Acta Universitatis Lodziensis

FOLIA OECONOMICA

2(353)
2021

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

Acta Universitatis Lodziensis

FOLIA OECONOMICA
2(353)
2021

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

Acta Universitatis Lodziensis

2(353)
2021

FOLIA OECONOMICA

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

Łódź 2021

COPE
Member since 2018
JM13703

ISSN 0208-6018
e-ISSN 2353-7663

Czasopismo afiliowane przy Wydziale Ekonomiczno-Socjologicznym Uniwersytetu Łódzkiego

WYDZIAŁ EKONOMICZNO-SOCJOLOGICZNY
Uniwersytet Łódzki

Folia Oeconomica
Acta Universitatis Lodziensis
ISSN 0208-6018 e-ISSN 2353-7663
www.czasopisma.uni.lodz.pl/foe/
2(353) 2021

Contents

Folia Oeconomica
Acta Universitatis Lodziensis
ISSN 0208-6018 e-ISSN 2353-7663
www.czasopisma.uni.lodz.pl/foe/
2(353) 2021

DOI: https://doi.org/10.18778/0208-6018.353.01

Agnieszka Palma
University of Łódź, Faculty of Economics and Sociology, Department of Demography and Social Gerontology, Łódź, Poland, agnieszka.palma@uni.lodz.pl

Fertility Trends in European Countries

Abstract: This work is intended as an attempt to illustrate and compare the pattern of fertility in European countries: Belarus, Croatia, Hungary, Poland, Portugal, Spain, Sweden, and Switzerland. It deals with the analysis of fertility trends, with an emphasis on birth by parity. Using data from the Human Fertility Database (HFD) from the year 2016, it has considered the parameters of parity progression ratios (PPR), projected parity progression ratios (PPPR), age-specific fertility rates (ASFR), age-order specific fertility rates (AOSFR), and cumulated order-specific fertility rates accordingly analysed. We have applied indicators known as the projected parity progression ratios to estimate trends of fertility. These offer a more detailed view of the family formation process than the traditional total fertility rate (TFR).

Keywords: fertility rates, average parity, parity progression ratios, projected parity progression ratios

JEL: J11, J12, J13

1. Introduction

Fertility analysis is a crucial subject in demography and it has attracted the attention of many researchers. Demographers have designed many various methods to measure this phenomenon. Most studies have used common measures, such as age-specific fertility and total fertility rates, to analyse fertility trends. These measures provide aggregated estimates of fertility. They do not provide enough detail on exactly how fertility is changing within parity. The analysis of fertility by parity helps us understand the fertility transition in a given population.

Parity progression ratios (PPRs) have been one of the most frequently used demographic tools in analysing data by parity. Parity progression ratios are well described in the demographic literature (Preston, Heuveline, Guillot, 2001). Louis Henry (1980) was the first to introduce parity progression ratios in 1953. They were further refined by other authors: Rodriguez and Hobcraft (1980), Brass and Juarez (1983), Feeney (1983), Brass (1985), Yadava and Bhattacharya (1985), Feeney and Jingyuan (1987), Aoun and Airey (1988), Yadava, Pandey, and Saxena (1992), Sloggett et al. (1994), Islam and Yadava (1997), Bhardwaj, Sharma, and Kumar (2010), as well as Yadava and Kumar (2011). By definition, a parity progression ratio refers to the proportion of women who progress from parity i to $i + 1$ (Feeney, Jingyuan, 1987; Preston, Guillot, Heuveline, 2001). PPRs can be calculated on a cohort or period basis depending on the data available.

Frejka (2008) wrote that the decreasing PPR for first and second births played a key role in the fertility decline of European women born after 1955. In Central and Eastern Europe, the fertility decline was mainly caused by the decreasing PPR for second births. Kohler, Billari, and Ortega (2002) and Billari and Kohler (2004) suggest that the lowest fertility in Europe in the 1990s was the result of delayed childbearing, especially for first births, as well as a low probability of progression after the first child.

Brass (1985) described a method of using the current distribution of age-order specific fertility rates to estimate the future trends of parity progression ratios known as the projected parity progression ratio method. PPPRs denote the proportions of women who are expected to progress from parity i to parity $i + 1$ by the end of their childbearing years. This study is based on this method. Brass used it in the Seychelles, and was subsequently applied by Moultrie et al. (2013) in Cambodia, and by Rossa and Palma (2020) in Poland, with some improvements.

The main purpose of this study is to investigate the tempo and pattern of fertility in European countries and the parity progression schedule. The specific objectives are to estimate parity progression ratios for each parity and to examine the differentials in these ratios in the considered countries. Taking into account the fact that there are very few women with high birth order, and therefore no significant changes in the distribution of children in the family, we focused on parities from one to four for simplification.

In this paper, we also make some ex-post comparisons, i.e. between the observed and predicted PPR for women in different age groups and between the PPPR and completed PPR observed for women aged 49 and over in a given calendar year. They allow us to assess the prediction accuracy of the modified Brass method as well as to explain future changes in the distribution of parities over a ten-year time horizon.

Input data were sourced from the Human Fertility Database. They present age-parity exposure of the female population in the year 2016 as well as counts of live births by birth order and age of mothers in the same year. The choice of countries was dictated by TFR size, data availability and geographical location. The database is available online at www.humanfertility.org.

2. Model for estimating parity progression ratios

2.1. Notation

For the convenience of the reader, we repeat the relevant material (Moultrie et al., 2013), thus making our exposition self-contained.

Let us assume the following notation:

N_x – population exposure for females in one-year age interval $[x, x + 1)$,

$N_x(i)$ – population exposure for females in one-year age interval $[x, x + 1)$ and of parity i,

N – total female population exposure in the reference year,

$N(i)$ – total population exposure in the reference year for women of parity i.

The following relations hold:

$$N_x = \sum_{i=0}^{\pi} N_x(i), \quad N(i) = \sum_{x=\alpha}^{\beta} N_x(i), \quad N = \sum_{i=1}^{\pi} N(i) = \sum_{x=\alpha}^{\beta} N_x, \tag{1}$$

where π is the highest parity in the data set, and α, β define the limits of the reproductive age range $[\alpha, \beta + 1)$. Further, we will assume that $\alpha = 15$ and $\beta = 49$.

Let us also denote:

B_x – the number of births to women aged $[x, x + 1)$,

$B_x(i)$ – the number of births to women in one-year age interval $[x, x + 1)$ and of parity i,

B – the total number of births in the reference year,

$B(i)$ – the total number of births in the reference year to women of i-th parity.

Similarly as in (1), we obtain the following relations:

$$B_x = \sum_{i=0}^{\pi} B_x(i),\ B(i) = \sum_{x=\alpha}^{\beta} B_x(i),\ B = \sum_{i=1}^{\pi} B(i) = \sum_{x=\alpha}^{\beta} B_x. \quad (2)$$

Now we define the Age-Specific Fertility Rate (ASFR) and the Age-Order Specific Fertility Rate (AOSFR) for women at the age between x and $x + 1$:

$$ASFR_x = \frac{B_x}{N_x} \text{ and } AOSFR_x(i) = \frac{B_x(i)}{N_x}. \quad (3)$$

Observe that:

$$ASFR = \sum_{i=1}^{\pi} AOSFR_x(i). \quad (4)$$

The Total Fertility Rate (TFR) and the Total Order Fertility Rates (TOFR's) are defined as:

$$TFR = \sum_{x=\alpha}^{\beta} ASFR_x \text{ and } TOFR(i) = \sum_{x=\alpha}^{\beta} AOSFR_x(i). \quad (5)$$

The Cumulated Age-Order Fertility Rates for parity i are derived by summing age-order specific fertility rates up to the desired age y. Thus, we have:

$$TOFR_y(i) = \sum_{x=\alpha}^{y} AOSFR_x(i). \quad (6)$$

Average parity P in the population is calculated by dividing the total number of children ever born by the number of women N:

$$P = \frac{1}{N} \sum_{j=0}^{\pi} j \cdot N(j). \quad (7)$$

Now, let $W(i)$ be the number of women in the population having attained parity i or higher. Note that:

$$W(i) = \sum_{j=i}^{\pi} N(j). \quad (8)$$

Then the proportion $M(i)$ of women ever-attaining parity i, can be expressed as:

$$M(i) = \frac{W(i)}{N}. \quad (9)$$

Obviously, the corresponding proportion at parity zero and over is $M(0) = N/N = 1$.

By analogy, we obtain $W_x(i)$ and $M_x(i)$.

It can be shown that the average parity (7) can be written as follows:

$$P = \frac{1}{N}\sum_{j=1}^{\pi} W(j) = \sum_{j=1}^{\pi} M(j). \tag{10}$$

2.2. (Projected) Parity Progression Ratios

Now we can derive the proportions of women who have progressed from parity i to $i + 1$, which are the parity progression ratios. These are obtained by dividing the proportion of women with $i + 1$ children by that with i children. Thus the parity progression ratios $PPR(i)$ and $PPR_x(i)$ are given by:

$$PPR(i) = \frac{M(i+1)}{M(i)}, PPR_x(i) = \frac{M_x(i+1)}{M_x(i)}. \tag{11}$$

It is important to know that age-specific parity progression ratios $PPR_x(i)$ calculated for young women who have not yet completed their childbearing life should be treated as current age-specific PPRs. These ratios can change rapidly when women move into higher parities in the future. In such cases, it is reasonable to calculate projected parity progression ratios to predict parity progression of young women by the end of their childbearing age.

Observe that proportions (9) can be expressed in terms of respective parity progression ratios (11) for lower birth orders as follows:

$$M(i + 1) = PPR(i) \cdot M(i) = PPR(i) \cdot PPR(i - 1) \cdot PPR(i - 2) \ldots PPR(0). \tag{12}$$

Projected proportions attaining parity i, M_x^* denote the proportions of women who are expected to attain at least parity i by the end of their childbearing life,

$$M_x^*(i) = M_x(i) + TOFR(i) - TOFR_x(i), \tag{13}$$

where the difference between Total Order Fertility Rate $TOFR(i)$ given in (5) and Cumulated Age-Order Fertility Rate $TOFR_x(i)$ defined in (6) can be treated as an estimate of an additional proportion of women aged x expected to achieve parity i by the end of their childbearing years. However, this interpretation is admissible under the assumption that current fertility will remain constant until the end of women's reproductive life span and that in a given year women have had at most one birth (Moultrie et al., 2013: 74).

Finally, the Projected Parity Progression Ratios (PPPRs) defined for one-year age intervals are as follows:

$$PPPR_x\left(i\right)=\frac{M_x^*\left(i+1\right)}{M_x^*\left(i\right)}. \tag{14}$$

These measures give the proportions of women who are expected to progress from parity i to parity $i + 1$ by the end of their childbearing years.

Two main assumptions are underlying the projected PPR method. According to Moultrie et al. (2013), women are assumed to have had at most one birth in a given year. The second assumption is that thus derived AOSFRs will remain constant in the future throughout the women's childbearing period.

3. Comparison of fertility patterns in European countries – Analysis based on Projected Parity Progression Ratios

3.1. Data, calculations, graphical illustration

Input data were sourced from the Human Fertility Database. The choice of countries was determined by TFR size, data availability, and geographical location. Data were not available for all European countries. Eight countries were selected for analysis. These are Belarus, Croatia, Hungary, Poland, Portugal, Spain, Sweden, and Switzerland. The reference year is 2016.

Period Specific Fertility Rates, Average Parity, and Total Fertility Rates

The period level of fertility is commonly measured by the total fertility rate (TFR), which gives the average number of children per woman. The global average fertility rate is just below 2.5 children per woman today. Over the last 50 years, the global fertility rate has halved.

More recently, according to the United Nations, the period total fertility rate in Europe reached 1.61 live births per woman in 2016. The lowest fertility rates were observed in Portugal (1.28 births per woman), Greece (1.31), Spain (1.32), Cyprus (1.34), Italy (1.36), and Poland (1.39). Conversely, France (1.89 births per woman) was the European country with the highest total fertility rate, followed by Sweden (1.87), Ireland (1.88), Russia (1.8), the United Kingdom (1.79), and Belarus (1.7).

The following figure illustrates this measure for the countries considered, Europe, and the World in 2016.

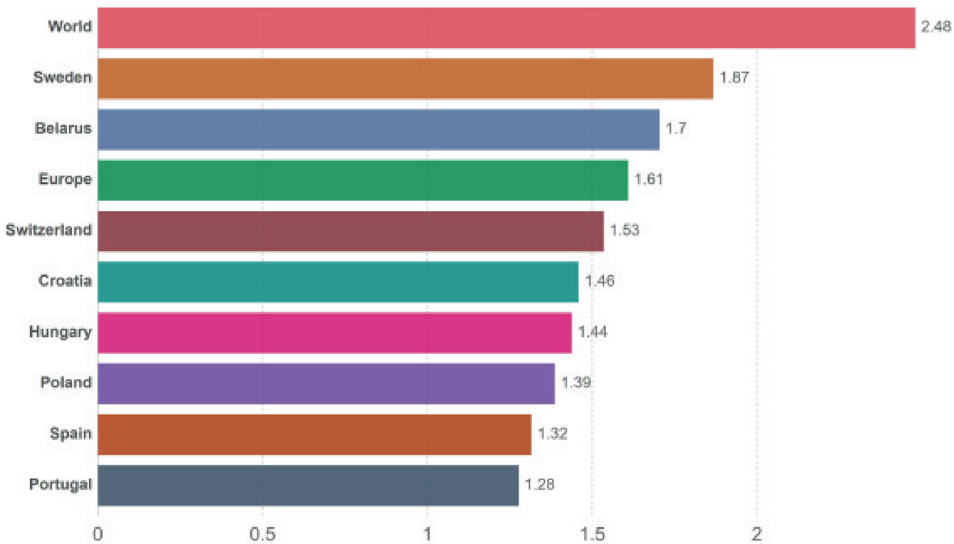


Figure 1. Total Fertility Rates in 2016

Source: Our World in Data

It is known that TFR does not give an accurate picture of the effects of fertility in family formation because it ignores the birth sequence, i.e. the fact that only women who are currently at parity zero, one, or two are at risk of having their first, second, or third child respectively. Thus, the same level of total fertility can be achieved by combining quite different distributions of birth by parity. Consequently, TFR does not inform about the influence of fertility on family formation.

In this analysis, the distribution of parities was estimated for women who had almost reached their reproductive age (40–45 years). Countries were sorted according to their TFR values. The differences in the distribution of birth parities between countries in 2016 are shown in Figure 2. It clearly illustrates the dissimilarities between countries in terms of the share of childless women as well as the varying weight of the first and the third child.

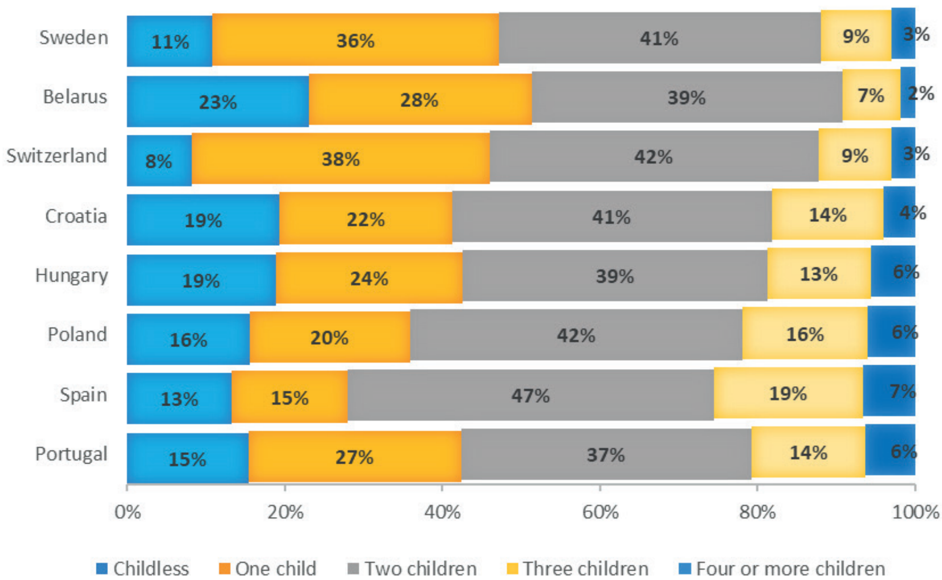


Figure 2. Women's birth parity distribution (age group 40–45) of European countries in 2016

Source: the author's own elaboration; The Human Fertility Database

Based on (3), we calculated the age-specific fertility rates ($ASFR_x$) and the age-order specific fertility rates ($AOSFR_x(i)$) for women aged between x and $x + 1$ and of parity $i = 1, 2, 3, 4$ for our countries. Age-specific fertility rates enable the analysis of the pattern of fertility by age of women, and the analysis of changes in the timing of childbearing. The calculation results are presented graphically below.

In Figure 3, we see that age-specific fertility rates vary greatly from one country to another. In most of the countries considered, the highest values of ASFR are attributed to older age groups. Only in Belarus, we notice a significant shift of the graph to the left, which means that in this country the highest ASFR values correspond to younger age groups. The uncharacteristic shape of the graph can be observed in Hungary, as ASFR values are higher for young women compared to those in the other analysed countries. Sweden, Belarus, and Switzerland have the highest rates, while fertility in the other countries barely reaches 0.1, even at the peak age of fertility.

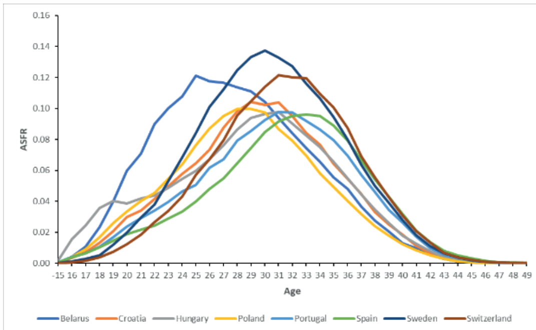


Figure 3. Age-Specific Fertility Rates, 2016

Source: the author's own elaboration; The Human Fertility Database

Then, we have a look at the AOSFR(*i*) charts for *i* = 1, 2, 3, 4. In most of the analysed countries, there is a tendency to delay the birth of children. This is different in Belarus, here the coefficients take the highest values for *i* = 1, and additionally for younger women than in other countries. There is also a decrease in the number of newborn children, and in some countries, such as Spain, Poland and Portugal, this decrease is very pronounced.

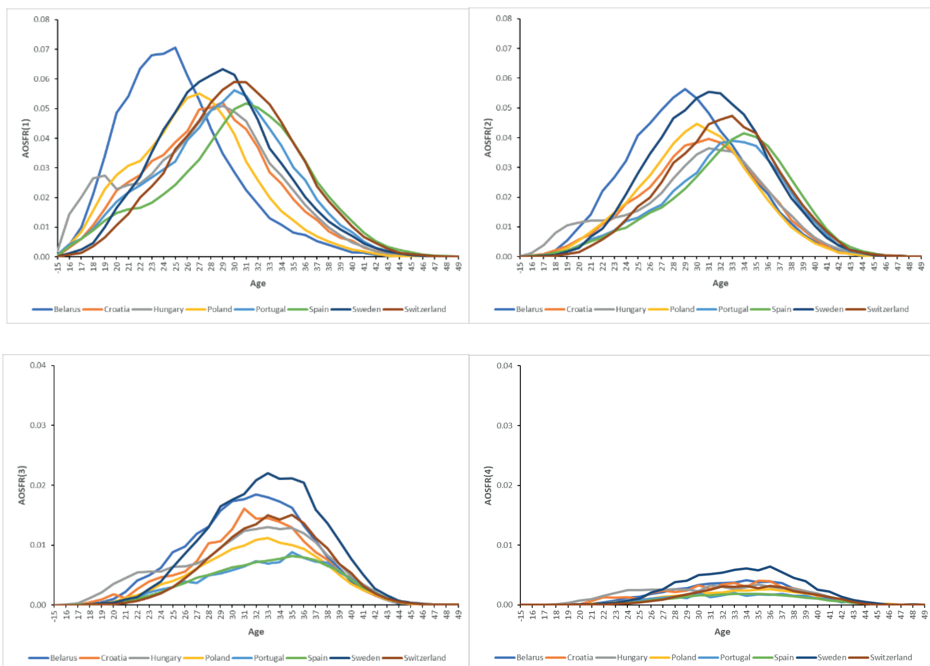


Figure 4. Age-Order Specific Fertility Rates

Source: the author's own elaboration; The Human Fertility Database

In the next step of the analysis, we computed the average parity of women of a given age *x*, P_x, for eight countries. Those were calculated according to (7) on the empirical data for selected countries from 2016 by dividing the total number of children ever born to women aged *x* in the reference year 2016 by the number of women aged *x* that year. Parameters P_x showed a typical pattern of average parities increasing with age, which we may see on the graph below. Compared to the countries surveyed, Spain has the lowest average parity, while Sweden has the highest. Belarus is characterised by higher values for young women.

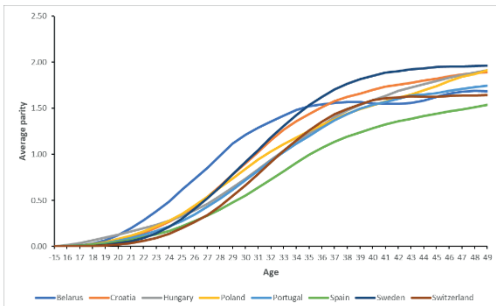


Figure 5. Average Parity, 2016

Source: the author's own elaboration; The Human Fertility Database

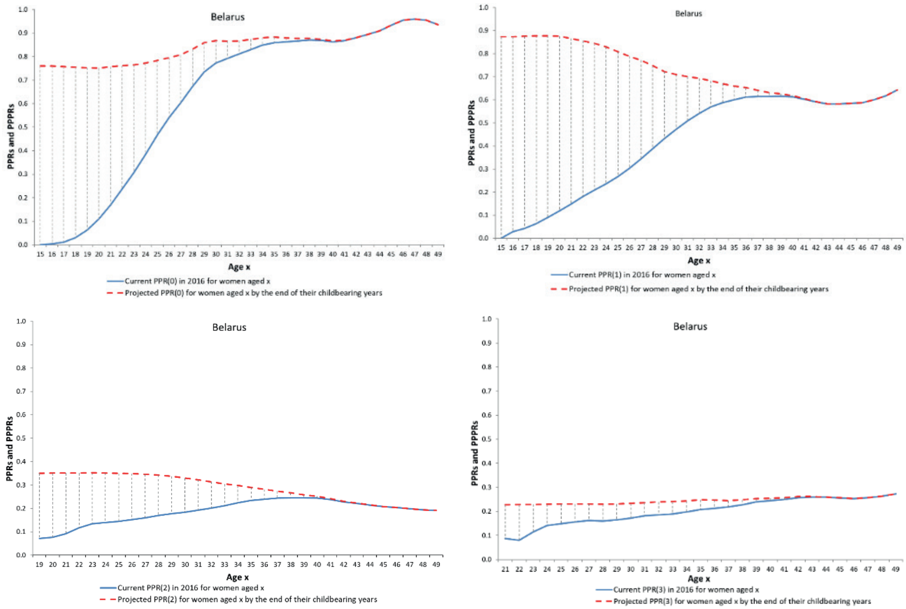
Projected and observed parity progression ratios

To explore parity progression ratios, we calculated proportions of women $M_x(i)$ aged between x and $x + 1$ having attained parity-order i or higher using formula (9). Finally, from (11), we obtained parity progression ratios $PPR_x(i)$.

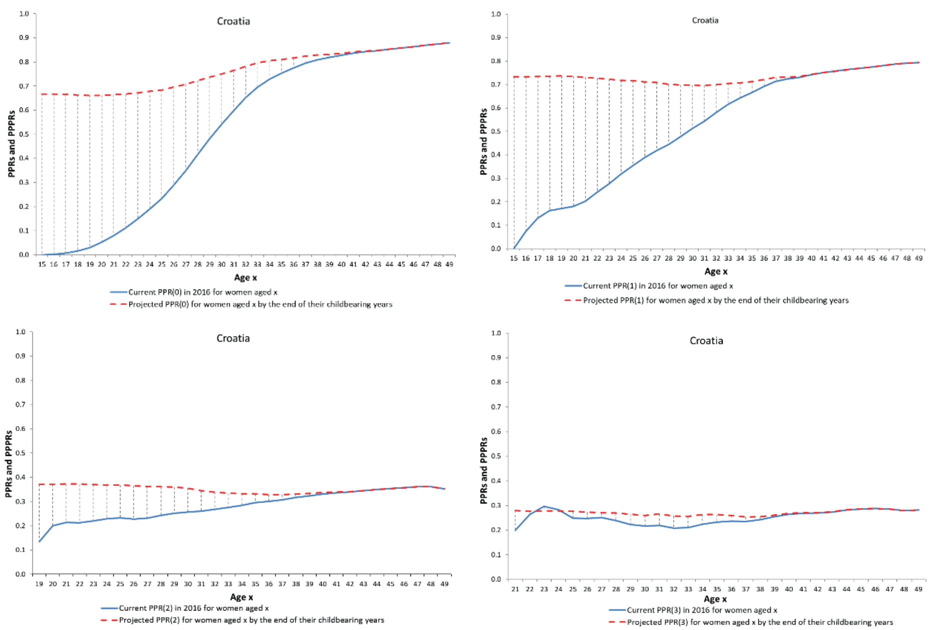
In the last step, we performed some additional calculations leading to projected parity progression ratios. First, cumulated age-order fertility rates $TOFR_x(i)$ must be computed. Next, we obtained the projected proportions of women $M_x^*(i)$ aged $[x, x + 1)$ who will attain at least parity i by the end of their childbearing years according to formula (13). Finally, the projected parity progression ratios between parity i and $i + 1$ were computed using formula (14). We will compare the projected (completed) and observed (uncompleted) *PPRs* by age and parity for every country considered. We present our results graphically.

In this part of the work, we will look more closely at the differences between the selected European countries by illustrating the projected and observed PPRs by age and parity. We can see typical shapes of observed (uncompleted) PPRs with curves increasing with age whereas projected PPRs tend to level. Both observed and projected PPRs decrease as parity increases. As expected, there are considerable differences between both types of PPRs, especially for young women, although differences vanish as they get older. Projected and observed parity progression ratios for $i = 0, 1, 2, 3$ are presented graphically below.

