979
20-24	0.995	0.990	**1.001**	**0.935**	**0.938**	**0.931**
25-29	0.980	0.974	0.986	**0.867**	**0.873**	**0.861**
30-34	**1.001**	0.999	**1.002**	**0.892**	**0.893**	**0.892**
35-39	0.995	0.994	0.996	**0.917**	**0.916**	**0.918**
40-44	0.988	0.983	0.993	0.926	0.921	0.931
45-49	0.983	0.975	0.990	0.930	0.921	0.939
50-54	0.969	0.955	0.983	0.929	0.913	0.944
55-59	0.941	0.918	0.963	0.913	0.889	0.936

Table 2. Survival rates for five year intervals between 2002 and 2011 census with and without migration, Poland (cont.)

Age	Survival Rates					
	Without migration			With migration		
	Total	Males	Females	Total	Males	Females
60-64	0.905	0.869	0.939	0.888	0.852	0.922
65-69	0.860	0.806	0.908	0.849	0.796	0.896
70-74	0.815	0.739	0.877	0.807	0.731	0.868
75-79	0.734	0.640	0.803	0.728	0.635	0.796
80+	0.456	0.392	0.490	0.451	0.389	0.485

Source: Estimates based on 2002 and 2011 Polish Population Census data, http://demografia.stat.gov.pl/bazademografia (Accessed 10 May 2014).

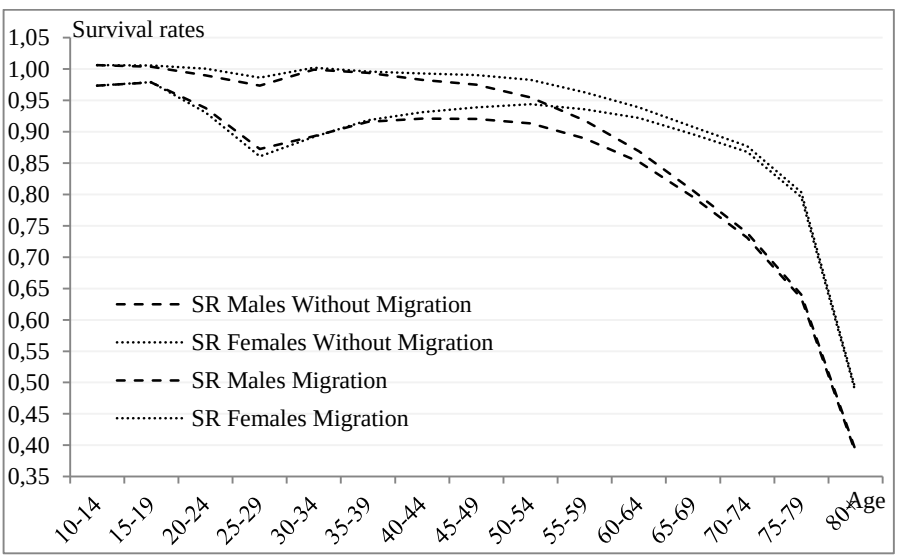

Figure 2. Survival rates for five year intervals between 2002 and 2011 census with and without migration, Poland

Source: Estimates based on 2002 and 2011 Polish Population Census data, http://demografia.stat.gov.pl/bazademografia (Accessed 10 May 2014).

Given that Poland is a source country for migrants, a significant decrease in the population was expected, and in fact introducing migration leads to a reduction of survival rates. It is especially pronounced in the age group with the highest mobility: 20-34 years old. For these age groups survival rates decreased by up to 11% (Tab. 2). We observe also a change in the relationship between survival ratios for men and women, reflecting differences in the intensity of migration by sex (Fig. 2). Survival rates without migration are higher for women than for men. This relationship is consistent with the survival probabilities of life

tables. The introduction of migration eliminates differences between the probabilities of the life tables and between censuses survival ratios. In addition, the reduction was so deep that we observe lower survival rates for women than men, as a consequence of migration. For women in the age group of 25-29 years, a drop in the survival ratios amounted nearly to 13%, while for men it was slightly above 10%. For subsequent age groups these differences are becoming less significant. Over 60 years of age, they do not exceed 2% and are rapidly converging to zero.

Returning to the assessment of census data quality, conducted with the use of demographic analysis, we compare 2011 census estimates (denoted as Census 2011 on Fig. 3) with estimates obtained by predicting population from 2002 census (denoted as Projection 2011 in Fig. 3). The population by sex and age of the 2002 census was adopted as a starting point for the projection. Survival probabilities from the life tables for the years 2002-2011 were used to obtain age and sex structure of the population for the following years, similarly as in population predictions. Live births by sex in years 2002-2011 from vital statistics evidence were incorporated and subjected to the ageing procedure. A two-stage approach was applied. The first stage used data from the records of infant's deaths by month. These data allowed estimating the number of children completing the first year of life. Next, the above estimates were subjected to the second stage of the ageing procedure that used survival probabilities of the life tables for subsequent years of age.

