

# THE SIZE OF THE SUBSTITUTION BIAS OF INFLATION MEASUREMENT IN RELATION TO THE LEVEL OF INNOVATIVENESS OF THE EUROPEAN UNION'S ECONOMIES

**Elżbieta Roszko-Wójtowicz**

Univeristy of Lodz, Lodz, Poland  
e-mail: elzbieta.roszko@uni.lodz.pl  
ORCID: 0000-0001-9337-7218

**Jacek Bialek**

Univeristy of Lodz, Lodz, Poland  
e-mail: jacek.bialek@uni.lodz.pl  
ORCID: 0000-0002-0952-5327

© 2018 Elżbieta Roszko-Wójtowicz, Jacek Bialek

*This is an open access article distributed under the Creative Commons Attribution-NonCommercial-NoDerivs license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>)*

DOI: 10.15611/ead.2018.4.05

JEL Classification: C43, C38, E31, 030

---

**Abstract:** The consumer price index (CPI) is a common measure of inflation. Similarly to the harmonised index of consumer prices (HICP), it is determined using the Laspeyres index, thus data on the consumption of the basket of goods do not have to be current. The Laspeyres index, using weights only from the base period, may not reflect changes in consumer preferences that occurred in the studied year. This is the reason for the formation of the so-called substitution bias in the measurement of inflation. The aim of the article is to assess the impact of the level of innovativeness of a given country's economy on the occurrence of the substitution effect. The empirical part of the article is based on basic innovation indices, i.e. the SII, IOI, and GII. The assessment of the relationship between the level of innovativeness and the scale of the substitution effect was carried out based on the methods of multidimensional statistical analysis (including cluster analysis, the PROFIT method).

**Keywords:** inflation measurement, substitution bias, innovativeness, innovation indices, cluster analysis, PROFIT method.

---

## 1. Introduction

The consumer price index (CPI) is a common measure of inflation [Hałka, Leszczyńska 2011]. Similarly to the harmonised index of consumer prices (HICP), it is determined using the Laspeyres index, thus data on the consumption of the basket of goods do

not have to be current [Camba-Mendez et al. 2002]. By assumption, the CPI and HICP are to approximate changes in the costs of household consumption in the broad sense, assuming their permanent utility, i.e. that they should approximate the so-called *Cost of Living Index* (COLI). The ideal approximation of COLI, however, are superlative indices, including the Fisher index, which use weights from both the base and the test period in their formula [Von der Lippe 2007]. Admittedly, it is possible to approximate the Fisher index by means of indices using only consumption data from the base period [Lent, Dorfman 2009; Białek 2017a; Białek 2017b], nonetheless most countries in the world continue to use the Laspeyres index to measure the CPI [White 1999]. The Laspeyres index, using weights only from the base period, may not reflect changes in consumer preferences that occurred in the studied year. This is the reason for the formation of the so-called substitution bias in the measurement of inflation, which is usually measured by the difference between the values of the Laspeyres and Fisher indices [Białek 2016]. In this article the authors focus only on the CPI substitution bias as the most recognised and the most often tested type of CPI bias, although in general there are distinguished as many as five main reasons for this type of bias [White 1999; Hałka, Leszczyńska 2011]. In the article the authors assume that the level of innovativeness of the economy, expressed by means of synthetic measures (e.g. the *Summary Innovation Index* (SII), the *Global Innovation Index* (GII), the *Innovation Output Indicator* (IOI)), remains in close relation with consumer behaviour, which is changing more and more dynamically. This is influenced by various factors including internationalisation processes, universal access to advanced information and communication technologies, as well as changes in the lifestyle of society. This results in a shortening of the product life cycle, intensification of competition and greater susceptibility to trends created in the mass media. Meeting consumers' needs and expectations [Baturó 2004, p. 743], not only the basic ones, is the cornerstone of their behaviour in the market. Over the years, consumers' expectations regarding business offers have become more sophisticated and complex. The measure of a company's success is therefore consumers' acceptance for the goods and services offered expressed by the high level of satisfying their needs. The more consumers are satisfied, the more often they reach for a given product or service [Olejniczak 2007, p. 48]. As consumers' interest in innovative products increases, the share from the sale of innovative products in the income of companies increases [Stępnia-Kucharska 2012; Hall, Mairesse 1995, pp. 263-293; Adams, Jaffe 1996, pp. 700-721]. Consequently, this leads to an improvement in the innovativeness of the entire economy. Innovation is therefore a determinant of the development of both the enterprise and the entire economy [Grębowiec 2017, p. 290]. Therefore innovations should be constantly created and implemented [Nowicka 2015, pp. 129-141].

A more innovative economy offers its consumers more innovative products and services. Thus, consumers' tendency to acquire these products and services as well as their awareness of the utility of new market products are greater than in

countries where the level of innovativeness is significantly lower. From the point of view of inflation measurement, it can be assumed that the applied measurement methodology does not keep up with the changes in consumer behaviour related to following innovations, which leads to a deepening substitution bias.

The aim of the article is to assess the impact of the level of innovativeness of a given country's economy on the occurrence of the substitution effect in measuring inflation. It should be assumed that the level of innovativeness of the economy is manifested in consumer preferences. This is especially reflected in the expectations of consumers regarding the functionality and utility of new goods and services, as well as the scope of use of advanced technologies. Due to this fact, the article presents the verification of the hypothesis that a higher level of innovativeness of the economy results in more dynamically changing consumer preferences, thus leading to a greater substitution bias in measuring inflation. The empirical part of the article is based on the basic innovation indices, i.e. the *SII*, *IOI*, and *GII*.

The empirical study consists of three main steps, briefly described as follows:

**Step 1.** At the beginning of the research, the significance of correlations between the effect of the substitution of goods in the CPI measurement and the analysed innovation indices together with their sub-indices is verified. A simple correlation analysis using Pearson's correlation coefficients is conducted.

**Step 2.** We continue the analysis by using k-means clustering in order to confirm the relations verified in Step 1 by observing clusters of the most and least innovative EU countries. The natural question that we want to answer is whether these groups of countries would be similar if we took into consideration also the level of the CPI substitution bias. Thus, in this step, the EU countries are categorised taking into account both innovativeness indices (also sub-indices) and the measures of the above-mentioned CPI bias.

**Step 3.** This step plays the role of the final confirmation of the previously obtained conclusions concerning the relationship between the innovativeness level and the effect of the CPI substitution. At this stage we apply the PROFIT method to observe the common impact of all the considered variables (connected with both innovativeness and the CPI substitution bias) on the grouping of the EU countries. In particular, according to the PROFIT concept, we draw regression lines corresponding to these variables on a two-dimensional perception map. Substantial, but not strong, connections between these variables are confirmed, i.e. some of the regression lines are quite convergent on the perception map which presents similarities and dissimilarities between the EU countries.

All data used in the article come from the Eurostat database. The analyses were carried out for data from 2016. The paper consists of a theoretical part which discusses the most common measures of innovativeness and the source of the substitution bias in the CPI measurement. The next part discusses the sources of statistical data used in the paper and the methods applied. The last section presents the results of the conducted analyses, along with the final conclusions.