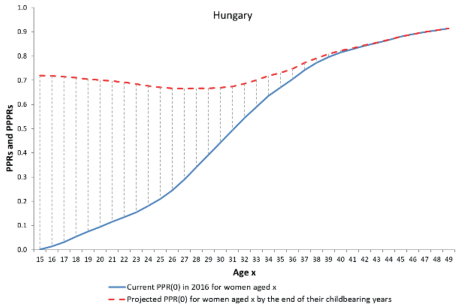
Belarus

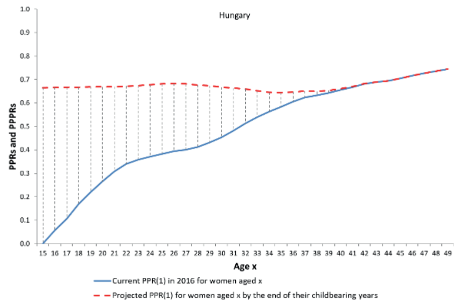


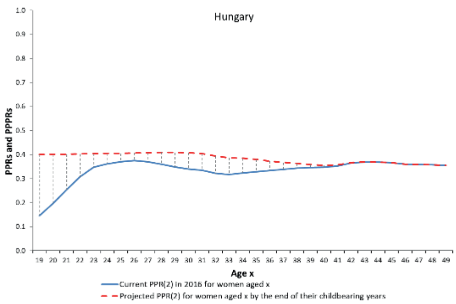
Croatia

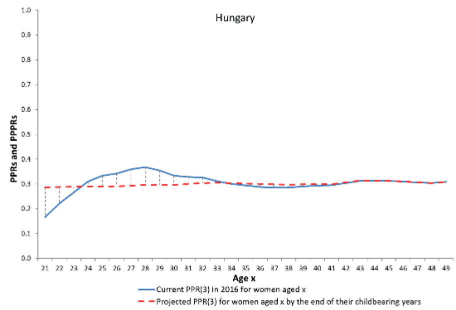


Hungary

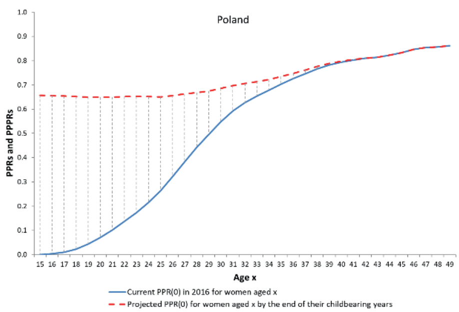


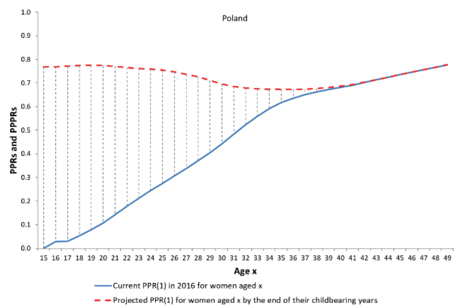


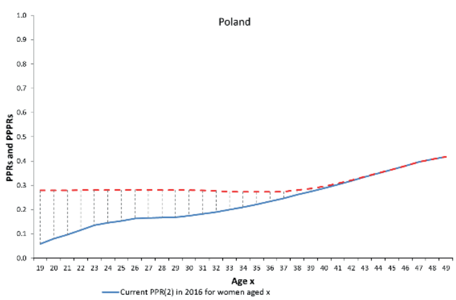


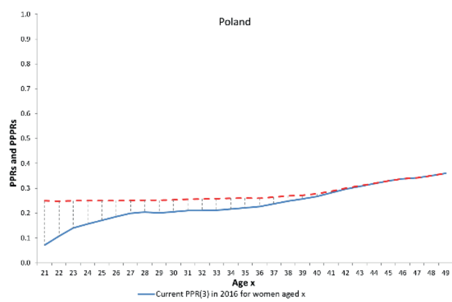


Poland

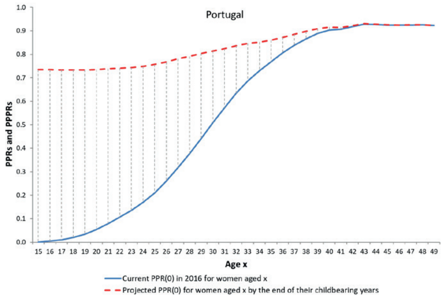


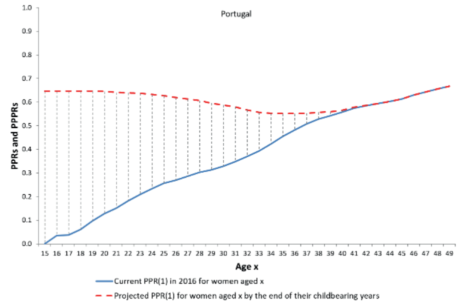


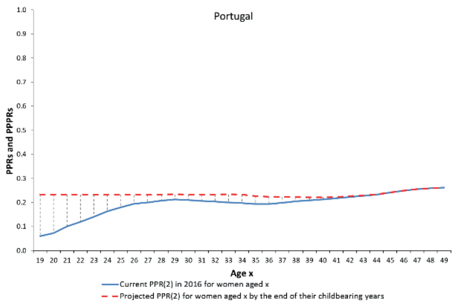


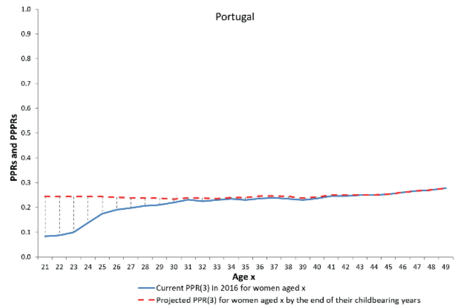


Portugal

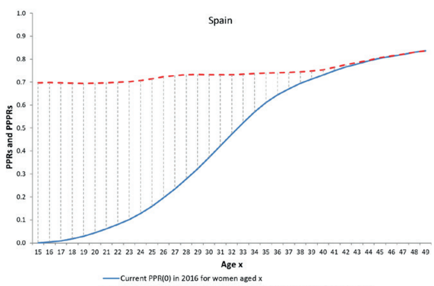


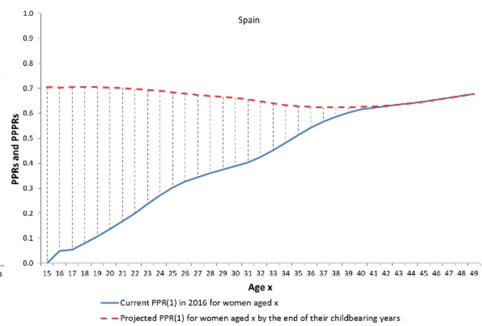


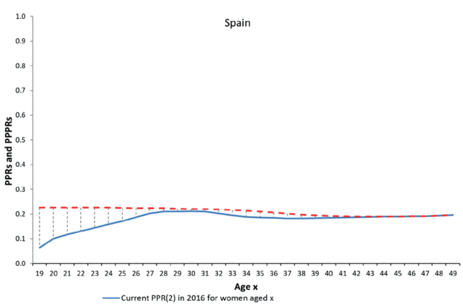


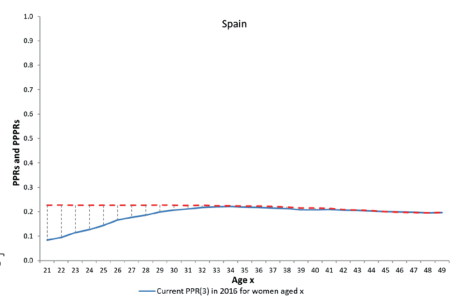


Spain

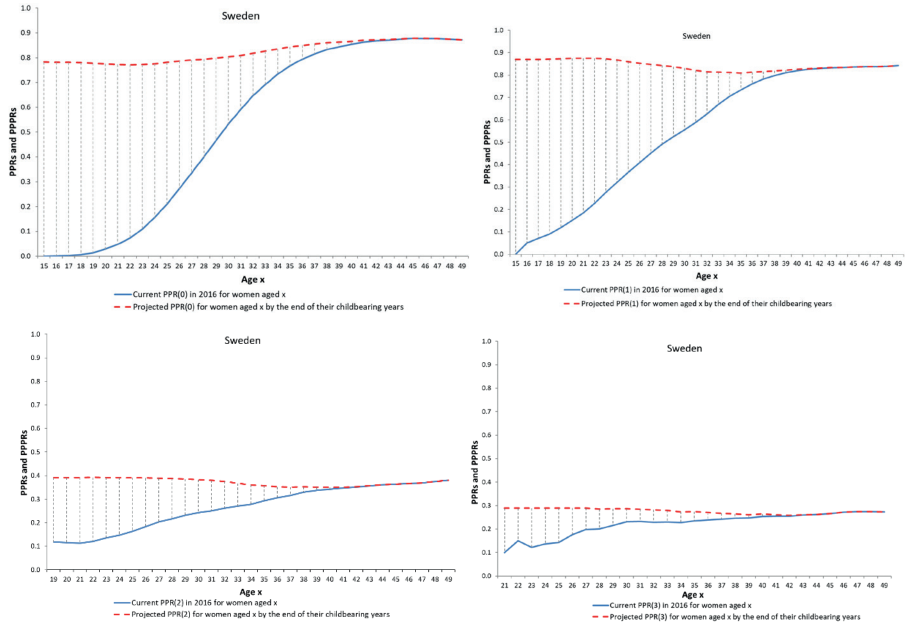








Sweden



Switzerland

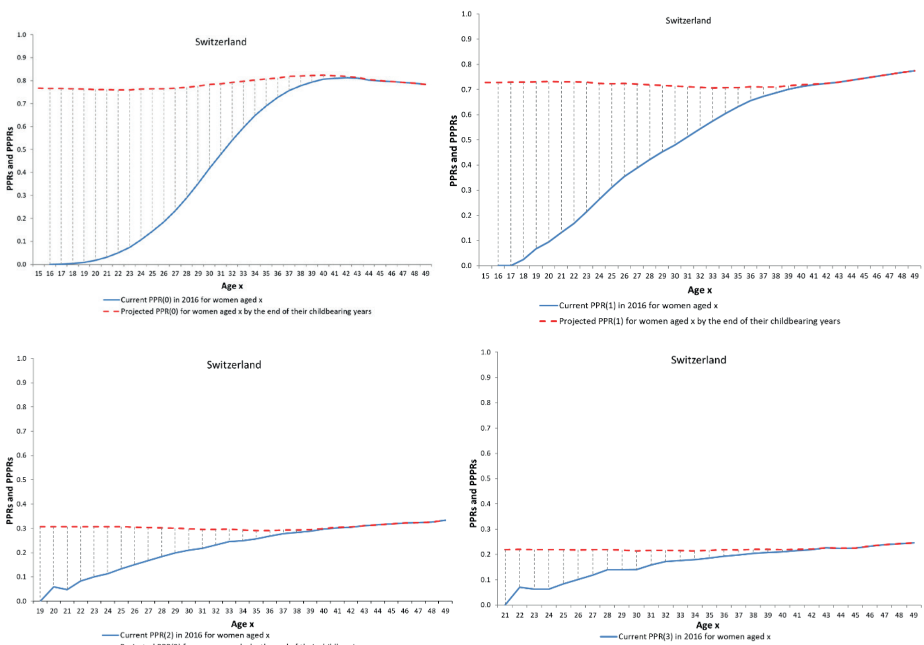


Figure 6. Projected and observed parity progression ratios

Source: the author's own elaboration; The Human Fertility Database

Comparing the individual graphs for the countries selected, we can see how considerable the differences between the values of the projected and observed PPRs are for parities $i = 0, 1, 2, 3$, respectively.

The following figures illustrate the graphs of PPRs for all the countries considered, depending on the parity.

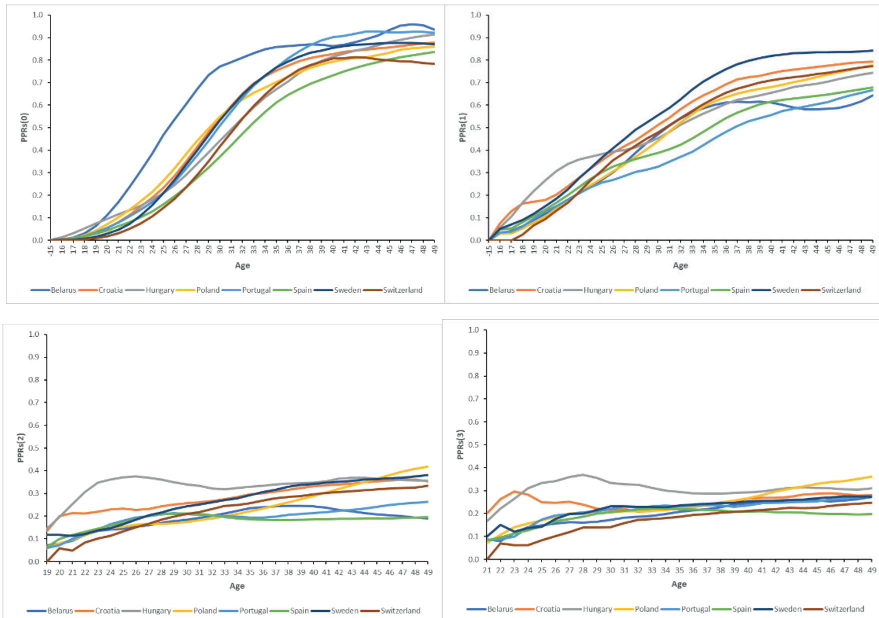


Figure 7. Parity Progression Ratios PPR(i) for $i = 0, 1, 2, 3$

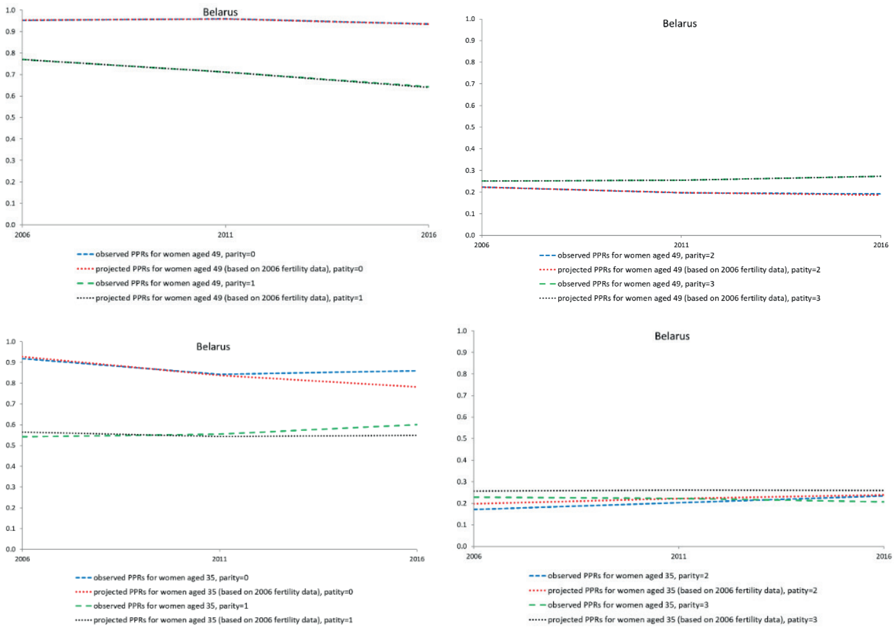
Source: the author's own elaboration; The Human Fertility Database

These parity progression ratios are calculated for all women. However, in this calculation, there are young women who have not yet completed their childbearing years. To obtain their parity progression ratios when they reach the end of childbearing, we need to calculate projected parity progression ratios.

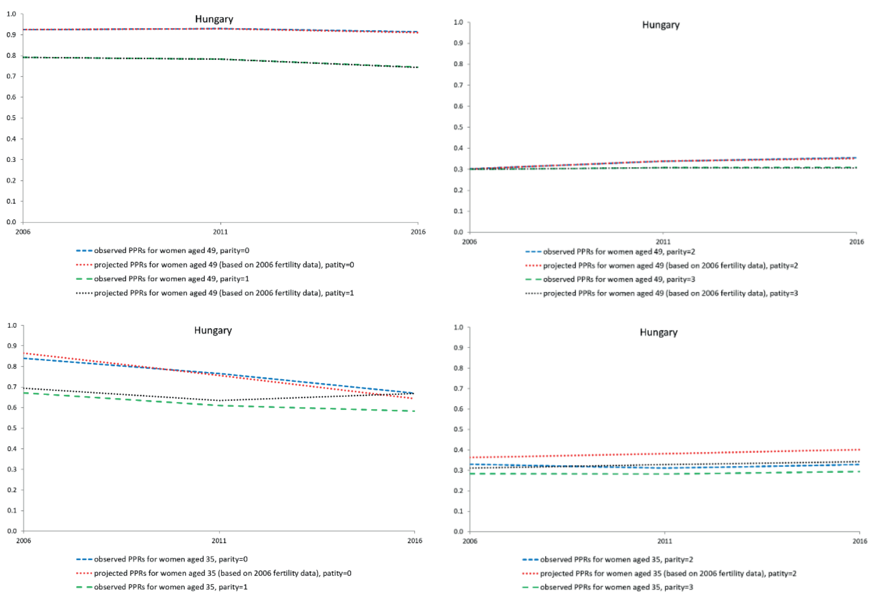
Prediction

The aim of this section is to see if the projected parity progression ratios give a good prediction of the completed parity progression ratios of younger women. To do this, the parity progression ratios for younger women in 2006 were projected forwards by ten years and then compared to the observed parity progression ratios in 2016. Figure 6 shows a comparison of the projected ratios for parities $i = 0, 1, 2, 3$ (predictions based on the 2006 fertility data) with the observed completed ratios for women aged 49 in the years 2006, 2011, and 2016 for the countries concerned. Analogous comparisons are made between the projected parity progression ratios (predictions based on 2006 fertility data) and the observed parity progression ratios for women aged 35 and of parities $i = 0, 1, 2, 3$ in the years 2006, 2011, and 2016. We could not make this comparison for Croatia because data from 2006 are not available.

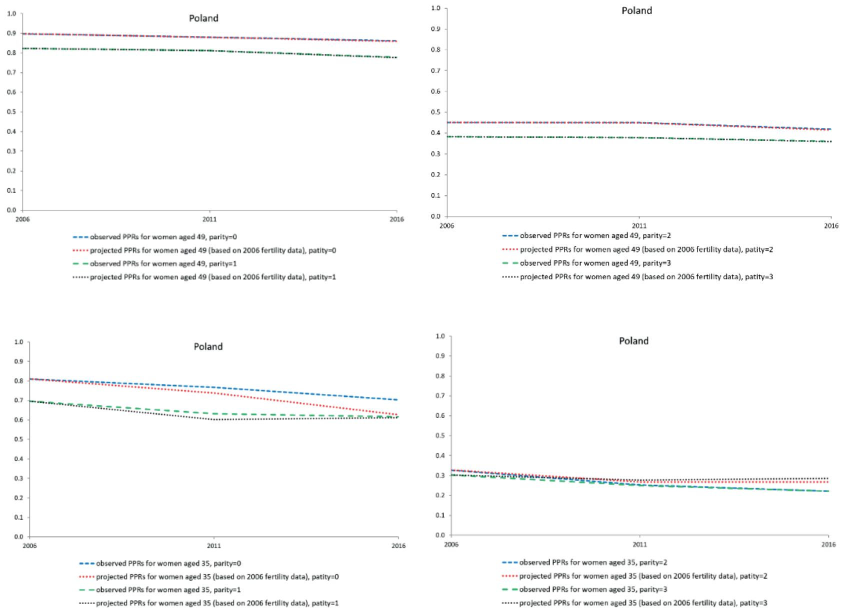
Belarus



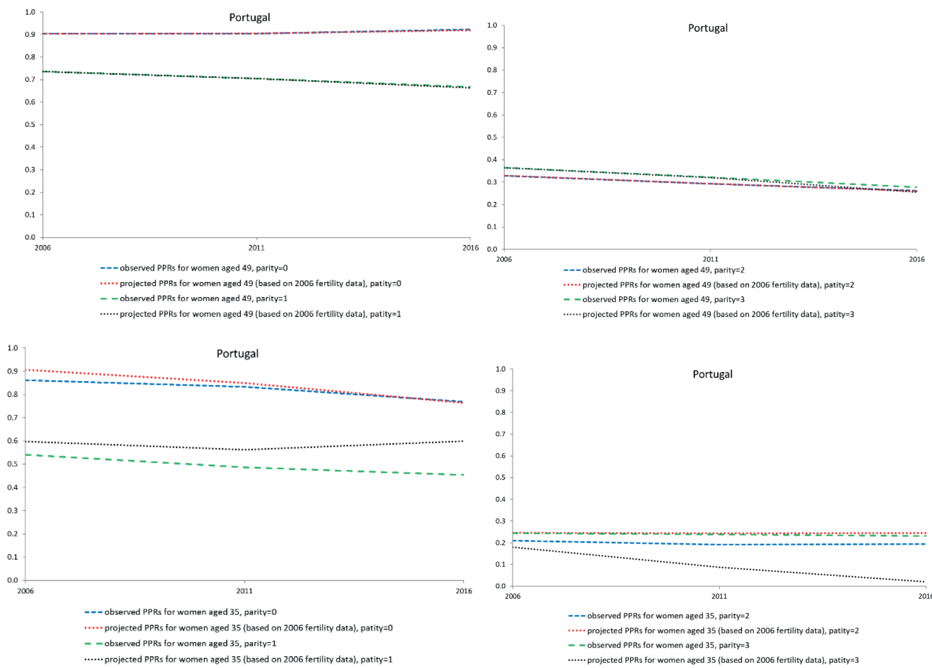
Hungary



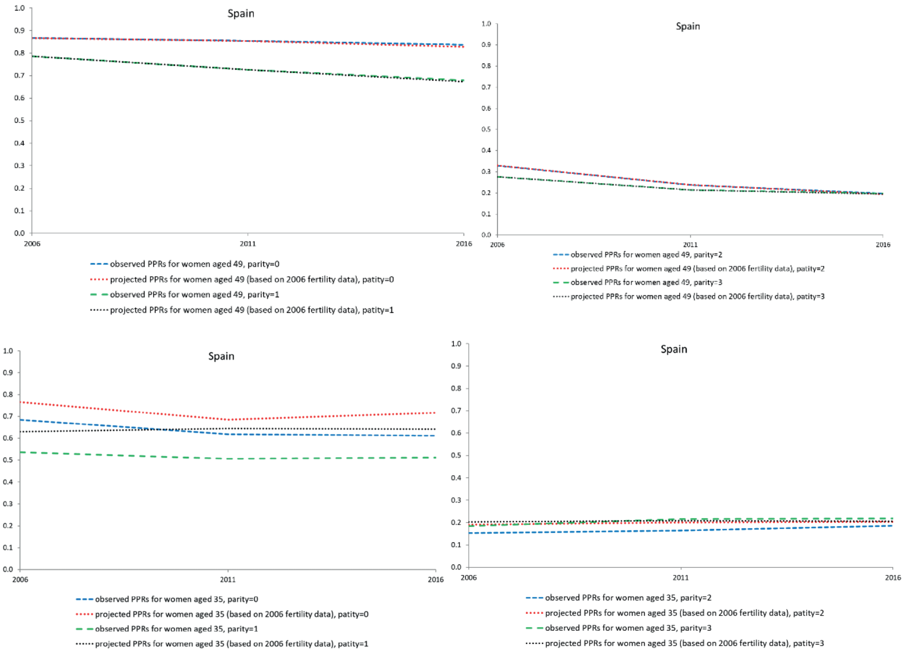
Poland



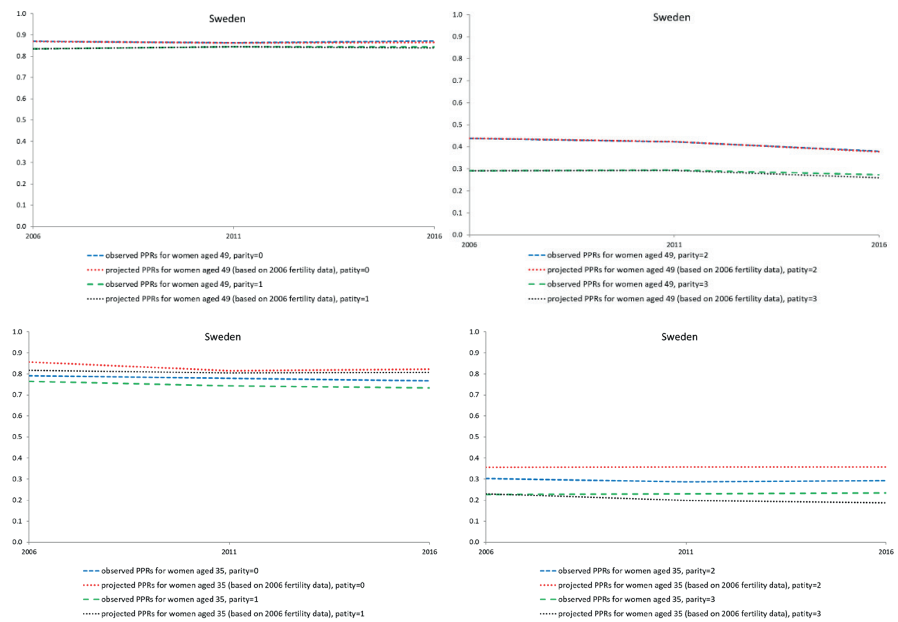
Portugal



Spain



Sweden



Switzerland

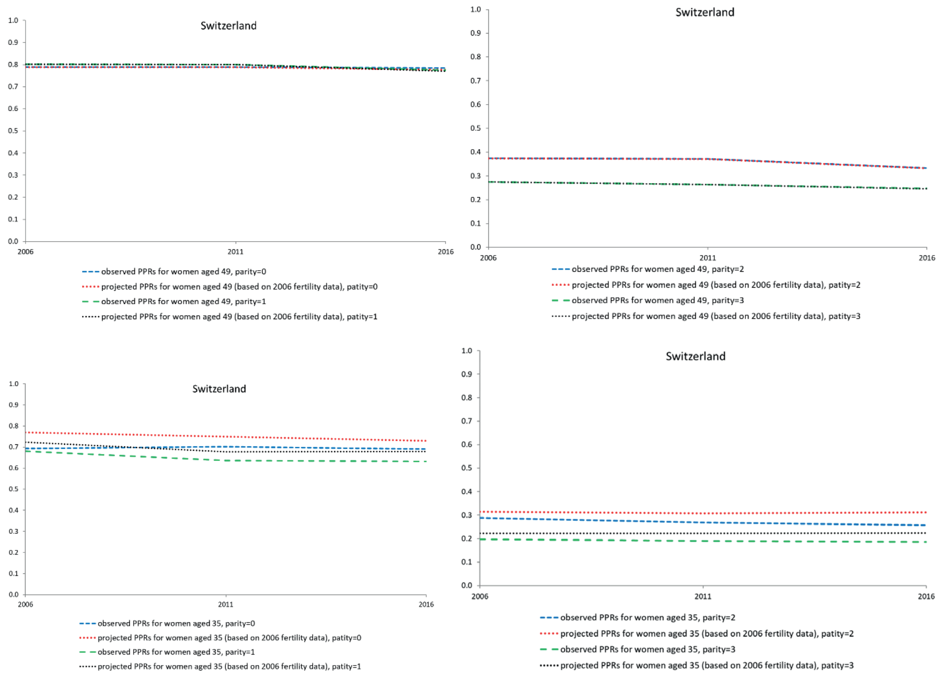


Figure 8. Projected and observed (completed) PPRs for women aged 49 and 35, parity $i = 0, 1, 2, 3$

Source: the author's own elaboration; The Human Fertility Database

It turns out that the projected PPR values are almost identical to the observed rates for women aged 49 in 2011 or 2016, correspondingly. On the other hand, comparing the projected and observed PPRs for women reaching age 35 in 2011 or 2016, respectively, we can find (cf. Figure 8) that the prognoses show some deviations for almost all the countries considered.

4. Conclusions

Parity progressive ratios give the probability that a woman of a given birth order ever proceeds to the next order. They have acquired a dominant place in fertility research. They have gained in importance as useful measures of fertility, especially when we want to compare the reproductive performance of two or more populations with a similar pattern of spacing but different desired family sizes.

In this work, we focused on the description and analysis of progressive fertility measures for the 2016 European female populations and we investigated the tempo and pattern of fertility in European countries. Then we used the projected parity progression ratio method to estimate the future trends of parity progression ratios. The objective of this study is to assess how well the projected parity progression ratio method works. Besides, traditional measures of fertility were also mentioned, age-specific fertility rates (ASFR), age-order specific fertility rates (AOSFR) and total fertility rates, which use age as the main structural feature of the female population. Based on the data from the HFD, we computed ASFR, AOSFR, PPR and PPPR for eight countries, next we presented them graphically. The results show that the projected parity progression ratio method produces a relatively good fit. In conclusion, it can be said that the PPPR method is useful in predicting the future trends of parity.

The fertility trends in the countries analysed are the result of many factors. The analysis shows that late decisions to have the first child and the tendency of women to have fewer children contribute to the fertility decline. In all countries, there is no problem with people wanting to have one child. The results indicate that most people also consider having a second child. In contrast, the situation is much worse when it comes to having further children. For example, the projected parity progression rates for Polish young women at the end of the reproductive period are expected to be about 20–40% lower than the completed PPR recorded in 2016.

References

Aoun S., Airey P. (1988), *Illustrative application of the use of projected parity progression ratios for the analysis of fertility*, London School of Hygiene Tropical Medicine, Centre for Population Studies, London.

Bhardwaj S.B., Sharma G.C., Kumar A. (2010), *Analysis of the Parity Progression Ratios*, "Journal of Reliability and Statistical Studies", vol. 3, no. 1, pp. 37–41.

Billari F.C., Kohler H. (2004), *Patterns of low and very low fertility in Europe*, "Population Studies", vol. 58, no. 2, pp. 161–176.

Brass W. (1985), *Advances in Methods for Estimating Fertility and Mortality from Limited and Defective Data*, London School of Hygiene and Tropical Medicine, Centre for Population Studies, London.

Brass W., Juarez F. (1983), *Censored cohort parity progression ratios from birth histories*, "Asian Pacific Census Forum", vol. 10, no. 1, pp. 5–13.

Feeney G. (1983), *Population Dynamics Based on Birth Intervals and Parity Progression*, "Population Studies", vol. 37, pp. 75–89.

Feeney G., Jingyuan Y. (1987), *Period parity progression measures of fertility in China*, "Population Studies", vol. 41, pp. 77–102.

Frejka T. (2008), *Parity distribution and completed family size in Europe Incipient decline of the two-child family model*, "Demographic Research", vol. 19(14), pp. 4–72.

Henry L. (1980), *Fertility of marriages: A new method of measurement*, "Population Studies Translation Series", no. 3, United Nations (French Edition Published in 1953).

Islam M., Yadava C. (1997), *On the estimation of parity progression ratio*, "Sankhya", vol. 58, series B, pp. 200–208.

Kohler H., Billari F.C., Ortega J.A. (2002), *The emergence of lowest-low fertility in Europe during the 1990s*, "Population and Development Review", vol. 28(4), pp. 641–680.

Moultrie T., Dorrington R., Hill A., Hill K., Timaus I., Zaba B. (2013), *Tools for demographic estimation*, International Union for the Scientific Study of Population, Paris.

Our World in Data, https://ourworldindata.org [accessed: 14.10.2020].

Preston S.H., Heuveline P., Guillot M. (2001), *Demography: Measuring and Modelling Population Processes*, Blackwell, Oxford.

Rodriguez G., Hobcraft J. (1980), *Illustrative analysis: Life table analysis of birth intervals in Colombia*, WFS Scientific Reports 16.

Rossa A., Palma A. (2020), *Predicting parity progression ratios for young women by the end of their childbearing life*, "Statistics in Transition New Series", vol. 21, no. 1, pp. 55–71.

Sloggett A., Brass W., Eldridge S. M., Timæus I.M. (1994), *Estimation of Demographic Parameters from Census Data*, Statistical Institute for Asia and the Pacific, Tokyo.

The Human Fertility Database, http://www.humanfertility.org [accessed: 14.10.2020].

Yadava R., Bhattcharya M. (1985), *Estimation of parity progression ratios from closed and open birth interval data*, Mimeo, Centre of Population Studies, Banaras Hindu University, Varanasi.

Yadava R., Kumar A. (2011), *On the estimation of parity progression ratios*, "Journal of Scientific Research", vol. 55, pp. 127–134.

Yadava R., Pandey A., Saxena N. (1992), *Estimation of parity progression ratios from the truncated distribution of closed and open birth intervals*, "Mathematical Biosciences", vol. 110, pp. 181–190.

Trendy płodności w krajach europejskich

Streszczenie: Analiza płodności jest w demografii tematem, który przyciąga uwagę wielu badaczy. Demografowie opracowali wiele różnych metod pomiaru tego zjawiska. Większość badań wykorzystuje tradycyjne miary. Nie dają one jednak szczegółowych informacji na temat tego, jak dokładnie zmienia się płodność w zależności od kolejności urodzenia. Głównym celem jest zbadanie tempa i wzorca płodności oraz zróżnicowania postępów w zakresie kolejności urodzeń w wybranych krajach europejskich.

Metoda zapoczątkowana przez Brassa pozwala oszacować progresywne współczynniki kolejności urodzeń. W badaniu tym dokonujemy porównań między obserwowanymi i przewidywanymi współczynnikami.

Wyniki pokazują, że metoda prognozowanych współczynników kolejności urodzeń daje stosunkowo dobre dopasowanie i jest przydatna do przewidywania przyszłych trendów płodności. Spadek dzietności w rozważanych krajach europejskich jest spowodowany późnymi decyzjami kobiet o posiadaniu pierwszego dziecka oraz ich tendencją do rodzenia mniejszej liczby dzieci.

Słowa kluczowe: płodność, współczynniki dzietności, dzietność kobiet, cząstkowe współczynniki płodności, wskaźnik przeciętnej kolejności urodzenia, progresywne współczynniki kolejności urodzenia

JEL: J11, J12, J13

OPEN ACCESS	
	Received: 2020-12-03; verified: 2021-02-02. Accepted: 2021-04-14
C O P E Member since 2018 JM13703	This journal adheres to the COPE's Core Practices https://publicationethics.org/core-practices

Folia Oeconomica
Acta Universitatis Lodziensis
ISSN 0208-6018 e-ISSN 2353-7663
www.czasopisma.uni.lodz.pl/foe/
2(353) 2021

DOI: https://doi.org/10.18778/0208-6018.353.02

Aneta Piechaczek
AGH University of Science and Technology, Faculty of Management
Independent Laboratory of Applied Mathematics in Economics
aneta.bech@gmail.com

Polish Domestic Tourism in the Face of SARS-CoV–2 Pandemic

Abstract: The paper focuses on the tourism market in the face of the SARS-CoV–2 pandemic. Its aim is to verify whether Polish tourism in this period actually recorded such a significant drop in interest. The research is based on secondary data from the reports of Statistics Poland concerning the use of accommodation facilities in Poland by both Poles and foreign tourists. Accommodation services offered by accommodation facilities constitute one of the basic tourist services included in the tourist market. Verifying their situation in the holiday season 2020 is therefore one of the easiest ways to estimate the impact of the pandemic on the tourism market in Poland. The paper also includes a comparison of data on Polish tourists in 2020 in relation to previous years. The research was conducted with the use of ANOVA analysis of variance.

Keywords: tourism, economic sectors, COVID–19, SARS-CoV–2, pandemic

JEL: Z31, Z32, Z39, L83

1. Introduction

The year 2020 brought many changes to the global economy. The SARS-CoV–2 virus pandemic, colloquially known as the coronavirus or COVID–19 pandemic, which arrived in Poland on 4th March, 2020, has affected the daily lives of people in almost the entire world. Its effects are also observed in world financial markets. The Asian Development Bank estimates the economic losses caused by the coronavirus at $3–6 trillion (Kornith, Ranasinghe, 2020: 987–990). Tourism is undoubtedly one of the sectors of the economy that is most affected by the SARS-CoV–2 virus outbreak. So far defined as one of the most important branches of the economy, generating annually 10% of the world's gross product (Widomski, 2020: 771–779), the tourism industry is currently facing a huge challenge. Lockdowns introduced by many countries, including Poland, limit the possibilities of moving not only beyond the borders of a given country but also within them. The first restrictions introduced in Poland on 15th March (ISAP, 2020) resulted in a drastic drop in the number of trips made by Poles in the following months. However, the situation improved somewhat in the summer when the government, observing a decline in the number of positive cases, gradually lifted the previously introduced restrictions. Attempts to improve the situation of Polish tourism were undertaken by the Ministry of Labour, Development and Technology by establishing the so-called *tourist voucher*.[1] By dint of such action, Polish tourism was supposed to bounce back from the bottom. Even though many Poles decided to travel in the summer, the numbers reported by Statistics Poland (GUS, 2020b) did not equal those observed in previous years. In consequence, the industry, which had been characterised by a steady growing trend in the last several years, had to face considerable losses. Furthermore, the second wave of COVID–19, which reached Poland in September 2020, did not improve the situation. The Polish economy, including tourism, had to face other effects of SARS-CoV–2, including the re-introduction of restrictions and the fear caused by an increasing number of infected citizens. However, a preliminary analysis of this situation raises some important questions. First of all, did the pandemic actually result in a significant drop in interest in tourism among Poles in the first half of 2020? How does it compare with previous years? Did lifting the restrictions and the government's strategy improve the situation of Polish tourism in the summer? Finally, did the second wave of the pandemic significantly worsen the situation in this branch of the economy again? The purpose of this paper is to find the answers to the above-raised questions.

1 'A tourist voucher is a new form of support for Polish families in a situation where the economy is weakened by the COVID–19 pandemic. It amounts to PLN 500 for each child up to the age of 18 and one additional benefit in the form of a voucher supplement, in the amount of PLN 500 for children with a disability certificate. With the help of the voucher, you can make payments for hotel services or tourist events carried out by a tourist entrepreneur or public benefit organization in the country' (Gov.pl, 2020a).

2. Literature review

Since the beginning of March 2020, when COVID–19 reached Poland, several valuable articles analysing the situation of Polish tourism in the face of the pandemic have been published. The authors drew attention to the economic effects suffered by entrepreneurs in the tourism industry (Walas, Kruczek, 2020: 79–95) and the tourist preferences of Poles in that period (Widomski, 2020: 771–779). However, those papers focused only on large cities that rely mainly on foreign tourists. Therefore, Polish regional tourism, which, according to Statistics Poland, so far comprised 80% of domestic tourists, was not adequately discussed in the current research. The topic of travelling in the time of the pandemic and how cultural tourism is changing under the influence of a spreading virus was discussed by participants at Gniezno Forum of Cultural Tourism Experts (von Rohrscheidt, Plichta, 2020). However, they mainly focused on health safety during travels. The subject of the coronavirus impact as an immediate threat on the condition of the Polish tourism economy was described by Panasiuk (2020: 55–70). He concentrated on activities that may support tourism in the long term.

Considerations on the short-term and long-term effects of SARS-CoV–2 for the entire economy, including tourism, were additionally addressed by Wąsiński and Wnukowski, emphasising the essence of international cooperation in this difficult period (Wąsiński, Wnukowski, 2020: 1–2). On the other hand, Niewiadomski writes about the temporary deglobalisation process, which will allow tourism to revive again after the pandemic. He claims that if only the opportunities created by the temporary tourist stagnation are exploited, this branch has a chance for even greater development (Niewiadomski, 2020: 651–656). However, for the time being, this is just wishful thinking.