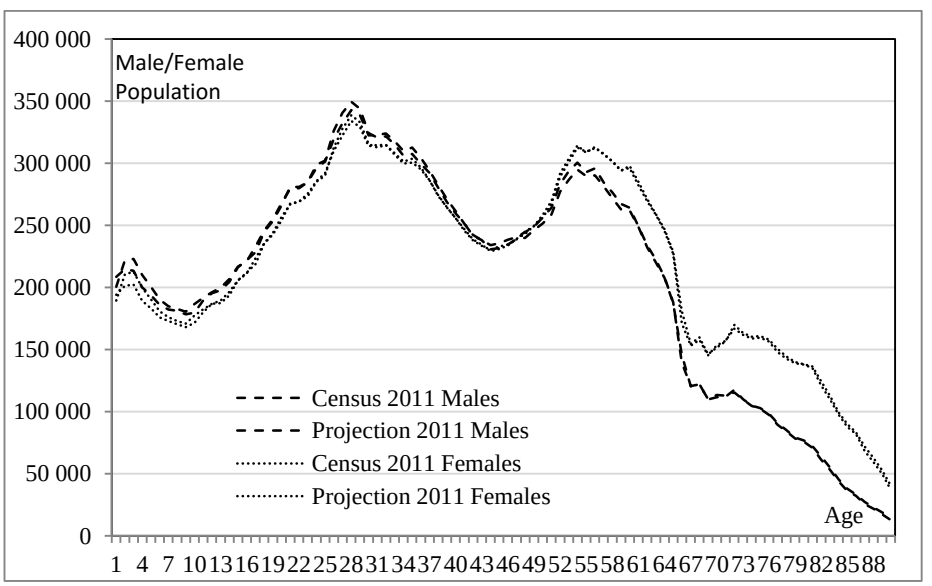

Figure 3. Population by age and sex: 2011 census estimates and projection, Poland

Source: 2002 and 2011 Population Census data, life tables and vital statistics years 2002-2011, http://demografia.stat.gov.pl/bazademografia (Accessed 10 May 2014).

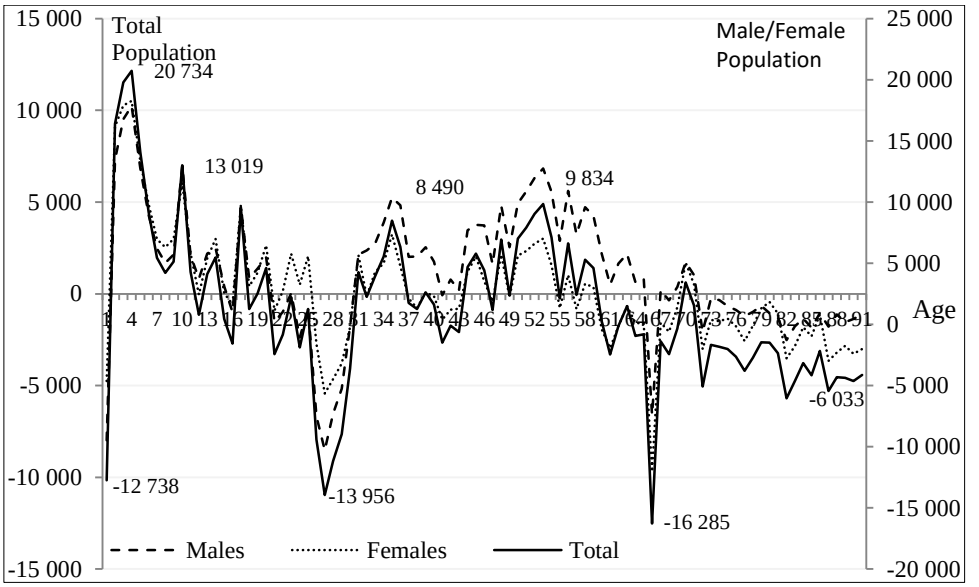

Figure 4. Differences between 2011 census population and projection, Poland

Source: 2002 and 2011 Population Census data, life tables and vital statistics years 2002-2011, http://demografia.stat.gov.pl/bazademografia (Accessed 10 May 2014).

Initially, this procedure did not consider migration as data on migration are not available for one age intervals. Additionally, as indicated above, the measurement of migration gives rise to considerable controversy (Fihel, Kaczmarczyk, Okólski 2006). This issue, along with the appropriate analysis requires a separate study. On the other hand, omission of migration allows evaluating the quality of census estimates. Generally, one can expect projection compliance with the 2011census data, except for those years of age in which intense migration was observed. The above presented algorithm for population estimates was applied on the national scale as well as for selected regions.

Comparison analysis of the two population estimates (Census 2011 and Projection 2011, Fig. 3) allows observing good compatibility at national scale. Difference between estimates of the 2011 census and the reference population is small and amounts about 100 thousands in total, which constitutes 0.26%. However, direction of this relation differs by sex. For women underestimation is observed while overestimation for men. This relation varies by age (Fig. 4) allowing identification of those age groups which require special attention: (i) infants and children 0-4 years old, (ii) young people: studying and starting their professional career, (iii) working age population, (iv) the elderly.

