STATISTICS IN TRANSITION-new series, August 2011
Vol. 12, No. 1, pp. 81—96

BOOTSTRAP METHOD WITH CALIBRATION FOR STANDARD ERROR ESTIMATORS OF INCOME POVERTY MEASURES

Agnieszka Zięba-Pietrzak[1], Jan Kordos[2], Robert Wieczorkowski[3]

ABSTRACT

The authors begin with calibration approach in sample surveys, focussing on the Eurostat approach. Next, the indicators of poverty and social exclusion are discussed as an essential tool for monitoring progress in the reduction of these problems. Most of these indicators are calculated according to the Eurostat recommendations, using data from European Statistics on Income and Living Conditions (EU-SILC). Complex sample design of the EU-SILC requires weighted analyses for estimates of population parameters and approximate methods of standard error estimation. In our study McCarthy and Snowden (1985) bootstrap method for standard errors estimation of income poverty measures is presented. In the next step the reweighting of bootstrap weights is applied and results of such calibration are discussed.

Key words: Auxiliary information; Bootstrap, Calibration, Complex sample surveys, EU-SILC, Poverty indicators; Weighting.

1. Introduction

Calibration provides a systematic way to incorporate auxiliary information in the estimation procedure. Calibration has established itself as an important methodological instrument in large-scale sample surveys. Several national statistical agencies have developed software designed to compute weights, usually calibrated to auxiliary information available in administrative registers and other accurate sources. In this paper calibration approach has been applied to estimation of standard errors of poverty measures, using calibration approach accepted by Eurostat (2002). This approach is generally presented later.

Standard error estimation indicates precision of estimators of unknown parameters and it is one of the most important issues in statistical inference.

[1] WSE. E-mail: a.zieba.pietrzak@gmail.com.
[2] WSE. E-mail: jan1kor2@aster.pl.
[3] GUS. E-mail: R.Wieczorkowski@stat.gov.pl.

Derivation of suitable estimator of the variance of the given statistic is difficult in case of non-linear structure of estimated measures and complex survey design (Sitter, 1992). Income poverty indicators are such kind of measures and are calculated according to the Eurostat recommendations, using data from European Statistics on Income and Living Conditions EU-SILC (Eurostat, 2009). Using this survey it is necessary to take into consideration that the selection of the survey sample is done by two-stage stratified sampling with different selection probabilities at the first stage. In this case it is not possible to obtain a closed-form algebraic expression for the estimated standard error. For such measures approximate standard error estimation is required. Various methods for obtaining variance estimates have been proposed in the literature (Kordos, Zięba-Pietrzak, 2010). One of the most common methods is bootstrap – a method, which has become an important tool in sample surveys since first research article about it (Efron, 1979).

Bootstrap methodology for sample surveys application has been extended for stratified sampling, involving a large number of strata with relatively few primary sampling units within strata when the parameter of interest is a nonlinear function (Rao, Wu, 1984). For stratified multistage designs McCarthy and Snowden proposed procedure which is an asymptotically valid method in assessing the variability of direct estimators (McCarthy, Snowden, 1985). This procedure has been applied in Polish official statistics since 2003 in the Labour Force Survey (Popiński, 2006). It is also proper for standard error estimation for poverty measures and it has been applied in the Polish EU-SILC survey since 2005 (GUS, 2009; Zieba & Kordos, 2010).

In this study McCarthy and Snowden (1985) bootstrap method for standard error estimators of income poverty measures is presented. In the next step the reweighting of bootstrap weights is applied and results of this integrated calibration method are discussed.

2. Income poverty measures

Poverty indicators used to monitor social exclusion phenomena are calculated on the basis of data from EU-SILC. Basic aim of this survey is the study, both at national and European level, of households' living conditions mainly in relation to their income. Four most common poverty measures and two additional are being focused on in this work. All of them are calculated using *equivalised income distribution* and therefore are called *income poverty measures* (Eurostat, 2005).