It is worth pointing out that the pandemic impact on tourism is not only considered by Polish but also foreign researchers. This problem was discussed in a special issue of Tourism Geographies. Nevertheless, most of the articles published there focus on speculations about the future of tourism in a post-pandemic reality. Authors indicate the need for transformation that the tourism industry will have to undergo (Benjamin, Dillette, Alderman, 2020: 476–483) and also the essence of its self-regeneration. The latter is understood by local support of tourism-related entities (Ateljevic, 2020: 467–475) and gradual lifting of travel bans (Hall, Scott, Gössling, 2020: 577–598). A local return to normal is expected to re-energise the economy. Moreover, the issue of returning to 'normality' in the sphere of international tourism is also raised in relation to the invention of the vaccine (Hall, Scott, Gössling, 2020: 577–598; Prideaux, Thompson, Pabel, 2020: 667–678). Same as Niewiadomski, the researchers write about the chances of tourism in the post-pandemic world (Brouder, 2020: 484–490; Higgins-Desbiolles, 2020: 610–623; Prideaux, Thompson, Pabel, 2020: 667–678). They also present the strategy of fighting the virus in the sphere of tourism, highlighting the importance of the government's actions in this area (Koh, 2020: 1015–1023; Yeh, 2020: 1–7). Gössling, Scott and Hall (2021: 1–20) suggest that travel bans, stay-at-home campaigns and border closures could be the reason for the drop in tourism. A small group of publications concentrate on the current state of tourism. These include the analysis of A. Carr concerning New Zealand tourism (Carr, 2020: 491–502) or V. Kumar's research describing the present state of tourism in India (Kumar, 2020: 179–185). Furthermore, the current influence of SARS-CoV–2 on the tourism of Nepal was described by N. Ulak (2020: 50–75). Additionally, M.R. Farzanegan et al. have proven a significant relationship between the decline in international tourism and the emerging cases of COVID–19 (Farzanegan et al. 2020: 1–6). Nonetheless, there are without doubt very few such articles compared to the number of papers examining the post-pandemic reality including also those written by A.A. Lew et al. (2020), U. Stankov, V. Filimonau and M.D. Vujičić (2020: 703–712) or S. Polyzos, A. Samitas and A.E. Spyridou (2020: 1–13).

Besides scientific studies, reports published by Statistics Poland are undoubtedly a valuable source of information on the impact of COVID–19 on the Polish tourism market. Table 1 presents the percentage decrease in the number of tourists for individual months of 2020 in relation to 2019.

Table 1. Percentage decrease in the number of tourists for individual months of 2020 in relation to 2019

Month	Percentage decrease in relation to 2019
March	65.0
April	96.5
May	88.1
June	62.7
July	33.2
August	25.7

Source: own elaboration[2]

Reports of Statistics Poland (Table 1) show that the largest drop in the number of tourists took place in April 2020, amounting to as much as 96.5% loss compared to 2019, which is unquestionably related to the travel ban introduced that month in Poland. On the other hand, the analyses of Statistics Poland have a key flaw – they only concern the quotations from 2020 in relation to 2019. It is worth remembering that tourism was on a long-lasting growing trend which achieved its peak in 2019. Therefore, these analyses do not provide a full view of the situation, as it should also be considered in relation to earlier years.

2 Based on the data contained in reports of Statistics Poland (GUS, 2020a).

3. Data and methods

Based on current knowledge, it seems reasonable to conduct the research on data that directly concern the activity of Polish domestic tourists in recent years. Such information was included in the reports published by Statistics Poland on the number of domestic tourists staying in tourist accommodation establishments with ten or more beds in Poland, aggregated monthly from January to September.[3] The period under consideration is seven consecutive years from 2014 to 2020.

One-way analysis of variance (one-way ANOVA) was chosen as the investigating method to compare the average number of Polish domestic tourists over the years. It was assumed that the observations constituted a set of functional data (analogous to the correlated biological or geological data) that had been collected independently and on that basis it was possible to use one-way ANOVA. The one-way ANOVA was applied here also due to the fact that the analysis was aimed only at the verification of how the analysed number of tourists had changed in relation to previous years, and not what additional impact the individual months for which the data were collected could have had. This choice was also justified by a relatively small sample size and the fact that the investigated time series was not complete (data from October, November and December were missing for the measurements to be a consistent annual representation). Collected observations can be treated as functional data, transformed into numbers and analysed performing classical ANOVA according to T. Mrkvička et al. (2020: 433). The ANOVA assumptions were verified using the Jarque-Bera test (normal residual distribution) and the Bartlett test (equality of variance in subgroups). The post-hoc verification was performed by Tukey's test.

Tukey's HSD test is based on the studentised range distribution (Benjamini, Brown, 2002: 1580). Unlike ANOVA, which shows only if results are significant overall, it will show exactly where differences lie. For measurement or analysis, the HSD for each pair of means was calculated by using the below presented formula (Nanda et al., 2021: 60):

$$HSD = \frac{M_i - M_j}{\sqrt{\frac{MS_w}{N}}},$$

where:

$M_i - M_j$ is the difference between the pair of means (to calculate this, M_i should be larger than M_j);

MS_w – is the Within Mean Square of the group;

N – is the number in the group.

The significance level was established as α = 5%.

3 Based on the data contained in statistical statements of Statistics Poland (GUS, 2020a).

4. Results

The study was divided into three stages related to the government's policy concerning counteracting the SARS-CoV–2 pandemic as well as to the holiday season in July and August. Phase one included data from January to June 2020. At that time, the coronavirus was slowly spreading throughout Poland as part of the first wave of infections. On 15th March, the government introduced the first lockdown, closing schools, restaurants, and shopping centres. Moreover, a domestic travel ban was imposed, and the country's borders were also closed. The restrictions were in force until mid-July, but first restrictions started to be lifted in May (Gov.pl, 2020b; Olszewska, 2020). Clearly rising numbers of tourists in that period can be observed for 2020 in Figure 1.

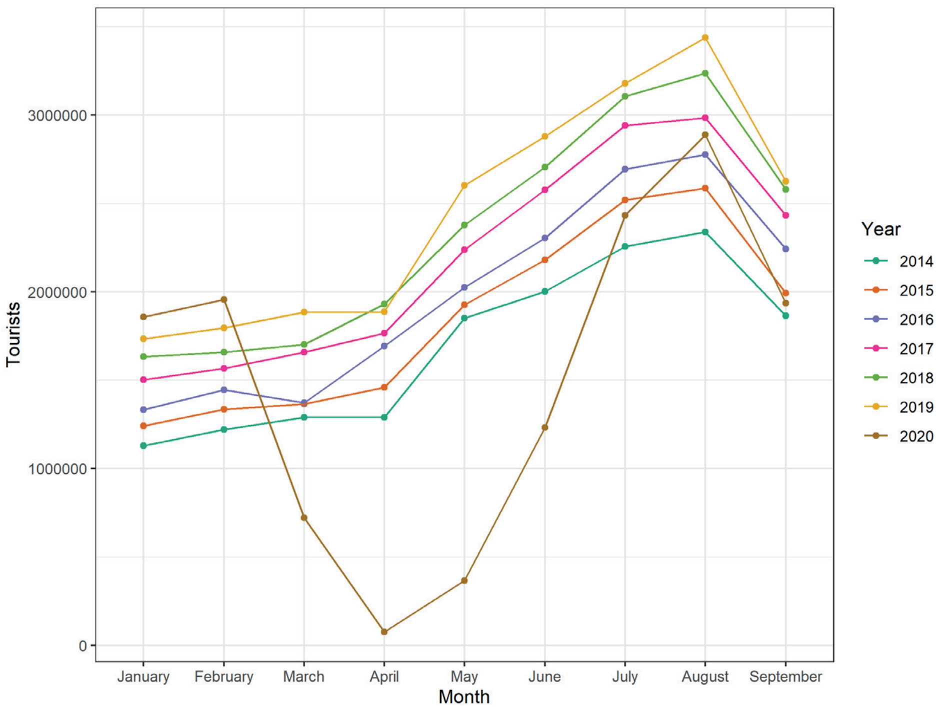


Figure 1. Number of domestic tourists staying in tourist accommodation establishments with ten or more beds in Poland in 2014–2020 by month

Source: own elaboration

The decreasing number of positive cases and the so-called *tourist voucher*, introduced in mid-July by the Ministry of Labour, Development and Technology, was meant to encourage Poles to travel again. The phase two sample was extended by these two summer months. According to the data presented in Table 1 and

Figure 1, it can be seen that the number of Polish citizens travelling during that period was slowly catching up to the numbers observed in previous years. Unfortunately, when looking at the data from September, it should be noted that the number of domestic tourists in Poland dropped again. Phase three of this study covered all months from January to September 2020, when the second wave of the pandemic reached Poland (Medonet.pl, 2020) and restrictions were re-introduced (Medexpress.pl, 2020).

4.1. Phase I

The first half of 2020 was characterised primarily by a slowdown in the economy due to the lockdown introduced in Poland. The impact of the travel ban is particularly visible for April 2020 (Figure 1). The decrease in the number of tourists was undoubtedly influenced by social moods, such as ostracism towards citizens from regions with more COVID–19 cases, as well as social anxiety. The significance of those differences in relation to the previous years was confirmed by the analysis of variance. ANOVA results presented in Table 2 unequivocally allow us to conclude that the average number of tourists in the individual years 2014–2020 was significantly different from each other at the significance level of $\alpha = 5\%$ (the verification of ANOVA assumptions is presented in Table 3). The post-hoc analysis done by Tukey's test (Table 4) shows that significant differences can be observed in the comparison of the average for the same periods of time for 2020 and 2019 (an average decrease of 1,096,160 tourists per month) and 2020 and 2018 (an average decrease of 966,195 tourists per month). Thus, the results of phase one analysis allow us to answer the first of the research questions stated in the introduction. The SARS-CoV–2 pandemic, through the first lockdown and because of the fear of a previously unknown virus, caused a significant decrease in the number of domestic tourists in Poland in the first half of 2020, not only referring to 2019, which was reported by Statistics Poland, but also in relation to 2018.

Table 2. ANOVA test results for data from the first half of 2020

F	P-value
3.474	0.0085

Source: own elaboration

Table 3. Results of tests verifying the ANOVA assumptions for data from the first half of 2020

Assumption	Test	Distribution	Statistics	P-value
Normality	Jarque-Bera	χ^2	1.8752	0.3916
Homogeneity of variance	Bartlett	K^2	5.5064	0.4807

Source: own elaboration

Table 4. Tukey's test results for data from the first half of 2020

	2020	2019	2018	2017	2016	2015
2019	–1,096,160.50					
	0.0068					
2018	–966,195.80	129,964.70				
	0.0230	0.9992				
2017	–849,886.70	246,273.80	116,309.20			
	0.0627	0.9736	0.9996			
2016	–660,720.30	435,440.20	305,475.50	189,166.30		
	0.2480	0.7124	0.9273	0.9932		
2015	–549,925.30	546,235.20	416,270.50	299,961.30	110,795.00	
	0.4568	0.4648	0.7523	0.9330	0.9997	
2014	–428,581.30	667,579.20	537,614.50	421,305.30	232,139.00	121,344.00
	0.7269	0.2375	0.4838	0.7420	0.9804	0.9994

Source: own elaboration

4.2. Phase II

In the next phase of the research, the set of observations was extended to include data from July and August. As it was summer holiday in Poland, the government lifted the previously introduced restrictions and additionally encouraged citizens to travel by offering a *tourist voucher*. Success of that strategy was confirmed by the results of the analysis of variance, included in Table 5. Assuming a significance level of 5%, we have no grounds to reject the null hypothesis of the equality of means in the compared years (p-value = 0.0641). The results of the verification of ANOVA assumptions are presented in Table 6. This means that after the collapse of tourism in the first half of 2020, two months of increased domestic tourist movement resulted in a significant improvement in the situation of tourism in Poland. Taking into account the two additional months, the average monthly number of domestic tourists for 2020 did not differ statistically from other analysed years. Moreover, tourists were not discouraged by the day-to-day increase in the number of coronavirus cases in August (a total of 21,684 cases of the disease in August, more than twice as many as in April – 10,566, when tourist movement practically stopped). This confirms the author's assumptions that the main reason for the decline in tourist movement was not only fear but the inability to travel due to restrictions and the government's guidelines.

Table 5. ANOVA test results for data including July and August 2020

F	P-value
2.15	0.0641

Source: own elaboration

Table 6. Results of tests verifying the ANOVA assumptions for data including July and August 2020

Assumption	Test	Distribution	Statistics	P-value
Normality	Jarque-Bera	χ^2	1.8752	0.3916
Homogeneity of variance	Bartlett	K^2	5.5064	0.4807

Source: own elaboration

4.3. Phase III

The summer period, which filled the tourist industry with slight optimism, had to give way to autumn when the second wave of the coronavirus pandemic reached Poland. The increasing number of positive cases (at the time record-breaking 1,587 cases reported on 25th September, 2020), the change of the Minister of Health, and thus the strategy of fighting the pandemic, influenced Polish society. Tourism started to collapse again, as shown in Figure 1. This is also confirmed by the ANOVA result (Table 7) for the sample with data from January to September (p-value < 5%). The test assumptions were verified on the basis of the results presented in Table 8. This time, Tukey's test (Table 9) shows a significant difference in the comparison only for the pair of 2019 and 2020 (in phase one, significant differences were observed in two pairs: 2019/2020 and 2018/2020). The beginnings of the second wave of the pandemic, despite the monthly number of COVID–19 cases exceeding 20,000 since August, did not bring such a drastic decline in tourist movement as it was the case during the first spring wave. The reasons for that outcome can be seen in the restrictions and the government's policy again (this time not as drastic as in April, as it did not take into account the travel ban).

Table 7. ANOVA test results for data including September 2020

F	P-value
2.617	0.0262

Source: own elaboration

Table 8. Results of tests verifying the ANOVA assumptions for data including September 2020

Assumption	Test	Distribution	Statistics	P-value
Normality	Jarque-Bera	χ^2	1.8752	0.3916
Homogeneity of variance	Bartlett	K^2	5.5064	0.4807

Source: own elaboration

Table 9. Tukey's test results for data including September 2020

	2020	2019	2018	2017	2016	2015
2019	–951,387.00					
	0.0385					
2018	–829,082.90	122,304.10				
	0.1042	0.9996				
2017	–688,972.10	262,414.90	140,110.80			
	0.2682	0.9754	0.9992			
2016	–491,197.60	460,189.40	337,885.30	197,774.60		
	0.6634	0.7271	0.9187	0.9944		
2015	–349,007.80	602,379.20	480,075.10	339,964.30	142,189.80	
	0.9063	0.4262	0.6866	0.9165	0.9991	
2014	–197,187.70	754,199.30	631,895.20	491,784.40	294,009.90	151,820.10
	0.9945	0.1777	0.3681	0.6621	0.9571	0.9987

Source: own elaboration

5. Discussion and conclusions

Taking into consideration the results of the analysis, it has been clearly confirmed that the SARS-CoV–2 virus pandemic has had a major impact on the tourism industry. The collapse of this branch of economy in the spring months (March, April and May) led to statistically significant differences in the number of domestic tourists not only in relation to 2019 but also to 2018. The main reasons for such ratings should not be seen in the fear of the unknown virus but in the restrictions (previously noted by Gössling et al.) and the government's policy and guidelines (already highlighted in Yeh's work). This is confirmed by the results of the other two phases of this study. Despite the number of new COVID–19 cases in Poland being twice as high in August as in April, the number of Poles travelling around the country significantly increased. A reasonable explanation for this situation could be only the lifting of restrictions and the government's assurance that the situation was under control, along with the *tourist voucher* as an incentive to travel. The same conclusions confirm the results for phase three. The change in the government's strategy, the reintroduction of restrictions, although weakened, as well as the information about the coronavirus return resulted in a statistically significant drop in the number of Polish domestic travellers. However, the differences were not as big as those observed in the first half of the year.

This paper fills the previously identified research gap regarding Polish domestic tourism during the pandemic, not only for large cities, which rely mainly on foreign tourists. It also opens up an opportunity to revise the topic once the data for the last three months of 2020 are available. The number of COVID–19 cases increased at the end of the year. In October, the proportion of positive coronavirus tests averaged 18%, rising to 40.5% in November. According to Statistics Poland, the highest percentage of positive SARS-CoV–2 tests (59%) was recorded on 16th November, 2020.[4] This allows us to assume that the crisis in the tourism industry will only deepen, especially taking into account the restrictions on accommodation establishments. As already mentioned by Hall et al. as well as Prideaux et al., only a vaccine is a real chance for tourism, as its rollout will lead to the lifting of the current restrictions and will result in changing the government's policy regarding travel.

References

Ateljevic I. (2020), *Transforming the (tourism) world for good and (re)generating the potential 'new normal'*, "Tourism Geographies", vol. 22, no. 3, pp. 467–475, https://doi.org/10.1080/14616688.2020.1759134

Benjamini Y., Braun H. (2002), *John W. Tukey's contributions to multiple comparisons*, "The Annals of Statistics", vol. 30, no. 6, pp. 1576–1594.

Benjamin S., Dillette A., Alderman D.H. (2020), *"We can't return to normal": committing to tourism equity in the post-pandemic age*, "Tourism Geographies", vol. 22, no. 3, pp. 476–483, https://doi.org/10.1080/14616688.2020.1759130

Brouder P, (2020), *Reset redux: possible evolutionary pathways towards the transformation of tourism in a COVID–19 world*, "Tourism Geographies", vol. 22, no. 3, pp. 484–490, https://doi.org/10.1080/14616688.2020.1760928

Carr A. (2020), *COVID–19, indigenous peoples and tourism: a view from New Zealand*, "Tourism Geographies", vol. 22, no. 3, pp. 491–502, https://doi.org/10.1080/14616688.2020.1768433

Farzanegan M.R., Gholipour H.F., Feizi M., Nunkoo R., Andargoli A.E. (2020), *International Tourism and Outbreak of Coronavirus (COVID–19): A Cross-Country Analysis*, "Journal of Travel Research", vol. 60(3), pp. 687–692, https://doi.org/10.1177/0047287520931593

Gov.pl (2020a), *Bon turystyczny*, https://www.gov.pl/web/rozwoj-praca-technologia/bonturystyczny [accessed: 2.12.2020].

Gov.pl (2020b), *Znosimy część ograniczeń związanych z COVID–19: więcej osób na trybunach, targach i konferencjach*, https://www.gov.pl/web/koronawirus/znosimy-czesc-ograniczen-zwiazanych-z-covid-19-wiecej-osob-na-trybunach-targach-i-konferencjach [accessed: 2.12.2020].

Gössling S., Scott D., Hall C.M. (2021), *Pandemics, tourism and global change: a rapid assessment of COVID–19*, "Journal of Sustainable Tourism", vol. 29, no. 1, pp. 1–20, https://doi.org/10.1080/09669582.2020.1758708

4 Own calculations based on the data of the Polish Ministry of Health found on the websites: https://koronawirusunas.pl (2020) and https://www.termedia.pl (2020).

GUS (2020a), *Turystyka*, https://stat.gov.pl/obszary-tematyczne/kultura-turystyka-sport/turystyka/ [accessed: 20.12.2020].

GUS (2020b), *Wykorzystanie turystycznej bazy noclegowej w Polsce w lipcu i sierpniu 2020 r.*, https://stat.gov.pl/obszary-tematyczne/kultura-turystyka-sport/turystyka/wykorzystanie-turystycznej-bazy-noclegowej-w-polsce-w-lipcu-i-sierpniu-2020-r-,6,21.html [accessed: 2.12.2020].

Hall C.M., Scott D., Gössling S. (2020), *Pandemics, transformations and tourism: be careful what you wish for*, "Tourism Geographies", vol. 22, no. 3, pp. 577–598, https://doi.org/10.1080/14616688.2020.1759131

Higgins-Desbiolles F. (2020), *Socialising tourism for social and ecological justice after COVID–19*, "Tourism Geographies", vol. 22, no. 3, pp. 610–623, https://doi.org/10.1080/14616688.2020.1757748

ISAP (2020), Rozporządzenie Ministra Spraw Wewnętrznych i Administracji z dnia 13 marca 2020 r. w sprawie czasowego zawieszenia lub ograniczenia ruchu granicznego na określonych przejściach granicznych, https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20200000435 [accessed: 2.12.2020].

Koh E. (2020), *The end of over-tourism? Opportunities in a post-Covid–19 world*, "International Journal of Tourism Cities", vol. 6, no. 4, pp. 1015–1023, https://doi.org/10.1108/IJTC-04-2020-0080

Kornith B., Ranasinghe R. (2020), *COVID–19 Pandemic's Impact on Tourism in Poland in March 2020*, "Geo Journal of Tourism and Geosites", vol. 31, no. 3, pp. 987–990, https://doi.org/10.30892/gtg.31308-531

Koronawirus.pl (2020), *Koronawirus w Polsce (SARS-CoV–2)*, https://koronawirusunas.pl/ [accessed: 2.12.2020].

Kumar V. (2020), *Indian Tourism Industry and COVID–19: Present Scenario*, "Journal of Tourism and Hospitality Education", vol. 10, pp. 179–185, https://doi.org/10.3126/jthe.v10i0.28768

Lew A.A., Cheer J.M., Haywood M., Brouder P., Salazar N.B. (2020), *Visions of travel and tourism after the global COVID–19 transformation of 2020*, "Tourism Geographies", vol. 22, no. 3, pp. 455–466, https://doi.org/10.1080/14616688.2020.1770326

Medexpress.pl, (2020), *Koronawirus: Wracają obostrzenia. Jakie?*, https://www.medexpress.pl/koronawirus-wracaja-obostrzenia-jakie/78546 [accessed: 2.12.2020].

Medonet.pl (2020), *Coraz więcej zakażeń koronawirusem w Europie. Kolejne państwa z rekordami*, https://www.medonet.pl/koronawirus/koronawirus-w-europie,coraz-wiecej-zakazen-koronawirusem-w-europie--kolejne-panstwa-z-rekordami,artykul,54777827.html [accessed: 2.12.2020].

Mrkvička T., Myllymäki M., Jîlek M., Hahn U. (2020), *A one-way ANOVA test for functional data with graphical interpretation*, "Kybernetika", vol. 56, no. 3, pp. 432–458.

Nanda A., Mohapatra B.B., Mahapatra A.P.K., Mahapatra A.P.K., Mahapatra A.P.K. (2021), *Multiple comparison test by Tukey's honestly significant difference (HSD): Do the confident level control type I error*, "International Journal of Statistics and Applied Mathematics", vol. 6, no. 1, pp. 59–65.

Niewiadomski P. (2020), *COVID–19: from temporary de-globalisation to a re-discovery of tourism?*, "Tourism Geographies", vol. 22, no. 3, pp. 651–656.

Olszewska A. (2020), *Koronawirus. Rząd znosi obostrzenia: otwarte restauracje i salony kosmetyczne, zmiany w edukacji*, https://www.gdansk.pl/wiadomosci/koronawirus-rzad-znosi-obostrzenia-otwarte-restauracje-i-salony-kosmetyczne-zmiany-w-edukacji,a,171017 [accessed: 2.12.2020].

Panasiuk A. (2020), *Przyczynek do badań nad wpływem pandemii na stan gospodarki turystycznej*, [in:] K. Nessel (ed.), *Turystyka w naukach społecznych*, vol. 3, Instytut Przedsiębiorczości Uniwersytetu Jagiellońskiego, Kraków, pp. 55–70.

Polyzos S., Samitas A., Spyridou A.E. (2020), *Tourism demand and the COVID–19 pandemic: an LSTM approach*, "Tourism Recreation Research", vol. 46(2), pp. 175–187, https://doi.org/10.1080/02508281.2020.1777053

Prideaux B., Thompson M., Pabel A. (2020), *Lessons from COVID–19 can prepare global tourism for the economic transformation needed to combat climate change*, "Tourism Geographies", vol. 22, no. 3, pp. 667–678, https://doi.org/10.1080/14616688.2020.1762117

Rohrscheidt A.M. von, Plichta P. (2020), *Turystyka w czasach zarazy. Jak współczesne epidemie zmieniają turystykę kulturową?*, "Turystyka Kulturowa", no. 2, pp. 161–184, https://ruj.uj.edu.pl/xmlui/bitstream/handle/item/244242/plichta_hoffmann_et_al_turystyka_w_czasach_zarazy_2020.pdf?sequence=1&isAllowed=y [accessed: 2.12.2020].

Stankov U., Filimonau V., Vujičić M.D. (2020), *A mindful shift: an opportunity for mindfulness-driven tourism in a post-pandemic world*, "Tourism Geographies", vol. 22, no. 3, pp. 703–712, https://doi.org/10.1080/14616688.2020.1768432

Termedia.pl (2020), *Ile dziennie wykonuje się testów na koronawirusa?*, https://www.termedia.pl/mz/Ile-dziennie-wykonuje-sie-testow-na-koronawirusa-,39038.html [accessed: 2.12.2020].

Ulak N. (2020), *COVID–19 Pandemic and its Impact on Tourism Industry in Nepal*, "Journal of Tourism & Adventure", vol. 3, no. 1, pp. 50–75, https://doi.org/10.3126/jota.v3i1.31356

Walas B., Kruczek Z. (2020), *The Impact of COVID–19 on Tourism in Kraków in the Eyes of Tourism Entrepreneurs*, "Studia Periegetica", vol. 2, no. 30, pp. 79–95.

Wąsiński M., Wnukowski D. (2020), *Skutki pandemii COVID–19 dla gospodarki światowe*, "Biuletyn PISM", vol. 84, pp. 1–2.

Widomski M. (2020), *Turystyka krajowa a pandemia*, „Poszerzamy Horyzonty", vol. 21, no. 1, pp. 771–779.

Yeh S.S. (2020), *Tourism recovery strategy against COVID–19 pandemic*, "Tourism Recreation Research", vol. 46(2), pp. 188–194, https://doi.org/10.1080/02508281.2020.1805933

Polska turystyka krajowa w obliczu pandemii SARS-CoV–2

Streszczenie: Artykuł skupia się na tematyce rynku turystycznego w obliczu pandemii SARS-CoV–2. Jego celem jest zweryfikowanie, czy polska turystyka w tym okresie faktycznie odnotowała tak znaczący spadek zainteresowania. Badania oparte są na danych wtórnych pochodzących z raportów Głównego Urzędu Statystycznego, dotyczących wykorzystania baz noclegowych w Polsce zarówno przez Polaków, jak i turystów zagranicznych. Usługi noclegowe oferowane przez bazy noclegowe stanowią jedną z podstawowych usług turystycznych wchodzących w skład rynku turystycznego. Zweryfikowanie, jak wyglądała ich sytuacja w okresie wakacyjnym 2020, jest więc jedną z najprostszych możliwości oszacowania wpływu pandemii na rynek turystyki w Polsce. Praca obejmuje również porównanie danych dotyczących polskich turystów w 2020 roku w odniesieniu do lat ubiegłych. Badania przeprowadzono na podstawie analizy wariancji ANOVA.

Słowa kluczowe: turystyka, sektory gospodarcze, COVID–19, SARS-CoV–2, pandemia

JEL: Z31, Z32, Z39, L83

Folia Oeconomica
Acta Universitatis Lodziensis
ISSN 0208-6018 e-ISSN 2353-7663

www.czasopisma.uni.lodz.pl/foe/
2(353) 2021

DOI: https://doi.org/10.18778/0208-6018.353.03

Adam Piotr Idczak
University of Łódź, Faculty of Economics and Sociology, Department of Statistical Methods
Łódź, Poland, adam.idczak@uni.lodz.pl

Sentiment Classification of Bank Clients' Reviews Written in the Polish Language

Abstract: It is estimated that approximately 80% of all data gathered by companies are text documents. This article is devoted to one of the most common problems in text mining, i.e. text classification in sentiment analysis, which focuses on determining the sentiment of a document. A lack of defined structure of the text makes this problem more challenging. This has led to the development of various techniques used in determining the sentiment of a document. In this paper, a comparative analysis of two methods in sentiment classification, a naive Bayes classifier and logistic regression, was conducted. Analysed texts are written in the Polish language and come from banks. The classification was conducted by means of a bag-of-n-grams approach, where a text document is presented as a set of terms and each term consists of n words. The results show that logistic regression performed better.

Keywords: sentiment analysis, opinion mining, text classification, text mining, logistic regression, naive Bayes classifier

JEL: C81, M31

1. Introduction

Approximately 80% of all data gathered by companies has textual form (Sullivan, 2001), such as e-mails, memos, reports, research, reviews, strategy, and marketing plans, etc. All of these textual forms provide a rich and extensive source of valuable (but undiscovered) information. The amount of available data is overwhelming, hence analysing data manually by analysts might be ineffective or even impossible. On the other hand, such a collection of data cannot be processed with typical techniques because of their unstructured form. Fortunately, there are several *text mining* applications available for deriving high-quality information from text documents. This creates an opportunity to take advantage of data to improve decision-making processes in companies.

Text classification in *sentiment analysis* is one of text mining applications which can provide answers to questions such as: "Do clients like my product (or service)?" or "Which aspects of my product (or service) do clients like or not?" It is also helpful in tracking and evaluating customer satisfaction. This type of text analysis focuses on detecting an author's attitude (called *sentiment*) toward entities and their attributes.

In this paper, sentiment classification of bank clients' reviews written in the Polish language is examined in a comparative analysis of two methods. In Section 2, sentiment analysis and document sentiment classification are introduced. The next section presents the idea of a bag-of-n-gram approach, a naive Bayes classifier and logistic regression. Section 4 contains an algorithm for the evaluation of the above-mentioned methods, a data overview, and the results of the comparison conducted. Finally, conclusions are stated at the end.

2. Sentiment analysis

Sentiment analysis (*opinion mining*) focuses on analysing textual data in order to assess an author's attitude toward entities and their attributes. This type of analysis is interdisciplinary in its nature, as it combines research and applications in such fields as: *natural language processing* (NLP), *data mining, web mining*, and *information retrieval*. It is presumed that the terms *sentiment analysis* and *opinion mining* were first introduced in (Dave, Lawrence, Pennock, 2003; Nasukawa, Yi, 2003) respectively, but research regarding *sentiment* and *opinion* emerged a few years earlier (Wiebe, 2000; Das, Chen, 2001; Tong, 2001; Morinaga et al., 2002; Pang, Lee, Vaithyanathan, 2002; Turney, 2002).

It is worth mentioning that there is no clear distinction between *sentiment analysis* and *opinion mining* among researchers and practitioners. In this paper, these two terms will be used interchangeably.

Sentiment analysis can be performed with respect to its granularity level (Liu, 2015):

1) document level – the objective is to classify a whole opinion document into *positive* or *negative sentiment*;
2) sentence level – the main task is to assign sentiment (positive or negative) to each sentence. Sentences without an opinion are considered as *neutral*;
3) aspect level – this type of analysis is focused on finding opinions concerning entities or their aspects and then assigning sentiment to them; for example, opinion *I love this restaurant, but the prices are too high* has overall positive sentiment, but it does not mean that the author of the opinion is positive about all aspects of the restaurant; thus, to obtain such details, one needs to apply aspect level analysis.

2.1. Document sentiment classification

Document sentiment classification is one of the most studied topics in the field of sentiment analysis. Its task is to assess the overall sentiment about an entity based on the opinion document evaluating the entity. In other words, the goal of document sentiment classification is to assign one label (positive, negative or neutral) to a document. Document sentiment classification does not take into account all aspects in the opinion document or seek sentiments regarding them, hence it is considered as document level analysis. There is a great deal of research devoted to sentiment classification studying various types of data and various types of techniques. Turney (2002) used the data from Eopinios.com website that contain reviews sampled from four domains: reviews of cars, banks, movies, and travel destinations. He calculated Semantic Orientation (*SO*) of a term by means of the number of hits returned from the query engine[1] with the reference to words *poor* and *excellent*:

$$SO(term) = \log_2\left(\frac{hits(term\ NEAR\ \text{"excellent"})\,hits(\text{"}poor\text{"})}{hits(term\ NEAR\ \text{"poor"})\,hits(\text{"}excellent\text{"})}\right). \tag{1}$$

The document is labelled as positive if averaged *SO* was positive, otherwise the document was labelled as negative. Pang, Lee, and Vaithyanathan (2002) used film reviews from the Internet Movie Database (IMDb). Their study utilises mostly unigrams and bigrams with term presence as features. Na, Khoo, and Wu (2005) examined unigrams and unigrams with part-of-speech (POS) tags with different weighting schemes (term presence, term frequency, and term frequency inverse document frequency) using on-line product reviews downloaded from the Review Centre (https://www.reviewcentre.com/). Many researchers appreciate messages (*tweets*) from Twitter as a source of data, e.g. Asur and Huberman (2010) classified

1 AltaVista Advanced Search engine.

film reviews (tweets) from Twitter using an n-grams approach in order to improve forecasting box-office revenue of movies. Tweets regarding the Irish Great Election in 2011 were utilised in a uni-gram approach. Hanbury and Nopp (2015) employ sentiment analysis in risk assessment for Eurozone banks. The authors evaluated CEO letters and Outlook sections (usually part of management report) by means of sentiment finance-oriented words. Such a finance-specific list of words comes from Loughran and McDonald's (2011) work. Selected studies with methods and accuracy are given in the Table 1.