## 2. The theoretical background

### 2.1. Measuring innovativeness

Innovativeness is a multidimensional concept [Mamica 2007, p. 122], and thus difficult to define unequivocally, with blurred borders [Anthony et al. 2014, p. 264]. The narrow and broad approach to innovativeness makes measuring innovativeness of the economy a complicated task [Łukasik 2017]. First of all, this is due to the fact that innovativeness is a very complex phenomenon influenced by many diverse economic, social and technological factors [Staśkiewicz 2013, pp. 576-577]. The very identification of the factors that have the strongest impact on innovativeness is a great substantive challenge. In the analysis of innovativeness, researchers rarely limit themselves to one measure and are willing to use many indicators, thus obtaining a more complete picture of innovativeness. There is still a lack of a single universal measure that could be used to quantify innovativeness and make comparisons between countries. Synthetic measures, which present a complex and multidimensional issue of innovativeness, are of great value in this respect, As firstly they allow us to organise the examined objects, usually countries, due to their level of innovative advancement.

In this section, selected synthetic measures of innovativeness will be presented, i.e. the *Global Innovation Index*, the *Summary Innovation Index* and the *Innovation Output Indicator*. Differences in the approach to measuring innovativeness in each of these indices have a significant impact on the final ranking of countries.

**Table 1.** List of individual innovation indices

	Index name		
	Global Innovation Index	Innovation Output Indicator	Summary Innovation Index
Source	JOHNSON Cornell University, The Business School of the World, The World Intellectual Property Organization	European Commission	European Commission
No. of variables	81	5	25
No. of countries	127	38	34
Year of the first publication, frequency	Since 2007, annual publication cycle	Since 2010, annual publication cycle	Since 2001, annual publication cycle

Source: [European Commission 2013; 2015; Vertesy and Deiss2016; Dutta et al. 2017].

It is worth emphasising that the methodology of innovation indices continues to evolve and is subject to numerous modifications over time. Due to the extent of

the issue, changes within the synthetic measures of innovativeness will be discussed only in relation to the *Global Innovation Index* [Dutta et al. 2017]. In the case of other indices, the authors will limit themselves to presenting rankings in a dynamic approach.

Methodological changes occur, among others, within the structure of the synthetic index as well as the variables used for its construction. Therefore the number of variables and sub-indices used to build the final indices is not constant. The emergence of new diagnostic features is an expression of following the latest trends and reflects the changes that are taking place in the field of research on the measurement of innovativeness.

The Global Innovation Index (GII) is an evolving project that builds on the experience and results of its previous editions. Currently, five pillars from the Input sphere and two pillars from the Output sphere are distinguished in the construction of the index. The five Input pillars that enable innovative activities are: (1) Institutions, (2) Human capital and research, (3) Infrastructure, (4) Market sophistication, (5) Business sophistication, while the two Output pillars capture the actual evidence of innovation outputs: (6) Knowledge and technology outputs and (7) Creative outputs. Each of the pillars consists of lower levels of aggregation, referred to as sub-pillars, and each sub-pillar is composed of individual indicators. How is the final ranking created? Sub-pillar scores are calculated as the weighted average of individual indicators, while pillar scores are calculated as the weighted average of sub-pillar scores.

The *Global Innovation Index (GII)* is one of the world's most famous studies of innovativeness. The report is published on an annual basis and experts from the following institutions work on its preparation: Johnson Cornell University which is one of the largest management and business schools in the world, the Business School of the World (INSEAD), and the World Intellectual Property Organisation (WIPO).

The tenth edition of the report presents 127 economies, the study covered 92.5% of the world's population and 97.6% of global GDP (US\$). 81 individual indicators characterising innovativeness were included in the ranking [Dutta et al., p. 12]. Thus, in relation to the previous edition, the number of variables used in the construction of the synthetic measure has not changed. However the set of variables has been extended in relation to previous *GII* rankings by two additional variables. According to the latest data presented in the *Global Innovation Index 2017*, Innovation Feeding the World, European countries dominate at the top of the ranking [Dutta et al. 2017].

Among the ten best economies in terms of innovativeness, as many as eight represent the European continent. Switzerland ranks invariably first, followed by Sweden and the Netherlands. The United Kingdom, which ranks fifth, is preceded in this edition by the United States. Next are Denmark, Finland, Germany and Ireland. This time, Luxembourg is outside the top ten. In the *GII 2017* ranking, Poland is 38th in the overall classification, slightly ahead of Hungary, Lithuania, Croatia

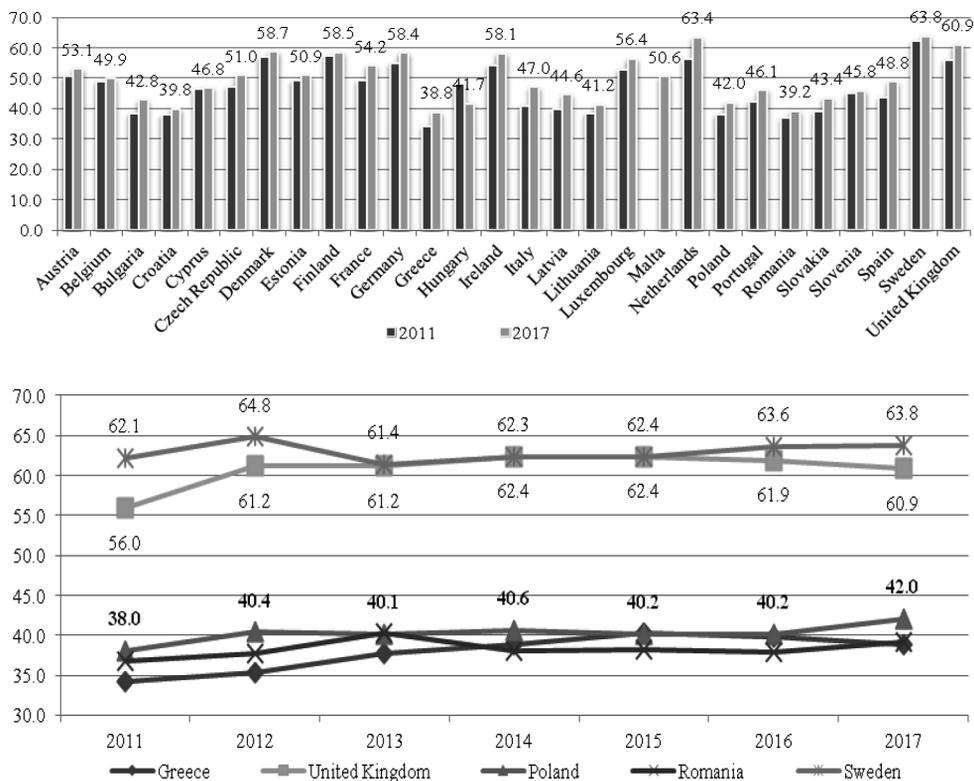


Fig. 1. Results of the *Global Innovation Index* – comparison of 2017 and 2011

Source: own elaboration based on the Knoema database, <https://knoema.com/>.

and Romania. Among the European countries in the *GII* 2017 ranking, Bulgaria, Slovakia and Latvia are slightly ahead of Poland. In the years 2011-2017, in the *Global Innovation Index* the highest positions are occupied by Sweden and the United Kingdom. The worst performers in the analysed period are Greece and Romania.