Equivalised disposable income (EQ_INC*i*) is defined as the household's total disposable income divided by its "equivalent size" to take account of the size and composition of the household, and attributed to each household member (Eurostat 2005). The equivalised household size is defined according to the modified OECD (Organisation for Economic Co-operation and Development) scale which gives a weight of 1.0 to the first adult, 0.5 to other household members aged 14 or over and 0.3 to household members aged less than 14. Finally, equivalised household size (*EQ_SS*) is calculated as:

$$EQ\_SS = 1 + 0{,}5 \cdot (HM_{14+} - 1) + 0{,}3 \cdot HM_{13-} \qquad (1)$$

where:

HM_{14+} – number of household members aged 14 and over
HM_{13-} – number of household members aged 13 or less.

The equivalence scales are the parameters which allow comparing the conditions of households of different sizes and different demographic structures.

The formulas for selected poverty indicators

1) MED (MED_{EQINC}) – *national median equivalised income* – it is a median in distribution of yearly equivalent net disposable income in households.
2) ARPR – *At-risk-of-poverty rate after social transfers* is defined as a share of poor persons in all persons. A poor person is defined as the one with an equivalised disposable income below the poverty threshold (ARPT)

$$ARPR = \frac{\sum_{i|EQINC_i \le ARPT} w_i}{\sum_{i=1}^{n} w_i} \qquad (2)$$

where:
w_i – the survey weight attached to i^{th} sample unit (individual),
ARPT – At-risk-of-poverty threshold after social transfers.

The poverty threshold ARPT equals 60% of the national median equivalised disposable income

$$ARPT = 0{,}6 \cdot MED_{EQINC} \qquad (3)$$

Disposable income is defined as a sum of net (after social transfers) monetary incomes gained by all the household members reduced by: property tax, inter-household cash transfers paid and statements for the Treasury Office.

3) RRPG – *Relative median poverty risk gap* is defined as a difference between the median equivalised disposable income of persons below the at-risk-of poverty threshold (MED_{POOR}) and the threshold itself (ARPT), expressed as a percentage of the at-risk-of poverty threshold:

4)

$$RRPG = \frac{ARPT - MED_{POOR}}{ARPT} \tag{4}$$

5) GINI – *Gini coefficient* measures inequality of income distribution and this is the relationship between cumulative shares of the population arranged according to the level of income and the cumulative share of the equivalised total income received by them.

To calculate this indicator persons have to be sorted according to EQINC from the lowest to the highest value. If w_i denote the sample weight of a unit *i-th sample unit* in the ascending income sorted distribution, then we have:

$$GINI = 100 \cdot \left(\frac{2 \cdot \sum_{i=first\_person}^{last\_person} \left(w_i \cdot EQINC_i \cdot \sum_{first\_person}^{person\_i} w_i \right) - \sum_{i=first\_person}^{last\_person} \left(w_i^2 \cdot EQINC_i \right)}{\left(\sum_{i=first\_person}^{last\_person} w_i \right) \cdot \sum_{i=first\_person}^{last\_person} \left(w_i \cdot EQINC_i \right)} - 1 \right) \tag{5}$$

where:

w_i – the survey weight attached to *i-th sample unit* (individual).

6) S80/S20 – *S80/S20 income quintile share ratio* measures inequality of income distribution – it is the ratio of sum of equivalised disposable income received by the 80% of the country's population with the highest income (top quintile) to that received by the 20% of the country's population with the lowest income (lowest quintile):

7)

$$S80/S20 = \frac{\sum_{i|EQINC_i \le Q_{80}} w_i EQINC_i}{\sum_{i|EQINC_i \le Q_{20}} w_i EQINC_i} \tag{6}$$

where:
w_i – the survey weight attached to *i-th sample unit* (individual).