Table 1. Selected studies on sentiment classification

No.	Author/Authors	Data set	Method	Accuracy (%)
1	Turney (2002)	Reviews of: – cars – banks – films – tours	Semantic Orientation	 84.0 80.0 65.8 70.5
2	Pang, Lee, and Vaithyanathan (2002)	Film reviews	NB[a]	81.0[g]/77.3[h]
			ME[b]	80.4[g]/77.43[h]
			SVM[c]	82.9[g]/77.13[h]
3	Na, Khoo, and Wu (2005)	On-line product reviews	SVM[c]	75.5[g]
4	Asur and Huberman (2010)	Tweets with film reviews	DynamicLMClassifier	98.0
5	Bermingham and Smeaton (2011)	Tweets regarding the Irish Great Election in 2011.	MNB[d] ADA-MNB[e] SVM[c] ADA-SVM[f]	62.94 65.09 64.82 64.28
6	Hanbury and Nopp (2015)	CEO letters	NB[a]	70.3[i]/75.0[j]
			SVM[c]	70.3[i]/79.2[j]
		Outlook sections of Eurozone banks	NB[a]	56.3[i]/70.4[j]
			SVM[c]	70.3[i]/70.4[j]

[a] Naive Bayes.
[b] Maximum Entropy.
[c] Support Vector Machines.
[d] Multinominal Naive Bayes.
[e] Adaboost M1 Multinominal Naive Bayes.
[f] Adaboost M1 Support Vector Machines.
[g] Unigram (binary).
[h] Bigram (binary).
[i] Lexicon-based approach.
[j] Document frequency and information gain.

Source: own elaboration

3. Classification algorithms

To employ a particular classification algorithm, the opinion documents analysed were expressed in bag-of-n-grams fashion. In this kind of document representation, a document consists of a set of terms (features) where n stands for the number of words in this particular term, e.g. uni-gram, bi-gram, etc. Given this, the documents can be presented as the following document-term matrix (DTM):

$$\boldsymbol{x} = \left[x_{ij}\right], \tag{2}$$

where:
x – is the document-term matrix,
x_{ij} – is the number of times that the j-th term occurred in the i-th document,
$i = 1, \ldots, I$ (I is the total number of documents in a training set),
$j = 1, \ldots, J$ (J is the total number of terms in a training set).

Features from matrix (2) can be transformed in various ways (Pang, Lee, Vaithyanathan, 2003; Na, Khoo, Wu, 2005):

1) term presence (binary):

$$x_{ij}^{*} = \begin{cases} 0, & when\ x_{ij} = 0 \\ 1, & when\ x_{ij} > 0 \end{cases}, \tag{3}$$

2) term frequency (TF):

$$x_{ij}^{*} = x_{ij}, \tag{4}$$

3) term frequency inverse document frequency (TFIDF):

$$x_{ij}^{*} = \begin{cases} 0, when\ x_{ij} = 0 \\ (1 + \log\left(x_{ij}\right))^{*}\log\left(\dfrac{I}{df_{j}}\right), when\ x_{ij} > 0 \end{cases}, \tag{5}$$

where:
I – is the number of all documents,
df_j – is the number of documents where the j-th term occurred.

3.1. Naive Bayes

Bayes' rule (Domański, Pruska, 2000) for document sentiment classifications defines conditional probability that the $\boldsymbol{x}_i$ document belongs to the C_k class:

$$\mathrm{P}\left(C_k|\boldsymbol{x}_i\right)=\frac{p_k f(\boldsymbol{x}_i \mid C_k)}{\sum_{k=1}^{K} p_k f(\boldsymbol{x}_i \mid C_k)}, \tag{6}$$

where:
C_k – is the k-th class, $k = 1, \ldots, K$,
$\boldsymbol{x}_i$ – is the i-th document with J features,
p_k – is the *a priori* probability that the document belongs to the C_k class,
$f(\boldsymbol{x}_i|C_k)$ – is a probability of occurrence of the $\boldsymbol{x}_i$ document, given it belongs to the C_k class.

A naive Bayes (NB) classifier assigns the $\boldsymbol{x}_i$ document to the class C_k if equation (7) is satisfied:

$$\mathrm{P}\left(C_k|\boldsymbol{x}_i\right)=\max_k \mathrm{P}\left(C_k|\boldsymbol{x}_i\right), \tag{7}$$

which is equivalent for:

$$\mathrm{P}\left(C_k|\boldsymbol{x}_i\right)=\max_k\left[p_k f\left(\boldsymbol{x}_i|C_k\right)\right]. \tag{8}$$

The above-mentioned classification rule assumes that terms $\boldsymbol{x}_j$ are independently distributed given the k-th class:

$$f\left(\boldsymbol{x}|C_k\right)=\prod_{j=1}^{J} f\left(x_j|C_k\right). \tag{9}$$

In order to train a naive Bayes classifier, p_k will be calculated using relative-frequency estimation:

$$\hat{p}_k=\frac{n_k}{I}, \tag{10}$$

where n_k is the number of documents given that belong to the k-th class, while $f(\boldsymbol{x}_i|C_k)$ will be calculated using relative-frequency estimation (for term presence or TF):

$$\hat{p}\left(x_j=x_{ij} \mid C_k\right)=\frac{n_{ijk}}{n_{jk}}, \tag{11}$$

or fitting a normal distribution (for TFIDF):

$$\hat{f}\left(x_j|C_k\right)=\left(\widehat{\sigma_{\mathrm{jk}}}\sqrt{2\pi}\right)^{-1} * \exp\left(-\frac{\left(x_j-\widehat{\mu_{\mathrm{jk}}}\right)^2}{2\widehat{\sigma_{\mathrm{jk}}^2}}\right), \tag{12}$$

where:
n_{ijk} – frequency of the i-th value of the j-th term in the k-th class,
n_{jk} – frequency of the j-th term in the k-th class,

$\widehat{\mu_{jk}}$ – mean of TFIDF for the *j*-th term in the *k*-th class,
$\widehat{\sigma_{jk}}$ – standard deviation of TFIDF for the *j*-th term in the *k*-th class.

3.2. Logistic regression

Let us assume that C be the Bernoulli random variable:

$$C \sim Bernoulli(p), \tag{13}$$

that can take one of two values:

$$C = \begin{cases} 0, & when\, the\, sentiment\, of\, a\, document\, is\, negative, \\ 1, & when\, the\, sentiment\, of\, a\, document\, is\, positve, \end{cases} \tag{14}$$

then the logistic regression (Hosmer, Lemeshow, Sturdivant, 2013) can be written as follows:

$$p = p(C = 0|\boldsymbol{x}_i) = \frac{e^{\beta_0 + \beta^t \boldsymbol{x}_i}}{1 + e^{\beta_0 + \beta^t \boldsymbol{x}_i}}, \tag{15}$$

where:
β_0 is an intercept and **β** is a vector of estimated parameters.

It is convenient to apply *logit transformation* on (15) to obtain some desirable properties of a linear model:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta^t \boldsymbol{x}_i, \tag{16}$$

in particular, the above-mentioned equation is linear in its parameters, hence betas have a handy interpretation in terms of *odds ratio* $\left(\frac{e^{\beta_0 + \beta^t \boldsymbol{x}'_i}}{e^{\beta_0 + \beta^t \boldsymbol{x}_i}}\right) e^{\beta_0 + \beta^t \boldsymbol{x}'_i}$2, i. e. if the x_j feature increases by 1 unit (ceteris paribus), the odds ratio will increase by e^{β_j}. This means that the odds that a document has positive sentiment (given the increased x_j) has increased (decreased) by $(e^{\beta_j} - 1) * 100\%$.

Probability $p(C = 0|\boldsymbol{x}_i)$ in (15) is a probability that the document $\boldsymbol{x}_i$ has positive sentiment, thus a probability that the document $\boldsymbol{x}_i$ has negative sentiment is calculated by the following equation:

$$p(C = 1|\boldsymbol{x}_i) = 1 - p(C = 0|\boldsymbol{x}_i). \tag{17}$$

The $\boldsymbol{x}_i$ document is classified as negative if the following equation is satisfied:

$$\mathrm{P}(C = 0|\boldsymbol{x}_i) = \max\left[p(C = 0|\boldsymbol{x}_i), p(C = 1|\boldsymbol{x}_i)\right], \tag{18}$$

otherwise, the document is considered as positive.

2 $e^{\beta_0 + \beta^t \boldsymbol{x}'_i}$ denotes the odds for the x_j feature to be increased by 1 unit.

Parameters from equation (15) can be estimated by means of the *maximum likelihood method* by maximising the following equation:

$$\mathrm{L}(\beta) = \prod_{i=1}^{I} p(C_1|\boldsymbol{x}_i)^{C_i} \left[1 - p(C_1|\boldsymbol{x}_i)\right]^{1-C_i}, \tag{19}$$

with respect to parameters β_0 and $\boldsymbol{\beta}$:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmax}}\, \mathrm{L}(\beta). \tag{20}$$

4. Evaluation

4.1. Experimental set-up

In order to evaluate a naive Bayes classifier and logistic regression in document sentiment classification, experiment is conducted in line with the algorithm presented in Figure 1. All calculations are made in R software. First, the documents analysed are read into the memory, and then they are initially processed, i.e. unwanted numbers, punctuations and words are deleted. Also, *lemmatisation* is a very important part of this step. The process of lemmatisation groups together the inflected forms of the word so that they can be analysed as a single item (*word's lemma*), e.g. *płakać* is lemma for *płakał*, *płakaliśmy*, *płacze*. It is especially important in the case of the Polish language, which is inflected. Lemmatisation is done by means of tm package in R. This step can have a crucial impact on features (and on the number of features) in the document-term matrix. For the purpose of this study, unigrams and bigrams will be considered. The DTM matrix is calculated by the use of hashmap, tm and tex2vec package. After the DTM is created, the three versions of the document-term matrix are calculated (binary, TF and TFIDF) employing RWeka and tm package. Then the matrix is used in 10-fold cross validation, according to Figure 1, where a naive Bayes classifier and logistic regression are learnt on a training sample and classification is evaluated on a validation sample. This part of algorithm is handled by e1071 and gmodels package. The classification is evaluated by means of accuracy:

$$accuracy = \frac{TP + TN}{I}, \tag{21}$$

where:
TP – the number of documents with positive sentiment classified as positive,
TN – the number of documents with negative sentiment classified as negative,
I – the number of all documents.

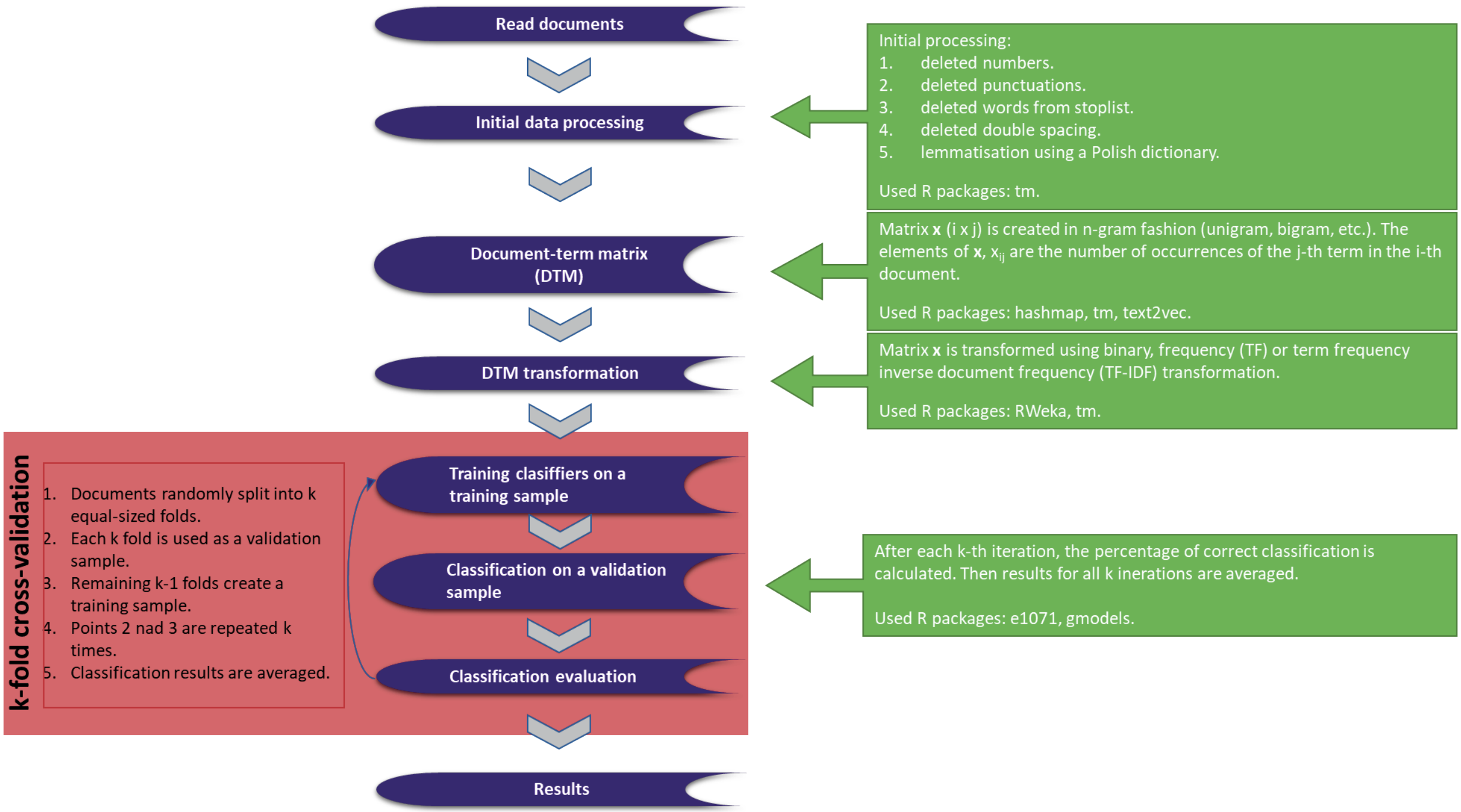


Figure 1. Algorithm

Source: own elaboration

4.2. The data

The data consist of 1,559 documents that are clients' reviews concerning one of Polish banks. Each document is labelled with positive or negative sentiment (positive or negative class). These labels were assigned manually by an opinion holder (by choosing a sad or happy face icon). There were 786 negative and 773 positive documents. Words with the highest frequency in each class (red for negative and green for positive) are shown in Figure 2.

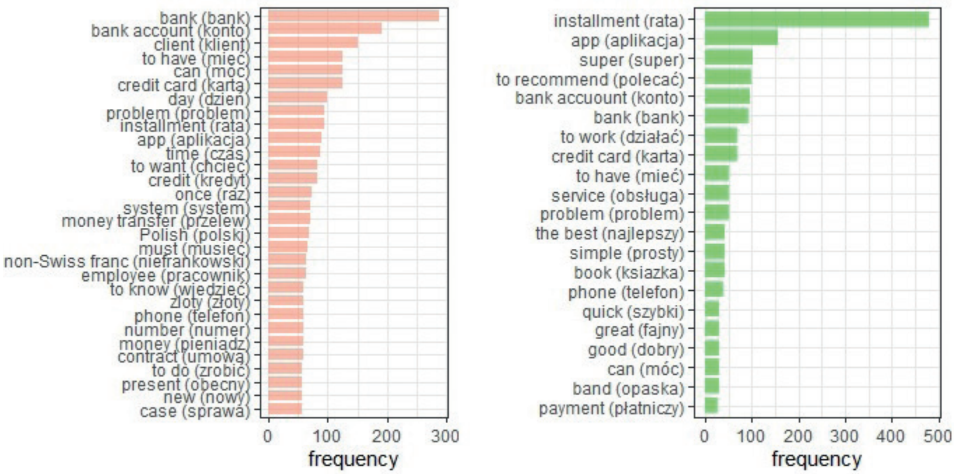


Figure 2. The most frequent words in each class (negative and positive)

Source: own calculations

4.3. Results

Figures 3 and 4 show results of classification of the above-mentioned data set for unigrams and bigrams respectively. Document sentiment classification was conducted by means of naive Bayes classifier (NB) and logistic regression (GLM). It turns out that the considered classification methods outperformed the 50% random-choice baseline and the results ranged from 51.06% to 82.81%. The highest accuracy was observed for logistic regression (unigram DTM with TFIDF) and the lowest was achieved for a naive Bayes classifier (bigram DTM with TFIDF).

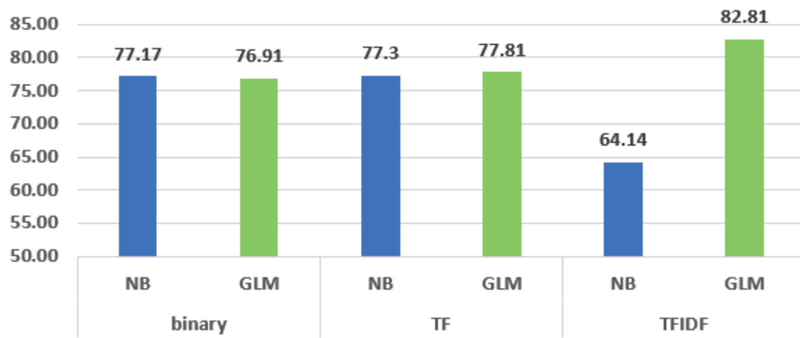


Figure 3. Accuracy (%) of unigrams

Source: own calculations

Results for unigrams are quite similar for binary and TF transformation and range from 76.91% to 77.81% but for TFIDF differences are greater, i.e. a Naive Bayes classifier with TFIDF (64.14%) performs worse than NB and GLM with binary or TF. Also, in terms of accuracy, NB is worse than any logistic regression. In fact, GLM with TFIDF has the highest percentage of correctly classified documents (82.81%).

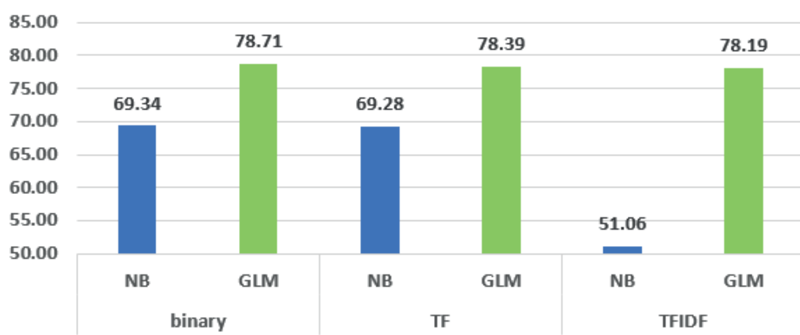


Figure 4. Accuracy (%) of bigrams

Source: own calculations

As for bigrams, logistic regression performed better than a naive Bayes classifier, yielding roughly 78% of correctly classified documents. Accuracy of NB was about –9 p.p. worse than GLM for binary and TF. NB with TFIDF has the lowest accuracy (only 51.06%), yielding performance only about 1 p.p. above the random-choice baseline.

5. Conclusions

In this paper, a naive Bayes classifier and logistic regression were examined in document sentiment classification performed for the Polish language. This problem was found by researchers (Pang, Lee, Vaithyanathan, 2003) to be more challenging than traditional topic-based classification, which concerns keywords that help identify topics. Document sentiment classification is more complex because sentiment (rather than topics) can be expressed in a more subtle manner.

The results produced in section 4.3 indicate that the performance of naive Bayes classifier and logistic regression applied to the customer reviews written in Polish is high. In all cases, the accuracy is higher than the random-choice baseline, and it also fits in the accuracy that the researchers obtained in their studies (see Table 1). Logistic regression with TF-IDF yielded the highest accuracy, i.e. 82.81%.

When it comes to TFIDF transformation, the accuracy for a naive Bayes classifier was undoubtedly poorer than in the case of the other approaches. The reason for such drop in performance is that the distribution of TFIDF features does not necessarily follow the density function when $f(\boldsymbol{x}_i|C_k)$ is a normal distribution.

It is worth mentioning that estimates of parameters of the above-mentioned methods are highly influenced by sparsity of the DTM matrix. Thus, performance of considered classifiers is driven by non-occurrence rather than occurrence of features obtained from the training set. Saif, He and Alani (2012) proposed two effective approaches to deal with sparsity of the DTM matrix.

The results (considered as high in terms of accuracy) presented in this article cannot be generalised to all types of documents written in the Polish language due to the fact that: (1) each type of data has its own specific way of expressing sentiment, (2) most of document sentiment classification research is conducted on documents written in the English language, whereas the Polish language is inflected, which affects the DTM matrix and can possibly add some complexity to expressing the sentiment. All in all, it seems that more studies on documents in the Polish language are needed.

References

Asur S., Huberman B.A. (2010), *Prediction the Future with Social Media*, https://www.researchgate.net/publication/45909086_Predicting_the_Future_with_Social_Media [accessed: 10.02.2021].

Bermingham A., Smeaton A.F. (2011), *On Using Twitter to Monitor Political Sentiment and Predict Election Results*, "Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP)", pp. 2–10, https://www.aclweb.org/anthology/W11-3702.pdf [accessed: 10.02.2021].

Das S., Chen M. (2001), *Yahoo! For Amazon: Extracting Market Sentiment from Stock Message Boards*, "Proceedings of APFA–2001".

Dave K., Lawrence S., Pennock D.M. (2003), *Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews*, "Proceedings of International Conference on World Wide Web (WWW–2003)", https://www.researchgate.net/publication/2904559_Mining_the_Peanut_Gallery_Opinion_Extraction_and_Semantic_Classification_of_Product_Reviews [accessed: 10.02.2021].

Domański Cz., Pruska K. (2000), *Nieklasyczne metody statystyczne*, PWE, Warszawa.

Hanbury A., Nopp C. (2015), *Detecting Risks in the Banking System by Sentiment Analysis*, "Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing", pp. 591–600, https://www.aclweb.org/anthology/D15-1071.pdf [accessed: 15.02.2021].

Hosmer D.W., Lemeshow S., Sturdivant R.X. (2013), *Applied Logistic Regression*, 3rd ed., John Wiley & Sons, New Jersey.

Liu B. (2015), *Sentiment Analysis. Mining Opinions, Sentiments, and Emotions*, Cambridge University Press, New York.

Loughran T., McDonald B. (2011), *When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks*, "Journal of Finance", vol. 66, no. 1, pp. 35–65, https://www.uts.edu.au/sites/default/files/ADG_Cons2015_Loughran%20McDonald%20JE%202011.pdf [accessed: 19.02.2021].

Morinaga S., Yamanishi K., Tateishi K., Fukushima T. (2002), *Mining Product Reputations on the Web*, "Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD–2002)", https://www.researchgate.net/publication/200044311_Mining_product_reputations_on_the_Web [accessed: 10.02.2021].

Na J.Ch., Khoo C., Wu P.H.J. (2005), *Use of negation phrases in automatic sentiment classification of product reviews*, "Library Collections, Acquisitions & Technical Services", no. 29, pp. 180–191, https://ccc.inaoep.mx/~villasen/bib/Use%20of%20negation%20phrases%20in%20automatic%20sentiment%20classification.pdf [accessed: 11.02.2021].

Nasukawa T., Yi J. (2003), *Sentiment Analysis: Capturing Favorability Using Natural Language Processing*, "Proceedings of the K-CAP–03, 2nd International Conference on Knowledge Capture", pp. 70–77, https://www.researchgate.net/publication/220916772_Sentiment_analysis_Capturing_favorability_using_natural_language_processing [accessed: 15.02.2021].

Pang B., Lee L., Vaithyanathan S. (2002), *Thumbs up? Sentiment Classification using Machine Learning Techniques*, "Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)", pp. 79–86, https://www.cs.cornell.edu/home/llee/papers/sentiment.pdf [accessed: 8.02.2021].

Review Centre, https://www.reviewcentre.com/ [accessed: 25.02.2021].

Saif H., He Y., Alani H. (2012), *Alleviating data sparsity for Twitter sentiment analysis*, [in:] *2nd Workshop on Making Sense of Microposts (#MSM2012): Big things come in small packages at the 21st International Conference on the World Wide Web (WWW'12), 16 Apr 2012, Lyon, France*, CEUR Workshop Proceedings (CEUR-WS.org), pp. 2–9, https://www.researchgate.net/publication/228450062_Alleviating_Data_Sparsity_for_Twitter_Sentiment_Analysis [accessed: 25.02.2021].

Sullivan D. (2001), *Integrating Data and Document Warehouses*, "DM Review Magazine", http://www.dmreview.com/article_sub_articleId_3697.html [accessed: 18.02.2021].

Tong R.M. (2001), *An Operational System for Detecting and Tracking Opinions in on-Line Discussion*, "Proceedings of SIGIR Workshop on Operational Text Classification".

Turney P. D. (2002), *Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews*, "Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)", pp. 417–424, https://www.researchgate.net/publication/248832100_Thumbs_Up_or_Thumbs_Down_Semantic_Orientation_Applied_to_Unsupervised_Classification_of_Reviews [accessed: 22.02.2021].

Wiebe J. (2000), *Learning Subjective Adjectives from Corpora*, "Proceedings of National Conference on Artificial Intelligence (AAAI–2000)", pp. 735–740, https://www.aaai.org/Papers/AAAI/2000/AAAI00-113.pdf [accessed: 13.02.2021].

Analiza sentymentu na podstawie polskojęzycznych recenzji klientów banku

Streszczenie: Szacuje się, że około 80% wszystkich danych gromadzonych i przechowywanych w systemach informacyjnych przedsiębiorstw ma postać dokumentów tekstowych. Artykuł jest poświęcony jednemu z podstawowych problemów textminingu, tj. klasyfikacji tekstów w analizie sentymentu, która rozumiana jest jako badanie wydźwięku tekstu. Brak określonej struktury dokumentów tekstowych jest przeszkodą w realizacji tego zadania. Taki stan rzeczy wymusił rozwój wielu różnorodnych technik ustalania sentymentu dokumentów. W artykule przeprowadzono analizę porównawczą dwóch metod badania sentymentu: naiwnego klasyfikatora Bayesa oraz regresji logistycznej. Badane teksty są napisane w języku polskim, pochodzą z banków i mają charakter marketingowy. Klasyfikację przeprowadzono, stosując podejście *bag-of-n-grams*. W ramach tego podejścia dokument tekstowy wyrażony jest za pomocą podciągów składających się z określonej liczby *n* wyrazów. Uzyskane wyniki pokazały, że lepiej spisała się regresja logistyczna.

Słowa kluczowe: analiza sentymentu, klasyfikacja dokumentów, textmining, regresja logistyczna, naiwny klasyfikator Bayesa

JEL: C81, M31

Received: 2021-01-17; verified: 2021-03-30. Accepted: 2021-06-30

Folia Oeconomica
Acta Universitatis Lodziensis
ISSN 0208-6018 e-ISSN 2353-7663

www.czasopisma.uni.lodz.pl/foe/
2(353) 2021

DOI: https://doi.org/10.18778/0208-6018.353.04

Małgorzata Graczyk
Poznań University of Life Sciences, Faculty of Agronomy and Bioengineering
Department of Mathematical and Statistical Methods, malgorzata.graczyk@up.poznan.pl

Bronisław Ceranka
Poznań University of Life Sciences, Faculty of Agronomy and Bioengineering
Department of Mathematical and Statistical Methods, bronislaw.ceranka@up.poznan.pl

Selected Remarks on Highly D-efficient Spring Balance Weighing Designs

Abstract: Here, we consider a new construction method of determining highly D-efficient spring balance weighing designs in classes in which a D-optimal design does not exist. We give some conditions determining the relations between the parameters of such designs and construction examples.

Keywords: D-efficient design, spring balance weighing design

JEL: C02, C18, C90

1. Introduction

Let us consider $\boldsymbol{\Psi}_{n\times p}(0,1)$, a class of $n\times p$ matrices $\mathbf{X}=(x_{ij})$, $i=1,2,\dots,n$, $j=1,2,..,p$. Any matrix $\mathbf{X}\in\boldsymbol{\Psi}_{n\times p}(0,1)$ is called a design matrix of the spring balance weighing design. Originally, the name spring balance weighing design pertained to experiments connected with determining unknown weights of objects by the use of balance with one pan which is called a spring balance. Nowadays, such designs are applied in many branches of knowledge including economic survey, see Banerjee (1975), Ceranka and Graczyk (2014). Some aspects of applications of spring balance weighing designs in agriculture are given by Ceranka and Katulska (1987a; 1987b; 1989), and Graczyk (2013). The example of application of such designs in bioengineering is presented in Gawande and Patkar (1999). Various problems related to spring balance weighing designs are presented in the literature. They are focused on the optimality criteria of such designs. The classic works here are Jacroux and Notz (1983), Koukouvinos (1996). Another group of issues is concerned with determining new methods of construction of the design matrices satisfying optimality conditions. The best general references here are Gail and Kiefer (1982), Ceranka and Graczyk (2010; 2012), Katulska and Smaga (2010).

For any matrix $\mathbf{X}\in\boldsymbol{\Psi}_{n\times p}(0,1)$, we consider a linear model:

$$\mathbf{y}=\mathbf{Xw}+\mathbf{e}, \tag{1}$$

where y is an $n\times 1$ random vector of observed measurements. Moreover, w is a $p\times 1$ vector representing unknown measurements of objects and w is an $n\times 1$ vector of random errors. We make two standing assumptions: it is required that there are no systematic errors, i.e. $E(\mathbf{e})$ = 0_n, and that the errors are uncorrelated and have different variances, i.e. Var($\mathbf{e}$) = $\sigma^2\mathbf{G}$, where $\sigma>0$ is a known parameter, G is the $n\times n$ diagonal positive definite matrix of known elements.

For the estimation of the vector of unknown measurements of objects **w**, we use the normal equation $\mathbf{X'G^{-1}X\hat{w}}=\mathbf{X'G^{-1}y}$. Under the assumption that **G** is a known positive definite matrix, $\mathbf{X'G^{-1}X}$ is nonsingular if and only if **X** is of full column rank. In the case when $\mathbf{X'G^{-1}X}$ is nonsingular, the generalised least squares estimator of **w** is given by $\mathbf{\hat{w}}=(\mathbf{X'G^{-1}X})^{-1}\mathbf{X'G^{-1}y}$ and Var($\mathbf{\hat{w}}$) = $\sigma^2(\mathbf{X'G^{-1}X})^{-1}$.

The statistical problem considered here is how to determine the estimator of the vector of unknown measurements of objects **w** when the observations follow the model (1). Among several questions taken under consideration, the properties of this estimator are under considerations. The characteristic features are determined by the properties of the design. Especially here, it is expected that the product of the variance of the estimators has attained the lowest bound. Hence, the criterion of the D-optimality is considered. The design $\mathbf{X}_\mathrm{D}$ is D-optimal in the class of the designs $\boldsymbol{\Psi}_{n\times p}(0,1)$ if $\det\left(\mathbf{X'G}^{-1}\mathbf{X}\right)^{-1}=\min\left\{\det\left(\mathbf{X'G}^{-1}\mathbf{X}\right)^{-1}:\mathbf{X}\in\boldsymbol{\Psi}_{n\times p}(0,1)\right\}$.

It is known that $\det\left(\mathbf{X'G}^{-1}\mathbf{X}\right)^{-1}$ is minimal if and only if $\det\left(\mathbf{X'G}^{-1}\mathbf{X}\right)$ is maximal. The concept of D-optimality was considered in the books of Raghavarao (1971), Banerjee (1975), as well as Shah and Sinh (1989). Although theoretical studies on providing knowledge to guide the selection of optimal designs are not scarce, we are still unable to determine a regular D-optimal design for any combination of the number of objects and the number of measurements.

In such a case, a highly D-efficient design is considered. For details, we refer the reader to Bulutoglu and Ryan (2009). In Ceranka and Graczyk's (2018) paper, the definition of D-efficiency is given. We indicate a highly D-efficient design when $\mathrm{D}_{\mathrm{eff}}=\left[\frac{\det\left(\mathbf{X'X}\right)}{\det\left(\mathbf{Y'Y}\right)}\right]^{1/p}\geq 0.95,$ where $\mathbf{Y}$ is the matrix of D-optimal spring balance weighing design.

The aim of this paper is to develop new construction methods related to D-optimal and highly D-efficient spring balance weighing designs for which random errors are uncorrelated and have different variances. An attempt has been made here to expand the theory of optimal designs. The aim of this research is to develop the results concerning new methods of determining optimal designs in classes in which they have not been determined in the literature so far.

