Economics and Business Review

Volume 7 (21) Number 1 2021

CONTENTS

Editorial introduction
Monika Banaszewska, Michał Pilc

ARTICLES

The minimal-time growth problem and turnpike effect in the stationary Gale economy
Emil Panek

Repeated weighting in mixed-mode censuses
Marcin Szymkowiak, Kamil Wilak

Germany’s attitude towards the enlargement of the eurozone
Sebastian Płóciennik

Fostering sustainable development through the European Digital Single Market
Ewa Latoszek

Assessing the gender wage gap: Turkey in the years 2002–2019
Zeynep Aktaş Koral, Murat Anıl Mercan

Factors influencing the use of trade credit in financing Polish listed companies
Aleksandra Dulnięc, Natalia Świda
Aims and Scope
The Economics and Business Review (EBR) is a quarterly journal focusing on theoretical, empirical and applied research in the fields of Economics and Corporate and Public Finance. The Journal welcomes the submission of high quality articles dealing with micro, mezzo and macro issues well founded in modern theories. The EBR focuses on new market economies.

The manuscript is a complete text, while in the second all document information identifying the author(s) should be removed from papers to allow them to be sent to anonymous referees. The manuscript should be submitted as a complete text, containing original, unpublished work not submitted for publication elsewhere.

1. Articles submitted for publication in the Economics and Business Review (EBR) must be in English and should conform to the EBR editorial style. Manuscripts are to be typewritten in Times New Roman font, double spaced and be aligned.

2. Manuscripts intended for publication should be written in English, edited in Word in accordance with the EBR editorial style. Manuscripts should be submitted to: secretary@ebr.edu.pl

3. Manuscripts are to be typewritten in Times New Roman font, double spaced and be aligned. The EBR 2017 editorial style is based on the American Psychological Association (APA) guidelines and sent to: secretary@ebr.edu.pl

4. Papers should have an abstract of about 100-150 words, keywords and the Journal of Economic Literature (JEL) classification code (classification code will be established in the name of the poet). For more information see APA Style used in EBR guidelines.

5. Authors should clearly declare the aim(s) of the paper. Papers should be divided into numbered (in Arabic numerals) sections.

6. Pages should be numbered. Maximum size of the paper should be up to 20 pages.

7. Manuscripts should be submitted as PDF files.

8. Manuscripts should be submitted with proper attention to the formats of manuscripts. More information and advice on the suitability and formats of manuscripts can be obtained from: secretary@ebr.edu.pl

9. Manuscripts should be submitted with Quoted texts should refer to particular section numbers: e.g.: See Section 1.4.

10. Footnotes should refer to particular section numbers: e.g.: See Section 1.4. Footnotes should appear as a separate footnote to the author’s name a, b, etc and should not be included in the main list of footnotes.

Acknowledgements

Thematic Editors


Language Editor: Owen Eastaeal • IT Editor: Marcin Regula

© Copyright by Poznań University of Economics and Business, Poznań 2021

ISSN 2392-1641
e-ISSN 2450-0097

POZNAŃ UNIVERSITY OF ECONOMICS AND BUSINESS PRESS
ul. Powstańców Wielkopolskich 16, 61-895 Poznań, Poland
phone +48 61 854 31 54, +48 61 854 31 55
www.wydawnictwo.ue.poznan.pl, e-mail: wydawnictwo@ue.poznan.pl
postal address: al. Niepodległości 10, 61-875 Poznań, Poland

Printed and bound in Poland by:
Poznań University of Economics and Business Print Shop

Circulation: 215 copies
Repeated weighting in mixed-mode censuses

Marcin Szymkowiak\textsuperscript{2}, Kamil Wilak\textsuperscript{3}

Abstract: The main aim of the paper is to use the repeated weighting (RW) method on data from the National Census of Population and Housing 2011 (NCPH) and Labour Force Survey (LFS) to ensure consistency between margins of final tables derived from different statistical sources. This technique, based on different data sources, would ensure consistency between estimates in final output tables. This is the first application of the RW approach on data from official statistics in Poland. The results obtained by applying the RW method to data from the NCPH and additional surveys (e.g. LFS) may be used by Statistics Poland for the formulation of conclusions and recommendations for the upcoming census in 2021. The method may be also considered as an important step towards the production of timely and more detailed statistical information in Poland based on multi-source data infrastructure in general\textsuperscript{4}.

Keywords: repeated weighting method, calibration, Generalised Regression Estimator, data linkage, National Census of Population and Housing 2011, Labour Force Survey.

JEL codes: C13, C83, J21.

Introduction

Weighting is a statistical technique commonly used and applied in practice to compensate for nonresponse and coverage errors (Särndal & Lundström, 2005). It is also used to make weighted sample estimates conform to known population external totals. In recent years a lot of theoretical work has been done in

\textsuperscript{1} Article received 2 January 2021, accepted 18 March 2021.
\textsuperscript{2} Poznań University of Economics and Business, Institute of Informatics and Quantitative Economics, Department of Statistics, al. Niepodległości 10, 61-875 Poznań, Poland, Statistical Office in Poznań, Wojska Polskiego 27/29, 60-624 Poznań, corresponding author: m.szymkowiak@ue.poznan.pl, ORCID: 0000-0003-3432-4364.
\textsuperscript{3} Poznań University of Economics and Business, Institute of Informatics and Quantitative Economics, Department of Statistics, al. Niepodległości 10, 61-875 Poznań, Poland, Statistical Office in Poznań, Wojska Polskiego 27/29, 60-624 Poznań, kamil.wilak@ue.poznan.pl, ORCID: 0000-0002-4305-6202.
\textsuperscript{4} The views expressed in this paper are those of the authors and do not necessarily reflect the policies of Statistics Poland.
the area of weighting and there has been a rise in the use of these methods in many statistical surveys conducted by National Statistical Offices around the world (Särndal, 2007). One of the most common techniques of weighting is calibration, which is a method of adjusting initial weights in surveys based on sampling in order to estimate known population totals of all auxiliary variables perfectly (Deville & Särndal, 1992). This method can also be used in surveys as a possible way of tackling unit nonresponse, providing gains in efficiency in terms of variance when there is a strong correlation between the variable of interest and auxiliary variables (Lundström & Särndal, 1999; Kott, 2006; Kott & Chang, 2010). It is worth noting that calibration is one of many weighting methods that can be used in practice. Others include poststratification, raking, GREG weighting (Generalised Regression Estimator), logistic regression weighting, mixture approach and logit weighting. A review of the weighting method with examples can be found in Kalton and Flores-Cervantes (2003).