While analysing population prediction based on 2002 census compared with 2011 census data, one must constantly bear in mind the methodological differences resulting from the method of the census: traditional and register-

based. However, regardless of the applied methodologies, we assess the compatibility of the final estimates. Note that if the difference between 2011 census and 2011 projection is negative, it means that the administrative registers do not cover a specific group of people - underestimation is observed. The opposite situation, when the 2011 census data are greater than expected from the projection, suggests that census data in fact show a non-existent population.

In each of the described cases, the justification of observed discrepancies would be desirable. In the previous discussion attention has been paid to the impact of migration, which is especially obvious for the age of 20-35 years. In the following discussion we turn attention to the possible explanations of differences observed for infants and children up to 4 years.

3.1. Coverage assessment – infants and children

A comparison of census population aged 0 completed years at the census critical moment (1 April 2011) with data on live births from vital statistics evidence allowed the evaluation of similarities and differences of these two data sources. With the register-based census approach, it is natural to expect full agreement with vital statistics. For the aim of this analysis, detailed data on births by month and sex (for five years before census) was incorporated and subjected to the ageing procedure. As described above, the two-stage approach was applied. Depending on availability of data on infants deaths by age in months and in time (by month of the year), vital statistics data was used to estimate the number of children completing the first year of life. In the analysis for subsequent years of age, more aggregate data was used based on survival probabilities of life tables for single year of age. To provide a comparable assessment for 2002 census, an identical procedure was applied to information from the years 1997-2002.

Data obtained for infants showed that 2011 census population was underestimated by nearly 13 thousand compared with the birth and death registers (Tab. 2). This represents 3.3% of census population. For the total number of children aged up to 4 years, an opposite situation was observed. This group of census population was overestimated by more than 58 thousand compared with the birth register. This result is difficult to explain, since census data showed children not included in the birth register. A common mistake is rather to underestimate the population, whereas the overall overestimation by 2.8% was observed in this case.

An underestimation of the number of infants is often explained by a delayed birth registration. An assessment of the population register, which was carried by Józefowski and Rynarzewska-Pietrzak (2010), indicated at least two-week delay in the transmission of information about new-born children. In addition, the authors raised the problem that the register does not take into account all the events as of the indicated date. This ‘outdate’ of the register means also the omission of infants. And it is particularly important that administrative records include all the events as of the indicated date – the critical moment of the census.

Table 2. Coverage assessment for infants and children: differences between 2011 census data and estimates based on vital statistics

Age	Total		Males		Females		Population aged 0-4 years and the difference between estimates		
	persons	%	persons	%	persons	%	2011 Census	Birth Register	Difference
0	-12 738	-3.3	-7 981	-4.0	-4 757	-2.5	2 057 998	1 999 725	58 273
1	16 414	3.8	7 297	3.3	9 117	4.3			2.83%
2	19 776	4.5	9 511	4.3	10 265	4.8			
3	20 734	5.1	10 229	4.9	10 505	5.3			
4	14 086	3.6	6 857	3.4	7 230	3.8			
5	9 192	2.5	4 304	2.3	4 888	2.7			

Source: Estimates based on 2002 and 2011 Polish Population Census data, life tables and vital statistics (records of births and deaths in the years 2002-2011), http://demografia.stat.gov.pl/bazademografia (Accessed 10 May 2014).

Table 3. Coverage assessment: differences between the number of children under one year of life according to the birth register and 2002 and 2011 population censuses

Data source	2011 Census			2002 Census		
	Total	Males	Females	Total	Males	Females
Census	389 903	200 592	189 311	351 662	180 116	171 546
Birth Register	402 641	208 573	194 068	357 096	183 440	173 656
Difference	-12 738	-7 981	-4 757	-5 434	-3 324	-2 110
Difference (%)	3.27	3.98	2.51	1.55	1.85	1.23

Source: Estimates based on 2002 and 2011 Polish Population Census data, life tables and vital statistics (records of births in the years 2002-2011), http://demografia.stat.gov.pl/bazademografia (Accessed 10 May 2014).

Similar differences were also observed earlier, in censuses conducted with the conventional approach (Paradysz, 2010). However, it might be expected that while conducting a census based on a population register, greater compatibility will be observed than in the case of a traditional census with independent field operations. Detailed information on the census of 2002 and 2011 indicates quite the opposite situation (Tab. 3). A discrepancy between the estimates of census and the birth records was identical in direction, but in 2011 it amounted to 3.27% and was higher by more than 100% in comparison to 2002, when it was equal to 1.55. These results imply a need for further work on the quality of the population registry in cooperation between responsible government authorities and public statistics and in consultation with the scientific community.

As already discussed, for children aged 1-4 years an opposite relation was observed. This time it was an overestimation. When we focus on children aged 1 completed year or older, the differences between census estimates, and the ones resulting from the birth register were positive. The biggest difference related to children aged 3 years (completed in 2011) exceeded 20 thousand (more than 5%). The overestimation was slightly higher for girls than for boys. For the total number of population aged 1-4, census population exceeded register data by 71 thousand, that is 4.3%.

