FACTA UNIVERSITATIS

SERIES **ECONOMICS AND ORGANIZATION**

FACTA UNIVERSITATIS

SERIES **ECONOMICS AND ORGANIZATION**
Vol. 18, Nº 1, 2021

UNIVERSITY OF NIŠ

INSTRUCTIONS FOR CONTRIBUTORS

Contributions should be (preferably) in English, French or German.

Under the paper title, the name(s) of the author(s) should be given while the full name, official title, institute or company affiliation and the like should be placed at the end of the paper together with the exact mail and e-mail address, as well as short (running) title of paper.

Manuscript format. A brief abstract of approximately 100 to 150 words in the same language and a list of up to six key words should precede the text body of the manuscript. Manuscripts should be prepared as doc. file, Word version 6.0 or higher. Manuscript should be prepared using a Word template (downloaded from web address http://casopisi.junis.ni.ac.rs/index.php/FUEconOrg/about/submissions#authorGuidelines).

Manuscript length. Brief articles and discussions (10 pages or less) are encouraged. Otherwise, papers should present well-focused arguments of approximately 16 pages.

Style requirements. **Letters, figures and symbols** should be clearly denoted.

Equations should be typewritten and, with the number, placed in parentheses at the right margin. References to equations should be in the form "Eq. (2)" or simply (2). For equations that cannot be entered in a single line, use the Equation Editor in MS Word. In equations and in the text, *italicize* symbols that are used to represent variables or parameters, including subscripts and superscripts. Only use characters and symbols that are available in the Equation Editor, in the *Symbol font* or in *Times New Roman*.

All illustrations (figures, photographs, line drawings, graphs) should be numbered in series and all legends should be included at the bottom of each illustration. All figures, photographs, line drawings and graphs, should be prepared in electronic form and converted in TIFF or JPG (max quality) file types, in 300 dpi resolution, for superior reproduction. Figures, line drawings and graphs prepared using elements of MS Drawing or MS Graph must be converted in form of pictures and unchangeable. All illustrations should be planned in advance so as to allow reduction to 12.75 cm in column width. Please review all illustrations to ensure that they are readable.

All **tables** should be numbered with consecutive Arabic numbers. They should have descriptive captions at the top of each table and should be mentioned in the text.

References should follow the **APA Style** convention, in alphabetical order at the end of the manuscript. The list of references should be formatted so that the second row in each entry is indented (Paragraph Indentation, Special - choose Hanging, by 0.5cm). Wherever possible, the DOI number should be provided, too, in addition to other reference data.

The APA style **citation** is applied in the text (according to the instructions that can be downloaded from the link http://www.apastyle.org/ or using samples given in the template). Citations in the text should be given in brackets, stating the author's surname, year of publication and, possibly, pages, if it is a direct quote).

Electronic submission. Papers for consideration should be submitted to the Series Editor in electronic form via the Journal's home page: http://casopisi.junis.ni.ac.rs/index.php/FUEconOrg.

FACTA UNIVERSITATIS

Series
Economics and Organization

Vol. 18, Nº 1, 2021

Contents

Original Scientific Paper

TESTING HECKSCHER-OHLIN-VANEK THEOREM BY USING NORMALIZED TRADE BALANCE APPROACH

UDC 332.053

Berislav Žmuk, Hrvoje Jošić

University of Zagreb, Faculty of Economics and Business, Croatia

Abstract. *The study described here introduces new approach for testing the Heckscher-Ohlin-Vanek (HOV) theorem based on the normalized trade balance concept. The intention was to include in the analysis all countries worldwide but due to the lack of data a certain number of countries had to be excluded. Overall 111 countries were observed according to region and income level for the year 2014. The HOV model was estimated using the sign test. It compared the expected sign of the normalized trade balance or net exports, according to the SITC 2 product classification, with the relative endowment of production factors intensively used in the production of a specific product. Production factors were divided into groups such as produced capital, labour force and natural resources further divided into forests, metals and minerals, oil, coal and gas, pastureland and cropland. Researchers in R&D per million people variable represented the impact of technological differences across countries. The results of the sign test have shown that the HOV theorem held in 55% of cases. The percentage of matched signs was highest for the non-OECD high income countries (75%) and lowest for the lower middle income and low income countries (below 50%).*

Key words: *Heckscher-Ohlin-Vanek model, sign test, normalized trade balance approach, the World*

JEL Classification: F1, F2

1. INTRODUCTION

Classical theories of international trade observed price determination and trade pattern only from the supply side. At that time labour was the only factor of production. Adam Smith's theory of absolute advantages determined the pattern of trade and specialization from

---

the different labour productivity in countries, while David Ricardo's theory of comparative advantages was based on different relative labour productivity. After the Industrial Revolution the capital was recognized as a factor of production as well. The neoclassical theory of international trade, established in 1920s and 1930s, was based on the Heckscher-Ohlin theory or also called the Factor Proportions Model (Heckscher, 1919 and Ohlin, 1924). In the Heckscher-Ohlin theory there are two factors of production, labour and capital. Comparative advantages are determined from the relative abundance of production factors. A country which is relatively abundant in a certain factor of production should export a commodity which intensively uses that factor of production. Leontief (1953) confronted the Heckscher-Ohlin theory with data for the United States using 1947 input-output tables. He came to the conclusion that the United States exported labour-intensive products and imported capital-intensive products, which was in contrast with the Heckscher-Ohlin theory, the result later named the Leontief paradox. The Heckscher-Ohlin theory has been subject of rigorous empirical investigation by many scientists but with little success. The results achieved on the sign test were often no better than flipping a coin. Vanek (1968) expanded the standard Heckscher-Ohlin model on a multi-country, multi-factor and multi-commodity framework and explained if a country's endowment in a certain factor of production exceeds that country's share in total World's GDP than that factor of production should be considered abundant. The so called Heckscher-Ohlin-Vanek (HOV) model or factor content model predicted that the content of relatively abundant factor embodied in export commodities should be larger than the content embodied in import commodities. The HOV theorem, however, did poorly in empirical research, primarily because of restrictive assumptions of the very theorem. According to Davis et al. (1997), the HOV theorem is a central theorem in international economics theory but empirically it is a flop. The empirical failure is owed extensively to examining the theory in its least realistic form.

The goal of this paper is to introduce a new approach for testing the Heckscher-Ohlin-Vanek theorem based on the normalized trade balance concept and its application by using sign test. The novelty of our approach is implementation of normalized trade balance concept alongside with the use of sign test instead of calculating the factor content of trade. Therefore, the main advantage of this approach is simplification of the overall model by replacing the complex and complicated input-output matrix calculation process with the normalized trade balance. In addition, this simplified approach allows the larger sample of countries to be included in the analysis. There are five chapters in the paper. After the introduction, the second chapter provides literature review about empirical findings on the HOV model. Data and methodology are presented in chapter three. In the fourth chapter the results of the HOV model analysis are presented and elaborated while the last chapter offers concluding remarks.