However, in practical applications, the calibration approach is often faced with a well-known problem: the inclusion of too many auxiliary variables may have several negative consequences (Szymkowiak, 2019). In particular, the use of too much auxiliary information may result in unstable estimates due to a large number of regression coefficients that have to be estimated from the sample. This problem is especially evident when an attempt is made to estimate a large set of tables from several data sources (sampling surveys, registers) in order to ensure that all estimates are mutually consistent, i.e. corresponding margins of any two estimated tables are equal (Boonstra, 2004). To obtain consistent estimates, Rennsen, Kroese and Willeboordse (2001) proposed a procedure of repeated weighting. This technique involves repeated application of the generalized regression estimator and generates a new set of weights for each table, which is estimated in order to achieve numerical consistency for margins of all tables. It is worth mentioning that so far the RW method has been successfully used only by Statistics Netherlands in their virtual census which is a set of integrated microdata files with coherent and detailed demographic and socio-economic data on persons, households, dwellings, jobs and benefits.

When data from different sources have to be reconciled the RW method may be considered (De Waal, Delden, & Scholtus, 2020). This statistical technique was developed for situations where information on one or several variables is available from more than a single data source (e.g. from administrative registers and sample surveys). The purpose of RW is to create appropriate weights for predefined census tables in order to ensure consistency of margins of all tables created with the use of various data sources supplying the census. This approach is in line with the growing need of linkage multiple data sources reported by National Statistical Institutions (NSIs) (Harron, Goldstein & Dibben, 2015). Nowadays data linking is an important cornerstone in the production of statistical information. Applications that use linked data are also a part of mainstream social science research (Chambers & Diniz da Silva, 2020). Thanks
to data linkage NSIs may avoid increasing the respondent burden. Moreover, another advantage of data linkage is the possibility of producing statistics with a higher degree of granularity than is typically the case in normal statistical procedure of data production (Luppes & Nielsen, 2020). It means that by combining data sources statistical agencies are able to produce more detailed and more timely statistics than using any single data source alone (Yang & Kim, 2020).

In the light of the above, the main aim of the study described in this article is to apply the RW method using data from the National Census of Population and Housing 2011 and the Labour Force Survey in order to ensure consistency of margins in the final tables. This article is organized as follows. Section 1 introduces the methodology of repeated weighting, including a brief description of the Generalised Regression Estimator which is the underlying concept of RW. Section 2 provides an overview of data from the National Census of Population and Housing 2011 and the Labour Force Survey in Poland. Finally, in Sections 3 and 4 the applications and results of the RW approach on data from official statistics in Poland are discussed. The paper ends with some concluding remarks.

1. Repeated weighting—theoretical aspects

With regard to sample surveys conducted by national statistical offices all over the world, it is assumed that each survey is carried out independently of the others. Therefore, a unique system of weights is created for each survey, which is used in the process of generalizing the results at the level of the entire population or for appropriately defined domains.

The final set of weights constructed for a given survey is used to create consistent tables but only within that particular survey. The situation becomes more complicated when multidimensional tables are created containing variables from more than one survey. Because of different systems of weights the final tables obtained from different surveys may not be consistent. This issue is particularly undesirable when one considers expectations of users of statistical data where it would be difficult to accept the situation of having different estimates of margins in the final tables for a variable defined in the same way in two different surveys relating to the same reference period.

Recognizing the problem outlined above, Statistics Netherlands implemented an estimation strategy alternative to calibration, referred to as the repeated weighting method, which aims to ensure consistent estimates for tables constructed with the use of data from various sources. This method has been described in several papers but its basic concept was formulated in Kroese and Renssen (1999). The method has been later developed theoretically and its practical applications have been described in Boonstra, van den Brakel, Knottnerus, Nieuwenbroek and Renssen (2003), Houbiers, Knottnerus, Kroese, Renssen

The main purpose of the repeated weighting method is to obtain a set of consistent tables when the counts are estimated using data from various sources (e.g. registers and selected sample surveys). Consistency of all tables means that margins for the same variables that can occur in at least two data sources used in the analysis are perfectly equal.

The repeated weighting method is based on the following four assumptions (Knottnerus & van Duin, 2006):
1) the reference period of the registers and surveys is the same;
2) the registers and surveys refer to the same population;
3) variables with the same name have the same definition for all relevant registers and surveys;
4) the categorical variables have hierarchical classifications, i.e. each class of a more detailed classification is nested within one class of a less detailed classification.

The repeated weighting method proposed by Statistics Netherlands consists of three main steps which can be synthetically described by the following algorithm.

In order to describe in detail this algorithm of the repeated weighting method a scenario with three data sources will be considered: a full enumeration (FE) and two sample surveys (SS1, SS2) (see Figure 1).

In Figure 1 block $S_0$ corresponds to the full enumeration FE which covers the entire population and include variables $A$, $B$, $C$. Block $S_1$ consists of respondents from the sample survey SS1 and includes variables $A$, $B$, $C$, $D$ from both FE and SS1. Similarly block $S_2$ includes respondents from the sample survey SS2 and variables $A$, $B$, $C$, $E$ from FE and SS2. Block $S_3$ contains units that were covered simultaneously by both sample surveys SS1 and SS2 and contain variables $A$, $B$, $C$, $D$, $E$ from all three surveys FE, SS1 and SS2.

As mentioned in Algorithm 1, the first step involves specifying the output tables in which appropriate counts are subject to the estimation process. Then, for the output tables defined in this way, their margins are added. The margins in a table are obtained by (i) aggregating the counts of one or more categorical variable of a multivariate table or by (ii) using less defined variants of the categorical variable. For example, two-dimensional table $D \times E$ contains margins for table $C \times D \times E$, but also for table $D \times E^{(2)}$, where variable $E^{(2)}$ has fewer levels than variable $E$ and levels of variable $E^{(2)}$ are obtained by aggregating levels of variable $E$. The tables are then ordered in the estimation process in such a way that the table with margins always precedes the more complex one.
Algorithm 1. Repeated weighting method

Step 1: Specification of output tables
Define all final tables and degree of nesting for individual variables.