The observed discrepancies might be associated with intensive migration and an increasing number of births given by Polish women abroad, especially in the United Kingdom (Janta 2013, Waller *et al.*, 2014, Zumpe *et al.*, 2012). Where are the infants registered as born in Poland, but not enumerated during the census? Where are the children aged 1-4 enumerated by the census, who were not listed in the Polish Birth Register? Answers to these questions are beyond the scope of this paper. Probably, an indefinite life situation might suggest Polish migrants to enumerate their children in the census survey in Poland. Some of them not only decided to enumerate their children in Poland, but also registered them in the home country. This might be logically explained, as children born abroad to Polish parents need to be also registered in Poland in order to gain Polish passports. And a large number of children born in England and Wales to parents of Polish citizenship obtain Polish passport (ONS 2013 p.23). The ONS data for the youngest age group (0-4 years), show difference amounting almost to 50 thousand between Polish-born and Polish nationals. The ONS data show also 74 456 live births in UK to Polish women in 2007-2010 (Zumpe *et al.* 2012 p. 24). An in-depth mirror statistics might reveal some trends, but it is basically impossible to provide exact numbers.

3.2. Coverage assessment – young people

The analysis referring to young people receiving education and starting their professional careers identifies another important problem. As showed above, the census underestimates population of young people aged 25-30. However, the analysis carried out at regional level gives various examples, either confirming or not confirming this observation.

It is worth noting that one of the primary purposes of the census is to provide information on population by age and sex in detailed territorial division. In the census based on administrative records, this task is fulfilled with respect to information that comes from registers. In relation to those characteristics of the population that can be possessed only from a sample survey, the problem of estimating for small areas arises. This means verification of compliance of definitions and classifications, data integration, examining the relation between different characteristics to choose auxiliary variables, methodological studies on estimation for small domains, assessment of consistency, calibration, etc. Bearing in mind all the above problems and their impact and consequences for the assessment of the census quality, we confine ourselves with the coverage analysis.

In regional dimension, the coverage is obviously different from that for the whole state. Poland is a country characterized by large regional differences. Therefore, a comprehensive evaluation of census data by territory will certainly provide a wide variety of information that would be extremely valuable in a regional development strategy. For this reason, the evaluation of regional census data would require a separate study. In this paper we focused on assessing the compatibility of the estimates for young people aged 20-35 years. The overall assessment for the whole country showed a significant underestimation of this group of people. Data at the country level is a balance of regional assessments. There are such territorial units for which the indicated underestimation would be even greater. But there are also such regions where we obtain contrary information, with the case of large cities as an example. Exemplary considerations apply to the population of Poznan - the fifth largest city in Poland with half million inhabitants (554 696). At regional level, one may notice greater discrepancies between 2011 census data and the projection than on the national scale. The total number of residents of the city was underestimated by more than 20 thousand, which gives 3.7% of the 2011 census population.

The differences between 2011 census and the projection are widely disparate according to age (Tab. 4). The biggest underestimate of 11.7 ths. (23.1%) refers to Poznan residents aged 30-34 years. On the other hand, the greatest overestimation of 10.5 ths. (24.1%) was observed for the age of 20-24. The difference between the number of people in a given age group according to 2011 census and corresponding population of respectively younger age group according to 2002 census shows unusual trends. The observed changes do not result from natural demographic processes, births and deaths, but internal migration and the suburbanization process.

Table 4. Differences between 2011 census data and projection based on 2002 census, children and youth, Poznan

Age	2011 census - adequate 2002 census population*		2011 census population – projection based on 2002 census	
	Absolute	Relative (%)	Absolute	Relative (%)
10-14	-2 206	-9.7	-2 360	-11.4
15-19	1 198	4.8	221	0.9
20-24	11 887	37.6	10 471	24.1
25-29	8 982	20.3	4 318	8.1
30-34	-12 595	-19.9	-11 684	-23.1
35-39	-10 957	-21.1	-7 058	-17.2

Note: *Difference between the number of people in a given age group according to 2011 census and corresponding population of respectively younger age group according to 2002 census

Source: Estimates based on 2002 and 2011 Polish Population Census data, life tables and vital statistics (records of births and deaths in years 2002-2011), http://demografia.stat.gov.pl/bazademografia (Accessed 10 May 2014).

Similar discrepancies for the age group 20-35 were observed by T. Józefowski and B. Rynarzewska -Pietrzak (2011), who studied the quality of the population register. They indicated that the reason for these discrepancies (amounting even to 34%) is the relationship between the actual population (census) and permanent residents (register), which results from the fact that Poznan functions as a university centre. As the capital of the region, Poznan is a city of almost 160 thousand students of different forms of studies, including nearly 90 thousand of regular daily students (in the 2011/2012 academic year, there were 88 349 regular daily students in Poznan, CSO 2013). These young people usually are not registered as Poznan citizens. But after graduation they usually decide to stay in the city and take a job there. In the 2011 census a full analogy was observed. Estimates based on the projection for the age group of 20-27 years were much lower than census data, as they refer to indigenous inhabitants of the city, who nine years earlier were 11-18 years old, and they do not include students who came to Poznan from other places.