2. LITERATURE REVIEW

In this chapter, literature review of empirical studies on the HOV model after Vanek's (1968) seminal paper are to be presented and elaborated. Bowen, Leamer and Sveikauskas (1987) examined the factor content of trade in a multi-factor and multi-country framework. The HOV theorem was tested using sign and rank test on twelve factors of production for 27 countries for the year 1967 by using the 1966 United States technology matrix. The number of correctly matched sign on the sign test was greater than 50% for eleven out of twelve factors

and but greater than 70% for only four factors of production. The authors found that the main reasons why the HOV theorem has no strong support in data are disproportionate consumption, technological differences across countries and measurement errors. Kim (1991) used the factor content model to evaluate the trade patterns of Korea in trade with the United States and Japan using 1978 Korea's and the United States' total input requirements. He observed whether the factor abundance determines the sign of net exports. It was shown that the HOV theorem does not perform well in predicting trade patterns when differences in technological capabilities are ignored. Trefler (1995) investigated the features of data that led to the poor performance of the HOV theorem identifying pronounced patterns in the deviations from the HOV model. Important facts such as the case of missing trade and the endowments paradox have often gone unnoticed. Davis et al. (1997) used Japanese regional data to test the HOV model. The strict HOV theorem performed poorly. The authors made two modifications regarding the general model; they abandoned the notion of identical technologies across countries and focused on production and absorption instead of using trade data. Furthermore, when the assumption of universal factor price equalization was relaxed, the HOV model performed remarkably well. Maskus and Webster (1999) developed a version of the HOV theorem with parametric technological differences. The econometric model on factor contents of trade data, output and consumption for the United States and the United Kingdom allowed for factor-specific and industry-specific productivity differences. Davis and Weinsten (2001) emphasized the importance of intermediates, aggregation bias and differences in patterns of absorption. Choi (2004) relaxed the assumption of ideal factor price equalization deriving a modified HOV theorem to predict the factor content of trade. The modified HOV theorem used input-output coefficients of the source country for each traded good which resulted in approval of theorem validity. Romalis (2004) derived and examined the factor proportions model in commodity markets. He made modifications in the standard Heckscher-Ohlin model by introducing transport costs and monopolistic competition. There was support in data for the Rybczynski effect for fast-growing economies. Chakrabarti (2005) demonstrated the importance of accounting for the absorption of inputs in a factor augmenting model of international productivity differences. The absence of such accounting can lead to biases in calculation of productivity parameters and could raise concerns about the decision about the HOV theorem acceptance or rejection. Requena et al. (2005) studied the HOV model empirically using Spanish regional data. After relaxing the strict assumptions of the HOV theorem such as factor price equalization, identical homothetic preferences and Hicks neutral technological differences, the model performed poorly.

Nishioka (2005) explored the international trade within the HOV model for the developed OECD countries. The knowledge factor was introduced into the HOV framework. It played an important role in determining comparative advantages for technologically advanced OECD countries. The strict HOV model with the inclusion of knowledge capital held on its own. Maskus and Nishioka (2006) estimated factor productivities from individual technology data for 15 OECD countries. The HOV theorem showed ability to explain North-South factor trade depending on factor abundance and productivity gaps. Factor-augmenting productivity differences were found as appropriate modification of the HOV theorem. Artal-Tur et al. (2008) built an assemble dataset for 17 Spanish regions for the years 1995, 2001 and 2004 by employing regional specific input-output tables. The inclusion of intermediate inputs in the computation of technology matrix slightly improved the number of correct matches on the sign test. Lu, Milner and Yu (2009) applied factor content tests using data for 58 countries and six factors of production. The results in general showed weak support for the HOV model with minor improvements achieved after adjusting for technology differences across countries.

Srivastava (2012) tackled the HOV theorem with the help of the excess supply approach by examining trade performances of ten manufacturing industries in 46 countries for the year 2009. The major source of comparative advantages were capital stock and secondary and higher educated labour. The HOV theorem proved to be valid in more than 60% of cases. Srivastava and Mathur (2014) performed partial and complete tests to investigate the validity of the HOV theorem using India's industry level data from 1989 to 2008 and five factors of production (primary educated labour force, secondary and tertiary level of educated labour force, capital and arable land). Measured signs were correct in more than 50% of cases. Zimring (2015) observed a large and rapid expansion of labour force in West Bank due to near-elimination of commuting into Israel. Production shifted to more labour-intensive industries (the Rybczynski effect). Allowing for district specific deviations the changes in production were consistent with the HOV model of trade. Jošić (2016) tested the factor proportions model in the case of Croatia based on the bilateral merchandise trade data between Croatia and the countries of the European Union and worldwide using the sign test. The results of the sign test have shown that Croatia does not use its comparative advantages effectively, leading to the rejection of the factor proportions model. Wu et al. (2017) investigated the greenhouse gas emissions intensities in Canadian agriculture and processed food industry. Natural resources were found to be the determining factor of Canadian agricultural structure whereby Canadian exports were more capital-intensive than imports.

3. DATA AND METHODOLOGY

The intention was to include in the analysis all countries worldwide. Unfortunately, due to the lack of data a certain number of countries had to be excluded from the analysis. Despite that, 111 countries were observed overall. Table 1 reveals that all parts of the World are well represented regionally. Table 2 shows the distribution of observed countries according to their income. The full list of observed countries is displayed in Table 7.

Table 1 Distribution of observed countries according to their geographic region

Region	Number of countries
East Asia & Pacific	13
Europe & Central Asia	37
Latin America & Caribbean	16
Middle East & North Africa	12
North America	2
South Asia	4
Sub-Saharan Africa	27
Total	111

Source: authors according to World Bank (2018, 2019a-d) and Trade Map (2019).

Table 2 Distribution of observed countries according to their income

Income level	Number of countries
Low income	17
Lower middle income	27
Upper middle income	24
High income: non-OECD	14
High income: OECD	29
Total	111

Source: authors according to World Bank (2018, 2019a-d) and Trade Map (2019).

In order to perform the factor endowment analysis, the data have been collected for 10 different variables representing factors of production, GDP and technology differences. The starting point for the variable selection were the previous findings in this field of research with papers presented in the literature review. The final decision upon the list of variables which were included in the analysis, was made based on data availability. World Bank (2018, 2019a-d) and Trade Map (2019) databases were used as data sources. The complete list of observed variables is displayed in Table 3.