Step 2: GREG estimation of the tables
Estimate the counts in tables using the generalized linear regression estimator and the most appropriate data set.

Step 3: Recalibration (the reweighting step)
Ensure numerical consistency between margins for different tables i.e. make sure that the margins of the reweighted table are consistent with their estimates from a preceding table or their known counts from a register.

Figure 1. An illustrative set of surveys and the construction of blocks
Source: Based on (Knottnerus and van Duin, 2006).

In the second step of Algorithm 1 the counts in each specified table are estimated using the generalized linear regression estimator (GREG)\(^5\) and the most appropriate data set, which is called a block and is denoted by \(S\). In most cases it will be the biggest survey in terms of sample size or a combination of surveys containing the variables required to derive the output variable. For example, if one wants to construct a two-dimensional table including the counts for levels of variables \(C\) and \(D\), the \(S_1\) block should be used, as presented in Figure 1.

\(^5\) Other calibration estimators based on different distance functions may be applied. Examples of such distance functions may be found in Deville and Särndal (1992), Haziza and Lesage (2016) or Wu and Lu (2016).
At this point, it should be emphasized that if all variables used in the process of constructing a table come from the register (in the case of the example shown in Figure 1, say, table $A \times B \times C$) then there is no need to estimate the appropriate counts. In this situation the table is built by simply summing up the counts for levels of individual variables in the register.

Let $x_k = (x_{k1}, \ldots, x_{kJ})^T$ denote a vector composed of values of auxiliary variables for the $k$-th unit, where $J$ denotes the number of auxiliary variables, and $k$ denotes units in the appropriate $S$ block, i.e. $k \in S$. Moreover let $\tau_x = \sum_{k \in U} x_k$ denote a vector composed of totals of all auxiliary variables, and $\tau_y = \sum_{k \in U} y_k$ – a vector composed of the numbers in the appropriate cells of the multidimensional table corresponding to a categorical variable $Y$ that has $P$ mutually exclusive levels. The vector $y_k$ corresponding to $\tau_y$ can then be perceived as a $P$-dimensional vector consisting of all zeros except for one unit value denoting the fact of belonging to the appropriate class of the $k$ element in relation to the categorical variable $Y$. Assuming the above notation the estimator of $\tau_y$ can be defined as follows:

\[
\hat{\tau}_y^{GREG(S)} = \hat{\tau}_y^{HT(S)} + \hat{\beta}_S^T (\tau_x - \hat{\tau}_x^{HT(S)}) = \sum_{k \in S} d^{(S)}_k y_k + \hat{\beta}_S^T \left( \tau_x - \sum_{k \in S} d^{(S)}_k x_k \right),
\]

where:

\[
\hat{\beta}_S = \left( \sum_{k \in S} d^{(S)}_k x_k x_k^T \right)^{-1} \sum_{k \in S} d^{(S)}_k x_k y_k^T,
\]

and $\hat{\tau}_y^{HT(S)}$ and $\hat{\tau}_x^{HT(S)}$ denote estimates obtained by using a well-known Horvitz-Thompson (HT) estimator for data from the $S$ block.

The weights $d^{(S)}_k$ in formulas (1) and (2) are selected because of the construction of the $S$ block. Given this setup (see Figure 1), two situations are possible:

1. $S$ is the same as the sample selected in one of the sample surveys: in this case, the weights from this survey are used as weights $d^{(S)}_k$;
2. $S$ is the common part of two blocks $S_1, S_2$ corresponding to two sample surveys, i.e. $S = S_1 \cap S_2$: then the weights are constructed as the product of the weights from both surveys:

\[
d^{(S)}_k = d^{(S_1)}_k d^{(S_2)}_k,
\]

which corresponds to the inverse of the probability that the $k$-th unit was selected for both surveys.

\[\text{These can be the original weights but, in practice, they tend to be calibration weights.}\]
The generalized regression estimator of the form (1), which is used in the repeated weighting method in the second step of Algorithm 1 can also be written in an equivalent way:

\[
\hat{\tau}_{Y}^{\text{GREG}(S)} = \sum_{k \in S} w^{(S)}_k y_k,
\]

where weights \( w^{(S)}_k \) can be expressed as follows:

\[
w^{(S)}_k = d^{(S)}_k \left\{ 1 + x_k^T \left( \sum_{k \in S} d^{(S)}_k x_k x_k^T \right)^{-1} \left( \tau_x - \hat{\tau}^{HT(S)}_x \right) \right\}.
\]

It may happen that the margins for a table estimated in the second step do not match the values in the previously created table in terms of shared variables. In this case, it is necessary to recalibrate the estimates to ensure the consistency of the estimated tables, which is exactly what happens in the third step of the RW method. Recalibration should therefore be understood as the process of adjusting values of a more detailed table to ensure that its margins are consistent with the margins of the previous table. For this purpose, the GREG estimator is used again.

Let \( m \) denote a vector of all linearly independent auxiliary variables corresponding to the margins of a table whose counts are contained in \( \tau_y \). The estimator of the parameter \( \tau_y \), based on the data from the S block, which is obtained in the third step of the RW method can be expressed as follows:

\[
\hat{\tau}_Y^{RW} = \hat{\tau}_Y^{\text{GREG}(S)} + \hat{\beta}_s^{RW T} \left( \hat{\tau}_m^{RW} - \hat{\tau}_m^{\text{GREG}(S)} \right)
\]

where:

\[
\hat{\beta}_s^{RW} = \left( \sum_{k \in S} w^{(S)}_k m_k m_k^T \right)^{-1} \sum_{k \in S} w^{(S)}_k m_k y_k^T.
\]

The elements of the vector \( \hat{\tau}_m^{RW} \) are estimates from the previous table or known counts from the register. Also, in this case, the estimator of the RW type, which is expressed by formula (6), can be represented in an alternative form using appropriate weights, i.e.:

\[
\hat{\tau}_Y^{RW} = \sum_{k \in S} r^{(S)}_k y_k
\]

where:

\[
r^{(S)}_k = w^{(S)}_k \left\{ 1 + m_k^T \left( \sum_{k \in S} w^{(S)}_k m_k m_k^T \right)^{-1} \left( \hat{\tau}_m^{RW} - \hat{\tau}_m^{\text{GREG}(S)} \right) \right\}.
\]
Another important issue in the repeated weighting method is the way of estimating the variance of the $\hat{t}_Y^{RW}$ estimator. How this variance estimator is constructed is described in the article by Kottnerus and van Duin (2006), in which the authors focus on the case of one register and two independent sample surveys.