2. The main result

We present the theorem determining the parameters of the highly D-efficient design given in Ceranka and Graczyk (2018; 2019).

Theorem 2.1. Let p be even. In any non-singular spring balance weighing design $\mathbf{X}\in\boldsymbol{\Psi}_{n\times p}(0,1)$ having $0.5p$ ones in each row and with the variance matrix of errors $\sigma^2\mathbf{I}_n$, $\det\left(\mathbf{X'X}\right)\leq(p-1)\left(\frac{np}{4(p-1)}\right)^p$, an upper bound is attained if and only if $\mathbf{X'X}=\frac{n}{4(p-1)}\left(p\mathbf{I}_p+(p-2)\mathbf{1}_p\mathbf{1}_p'\right)$, where $0.25n(p-2)(p-1)^{-1}$ and $0.25np(p-1)^{-1}$ are integers.

The design $\mathbf{X}\in\boldsymbol{\Psi}_{n\times p}(0,1)$ having the form given in Theorem 2.1 is considered as highly D-efficient.

2.1. Addition of one measurement

Let $\mathbf{X}_1 \in \mathbf{\Psi}_{(n-1)\times p}(0,1)$ be the design of the highly D-efficient spring balance weighing design. Now, let us consider the design $\mathbf{X} \in \mathbf{\Psi}_{n\times p}(0,1)$ in the form:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{x}_1' \end{bmatrix}, \tag{2.1}$$

where $\mathbf{x}_1$ is any $p\times 1$ vector of elements 1 or 0, $\mathbf{x}_1'\mathbf{1}_p = t_1$, $1 \le t_1 \le p$. So, the variance matrix of errors is given as:

$$\mathbf{G} = \begin{bmatrix} \mathbf{I}_{n-1} & \mathbf{0}_{n-1} \\ \mathbf{0}_{n-1}' & g_1^{-1} \end{bmatrix}. \tag{2.2}$$

Furthermore, we study the function $\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)$.

Because $\mathbf{X}_1'\mathbf{X}_1 = \frac{n-1}{4(p-1)}\left(p\mathbf{I}_p + (p-2)\mathbf{1}_p\mathbf{1}_p'\right)$, then $\det\left(\mathbf{X}_1'\mathbf{X}_1\right) \le (p-1)\left(\frac{(n-1)p}{4(p-1)}\right)^p$

and the determinant is maximal if and only if $\frac{(n-1)p}{4(p-1)}$ and $\frac{(n-1)(p-2)}{4(p-1)}$ are integers.

Thus for the design matrix $\mathbf{X} \in \mathbf{\Psi}_{n\times p}(0,1)$, we have

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) \le (p-1)\left(\frac{(n-1)p}{4(p-1)}\right)^p \left(1 + g_1\mathbf{x}_1'\left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\mathbf{x}_1\right).$$

Owing to the fact that $\left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1} = \frac{4(p-1)}{(n-1)p}\left(\mathbf{I}_p - \frac{p-2}{p(p-1)}\mathbf{1}_p\mathbf{1}_p'\right)$, we obtain:

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) \le (p-1)\left(\frac{(n-1)p}{4(p-1)}\right)^p \left(1 + g_1\mathbf{x}_1'\frac{4(p-1)}{(n-1)p}\left(\mathbf{I}_p - \frac{p-2}{p(p-1)}\mathbf{1}_p\mathbf{1}_p'\right)\mathbf{x}_1\right) =$$
$$(p-1)\left(\frac{(n-1)p}{4(p-1)}\right)^p \left(1 + \frac{4(p-1)g_1}{(n-1)p}\left(\mathbf{x}_1'\mathbf{x}_1 - \frac{p-2}{p(p-1)}\mathbf{x}_1'\mathbf{1}_p\mathbf{1}_p'\mathbf{x}_1\right)\right).$$

In order to maximise the expression to the right of the inequality sign, we observe that $t_1 - \frac{p-2}{p(p-1)}t_1^2 \le \frac{p^3+8}{4p(p-1)}$ and the equality in the above-presented inequality holds if and only if $t_1 = \frac{p+2}{2}$. Hence,

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)\leq(p-1)\left(\frac{(n-1)p}{4(p-1)}\right)^{p}\left(1+\frac{4(p-1)g_1}{(n-1)p}\cdot\frac{p^3+8}{4p(p-1)}\right)=$$
$$(p-1)\left(\frac{(n-1)p}{4(p-1)}\right)^{p}\left(1+\frac{g_1\left(p^3+8\right)}{(n-1)p^2}\right).$$

The above-presented equality is fulfilled if and only if $t_1=\frac{p+2}{2}$ and $\frac{(n-1)p}{4(p-1)}$ and $\frac{(n-1)(p-2)}{4(p-1)}$ are integer numbers.

Theorem 2.2. Let p be even. Any spring balance weighing design $\mathbf{X}\in\mathbf{\Psi}_{n\times p}(0,1)$ in the form (2.1) with the variance matrix of errors in the form (2.2) is highly D-efficient in the class $\mathbf{\Psi}_{n\times p}(0,1)$ if and only if $t_1=\frac{p+2}{2}$ and $\frac{(n-1)p}{4(p-1)}$ and $\frac{(n-1)(p-2)}{4(p-1)}$ are integer numbers.

Definition 2.1. Any spring balance weighing design $\mathbf{X}\in\mathbf{\Psi}_{n\times p}(0,1)$ in the form (2.1) with the variance matrix of errors in the form (2.2) is highly D-efficient if:

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)\leq(p-1)\left(\frac{(n-1)p}{4(p-1)}\right)^{p}\left(1+\frac{g_1\left(p^3+8\right)}{(n-1)p^2}\right).$$

Example 2.1. Let us consider the variance matrix of errors $\sigma^2\mathbf{G}$, where $\mathbf{G}^{-1}=\begin{bmatrix}\mathbf{I}_6 & \mathbf{0}_6\\ \mathbf{0}_6' & 5\end{bmatrix}$. We determine a highly D-efficient design in the class $\mathbf{X}\in\mathbf{\Psi}_{7\times 4}(0,1)$. So, take the highly D-efficient spring balance weighing design $\mathbf{X}_1\in\mathbf{\Psi}_{6\times 4}(0,1)$ in the form $\mathbf{X}_1=\begin{bmatrix}1&1&0&0\\1&0&1&0\\1&0&0&1\\0&1&1&0\\0&1&0&1\\0&0&1&1\end{bmatrix}$. If $\mathbf{x}_1'=[1\ 1\ 1\ 0]$. So, $\mathbf{X}=\begin{bmatrix}\mathbf{X}_1\\ 1\ 1\ 1\ 0\end{bmatrix}\in\mathbf{\Psi}_{7\times 4}(0,1)$ is a highly D-efficient spring balance weighing design.

Example 2.2. Let us consider the variance matrix of errors $\sigma^2\mathbf{G}$, where $\mathbf{G}^{-1}=\begin{bmatrix}\mathbf{I}_{10} & \mathbf{0}_{10}\\ \mathbf{0}_{10}' & 5\end{bmatrix}$. We determine a highly D-efficient design in the class $\mathbf{X}\in\mathbf{\Psi}_{11\times 6}(0,1)$. So, take the highly D-efficient spring balance weighing design

$\mathbf{X}_1 \in \mathbf{\Psi}_{10\times 6}(0,1)$ in the form $\mathbf{X}_1 = \begin{bmatrix} 1&1&0&0&1&0\\ 1&1&0&0&0&1\\ 1&0&1&1&0&0\\ 1&0&1&0&0&1\\ 1&0&0&1&1&0\\ 0&1&1&1&0&0\\ 0&1&1&0&1&0\\ 0&1&0&1&0&1\\ 0&0&1&0&1&1\\ 0&0&0&1&1&1 \end{bmatrix}$. If $\mathbf{x}_1^{'} = [1\ 1\ 1\ 1\ 0\ 0]$, then

$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ 1\ 1\ 1\ 1\ 0\ 0 \end{bmatrix} \in \mathbf{\Psi}_{11\times 6}(0,1)$ is a hhly D-efficient spring balance weighing design.

2.2. Addition of two measurements

Let $\mathbf{X}_1 \in \mathbf{\Psi}_{(n-2)\times p}(0,1)$ be the design of the highly D-efficient spring balance weighing design. Now, let us consider the design $\mathbf{X} \in \mathbf{\Psi}_{n\times p}(0,1)$ in the form:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{x}_1^{'} \\ \mathbf{x}_2^{'} \end{bmatrix}, \tag{2.3}$$

where $\mathbf{x}_h$ is a $p\times 1$ vector of elements 1 or 0, $\mathbf{x}_h^{'}\mathbf{1}_p = t_h, h = 1,2$, $\mathbf{x}_1^{'}\mathbf{x}_2 = u_{12}$, $1 \le t_h \le p$. So, the variance matrix of errors is given as:

$$\mathbf{G} = \begin{bmatrix} \mathbf{I}_{n-2} & \mathbf{0}_{n-2} & \mathbf{0}_{n-2} \\ \mathbf{0}_{n-2}^{'} & g_1^{-1} & 0 \\ \mathbf{0}_{n-2}^{'} & 0 & g_2^{-1} \end{bmatrix}. \tag{2.4}$$

Furthermore, we study the function $\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)$.

Because $\mathbf{X}_1^{'}\mathbf{X}_1 = \frac{n-2}{4(p-1)}\left(p\mathbf{I}_p + (p-2)\mathbf{1}_p\mathbf{1}_p^{'}\right)$, then $\det\left(\mathbf{X}_1^{'}\mathbf{X}_1\right) \le (p-1)\left(\frac{(n-2)p}{4(p-1)}\right)^p$ and the maximum value is attained if and only if $\frac{(n-2)p}{4(p-1)}$ and $\frac{(n-2)(p-2)}{4(p-1)}$ are integers. Thus, $\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) \le (p-1)\left(\frac{(n-2)p}{4(p-1)}\right)^p \det\left(\mathbf{I}_2 + \begin{bmatrix} g_1\mathbf{x}_1^{'} \\ g_2\mathbf{x}_2^{'} \end{bmatrix}\left(\mathbf{X}_1^{'}\mathbf{X}_1\right)^{-1}\mathbf{x}_1\mathbf{x}_2\right)$.

Since $\left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1} = \frac{4(p-1)}{(n-2)p}\left(\mathbf{I}_p - \frac{p-2}{p(p-1)}\mathbf{1}_p\mathbf{1}_p'\right)$, then we have

$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) \le (p-1)\left(\frac{(n-2)p}{4(p-1)}\right)^p \det(\mathbf{T}_2)$, here:

$$\mathbf{T}_2 = \mathbf{I}_2 + \begin{bmatrix} g_1\mathbf{x}_1' \\ g_2\mathbf{x}_2' \end{bmatrix}\frac{4(p-1)}{(n-2)p}\left(\mathbf{I}_p - \frac{p-2}{p(p-1)}1_p1_p'\right)[\mathbf{x}_1\mathbf{x}_2]$$

$$\det(\mathbf{T}_2) = \left[1 + \frac{4(p-1)g_1}{(n-2)p}\left(t_1 - \frac{p-2}{p(p-1)}t_1^2\right)\right]\cdot\left[1 + \frac{4(p-1)g_2}{(n-2)p}\left(t_2 - \frac{p-2}{p(p-1)}t_2^2\right)\right] - \left(\frac{4(p-1)}{(n-2)p}\right)^2 g_1g_2\left(u_{12} - \frac{p-2}{p(p-1)}t_1t_2\right)^2.$$

$\det(\mathbf{T}_2)$ takes the maximum value if $t_h - \frac{p-2}{p(p-1)}t_h^2 \le \frac{p^3+8}{4p(p-1)}$ and the equality is fulfilled if and only if $t_h = \frac{p+2}{2}, h = 1,2$.

Hence,

$$\det(\mathbf{T}_2) \le \left[1 + \frac{4(p-1)g_1}{(n-2)p}\frac{p^3+8}{4p(p-1)}\right]\cdot\left[1 + \frac{4(p-1)g_2}{(n-2)p}\frac{p^3+8}{4p(p-1)}\right] - \left(\frac{4(p-1)}{(n-2)p}\right)^2 g_1g_2\left(u_{12} - \frac{(p-2)(p+2)^2}{4p(p-1)}\right)^2 =$$

$$\left(1 + \frac{g_1(p^3+8)}{(n-2)p^2}\right)\left(1 + \frac{g_2(p^3+8)}{(n-2)p^2}\right) - \left(\frac{4(p-1)}{(n-2)p}\right)^2 g_1g_2\left(u_{12} - \frac{(p-2)(p+2)^2}{4p(p-1)}\right)^2.$$

u_{12} is an integer number. Now, we consider two cases: $p \equiv 0 \mod 4$ and $p+2 \equiv 0 \mod 4$.

If $p \equiv 0 \mod 4$, then $u_{12} = 0.25(p+4)$ and

$$\left(u_{12} - \frac{(p-2)(p+2)^2}{4p(p-1)}\right)^2 = \left(\frac{p+4}{4} - \frac{(p-2)(p+2)^2}{4p(p-1)}\right)^2 = \left(\frac{p^2+8}{4p(p-1)}\right)^2.$$

In this situation,

$\det(\mathbf{T}_2) \le \left(1 + \frac{g_1(p^3+8)}{(n-2)p^2}\right)\cdot\left(1 + \frac{g_2(p^3+8)}{(n-2)p^2}\right) - \left(\frac{4(p-1)}{(n-2)p}\right)^2 g_1g_2\left(\frac{p^2+8}{4p(p-1)}\right)^2$ and

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) \le (p-1)\left(\frac{(n-2)p}{4(p-1)}\right)^{p}$$
$$\left\{\left(1+\frac{g_1\left(p^3+8\right)}{(n-2)p^2}\right)\cdot\left(1+\frac{g_2\left(p^3+8\right)}{(n-2)p^2}\right)-g_1g_2\left(\frac{p^2+8}{p^2(n-2)}\right)^2\right\}.$$

Theorem 2.3. Let $p \equiv 0 \mod 4$. Any spring balance weighing design $\mathbf{X} \in \mathbf{\Psi}_{n\times p}(0,1)$ in the form (2.3) with the variance matrix of errors in the form (2.4) is highly D-efficient in the class $\mathbf{\Psi}_{n\times p}(0,1)$ if and only if $t_h = \frac{p+2}{2}$ and $\frac{(n-1)p}{4(p-1)}$ and $\frac{(n-2)(p-2)}{4(p-1)}$ are integer numbers, $u_{12} = 0.25(p+4), h = 1,2$.

Definition 2.2. Let $p \equiv 0 \mod 4$. Any spring balance weighing design $\mathbf{X}$ in the form (2.3) with the variance matrix of errors in the form (2.4) is highly D-efficient if

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) = (p-1)\left(\frac{(n-2)p}{4(p-1)}\right)^{p}$$
$$\left\{\left(1+\frac{g_1\left(p^3+8\right)}{(n-2)p^2}\right)\cdot\left(1+\frac{g_2\left(p^3+8\right)}{(n-2)p^2}\right)-g_1g_2\left(\frac{p^2+8}{p^2(n-2)}\right)^2\right\}.$$

Now, we turn to the case $p+2 \equiv 0 \mod 4$.

If $p+2 \equiv 0 \mod 4$, then $u_{12} = 0.25(p+2)$ and

$$\left(u_{12}-\frac{(p-2)(p+2)^2}{4p(p-1)}\right)^2 = \left(\frac{p+2}{4}-\frac{(p-2)(p+2)^2}{4p(p-1)}\right)^2 = \left(\frac{(p+2)(p-4)}{4p(p-1)}\right)^2.$$

In this case,

$$\det\left(\mathbf{T}_2\right) \le \left(1+\frac{g_1\left(p^3+8\right)}{(n-2)p^2}\right)\cdot\left(1+\frac{g_2\left(p^3+8\right)}{(n-2)p^2}\right)-\left(\frac{4(p-1)}{(n-2)p}\right)^2 g_1g_2\left(\frac{(p+2)(p-4)}{4p(p-1)}\right)^2$$

and

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) \le (p-1)\left(\frac{(n-2)p}{4(p-1)}\right)^{p}$$
$$\left\{\left(1+\frac{g_1\left(p^3+8\right)}{(n-2)p^2}\right)\cdot\left(1+\frac{g_2\left(p^3+8\right)}{(n-2)p^2}\right)-g_1g_2\left(\frac{(p+2)(p-4)}{p^2(n-2)}\right)^2\right\}.$$

Theorem 2.4. Let $p+2 \equiv 0 \mod 4$. Any spring balance weighing design $\mathbf{X} \in \mathbf{\Psi}_{n\times p}(0,1)$ in the form (2.3) with the variance matrix of errors in the form (2.4) is highly D-efficient in the class $\boldsymbol{\Psi}_{n\times p}(0,1)$ if and only if $t_h = \frac{p+2}{2}$ and $\frac{(n-2)p}{4(p-1)}$ and $\frac{(n-2)(p-2)}{4(p-1)}$ are integer numbers, $u_{12} = 0.25(p+2), h = 1,2$.

Definition 2.3. Let $p+2 \equiv 0 \, mod \, 4$. Any spring balance weighing design **X** in the form (2.3) with the variance matrix of errors in the form (2.4) is highly D-efficient if

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) = (p-1)\left(\frac{(n-2)p}{4(p-1)}\right)^p$$

$$\left\{\left(1+\frac{g_1\left(p^3+8\right)}{(n-2)p^2}\right)\cdot\left(1+\frac{g_2\left(p^3+8\right)}{(n-2)p^2}\right)-g_1g_2\left(\frac{(p+2)(p-4)}{p^2(n-2)}\right)^2\right\}.$$

Example 2.3. Let us consider the variance matrix of errors

$$\sigma^2\mathbf{G}, \text{ where } \mathbf{G}^{-1} = \begin{bmatrix} \mathbf{I}_6 & \mathbf{0}_6 & \mathbf{0}_6 \\ \mathbf{0}_6' & 5 & 0 \\ \mathbf{0}_6' & 0 & 3 \end{bmatrix}.$$

We determine a highly D-efficient design in the class $\mathbf{X} \in \boldsymbol{\Psi}_{8\times 4}(0,1)$. So, take the highly D-efficient spring balance weighing design $\mathbf{X}_1 \in \boldsymbol{\Psi}_{6\times 4}(0,1)$ in the form

$$\mathbf{X}_1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}. \text{ If } \mathbf{x}_1' = [1\ 1\ 1\ 0],\ \mathbf{x}_2' = [1\ 1\ 0\ 1], \text{ then } \mathbf{X} = \begin{bmatrix} & \mathbf{X}_1 & & \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix} \in \boldsymbol{\Psi}_{8\times 4}(0,1)$$

is a highly D-efficient spring balance weighing design.

Example 2.4. Let us consider the variance matrix of errors $\sigma^2\mathbf{G}$, where $\mathbf{G}^{-1} = \begin{bmatrix} \mathbf{I}_{10} & \mathbf{0}_{10} & \mathbf{0}_{10} \\ \mathbf{0}_{10}' & 5 & 0 \\ \mathbf{0}_{10}' & 0 & 3 \end{bmatrix}$. We determine a highly D-efficient design in the class $\mathbf{X} \in \boldsymbol{\Psi}_{12\times 6}(0,1)$. So, take the highly D-efficient spring balance weighing design

$\mathbf{X}_1 \in \mathbf{\Psi}_{10\times 6}(0,1)$ in the form $\mathbf{X}_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$. If $\mathbf{x}_1' = [1\ 1\ 1\ 1\ 0\ 0]$, $\mathbf{x}_2' = [1\ 1\ 0\ 0\ 1\ 1]$, then $\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ 1\ 1\ 1\ 1\ 0\ 0 \\ 1\ 1\ 0\ 0\ 1\ 1 \end{bmatrix} \in \mathbf{\Psi}_{12\times 6}(0,1)$ is a highly D-efficient spring balance weighing design.

2.3. Addition of three measurements

Let $\mathbf{X}_1 \in \mathbf{\Psi}_{(n-3)\times p}(0,1)$ be the design of the highly D-efficient spring balance weighing design. Now, let us consider the design $\mathbf{X} \in \mathbf{\Psi}_{n\times p}(0,1)$ in the form:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{x}_1' \\ \mathbf{x}_2' \\ \mathbf{x}_3' \end{bmatrix}, \tag{2.5}$$

where $\mathbf{x}_h'\mathbf{1}_p = t_h$, $\mathbf{x}_h'\mathbf{x}_s = u_{hs}$, $h,s = 1,2,3, h \neq s$. So, the variance matrix of errors is given as:

$$\mathbf{G} = \begin{bmatrix} \mathbf{I}_{n-3} & \mathbf{0}_{n-3} & \mathbf{0}_{n-3} & \mathbf{0}_{n-3} \\ \mathbf{0}_{n-3}' & g_1^{-1} & 0 & 0 \\ \mathbf{0}_{n-3}' & 0 & g_2^{-1} & 0 \\ \mathbf{0}_{n-3}' & 0 & 0 & g_3^{-1} \end{bmatrix}. \tag{2.6}$$

We study the function $\det\left(\times'\mathbf{G}^{-1}\times\right)$. Because $\mathbf{X}_1'\mathbf{X}_1 = \frac{n-3}{4(p-1)}\left(p\mathbf{I}_p + (p-2)\mathbf{1}_p\mathbf{1}_p'\right)$, then $\det\left(\mathbf{X}_1'\mathbf{X}_1\right) \le (p-1)\left(\frac{(n-3)p}{4(p-1)}\right)^p$ and the maximum is attained if and only if $\frac{(n-3)p}{4(p-1)}$ and $\frac{(n-3)(p-2)}{4(p-1)}$ are integers. Consequently,

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) \le (p-1)\left(\frac{(n-3)p}{4(p-1)}\right)^p \det\left(\mathbf{I}_3 + \begin{bmatrix} g_1\mathbf{x}_1' \\ g_2\mathbf{x}_2' \\ g_2\mathbf{x}_3' \end{bmatrix}\left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\left[\mathbf{x}_1\mathbf{x}_2\mathbf{x}_3\right]\right).$$

Since $\left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1} = \frac{4(p-1)}{(n-3)p}\left(\mathbf{I}_p - \frac{p-2}{p(p-1)}\mathbf{1}_p\mathbf{1}_p'\right)$, then

$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) \le (p-1)\left(\frac{(n-3)p}{4(p-1)}\right)^p \det\left(\mathbf{T}_3\right)$, where

$\mathbf{T}_3 = \mathbf{I}_3 + \begin{bmatrix} g_1\mathbf{x}_1' \\ g_2\mathbf{x}_2' \\ g_2\mathbf{x}_3' \end{bmatrix}\frac{4(p-1)}{(n-3)p}\left(\mathbf{I}_p - \frac{p-2}{p(p-1)}\mathbf{1}_p\mathbf{1}_p'\right)\left[\mathbf{x}_1\mathbf{x}_2\mathbf{x}_3\right]$. By extension, we obtain

$\det\left(\mathbf{T}_3\right) = \mathrm{A}_1 + 2\mathrm{B}_1\mathrm{C}_1 - \mathrm{D}_1$, where $\mathrm{A}_1 = \prod_{h=1}^{3}\left(1 + \frac{4(p-1)g_h}{(n-3)p}\left(t_h - \frac{p-2}{p(p-1)}t_h^2\right)\right)$,

$\mathrm{B}_1 = \prod_{h=1}^{3}\frac{4(p-1)g_h}{(n-3)p}$, $\mathrm{C}_1 = \prod_{h,s}\left(u_{hs} - \frac{p-2}{p(p-1)}t_h t_s\right)$, $h,s = 1,2,3, h \ne s$,

$\mathrm{D}_1 = \sum_{h,s,z}\left(\frac{4(p-1)g_h}{(n-3)p}\right)^2\left(u_{hs} - \frac{p-2}{p(p-1)}t_h t_s\right)^2\left(1 + \frac{4(p-1)g_z}{(n-3)p}\left(t_z - \frac{p-2}{p(p-1)}t_z^2\right)\right)$,

$h,s,z = 1,2,3, h \ne s \ne z$. Because $t_h - \frac{p-2}{p(p-1)}t_h^2 \le \frac{p^3+8}{4p(p-1)}$, then the equality is fulfilled if and only if $t_h = \frac{p+2}{2}, h = 1,2,3$. Hence, $\det\left(\mathbf{T}_3\right) \le \mathrm{A}_2 + 2\mathrm{B}_2\mathrm{C}_2 - \mathrm{D}_2$,

where $\mathrm{A}_2 = \prod_{h=1}^{3}\left(1 + \frac{4(p-1)g_h}{(n-3)p}\frac{p^3+8}{4p(p-1)}\right)$, $\mathrm{B}_2 = \left(\frac{4(p-1)}{(n-3)p}\right)^3\prod_{h=1}^{3}g_h$,

$$\mathrm{C}_2 = \prod_{h,s} \left(u_{hs} - \frac{(p-2)(p+2)^2}{4p(p-1)} \right),$$

$$\mathrm{D}_2 = \left(\frac{4(p-1)}{(n-3)p} \right)^2 \sum_{h,s,z} g_h g_s \left(u_{hs} - \frac{(p-2)(p+2)^2}{4p(p-1)} \right)^2 \left(1 + \frac{4(p-1)g_z}{(n-3)p} \left(\frac{p^3+8}{4p(p-1)} \right) \right),$$

$h,s,z = 1,2,3, h \neq s \neq z$. After calculations, we have:

$$\det(\mathbf{T}_3) = \mathrm{A}_3 + 2\mathrm{B}_2\mathrm{C}_2 - \mathrm{D}_3$$

$$\mathrm{A}_3 = \prod_{h=1}^{3} \left(1 + \frac{g_h (p^3+8)}{(n-3)p^2} \right),$$

$$\mathrm{D}_3 = \left(\frac{4(p-1)}{(n-3)p} \right)^2 \sum_{h,s,z} g_h g_s \left(u_{hs} - \frac{(p-2)(p+2)^2}{4p(p-1)} \right)^2 \left(1 + \frac{g_z (p^3+8)}{(n-3)p^2} \right).$$

In order to maximise $\det(\mathbf{T}_3)$, we need to calculate a maximum for integer number u_{hs}.

Now, we consider two cases: $p \equiv 0 \, mod \, 4$ and $p+2 \equiv 0 \, mod \, 4$.

If $p \equiv 0 \, mod \, 4$, then minimum value of $u_{hs} - \frac{(p-2)(p+2)^2}{4p(p-1)}$ equals

$\left(\frac{p^2+8}{4p(p-1)} \right)^2$ for $u_{hs} = 0.25(p+4)$. In this case, $\det(\mathbf{T}_3) \leq \mathrm{A}_3 + 2\mathrm{B}_2\mathrm{C}_4 - \mathrm{D}_4$, where

$\mathrm{C}_4 = \left(\frac{p^2+8}{4p(p-1)} \right)^3$, $\mathrm{D}_4 = \left(\frac{p^2+8}{(n-3)p^2} \right)^2 \sum_{h,s,z} g_h g_s \left(1 + \frac{g_z (p^3+8)}{(n-3)p^2} \right)$ and

$\det(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}) \leq (p-1) \left(\frac{(n-3)p}{4(p-1)} \right)^p \{\mathrm{A}_4 + 2\mathrm{B}_2\mathrm{C}_4 - \mathrm{D}_4\}$. So, we can formulate the following theorem.

Theorem 2.5. Let $p \equiv 0 \, mod \, 4$. Any spring balance weighing design $\mathbf{X} \in \mathbf{\Psi}_{n \times p}(0,1)$ in the form (2.5) with the variance matrix of errors in the form (2.6) is highly D-efficient in the class $\boldsymbol{\Psi}_{n\times p}(0,1)$ if and only if $t_h=\frac{p+2}{2}$ and $\frac{(n-3)p}{4(p-1)}$ and $\frac{(n-3)(p-2)}{4(p-1)}$ are integer numbers, $u_{hs}=0.25(p+4), h=1,2,3$.

Definition 2.4. Let $p\equiv 0\,mod\,4$. Any spring balance weighing design $\mathbf{X}\in\boldsymbol{\Psi}_{n\times p}(0,1)$ in the form (2.5) with the variance matrix of errors in the form (2.6) is highly D-efficient if;

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)=(p-1)\left(\frac{(n-3)p}{4(p-1)}\right)^p$$

$$\left\{\begin{array}{l}\prod_{h=1}^{3}\left(1+\frac{g_h\left(p^3+8\right)}{(n-3)p^2}\right)+2\left(\frac{p^2+8}{(n-3)p^2}\right)^3\prod_{h=1}^{3}g_h-\\ \left(\frac{p^2+8}{(n-3)p^2}\right)^2\sum_{h,s,z}g_hg_s\left(1+\frac{g_z\left(p^3+8\right)}{(n-3)p^2}\right)\end{array}\right\}.$$

Now, we turn to the case $p+2\equiv 0\,mod\,4$.

If $p+2\equiv 0\,mod\,4$, then $u_{12}=0.25(p+2)$ and

$$u_{12}-\frac{(p-2)(p+2)^2}{4p(p-1)}=\frac{p+2}{4}-\frac{(p-2)(p+2)^2}{4p(p-1)}=-\frac{(p+2)(p-4)}{4p(p-1)}.$$

In this case, $\det(\mathbf{T}_3)\le \mathrm{A}_3+2\mathrm{B}_2\mathrm{C}_5-\mathrm{D}_5$, where $\mathrm{C}_5=\left(\frac{(p+2)(p-4)}{4p(p-1)}\right)^3$,

$$\mathrm{D}_5=\left(\frac{(p+2)(p-4)}{(n-3)p^2}\right)^2\sum_{h,s,z}g_hg_s\left(1+\frac{g_z\left(p^3+8\right)}{(n-3)p^2}\right) \text{ and}$$

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)\le(p-1)\left(\frac{(n-3)p}{4(p-1)}\right)^p\left\{\mathrm{A}_5+2\mathrm{B}_5\mathrm{C}_5-\mathrm{D}_5\right\}.$$

Theorem 2.6. Let $p+2\equiv 0\,mod\,4$. Any spring balance weighing design $\mathbf{X}\in\boldsymbol{\Psi}_{n\times p}(0,1)$ in the form (2.5) with the variance matrix of errors in the form (2.6) is highly D-efficient if and only if $t_h=\frac{p+2}{2}$ and $\frac{(n-3)p}{4(p-1)}$ and $\frac{(n-3)(p-2)}{4(p-1)}$ are integer numbers, $u_{12}=0.25(p+2), h=1,2,3$.

Definition 2.5. Any spring balance weighing design $\mathbf{X}\in\boldsymbol{\Psi}_{n\times p}(0,1)$ in the form (2.53) with the variance matrix of errors in the form (2.6) is highly D-efficient if:

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)=(p-1)\left(\frac{(n-3)p}{4(p-1)}\right)^{p}$$
$$\left\{\begin{array}{c}\prod_{h=1}^{3}\left(1+\frac{g_h\left(p^3+8\right)}{(n-3)p^2}\right)+2\left(\frac{(p+2)(p-4)}{(n-3)p^2}\right)^{3}\prod_{h=1}^{3}g_h-\\ \left(\frac{(p+2)(p-4)}{(n-3)p^2}\right)^{2}\sum_{h,s,z}g_h g_s\left(1+\frac{g_z\left(p^3+8\right)}{(n-3)p^2}\right)\end{array}\right\}$$

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)=(p-1)\left(\frac{(n-3)p}{4(p-1)}\right)^{p}$$
$$\left\{\begin{array}{c}\prod_{h=1}^{3}\left(1+\frac{g_h\left(p^3+8\right)}{(n-3)p^2}\right)+2\left(\frac{(p+2)(p-4)}{(n-3)p^2}\right)^{3}\prod_{h=1}^{3}g_h-\\ \left(\frac{(p+2)(p-4)}{(n-3)p^2}\right)^{2}\sum_{h,s,z}g_h g_s\left(1+\frac{g_z\left(p^3+8\right)}{(n-3)p^2}\right)\end{array}\right\}.$$

Example 2.5. Let us consider the variance matrix of errors $\sigma^2\mathbf{G}$, where $\mathbf{G}^{-1}=\begin{bmatrix}\mathbf{I}_6 & \mathbf{0}_6 & \mathbf{0}_6 & \mathbf{0}_6\\ \mathbf{0}_6' & 5 & 0 & 0\\ \mathbf{0}_6' & 0 & 3 & 0\\ \mathbf{0}_6' & 0 & 0 & 7\end{bmatrix}$. We determine a highly D-efficient design in the class $\mathbf{X}\in\mathbf{\Psi}_{9\times4}(0,1)$. So, take the highly D-efficient spring balance weighing design $\mathbf{X}_1\in\mathbf{\Psi}_{6\times4}(0,1)$ in the form $\mathbf{X}_1=\begin{bmatrix}1&1&0&0\\1&0&1&0\\1&0&0&1\\0&1&1&0\\0&1&0&1\\0&0&1&1\end{bmatrix}$. If $\mathbf{x}_1'=[1\ 1\ 1\ 0]$, $\mathbf{x}_2'=[1\ 1\ 0\ 1]$, $\mathbf{x}_3'=[1\ 0\ 1\ 1]$, then $\mathbf{X}=\begin{bmatrix}\mathbf{X}_1\\ 1\ 1\ 1\ 0\\ 1\ 1\ 0\ 1\\ 1\ 0\ 1\ 1\end{bmatrix}\in\mathbf{\Psi}_{9\times4}(0,1)$ is a highly D-efficient spring balance weighing design.