Table 3 List of observed variables

Variable group	Variable code	Variable
Income	GDP	Gross domestic product (GDP) (in USD)
Production factors	PCAP	Produced capital (in USD)
	LABF	Labour force (number of persons)
Natural resources	FOR	Forests (in USD)
	MMIN	Metals and minerals (in USD)
	OCNG	Oil, coal and natural gas (in USD)
	PAST	Pastureland (in USD)
	CROP	Cropland (in USD)
	FISH	Fishing (in metric tons)
Technology	R&D	Number of researchers in R&D per million people

Source: authors according to World Bank (2018, 2019a-d) and Trade Map (2019).

Furthermore, the data availability determined the observed period as well. It has been decided that data which are available for the most recent period will be collected. According to the observed databases it turned out that the data for the year 2014 are the most recent one for the most of observed variables. Unfortunately, the R&D variable had missing data for 46 countries (41.44%). In these cases, the data for the period closest to the year 2014 were used as an approximation for 2014 data. The fact that the missing values were imputed by using data from different periods should be taken as a limitation of the research. Therefore, the results where R&D variable was included in the analysis should be observed and discussed with special attention. According to Erlat and Erlat (2003) products can be grouped into five groups with a different product factor intensity level. The goods have been classified into five product groups according to their product intensity. Those are raw material intensive goods (RMIG), labour-intensive goods (LIG), capital-intensive goods (CIG), easy-to-imitate research-intensive goods (EIRIG) and difficult-to-imitate research-intensive goods (DIRIG). The HOV model is defined as follows (Feenstra, 2003):

$$F_i = V_i - s_i V_w \tag{1}$$

where F_i is the factor content of trade of country i, V_i is the factor abundance of the country i, s_i is the share of i-th country's GDP in the World GDP and V_w is the World factor abundance. In the traditional HOV model the signs on the left and right side of the Equation 1 are compared. Standard sign tests of the HOV theorem go roughly as follows: (1) calculate the imports and exports of a country in terms of factors embodied in the goods that are traded, (2) compare the country's share of World's GDP to a country's share of each factor of endowment in the total World's endowment of that factor and (3) a country should be a net exporter of products that intensively uses abundant factor of production.

This paper implements new approach to explore the validity of the Heckscher-Ohlin-Vanek theorem by using normalized trade balance concept on the left side of the Equation 1. Therefore, instead of calculating net factor content of trade from input-output tables, the normalized trade balance was calculated. The normalized trade balance is calculated using the following equation:

$$TB_{ij} = \frac{E_{ij} - I_{ij}}{E_{ij} + I_{ij}} \tag{2}$$

where TB_{ij} is trade balance of country i for the product group j, E_{ij} is export of country i for product group j and I_{ij} is the import of country i for the product group j. According to the Equation 2, if the export of products in a country i is larger than the import of the same product groups, the resulting sign is positive and vice versa. The Equation 2 assumes identical technology and factor content of imports and exports, which can be observed as a disadvantage of new the approach. The normalized trade balance alone, however, does not measure the factor content of trade. It has been used as a concept in the literature for decades as an alternative measure for the revealed comparative advantage. Therefore, in the paper the sign test is conducted by comparing the expected sign of the net exports of SITC 2 product classification with the relative endowment of production factors intensively used in the production of a specific product, as given here:

$$Sign(TB) = Sign\left(V_i^k - s_i V_w^k\right) \tag{3}$$

This paper takes a different approach, testing the HOV theorem good-by-good. That is, instead of testing whether country's total trade in a factor is as we would expect from the country's factor abundance, it tests, for each good the country trades, whether it is traded (on net) in the right direction. For example, if a country is labour abundant, it tests whether each good that is produced with the labour intensive technology is (on net) exported. The whole analysis procedure can be briefly described as follows. Firstly, the standardized trade balances for each of 96 product groups (there are no groups of products with codes 00, 77 and 98 whereas the product group 99 includes everything that was not classified before and because of that it is omitted from the analysis) of all 111 countries are calculated separately. Afterwards, the values of variables for each observed country are compared to the World value and the corresponding share (proportion) is calculated. Due to the specific characteristics of the R&D variable, for this variable the countries values are compared to the World average value. The resulting shares or proportions are then compared to the GDP share of the observed country in the overall World GDP value. If the calculated share is higher than the GDP share, it is assumed that the observed country is abundant in that factor of production. Consequently, the conclusion is that the country should export product which intensively uses the abundant factor of production and in that case the positive sign will be achieved on the sign test. Finally, two estimated signs are compared and it is checked whether they match or not. It is assumed that the sign test will result in a match in at least 50% of cases.

4. RESULTS AND DISCUSSION

The sign test will be conducted by using the total of 10 variables observed in the data and methodology section. In order to get the insight about the distributions of the observed variables in Table 4, the basic descriptive statistics of results is provided.

Table 4 Descriptive statistics results of the observed variables, n=111 selected countries, data for 2014

Variable	Unit	Statistics							
		Average	St. Dev.	Coeff. var.	Min	1st quart.	Median	3rd quart.	Max
GDP	bil. USD	669	2,038	304	1	20	66	405	17,428
PCAP	bil. USD	2,688	8,164	304	3	66	236	1,425	68,943
LABF	mill. per.	27	89	332	0	2	6	19	787
MMIN	bil. USD	87	315	362	0	0	3	21	2,101
FOR	bil. USD	21	52	251	0	1	5	16	353
OCNG	bil. USD	275	765	279	0	0	4	80	4,952
PAST	bil. USD	119	350	295	0	10	29	75	2,848
CROP	bil. USD	225	963	427	0	10	43	123	9,676
FISH	mill. MT	2	8	479	0	0	0	1	76
R&D	no. per mill.	1,566	2,024	129	7	70	565	2,640	7,311

Source: authors' calculation.

According to the descriptive statistics results from the Table 4 it can be easily concluded that there are huge differences between 111 observed countries for each of the selected variables. The lowest variation in data, according to the coefficient of variation, seems to be for the R&D variable (129%) whereas the largest is for the FISH variable (479%). The comparison of differences between the minimum and the maximum values shows that data ranges are very wide. If the values of quartiles are observed, it can be concluded that all variables are highly positively skewed.

The main descriptive statistics results of calculated shares for the observed countries in the whole World value are shown in Table 5. As expected, according to the results from Table 4, huge differences in shares (proportions) are present here as well.