To conclude this section, it is worth emphasizing that the repeated weighting method is not the only technique that can be used in the process of constructing tables that preserve the consistency between respective margins. Some alternatives to the RW method include imputation-based approaches (mass imputation, the repeated imputation technique) and macro-integration. They are described in detail in De Waal (2016).

2. The data

In order to present the possibility of using the repeated weighting method in the NCPH and to construct appropriate census tables, data from the Labour Force Survey were additionally considered. The main objective of the LFS is to collect information about the working, unemployed and economically inactive population. Because the reference point of the NCPH was March 31, 2011, LFS data from for the first quarter of 2011 were used in the analysis (85,275 observations).

NCPH 2011 was a mixed mode census, i.e. based on data from administrative sources as well as those collected directly from the population, partly by full enumeration and partly through a sample survey (Statistics Poland, 2014). Collected information about permanent residents living in Poland is contained in two datasets: one with 31,957,682 observations from full enumeration and the second one, with 6,694,220 records, from the sample survey.

The repeated weighting method requires that data from different sources supplying a census (registers and sample surveys) are properly linked. Deterministic or stochastic methods of data linkage may be used to combine information from different sources to identify and bring together records from separate files, which correspond to the same entities (Rässler, 2012; Roszka, 2013). In most cases, it is not a simple procedure and linkage errors, false and missed links are unavoidable (Zhang & Tuoto, 2020).

In the deterministic method unique identifiers for each record are compared to determine a match (Harron et al., 2015). In administrative sources relating to persons the PESEL number (Universal Electronic System for Registration of the Population in Poland) is typically used as the matching key. In practice, however, it often happens that there is no common key in datasets to be linked. In such situations stochastic methods are used, including probabilistic record linkage which makes it possible to integrate databases containing information about the same units with no shared unique key (Sayers, Ben-Shlomo, Blom...
This method uses the so-called pairing variables, which are used to estimate the probability that a pair of records from different datasets refer to the same unit. Pairing variables are selected from among the variables present in both integrated datasets; in the case of persons, likely candidates include the address of residence, sex or birth date.

Owing to the lack of access to unique identifiers shared by the NCPH and the LFS deterministic integration could not be used. This is why probabilistic record linkage, mentioned above, was applied. In order to optimize the integration algorithm by reducing the number of record pairs to be compared, the following variables were used as blocking variables: the commune of residence (NUTS5 level), the category of the place of residence (urban / rural), sex and ten-year age groups. Record matching was performed using the birth year as a pairing variable. The reclin package (van der Laan, 2018) of the R program (R Core Team, 2019) was used for the stochastic process of data integration.

Following the integration only some respondents in the LFS (18,481 observations) were linked with both the full and sample part of the NCPH (see Figure 2). A much larger group of LFS respondents could only be linked with the full part of the NCPH (66,794 observations). The inclusion relationship between sets $FS$, $SS_1$, $SS_2$ will be used for the construction of the so-called blocks, i.e. sets that will be the basis for estimating total values in the specified output tables.

![Figure 2. Integration of datasets from the NCPH and the LFS using probabilistic record linkage](image)
Source: Own study.
3. The estimation procedure

Four blocks were extracted from the datasets obtained by integrating the NCPH and the LFS (see Figure 3). In Figure 3 block $S_0$ corresponds to the full enumeration in the NCPH—it covers the entire population and includes only those variables that were observed in the full enumeration in the NCPH. Block $S_1$ consists of respondents from the sample part of the NCPH and includes variables from both the full enumeration and the sample part of the NCPH. Block $S_2$ is an extension of the LFS dataset with variables from the full enumeration of the NCHP. Block $S_3$ is the common part of the sample part of the NCPH and LFS ($S_3 = S_1 \cap S_2$). On the one hand, $S_3$ is the smallest block in terms of the number of observations. On the other hand, it contains variables from all three surveys and therefore is the richest due to the scope of information.

According to the algorithm described above, the estimation of totals using the RW estimator is performed in the following three steps. In the first step, the final tables are defined, i.e. cross-classifications for which totals will be estimated. For illustrative purposes the following variables were used: urban /
rural ($U$), sex ($G$), age group ($A$), biological disability ($D$) and work experience ($E$). Their coding is presented in Table 1.

Table 1. Variables used for constructing the census tables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>urban/rural</td>
<td>$U$</td>
<td>1 urban</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 rural</td>
</tr>
<tr>
<td>sex</td>
<td>$G$</td>
<td>1 male</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 female</td>
</tr>
<tr>
<td>age group</td>
<td>$A$</td>
<td>1 15–17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 18–19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 20–24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 25–29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 30–34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 35–39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7 40–44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8 45–49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9 50–54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 55–59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11 60–64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12 65+</td>
</tr>
<tr>
<td>biological disability*</td>
<td>$D$</td>
<td>1 yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 no</td>
</tr>
<tr>
<td>work experience</td>
<td>$E$</td>
<td>1 yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 no</td>
</tr>
</tbody>
</table>

*Category 1 (yes) is assigned to people who reported biological disability, while Category 2 (no) applies to all others, which includes people who did not answer the question about biological disability. This method of coding and presentation of results was adopted by Statistics Poland.

Source: Own study.

The values of urban / rural, sex and age group variables are known for all respondents in the NCPH, while information on biological disability was collected only for respondents in the sample part of the NCPH and information regarding work experience is available only for LFS respondents.

In the presented example, we estimate table $D \times E \times G$ containing counts of people in cross-classifications defined by levels of biological disability, work experience and sex, as well as tables resulting from the aggregation of at
least one of its dimensions. The final set of tables to be estimated is as follows: \( D, E, D \times E, D \times G, E \times G, D \times E \times G. \)

The \( G \) table, containing counts of people by sex, is not estimated because it can be calculated by counting observations from the \( S_0 \) block covering the entire population. It should be noted that biological disability \( D \) and work experience \( E \) are not simultaneously observed in any of the surveys considered. Therefore, it is not possible to estimate \( D \times E \times G \) and \( D \times E \) based on one data source only. Hence, the \( S_1 \) block, which is a common part of the NCPH and the LFS, is used for this purpose.