A similar situation can be seen in the case of the *Summary Innovation Index* ranking, where in the years 2010-2017, Sweden ranked the highest, followed closely by Denmark. The lowest positions in the *SII* ranking in those years were filled by Bulgaria and Romania.

In 2013 the European Commission introduced a new indicator for measuring innovativeness, characterised by the fact that it focuses on measuring the effects of innovative activities [European Commission 2013]. It was created in response to the aim of increasing R&D expenditure formulated as part of the Europe 2020 Strategy [European Commission 2010, p. 8]. The new indicator is to assess the progress of the EU Member States in implementing the set benchmark. It was assumed that 3% of the EU’s GDP should be invested in R&D by the end of 2020.

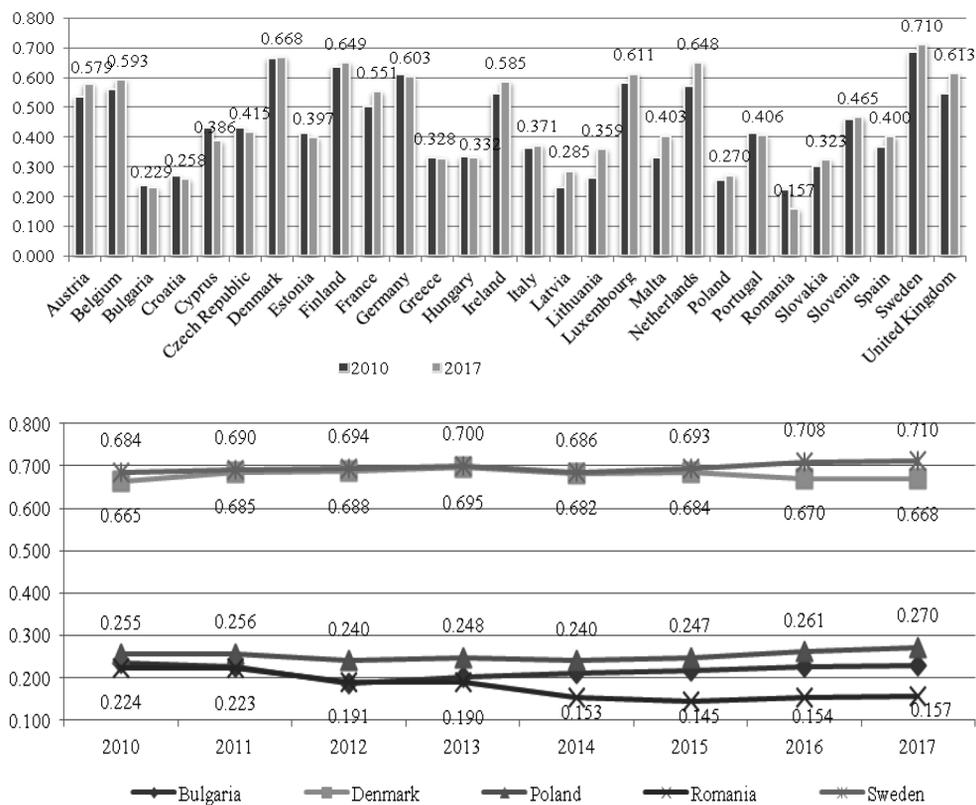
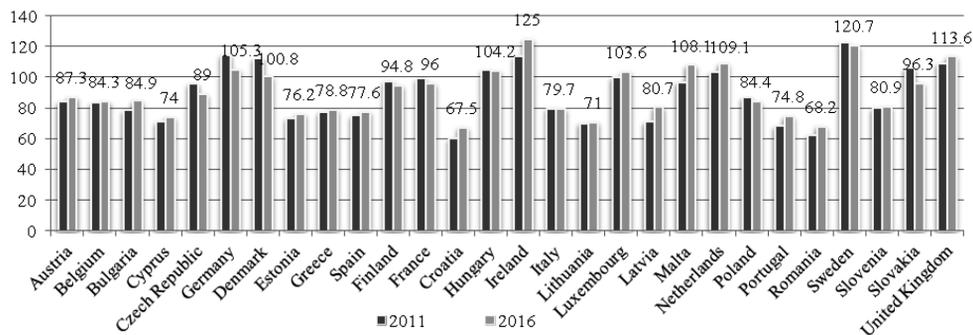
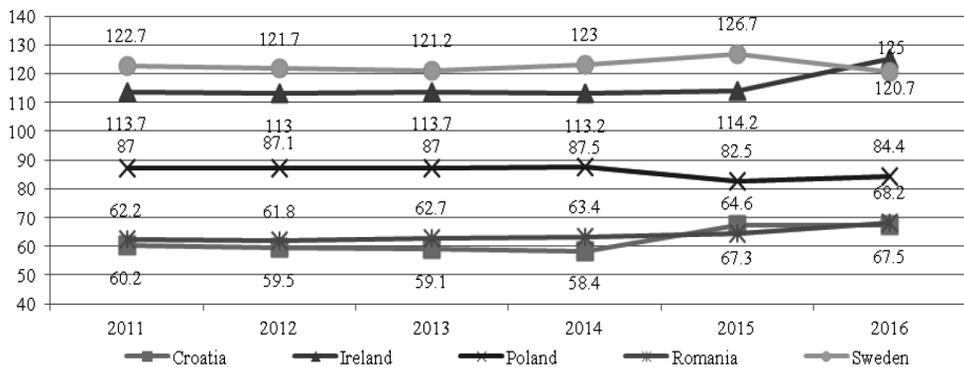


Fig. 2. Results of the Summary Innovation Index – comparisons of 2017 and 2010

Source: own elaboration based on the European Innovation Scoreboard, [https://ec.europa.eu/growth/industry/innovation/facts-figures/scoreboards\\_en](https://ec.europa.eu/growth/industry/innovation/facts-figures/scoreboards_en).





**Fig. 3.** Results of the *Innovation Output Indicator* – comparison of 2017 and 2010

Source: own elaboration based on [Vertesy 2017, p. 33; [http://publications.jrc.ec.europa.eu/repository/bitstream/JRC108942/jrc108942\\_ioi\\_2017\\_report\\_final.pdf](http://publications.jrc.ec.europa.eu/repository/bitstream/JRC108942/jrc108942_ioi_2017_report_final.pdf)].

Although methodologically the *IOI* differs significantly from the above-mentioned indices, the results of the ranking, especially in relation to the leader of the ranking and the country closing the ranking, are not surprising. In the years 2010-2016, Sweden and Ireland were the best in the *IOI* ranking, while Croatia and Romania closed the ranking.