Table 5 Descriptive statistics of calculated shares for the observed countries in the whole World value, n=111 selected countries, data for 2014

Statistics	Variable									
	GDP	PCAP	LABF	MMIN	FOR	OCNG	PAST	CROP	FISH	R&D
Average	0.0085	0.0089	0.0081	0.0086	0.0087	0.0070	0.0085	0.0087	0.0081	1.0628
St.Dev.	0.0258	0.0269	0.0268	0.0310	0.0218	0.0196	0.0250	0.0372	0.0390	1.3742
Coeff.Var.	304	304	332	362	251	279	295	427	479	129
Min	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0049
1st quar.	0.0002	0.0002	0.0007	0.0000	0.0004	0.0000	0.0007	0.0004	0.0001	0.0476
Median	0.0008	0.0008	0.0019	0.0003	0.0019	0.0001	0.0021	0.0016	0.0009	0.3832
3rd quar.	0.0051	0.0047	0.0058	0.0021	0.0067	0.0020	0.0054	0.0048	0.0032	1.7920
Max	0.2202	0.2271	0.2360	0.2069	0.1471	0.1267	0.2036	0.3736	0.3934	4.9628

Source: authors' calculation.
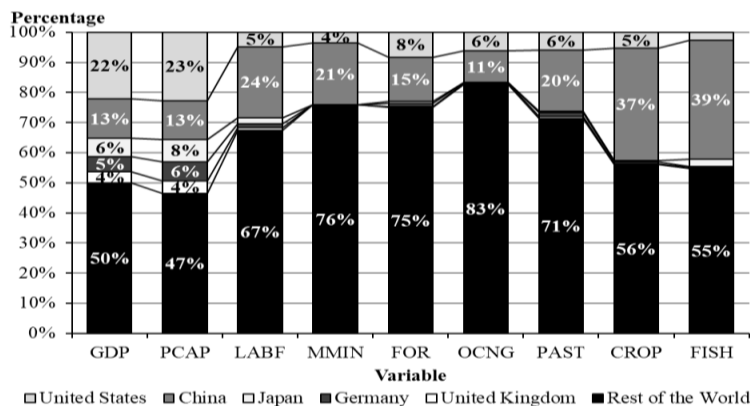
Fig. 1 Factor endowment shares, top five countries according to the GDP share in the World value and the rest of the World, data from 2014

In Figure 1 the top five countries according to the share of GDP value in the total World GDP value are emphasized. According to the Figure 1, those five countries (the United States of America, China, Japan, Germany and the United Kingdom) together encompass the half of the World's total GDP. What's more intriguing, they together dispose with more than half of the World's total produced capital. In other observed variables, the main contributor to the share of those five countries in the total World's values is China meaning that China is abundant in these factors of production. Whereas in Figure 1 five countries with the largest GDP share in the World value were pointed out, in Table 6 countries with the lowest shares of each variable in the whole World value were displayed.

Table 6 The last five observed countries according to observed variables values compared to the World level, data for 2014

Variable	Statistics	Rank 111	Rank 110	Rank 109	Rank 108	Rank 107
GDP	Country	Gambia	Swaziland	Togo	Malawi	Moldova
	Value	1.5911E-05	5.5317E-05	5.7738E-05	7.6427E-05	1.0089E-04
PCAP	Country	Gambia	Togo	Malawi	Rwanda	Madagascar
	Value	9.8129E-06	4.6023E-05	5.1648E-05	5.7452E-05	7.1365E-05
LABF	Country	Iceland	Malta	Luxembourg	Swaziland	Mauritius
	Value	5.9652E-05	5.9865E-05	8.2163E-05	1.2904E-04	1.7930E-04
MMIN	Country	Bahrain, Belgium, Cambodia, Denmark, El Salvador, Estonia, Gambia, Iceland, Italy, Kuwait, Latvia, Lithuania, Malta, Mauritius, Moldova, Nepal, Netherlands, Paraguay, Qatar, Singapore, Slovenia, Swaziland, United Arab Emirates, West Bank and Gaza				
	Value	0				
FOR	Country	Malta, West Bank and Gaza		Bahrain	Iceland	Oman
	Value	0		2.1372E-07	5.3979E-07	8.9456E-07
OCNG	Country	Belgium, Burkina Faso, Cambodia, Costa Rica, El Salvador, Finland, Gambia, Honduras, Iceland, Kenya, Latvia, Luxembourg, Madagascar, Mali, Malta, Mauritius, Namibia, Nicaragua, Panama, Paraguay, Portugal, Rwanda, Singapore, Sri Lanka, Sweden, Switzerland, Togo, Uganda, Uruguay, West Bank and Gaza				
	Value	0				
PAST	Country	Singapore	Malta	Bahrain	Mauritius	Qatar
	Value	1.4069E-07	1.4813E-05	3.4674E-05	3.5316E-05	6.5091E-05
CROP	Country	Iceland	Singapore	Bahrain	Malta	Luxembourg
	Value	2.4472E-06	6.9917E-06	1.0849E-05	1.7096E-05	1.7712E-05
FISH	Country	Luxembourg	Swaziland	Botswana	Macedonia, FYR	Jordan
	Value	0	8.5295E-07	6.0379E-06	7.7334E-06	9.0878E-06
R&D	Country	Congo, Dem. Rep.	Niger	Rwanda	Lao PDR	Tanzania
	Value	4.9069E-03	5.0343E-03	8.3818E-03	1.0743E-02	1.2448E-02

Source: authors' calculation.

Table 7 Results of the sign tests according to the observed countries, data for 2014