Due to the way quintiles are calculated, the number of people with $EQINC_i \leq Q_{20}$ can differ from the number of people with $EQINC_i \geq Q_{80}$, in which case the following correction is necessary:

$$S80/S20 = \frac{\sum_{i|EQINC_i \geq Q_{80}} w_i EQINC_i \Big/ \sum_{i|EQINC_i \geq Q_{80}} w_i}{\sum_{i|EQINC_i \leq Q_{20}} w_i EQINC_i \Big/ \sum_{i|EQINC_i \leq Q_{20}} w_i} \tag{7}$$

8) MEAN_EQINC – *mean equivalised income* – it is an average yearly equivalent net disposable income in households.

9)

$$MEAN\_EQINC = \frac{\sum_{i=1}^{n} wiEQINC_i}{\sum_{i=1}^{n} wi} \tag{8}$$

where:

w_j – the survey weight attached to *i-th sample unit* (individual).

3. Sample design in the Polish EU-SILC

The Polish EU-SILC sample is two-stage stratified sampling with unequal selection probabilities at the first stage. The first-stage sampling units (primary sampling units – PSU) are enumeration census areas, while at the second stage dwellings are selected. All the households from the selected dwellings are supposed to enter the survey. The basis for creating strata is the size of a locality (measured by the number of dwellings), and strata are built up within a voivodship. A voivodship is NUTS2 level – a territorial system of classification worked out by Eurostat. According to the sample design, proper weights are constructed. Every statistic and its variability are estimated using these weights.

The selection of the EU-SILC sample is done by two-stage stratified sampling with unequal selection probabilities at the first stage. The first-stage sampling units (primary sampling units – PSU) are enumeration census areas, while at the second stage dwellings are selected. All the households from the selected dwellings are supposed to enter the survey. The basis for creating strata is the size of a locality (measured by the number of dwellers), and strata are built up within a voivodship. A voivodship is NUTS2 level at The Nomenclature of Territorial Units for Statistics (NUTS) – a territorial system of classification worked out by Eurostat. There are 16 NUTS2 units in Poland.

In EU-SILC carried out by Central Statistical Office of Poland in 2008 the final sample contained 13 984 households with 41 200 persons (33 801 persons aged 16 and more were interviewed) from 5093 PSUs (primary sampling units) allocated in 211 strata.

Weights construction in Polish EU-SILC 2008

According to EU-SILC sample, structure proper weights are calculated. Final weights, DB090 (for households) and RB050 (for household members), are constructed in complex and multi-stage process (for more detail see GUS, 2009). During this process the *integrated calibration method* recommended by Eurostat is applied (Eurostat, 2004). It means that final weights are calculated with the use of demographic data from other sources. This method ensures consistency of the generalized results with the external demographic data available. Every statistic and its variability are estimated using these weights.

4. The bootstrap method (McCarthy and Snowden, 1985)

In Polish practice bootstrap method (McCarthy and Snowden, 1985) has been used in Polish official statistics since 2003 till present time (Kordos & Zieba-Pietrzak, 2010). This method makes possible treating different estimators uniformly and eliminates the necessity of derivation of complicated analytical formulas (Popiński, 2006). It is applied separately in each stratum and it is proper for stratified multistage sample design. For such design this bootstrap procedure is an asymptotically valid method in assessing the variability of direct estimators (Shao, 2003).

Applying this method, every bootstrap sample is obtained drawing with-replacement random sample of $a_h - 1$ PSUs out the a_h, sampled in each stratum h ($h = 1,2,...,L$). All items (secondary sampling units) from sampled PSUs are included in the new bootstrap sample. After every resampling, the original weights are properly rescaled:

$$w_j(b) = w_j \frac{a_h}{a_h - 1} m_j(b) \tag{9}$$

where:

a_h – number of PSUs in stratum h

$w_j(b)$ – weight for person from *j-th* household in *b-th* bootstrap sample,

w_j – original weight for person from *j-th* household,

$m_j(b)$ – number of times PSU from *j-th* household is included in b^{th} bootstrap sample ($b=1,2,...,B$).