Example 2.6. Let us consider the variance matrix of errors $\sigma^2\mathbf{G}$, where $\mathbf{G}^{-1}=\begin{bmatrix}\mathbf{I}_{10} & \mathbf{0}_{10} & \mathbf{0}_{10} & \mathbf{0}_{10}\\ \mathbf{0}_{10}^{'} & 5 & 0 & 0\\ \mathbf{0}_{10}^{'} & 0 & 3 & 0\\ \mathbf{0}_{10}^{'} & 0 & 0 & 7\end{bmatrix}$. We determine a highly D-efficient design in the class $\mathbf{X}\in\mathbf{\Psi}_{13\times 6}(0,1)$. So, take the highly D-efficient spring balance weighing design $\mathbf{X}_1\in\mathbf{\Psi}_{10\times 6}(0,1)$ in the form $\mathbf{X}_1=\begin{bmatrix}1&1&0&0&1&0\\1&1&0&0&0&1\\1&0&1&1&0&0\\1&0&1&0&0&1\\1&0&0&1&1&0\\0&1&1&1&0&0\\0&1&1&0&1&0\\0&1&0&1&0&1\\0&0&1&0&1&1\\0&0&0&1&1&1\end{bmatrix}$. If $\mathbf{x}_1^{'}=[1\ 1\ 1\ 1\ 0\ 0]$, $\mathbf{x}_2^{'}=[1\ 1\ 0\ 0\ 1\ 1]$, $\mathbf{x}_3^{'}=[0\ 0\ 1\ 1\ 1\ 1]$, then $\mathbf{X}=\begin{bmatrix}\mathbf{X}_1\\ 1\ 1\ 1\ 1\ 0\ 0\\ 1\ 1\ 0\ 0\ 1\ 1\\ 0\ 0\ 1\ 1\ 1\ 1\end{bmatrix}\in\mathbf{\Psi}_{13\times 6}(0,1)$ is a highly D-efficient spring balance weighing design.

3. Conclusions

Here, the theory and practice of the spring balance weighing design is presented. It is not possible to determine the D-optimal spring balance weighing design in any class $\mathbf{\Psi}_{n\times p}(0, 1)$. Because of this, new construction methods of highly D-efficient designs in such classes are considered. The methods of determining highly D-efficient designs in classes in which D-optimal spring balance weighing designs have not been determined so far in the literature are presented in the examples. It is worth noting that in the highly D-efficient spring balance weighing design we are able to determine unknown measurements of objects with the minimum product of variances of their estimators.

It is worth emphasising that spring balance weighing designs can be applied in all experiments in which the experimental factors are at two levels, see, for example, Ceranka and Graczyk (2014). Let us suppose that we study the real estate market and we are interested in the influence of the following factors: the prospect of further price increases in the local housing market, availability of loans for the purchase of apartments, the prospect of increasing VAT, fears related to the liquidation of the interest relief, the current price increase observed in the local market, availability of housing in the secondary market (each at two levels coded with 1 or 0). From the statistical point of view, we are interested in determining the influences of these factors using twenty different combinations. In the notation of weighing designs, we determine unknown measurements of $p = 6$ objects in $n = 13$ surveys, so we consider the class $\mathbf{\Psi}_{13\times 6}(0,1)$. The scheme of determination of the measurements, i.e. the design matrix, is given in the example 2.6. Possible applications of the discussed designs should be sought wherever the measurement results can be written as a linear combination of unknown object measures with coefficients equal to 0 or 1.

This paper was presented at the conference MSA 2019 which financed its publication. Organisation of the international conference "Multivariate Statistical Analysis 2019" (MSA 2019) was supported by resources for the popularisation of scientific activities of the Minister of Science and Higher Education in the framework of agreement No 712/PDUN/202019.

References

Banerjee K. S. (1975), *Weighing Designs for Chemistry, Medicine, Economics, Operation Research, Statistics*, Marcell Dekker Inc., New York.

Bulutoglu D. A., Ryan K. J. (2009), *D-optimal and near D-optimal 2^k fractional factorial designs of Resolution V*, "Journal of Statistical Planning and Inference", no. 139, pp. 16–22.

Ceranka B., Graczyk M. (2010), *Some construction of optimum weighing designs*, "Acta Universitatis Lodziensis. Folia Oeconomica", no. 235, pp. 235–239.

Ceranka B., Graczyk M. (2012), *Notes on the optimum chemical balance weighing designs*, "Acta Universitatis Lodziensis. Folia Oeconomica", no. 269, pp. 91–101.

Ceranka B., Graczyk M. (2014), *On certain A-optimal biased spring balance weighing designs*, "Statistics in Transition. New Series", Spring, vol. 15(2), pp. 317–326.

Ceranka B., Graczyk M. (2018), *Highly D-efficient designs for even number of objects*, "Revstat-Statistical Journal", no. 6, pp. 475–486.

Ceranka B., Graczyk M. (2019), *A highly D-efficient spring balance weighing designs for an even number of objects*, "Acta Universitatis Lodziensis. Folia Oeconomica", no. 344, pp. 17–27.

Ceranka B., Katulska K. (1987a), *Zastosowanie optymalnych sprężynowych układów wagowych*, [in:] *Siedemnaste Colloquium Metodologiczne z Agro-Biometrii*, PAN, Warszawa, pp. 98–108.

Ceranka B., Katulska K. (1987b), *Zastosowanie teorii sprężynowych układów wagowych do analizy doświadczeń z mieszankami*, "Listy Biometryczne", no. XXIV, pp. 17–26.

Ceranka B., Katulska K. (1989), *Application of the biased spring balance weighing theory to estimation of differences of line effects for legume content*, "Biometrical Journal", no. 31, pp. 103–110.

Gail Z., Kiefer J. (1982), *Construction methods for D-optimum weighing designs when n ≡ 3 mod 4*, “The Annals of Statistics”, no. 10, pp. 502–510.
Gawande B.N., Patkar A.Y. (1999), *Application of factorial design for optimization of Cyclodextrin Glycosyltransferase production from Klebsiella Pneumoniae AS–22*, “Biotechnology and Bioengineering”, no. 64, pp. 168–173.
Graczyk M. (2013), *Some applications on weighing designs*, “Biometrical Letters”, vol. 50(1), pp. 15–26.
Jacroux M., Notz W. (1983), *On the Optimality of Spring Balance Weighing Designs*, “The Annals of Statistics”, no. 11, pp. 970–978.
Katulska K., Smaga Ł. (2010), *On some construction of D-optimal chemical balance weighing designs*, “Colloquium Biometricum”, no. 40, pp. 37–45.
Koukouvinos Ch. (1996), *Linear models and D-optimal designs for n ≡ 2 mod 4*, “Statistics and Probability Letters”, no. 26, pp. 329–332.
Raghavarao D. (1971), *Constructions and Combinatorial Problems in Designs of Experiments*, John Wiley Inc., New York.
Shah K.R., Sinh B.K. (1989), *Theory of Optimal Designs*, Springer Verlag, Berlin.

Uwagi o wysoce D-efektywnych sprężynowych układach wagowych

Streszczenie: W artykule rozważamy nowe metody konstrukcji wysoce D-efektywnych sprężynowych układów wagowych w klasach, w których nie istnieje układ D-optymalny. Podajemy warunki wyznaczające relacje pomiędzy parametrami tych układów oraz przykłady konstrukcji.

Słowa kluczowe: sprężynowy układ wagowy, układ D-efektywny

JEL: C02, C18, C90

Received: 2020-03-29; verified: 2020-10-02. Accepted: 2021-07-01

Folia Oeconomica
Acta Universitatis Lodziensis
ISSN 0208-6018 e-ISSN 2353-7663
www.czasopisma.uni.lodz.pl/foe/
2(353) 2021

DOI: https://doi.org/10.18778/0208-6018.353.05

Jakub Karnowski
SGH Warsaw School of Economics, Collegium of Economic Analysis
Department of International Comparative Studies, Warsaw, Poland, kubakarn@sgh.waw.pl

Andrzej Rzońca
SGH Warsaw School of Economics, Collegium of Economic Analysis
Department of International Comparative Studies, Warsaw, Poland, andrzej.rzonca@gmail.com

Directions for the Reconstruction of the Tax System in Poland – a Growth-Enhancing Proposal

Abstract: The paper aims to present directions for the growth-enhancing reconstruction of the tax system in Poland. It presents a diagnosis of the main strengths and weaknesses of that system. Based on this diagnosis and a review of the literature, the authors propose a package of recommendations whose introduction would be conducive to economic growth. The recommendations include: shifting the burden of taxation from income, in particular low labour income, to consumption; exempting low earners from a part of social security contributions; the introduction of the possibility for local governments to increase the PIT-free allowance above the centrally set base amount; the unification of the basis for the PIT, National Health Fund and Social Insurance Institution contributions; the elimination of differences in contributions for different types of contracts on the basis of which work is performed; the extension of one-off amortisation to all machine investments; and the elimination of sectoral taxes.

Keywords: tax system, taxes, pro-growth proposal, impact of taxes on economic growth

JEL: H20, H21, H24, H25, H29, O10, O11, O43

1. Introduction

After the fall of socialism, Poland achieved economic success without precedent in its history. In 1990–2010, the increase in per capita GDP was about six times as strong as in 1918–1938.[1] Since 1989 per capita GDP has increased by more than two and a half times. Poland has repeated the economic miracle of Germany: since 1989 per capita GDP in the former country has followed almost the same path as in the latter country after 1955.[2] Poland is now more or less as far away from Germany in terms of per capita GDP as it was from Hungary in 1990, which it surpassed in 2012.[3] Poland has also surpassed two countries of the 'old' European Union (EU): Greece and Portugal. Nevertheless, it still lags far behind the EU average.

In order to eliminate that gap, Poland has to strengthen the elements that contributed to its economic success after 1989. Those were free market reforms, introduced earlier than in other post-socialist countries (see, e.g., Balcerowicz, 1995 or Aslund, Djankov, 2014) and macroeconomic policies that avoided large imbalances (see, e.g., Bakker, Gulde, 2010). If Poland does not continue with its reforms or leaves itself little room for manoeuvre in macroeconomic policy, the Polish economic miracle will come to an end. Even before the pandemic, long-term forecasts indicated that economic growth in Poland would slow down to just 1% after 2040 without further reforms (European Commission, 2015a). Although Poland would remain a country with a fairly high per capita income, it would never catch-up to even poorer countries of the 'old' EU, such as Spain.

Taxation is an area where growth-enhancing reforms are possible and particularly desirable for at least three reasons.

Firstly, taxes are an important factor in determining economic growth. Their burden depends on productive behaviours: whether and how much the economic agent (taxpayer) works, acquires new skills, saves, invests, innovates, etc.

Secondly, the effects of changes in taxes on growth occur more rapidly than those of many other structural reforms. In the past, they were estimated to become clearly visible after five to ten years, i.e., only after one or even two parliamentary terms (see Kneller, Stevens, 2006). However, according to more recent studies, this period is much shorter (and does not exceed three to five years).

Thirdly, the negative impact of taxes on economic growth is strongly felt in Poland.[4] Entrepreneurs complain that taxes are the main barrier to their development

1 The authors' calculations based on data from Maddison-Project, http://www.ggdc.net/maddison/maddison-project/home.htm [accessed: 3.05.2021].

2 The authors' calculations based on data from The Conference Board Total Economy Database™.

3 The authors' calculations based on the IMF World Economic Outlook Database.

4 Compare, for example, data from the Executive Opinion Survey conducted annually by the World Economic Forum and published in subsequent editions of 'The Global Competitiveness Report'.

more frequently than on average in the countries of the region. They particularly and frequently point to tax regulations as a barrier to development (see Figure 1). Moreover, the percentage of entrepreneurs complaining about the amount of tax burden increases. Recently, it has once again exceeded the average for countries in the region (see Figure 2). All in all, taxes are growing in importance as a barrier to development compared to other barriers, such as the quality of infrastructure and corruption (see Figures 3 and 4).

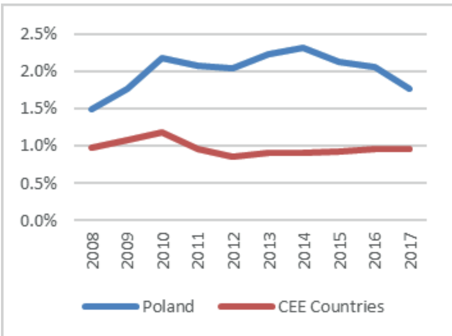


Figure 1. Tax regulations as a barrier to business development (illustrated as a percentage of responses)

Source: World Economic Forum, 2016

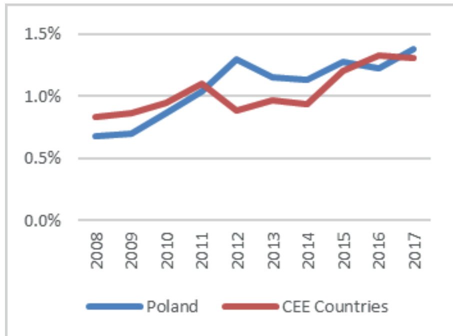


Figure 2. Tax burden as a barrier to business development (illustrated as a percentage of responses)

Source: World Economic Forum, 2016

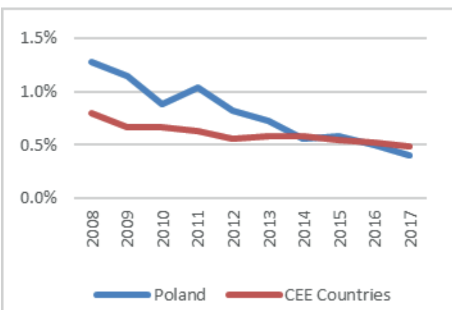


Figure 3. Insufficient infrastructure as a barrier to business development (illustrated as a percentage of responses)

Source: World Economic Forum, 2016

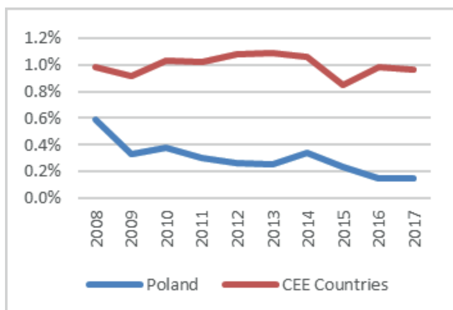


Figure 4. Corruption as a barrier to business development (illustrated as a percentage of responses)

Source: World Economic Forum, 2016

The aim of the article is to present directions for the growth-enhancing reconstruction of the tax system in Poland. We are aware that there is a wide range of tax reform proposals in Poland. However, they are most often weakly related to the literature concerning effects of taxes on economic growth (with only few exceptions – see, e.g. Bukowski, Morawski, Trzeciakowski, 2016). We seek to fill this gap.

The analysis presented below is conducted in three steps that correspond to the main sections of the article.

In section two, we summarise how taxes affect growth. We do that by referring to a simple endogenous growth model. We also define the possibilities of changes in the tax system in Poland in the coming years from the perspective of the sustainability of general government.

In section three, we diagnose the tax system in Poland. We present its strengths that should be preserved or augmented if one wants to enhance economic growth. Above all, however, we identify its weaknesses that need to be removed or at least mitigated.

In section four, we outline the general directions of desired changes in the tax system in Poland. The changes meet three conditions resulting from the analysis carried out in the previous sections: firstly, they are growth enhancing; secondly, they are potentially sustainable as they do not undermine the sustainability of general government; thirdly, they enhance the strengths of the tax system in Poland or mitigate its weaknesses. The last two conditions imply that the shape of the tax system presented in the article does not maximise economic growth as only changes achievable in the foreseeable future are considered.

Section five concludes the analysis.

2. Effect of taxes on economic growth

Taxes are hereafter understood as (most often) monetary, compulsory, and non-refundable benefits borne to the State (see Figure 5). The monetary form is the only form of tax payment recognised by Polish law. This manner of regulating taxes is the most convenient for the majority of citizens as well as for the State.[5] The compulsory nature of taxation means that its payment, as well as the amount, does not depend on the benevolence of citizens. If some people try not to pay the tax, the State has the right to force them to pay, applying appropriate sanctions. If taxes were not compulsory, the State would quickly stop earning most of its revenue, as the so-called free rider problem would appear. Finally, it follows from the non-refundability of taxes that the State, having collected taxes in the amount provided for by law, does not have to reimburse any portion of it after some time. In the article, benefits that have the above-mentioned characteristics are treated as taxes, regardless of their name. According to this definition, taxes are all types of compulsory contributions that finance the State's expenditures.[6]

5 In the past, tax also took other forms – it was very often collected, for example, in agricultural crops.

6 Sometimes the definition of taxes includes another feature, i.e. their non-reciprocal nature. That feature means absence of a direct 'quid-pro-quo' between taxpayers and the State wherein

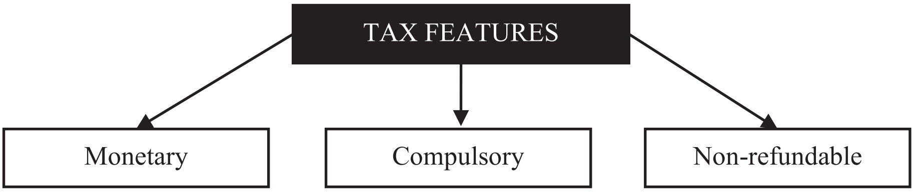


Figure 5. Tax features
Source: authors' elaboration

Taxes can generally be divided into two groups: direct and indirect (see Figure 6). The former one is imposed on private individuals' income (possibly also on their property), while the latter is charged on expenses – most often consumption expenses.

Taxes can distort incentives for productive behaviours: whether and how much to work, acquire new skills, save and invest, or innovate. In particular, direct taxes are distortionary, especially in an economy open to capital and labour flows.[7] However, even lump-sum taxation, which serves as a reference for estimating deadweight losses from other types of taxes, does not affect incentives for productive behaviours only as long as economic agents do not expect any change in their tax burden. Any reference of a formally lump-sum tax to variables that depend on productive behaviours makes it a type of tax that distorts decisions concerning these behaviours.[8]

Each of these decisions, in turn, influences economic growth (see Figure 7). The higher the percentage of people working in an economy, the higher per capita income the economy can achieve. A similar effect occurs with an improvement in the efficiency with which labour and the existing capital stock are used. Innovations are even more powerful, as they can influence not only the level of income but its growth rate as well. The more innovations an economy introduces, the faster it grows. In turn, intensity of innovation is partly linked to investments, as a part of innovation is embodied in new machines and devices. Besides, investments are

paying taxes does not give taxpayers the right to make any claims against the State. This characteristic is not met by some of the contributions (e.g. by paying a pension contribution one acquires the right to a pension in the future, while paying the health insurance contribution, one can use the services of healthcare institutions that have concluded an appropriate contract with the National Health Fund, etc.). However, their impact on development is qualitatively similar to other taxes. For this reason, a broader definition of tax has been adopted in the article.

7 In some studies, only their impact proves to be statistically significant (see Easterly, Rebelo, 1993).

8 At the same time, even a lump-sum tax fixed over time, if it were high enough, could have an impact on incentives and, as a result, economic growth. This would happen if lump-sum tax exceeded the level of consumption (before tax introduction) of at least one economic agent, as a result of which that agent would have to cover part of it by reducing savings.

a source of capital, which, if broadly defined, also depends upon qualifications by employees (human capital). Capital is another factor determining the level of per capita income. Finally, investments require financing, and national savings are the most stable source of financing. The higher the savings in a country, the easier it is to finance investments.

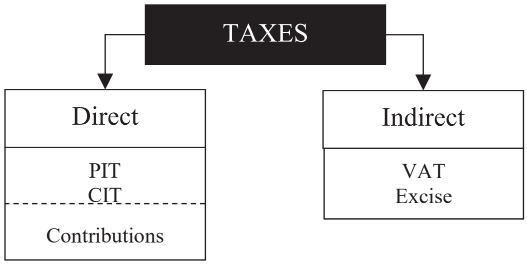


Figure 6. Tax breakdown
Source: authors' elaboration

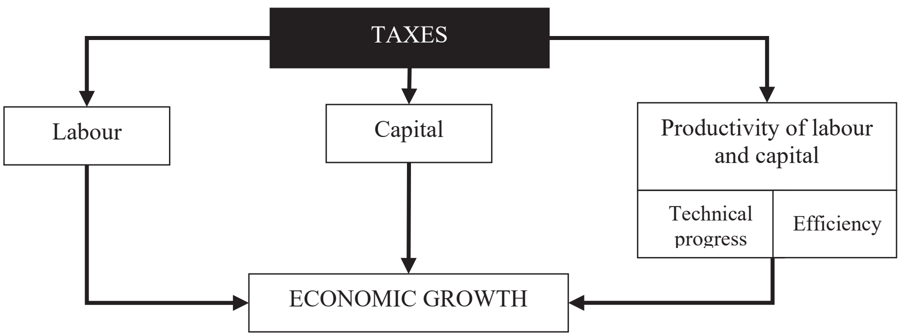


Figure 7. Influence channels of taxes on economic growth
Source: authors' elaboration

Even if taxes distort productive behaviours, and thereby negatively affect economic growth, they are indispensable.

The State cannot be replaced by the private sector in conducting some tasks fundamental to economic growth. These tasks include protection of individuals and their property (Keefer, Knack, 1995; Sala-i-Martin, 1997), national defence, as long as it does not consume more resources than are needed to convince people that the country is not seriously threatened from abroad (Landau, 1996; Baffes, Shah, 1998; Aizeman, Glick, 2003), and basic research (see, e.g., Jaumotte, Pain, 2005).

Some tasks that the State carries out support economic growth but can also be carried out by the private sector. However, the private sector will not carry out the tasks to the extent that will provide maximum benefits because it only takes into account the part of the benefit for which it can be paid and omits indirect benefits for which it cannot bill their recipients, i.e. positive externalities. These tasks (referred to as merit goods) include ensuring wide access to quality infrastructure, education, and health care. Their implementation by the State accelerates economic growth as long as the spending of tax money is guided by economic rational and not by political calculation (Aschauer, 1989a; 1989b; 1989c; 2000a; 2000b; Baffes, Shah, 1998; Ramirez, Ranis, Stewart, 2000; Bleaney, Gemmell, Kneller, 2001; Miller, Tsoukis, 2001; Gyimah-Brempong, Wilson 2004). For example, the State builds roads where there will be car traffic, and provides education for children, not only a 'school stay'.

There are also tasks of the State that do not, in themselves, affect economic growth and yet increase welfare. In particular, pensions for the elderly or for disabled do not accelerate economic growth, but society is better-off in those countries where they exist.

That said, the State also undertakes tasks that are detrimental to economic growth and welfare, even when the costs of taxes financing these tasks are not taken into account. For example, benefits for people who are capable of work can incentivise them to exit the labour market. Even worse, this behaviour can be inherited because children observe first hand that it is possible to get by without work (Hansson, Henrekson, 1994; Atkinson, 1999; Mares, 2007; Afonso, Allegre, 2011; Clemente, Marcuello, Montañés, 2012; Afonso, Jalles, 2014). In turn, subsidies to unprofitable sectors or companies inhibit the reallocation of capital and labour from where they are not productive to where they would be (Aiginger, Falk, 2005).

The increase in possible benefits for economic growth (or welfare) derived from government spending decreases with increasing levels of spending (Hulten, 1996; Afonso, Schuknecht, Tanzi, 2003; 2006; Baldacci et al., 2004). On the one hand, proper supervision of expenditures becomes more difficult and costly as the bureaucracy grows. On the other hand, the possibilities of productive use of expenditures become limited. Moreover, there is a growing risk of them being misdirected.[9] Meanwhile, the tax burden from which they are financed increases. The accelerating cost of taxes in combination with the slower-growing benefits derived from government spending mean that, at a certain level of expenditure, the costs of taxes are equal to the benefits of expenditure.

This level maximises economic growth. According to different estimates, it varies between 17% and 25% of GDP (see, e.g. Smith, 2006; 2016; Skrok, 2013). In particular, it is lowered by the politicisation of the public sector, while it is increased by the efficiency of this sector and the ability of the State to obtain income

9 For example, in only six OECD countries, mostly with government spending below the OECD median (i.e. the Czech Republic, Canada, South Korea, and Switzerland), one tenth of the lowest-income households receive more than one tenth of social transfers. In five countries, this share is 4% or less. In three of these countries (Greece, Portugal and Italy), government expenditures are high (OECD, 2017).

without raising tax rates to levels that significantly weaken the incentives for productive behaviours. These possibilities increase with per capita income. The structure of the economy is changing in a way that makes tax evasion more difficult, at the same time reducing the relative benefits of engaging in those economic activities where such evasion is easy. Government expenditures of Asian Tigers (Hong Kong, South Korea, Singapore, and Taiwan) range from 17 to 25% of GDP (see Figure 8). They were similar in the West until the 1960s (Tanzi, Schuknecht, 2000).[10] Among Western countries, they returned to those levels in Ireland (see Figure 8).

The range of government spending that allows an economy to maximise social welfare, resulting both from growth and a sense of social security, is higher and amounts to 30%–35% of GDP (cf. Tanzi, Schuknecht, 2000; Tanzi, 2008; Smith, 2016). This is the range of government expenditures in New Zealand, Switzerland, and Lithuania; government expenditures are slightly higher in Australia, the United States, and in Latvia in the CEE region (see Figure 8).

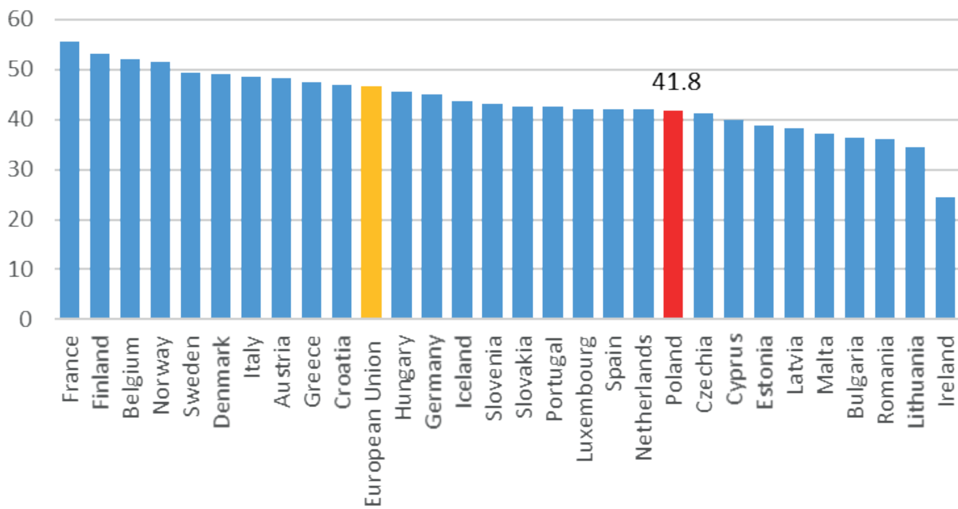


Figure 8. General government expenditures in the EU Member States. Data for 2019 (as a percentage of GDP)

Source: Eurostat, 2021a

10 Low government spending does not guarantee rapid economic growth as money does not have to be spent on objectives that favour it. Nor does it have to mean lower taxes. Low tax revenues may be the result of low tax collection of taxes that are set at a high level or, worse, at a discretionary level by officials. High or arbitrarily set official taxes give officials a wide range of possibilities for the extortion of bribes, i.e. a corruption tax. As a result of a high corruption tax, total forced payments become large and uncertain (Frye, 2001).

In Poland, government expenditure amounts to almost 43% of GDP.[11] It is, therefore, about 20% of GDP above the level that allows for maximising economic growth and about 10% of GDP above the level allowing it to maximise social welfare. In the region, only Slovenia and Hungary have higher government expenditures. They are also lower in most of the developed countries to which Poland is still aspiring to catch up. At the same time, even fewer of those countries had equally high expenditures at Poland's current stage of development.

Leibfritz, Thornton, and Bibbee (1997) conduct a simple cross-sectional study isolating the impact of non-tax factors on the growth of the economy. They show that the increase in the tax-to-GDP ratio in the OECD countries, from about 30% in the 1960s to about 40% in the 1990s, reduced the average annual growth rate by about 0.5 pps or by approximately one fifth. A review of more recent empirical studies confirms the conclusion that an increase (reduction) of the tax-to-GDP ratio of 10 pps slows down (accelerates) the growth rate on average by 0.5–1 pps (see, e.g., Bergh, Henrekson, 2011; 2016). Calibrating endogenous growth models to the characteristics of OECD economies gives results that depend on the model design, but they usually turn out to be even stronger (see, e.g., King, Rebelo, 1990). Nevertheless, it should be noted that there is no consensus among economists whether taxes affect the long-term growth rate or only the level of income in the long term (see, e.g., Koester, Kormendi, 1989).

Research on the countries in our region indicates a significantly stronger impact of taxes on economic growth than in developed countries, on which most research focuses. In particular, an analysis carried out by Skrok (2013) for the new EU Member States shows that a reduction in the tax-to-GDP ratio of 1pps led to an increase in their growth ranging from 0.2–0.5 pps, provided that the reduction concerned income taxes, and the loss of revenue in the budget was matched by a reduction in social spending.

One may summarise the main points of the above-presented discussion by referring to the learning-by-doing model, i.e. a simple endogenous growth model (cf. Rzońca, 2005). The following assumptions are made in the model:

The production function of an individual entrepreneur is as follows:

$$Y_i(t) = (K_i(t))^{\alpha} \left(B\left(K(t)/L(t)\right)^{\beta} L_i(t) \right)^{1-\alpha}, \tag{1}$$

where: $B\left(K(t)/L(t)\right)^{\beta} = A(t)$ – level of technology ('knowledge') in the economy; Y_i – individual output; K_i – individual capital input; L_i – individual labour input; K – aggregate capital input; L – aggregate labour input; $0 < B < e^{1/}\alpha$; $0 < \alpha < 1$; $\beta \geq 0$.

This form of the microeconomic production function reflects two basic assumptions of the learning-by-doing model: the production process provides companies

11 Data acquired from the IMF World Economic Outlook Database. The information refers to the pre-pandemic period. In 2020, due to economic costs of COVID–19, general government expenditure in Poland exceeded 50% of GDP. Yet, this increase is likely to be short lived.

with knowledge how to produce more efficiently and the knowledge spillovers freely throughout the economy. Note that the level of technology ('knowledge') is slightly modified compared to macroeconomics textbooks. It depends on capital per unit of labour, instead of capital. Due to such a modification, the production function shows constant economies of scale both at the micro and macro level, regardless of the adopted product elasticities with respect to individual production factors.

The aggregate production function is easily obtained as a result of summing up the production functions of all (m) companies in the economy.

$$Y=\sum_{i=1}^{m}Y_i=\sum_{i=1}^{m}\Big((A(t))^{1-\alpha}(k(t))^{\alpha}L_i(t)\Big)=(A(t))^{1-\alpha}(k(t))^{\alpha}\sum_{i=1}^{n}\big(L_i(t)\big)=$$
$$=(A(t))^{1-\alpha}(k(t))^{\alpha}L(t)=(A(t))^{1-\alpha}\left(\frac{K(t)}{L(t)}\right)^{\alpha}L(t)=K(t)^{\alpha}\big(A(t)L(t)\big)^{1-\alpha}, \tag{2}$$

where: $k=K/L$. Labour input grows at a constant rate (n).

$$\dot{L}(t)=nL(t), \tag{3}$$

where: $n\geq 0$.

All households are the same. Their number is equal to the labour input (each household has a unit of labour). Households own companies; hence, all income generated by the companies goes to them. Households pay tax on their income. They maximise their utility, which is a function of their consumption over an infinite time horizon. The utility function is defined by the following formula:

$$U_s=\int_s^{\infty}e^{-\rho(t-s)}u(c)\,dt, \tag{4}$$

where: $u(c)$ – instantaneous utility; c – consumption of an individual household (consumption per unit of labour); ρ – discount rate.

The State collects income tax. Its rate (τ) depends neither on the source nor on the amount of income of an economic agent. All government revenues come from the income tax. Thus its rate also determines the degree of fiscalism.

$$\tau(t)=\frac{T(t)}{Y(t)}=\frac{\tau(t)Y(t)}{Y(t)}=\frac{G(t)}{Y(t)}, \tag{5}$$

where: $\tau\in[0;1)$; T – tax revenue; G – government expenditure.