Country	Sign matching Yes	No	Share of matched signs	Country	Sign matching Yes	No	Share of matched signs
Albania	27	68	28%	Madagascar	52	44	54%
Argentina	75	21	78%	Malawi	47	49	49%
Australia	57	39	59%	Malaysia	47	49	49%
Austria	58	38	60%	Mali	39	57	41%
Bahrain	84	12	88%	Malta	71	25	74%
Belgium	47	49	49%	Mauritius	59	37	61%
Bolivia	49	47	51%	Mexico	60	36	63%
Bosnia and Herzegovina	37	59	39%	Moldova	41	54	43%
Botswana	50	46	52%	Morocco	35	61	36%
Brazil	65	31	68%	Mozambique	40	56	42%
Bulgaria	35	61	36%	Namibia	54	42	56%
Burkina Faso	46	48	49%	Nepal	48	48	50%
Cambodia	50	45	53%	Netherlands	62	34	65%
Canada	54	42	56%	Nicaragua	39	57	41%
Chad	42	53	44%	Niger	23	70	25%
Chile	74	22	77%	Nigeria	36	60	38%
China	61	35	64%	Norway	65	31	68%
Colombia	48	48	50%	Oman	87	8	92%
Congo, Dem. Rep.	30	66	31%	Pakistan	57	39	59%
Congo, Rep.	23	72	24%	Panama	79	17	82%
Costa Rica	49	47	51%	Papua New Guinea	43	53	45%
Cote d'Ivoire	48	48	50%	Paraguay	49	47	51%
Croatia	63	33	66%	Philippines	52	44	54%
Denmark	49	47	51%	Poland	47	49	49%
Ecuador	44	52	46%	Portugal	44	52	46%
Egypt, Arab rep.	49	47	51%	Qatar	88	8	92%
El Salvador	45	51	47%	Romania	34	59	37%
Estonia	57	39	59%	Russian Federation	63	33	66%
Ethiopia	48	47	51%	Rwanda	43	53	45%
Finland	64	32	67%	Senegal	45	51	47%
France	59	37	61%	Singapore	67	29	70%
Gambia, The	41	53	44%	Slovak Republic	59	37	61%
Georgia	35	61	36%	Slovenia	58	38	60%
Germany	51	45	53%	South Africa	36	59	38%
Ghana	41	55	43%	Spain	45	51	47%
Greece	54	42	56%	Sri Lanka	60	36	63%
Guatemala	43	53	45%	Swaziland	35	61	36%
Honduras	39	57	41%	Sweden	63	33	66%
Hungary	45	51	47%	Switzerland	63	33	66%
Iceland	57	39	59%	Tanzania	50	46	52%
India	62	34	65%	Thailand	63	33	66%
Indonesia	59	37	61%	Togo	41	51	45%
Iraq	87	3	97%	Tunisia	43	53	45%
Ireland	66	30	69%	Turkey	64	32	67%
Italy	44	52	46%	Uganda	50	46	52%
Japan	68	28	71%	Ukraine	39	56	41%
Jordan	54	41	57%	United Arab Emirates	70	26	73%
Kazakhstan	72	23	76%	United Kingdom	68	28	71%
Kenya	50	46	52%	United States	59	37	61%
Korea, Rep.	64	32	67%	Uruguay	63	33	66%
Kuwait	89	7	93%	Venezuela, RB	75	21	78%
Lao PDR	45	51	47%	Vietnam	53	42	56%
Latvia	54	42	56%	West Bank and Gaza	30	62	33%
Lithuania	58	38	60%	Zambia	32	64	33%
Luxembourg	61	35	64%	Zimbabwe	43	53	45%
Macedonia, FYR	41	54	43%				

Source: authors' calculation.

The calculated shares for the observed countries in the whole World value are paired with a certain product groups and are compared with the GDP share. In the case when the calculated share for a certain product group in a country is larger than the share of GDP in

total World's GDP value, a positive sign is attached to that product group. Namely, in that case, it is expected that the country should have a positive net exports in that product. The Equation 2 is used to calculate the normalized trade balances for each product groups and for each country. If the normalized trade balance results in a positive score, a positive sign should also be obtained for country's factor endowment. The sign test is conducted by checking whether the two estimated signs match or not. The assumption is that the factor endowment signs and normalized trade balance signs would match in at least 50% of cases.

In Table 7 the results of the sign tests for each of the observed countries are shown. Overall sign matching rate turned out to be 55% (5,845 matched signs and 4,774 unmatched signs) which is similar to what has been found out within the standard signs test in previous studies leading to the conclusion that HOV theorem does not offer a very good description of reality. The countries with the highest sign matching rates, which are above 90%, are Iraq (97%), Kuwait (93%), Qatar (92%) and Oman (92%). On the other hand, countries with the lowest sign matching rates, which are below 30%, are the Republic of the Congo (24%), Niger (25%) and Albania (28%).

Table 8 Results of the sign tests according to the groups of products, data for 2014

| Code	Product	Sign matching		Share of matched signs
		Yes	No	
01	Live animals	44	67	40%
02	Meat and edible meat offal	52	59	47%
03	Fish and crustaceans, molluscs and other aquatic invertebrates	75	36	68%
04	Dairy produce; birds' eggs; natural honey; edible products of animal origin	45	66	41%
05	Products of animal origin, not elsewhere specified or included	55	56	50%
06	Live trees and other plants; bulbs, roots and the like; cut flowers etc.	65	46	59%
07	Edible vegetables and certain roots and tubers	68	43	61%
08	Edible fruit and nuts; peel of citrus fruit or melons	77	34	69%
09	Coffee, tea, maté and spices	77	34	69%
10	Cereals	49	62	44%
11	Products of the milling industry; malt; starches; inulin; wheat gluten	48	63	43%
12	Oil seeds and oleaginous fruits; miscellaneous grains, seeds and fruit; etc.	67	44	60%
13	Lac; gums, resins and other vegetable saps and extracts	54	57	49%
14	Vegetable plaiting materials; vegetable products not elsewhere spec. etc.	78	31	72%
15	Animal or vegetable fats and oils and their cleavage products; etc.	64	47	58%
16	Preparations of meat, of fish or of crustaceans, molluscs etc.	73	38	66%
17	Sugars and sugar confectionery	60	51	54%
18	Cocoa and cocoa preparations	58	53	52%
19	Preparations of cereals, flour, starch or milk; pastry cooks' products	46	65	41%
20	Preparations of vegetables, fruit, nuts or other parts of plants	64	47	58%
21	Miscellaneous edible preparations	43	68	39%
22	Beverages, spirits and vinegar	65	46	59%
23	Residues and waste from the food industries; prepared animal fodder	56	55	50%
24	Tobacco and manufactured tobacco substitutes	53	58	48%
25	Salt; sulphur; earths and stone; plastering materials, lime and cement	64	47	58%
26	Ores, slag and ash	81	30	73%
27	Mineral fuels, mineral oils and products of their distillation; etc.	95	16	86%
28	Inorganic chemicals; organic or inorganic compounds etc.	68	43	61%
29	Organic chemicals	76	35	68%
30	Pharmaceutical products	85	26	77%
31	Fertilisers	64	47	58%
32	Tanning or dyeing extracts; tannins and their derivatives; dyes, etc.	78	33	70%
33	Essential oils and resinoids; perfumery, cosmetic or toilet preparations	65	46	59%
34	Soap, organic surface-active agents, washing preparations, etc.	70	41	63%