The bootstrap variance estimate of the corresponding indicator is obtained by the usual Monte Carlo approximation based on the independent bootstrap replicates (the sampling procedure is replicated B times).

At national level variance of direct estimator is estimated in the following way:

$$V(\hat{\theta}) = \frac{1}{B-1}\sum_{b=1}^{B}(\hat{\theta}_b^* - \hat{\theta}^*)^2 \tag{10}$$

where:

$\hat{\theta}$ – the direct estimator of indicator

$\hat{\theta}_b^*$ – the bootstrap estimator obtained using modified weights from b^{th} bootstrap sample

$\hat{\theta}^*$ – is computed as: $\hat{\theta}^* = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}^{*b}$.

Finally, the estimator of standard error for direct estimator is defined as:

$$SE(\hat{\theta}) = \sqrt{V(\hat{\theta})} . \tag{11}$$

5. Bootstrap method with calibration of sampling weights

As it has been stressed at the beginning, calibration in its several forms has become an essential and popular tool in sample surveys. There are many reasons why the accuracy of the estimates should be improved by applying it. One of these reasons is that particular estimates of the calibrated data should coincide with known population totals (Gambino, 1999). In EU-SILC this issue is crucial also in case of subpopulations. The original totals, based on demographic variables such as age, sex, geographic location and household size should be estimated without error. The “sampling weights” are not enough for safeguarding this accuracy and calibration is necessary.

A mainstream practice is the use of auxiliary information (e.g. from the census) for the strengthening of the survey results (Eurostat, 2002). By applying calibration (reweighting of the sample weights) agreement with known population margins results is provided. This method is also called *post-stratification*. Appropriate calculations are made and each statistical unit (individual, household or business) is re-assigned a unique weight. Calibrated estimates are subsequently based on these weights.

5.1. Calibration algorithm used in EU-SILC

The principle of the calibration is the following: starting from an initial weighting variable, and auxiliary variables, final weighting variable is constructed, which will perfectly estimate the totals on whichever of the auxiliary variable in such a way that its weight values for each unit in the sample are as close as possible to the initial weights (Eurostat, 2004).

In basic idea of calibration it is assumed that the totals of the auxiliary variables (called calibration variables) over the whole population are known and come from sufficiently reliable external data sources such as, for instance, census.

The calibration procedure is to calculate new household weights (depending on the sample of households), called "calibrated" (or final) weights, such that two following properties are satisfied:

the "calibrated" weights are "not very different" from the initial weights (ratios between the final weights and the initial weights, called the g-weights, is "not very different" from 1),

the totals of the auxiliary variables are estimated *with zero variance* from the final weights (Eurostat, 2004).

From the mathematical point of view, a *"function of distance"* G (.) is considered. This function will control that the *g-weights* do not vary too much from 1. G is supposed to be (Deville et al., 1993):

positive,

continuously differentiable in a neighborhood of 1,

strictly convex,

- G'(*1*)=*0*.

Under these assumptions, the "calibrated" weights form (i.e. different methods) the following problem (*P*) of optimization subject to constraints:

$$(P)\begin{cases} \min \sum_{k\in S} d_k \cdot G\left(\frac{w_k}{d_k}\right) \\ \text{subject\_to}: \sum_{k\in S} w_k \cdot x_{rk} = X_r \quad for\ \ r = 1,\dots,p \end{cases} \tag{12}$$

where:

d_k – the *k-th* household design weight (before calibration),

w_k – the *k-th* household final weight (after calibration),

$x_1, x_2 \ldots x_p$ – auxiliary variables,

x_{rk} – the value of the auxiliary variable xr ($r=1...p$) on the household k,

X_r – total of *r-th* auxiliary variable in population,

S – sample of households.