Current findings seem to be consistent with the analysis conducted for previous census data. But it is worth noting that a distinct change in the relationship between census data and projection estimates was observed for the age of 29 years. It should be emphasized at this point that the projection was made on the basis of the census carried out by traditional method in 2002. This means that the increase in data was due to the number of students studying in Poznan nine years earlier. In 2011 census, the estimates referred to data from the registry, which did not include students. Thus, the drop observed for the age of 29 indicates the 'loss' of the population that had studied at the universities nine years earlier. The deficiencies in the register-based 2011 census suggest that students studying in the city nine years earlier had not decided to stay after graduation. Of course, we do not know whether they returned to their place of residence prior to the studies, or emigrated abroad. Nor do we know whether they were people who were successful in their professional career in Poznan or surroundings, and decided to live in suburban areas rather than in the city centre. The deficiencies of 2-3 thousand people in each of the subsequent year of age show that the city was not able to keep the potential of young and educated people. All in all, this is a group of about 20 thousand people, that is 4% of the city population at the age of the most intense economic, matrimonial and reproductive activities.

This analysis may also indicate another problem. It is the decreasing number of city residents, which is not only related to foreign emigration, but also to the process of suburbanization. Within a radius of 20-25 km around the city, new settlements are created and inhabited mostly by young, educated people, who after graduation and marriage, change a student’s flat in the city for a house near the city (Klimanek, 2012). At the moment, the problem of suburbanization is becoming increasingly important for the development of the city and its surroundings.

4. Conclusion

The preparation and implementation of a new census methodology require time. It is very important to emphasize here the introduction of relevant legislative regulations and the process for reviewing and improving the quality of administrative records. Extensive work on the evaluation of the quality of administrative records and their use by public statistics should be considered as a great achievement of the 2011 census in Poland. Of course, the process of evaluation and improvement of records is a continuous one, and work on this has barely started.

The above discussion is an attempt made to evaluate the results of the census, and also the population register indirectly. Using the methods of demographic analysis as concerns fertility, mortality, migration and projections, a comparative analysis was conducted. Different cross-sections according to sex, age and place of residence were considered. In each of the separate domains, adequacy and divergence assessments were provided and accompanied by substantive explanations. Among the results obtained, we can specify:

The survival rates for young people (age groups: 9-13, 16-19, 30 years and 33-35 years) between censuses were exceeding one.

This situation may suggest that not all the people were enumerated in 2002 census. In particular, the lack of infants in the 2002 census was already observed in earlier studies (Paradysz 2010). The deficiencies observed for people aged 30-35 years may reflect emigrants who were not counted in the census conducted by the conventional method, but were included in the register-based census (as they are not removed from the register). The population register is not free from erroneous enumerations, repetitions and omissions and others. Thus, further work on its improvement is necessary.

For better identification of the usually resident population, it should be considered to extend the residence analysis. Additional methods of assessing compliance with the definition criteria might be introduced, for example by examining activity of individuals in other registers to ensure that only individuals with a pre-defined number of "signs of life" are counted in the census.

1. The lack of children aged 0 completed years observed in the census was most likely due to a delay in the registry. A two-week delay corresponds to about 1/3 of the monthly number of births and the missing number of infants - which is consistent with the statutorily specified time to register one's child. This shortcoming requires adequate solutions.
2. The surplus of children aged 1-4 years observed in the census can be explained by children born in exile and registered in Poland to obtain Polish citizenship. This is facilitated by a simple registration procedure. This was partially confirmed by mirror statistical analysis for England and Wales.
3. Trends characterizing changes in population aged 20-30-40 years could be explained through an in-depth comparative analysis of data from successive

censuses and current population statistics. Diversity of data sources and the principles of demographic analysis allowed a discussion on the 'surplus' of population aged 20-29 years and the 'shortage' of the population aged 29-40 years in the cross-section of large cities and surrounding areas.

A quality assessment system is built into procedures of conducting a census, which is a survey that uses various data sources. The implementation of the mechanism for mutual control, research compliance and comparative analyses results in more reliable information. The use of variety of sources promotes their in-depth exploration also in terms of demographic analyses. On the other hand, the use of multiple sources of information makes it possible to obtain inconsistent results being a 'natural' danger. In consequence, divergent estimates force attempts to provide consistent estimates and explain reasons of the differences. The 2011 census was a complex procedure, which for a single individual combined information from two different types of sources: registers and sample surveys. The analysis using integrated data (register and sample survey) requires the development of new theoretical concepts (Zhang, 2011, 2012). The census based on multiple data sources enforced application of modern methodology. In the case of 2011 census in Poland it meant great scientific work related to the development of modern statistical methods such as calibration, statistical data integration, GIS, estimation for small domains, etc. The advantages of the applied methodology are not only difficult to measure and assess, but they should be considered in terms of a precondition for further development of statistics in the most desirable sense: the development of science in response to the needs.

REFERENCES

AL, P. G., BAKKER, B. F. M., (2000). Re-engineering Social Statistics by Micro-integration of Different Sources; An Introduction. "Netherlands Official Statistics". Volume 15. Summer 2000. Special issue: Integrating Administrative Registers and Household Surveys. Statistics Netherlands.