The State allocates some part of its revenue to finance the supply of merit goods. Let us call this part the government capital accumulation rate. The capital accumulation rate in the economy is the average of the capital accumulation rates in government and private sectors weighted by the share of these sectors in the total income.

$$s = s_p(1-\tau) + s_s\tau, \tag{6}$$

where: s_p – the capital accumulation rate in the private sector; s_s – the capital accumulation rate in the government sector.

The State can also provide public goods that can be treated as an additional production factor.

$$Y_i(t) = h(g)(K_i(t))^{\alpha}\left[A(t)L_i(t)\right]^{1-\alpha}, \tag{7}$$

where: g – spending on public goods to output ratio (as it is not so much the absolute volume of supply of these goods that is important for the production process but their availability); h – the function, which is non-negative and growing slower and slower, which reflects falling albeit always positive marginal productivity of public goods. For example, the following function meets these conditions:

$$h(g) = (a + bg)^{\gamma}, \tag{8}$$

where: a, b, γ – parameters; $a \geq 0$; $b > 0$; $0 < \gamma < 1$.

The model has stable and non-zero growth, only if $\beta = 0$ (i.e., externalities of capital accumulation are large enough to neutralise the decreasing marginal productivity of capital at the micro level, but not larger). If $\beta < 0$, the growth rate is converging towards zero, due to decreasing marginal productivity of capital. The model simplifies to the neoclassical growth model (with no technical progress). In that case, long term growth is unaffected by any of the fiscal variables. They can only influence the output level per unit of labour in the steady state. As this influence is qualitatively quite similar to their effects on growth when $\beta = 0$, we focus on that case.[12]

When the State does not provide public goods, then growth of output per unit of labour is given by the following formula:

$$\frac{\dot{y}(t)}{y(t)} = \frac{(1-\tau)\alpha B^{1-\alpha} + s_s\tau\alpha B^{1-\alpha} - \rho - n}{\theta}, \tag{9}$$

where: $\theta = -\frac{c(t)u''(c(t))}{u'(c(t))} = \text{const}$ – relative risk aversion (greater than zero and not equal to one).

It follows that income tax reduces the growth rate, except when the State spends all tax revenues on merit goods. The greater the fiscalism, the slower the growth. However, if the tax were not an income tax but a lump sum (not distorting

12 For $\beta > 0$, growth of output per unit labour would be exploding. As this result hardly corresponds to reality, at least in the long term, it is not analysed further.

the marginal product of capital), then the tax burden (at least up to a point) would not affect growth.

With strongly distortionary taxation, the State is hardly able to accelerate growth. This is because capital accumulation by the private sector is a decreasing function not only of distortionary taxation but also of capital accumulation by the State.

$$\frac{S_p}{Y}=\frac{(1-\tau)\alpha}{\theta}+\frac{n(\theta-1)-\rho}{\theta\beta^{1-\alpha}}-\frac{s_s\tau\cdot(\theta-\alpha)}{\theta}. \tag{10}$$

Government spending on merit goods crowds out some private investment. Economic agents postpone their consumption less if the State takes on the burden of accumulating the capital (broadly defined) necessary to ensure an adequate level of future consumption. Note however, that the crowding-out effect would be weaker if capital accumulated by the State were not a perfect substitute for capital accumulated by the private sector.

The State in the model could raise the capital accumulation only if it set its investment at such a level that would reduce private investment to zero. As a result, the State would take over the entire capital stock in the economy. However, as international experience shows, with State ownership of capital, the process of learning by doing is at least slower (β drops below unity). The initial growth may be fast but it slows to zero over time – no matter how much the rate of capital accumulation has increased.

If the State supplies public goods, the growth equation changes to the following form:

$$\frac{\dot{y}(t)}{y(t)}=\frac{(1-\tau)(a+bg)^{\gamma}\alpha B^{1-\alpha}+s_s\tau(a+bg)^{\gamma}\alpha B^{1-\alpha}-\rho-n}{\theta}. \tag{11}$$

This equation shows that there is a range within which increasing the share of expenditure on public goods in the output accelerates the long-term growth more than it is decelerated by the increase in distortionary taxation necessary to finance the increase in expenditure. For the adopted form of the h function, the right-hand border of this interval is the expenditure on public goods relative to the output (g) that maximises the following expression: $(1-g)(a+bg)^{\gamma}$. That is:

$$\frac{d\left[(1-g)(a+bg)^{\gamma}\right]}{dg}0 \Leftrightarrow g=\frac{\gamma b-a}{\gamma b+b}<0.5. \tag{12}$$

At the end of this section, it should be emphasised that one should not think about a tax reform that would result in an increase in general government deficit. A deficit usually means higher taxes in the future. This is implied by dynamic efficiency, which requires the interest rate in the long term to be higher than the GDP growth rate (see, e.g., Romer, 2000). With such a relationship between the interest rate and the growth rate, a deficit even in a single period means, on average, a higher tax burden in other periods. Additional tax revenue is necessary to cover at least part of the interest on the sovereign debt incurred to finance the deficit.

Deficits have a negative impact on economic growth also through mechanisms other than future tax increases. Firstly, a deficit worsens the structure of government expenditures and makes it easier to direct them towards goals that do not benefit society. Secondly, by consuming private savings, a deficit makes it difficult to finance investments. Although the issue of Ricardian equivalence is a subject of dispute among economists, most empirical studies indicate that households increase their savings by 20% to 50% of the deficit increase (see, e.g., Gale, Orszag, 2003). Thus, the pool of savings that can finance investments is reduced by 50% to 80% of the deficit. Thirdly, by raising the interest rate and increasing uncertainty, including uncertainty about future tax burdens, a deficit discourages economic agents from making investments. Fourthly, by increasing exchange rate volatility, it hinders both exports and imports and, as a result, weakens the country's connection to the world economy. Finally, sometimes it ends in a crisis (for more, see, e.g. Ciżkowicz, Rzońca, 2011).

Therefore, tax reconstruction has a chance to speed up economic growth, provided that it does not increase the budget deficit. If it were to include tax cuts, it should be accompanied by a reduction in government spending that is detrimental, neutral, or insignificant to economic growth. If the reconstruction of taxes deepens the budget deficit, taxes will have to be increased in the following periods. Due to additional mechanisms, including uncertainty about the level of taxes, the deficit may be more damaging to economic growth than increases in taxes that eventually appear due to the deficit.

In the foreseeable future, there is no room for tax changes that would deplete budget revenue. There is a high general government deficit. According to early estimates by the Central Statistical Office, in 2020, the year in which the COVID–19 pandemic broke out, it amounted to almost 7% of GDP. Tax revenues decreased, and expenditures increased significantly. According to the European Commission,[13] in 2021, only the structural part of the deficit will exceed 4% of GDP. This heralds serious fiscal tensions in the near future.

They may be exacerbated by interest rate increases. Currently, interest rates in Poland are at their historical low, so that in spite of high sovereign debt perilously close to the constitutional limit of 60% of GDP, the cost of debt servicing consumes a record-low percentage of GDP (in 2021, it will be 1.4%, almost half of that in 2012,[14] and in 2022, it is expected to fall even further to 1.3% of GDP[15]).

13 European Economic Forecast, Institutional Paper 136, November 2020.

14 Data from the Ameco European Commission database.

15 European Economic Forecast, Institutional Paper 136, November 2020.

However, following a period of low rates, interest rate increases always occur. Growing inflation is likely to bring Poland closer to that point.

The already adopted regulations ensuring an expenditure increase for a specific purpose in the coming years ahead of economic growth are yet another factor that will exacerbate tensions in general government. The increase includes expenditures on pensions (an outcome of lowering the retirement age and the introduction of 13th and 14th pension), national defence (from 2% of GDP to 2.5% of GDP in 2030), health care (from 4.7% to 6.5% of GDP in 2025), and science (from 0.4% of GDP to 1.0% of GDP).

This does not mean that nothing should be done about taxes until the deficit has been reduced. Not all taxes have the same effect on work, savings, investments, and innovations. At the same time, the influence of taxes on these behaviours depends not only on the amount of taxes but also on the uncertainty about that amount, and the costs of fulfilling tax obligations. Change in the structure of taxes, the reduction of uncertainty about their amount, the facilitation of tax payment, and the reduction of tax compliance costs are areas that can have a positive impact on economic growth, and they are the focus of changes proposed by the authors.

The acceleration of economic growth that would ensure these changes would facilitate the reduction of the government expenditure-to-GDP ratio. Its decline after the elimination of the deficit would create room for tax cuts and further growth benefits. Therefore, the presented proposals should be treated as the first stage of tax reform in a direction that would maximise economic growth and welfare in Poland.

3. Strengths and weaknesses of the tax system in Poland

Taxes in Poland, assessed in terms of their impact on economic growth, have both strengths and weaknesses.

Their main strength is the relatively high proportion of indirect taxes from consumption in government revenues, including levies from a VAT (see Figure 9). This feature is more important than on average in the European Union. However, for most new Member States, its importance is even greater.[16] Thus, there is a room for improvement of this positive feature in Poland's tax structure. The high proportion of consumption taxes in the government revenues is beneficial because these taxes weaken the incentives for work, investments and entrepreneurship to a lesser extent than other type of taxes.

16 Data from the European Commission.

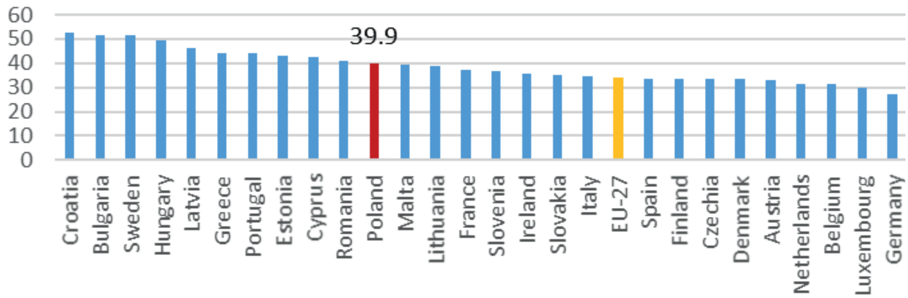


Figure 9. Share of indirect taxes in general government revenues (%). Data for 2019
Source: European Commission, 2021

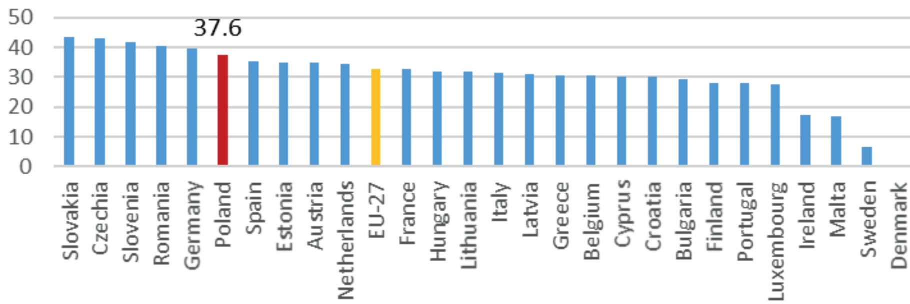


Figure 10. Share of social security contributions in general government revenues (%). Data for 2019
Source: European Commission, 2021

In theory, taxation of consumption should be equivalent to the taxation of labour income (Layard, Nickell, Jackman, 1996; Stiglitz, 2004: 605–609). Two crucial factors determine utility maximised by households: consumption and leisure. Consumption is determined by income, which is primarily derived from work. Work, however, consumes time. Both taxation of labour income and taxation of consumption distort the 'price' relationship between leisure and consumption in the same way. The higher the taxes on consumption or labour income, the more expensive consumption becomes relative to leisure. As a result, both taxes weaken people's incentive to work.[17] However, in practice there is no such equivalence, if only because taxes on consumption affect the expenses of everyone, including those who do not work and those who consume imported goods, i.e., goods produced by working abroad, while taxation of labour income is by definition imposed on those who work only within the borders of a given country. This difference would not be of great importance if unemployed persons lived off the savings they had accumulated

17 One can imagine a situation where a VAT tax would increase the incentive to work. This would be the case if it was imposed only on goods complementary to leisure time (e.g. tourist services), and income obtained in this way was used to subsidise goods complementary to labour (e.g. childcare) – see Corlett and Hague (1953).

during their employment and if consumers purchased only domestically produced goods. In practice, the majority of unemployed persons make a living from other people's work, and consumers buy at least some imported goods. While taxation of consumption does not distort the relations between the prices of consumption and labour, including work performed abroad, the taxation of income from labour limits the consumption possibilities only of those working in the country.

Martinez-Vazquez, Vulovic, and Liu (2009) provide a broad overview of the effects of shifts between taxes on consumption and labour income, respectively. A smaller negative impact of consumption taxes on economic growth than taxes on labour income is pointed out by Widmalm (2001), Dahlby (2003), International Monetary Fund (2004), European Commission (2006), Arnold (2008), OECD (2010), Gemmel, Kneller, and Sanz (2011; 2014), Thomas and Picos-Sánchez (2012), Pestel and Sommer (2017).

Another positive feature of taxes in Poland is a moderate, compared to most EU countries, level of the upper income tax rates, both for companies (CIT) and on personal income (PIT) (see Figure 11). However, the difference in marginal CIT between Poland and the EU on average has been systematically shrinking in recent years. Still, the rates of both types of income taxes alone are less of a disincentive to productive behaviours than in the EU countries. In particular, the lower taxation of high personal income than in the West is beneficial for the accumulation of human capital, i.e., for improving one's own qualifications.

High levels of skilled labour support economic growth in at least four ways. Firstly, they encourage the search for new and more efficient production techniques. Secondly, they are often a necessary condition for the introduction and development of technologies invented abroad (see, e.g., Nelson, Phelps, 1966; Griffith, Redding, Van Reenen, 2005). Thirdly, they facilitate productivity growth of low-skilled workers. This is achieved not only by improving the organisation of work, which is made possible by high skills of their superiors, but also by transferring knowledge, even unknowingly, regarding how to produce goods and services more efficiently (see, e.g., Feldstein, 1973). Fourthly, they provide an income that gives the opportunity to increase savings (see, e.g., Kaldor, 1956), which are a source of investment financing.[18]

Individuals with high skills, and consequently income, constitute a group whose professional activity is particularly sensitive to the level of taxation (Disney, 2000 and works cited therein). Satisfied consumption needs increase the importance they attach to leisure. Self-employment, which they choose more often than other society members, gives them the freedom to shape their working time, which is not available to contract workers (Showalter, Thurston, 1997). Simultaneously, work of highly skilled individuals affects the demand for unskilled labour. Their professional work forces them to buy many services that they cannot perform

18 For example, in Canada and Germany, the wealthiest one-fifth of the population puts aside one fifth of their income, and in the UK, one fourth. In all these countries, the poorest one-fifth of the population has no savings at all.

on their own due to lack of free time – they are able to do so because of their higher income (Feldstein, 1973). The importance of relatively low taxation on highly qualified persons (experts, managers, specialists, etc.) in terms of economic growth is increased by their freedom to work abroad. The low tax rate decreases the incentive for economic emigration, which would otherwise be strong as long as labour productivity and, as a result, wages in Poland differ significantly from those in the West.

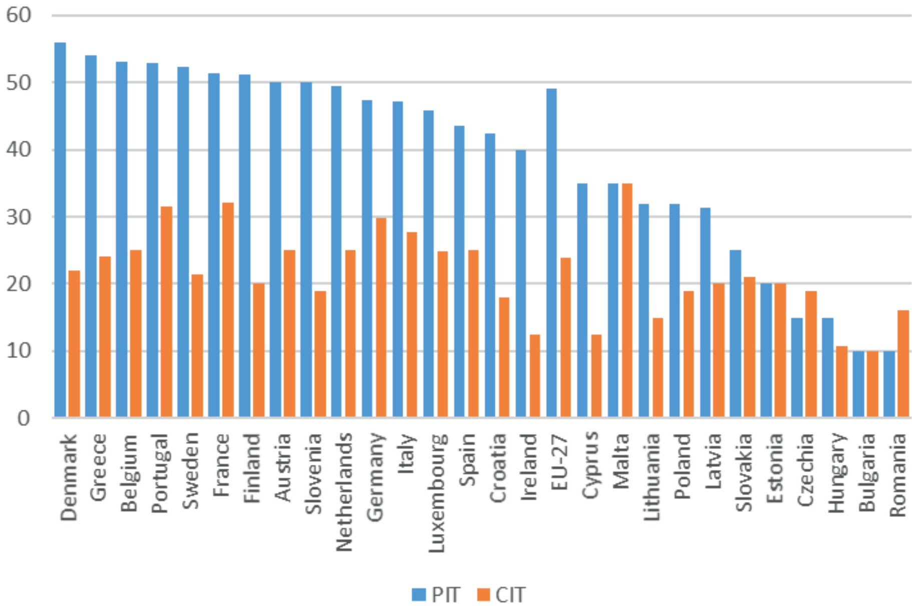


Figure 11. The highest CIT and PIT rate in the EU countries in 2020 (%)

Source: European Commission, 2021

The negative aspects of Poland's tax structure include a very high proportion of government revenues from social security contributions, which can be treated as a form of taxation of labour income (see Figure 10). This proportion is clearly higher than in the countries of our region and higher than in the West.[19] The level of taxation of labour income is mitigated by a significantly lower personal income tax than in the West.[20] However, due to the low PIT-free allowance compared to other countries, this mitigation applies to a small extent to low income from work. While the combined PIT and contributions on the average wage in Poland is clearly below the OECD average, in the case of low wages (at the level of half the average wage), it corresponds to the average of OECD countries and is much higher than in countries such as Ireland, the UK, or the Netherlands.[21]

19 Data from the European Commission.

20 Data from the European Commission.

21 OECD data.

Individuals with low income comprise the group whose work is the most sensitive to taxation (Disney, 2000). Low income is usually either a reflection of low skills or a result of family duties that make it difficult to adapt time to employers' requirements. For such people, a financially attractive alternative may be living off benefits or on the salary of a family member, including parents, the more so as working involves some costs (e.g., travel or purchase of services previously performed by themselves).

In Poland, these costs, and as a consequence, the impact of taxation on incentives of unskilled labour to work are reinforced by the structure of social benefits. Seeking employment results in either a total loss of benefits or a reduction of benefits by an amount equal to the increase in income from employment. As a result, the average effective taxation of income of people moving from unemployment to employment exceeds 80% and this is the seventh highest value in the EU.[22]

Another weakness of taxes in Poland is higher taxation of income from investment in machinery than in the region. While the effective taxation of profits in general does not differ significantly from the average in the region, the difference is clear in the case of profits from investment in machinery. All countries in the region tax those profits less than Poland (Spengel, Elschner, Endres, 2012).

On average, investments in machinery are more conducive to economic growth than other types of investment (De Long, Summers, 1991). These investments more often lead to technical progress, which, in addition to improving efficiency, is a source of productivity growth in the economy. Investment and productivity growth are also hindered by further tax system weaknesses in Poland.

One of the hindrances to the growth of investment and productivity is the complexity of taxes. Poland's taxes are regulated by 11 Acts and 292 Regulations for a total of 5,789 pages. It would take almost 284 hours to read them (Grant Thornton, 2017a). In the opinion of entrepreneurs, taxes are not as complex in any EU country as they are in Poland. Moreover, the degree of their complexity is a barrier to development, which negatively distinguishes Poland from the countries in the region (see the Introduction).

A manifestation of this complexity is the large amount of time that an average entrepreneur has to allocate to fulfilling tax obligations. In Poland entrepreneurs utilise 260 hours per year, compared to 161 hours on average in the EU and EFTA countries, and 50–82 hours in countries such as Estonia, Ireland, and Luxembourg. Only two EU countries, Hungary and Bulgaria, require more time devoted to taxes than Poland (PwC UK, World Bank Group, 2017). Micro-enterprises in Poland have to submit an average of 30 different tax returns per year, medium enterprises – 41, and large one – 67.

Taxes are complicated by numerous exceptions and tax reliefs. Poland has the second highest share of tax preferences in VAT in the EU after Spain, even higher

22 Data from the European Commission.

than Greece. In terms of forgone revenue from this tax, they are almost twice as expensive as the EU average (CASE, 2016). If not for these preferences, the VAT rate could be less than 17%,[23] instead of 23%. The same good may be taxed at different VAT rates depending on its purpose, the type of activity of the entrepreneur, or minor differences in composition or state. Reduced rates are applied not only by Poland, but also by the other EU countries, with the exception of Denmark. However, two such rates are not equally common. There is only one reduced rate in eight EU countries (see Appendix).

In turn, if it were not for the numerous CIT reliefs, its rate could be 13% instead of 19%.[24] In relation to PIT, such a calculation is more difficult due to the different charging forms of this tax. In the case of taxpayers accounting for the tax scale, the average income burden after deducting social security contributions is less than 9%. Exceptions and preferences also apply to social security contributions. Thus, depending on the legal form of the agreement under which labour income is received, it may be charged with very different overall amount of taxes. For the minimal remuneration of PLN 2,800, that amount may stands at 0% (a mandate contract, an employee is a student below the age of 26) or 41% (an employment contract, an employee over the age of 26, a participant of the Employee Capital Schemes (PPK)). For the remuneration of PLN 20,000, it may amount to 16% (self-employment) through 23% to 45% (an employment contract, an employee is over the age of 26, a participant of the PPK). At the same time, individual taxes and contributions have different bases for calculating their amount.

Tax preferences distort the demand for individual goods and services as well as their production costs. In this way, they create barriers to the reallocation of labour and capital to their most efficient use. The directions of investments and changes in employment are not entirely determined by the profitability of individual projects, but to a large extent, by differences in taxation.

The complicated law causes a great deal of confusion. The tax administration in Poland issues over 30,000 tax interpretations per year, with the average length being five pages (Grant Thornton, 2017a). A single person does not have the physical ability to check all of these documents. At the same time, one cannot count on a friendly approach of the tax administration to problems with the interpretation of unclear laws. Although the clause for resolving doubts of interpretation is in favour of the taxpayer (*in dubio pro tributario*), the tax administration very rarely uses it (Grant Thornton, 2017c).

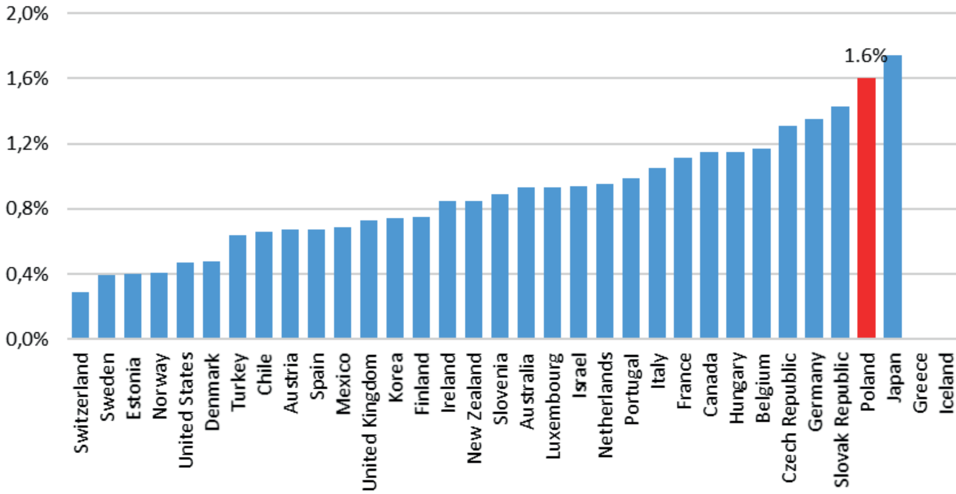
The tax administration is often wrong in its interpretations. Court challenges occur in about one third of the unfavourable decisions to taxpayers. It takes a great deal of time for taxpayers wrongly accused of tax understatement to clear themselves of charges. On average, the time frame is 56 months, with the longest trial

23 FOR estimates based on MF and CASE data (2016).

24 FOR estimates based on MF data.

lasting 20 years. Moreover, before taxpayers receive a favourable judgment, they must pay the disputed amount (Grant Thornton, 2017a).

The frequent overruling of tax administration decisions by the courts is a symptom of a more general problem – low efficiency of tax administration. Among the OECD countries, only in Japan is the cost of tax collection (in relation to tax revenues) higher than in Poland (see Figure 12).



No data referring to Greece and Iceland

Figure 12. Tax collection costs in the OECD countries (as a percentage of tax revenue). Data for 2013 (Latest available)

Source: OECD, 2015

Disputes with the tax administration reduce average capacity utilisation, and thus productivity of labour and capital. Often, disputes limit the scope of companies' activities and sometimes result in a permanent cessation of business. The risk of liquidation, regardless of which party of the dispute is right, the entrepreneur or the tax administration, increases with the length of the court proceedings. Some of the resources available to companies (e.g., knowledge of how to reach customers or use machines) may be so specific that it is difficult for other companies to take over and use them fully. Moreover, some companies may be reluctant to enter into transactions with entrepreneurs who have serious problems with the tax administration.[25] Non-use of resources, if long enough, further restricts their use in new applications – unmaintained machines break down, buildings deteriorate, and workers lose their qualifications along with the ability to improve them.

25 This does not mean that economic efficiency is in conflict with the pursuit of tax fraud. The truth is precisely the opposite. However, the State should not only prosecute tax fraud but above all reduce the incentives to commit it (see, e.g. Smith, 1954, vol. II: 586–587).

The complicated tax law reduces productivity in the economy also by the unproductive behaviour it causes (cf. Slemrod, Yitzhaki, 2002). Meeting tax obligations requires unnecessary expenditures that could be used in a productive manner. With simple taxes, it would be less costly to keep proper records. There would be fewer unclear or contentious cases, so tax audits would be less frequent and shorter. Taxpayers would have to appeal to courts against unfavourable decisions of the tax administration less frequently. Finally, the State expenditure on taxpayer control could be reduced without the risk of increased tax fraud.[26]

Above all, complicated tax laws lower productivity by inhibiting reallocation of labour and capital to the most efficient uses and reducing investment in high-performance technologies. Such laws encourage economic agents to develop activities in which they have as limited relationship with the tax administration as possible. In addition to the activities covered by the tax preferences already mentioned, this condition is met by small-scale and non-innovative activities; they are not conspicuous, and at the same time, they do not require large production assets against which the tax administration could pursue its claims.

A review of research on the determinants of companies' size confirms that it significantly depends on the quality of the tax system (Friesenbichler et al., 2014). The high percentage of micro-enterprises in countries with complicated taxes is not only the result of the choice made by entrepreneurs to reduce interactions with the tax administration. It also reflects the smaller possibilities of micro-businesses overcoming barriers to development, including covering the costs of complying with tax obligations. Taking Canada, New Zealand and the UK, for example, these costs consume 2% of annual sales in a company with revenues below $50,000, but only 0.04% in companies with sales above one million dollars (Government Accountability Office, 2011).

The small scale of operations reduces the productivity of companies because it does not allow them to achieve economies of scale, including investing in high-performance technologies, which often requires large scale production (see, e.g., Pagano, Schivardi, 2003; Castany, López-Bazo, Moreno, 2005; Van Biesebroeck, 2005; Leung, Meh, Terajima, 2008; Melitz, Ottaviano, 2008; Braguinsky, Branstetter, Regateiro, 2011). The impact on productivity is particularly negative if a business is run in the shadow economy. It is not possible there to ask the State to enforce claims against unreliable counterparties. Therefore, business is conducted mainly without deferred payments or with clients who are well known and trusted. The former limits the size of individual transactions, while the latter limits the circle of customers. Modern economic growth requires depersonalised transactions (North, 1993). In addition, due to the risk of both detection and coming across an unreliable customer, it is practically impossible to conduct large or long-term contracts that would guarantee the recovery of at least part of the investment.

26 Cf. Smith, 1954: vol. II, p. 587.

The greater the importance of economic activities that are hardly visible to the tax administration, the more oppressive the tax administration becomes towards other entities in order to obtain planned budget revenues. This further weakens incentives to scale up activities and invest. At the same time, there are more reasons to stay out of the market sector and produce various types of goods and services for personal use only. By definition, such production closes the possibilities of both specialisation and economies of scale.

The complicated system results in numerous loopholes continuously discovered by taxpayers. The legislators in turn attempt to fight these gaps by adding more laws. At the same time, the opaqueness of the tax system makes it easier for interest groups to introduce new preferences. All this, together with the casuistic nature of tax law, leads to the high variability of tax regulations, which is another weakness of the tax system in Poland. So far, each tax law has been amended on average about twice a year. In the years 2012–2014, Poland produced 56 times more tax regulations than in Sweden, where tax law is the most stable in the EU, and much more than in other countries of the region. For example, Poland produced 11 times more tax regulations than Lithuania and twice as many as Hungary. In 2016, the tax law inflation accelerated even more – twice as many regulations were produced than the average in the previous few years (Grant Thornton, 2017a). In total, 1,784 pages of laws and regulations were introduced, thus amending almost one third of the tax laws. One of the amendments to the Personal Income Tax Act was pushed through the entire legislative process, including the signature of the President, within one day (Grant Thornton, 2017b). In the following years, the prolific enactment of tax regulations continued.

The instability of tax law makes individuals uncertain about the tax burdens. The shape of the tax system becomes an additional risk. This risk not only discourages investment but also, once it has materialised, leads to a waste of some resources in the economy. If no changes in taxes are foreseen at the project implementation stage, and the estimates of its profitability prove to be overestimated, then the investment may become worthless. The revenue may no longer cover the costs and the capital may not be (fully) used in other applications due to its specific nature.

Companies can partially protect themselves against the threat of unstable taxes by replacing capital with labour from the shadow economy. This allows them to quickly adapt to changes in the environment that are difficult to predict. These companies may dismiss employees not covered by any contract or limit their wages (or even not pay them). Poorly equipped with machines, they are characterised by low productivity, which makes it difficult to cover labour costs. However, avoiding non-wage labour costs (which operating in the shadow economy enables) mitigates this difficulty.

It is worth noting that complicated and unstable tax regulations mutually reinforce their negative impact on investment and productivity in the economy. On the one hand, if the tax system was complicated but stable, it would require a one-off cost to understand. Where the regulations are continuously changed, the costs of tracking them never end. On the other hand, if taxes were frequently changed but simple, it would be possible to identify potential scenarios and prepare for each of them. When the number of parameters in the tax system subject to change is high, the analysis of potential scenarios becomes very difficult, if possible at all. This difficulty discourages investments in projects that cannot be abandoned without serious losses; by contrast, it does encourage small-scale operations without significant production assets, especially of specific nature.

Finally, the high degree of centralisation of the tax system in Poland should be considered a weakness. There are local taxes in Poland. Local governments have also a share in income tax revenues. However, their share in general government revenues is lower in Poland than in the average EU country (see Figure 13).

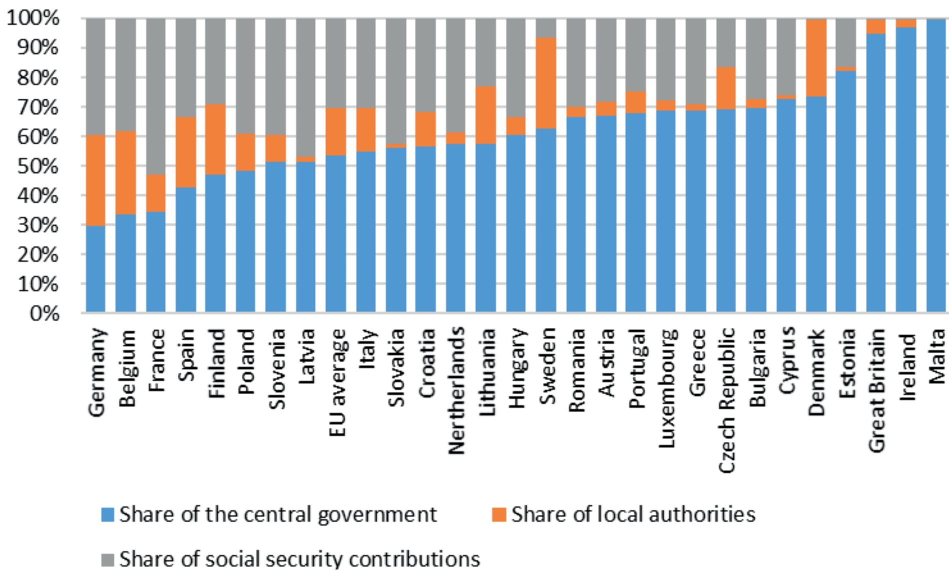


Figure 13. Shares in general government revenues (%). Data for 2019

Source: Eurostat, 2021b

Research indicates that if the expenditure of local governments is not based on their own revenues but on transfers, local authorities do not compete with each other in creating the best possible conditions to attract economic agents (Besley, Coate, 2003). These conditions do not determine their ability to spend. This ability depends on the central government’s favouritism. As a result, the efforts of local authorities to broaden their tax base weaken – the more so when the transfers received from the central budget are (negatively) linked to their own revenues or, more generally, per capita income in a given area (Bird, Smart, 2002). Simultaneously, the resilience of local governments to shocks may weaken because they may rely on the central government and they have no provisions for self-remedy. Finally, fiscal discipline also weakens. The more central government finances local authorities, the higher the political costs it faces if it refuses to support a unit in financial difficulties (Rodden, 2005). The negative effects of transfers are more serious if the transfer system is more complicated. High complexity of transfers limits the ability of voters to evaluate public services provided by local governments (Kotsogiannis, Schwager, 2008).