In the second step, the totals contained in the tables defined in the first step are estimated based on the GREG estimator described in Section 1. Every table is estimated based on the most numerous block containing the variables defining the cross-classifications (see Table 2). As regards auxiliary information for the estimation process, a vector of indicator variables \( x_k^{UAG} = (x_{1,k}, \ldots, x_{48,k})^T \) was used, which determines membership in one of the 48 groups defined by interactions of the following variables: urban / rural (\( U \)), sex (\( G \)) and age group (\( A \))^7.

![Table 2. Final tables and corresponding blocks](image)

<table>
<thead>
<tr>
<th>Block</th>
<th>Output table</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_1 )</td>
<td>( D )</td>
</tr>
<tr>
<td>( S_1 )</td>
<td>( G \times D )</td>
</tr>
<tr>
<td>( S_2 )</td>
<td>( E )</td>
</tr>
<tr>
<td>( S_2 )</td>
<td>( G \times E )</td>
</tr>
<tr>
<td>( S_3 )</td>
<td>( D \times E )</td>
</tr>
<tr>
<td>( S_3 )</td>
<td>( G \times D \times E )</td>
</tr>
</tbody>
</table>

Source: Own study.

The \( S_1 \) block was used to estimate the number of people by biological disability (\( \hat{D} \)) as well as biological disability and sex (\( D \times G \)). For these domains, the following GREG estimators were used:

\[
\hat{\tau}_D^{GREG(S_1)} = \hat{\tau}_D^{HT(S_1)} + \hat{\beta}_{S_1,D,UGA}^T (\tau_{UGA} - \hat{\tau}_{UGA}^{HT(S_1)}),
\]

\[
\hat{\tau}_{DG}^{GREG(S_1)} = \hat{\tau}_{DG}^{HT(S_1)} + \hat{\beta}_{S_1,DG,UGA}^T (\tau_{UGA} - \hat{\tau}_{UGA}^{HT(S_1)}),
\]

where the final weights from the sample part of the NCPH were used as design weights \( d_k^{(S)} (k \in S_1) \) with auxiliary variables defined as shown above.

---

7 This set of auxiliary variables was originally used in the process of weight calibration in the LFS by Statistics Poland.
The $S_2$ block was used to estimate totals by work experience ($E$) as well as work experience and sex ($E \times G$). GREG estimators for these tables are given by:

$$
\hat{\tau}_{S2, E}^{GREG(S_2)} = \hat{\tau}_{E}^{HT(S_2)} + \hat{\beta}_{S2, E, UGA}^T (\tau_{UGA} - \hat{\tau}_{UGA}^{HT(S_2)}),
$$
(12)

$$
\hat{\tau}_{S2, EG}^{GREG(S_2)} = \hat{\tau}_{EG}^{HT(S_2)} + \hat{\beta}_{S2, EG, UGA}^T (\tau_{UGA} - \hat{\tau}_{UGA}^{HT(S_2)}),
$$
(13)

where the original LFS weights were used as design weights along with auxiliary variables defined as shown above.

The $S_3$ block was used to estimate totals in domains defined by biological disability and work experience ($D \times E$) as well as biological disability, work experience and sex ($D \times E \times G$) using the following GREG estimators:

$$
\hat{\tau}_{S3, DE}^{GREG(S_3)} = \hat{\tau}_{DE}^{HT(S_3)} + \hat{\beta}_{S3, DE, UGA}^T (\tau_{UGA} - \hat{\tau}_{UGA}^{HT(S_3)}),
$$
(14)

$$
\hat{\tau}_{S3, DEG}^{GREG(S_3)} = \hat{\tau}_{DEG}^{HT(S_3)} + \hat{\beta}_{S3, DEG, UGA}^T (\tau_{UGA} - \hat{\tau}_{UGA}^{HT(S_3)}). 
$$
(15)

The $S_3$ block is the common part of two sample surveys, hence input weights were determined as the product of weights from the sample part of the NCPH and LFS weights: $d_k^{(S_3)} = d_k^{(S_1)} d_k^{(S_2)}$ ($k \in S_3 = S_1 \cap S_2$). The same vector of auxiliary variables was selected.

By applying this approach separately for each of the $S_i$ blocks ($i = 1, 2, 3$) one set of weights $w_k^{(S_i)}$ is obtained, which can then be used in the process of estimating totals as follows:

$$
\hat{\tau}_{y}^{GREG(S_i)} = \sum_{k \in S_i} w_k^{(S_i)} y_k,
$$
(16)

where $y_k$ is a vector of indicator variables specifying the cross-classifications in the estimated table. The weights $w_k^{(S_i)}$ can be expressed using formula (5), where the $S$ block is replaced with the $S_i$ block and the vector of auxiliary variables is defined the same way as previously, i.e. $x_k = x_k^{UGA}$.

The use of one vector of weights $w_k^{(S_i)}$ in the given $S_i$ block ensures consistency between the output tables estimated using this block. It is also easy to show that by using sex ($G$) as the auxiliary variable in the second step of the estimation process consistency was also achieved between tables estimated on the basis of the $S_1$ and $S_2$ blocks. Therefore, the estimates of totals in the domains: $D, D \times G, E$ and $E \times G$ obtained using the estimators $\hat{\tau}_{D}^{GREG(S_1)}, \hat{\tau}_{DG}^{GREG(S_1)}, \hat{\tau}_{E}^{GREG(S_2)}, \hat{\tau}_{EG}^{GREG(S_2)}$ can be regarded as consistent. Therefore they will be treated as the final RW estimates.

A problem arises in the case of tables estimated from the $S_3$ block. If the counts of people by biological disability $D$ and by work experience $E$ were to be calculated by aggregating the appropriate totals contained in $\hat{\tau}_{DE}^{GREG(S_3)}$ and
results could be obtained that were different from the corresponding totals given by the estimate of \( \hat{\tau}_{D}^{GREG(S)} \) and \( \hat{\tau}_{E}^{GREG(S)} \). In order to avoid this potential inconsistency it is necessary to recalibrate the output tables.