Code	Product	Sign matching		Share of matched signs
		Yes	No	
35	Albuminoidal substances; modified starches; glues; enzymes	85	26	77%
36	Explosives; pyrotechnic products; matches; pyrophoric alloys; etc.	74	37	67%
37	Photographic or cinematographic goods	81	30	73%
38	Miscellaneous chemical products	82	29	74%
39	Plastics and articles thereof	85	26	77%
40	Rubber and articles thereof	35	76	32%
41	Raw hides and skins (other than fur skins) and leather	62	49	56%
42	Articles of leather; saddlery and harness; travel goods, handbags etc.	55	56	50%
43	Fur skins and artificial fur; manufactures thereof	45	60	43%
44	Wood and articles of wood; wood charcoal	78	33	70%
45	Cork and articles of cork	49	62	44%
46	Manufactures of straw, of esparto or of other plaiting materials; etc.	70	41	63%
47	Pulp of wood or of other fibrous cellulosic material; recovered etc.	60	50	55%
48	Paper and paperboard; articles of paper pulp, of paper or of paperboard	39	72	35%
49	Printed books, newspapers, pictures and other products of the printing etc.	37	74	33%
50	Silk	51	58	47%
51	Wool, fine or coarse animal hair; horsehair yarn and woven fabric	38	73	34%
52	Cotton	55	56	50%
53	Other vegetable textile fibres; paper yarn and woven fabrics of paper yarn	54	57	49%
54	Man-made filaments; strip and the like of man-made textile materials	44	67	40%
55	Man-made staple fibres	46	65	41%
56	Wadding, felt and nonwovens; special yarns; twine, cordage, ropes etc.	43	68	39%
57	Carpets and other textile floor coverings	50	61	45%
58	Special woven fabrics; tufted textile fabrics; lace; tapestries; etc.	42	69	38%
59	Impregnated, coated, covered or laminated textile fabrics; etc.	36	75	32%
60	Knitted or crocheted fabrics	36	74	33%
61	Articles of apparel and clothing accessories, knitted or crocheted	71	40	64%
62	Articles of apparel and clothing accessories, not knitted or crocheted	70	41	63%
63	Other made-up textile articles; sets; worn clothing, etc.	59	52	53%
64	Footwear, gaiters and the like; parts of such articles	61	50	55%
65	Headgear and parts thereof	57	54	51%
66	Umbrellas, sun umbrellas, walking sticks, seat-sticks, whips, etc.	49	61	45%
67	Prepared feathers and down and articles made of feathers or of down; etc.	60	51	54%
68	Articles of stone, plaster, cement, asbestos, mica or similar materials	42	69	38%
69	Ceramic products	47	64	42%
70	Glass and glassware	42	69	38%
71	Natural or cultured pearls, precious or semi-precious stones, etc.	65	45	59%
72	Iron and steel	70	41	63%
73	Articles of iron or steel	72	39	65%
74	Copper and articles thereof	72	39	65%
75	Nickel and articles thereof	66	45	59%
76	Aluminium and articles thereof	68	43	61%
78	Lead and articles thereof	55	55	50%
79	Zinc and articles thereof	72	38	65%
80	Tin and articles thereof	59	50	54%
81	Other base metals; cermets; articles thereof	59	49	55%
82	Tools, implements, cutlery, spoons and forks, of base metal; etc.	41	70	37%
83	Miscellaneous articles of base metal	42	69	38%
84	Machinery, mechanical appliances, nuclear reactors, boilers; parts thereof	87	24	78%
85	Electrical machinery and equipment and parts thereof; etc.	84	27	76%
86	Railway or tramway locomotives, rolling stock and parts thereof; etc.	81	30	73%
87	Vehicles other than railway or tramway rolling stock, and parts etc.	78	33	70%
88	Aircraft, spacecraft, and parts thereof	75	35	68%
89	Ships, boats and floating structures	76	34	69%
90	Optical, photographic, cinematographic, measuring, checking, etc.	91	20	82%
91	Clocks and watches and parts thereof	74	37	67%
92	Musical instruments; parts and accessories of such articles	54	57	49%
93	Arms and ammunition; parts and accessories thereof	32	66	33%
94	Furniture; bedding, mattresses, mattress supports, cushions etc.	50	61	45%
95	Toys, games and sports requisites; parts and accessories thereof	52	59	47%
96	Miscellaneous manufactured articles	43	68	39%
97	Works of art, collectors' pieces and antiques	55	55	50%
Total		5,845	4,774	55%

Note: due to the limited space, the titles of product groups have been cut.

Source: authors' calculation.

Table 8 shows the results of the sign tests according to the product groups. The product groups with the highest sign matching rates, which are above 80%, are Mineral fuels, mineral oils and products of their distillation; bituminous substances; etc. – code 27 (86%) and Optical, photographic, cinematographic, measuring, checking, precision, medical or surgical etc. – code 90 (82%). On the other hand, the product groups with the lowest sign matching rates, which are below 35%, are Rubber and articles thereof – code 40 (32%), Impregnated, coated, covered or laminated textile fabrics; textile articles of a kind suitable etc. – code 59 (32%), Arms and ammunition; parts and accessories thereof – code 93 (33%), Knitted or crocheted fabrics – code 60 (33%), Printed books, newspapers, pictures and other products of the printing industry; manuscripts – code 49 (33%) and Wool, fine or coarse animal hair; horsehair yarn and woven fabric – code 51 (34%).

In Figure 2 the percentages of matched signs according to geographical region of a country are shown. The most successful matched signs rate was achieved in countries from the Middle East and North Africa (69%). The key finding is that HOV mainly holds in oil-rich economies. One of the reasons for this could be a focus on exports of only one essential product, oil, but a more detailed analysis should be conducted to validate this result. On the other hand, only countries from Sub-Saharan Africa achieved the percentage of matched signs below 50%.
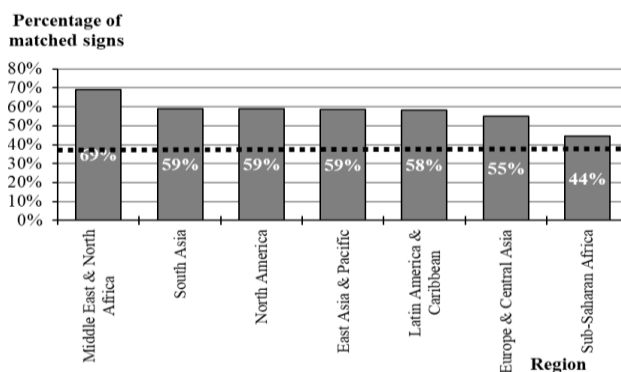
Fig. 2 Percentage of matched signs according to geographical region of a country, n=111 observed countries, data for 2014
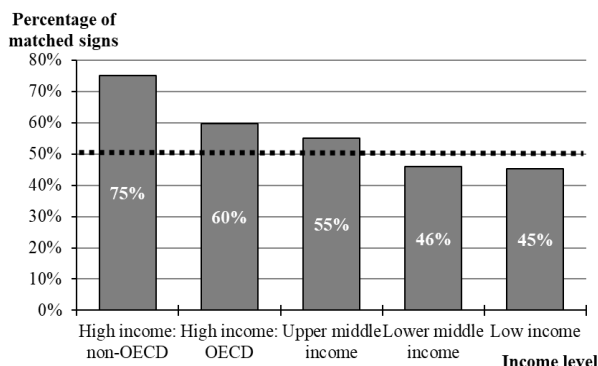
Fig. 3 Percentage of matched signs according to income level of countries, n=111 observed countries, data for 2014