4. Recommendations

The following proposals for changes to the tax system in Poland meet, as indicated in the Introduction, three conditions: they foster economic growth; they are potentially sustainable; and they improve the strengths of the tax system in Poland or mitigate its weaknesses.

An essential element of the proposed changes is to shift the burden of taxation from income, in particular low income from labour, to consumption. Shifting the tax burden instead of reducing it (at least in the initial stage of tax system reconstruction) is associated with the need to meet the second of the conditions required for such reconstruction.

Based on various models, it can be estimated that shifting the tax burden from income to consumption alone could increase GDP per capita by 1 to 8% over 10 years. Thus, the average growth rate of the economy over that period would accelerate by about 0.1–0.8 percentage points. The lower limit of this range is obtained from the European Commission's (European Commission, 2006) general equilibrium model. Other tools imply benefits close to the upper limit of this range (compare International Monetary Fund, 2004; Arnold, 2008; Gemmell, Kneller, Sanz, 2011).

This shift should be done by limiting preferences in VAT, the revenue from which would be sufficient to cover the cost of increasing the PIT-free allowance and exempting low earners from a part of their social security contributions. The range of contributions covered by the exemption should be selected in such a way that the total taxation of the highest possible labour income is as close as possible to the taxation of income from economic activity. Such a solution would help to reduce the shadow economy, as the tax benefit of illegal employment would disappear, or at least decrease. At the same time, the solution would reduce the taxation of labour income of those groups whose incentives to take up employment are most sensitive to taxes (i.e., young people, low skilled workers or those unable to work full-time). In order to further strengthen the incentive to work, it would be necessary to make the receipt of benefits by persons able to work dependent on their professional activity and to change the way in which benefits are taken away from a person whose income increased, so that the loss of benefits is gradual, rather than sudden. In particular, the allowance 500+ should be covered by this principle.

With a broad reduction in VAT preferences, including those regarding goods that are mainly consumed by the poorest individuals or those with children, the budget should allocate part of the additional revenue to compensate for them. Compensation can be made in such a way that shifting the burden of taxation from income to consumption, regardless of its scope, does not lead to a worsening of the material situation of any social group even in the short term, i.e., before its positive effects on legal employment and economic growth become apparent.

The elimination of differences in contributions for different types of contracts on the basis of which work is performed could serve as an additional source of coverage of the costs of lowering labour income taxation especially for low earners. An additional benefit of this elimination would be the simplification of the social security system and the removal of distortions when the main condition for choosing a given form of contract is not its legal and economic sense but its non-contributory nature. There would be an end to lawsuits concerning, for example, whether digging a ditch is a contract for specific work.

A significant increase in the PIT-free allowance could be achieved by transferring the entire income tax revenue to local authorities and allowing them to compete with the allowance amount. This solution would remove another flaw in the tax system in Poland, namely its excessive centralisation. Local authorities (municipal councils) could increase the allowance above the base level defined at the central level. Such tax competition between local authorities would avoid the inefficient flow of capital from poor parts of the country to the rich ones. If it contributed to the movement of people (especially the poor, i.e., those without capital) from poor parts of the country to the wealthy ones, it would strengthen the mechanism of equalisation of income per capita (Barro, Sala-i-Martin, 1991).

These changes could be complemented by an exemption from social security contributions for older and younger people. As they are characterised by lower than average labour activity in Poland (compared to the EU), this solution would not pose a risk of a significant deadweight loss. Making contributions dependent on age would not cause significant distortions in the choices made by the economic agents because age is not the subject of their choice. This dependency could shift the demand of entrepreneurs for labour towards young and old. However, in a situation of shortage of labour, this shift should not significantly reduce the chances of other age groups for employment. By exempting older and younger people from social security contributions, their net income could significantly increase while the cost of their employment for entrepreneurs would decrease.

Another proposed change is to extend the possibility of 'one-off' amortisation to all investments in machines. It would strengthen the incentive for such investments. Although they are the most conducive to economic growth, their low level mostly contributes to the low investment rate in Poland and negatively distinguishes Poland from other countries in the region (Forum Obywatelskiego Rozwoju, 2015). The extension would result in a temporary decrease in CIT revenue. Thus its implementation should depend on the scale of the preference reduction in VAT. If the scale were small, the implementation of this proposal would have to be postponed until the general government deficit has been significantly reduced.

The proposals presented above would not only alleviate the main weaknesses of taxes in Poland (i.e., a high share of social security contributions in government revenues and, as a result, a fairly high overall taxation of low labour income; higher taxation of income from investment in machinery than from other investments; a fairly high degree of centralisation of taxes) but also augment one of their strengths and maintain another one (respectively, the high share of consumption tax in government revenues and moderate marginal income taxation). They would also reduce the complexity of the tax system, and a reduction in VAT preferences could radically simplify it. If it were broad enough, it would at the same time significantly narrow the scope for fraud, significantly narrowing the VAT gap in Poland, which until recently was one of the largest in the EU (CASE, 2016). This gap, despite many measures taken in recent years to reduce it, is still quite large.

A significant simplification of taxes and/or their stabilisation would also be achieved by unifying the basis for the PIT, National Health Fund and Social Insurance Institution contributions. Consideration should also be given to replacing the casuistic definition of revenues and costs in economic activities by more general definitions. The first of these proposals would reduce the number of mathematical operations taken to calculate PIT and contributions to the National Health Fund and Social Insurance Institution from sixteen to three. It would thus reduce the risk of confusion and administrative costs for both entrepreneurs and the tax authorities. The second proposal would interrupt the continuous struggle between the Minister of Finance and entrepreneurs in finding new loopholes in the legislation. The implementation of this proposal should be accompanied by a change in the interpretation of tax law by the tax administration. This would consist of two elements. The first is a clear increase in the availability of general interpretations of tax law. The second is a gradual building of a structuralised and searchable database based on conclusions resulting, in particular, from judicial decisions.

A phasing out of sectoral taxes would also serve as a simplification, and it would create a sense of stability. These taxes are fiscally inefficient in that they generate little revenue compared to the distortions caused, weakening economic growth. In particular, these taxes are similar in nature to tax surcharges, the fears of which result in a significant reduction or dispersion of investment between different tax jurisdictions (cf. Janeba, 2000). These fears are not only limited to sectors that are already covered by this type of taxes but concern also other sectors as such taxes undermine the credibility of the government's commitment not to impose confiscatory burdens.

5. Conclusions

Almost every type of tax distorts economic agents' incentives to pursue productive behaviours, i.e., working, acquiring skills, saving and investing, as well as innovating. Simultaneously, any type of tax may be conducive to economic growth if the State allocates the revenue from it to purposes that sufficiently promote growth. Potentially growth-enhancing government expenditure includes expenditure on public goods and merit goods as well as on social expenditure that encourages/facilitates labour activity.

As government expenditure rises, the potential benefits from it fall, while the tax burden and distortions it causes increase. That said, more developed countries have larger possibilities to raise government revenue without significant distortions. With the normal relationship between interest rates and economic growth, lowering taxes without corresponding reductions in government expenditure results in the need to increase taxes in the following periods. Due to additional mechanisms, including uncertainty about the level of taxes, a deficit may be more damaging to economic growth than the eventual increases in taxes. Therefore, the reconstruction of the tax system should preferably not lead to an increase in the deficit.

Not all taxes are similarly distortionary. As a result, not only the level of taxation but also its structure is important for economic growth. This gives an opportunity to rebuild the tax system, and thereby enhance economic growth, without waiting until the general government deficit is eliminated.

Even if there is equivalence in theory between the selected types of taxes (e.g., taxation of labour and consumption), in reality, it does not work. In particular, it is blocked by the heterogeneity of economic agents. Therefore, in practice, shifting the tax burden from labour to consumption promotes economic growth.

Some taxes, which are not significant for fiscal revenues, may be very distortionary and, as a result, have an extremely adverse effects on economic growth. They include, in particular, taxation of company profits, or reliefs and exemptions from this tax, as well as sectoral taxes.

The magnitude of the impact of a given type of tax on economic growth depends not only on its amount, but also on detailed solutions that determine its shape. The magnitude of this impact is influenced mainly by two elements. The first is the cost of compliance by economic agents with their tax obligations, which affects the incentives to avoid these obligations, and the other is the possibility of avoiding a given tax, including the use of tax reliefs and exemptions. In turn, the evasion or avoidance of a given tax contributes to its frequent changes. The instability of taxes reinforces their negative impact on economic growth. In order to mitigate this impact, the reconstruction of taxes should simplify them, which in turn should promote their future greater stability.

The tax system in Poland has two main strengths, the importance of consumption tax in government revenues and moderate marginal income taxation. It also has five major weaknesses: a high share of social security contributions in government revenues and the associated relatively high total taxation of low labour income, higher taxation of income from investment in machinery than from other investments, high complexity, instability, and a fairly high level of centralisation.

The reconstruction of the tax system in Poland should preserve or enhance its basic strengths and eliminate or at least alleviate its weaknesses. For this purpose, the following should be considered:

- shifting the burden of taxation from income, in particular low labour income, to consumption; this shift should be done by reducing VAT preferences, the revenue from which would first cover the cost of raising the PIT-free allowance;
- exempting low earners from a part of their social security contributions;
- the introduction of the possibility for local governments to increase the PIT-free allowance above the centrally set base amount;
- the unification of the basis for the PIT, National Health Fund and Social Insurance Institution contributions, which would reduce the number of mathematical activities when calculating the PIT, National Health Fund and Social Insurance Institution contributions from sixteen to three;
- the elimination of differences in contributions for different types of contracts on the basis of which work is performed;
- the extension of one-off amortisation to all machine investments;
- the elimination of sectoral taxes.

References

Afonso A., Allegre J.G. (2011), *Economic Growth and Budgetary Components: A Panel Assessment for the EU*, “Empirical Economics”, vol. 41(3), pp. 703–723.

Afonso A., Jalles J.T. (2014), *Fiscal composition and long-term growth*, “Applied Economics”, vol. 46(3), pp. 349–358.

Afonso A., Schuknecht L., Tanzi V. (2003), *Public sector efficiency: an international comparison*, “ECB Working Paper”, no. 242.

Afonso A., Schuknecht L., Tanzi V. (2006), *Public Sector Efficiency: Evidence for New EU Member States and Emerging Markets*, “ECB Working Paper”, no. 581.

Aiginger K., Falk M. (2005), *Explaining Differences in Economic Growth Among OECD Countries*, “Empirica”, no. 32, pp. 19–43.

Aizeman J., Glick R. (2003), *Military Expenditure, Threats, And Growth*, “FRBSF Working Paper”, 2003–2008.

Arnold J. (2008), *Do Tax Structures Affect Aggregate Economic Growth? Empirical Evidence From a Panel of OECD Countries*, “OECD Economics Department Working Paper”, no. 643.

Aschauer D.A. (1989a), *Is Public Expenditure Productive?*, "Journal of Monetary Economics", vol. 23(2), pp. 177–200.

Aschauer D.A. (1989b), *Public Investment and Productivity Growth in the Group of Seven*, "Economic Perspectives", vol. 13(5), pp. 17–25.

Aschauer D.A. (1989c), *Does Public Capital Crowd Out Private Capital?*, "Journal of Monetary Economics", vol. 24(2), pp. 171–188.

Aschauer D.A. (2000a), *Public Capital and Economic Growth: Issues of Quantity, Finance, and Efficiency*, "Economic Development and Cultural Change", vol. 48(2), pp. 391–406.

Aschauer D.A. (2000b), *Do States Optimize? Public Capital And Economic Growth*, "Annals of Regional Science", vol. 34(3), pp. 343–363.

Aslund A., Djankov S. (2014), *The Great Rebirth: Lessons from the Victory of Capitalism Over Communism*, Peterson Institute for International Economics, Washington.

Atkinson A.B. (1999), *The Economic Consequences of Rolling Back the Welfare State*, MIT Press, Cambridge.

Baffes J., Shah A. (1998), *Productivity of Public Spending, Sectoral Allocation Choices, and Economic Growth*, "Economic Development and Cultural Change", vol. 46(2), pp. 291–303.

Bakker B.B., Gulde A.M. (2010), *The Credit Boom in the EU New Member States: Bad Luck or Bad Policies?*, "IMF Working Paper", no. 10/130.

Balcerowicz L. (1995), *Socialism, Capitalism, Transformation*, Central European University Press, Budapest.

Baldacci E., Clemens B., Gupta S., Cui Q. (2004), *Social Spending, Human Capital, and Growth in Developing Countries: Implications for Achieving MDGs*, "IMF Working Paper", no. 04/217.

Barro R.J., Sala-i-Martin X. (1991), *Convergence Across States and Regions*, "Brookings Papers on Economic Activity", vol. 22(1), pp. 107–182.

Bergh A., Henrekson M. (2011), *Government Size and Growth: A Survey and Interpretation of the Evidence*, "Journal of Economic Surveys", vol. 25(5), pp. 872–897.

Bergh A., Henrekson M. (2016), *Government Size And Growth: A Rejoinder*, "Journal of Economic Surveys", vol. 30(2), pp. 393–396.

Besley T., Coate S. (2003), *Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Approach*, "Journal of Public Economics", vol. 87(12), pp. 2611–2637.

Bird R.M., Smart M. (2002), *Intergovernmental Fiscal Transfers: International Lessons for Developing Countries*, "World Development", vol. 30(6), pp. 899–912.

Bleaney M., Gemmell N., Kneller R. (2001), *Testing the Endogenous Growth Model: Public Expenditure, Taxation, and Growth Over Long Run*, "Canadian Journal of Economics", vol. 34(1), pp. 26–57.

Braguinsky S., Branstetter L., Regateiro A. (2011), *The Incredible Shrinking Portuguese Firm*, "NBER Working Paper", no. 17265.

Bukowski M., Morawski I., Trzeciakowski R. (2016), *Zrozumiały i prosty VAT. Korzyści z ujednoliconych stawek*, WiseEuropa, Warszawa.

CASE (2016), *Study and Reports on the VAT Gap in the EU–28 Member States: 2016 Final Report*, Center for Social and Economic Research, Warsaw.

Castany L., López-Bazo E., Moreno R. (2005), *Differences in total factor productivity across firm size. A distributional analysis*, Paper presented at ERSA 2005 – 45th Congress of the European Regional Science Association, Amsterdam, 23–27 August.

Ciżkowicz, P., Rzońca, A. (2011), *Mechanizmy oddziaływania deficytu fiskalnego na wzrost gospodarki,* "Gospodarka Narodowa. The Polish Journal of Economics", vol. 251(10), pp. 1–20.

Clemente J., Marcuello C., Montañés A. (2012), *Government Social Spending and GDP: Has There Been a Change in Social Policy?*, "Applied Economics", vol. 44(22), pp. 2895–2905.

Corlett W.J., Hague D.C. (1953), *Complementarity and The Excess Burden of Taxation*, "Review of Economic Studies", vol. 21(1), pp. 21–30.

Dahlby B. (2003), *Restructuring The Canadian Tax System By Changing The Mix Of Direct And Indirect Taxes*, [in:] H.G. Grubel (ed.), *Tax Reform in Canada: Our Path to Greater Prosperity*, Fraser Institute, Vancouver, pp. 77–108.

De Long B., Summers L. (1991), *Equipment Investment and Economic Growth*, "Journal of Economics", vol. 106(2), pp. 445–502.

Disney R. (2000), *The impact of Tax and Welfare Policies on Employment and Unemployment in OECD Countries*, "IMF Working Paper", no. 00/164.

Easterly W., Rebelo S. (1993), *Fiscal Policy and Economic Growth: an Empirical Investigation*, "NBER Working Paper", no. 4499.

European Commission (2006), *Macroeconomic Effects of a Shift From Direct To Indirect Taxation: A Simulation for 15 EU Member States*, Note presented by the European Commission services (DG TAXUD) at the 72nd meeting of the OECD Working Party No. 2 on Tax Policy Analysis and Tax Statistics, Paris, 14–16 November.

European Commission (2015a), *The 2015 Ageing Report: Economic And Budgetary Projections for The EU–27 Member States 2013–2060*, Brussels.

European Commission (2015b), *VAT rates*, https://ec.europa.eu/taxation_customs/business/vat/eu-vat-rules-topic/vat-rates_en [accessed: 3.05.2021].

European Commission (2021), *Taxation and customs union. Data on taxation*, https://ec.europa.eu/taxation_customs/business/economic-analysis-taxation/data-taxation_en [accessed: 3.05.2021].

Eurostat (2021a), *General Government Expenditure By Function*, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Government_expenditure_by_function_%E2%80%93_COFOG [accessed: 3.05.2021].

Eurostat (2021b), *Government revenue, expenditure and main aggregates*, https://ec.europa.eu/eurostat/databrowser/view/gov_10a_main/default/table?lang=en [accessed: 3.05.2021].

Feldstein M. (1973), *On The Optimal Progressivity of the Income Tax*, "Journal of Public Economics", vol. 2(4), pp. 357–376.

Forum Obywatelskiego Rozwoju (2015), *Następne 25 lat: Jakie reformy musimy przeprowadzić, by dogonić Zachód?*, Warszawa.

Friesenbichler K., Fritz O., Hölzl W., Misch F., Streicher G., Yeter M. (2014), *The Efficiency Of EU Public Administration in Helping Firms Grow*, European Commission, DG Enterprise and Industry, Brussels.

Frye T. (2001), *Keeping Shop. The Value of the Rule of Law in Warsaw and in Moscow*, [in:] P. Murell (ed.), *Assessing the Value of the Law in Transition Economies*, University of Michigan Press, Ann Arbor, pp. 229–248.

Gale W.G., Orszag P.R. (2003), *The economic effects of long-term fiscal discipline*, "Urban-Brookings Tax Policy Centre Discussion Paper", no. 8.

Gemmell N., Kneller R., Sanz I. (2011), *The Timing and Persistence of Fiscal Policy Impacts on Growth: Evidence from OECD Countries*, "Economic Journal", vol. 121(1), pp. F33–F58.

Gemmell N., Kneller R., Sanz I. (2014), *The Growth Effects of Tax Rates In The OECD*, "Canadian Journal of Economics", vol. 47(4), pp. 1217–1255.

Government Accountability Office (2011), *Value-Added Taxes. Potential Lessons For The United States From Other Countries' Experiences*, United States Government Accountability Office Report, GAO–11–867T.

Grant Thornton (2017a), *Podatki w Polsce*, Warszawa.

Grant Thornton (2017b), *Produkcja prawa w Polsce największa w historii*, Warszawa.

Grant Thornton (2017c), *Fiskus nadal nie ma wątpliwości*, Warszawa.

Griffith R., Redding S.J., Van Reenen J.M. (2005), *Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries*, "Review of Economics and Statistics", vol. 86(4), pp. 883–895.

Gyimah-Brempong K., Wilson M. (2004), *Health, Human Capital And Economic Growth in Sub-Saharan African and OECD Countries*, "Quarterly Review of Economics and Finance", no. 44, pp. 296–320.

Hansson P., Henrekson M. (1994), *A New Framework for Testing the Effect of Government Spending on Growth and Productivity*, "Public Choice", vol. 81(3–4), pp. 381–401.

Hulten Ch.R. (1996), *Infrastructure Effectiveness as a Determinant of Economic Growth: How Well You Use it May Be More Important than How Much You Have*, "NBER Working Paper", no. 5847 (Revised 2005).

International Monetary Fund (2004), *IMF Article IV Review of Austria, Selected Issues: The Impact of Government Size and the Composition of Revenue and Expenditure on Growth*, Washington.

Janeba E. (2000), *Trade, Income Inequality, and Government Policies: Redistribution of Income or Education Subsidies?*, "NBER Working Paper", no. 7485.

Jaumotte F., Pain N. (2005) *An Overview of Public Policies To Support Innovation*, "OECD Economics Department Working Papers", no. 456.

Kaldor N. (1956), *Alternative Theories of Distribution*, "Review of Economic Studies", no. 23, pp. 83–100.

Keefer P., Knack S. (1995), *Institutions, And Economic Performance: Cross-Country Tests Using Alternative Institutional Measures*, "Economics and Politics", vol. 7(3), pp. 207–227.

King R.G., Rebelo S. (1990), *Public Policy and Economic Growth: Developing Neoclassical Implications*, "Journal of Political Economy", vol. 98(5), pp. 126–150.

Kneller R., Stevens P.A. (2006), *Frontier Technology and Absorptive Capacity: Evidence from OECD Manufacturing Industries*, "Oxford Bulletin of Economics and Statistics", vol. 68(1), pp. 1–21.

Koester R.B., Kormendi R.C. (1989), *Taxation, Aggregate Activity and Economic Growth: Cross-Country Evidence on Some Supply-Side Hypotheses*, "Economic Inquiry", vol. 27(3), pp. 367–386.

Kotsogiannis Ch., Schwager R. (2008), *Accountability and Fiscal Equalization*, "Journal of Public Economics", vol. 92(12), pp. 2336–2349.

Landau D. (1996), *Is One of The 'Peace Dividends' Negative? Military Expenditure and Economic Growth in The Wealthy OECD Countries*, "Quarterly Review of Economics and Finance", vol. 36(2), pp. 183–196.

Layard R., Nickell S., Jackman R. (1996), *Combating Unemployment: Is Flexibility enough?*, 'Centre For Economic Performance Discussion Paper", no. 293, London School of Economics and Political Science, London.

Leibfritz W., Thornton J., Bibbee A. (1997), *Taxation and Economic Performance*, OECD, Paris.

Leung D., Meh C., Terajima Y. (2008), *Firm size and productivity*, "Bank of Canada Working Paper", no. 2008–45.

Maddison-Project, http://www.ggdc.net/maddison/maddison-project/home.htm [accessed: 3.05.2021].

Mares I. (2007), *The Economic Consequences of the Welfare State*, "International Social Security Review", vol. 60(2–3), pp. 65–81.

Martinez-Vazquez J., Vulovic V., Liu Y. (2009), *Direct versus Indirect Taxation: Trends, Theory and Economic Significance*, "International Studies Program Working Paper", no. 09–11.

Melitz M.J., Ottaviano G.I.P. (2008), *Market Size, Trade, and Productivity*, "Review of Economic Studies", vol. 75(1), pp. 295–316.

Miller N.J., Tsoukis Ch. (2001), *On the Optimality of Public Capital for Long-Run Economic Growth: Evidence from Panel Data*, "Applied Economics", vol. 33(9), pp. 1117–1129.

Nelson R.R., Phelps E.S. (1966), *Investment in Humans, Technological Diffusion, and Economic Growth*, "American Economic Review", vol. 56(1/2), pp. 69–75.

North D.C. (1993), *The New Institutional Economics and Development*, "Economic History 9309002", University Library of Munich, Munich.

OECD (2010), *Tax Policy Reform and Economic Growth*, "OECD Tax Policy Studies", no. 20.

OECD (2015), *Tax Database. Tax Collection Costs*, https://data.oecd.org/inequality/poverty-rate.htm [accessed: 3.05.2021].

OECD (2017), *OECD Economic Survey of Italy*, Paris.

Pagano P., Schivardi F. (2003), *Firm Size Distribution and Growth*, "Scandinavian Journal of Economics", vol. 105(2), pp. 255–274.

Pestel N., Sommer E. (2017), *Shifting Taxes from Labor to Consumption: More Employment and More Inequality*, "Review of Income and Wealth", vol. 63(3), pp. 542–563.

PwC UK, World Bank Group, (2016), *Paying Taxes 2017*, London.

Ramirez A., Ranis G., Stewart F. (2000), *Economic Growth and Human Development*, "World Development", vol. 28(2), pp. 197–219.

Rodden J.A. (2005), *Hamilton's Paradox: The Promise and Peril of Fiscal Federalism (Cambridge Studies in Comparative Politics)*, Cambridge University Press, Cambridge.

Romer D. (2000), *Makroekonomia dla zaawansowanych*, Wydawnictwo Naukowe PWN, Warszawa.

Rzońca A. (2005), *Finanse publiczne a wzrost gospodarki w długim okresie. Analiza przy wykorzystaniu zmodyfikowanego modelu nabywania wiedzy przez praktykę*, "Materiały i Studia", no. 198, pp. 1–39.

Sala-i-Martin X. (1997), *I Just Ran Two Million Regressions*, "American Economic Association Papers and Proceedings", vol. 87(2), pp. 178–183.

Showalter M.H., Thurston N.K. (1997), *Taxes And Labour Supply of High Income Physicians*, "Journal of Public Economics", vol. 66(1), pp. 73–97.

Skrok E. (2013), *Podatki a wzrost gospodarczy: wnioski dla Polski wynikające z teorii i doświadczeń międzynarodowych*, Warszawa (mimeo).

Slemrod J., Yitzhaki S. (2002), *Tax Avoidance, Evasion, and Administration*, [in:] A.J. Auerbach, M. Feldstein (eds.), *Handbook of Public Economics*, vol. 3, Elsevier, Amsterdam–London–New York, pp. 1425–1470.

Smith A. (1954), *Badania nad naturą i przyczynami bogactwa narodów*, Państwowe Wydawnictwo Naukowe, Warszawa.

Smith D.B. (2006), *Living with Leviathan: Public Spending*, "Taxes and Economic Performance. IEA Hobart Paper", no. 158.

Smith D.B. (2016), *Spending, Tax and Economic Welfare*, [in:] P. Booth (ed.), *Taxation, government spending and economic growth*, Institute of Economic Affairs, London, pp. 87–99.

Spengel Ch., Elschner Ch., Endres D. (2012), *Effective Tax Levels Using the Devereux/Griffith Methodology*, Project for the EU Commission, TAXUD/2008/CC/099.

Stiglitz J.E. (2004), *Ekonomia sektora publicznego*, Wydawnictwo Naukowe PWN, Warszawa.

Tanzi V. (2008), *Regulating for the New Economic Order: "The Good, The Bad, and The Damaging*, Politeia, London.

Tanzi V., Schuknecht L. (2000), *Public Spending in the 20th Century. A Global Perspective*, Cambridge University Press, Cambridge.

Thomas A., Picos-Sánchez F. (2012), *Shifting from Social Security Contributions to Consumption Taxes: The Impact on Low-Income Earner Work Incentives*, "OECD Taxation Working Paper", no. 11.

Van Biesebroeck J. (2005), *Firm Size Matters: Growth and Productivity Growth in African Manufacturing*, "Economic Development and Cultural Change", vol. 53(3), pp. 545–583.

Widmalm F. (2001), *Tax Structure And Growth: Are Some Taxes Better Than Others?*, "Public Choice", vol. 107(3), pp. 199–219.

World Economic Forum (2016), *The Global Competitiveness Report 2016–2017*, Geneva.

Appendix. Amount of reduced VAT rates in the EU countries in the period 2002–2021

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Belgium	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12
Bulgaria	n/a	n/a	n/a	n/a	n/a	7	7	7	7	9	9	9	9	9	9	9	9	9	9	9
Czech Republic	5	5	5	5	5	5	9	9	10	10	14	15	15	10/15	10/15	10/15	10/15	10/15	10/15	10/15
Denmark	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Germany	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
Estonia	5	5	5	5	5	5	5	9	9	9	9	9	9	9	9	9	9	9	9	9
Ireland	12.5 (4.3)	13.5 (4.3)	13.5 (4.4)	13.5 (4.8)	13.5 (4.8)	13.5 (4.8)	13.5 (4.8)	13.5 (4.8)	13.5 (4.8)	9/13.5 (4.8)	9/13.5 (4.8)	9/13.5 (4.8)	9/13.5 (4.8)	9/13.5 (4.8)	9/13.5 (4.8)	9/13.5 (4.8)	9/13.5 (4.8)	9/13.5 (4.8)	9/13.5 (4.8)	9/13.5 (4.8)
Greece	8 (4)	8 (4)	8 (4)	9 (4.5)	9 (4.5)	9 (4.5)	9 (4.5)	9 (4.5)	5.5/11	6.5/13	6.5/13	6.5/13	6.5/13	6/13	6/13	6/13	6/13	6/13	6/13	6/13
Spain	7 (4)	7 (4)	7 (4)	7 (4)	7 (4)	7 (4)	7 (4)	7 (4)	8 (4)	8 (4)	8 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)
France	5.5 (2.1)	5.5 (2.1)	5.5 (2.1)	5.5 (2.1)	5.5 (2.1)	5.5 (2.1)	5.5 (2.1)	5.5 (2.1)	5.5 (2.1)	5.5 (2.1)	5.5/7 (2.1)	5.5/7 (2.1)	5.5/10 (2.1)	5.5/10 (2.1)	5.5/10 (2.1)	5.5/10 (2.1)	5.5/10 (2.1)	5.5/10 (2.1)	5.5/10 (2.1)	5.5/10 (2.1)
Croatia	0	0	0	0	10 (0)	10 (0)	10 (0)	10 (0)	10 (0)	10 (0)	10 (0)	5/10	5/13	5/13	5/13	5/13	5/13	5/13	5/13	5/13
Italy	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10 (4)	10/5 (4)	10/5 (4)	10/5 (4)	10/5 (4)	10/5 (4)	10/5 (4)
Cyprus	5	5	5	5/8	5/8	5/8	5/8	5/8	5/8	5/8	5/8	5/8	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9
Latvia	n/a	9	5	5	5	5	5	10	10	12	12	12	12	12	12	12	5/12	5/12	5/12	5/12
Lithuania	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9
Luxembourg	3/6	3/6	3/6	3/6	3/6	3/6	3/6	3/6	3/6	3/6	3/6	3/6	3/6	3/8	3/8	3/8	3/8	3/8	3/8	3/8
Hungary	12 (0)	12 (0)	5/15	5/15	5/15	5	5	5/18	5/18	5/18	5/18	5/18	5/18	5/18	5/18	5/18	5/18	5/18	5/18	5/18
Malta	5	5	5	5	5	5	5	5	5	5/7	5/7	5/7	5/7	5/7	5/7	5/7	5/7	5/7	5/7	5/7
Netherlands	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	9	9	9
Austria	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10/13	10/13	10/13	10/13	10/13	10/13

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Poland	7 (3)	7 (3)	7 (3)	7 (3)	7 (3)	7 (3)	7 (3)	7 (3)	7 (3)	5/8	5/8	5/8	5/8	5/8	5/8	5/8	5/8	5/8	5/8	5/8
Portugal	5/12	5/12	5/12	5/12	5/12	5/12	5/12	5/12	6/13	6/13	6/13	6/13	6/13	6/13	6/13	6/13	6/13	6/13	6/13	6/13
Romania	n/a	n/a	9	9	9	9	9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9	5/9
Slovenia	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	5/9.5	5/9.5
Slovakia	10	14	n/a	n/a	n/a	10	10	10	6/10	10	10	10	10	10	10	10	10	10	10	10
Finland	8/17	8/17	8/17	8/17	8/17	8/17	8/17	8/17	9/13	9/13	9/13	10/14	10/14	10/14	10/14	10/14	10/14	10/14	10/14	10/14
Sweden	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12	6/12

* n/a – non-applicable.

Source: European Commission, 2015b

Kierunki przebudowy systemu podatkowego w Polsce. Propozycja prowzrostowa

Streszczenie: Celem artykułu jest przedstawienie kierunków przebudowy systemu podatkowego w Polsce, która wzmacniałaby wzrost gospodarczy. Opracowanie zawiera diagnozę głównych silnych stron i słabości tego systemu. Na podstawie tej diagnozy oraz przeglądu literatury autorzy proponują rekomendacje, których uwzględnienie przy przebudowie podatków sprzyjałoby wzrostowi gospodarczemu. Obejmują one: przesunięcie ciężaru opodatkowania z dochodów, zwłaszcza z nisko płatnej pracy, na konsumpcję; objęcie części składek na ubezpieczenie społeczne kwotą wolną; umożliwienie samorządom podwyższenia kwoty wolnej od podatku PIT powyżej centralnie ustalonej kwoty bazowej; ujednolicenie podstawy wymiaru podatku PIT, składek NFZ i ZUS; usunięcie różnic w oskładkowaniu różnych typów umów, na podstawie których jest wykonywana praca; rozszerzenie możliwości jednorazowej amortyzacji na wszystkie inwestycje maszynowe; eliminację podatków sektorowych.

Słowa kluczowe: podatki, system podatkowy, propozycja prowzrostowej zmiany

JEL: H20, H21, H24, H25, H29, O10, O11, O43

Received: 2021-01-17; verified: 2021-05-04. Accepted: 2021-07-06