BAFFOUR, B., KING, T., VALENTE, P., (2013). The Modern Census: Evolution, Examples and Evaluation. "International Statistical Review" (2013), 81, 3, pp. 407–425, DOI: 10.1111/insr.12036

BALDRIGE, M., BROWN, C. J., JONES, S., KEANE, J. G., (1985). Evaluating Censuses of Population and Housing. Department of Commerce. United States of America/US Bureau of the Census.

BARWIŃSKI, M., (2014). Ethnic Structure of Polish Population in the Light of the Results of 2011 Census [Struktura narodowościowa Polski w świetle wyników spisu powszechnego z 2011 roku]. „Przegląd Geograficzny", 86 (2), pp. 217–241.

BERKA, C., HUMER, S., LENK, M., MOSER, M., RECHTA, H., SCHWERER, E., (2010). A quality framework for statistics based on administrative data sources using the example of the Austrian census 2011, "Austrian Journal of Statistics", No. 39.

BERKA, C., HUMER, S., LENK, M., MOSER, M., RECHTA, H., SCHWERER, E., (2012). Combination of evidence from multiple administrative data sources: quality assessment of the Austrian register-based census 2011. "Statistica Neerlandica", Vol. 66, Issue 1, 18–33.

BETHLEHEM, J., (2009). The Rise of Survey Sampling, Discussion paper (09015), Statistics Netherlands.

BORKET, M., PENNINX, R., (2011). Policymaking in the Field of Migration and Integration in Europe: An Introduction. [in:] G. Zincone, R. Penninx, M. Borkert (ed.), Migration Policymaking in Europe. The Dynamics of Actors and Contexts in Past and Present, Amsterdam: IMISCOE Research. Amsterdam University Press, pp. 7–21.

BROWN, J., ABBOTT, O., DIAMOND, I., (2006). Dependence in the 2001 one-number census project. "Journal of Royal Statistical Society", A, 169, pp. 883–902.

CBOS, (2013). Was It Worth It to Change the System? Assessment of Social Change After 1989. [Czy warto było zmieniać ustrój? Społeczna ocena przemian po 1989 roku]. Statement from research, BS/73/2013.

CHATFIELD, C., (2002). Confession of a Pragmatic Statistician. "The Statistician", Vol. 51, No. 1. (2002), pp. 1–20.

CSO, (2012). The Report on the Results. The National Census of Population and Housing 2011 [Raport z wyników. Narodowy Spis Powszechny Ludności I Mieszkań 2011], Warsaw.

CSO, (2013). Higher Education Institutions and their Finances in 2012 [Szkoły wyższe i ich finanse w 2012 roku], Warsaw.

FIHEL, A., KACZMARCZYK, P., OKÓLSKI, M., (2006). Labour Mobility in the Enlarged European Union. CMR Working Paper, No. 14/72.

FOSU, G. B., (2001). Evaluation of population census data through demographic analysis. Symposium on Global Review of 2000 Round of Population and Housing Censuses: Mid-Decade Assessment and Future Prospects Statistics Division Department of Economic and Social Affairs United Nations Secretariat New York, 7-10 August 2001.

GOŁATA, E., (2009). Economic Activity in Population Census 2011 and Administration Resources [Aktywność ekonomiczna ludności w NSP'2011 a zasoby rejestrów administracyjnych] [in:] Methods and Sources of Information in the public statistics [Metody i Źródła Pozyskiwania Informacji w Statystyce Publicznej], Published by: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznan.

GOŁATA, E., (2012). Population Census And Truth (Spis ludności i prawda). „Studia Demograficzne". Polish Academy of Sciences, 161 (1), pp. 23–55.

HOLT, T., (2007). The Official Statistics Olympic Challenge: Wider, Deeper, Quicker, Better, Cheaper. Discussion. "The American Statistician", Vol. 61 (1), pp. 1–8.

JANTA, B., (2013). Polish Migrants' Reproductive Behaviour in the United Kingdom. "Studia migracyjne – Przegląd Polonijny", No. 3, 63–96.

JOŃCZY, R., (2010). International Migration From Rural Areas of Opolskie Province After the Polish Accession to the European Union. Selected Economic and Demographic Aspects [Migracje zagraniczne z obszarów wiejskich województwa opolskiego po akcesji Polski do Unii Europejskiej. Wybrane aspekty ekonomiczne i demograficzne]. Published by: Wydawnictwo Instytut Śląski Sp. z o.o. Opole–Wroclaw.

JÓZEFOWSKI, T., RYNARZEWSKA-PIETRZAK, B., (2010). Evaluation of the Possibilities to Use the PESEL Register in the Population Census [Ocena możliwości wykorzystania rejestru PESEL w spisie ludności]. [in:] The Measurement and Information in the Economy [Pomiar i informacja w gospodarce]. Scientific Notebook of the Faculty of Informatics and Electronic Economy. E. Gołata (ed.). Published by: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznan.

KICINGER, A., KORYŚ, I., (2011). A new Field to Conquer and Manage: Migration Policymaking in Poland after 1989, [in:] G. Zincone, R. Penninx, M. Borkert (ed.), Migration Policymaking in Europe. The Dynamics of Actors and Contexts in Past and Present, Amsterdam: IMISCOE Research. Amsterdam University Press, pp. 347–376.