In order to derive calibration weights the significant number of matrix calculations should be made and specialized software is needed. For this need the calibration algorithm was implemented and applied using SAS software (GUS, 2007). The software finds a solution in iterative way, function G can take different form (different method can be used: the linear method, the raking ratio method, the “logit” method, the linear truncated method, hyperbolic sinus, etc.) (Eurostat, 2004).

In this work the hyperbolic sinus method was applied. In practice this method yields final weights very close to initial ones. The formula of function G is as follows:

$$G_{\alpha}(x) = \frac{1}{2\alpha}\int_{1}^{x} sh\left[\alpha\left(\frac{1-t}{t}\right)\right]dt \qquad (13)$$

where:

α – parameter which is positive (it allows to control a degree of dispersion of calibrated weights in relation to original weights, its initial value equals 1)

sh(x) – the hyperbolic sinus function: $sh(x) = \frac{\exp(x) - \exp(-x)}{2}$.

5.2.Integrated calibration method

The idea of integrated calibration method, recommended by Eurostat, is to use calibration variables defined at both household and individual levels. The individual variables are aggregated at the household level by calculating household totals (such as the number of males/females in the household, the number of persons aged of 16 and over, etc.). After that, the calibration is done at the household level using the household variables and the individual variables in their aggregate form. This technique ensures consistency between households and individual estimates because (see the next step) all household members will have the same household cross-sectional weight as personal cross-sectional weight.

The weights obtained after this procedure of integrated calibration are the household cross-sectional weights (target variable DB090). The personal cross-sectional weight RB050 is equal for all the members of a given household and to the household cross-sectional weight DB090: (Eurostat, 2004).

Auxiliary information in Polish EU-SILC

In EU-SILC calibration procedure the additional variables comprised data on the number of households according to 4 size classes (1-person, 2-person, 3-person and 4 and more person household) in voivodship (NUTS2 regions) breakdown and place-of-residence (urban/rural). As regards individuals, the data were presented as the number of persons by gender and 14 age groups (under 16 years, 16-19 years, eleven 5-year groups, 75 years and over) in voivodship breakdown.

These additional variables were derived from the current demographic estimates and the 2002 Census, and were specific for each year of the survey.

Integrated calibration method in bootstrap method for standard error estimation

In bootstrap algorithm of standard error estimation procedure the construction of every resample will in general mean that the post-strata marginal totals are no longer satisfied. It means that every bootstrap sample should be reweighted. Bootstrap data may be reweighted by applying to it the usual algorithm with the original population marginal totals (Canty, Davison, 1999). The results of such calibration are presented below.

5.3. Empirical results - bootstrap method with calibration

Integrated calibration method for Poland

Every bootstrap sample was reweighted using additional variables, derived from the current demographic estimates (the same as original weights). After that standard error estimation was conducted. Table 1 presents results before and after weights calibration.

To draw a comparison a coefficient of variation was used:

$$cv = \frac{SE(\hat{\theta})}{\hat{\theta}} \tag{14}$$

where:

$SE(\hat{\theta})$ – standard error of a statistic under the given sample design,

$\hat{\theta}$ – the direct estimator of indicator.

Table 1. Comparison of coefficients of variation for Poland using bootstrap (B=1000) before and after weights calibration.

Indicator	Estimate	Without calibration		With calibrated weights	
		SE	cv (in %)	SE_calib	cv_calib (in %)
ARPR (%)	16,88	0,4	2,38	0,39	2,33
RRPG (%)	20,54	0,72	3,51	0,72	3,48
Gini (%)	32,02	0,42	1,32	0,41	1,27
S80_S20	5,11	0,1	1,93	0,1	1,88
Mean_Inc (EUR)	4936,68	42,74	0,87	39,72	0,8
Median (EUR)	4153,96	31,74	0,76	28,9	0,7

Source: Derived from data provided by GUS: European Statistics on Income and Living Conditions 2008

Comparing results one can observe that values of standard error after weights calibration are lower than before but differences are very small. It means that at country level reweighting has statistically insignificant effect for the bootstrap.