The third step is performed in two stages (3a and 3b). In stage 3a, the GREG estimator is used again to ensure consistency between the following output tables: \( D, E, D \times E \). A vector of auxiliary variables was constructed as follows:

\[
x_{k}^{D+E} = (x_{1,k}^{D}, x_{2,k}^{D}, x_{1,k}^{E})^{T},
\]

(17)

where \( x_{1,k}^{D} \) takes the value 1, when the \( D \) variable for the \( k \)-th unit is the first category and 0 otherwise. The remaining indicator variables are defined in the analogous fashion. In order to avoid collinearity in \( x_{k}^{D+E} \), the variable \( x_{2,k}^{E} \) was omitted. The weights \( w_{k}^{(3)} \) from the second step were used as input weights. Finally, the estimates of \( D \times E \) table were obtained using the following estimator:

\[
\hat{\tau}_{DE}^{RW} = \hat{\tau}_{DE}^{GREG(S)} + \left( \hat{\beta}_{w}^{(3), DE} \right)^{T} \left( \hat{\tau}_{D}^{GREG(S)} - \hat{\tau}_{D}^{GREG(S)} \right).
\]

(18)

The purpose of stage 3b was to ensure consistency between estimates in \( D \times E \times G, D \times E, D \times G, E \times G \) tables. In this case, the vector of auxiliary variables was constructed as follows:

\[
x_{k}^{DG+EG+D+E} = (x_{1,1,k}^{DG}, x_{1,2,k}^{DG}, x_{2,1,k}^{DG}, x_{2,2,k}^{DG}, x_{1,1,k}^{EG}, x_{2,1,k}^{EG}, x_{1,1,k}^{DE})^{T},
\]

(19)

where \( x_{1,1,k}^{DG} \) is an indicator variable taking the value 1, when the \( k \)-th unit simultaneously has the first category of the \( D \) variable and the first level of the \( G \) variable, and 0 otherwise. The remaining indicator variables are defined in the analogous way. The weights \( w_{k}^{(3)} \) from the second step were used as input weights. Finally, the estimator of table \( D \times E \times G \) is given by:

\[
\hat{\tau}_{DE}^{RW} = \hat{\tau}_{DE}^{GREG(S)} + \left( \hat{\beta}_{w}^{(3), DE, DG} \right)^{T} \left( \hat{\tau}_{DG}^{GREG(S)} - \hat{\tau}_{DG}^{GREG(S)} \right) + \left( \hat{\beta}_{w}^{(3), DE, EG} \right)^{T} \left( \hat{\tau}_{EG}^{GREG(S)} - \hat{\tau}_{EG}^{GREG(S)} \right) + \left( \hat{\beta}_{w}^{(3), DE, D+E} \right)^{T} \left( \hat{\tau}_{D+E}^{GREG(S)} - \hat{\tau}_{D+E}^{GREG(S)} \right).
\]

(20)

The RW estimators described by formulas (18)–(20) are identical to calibration estimators given by:

\[
\hat{\tau}_{DE}^{RW} = \sum_{k \in S_{y}} \hat{y}_{k}^{(S_{y},1)} y_{k}.
\]

(21)
\[ \hat{t}_{\text{RW}}^{(S_i, 2)} = \sum_{k \in S} r_k^{(S_i, 2)} y_k. \]  

The weights \( r_k^{(S_i, 1)} \) and \( r_k^{(S_i, 2)} \) are given by formula (9), where the vectors of auxiliary variables \( m_k \) are \( x_k^{D+} = x_k^{D+ + E} \) and \( x_k^{DG+ + D+} = x_k^{D+ + E + D+} \) respectively, and inputs weights \( w_k^{(S)} \) are weights \( w_k^{(S)} \), obtained in the second step.

4. Results

Before applying the repeated weighting method, which consists of the three steps described above, a table containing the structure of the population aged 15+ by sex was constructed (see Table 3). It was constructed using data from the NCPH (full enumeration), i.e. was based on the \( S_0 \) block. Values in Table 3 do not have to be estimated because they can be obtained by simply counting observations (respondents) from the \( S_0 \) block covering the entire population. Such a table is constructed in order to verify the consistency of the row and column totals in the remaining tables, where sex is one of the variables.

**Table 3. The structure of the population aged 15+ by sex (G)**

<table>
<thead>
<tr>
<th></th>
<th>( G_1 )</th>
<th>( G_2 )</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>15,272,239</td>
<td>16,685,443</td>
<td>31,957,682</td>
</tr>
</tbody>
</table>

Source: Calculation based on integrated datasets from NCPH and LFS.

In the first step of applying the RW method the following tables were constructed: \( D, D \times G, E, E \times G, D \times E \) and \( D \times E \times G \) according to formulas (10)-(15) from the previous section. The totals estimated in the second step are presented in Tables 4-9. The margins for sex (G) and biological disability (D) in Table 5 (\( D \times G \)) and for sex (G) and work experience (E) in Table 7 (\( E \times G \)) are consistent with values shown in Tables 3, 4 and 6 (\( D, G \) and \( E \)). As noted earlier consistency issues arise with Tables 8 and 9 (\( D \times E \) and \( D \times E \times G \)). The margins for biological disability (D) and work experience (E) are not consistent with those presented in Tables 4, 5, 6 and 7 (\( D, D \times G, E \) and \( E \times G \)). Consistency is only achieved in the case of sex (G), which results from its being used as one of the auxiliary variables in the estimation process.

**Table 4. The structure of the population aged 15+ by biological disability (D)**

<table>
<thead>
<tr>
<th>( D_1 )</th>
<th>( D_2 )</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>4,015,414</td>
<td>27,942,268</td>
<td>31,957,682</td>
</tr>
</tbody>
</table>

Source: Calculation based on integrated datasets from NCPH and LFS.
Table 5. The structure of the population aged 15+ by biological disability and sex \((D \times G)\)

<table>
<thead>
<tr>
<th>G_i</th>
<th>D_1</th>
<th>D_2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>G_1</td>
<td>1,815,894</td>
<td>13,456,345</td>
<td>15,272,239</td>
</tr>
<tr>
<td>G_2</td>
<td>2,199,520</td>
<td>14,485,923</td>
<td>16,685,443</td>
</tr>
<tr>
<td>Total</td>
<td>4,015,414</td>
<td>27,942,268</td>
<td>31,957,682</td>
</tr>
</tbody>
</table>

Source: Calculation based on integrated datasets from NCPH and LFS.