The successfulness of signs matching was observed according to the country income level as well. Figure 3 shows that the highest percentage of matched signs is achieved for countries having high income and which are not members of the Organisation for Economic Cooperation and Development (OECD). Countries with lower middle income (46%) and low income (45%) achieved lower percentage of matched signs in comparison to the benchmark value of 50%. It could be asked: why do rich countries have higher matching rates than poor countries? The reason could be more efficient use of country's comparative advantages and the specialization in production and export.
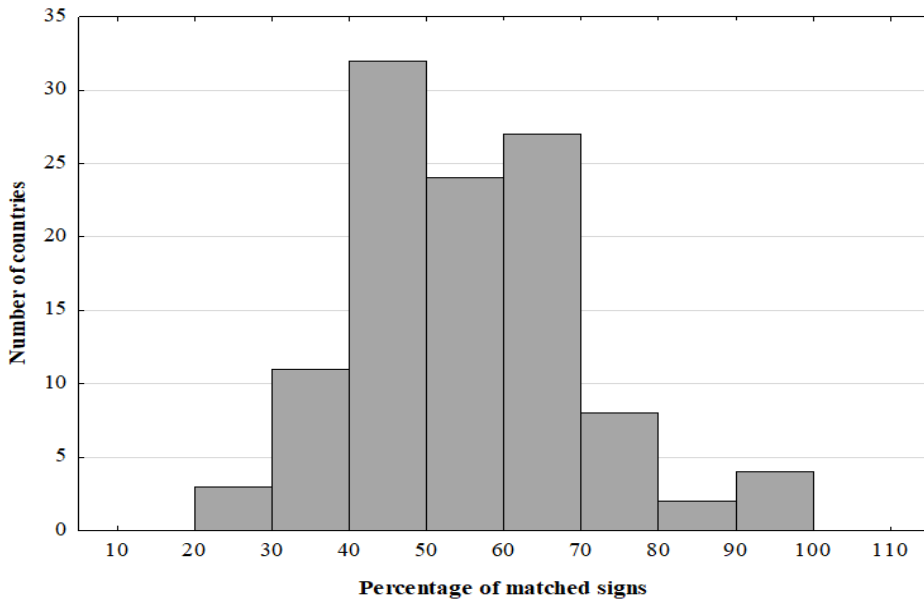
Fig. 4 Histogram of distribution of matched signs, n=111 observed countries, data for 2014

Finally, in Figure 4 the histogram of distribution of matched signs across countries is shown. It turned out that most countries (32) have a percentage of matched signs between 40% and 50%. It has to be emphasized that three countries, out of thirty-two, have a percentage of matched signs that equals 50%. However, 68 countries or 61% of observed countries have a percentage of matched signs above the benchmark value of 50%.

Previous empirical tests mainly failed to prove HOV theorem, or did so only in a certain extent, which provides ground for further (amended) tests. We should keep in mind that completely random pattern of signs, such as obtained by flipping a coin, would still generate correct signs 50% of the time in a large sample. Overall result of the sign test conducted in this paper is only 55% of matched signs, which is not enough to present it as a proof of HOV theorem validity. Therefore, the sign test must do considerably better in order to be concluded that the HOV theorem is empirically sound. The advantage of this approach is that it allows for the larger sample of countries to be included in the analysis. Disadvantage of this approach is that it assumes the identical technology in countries despite the inclusion of the R&D variable. Also gross value of trade flows does not reveal country's value added comprised in these flows. For example, China's huge exports of hi-

tech electronics do not represent its net exports of capital intensive and R&D intensive products, because its value added in this sector is different from value added upward in the global value chain.

5. CONCLUSIONS

This paper introduces new approach for testing the Heckscher-Ohlin-Vanek theorem. The HOV theorem was tested in the case of 111 countries worldwide. Instead of calculating the factor content requirements based on the input-output tables, the normalized trade balance was calculated. The products or goods are divided into five groups according to their production intensity (labour-intensive goods, capital intensive goods, natural-resources intensive goods, easy-to-imitate technology intensive goods and hard-to-imitate technology intensive goods). Furthermore, the countries' factors of production are divided into produced capital, labour force and natural resources further divided into six factors. The sign test inspected correct matching signs between normalized trade balance indices and factor endowments for each country and product according to the SITC 2 classification of products. Overall sign matching rate turned out to be 55% (5,845 matched signs and 4,774 unmatched signs). Countries with the highest sign matching rates of above 90% were Iraq (97%), Kuwait (93%), Qatar (92%) and Oman (92%). Countries with the lowest sign matching rates, of below 30% were the Republic of the Congo (24%), Niger (25%) and Albania (28%). It seems that countries with a higher income level have higher percentage of matched signs (high-income non-OECD and high-income OECD countries with 75% and 60% of sign matches respectively). On the other hand, lower-middle income and low-income countries had a percentage of matched signs below 50%. The fact that the missing values were imputed by using data from other periods should be taken as a limitation of research so the results where the R&D variable was included in the analysis should be carefully discussed. Further research should be made by using this new approach. Improvement could be made by including productivity differences among countries by calculating the effective factor endowments.

REFERENCES

Artal-Tur, A., Castillo-Giménez, J., Requena-Silvente, R., & Llano-Verduras, C. (2008). Extending the Heckscher-Ohlin-Vanek model in a regional setting: the role of intermediates, technological differences and space. http://www.etsg.org/ETSG2009/papers/artal.pdf

Bowen, H. P., Leamer, E. E., & Svikauskas, L. (1987). Multi-country, Multi-factor Tests of the Factor Abundance Theory. *The American Economic Review*, 77(5), 791-809.

Chakrabarti, A. (2005). The Heckscher-Ohlin-Vanek Theorem and International Factor Price Differences. *Open economies review*, 16, 407–412.