## 2.2. The Consumer Price Index biases. The effect of substitution of goods

The *Consumer Price Index* (CPI) is commonly used as a basic measure of inflation, and in practice the Laspeyres price index (the weighted arithmetic mean of price relatives) is used to measure the CPI [White 1999; Clements, Izan 1987]. Most researchers dealing with inflation measurement issues mention five potential sources of the Consumer Price Index biases [White 1999; Hałka, Leszczyńska 2011].

*Commodity substitution bias.* This bias results from changes in relative prices of particular goods included in the CPI basket. The substitution effect results from the fact that consumers react to price changes by changing those consumer goods or services which are relatively more expensive for relatively cheaper ones.

*Outlet substitution bias.* This bias results from the migration of consumers towards more attractive, often just emerging, markets for purchases. Such a new market may be, for example, an online wholesale shop or a mail order outlet.

*New goods bias.* New goods which consumers have started to use in the period covered by the inflation survey are the source of this type of bias. Most often these are products that have been innovative, resulting from the introduction of a new technology for their manufacture and that have just come into common use, and thus have not been included in the CPI basket.

*Quality adjustment bias.* It is this type of CPI estimation bias which results from the changing (along with, e.g. growing customer/consumer expectations) of

the quality of goods offered by the market. By definition, the CPI should measure changes in prices of goods and services assuming that their features have not changed in relation to the base period. In reality, however, products from the basket of goods change – they are improved, modified and/or often simply withdrawn [Nordhaus 1998].

*Formula bias (elementary index bias)*. This bias may arise as a result of applying a given index formula at the lowest level of data aggregation. In cases when the arithmetic mean of all price indices of a given good is used for calculating the price index of a given product, the price index will be overestimated [Hałka, Leszczyńska 2011]. If, on the other hand, the average price of a good for a given period is calculated from all outlets and then it is related to the average price of that good in the previous period, the price index should not show this type of bias [Ducharme 2000]. This bias is all the greater, the greater the volatility of the prices analysed [Silver, Heravi 2007].

The CPI approximates changes in the costs of household consumption assuming constant utility (COLI, *Cost of Living Index*). The Lapeyres formula, which is used in CPI calculations and which uses budget shares from the base period as weights, does not take into account changes in the structure of consumption which occur as a result of price changes in a given time interval. This leads to the conclusion that the Laspeyres index can be biased due to *commodity substitution*. Many economists consider *superlative indices* (such as the Fisher index or the Törnqvist index) to be the best approximation of COLI [Von der Lippe 2007]. The difference between the Laspeyres index and the given superlative index should approximate the value of the commodity substitution bias. In this paper, we measure the effect of substitution as the difference between the Laspeyres price index and the Fisher index which can be calculated as a geometric mean of the Laspeyres and Paasche indices<sup>1</sup>.

### 3. Data and methodology

#### 3.1. Data set

Data used in the article come from several sources. Eurostat proved to be the most useful in terms of data on measuring inflation in the individual EU Member States. Annual data on average price indices presented at various levels of classification (COICOP 3 and COICOP 4) are derived from the Economy and Finance field. It is worth adding that the COICOP classification consists of the following levels: two-digit – 12 divisions, three-digit – 44 groups, four-digit – 110 classes and five-digit [Eurostat 2013]. [Aggregation at the COICOP 3 level allows the division into 44 groups, while at the lower level of aggregation – COICOP 4 – data are presented broken down into 110 groups]. Data on the level of innovativeness in the individual

---

<sup>1</sup> For more details, see for instance Von der Lippe [2007].

countries were taken from thematic reports (*Innovation Output Indicator*) or directly from on-line databases (*Summary Innovation Index, Global Innovation Index*).

Due to the availability of data on the level of inflation and weights assigned to individual groups, the analyses carried out in the empirical part of the article refer to 2016. Nevertheless, the section describing the measurement of innovativeness also includes information on the latest available data on the individual innovation indices. Thus, the information on the level of innovativeness in the EU Member States in 2017 was also taken into account.

**Table 2.** Main characteristics of selected variables

Symbol	Description
1	2
SUB3	Annual data on average indices (annual average indices) form the basis for calculating the substitution bias at the COICOP3 level. The basis of comparisons for the average value of the index at the COPICOP3 level is the value of the index from the previous year. For the purpose of the analysis, it was necessary to convert single-base indices (2015 = 100) into chain indices. The analysis was carried out for 2016 on the basis of data from the Eurostat database. The substitution bias was measured by the difference between the Laspeyres index and the superlative Fisher index. Time coverage: 2016.
MSUB3	Absolute value of SUB3. Time coverage: 2016.
SUB4	Annual data on average indices (annual average indices) form the basis for calculating the substitution bias at the COICOP4 level. The basis of comparisons for the average value of the index at the COPICOP4 level is the value of the index from the previous year. For the purpose of the analysis, it was necessary to convert single-base indices (2015 = 100) into chain indices. The analysis was carried out for 2016 on the basis of data from the Eurostat database. The substitution bias was measured by the difference between the Laspeyres index and the superlative Fisher index. Time coverage: 2016.
MSUB4	Absolute value of SUB4. Time coverage: 2016.
SII	Summer Innovation Index – data were derived directly from the thematic reports and the Eurostat database. For the purpose of clustering and the PROFIT method, all data used in the analysis were standardised. Time coverage: 2016.
GII	Global Innovation Index – data were derived directly from the thematic reports and the Knoema database. For the purpose of clustering and the PROFIT method, all data used in the analysis were standardised. Time coverage: 2016.
IOI	Innovation Output Indicator – data were derived directly from the thematic reports and the Eurostat database. For the purpose of clustering and the PROFIT method, all data used in the analysis were standardised. Time coverage: 2016.
RS	Research systems – data were derived directly from the thematic reports and the Knoema database. For the purpose of clustering and the PROFIT method, all data used in the analysis were standardised. Time coverage: 2016.
SIMP	Sales impacts – data were derived directly from the thematic reports and the Knoema database. For the purpose of clustering and the PROFIT method, all data used in the analysis were standardised. Time coverage: 2016.

1	2
INF	Infrastructure – data were derived directly from the thematic reports and the Knoema database. For the purpose of clustering and the PROFIT method, all data used in the analysis were standardised. Time coverage: 2016.
INST	Institutions – data were derived directly from the thematic reports and the Knoema database. For the purpose of clustering and the PROFIT method, all data used in the analysis were standardised. Time coverage: 2016.
MS	Market sophistication – data were derived directly from the thematic reports and the Knoema database. For the purpose of clustering and the PROFIT method, all data used in the analysis were standardised. Time coverage: 2016.
OUTPUT	Innovation Output (GII) – data were derived directly from the thematic reports and the Knoema database. For the purpose of clustering and the PROFIT method, all data used in the analysis were standardised. Time coverage: 2016.
KT	Knowledge and technology outputs – data were derived directly from the thematic reports and the Knoema database. For the purpose of clustering and the PROFIT method, all data used in the analysis were standardised. Time coverage: 2016.