Integrated calibration method at NUTS2 level in Poland

At NUTS2 level weights calibration does not improve the outcomes – for some voivodships standard error decreases and in other cases it increases. That is why average values of coefficient of variation show very similar performance before and after reweighting for every measure (Table 2).

Table 2. Comparison of average values of estimates, standard errors and coefficient of variation calculated for six measures for Poland (16 NUTS2 units) using bootstrap (B=1000) before and after weights calibration.

Indicator		Estimate	Without calibration		With calibrated weights	
			SE	CV (in %)	SE_calib	CV_calib (in%)
ARPR (%)	MIN	12.4	0.9	7.2	0.9	6.8
	MAX	27.6	3.5	21.5	3.2	19.4
	MEAN	17.9	2.1	11.6	2.0	11.1
RRPG (%)	MIN	10.1	1.7	8.6	1.8	9.0
	MAX	24.7	9.4	39.6	8.9	44.1
	MEAN	20.0	3.1	16.1	3.1	16.5
Gini (%)	MIN	27.4	0.7	2.4	0.7	2.4
	MAX	38.5	2.2	8.0	2.7	9.9
	MEAN	30.2	1.3	4.3	1.3	4.3
S80_S20	MIN	4.1	0.2	4.0	0.2	4.0
	MAX	6.6	0.5	10.8	0.6	13.5
	MEAN	4.8	0.3	6.7	0.3	6.7
Mean_Inc (EUR)	MIN	3 997.8	92.9	1.8	83.5	1.6
	MAX	6 223.9	278.6	5.4	333.2	6.5
	MEAN	4 760.9	150.3	3.1	147.3	3.1
Median (EUR)	MIN	3 472.8	74.1	1.8	70.1	1.7
	MAX	4 618.6	228.2	5.9	221.2	5.7
	MEAN	4 089.2	137.5	3.4	131.6	3.2

Source: Derived from data provided by GUS: European Statistics on Income and Living Conditions 2008

There is no one tendency in differences between average values of coefficient of variation before and after weights calibration – in case of ARPR, Mean_Inc and Median it decreases, in case of RRPG it increases (Table 3).

Table 3. Comparison of average values of coefficient of variation for Poland (16 NUTS2 units) using bootstrap (B=1000) before and after weights calibration.

	mean_cv (%)	**mean_cv_calib (%)**	**cv-cv_calib (%)**
ARPR (%)	11.65	11.07	0.6
RRPG (%)	16.06	16.52	- 0.5
Gini (%)	4.28	4.29	- 0.0
S80_S20	6.68	6.70	- 0.0
Mean_Inc (EUR)	3.14	3.06	0.1
Median (EUR)	3.37	3.23	0.1

Source: Derived from data provided by GUS: European Statistics on Income and Living Conditions 2008 (EU-SILC), version of 01.03.2010

There are some significant differences in case of estimation at NUTS2 level – mainly in region with smaller sample size.

6. Concluding remarks

We would like to add that two important measurement concepts related to calibration are *precision* and *accuracy. Precision* refers to the variability in the parameter depending on size of sample, while *accuracy* refers to the actual amount of error that exists in calibration. All measurement processes used for calibration are subject to various sources of error. It is common practice to classify them as random or systematic errors. When measurement is repeated many times, the results will exhibit random statistical fluctuations which may or may not be significant, depending on the size of a sample. Systematic errors are offsets from the true value of a parameter and, if they are known, corrections are generally applied, eliminating their effect on the calibration. If they are not known, they can have an adverse effect on the accuracy of the calibration. High-accuracy calibrations are usually accompanied by an analysis of the sources of an error and a statement of the uncertainty of the calibration. Uncertainty indicates how much the accuracy of the calibration could be degraded as a result of the combined errors. In our case we estimated precision of the estimates.