KLIMANEK, T., (2012). The Results of the Study of Migratory Behavior of the Inhabitants of Poznan [Wyniki badania zachowań migracyjnych mieszkańców Poznania], [in:] Migration of the Residents of Large Cities [Migracje mieszkańców dużych miast], E. Gołata (ed.). Published by: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznan.

KOBUS, P., SMOLKA, M., NOWAKOWSKA, G., (2009). Report on the Works on the Files of Comprehensive Social Insurance System [Raport z prac na zbiorach kompleksowego systemu informatycznego ZUS], CSO: Central Bureau of Census [Centralne Biuro Spisów GUS], Warsaw.

KORDOS, J., (2007). Some Aspects of Post-Enumeration Surveys in Poland. „Statistics in Transition – new series", Vol. 8 (3). pp. 563–576.

KORDOS, J., (2012). The Interplay Between Sample Survey Theory And Practice In Poland [Współzależność pomiędzy rozwojem teorii i praktyki badań reprezentacyjnych w Polsce] „Przegląd Statystyczny" Special Issue 1 – 2012. pp. 61–68. Polish Academy of Sciences.

KUC, M., (2014). Najstarszy człowiek świata żyje w Polsce. W bazie PESEL. [The oldest man in the world lives in Poland. In the PESEL database.], "Gazeta Prawna". 4.06.2014. Warsaw. http://prawo.gazetaprawna.pl/artykuly/801115,najstarszy-czlowiek-swiata-zyje-w-polsce-w-bazie-pesel.html.

LONGVA, THOMSEN, SEVEREIDE, (1998). Reducing Costs of Censuses in Norway Through Use of Administrative Registers. Statistics Norway 1998.

MRYSKALA, P., (2011). Use of Registers and Administrative Data Sources for Statistical Purposes. Best Practices of Statistics Finland. A Handbook of Principles and Practices. Statistics Finland.

NEWELL, R., SMALLWOOD, S., (2010). A cross country review of the validation and/or adjustment of census data. Population Trends, No. 141, Office for National Statistics.

ONS, (2013). Detailed Country of Birth and Nationality Analysis from the 2011 Census of England and Wales.

PARADYSZ, J., (2002). The Non-random Errors in the Fertility Survey within the 1970 Population Census (O błędach nielosowych w badaniu dzietności kobiet w ramach Narodowego Spisu Powszechnego 1970). [in:] Population censuses in the Republic of Poland from 1921 to 2002. Selected Writings by Demographers [Spisy ludności Rzeczypospolitej Polskiej 1921–2002. Wybór pism demografów]. Z. Strzelecki. T. Toczyński (eds.). Polish Demographic Society. Central Statistical Office [Polskie Towarzystwo Demograficzne. Główny Urząd Statystyczny], Warsaw, pp. 479–482.

PARADYSZ, J., (2010). The Necessity for Use of Indirect Estimation in Population Censuses [Konieczność estymacji pośredniej na użytek spisów powszechnych]. [in:] The Measurement and Information in the Economy [Pomiar i informacja w gospodarce]. Scientific Notebook of the Faculty of Informatics and Electronic Economy. E. Gołata (ed.). Published by: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznan.

PLEWIS, I., SIMPSON, L., WILLIAMSON, P. (2011). Census 2011: Independent review of coverage assessment, adjustment and quality assurance. Available at: www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/the-2011-censusproject/independent-assessments/independent-review-of-coverage-assessment-- adjustment-and-quality-assurance/independent-review-final-report.pdf

REGULATION (EU), No 1260/2013 of the European Parliament and of the Council of 20 November 2013 on European demographic statistics Text with EEA relevance. Available at: http://eur-lex.europa.eu/legal content/EN/TXT/?uri=CELEX:32013R1260.

REPORT, (2011). The National Census of Population and Housing 2011 in Assessment of the Ukrainian Minority [Raport. Narodowy Spis Powszechny Ludności i Mieszkań 2011 w ocenie mniejszości ukraińskiej]. 2011. Association of Ukrainians in Poland, Warsaw.

ROSZKA, W., (2013). Statistical Data Integration in Socio-Economic Research [Statystyczna integracja danych w badaniach społeczno-ekonomicznych]. PhD Thesis. Poznan University of Economics. Available at: http://www.wbc.poznan.pl/Content/265243/Roszka_Wojciech_doktorat.pdf.

RUGGLES, S., (2014). Big Microdata for Population Research. "Demography" 51, 287–297.

SAKSON, B., (2002). The impact of "invisible" migration of the eighties on the demographic structure of Polish Population [Wpływ "niewidzialnych" migracji zagranicznych lat osiemdziesiątych na struktury demograficzne Polski]. Monographs and Studies, No. 481, Warsaw School of Economics. Warsaw.

ŚLESZYŃSKI, P., (2004). Regional Disparities Between Population According to the 2002 Population Census and Based on the Current Recording [Regionalne różnice pomiędzy liczbą ludności według narodowego spisu powszechnego w 2002 r. i rejestrowaną na podstawie ewidencji bieżącej], „Studia Demograficzne", 145 (1), pp. 93–103.