Source: own elaboration.

The table presents the individual variables used in the course of the research process described in the empirical part of the paper.

### 3.2. Clustering using *k*-means method

Cluster analysis means data segmentation or clustering, and it is one of the best known data mining methods. It is an advanced data analysis tool for clustering variables or objects into unknown groups. This analysis is a method of classification which consists in dividing a (usually multivariate) data set into groups (clusters) in such a way that the objects in the same group are similar to one another, and at the same time different from those in other groups [Gatnar, Walesiak 2004, pp. 310-317]. When performing cluster analysis, we divide a data set into groups using a variety of criteria that allow us to determine the similarity of objects in a given group [Grabiński et al. 1989]. This categorisation aims at distinguishing objects between groups to the highest degree, while maintaining their consistency – similarity – within the same group [Marek 1989, pp. 19-22].

The literature on the subject presents many clustering methods [Sneath, Sokal, 1973]. For the purpose of the conducted analyses, the authors decided to use the *k*-means method in the article. This method is the most classic algorithm of cluster analysis, and at the same time it is widely used in the analysis of data from various fields. The subject of the analysis is a set of  $n$  objects (e.g. countries) described by  $m$  features (the so-called diagnostic variables). The aim of the analysis is to find the optimal division of the output set of objects into  $k$  subsets. *K*-means clustering is an iterative procedure, where in each iteration some objects are transferred to other clusters in order to obtain the minimum intra-group and maximum intergroup heterogeneity.

The  $k$ -means algorithm can be described as follows:

1. The starting point is the division of a given set of objects into  $k$  subsets.
2. Then each object is assigned to the “nearest” pre-selected representative of  $k$  groups. The  $k$  number of representatives is chosen randomly in such a way that they are as far apart from one another as possible. The selected objects constitute the seeds of the groups.
3. On the basis of the arithmetic mean of the coordinates of objects belonging to a given group, the centres of each group in the space of diagnostic variables are determined.
4. We return to Step 2 if at least one object has been moved to another group. The determination of new cluster centres and the verification of the correctness of objects belonging to the groups will be carried out as long as there is a transfer of objects between the clusters [Jajuga 1993, pp. 55-60].

The algorithm of the  $k$ -means method can be treated as a kind of “inverse” analysis of variance. With its help we find a division of the studied population into groups which maximises intergroup variance, and consequently  $F$  statistics.

### 3.3. PROFIT method

PROFIT (PROpertyFITting) is a kind of external vector analysis of preference mapping. The standard reason for using this method is testing hypotheses about the attributes that influence people’s judgement of the similarities among a set of objects. Nevertheless there are no technical objections to use PROFIT for other cases, not only connected to human preferences. The PROFIT method is a two-step procedure which is a combination of multidimensional scaling and multiple regression analysis. To understand the idea of PROFIT, we should start from the multidimensional scaling (MDS) description. Multidimensional scaling (MDS) is a means of visualising the level of similarity of individual cases of a dataset. It refers to a set of related ordination techniques used in information visualisation, in particular to display the information contained in a distance matrix. It is a form of non-linear dimensionality reduction [Kruskal, Wish 1978; Young, Hamer 1987]. An MDS algorithm aims to place each object in  $N$ -dimensional space such that the between-object distances are preserved as well as possible. Each object is then assigned coordinates in each of the  $N$  dimensions. The number of dimensions of an MDS plot  $N$  can exceed 2 and is specified a priori. Choosing  $N = 2$  optimises the object locations for a two-dimensional scatterplot. MDS can have a metric or non-metric form and, as a rule, the non-metric MDS is used before PROFIT analysis [Kruskal 1964; Takane et. al. 1977]. In such case, the method finds both a non-parametric monotonic relationship between the dissimilarities in the object-object matrix and the Euclidean distances between objects, and the location of each object in the low-dimensional space. The relationship is typically found using isotonic regression: let  $x$  denote the vector of proximities,  $f(x)$  – a monotonic transformation

of  $x$  and  $d$  is the distance measure; then coordinates have to be found that minimise the so-called *stress* function defined as follows [Kruskal 1964]:

$$stress = \sqrt{\frac{\sum_{i,j} (d_{ij} - f(x_{ij}))^2}{\sum_{i,j} d_{ij}^2}},$$

where lower indices denote the  $i$ -th and the  $j$ -th point on the MDS map.

PROFIT analysis evaluates the correspondence between one or more object attributes and the location of objects in a multidimensional space. As mentioned above, PROFIT, which is a combination of multidimensional scaling and multiple regression analysis, consists of two phases. After the first step (MDS), we obtain a configuration of  $n$  points  $x = (x_1, x_2, \dots, x_n)^T$  in  $r$ -dimensional space ( $r$  is usually 2 or 3). In the second phase, PROFIT takes as input both a configuration of points  $x$  and a set of attribute preferences data  $p_k = (p_{k1}, p_{k2}, \dots, p_{kn})^T$ , where  $k = 1, 2, \dots, m$  is a number of attributes [Zaborski, Peřka 2013]. Then it performs a multiple regression using the coordinates of  $x$  as independent variables and the attributes as the dependent variables. The procedure provides a separate regression for each one attribute with “classical” regression coefficients:

$$a_k = (x^T x)^{-1} x^T p_k.$$

## 4. Results

The initial analysis indicates that there is a not very strong but statistically significant correlation between the substitution effect (measured using variables SUB3, MSUB3, SUB4 and MSUB4) and the analysed innovation indices. This is confirmed by both the determined correlation coefficients (Table 3) and the results of the clustering of the EU countries based on the mentioned variables. In the latter case, k-means clustering was used ( $k=2$  and  $k=3$ ) on data previously normalised by standardisation (cf. Figure 4). It can be seen that those clusters of countries that simultaneously show a higher rate of CPI substitution bias are characterised by greater values of innovation indices (cf. Table 4). After analysis of Table 3 and Figure 4, the relationship between the substitution effect and the *IOI* level is questionable.