In conclusion:

1. Reweighting (calibration of sampling weights), in our case, has not significant effect for the bootstrap estimates at country level.
2. At NUTS2 level weights calibration does not always reduce standard error estimates – in some cases standard error decreases and in other cases it increases. However, these differences are statistically not significant.
3. Calibration is time-consuming because of repeating calibration algorithm as many times as the number of bootstrap replicates (for example *b=1000*)

Acknowledgment

The presented work was done under the S.A.M.P.L.E. project (Small Area Methods for Poverty and Living Condition Estimates). This research programme was funded by the European Commission under the Seventh Framework (FP7) Programme of the European Union. (http://www.sample-project.eu/).

The authors would like to thank the Central Statistical Office of Poland for the production and provision of the survey data used, Mr. Bronisław Lednicki for sampling consultations, and Dr. Waldemar Popinski, the referee, for constructive corrections and comments.

Data sources provided by GUS come from European Statistics on Income and Living Conditions (EU-SILC 2008, version of 01.03.2010; EU-SILC 2007, version of 01.08.2009). The European Commission does not take responsibility for conclusions in this work.

REFERENCES

CANTY A.J., DAVISON A.C.,(1999), *Resampling-based Variance Estimation for Labour Force Survey*, Journal of the Royal Statistical Society. Series D (The Statistician), Vol. 48, No. 3 (1999), pp. 379-391.

DEVILLE, J.-C. and SARNDAL, C.E. and Sautory, O. (1993), *Generalized Ranking Procedures in Survey Sampling*, Journal of the American Statistical Association, Vol. 88, No. 423, pp. 1013-1020.

EFRON B., (1979) *Bootstrap methods: another look at the jackknife*, "Annals of statistics" 7, 1-26.

EUROSTAT, (2002), *Estimation Techniques in Statistics – A Methodological Note on Calibration.*

EUROSTAT, (2004), *Description of target variables: Cross-sectional and Longitudinal*, EU-SILC 065/04, pp. 31 – 36.

EUROSTAT, (2005), *Continuity of indicators between end-ECHP and start-SILC Algorithms to compute cross-sectional indicators of poverty and social inclusion adopted under the open method of coordination.*

EUROSTAT, (2009) *Methodological studies and quality assessment of EU-SILC*, Report SILC.04 15 April 2009, SAS programs for variance estimation of the measures required for Intermediate Quality Report.

GAMBINO J., (1999), *Discussion of "Issues in Weighting Household and Business Surveys"*, Proceedings of the 52th International Statistical Institute Session, Finland.

GUS, (2009), *Incomes and Living Conditions of the Population in Poland,* report from the EU-SILC survey of 2007 and 2008, Warsaw.

KORDOS J. , ZIĘBA-PIETRZAK A., (2010), *Development of standard errors estimation methods in complex household sample surveys in Poland*, Statistics in Transition –new series, Vol. 11, No.2, pp. 231–253.

MCCARTHY P. J., SNOWDEN C. B., (1985) T*he Bootstrap and Finite Population Sampling*, "Vital and Health Statistics", Series 2, no. 95, Public Health Service Publication 85-1369, U.S. Government Printing Office, Washington DC.

POPIŃSKI W.,(2006), *Development of the Polish Labour Force Survey*, Statistics in Transition, Vol. 7, No. 5, pp.. 1009-1030.

RAO J.N.K., WU C.F.J. (1984), *Bootstrap Inference for Sample Surveys*, Proceedings of the American Statistical Association Survey Research Methods Section, 106-112.

SHAO J., (2003) *Impact of the Bootstrap on Sample Surveys*, "Statististical Science" Vol. 18, No 2, 191-198.

SITTER R. R., (1992), *Comparing Three Bootstrap Methods for Survey Data*, The Canadian Journal of Statistics, Vol. 20, No. 2, 135-154.

ZIĘBA, A., KORDOS, J. (2010), Comparing Three Methods of Standard Error Estimation for Poverty Measures. In: J. Wywiał, W. Gamrot (Eds), Research in Social and Economic Surveys, Katowice, University of Economics.