Table 6. The structure of the population aged 15+ by work experience \((E)\)

<table>
<thead>
<tr>
<th>E</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E_1</td>
<td>E_2</td>
<td>Total</td>
</tr>
<tr>
<td>27,399,998</td>
<td>4,557,684</td>
<td>31,957,682</td>
<td></td>
</tr>
</tbody>
</table>

Source: Calculation based on integrated datasets from NCPH and LFS.

Table 7. The structure of the population aged 15+ by work experience and sex \((E \times G)\)

<table>
<thead>
<tr>
<th>G_i</th>
<th>E_1</th>
<th>E_2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>G_1</td>
<td>13,242,650</td>
<td>2,029,589</td>
<td>15,272,239</td>
</tr>
<tr>
<td>G_2</td>
<td>14,157,349</td>
<td>2,528,094</td>
<td>16,685,443</td>
</tr>
<tr>
<td>Total</td>
<td>27,399,999</td>
<td>4,557,683</td>
<td>31,957,682</td>
</tr>
</tbody>
</table>

Source: Calculation based on integrated datasets from NCPH and LFS.

Table 8. The structure of the population aged 15+ by biological disability and work experience \((D \times E)\)—intermediate estimates obtained in the second step

<table>
<thead>
<tr>
<th>D_i</th>
<th>E_1</th>
<th>E_2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>D_1</td>
<td>3,881,336</td>
<td>234,662</td>
<td>4,115,997</td>
</tr>
<tr>
<td>D_2</td>
<td>23,529,766</td>
<td>4,311,919</td>
<td>27,841,685</td>
</tr>
<tr>
<td>Total</td>
<td>27,411,102</td>
<td>4,546,581</td>
<td>31,957,682</td>
</tr>
</tbody>
</table>

Source: Calculation based on integrated datasets from NCPH and LFS.
Table 9. The structure of the population aged 15+ by biological disability, work experience and sex ($D \times E \times G$)—intermediate estimates obtained in the second step

<table>
<thead>
<tr>
<th></th>
<th>$E_1$</th>
<th>$E_2$</th>
<th>$G_1$</th>
<th>$G_2$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_1$</td>
<td>1,866,107</td>
<td>2,015,229</td>
<td>85,382</td>
<td>149,280</td>
<td>4,115,998</td>
</tr>
<tr>
<td>$D_2$</td>
<td>11,393,228</td>
<td>12,136,537</td>
<td>1,927,521</td>
<td>2,384,397</td>
<td>27,841,683</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>13,259,335</strong></td>
<td><strong>14,151,766</strong></td>
<td><strong>2,012,903</strong></td>
<td><strong>2,533,677</strong></td>
<td><strong>31,957,681</strong></td>
</tr>
</tbody>
</table>

Source: Calculation based on integrated datasets from NCPH and LFS.

In the third step of the repeated weighting algorithm, given the observed inconsistency between margins in the tables estimated in the second step, it was necessary to perform recalibration. As a result, the margins for biological disability ($D$) and work experience ($E$) in Tables 10 ($D \times E$) and 11 ($D \times E \times G$) match the values in Tables 4, 5, 6 and 7 ($D$, $D \times G$, $E$ and $E \times G$ respectively). Also aggregations by $D \times E$ in Table 11 ($D \times E \times G$) are consistent with the values in Table 10 ($D \times E$).

Table 10. The structure of the population aged 15+ by biological disability and work experience ($D \times E$)—final estimates obtained in the third step

<table>
<thead>
<tr>
<th></th>
<th>$E_1$</th>
<th>$E_2$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_1$</td>
<td>3,786,414</td>
<td>229,000</td>
<td>4,015,414</td>
</tr>
<tr>
<td>$D_2$</td>
<td>23,613,584</td>
<td>4,328,684</td>
<td>27,942,268</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>27,399,998</strong></td>
<td><strong>4,557,684</strong></td>
<td><strong>31,957,682</strong></td>
</tr>
</tbody>
</table>

Source: Calculation based on integrated datasets from NCPH and LFS.

Table 11. The structure of the population aged 15+ by biological disability, work experience and sex ($D \times E \times G$)—final estimates obtained in the third step

<table>
<thead>
<tr>
<th></th>
<th>$E_1$</th>
<th>$E_2$</th>
<th>$G_1$</th>
<th>$G_2$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_1$</td>
<td>1,737,104</td>
<td>2,049,311</td>
<td>78,790</td>
<td>150,209</td>
<td>4,015,414</td>
</tr>
<tr>
<td>$D_2$</td>
<td>11,505,546</td>
<td>12,108,038</td>
<td>1,950,799</td>
<td>2,377,885</td>
<td>27,942,268</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>13,242,650</strong></td>
<td><strong>14,157,349</strong></td>
<td><strong>2,029,589</strong></td>
<td><strong>2,528,094</strong></td>
<td><strong>31,957,682</strong></td>
</tr>
</tbody>
</table>

Source: Calculation based on integrated datasets from NCPH and LFS.