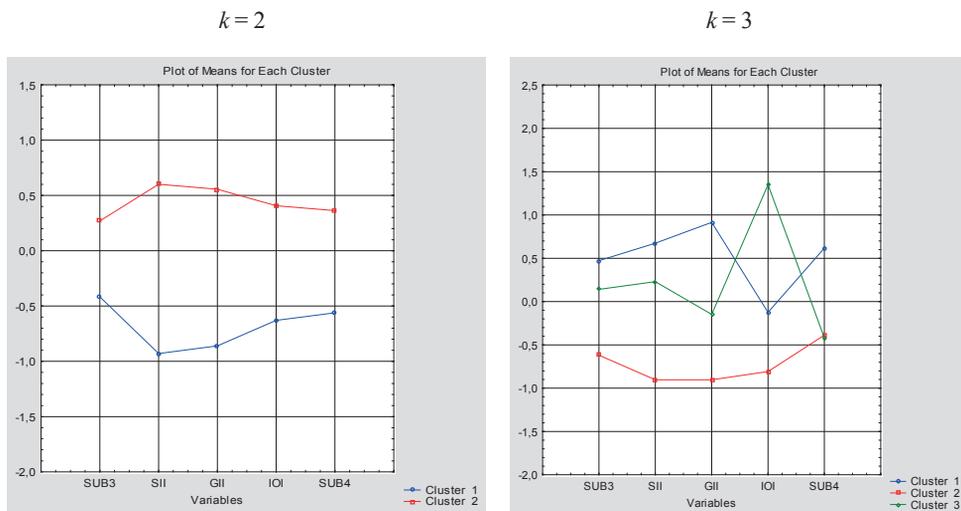
For further analysis, it was decided to use only those sub-indices related to the measurement of innovativeness which showed a statistically significant correlation with one of variables SUB3, MSUB3, SUB4 or MSUB4 (i.e. the variables measuring the substitution bias). The results for only significant correlations are presented in Table 5.

**Table 3.** Pearson correlation coefficients between measures of CPI substitution bias and innovation indices (EU countries, 2016)

Variable	Correlations; $N = 28$ (Casewise deletion of missing data)		
	IOI	GII	SII
SUB3	0.198	0.312*	0.145
MSUB3	0.021	-0.197	-0.231
SUB4	-0.004	0.281*	0.337*
MSUB4	0.048	-0.078	0.016

(\*) Marked correlations are significant at  $p < 0.05$ ,

Source: own elaboration based on Eurostat data and European Commission (2017). Calculations conducted in Statistica 11.



**Fig. 4.** Plot of means after  $k$ -means clustering depending on  $k$  parameter ( $k = 2$  or  $k = 3$ )

Source: own elaboration based on Eurostat data and European Commission (2017). Calculations conducted in Statistica 11.

**Table 4.** The EU countries as members of clusters after  $k$ -means clustering for  $k = 3$

Members of Cluster 1	Austria; Belgium; Cyprus; Denmark; Finland; France; Germany; Ireland; Malta
Members of Cluster 2	Bulgaria; Croatia; Estonia; Greece; Hungary; Latvia; Lithuania; Portugal; Romania; Slovakia
Members of Cluster 3	Czech Republic; Italy; Luxembourg; Netherlands; Poland Slovenia; the United Kingdom

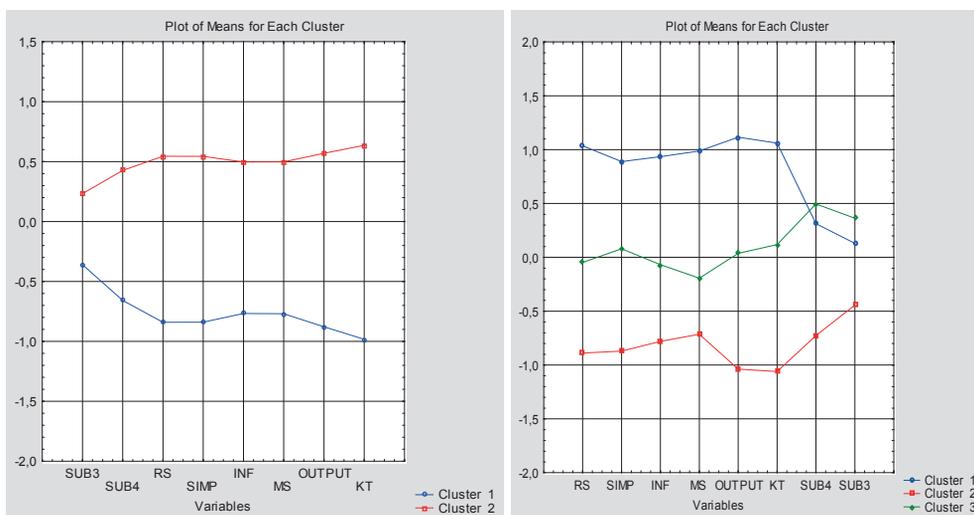
Source: own elaboration based on Eurostat data and European Commission (2017). Calculations conducted in Statistica 11.

**Table 5.** Pearson correlation coefficients between measures of CPI substitution bias and innovation sub-indices (EU countries, 2016)

Variable	Correlations						
	N = 28 (Casewise deletion of missing data)						
	RS	SIMP	INF	INST	MS	OUTPUT	KT
SUB3	0.138	0.312	-0.036	0.359	0.275	0.225	0.314
SUB4	0.384*	0.444*	0.112	0.388*	0.383*	0.398*	0.518*
MSUB3	-0.143	-0.221	-0.454*	-0.136	-0.007	-0.244	-0.166
MSUB4	0.100	0.057	-0.352	0.089	0.087	-0.038	0.101

(\*) Marked correlations are significant at  $p < 0.05$ .

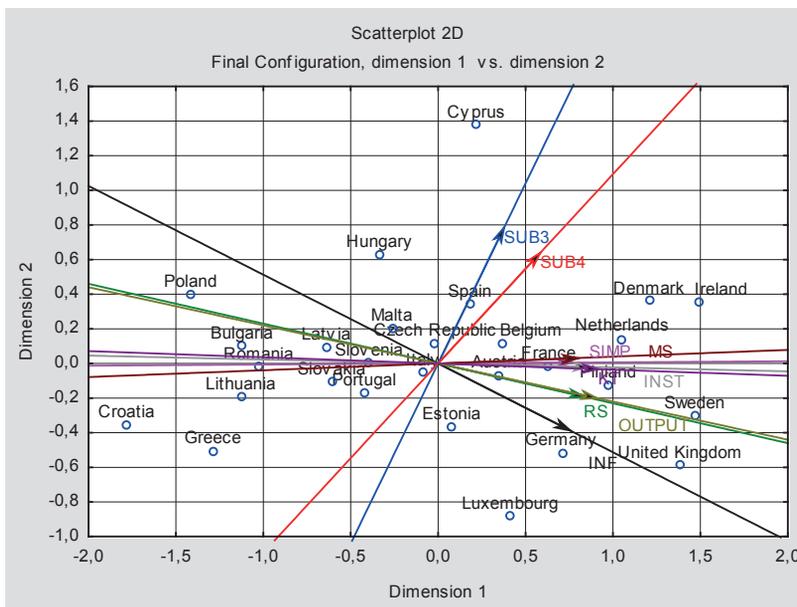
Source: own elaboration based on Eurostat data and European Commission (2017). Calculations conducted in Statistica 11.