ŚLESZYŃSKI, P., (2005). Differences in the Numer of Population Disclosed in the National Census 2002 [Różnice liczby ludności ujawnione w Narodowym Spisie Powszechnym 2002]. „Przegląd Geograficzny", 77 (2), pp.193–212.

Statistics Denmark, (1995). Statistics on Persons in Denmark – a register-based statistical system. Statistics Denmark.

STATISTICS FINLAND, (2004). Use of Registers and Administrative Data Sources for Statistical Purposes. Best Practices of Statistics Finland. Tilastokeskus. Statistikcentralen. Statistics Finland. Helsinki.

STATISTICS NETHERLAND, (2004). The Dutch Virtual Census of 2001. Analysis and Methodology. Ed. Eric Schulte Nordholt, Marijke Hartgers and Rita Gircour. Statistics Netherlands, Voorburg/Heerlen, 2004.

UN, (2006). Conference of European Statisticians Recommendations for the 2010 Censuses of Population and Housing. 2006. New York. Geneva. United Nations Economic Commissions for Europe. Statistical Office of the European Communities. ECE/CES/STAT/NONE/2006/4. UN Publications.

UN, (2007). Register-based Statistics in the Nordic Countries. Review of Best Practices with Focus on Population and Social Statistics. United Nations Economic Commission For Europe. United Nations. New York, Geneva, 2007.

UN, (2009). Manual on Census Evaluation. Post Enumeration Surveys. Demographic And Social Statistics Branch, United Nations Statistics Division.

UN, (2010). Post Enumeration Surveys Operational Guidelines. Technical Report. United Nations Statistics Division New York, April 2010.

UN, (2010a). Report on the Results of a Survey on Census Methods used by Countries in the 2010 Census Round. Department of Economic and Social Affairs United Nations. New York, Available at: http://unstats.un.org/unsd/census2010.htm.

UN, (2011). Using Administrative and Secondary Sources for Official Statistics. United Nations Economic Commission For Europe. United Nations. New York, Geneva, 2011.

UN, (2012). Overview of the 2010 round of population and housing censuses in the UNECE region. ECE/CES/ GE.41/2012/20, 16 May 2012.

UN, (2012a). Lessons learned from the Population and Housing Census in Poland. Note by the Central Statistical Office of Poland. Economic Commission for Europe.

UN, (2012b). Quality assessment of register-based census data in Austria, ECE/CES. Working Paper 16, 16 May 2012.

UN, (2013). Overview of National Experiences for Population and Housing Censuses of the 2010 Round. United Nations Statistics Division, 2013, New York June 2013.

UN, (2013a). Field operations. Legislation. Lessons Learned: Key Results of the UNECE Survey on National Census Practices and First Proposals About the CES Recommendations for the 2020 Census Round. Note by the UNECE Steering Group on Population and Housing Censuses. Geneva. 30 September – 3 October 2013. ECE/CES/GE.41/2013/6.

UN, (2013b). Census quality and coverage: Key results of the UNECE Survey on National Census Practices. Note by the UNECE Task Force on census coverage and quality. ECE/CES/ GE.41/2013/8. Geneva. 30 September – 3 October 2013.

WALLER, L., BERRINGTON A., RAYMER J., (2014). New Insights into the Fertility Patterns of Recent Polish Migrants in the United Kingdom. "Journal of Population Research". DOI: 10.1007/s12546-014-9125-5.

WALLGREN, A., WALLGREN, B., (2007, 2014). Register-based Statistics. Statistical Methods for Administrative Data. John Wiley & Sons. Ltd.

WALLGREN, A., WALLGREN, B., (2013). Quality Assessment in Systems with Registers and Sample Surveys. http://www.statistics.gov.hk/wsc/IPS078-P2-S.pdf.

VALENTE, P., (2010). Census taking in Europe: how are populations counted in 2010? "Population & Societies", No. 467.

VAN DER LAAN, P., (2000). Integrating administrative registers and household surveys, "Netherlands Official Statistics", Vol. 15, Summer 2000, Special issue: Integrating administrative registers and household surveys, Statistics Netherlands, Voorburg/Heerlen.

ZASĘPA, R., (1993). Use of Sampling Methods in Population Censuses in Poland, "Statistics in Transition", Vol. 1, Number 1, June 1993.

ZHANG, L.-C., (2011). A Unit-Error Theory for Register-Based Household Statistics. „Journal of Official Statistics", Vol. 27(3), pp. 415–432.

ZHANG, L.-C., (2012). Topics of Statistical Theory for Register-based Statistics and Data Integration. Statistica Neerlandica, Vol. 66, No. 1. pp. 41–63.

ZHANG, L.-C., (2013). Population Size Estimation Based on Multiple Lists. Uncertainty Analysis for Categorical Data Fusion. Open Lecture given at Poznan University of Economics.

ZUMPE, J., DORMON, O., JEFFERIES, J., (2012). Childbearing Among UK Born and Non-UK Born Women Living in the UK, ONS, 25 October 2012.