Another important issue in the repeated weighting method is the way of assessing the estimation precision of totals in the output tables (see Table 12).
Table 12. Relative estimation errors of totals in the output tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Domain</th>
<th>n</th>
<th>REE(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D$</td>
<td>$D_1$</td>
<td>2,647</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>$D_2$</td>
<td>15,834</td>
<td>0.0</td>
</tr>
<tr>
<td>$D \times G$</td>
<td>$D_1 \times G_1$</td>
<td>1,221</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>$D_1 \times G_2$</td>
<td>1,426</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>$D_2 \times G_1$</td>
<td>7,442</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>$D_2 \times G_2$</td>
<td>8,392</td>
<td>0.0</td>
</tr>
<tr>
<td>$E$</td>
<td>$E_1$</td>
<td>15,812</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>$E_2$</td>
<td>2,669</td>
<td>0.6</td>
</tr>
<tr>
<td>$E \times G$</td>
<td>$E_1 \times G_1$</td>
<td>7,453</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>$E_1 \times G_2$</td>
<td>8,359</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>$E_2 \times G_1$</td>
<td>1,210</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>$E_2 \times G_2$</td>
<td>1,459</td>
<td>0.8</td>
</tr>
<tr>
<td>$D \times E$</td>
<td>$D_1 \times E_1$</td>
<td>2,496</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>$D_1 \times E_2$</td>
<td>151</td>
<td>9.6</td>
</tr>
<tr>
<td></td>
<td>$D_2 \times E_1$</td>
<td>13,316</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>$D_2 \times E_2$</td>
<td>2,518</td>
<td>0.8</td>
</tr>
<tr>
<td>$D \times E \times G$</td>
<td>$D_1 \times E_1 \times G_1$</td>
<td>1,162</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>$D_1 \times E_1 \times G_2$</td>
<td>1,134</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td>$D_1 \times E_2 \times G_1$</td>
<td>59</td>
<td>48.9</td>
</tr>
<tr>
<td></td>
<td>$D_1 \times E_2 \times G_2$</td>
<td>92</td>
<td>42.8</td>
</tr>
<tr>
<td></td>
<td>$D_2 \times E_1 \times G_1$</td>
<td>6,291</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>$D_2 \times E_1 \times G_2$</td>
<td>7,025</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>$D_2 \times E_2 \times G_1$</td>
<td>1,151</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td>$D_2 \times E_2 \times G_2$</td>
<td>1,367</td>
<td>4.6</td>
</tr>
</tbody>
</table>

Source: Calculations based on integrated datasets from NCPH and LFS.

Table 12 presents relative estimation errors (REE) for all output tables, where estimator variance was evaluated using the method described by Knottnerus and van Duin (2006). In most cases, the obtained estimates are acceptable, with errors not exceeding 10%, except for estimates in table $D \times E \times G$ i.e.

---

8 Unfortunately, the applied approach does not take into account the randomness resulting from the stochastic method of data integration; hence the error measures presented in this article may be underestimated.
\[ D_1 \times E_2 \times G_1 \text{ and } D_1 \times E_2 \times G_2, \] for which exceeds this threshold. This is due to the small number of observations (59 and 92 respectively) used for estimating these totals.

**Conclusions**

The idea of the RW method is to ensure consistency of all final tables that created from various data sources supplying the census or to obtain numerically consistent estimates of tables from a combination of registers and surveys. The RW technique is based on repeated application of GREG and produces a new set of calibration weights for each table estimate. So far, it has been used successfully only by Statistics Netherlands in the regular estimation process to ensure consistency between different tables. Other National Statistical Offices, including Statistics Poland, in most cases use the calibration approach, which means that different systems of weights are used to produce tables based on different surveys.

In this study, the RW method was implemented based on data from NCPH and LFS from 2011. The goal was to construct tables defined by the status of biological disability, the work experience and sex. Thanks to the use of the RW method the following goals have been realised. First, this technique made it possible to construct a set of tables with consistent margins. This is, in fact, one of the most important goals in official statistics, which ensures that the end user obtains consistent data from various sources of information. Second, by combining information from many data sources, the RW method has been used to produce output tables that could not be obtained using only one source of information. This application has shown that the RW is particularly useful when the final tables have some variables in common.

Bearing in mind the advantages of the RW method and the fact that it has not been used so far in surveys carried out by Statistics Poland, the main goal of this article was to present the first application of this approach to real data from output tables. The use of the RW method and its testing on data from 2011 could pave the way for its implementation in the next Polish census scheduled for 2021 or in surveys where it is necessary to maintain consistency of estimates with results obtained from registers or other data sources.

**References**


Aims and Scope

The Economics and Business Review is a quarterly journal focusing on theoretical, empirical and applied research in the fields of Economics and Corporate and Public Finance. The Journal welcomes the submission of high quality articles dealing with micro, mezzo and macro issues well founded in modern theories and relevant to an international audience. The EBR's goal is to provide a platform for academicians all over the world to share, discuss and integrate state-of-the-art Economics and Finance thinking with special focus on new market economies.

The manuscript

1. Articles submitted for publication in the Economics and Business Review should contain original, unpublished work not submitted for publication elsewhere.
2. Manuscripts intended for publication should be written in English, edited in Word in accordance with the APA editorial guidelines and sent to: secretary@ebr.edu.pl. Authors should upload two versions of their manuscript. One should be a complete text, while in the second all document information identifying the author(s) should be removed from papers to allow them to be sent to anonymous referees.
3. Manuscripts are to be typewritten in 12” font in A4 paper format, one and half spaced and be aligned. Pages should be numbered. Maximum size of the paper should be up to 20 pages.
4. Papers should have an abstract of about 100-150 words, keywords and the Journal of Economic Literature classification code (JEL Codes).
5. Authors should clearly declare the aim(s) of the paper. Papers should be divided into numbered (in Arabic numerals) sections.
6. Acknowledgements and references to grants, affiliations, postal and e-mail addresses, etc. should appear as a separate footnote to the author’s name a, b, etc and should not be included in the main list of footnotes.
7. Footnotes should be listed consecutively throughout the text in Arabic numerals. Cross-references should refer to particular section numbers: e.g.: See Section 1.4.
8. Quoted texts of more than 40 words should be separated from the main body by a four-spaced indentation of the margin as a block.
10. Copyrights will be established in the name of the E&BR publisher, namely the Poznań University of Economics and Business Press.

More information and advice on the suitability and formats of manuscripts can be obtained from:

Economics and Business Review
al. Niepodległości 10
61-875 Poznań
Poland
e-mail: secretary@ebr.edu.pl
www.ebr.edu.pl
Subscription

Economics and Business Review (E&BR) is published quarterly and is the successor to the Poznań University of Economics Review. The E&BR is published by the Poznań University of Economics and Business Press.

Economics and Business Review is indexed and distributed in Scopus, Clarivate Analytics, DOAJ, ERIH plus, ProQuest, EBSCO, CEJSH, BazEcon, Index Copernicus and De Gruyter Open (Sciendo).

Subscription rates for the print version of the E&BR: institutions: 1 year – €50.00; individuals: 1 year – €25.00. Single copies: institutions – €15.00; individuals – €10.00. The E&BR on-line edition is free of charge.