**Fig. 5.** Plot of means after  $k$ -means clustering depending on  $k$  parameter ( $k = 2$  or  $k = 3$ )

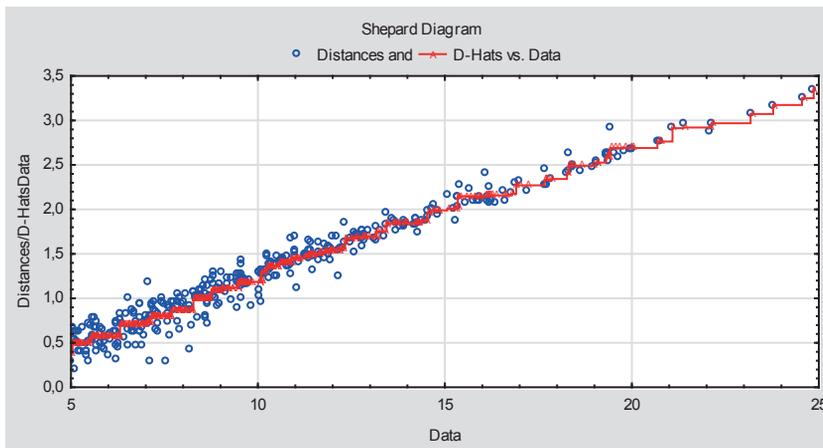
Source: own elaboration based on Eurostat data and European Commission (2017). Calculations conducted in Statistica 11.

Similarly as before, the EU countries were categorised taking into account both sub-indices of innovation (cf. Table 5) and the measures of CPI substitution bias. Figure 5 confirms the previous conclusion that along with the increase in innovativeness, the scale of the substitution effect also grows. However, in order to additionally confirm this conclusion, a complementary analysis of the EU countries was carried out using the PROFIT method and taking into account all the variables presented in Table 5. The results are presented in Figure 6. Both the STRESS function value (0.0809) and the Shepard diagram (Figure 7) confirm the sufficient accuracy



**Fig. 6.** Perception map with regression lines (PROFIT results) for the EU countries

Source: own elaboration based on Eurostat data and European Commission (2017). Calculations conducted in Statistica 11.



**Fig. 7.** Shepard’s diagram obtained in the PROFIT method

Source: own elaboration based on Eurostat data and European Commission (2017). Calculations conducted in Statistica 11.

obtained in the PROFIT method. Its application leads to the conclusion that although the impact vectors of variables related to innovativeness and variables related to

the substitution effect determined by regression lines are not perfectly convergent, they cannot be considered completely divergent (as they lie in one half-plane). In particular the interaction between variables SUB3 and SUB4 and SIMP, MS as well as the INST sub-indices seems to be noticeable on the perception map created as a result of the PROFIT method.

## 5. Conclusions

It would seem that there is a statistically significant positive correlation between the substitution bias when measuring the inflation of a given country and, broadly speaking, the level of innovativeness of its economy. The positive sign of this correlation is intuitive – the higher the level of innovativeness of the economy, the more dynamically the preferences of consumers in this country change (for example, due to emerging new goods, new technologies, new markets, etc.), the more likely that the Laspeyres formula, used to measure inflation (HICP) but using weights from the base period and hence out-of-date ones, ceases to effectively evaluate the level of inflation (the substitute bias of the measurement increases). In the study, only the indications of the IOI index do not seem to correlate with the substitution bias, while the remaining innovation indices seem to be in a weak but statistically significant relationship with variables SUB3 and SUB4 (measuring the substitution effect at different levels of data aggregation). By grouping the EU countries due to the level of substitution bias, and additionally taking into account the innovation indices and their selected sub-indices, it can be seen that generally a low level of innovativeness is observed in countries with a simultaneous low rate of substitution bias (and vice versa). This is particularly evident when using the k-means clustering method for  $k = 2$ , although the adoption of  $k = 3$  confirms this conclusion to a large extent. In turn, the application of the PROFIT method to selected sub-indices of innovation as well as variables SUB3 and SUB4 leads to the conclusion that although the impact vectors determined by regression lines are not perfectly convergent, they cannot be considered completely divergent. In particular, the interaction with sub-indices SIMP, MS and INST seems to be noticeable on the perception map created with the results of the PROFIT method. Interestingly, the sub-index which seems to have a negative correlation with the variables determining the level of substitute bias (nominal or absolute) is variable INF (cf. Table 5 and Figure 6).

## Bibliography

- Adams J.D., Jaffe A.B., 1996, *Bounding the effects of R&D: an investigation using matched establishment-firm data*, RAND Journal of Economics, vol. 27, no. 4, pp. 700-721.
- Anthony S.D., Johnson M.W., Sinfield J.V., Altman E.J., 2014, *Przez innowację do wzrostu – jak wprowadzać innowację przełomową*, Wolters Kluwer SA, Warszawa.

- Baturo W., 2004, *Nowa encyklopedia powszechna PWN*, Wydawnictwo Naukowe PWN, Warszawa.
- Białek J., 2016, *Reduction in CPI commodity substitution bias by using the modified lloyd-moulton index*, *Statistika – Statistics and Economy Journal*, vol. 96 (2), pp. 51-59, Czech Statistical Office, Praga.
- Białek J., 2017a, *Approximation of the Fisher price index by using the Lloyd-Moulton index: Simulation study*, *Communications in Statistics – Simulation and Computation*, 46(5), pp. 3588-3598.
- Białek J., 2017b, *Approximation of the Fisher price index by using Lowe, Young and AG Mean indices*, *Communications in Statistics – Simulation and Computation*, 46(8), pp. 6454-6467.
- Camba-Mendez G., Gaspar V., Wynne M., 2002, *Measurement issues in european consumer price indices and the conceptual framework of the HICP*, European Central Bank, Frankfurt.
- Clements K.W., Izan H.Y., 1987, *The measurement of inflation: a stochastic approach*, *Journal of Business and Economic Statistics*, 5(3), pp. 339-350.
- Ducharme L.M., 2000, *The Canadian CPI and the bias issue: present and future outlooks*, *Estadística Espanola*, vol. 42, no. 145, pp. 25-41.
- Dutta S., Lanvin B., Wunsch-Vincent S. (eds.), 2017, *The Global Innovation Index 2015. Innovation Feeding the World*, Fontainebleau, Ithaca, and Geneva: Cornell University, INSEAD, the World Intellectual Property Organization (WIPO), [http://www.wipo.int/edocs/pubdocs/en/wipo\\_pub\\_gii\\_2017.pdf](http://www.wipo.int/edocs/pubdocs/en/wipo_pub_gii_2017.pdf); assessed 15.08.2018.
- European Commission, 2010, *EUROPE 2020. A strategy for smart, sustainable and inclusive growth*, Communication from the Commission, COM(2010) 2020, <http://ec.europa.eu/eu2020/pdf/COMPLETE%20EN%20BARROSO%20%20%20007%20-%20Europe%202020%20-%20EN%20version.pdf>; assessed 20.08.2018.
- European Commission, 2013, *Commission launches new innovation indicator*. Brussels: European Commission. Retrieved September 13, 2013, from [http://europa.eu/rapid/press-release\\_IP-13-831\\_pl.htm](http://europa.eu/rapid/press-release_IP-13-831_pl.htm); assessed: 11.09.2018.
- European Commission, 2017, *European Innovation Scoreboard 2017*, Internal Market, Industry, Entrepreneurship and SMEs, Brussels.
- Eurostat, 2013, *COICOP Five-Digit Structure and Explanatory Notes*, [https://www.dst.dk/ext/4197663288/0/pris/COICOP\\_pdf](https://www.dst.dk/ext/4197663288/0/pris/COICOP_pdf); assessed 18.09.2018.
- Gatnar E., Walesiak M., 2004, *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*, Wydawnictwo Akademii Ekonomicznej, Wrocław.
- Grabiński T., Wydymus S., Zeliaś A., 1989, *Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych*, PWN, Warszawa.
- Grębowiec M., 2017, *Uwarunkowania akceptacji innowacyjnych produktów żywnościowych na rynku soków i napojów owocowych*, *Handel Wewnętrzny*, vol. 3(368), pp. 287-300.
- Hall B., Mairesse J., 1995, *Exploring the relationship between R&D and productivity in French manufacturing firms*, *Journal of Econometrics*, vol. 65, pp. 263-293.
- Hałka A., Leszczyńska A., 2011, *Wady i zalety wskaźnika cen towarów i usług konsumpcyjnych – szacunki obciążenia dla Polski*, *Gospodarka Narodowa*, 9(241), pp. 51-75.
- Jajuga K., 1993, *Statystyczna analiza wielowymiarowa*, Wydawnictwo Naukowe PWN, Warszawa.
- Kruskal J.B., 1964, *Nonmetric multidimensional scaling: A numerical method*, *Psychometrika*, 29(2), pp. 115-129.
- Kruskal J.B., Wish M., 1978, *Multidimensional Scaling*, Sage, California.
- Lent J., Dorfman A.H., 2009, *Using a weighted average of base period price indexes to approximate a superlative index*, *Journal of Official Statistics*, 25(1), pp. 139-149.
- Łukasik P., 2017, *Analiza problemów pomiaru innowacyjności przedsiębiorstwa, Nierówności Społeczne a Wzrost Gospodarczy*, vol. 52 (4/2017), pp. 416-423, <https://repozytorium.ur.edu.pl/bitstream/handle/item/3348/30%20%C5%82ukasik-analiza%20problem%C3%B3w.pdf?sequence=1&isAllowed=y>; assessed 10.09.2018.

- Mamica Ł., 2007, *Audyty innowacyjny firm jako narzędzie monitorowania innowacyjności gospodarki*, [in:] S. Mazur (ed.), *Ewaluacja funduszy strukturalnych: perspektywa regionalna*, Uniwersytet Ekonomiczny w Krakowie, Fundacja Uniwersytetu Ekonomicznego, Kraków.
- Marek T., 1989, *Analiza skupień w badaniach empirycznych*, PWN, Warszawa.
- Nordhaus W.D., 1998, *Quality change in price indexes*, *Journal of Economic Perspectives*, vol. 12, pp. 59-68.
- Nowicka A., 2015, *Innowacyjna gospodarka kultura uczenia się przez całe życie*, [in:] C. Schmidt, *Współczesne problemy zarządzania i ekonomii*, Poltext, Warszawa.
- Olejniczak T., 2007, *Innowacja produktowa na rynku żywności- postrzeganie, wpływ na zachowania nabywcze konsumentów*, [in:] T. Czuba, M. Reysowski, M. Skurczyński (ed.), *Innowacje w marketingu 4.0*, Wydawnictwo Uniwersytetu Gdańskiego, Sopot.
- Silver M., Heravi S., 2007, *Why elementary price index number formulas differ: Evidence on price dispersion*, *Journal of Econometrics*, 140(2), pp. 874-883.
- Sneath P.H.A., Sokal R.R., 1973, *Numerical Taxonomy*, W.H. Freeman and Co., San Francisco.
- Stępnia-Kucharska A., 2012, *Działalność innowacyjna przedsiębiorstw przemysłowych w Polsce*, *Studia Prawno-Ekonomiczne*, vol. LXXXVI, pp. 293-319.
- Staśkiewicz J., *Innowacyjność Polski na tle wybranych krajów w ujęciu wskaźników syntetycznych*, *Finanse, Rynki Finansowe, Ubezpieczenia*, vol. 57, pp. 575-589.
- Takane Y., Young F., de Leeuw J., 1976, *Non-metric individual differences multidimensional scaling: an alternating least squares method with optimal scaling features*, *Psychometrika*, 42(1), pp. 7-67.
- Vertesy D., *The Innovation Output Indicator 2017. Methodology Report*, EUR 28876 EN, Publications Office of the European Union, Luxembourg, 2017, doi:10.2760/971852, JRC108942, p. 33, [http://publications.jrc.ec.europa.eu/repository/bitstream/JRC108942/jrc108942\\_ioi\\_2017\\_report\\_final.pdf](http://publications.jrc.ec.europa.eu/repository/bitstream/JRC108942/jrc108942_ioi_2017_report_final.pdf), assessed 11.09.2018.
- Von der Lippe P., 2007, *Index Theory and Price Statistics*, Peter Lang, Frankfurt, Germany.
- White A.G., 1999, *Measurement biases in consumer price indexes*, *International Statistical Review*, 67(3), pp. 301-325.
- Young F.W., Hamer R.M., 1987, *Multidimensional Scaling: History, Theory and Application*, Lawrence Erlbaum Associates, New Jersey.
- Zaborski A., Pełka M., 2013, *Geometrical presentation of preferences by using profit analysis and R program*, *Folia Oeconomica. Acta Universitatis Lodzianensis*, 285, pp. 191-197.

## WIELKOŚĆ OBCIĄŻENIA SUBSTYTUCYJNEGO POMIARU INFLACJI W RELACJI DO POZIOMU INNOWACYJNOŚCI GOSPODAREK UNII EUROPEJSKIEJ

**Streszczenie:** Wskaźnik cen towarów i usług konsumpcyjnych (CPI) jest powszechną miarą inflacji. Podobnie jak zharmonizowany wskaźnik cen (HICP) jest wyznaczany za pomocą indeksu Laspeyresa, dzięki czemu dane o konsumpcji koszyka dóbr nie muszą być bieżące (z okresu badanego). Indeks Laspeyresa, wykorzystując wagi jedynie z okresu bazowego, może nie odzwierciedlać zmian w preferencji konsumentów, jakie nastąpiły w badanym roku. Jest to przyczyna powstawania tzw. obciążenia substytucyjnego przy pomiarze inflacji. Celem pracy jest ocena wpływu poziomu innowacyjności gospodarki danego kraju na występowanie efektu substytucyjnego w pomiarze inflacji. Część empiryczna artykułu bazuje na podstawowych indeksach innowacyjności, tzn. SII, IOI czy GII. Ocena powiązań poziomu innowacyjności i skali efektu substytucyjnego została przeprowadzona z wykorzystaniem metod wielowymiarowej analizy statystycznej (m.in. analizy skupień, metody PROFIT).

**Słowa kluczowe:** pomiar inflacji, innowacyjność, indeksy innowacyjności, analiza skupień, metoda PROFIT.