Choi, E. K. (2004). The Hescher-Ohlin-Vanek Theorem without Factor Price Equalization. Draft (April 19, 2004). https://pdfs.semanticscholar.org/325c/ad4db297024a046b5b22bd2d93f1e9db4658.pdf

Davis, D. R., Weinstein, D. E., Bradford, S. C., & Shimpo, K. (1997). Using International and Japanese Regional Data to Determine When the Factor Abundance Theory of Trade Works. *The American Economic Review*, 87, 421-446.

Davis, D., & Weinsten, D. (2001). The factor content of trade, Working Paper 8637. http://www.columbia.edu/~dew35/PDF/Hdbk_Empir.pdf

Erlat, G., & Erlat, H. (2003). Measuring Intra-Industry and Marginal Intra-Industry Trade: The Case for Turkey. *Emerging Markets Finance and Trade*, 39(6), 5-38.

Feenstra, R. C. (2003). *Advanced international trade: Theory and Evidence*. Princeton, NJ: Princeton University Press.

Heckscher, E. F. (1919). The effect of foreign trade on the distribution of income. Ekonomisk Tidskrift, *21*, 497-512. Reprinted in Ellis, H.S. & L.A. Metzler, (eds.), *Readings in the Theory of International Trade*. London: Allen and Undwin.

Jošić, H. (2016). Testing the factor proportions model for Croatia. *Ekonomski vjesnik/EconViews*, 2017(2), 353-370.

Kim, U. Y. (1991). Heckscher-Ohlin-Vanek theory and empirical study. *The Korean Economic Review*, *7*(2), 173-189.

Leontief, W. (1953). Domestic Production and Foreign Trade: The American Capital Position Re-Examined. *Proceedings of the American Philosophical Society*, 97, 332-349.

Lu, M., Milner, C., & Yu, Z. (2009). Testing the H-O-V Model in Bilateral Trade with Allowance for Technology Differences. European Trade Study Group (ETSG) Annual Conference at the University of Rome "Tor Vergata" (September 10-12, 2009).

Maskus, K. E., & Nishioka, S. (2006). Development-Related Biases in Factor Productivities and the HOV Model of Trade. *Discussion papers in economics*, Working Paper No. 06-08.

Maskus, K., E., & Webster, A. (1999). Estimating the HOV Model with Technology Differences Using Disaggregated Labor Skills for the United States and the United Kingdom. *Review of International Economics*, *7*(1), 8-19.

Nishioka, S. (2005). An Explanation of OECD Factor Trade with Knowledge Capital and the HOV Model. *Discussion papers in economics*, Working Paper No. 05-06. https://www.colorado.edu/economics/sites/default/files/attached-files/wp05-06.pdf

Ohlin, B. (1924). The Theory of Trade. Translated and reprinted in Flam, H., & Flanders, M. J. (editors and translators), *Heckscher-Ohlin Trade Theory*. Cambridge: MIT Press.

Requena, F., Castillo, J., & Artal, A. (2005). Multiregion, multifactor tests of the Heckscher-Ohlin-Vanek using Spanish regional data. XXX Simposio Análisis Económico. Murcia: Servicio de Publicaciones de la Universidad de Murcia, 15-17 Septembre 2005.

Romalis, J. (2004). Factor proportions and the structure of commodity trade. *American Economic Review*, *94*, 67-97.

Srivastava, A. (2012). Heckscher Ohlin Vanek Theorem: an excess supply approach, Munich Personal RePEc Archive. https://ideas.repec.org/p/pra/mprapa/38279.html

Srivastava, A., & Mathur, S. K. (2014). Validity of Heckscher-Ohlin-Vanek Hypothesis - A Complete and Partial Test Approach. *Korea and the World Economy*, *15*(3), 355-393.

Trade Map (2019). List of products imported and exported by certain country. https://www.trademap.org/Product_SelCountry_TS.aspx

Trefler, D. (1995). The Case of the Missing Trade and Other Mysteries. *American Economic Review*, 1029-1046.

Vanek, J. (1968). The factor proportions theory: The N-factor case. *Kyklos*, *2*, 749-754.

World Bank (2018). The Changing Wealth of Nations 2018: Country Tool. https://development-data-hub-s3-official.s3.amazonaws.com/ddhfiles/94641/cwon2018-country-tool-v1.0_0.xlsx

World Bank (2019a). Capture fisheries production (metric tons). https://data.worldbank.org/indicator/ER.FSH.CAPT.MT

World Bank (2019b). GDP (current US$). https://data.worldbank.org/indicator/NY.GDP.MKTP.CD

World Bank (2019c). Labor force, total. https://data.worldbank.org/indicator/SL.TLF.TOTL.IN

World Bank (2019d). Researchers in R&D (per million people). https://data.worldbank.org/indicator/SP.POP.SCIE.RD.P6

Wu, T., Tomassin, P. J., & Mukhopadhyay, K. (2017). Using H-O-V Theorem to Predict the Factor Intensities in Canadian Agricultural Trade. *International Journal of Applied Economics*, *14*(1), 45-64.

Zimring, A. (2015). Testing the Heckscher-Ohlin-Vanek Theory with a Natural Experiment. *Canadian Journal of Economics*, *52*(1), 58-92.

TESTIRANJE HECKSCHER-OHLIN-VANEKOVOG TEOREMA KORIŠĆENJEM PRISTUPA NORMALIZOVANOG TRGOVINSKOG BILANSA

Studija opisana ovde uvodi pristup normalizovanog trgovinskog bilansa za testiranje Heckscher-Ohlin-Vanekovog (HOV) teorema. Namera je bila da se sve zemlje širom sveta uključe u analizu, ali je zbog nedostatka podataka određeni broj zemalja morao da bude isključen iz analize. Zbog toga je ukupno 111 zemalja primećeno prema regionu i nivou prihoda u 2014. Pri proceni HOV modela koristio se test predznaka. Test je uporedio očekivani znak normalizovanog trgovinskog bilansa ili neto izvoza prema klasifikaciji proizvoda SITC 2 sa relativnom zadužbinom proizvodnih faktora koji se intenzivno

koriste u proizvodnji određenog proizvoda. Proizvodni faktori su podeljeni na grupe kao što su proizvedeni kapital, radna snaga i prirodni resursi koji su dodatno podeljeni na šume, metale i minerale, naftu, ugalj i gas, pašnjake i useve. Varijabla istraživači u R&D na milion ljudi predstavlja uticaj tehnoloških razlika širom zemalja. Rezultat testa predznaka je pokazao da HOV teorem drži u 55 odsto slučajeva. Procenat podudarnih predznakova je najveći za zemlje sa visokim prihodima koje nisu članice OECD-a (75%) i najniži za zemlje sa nižim srednjim prihodima i niskim prihodima (ispod 50%).

Ključne reči: *Heckscher-Ohlin-Vanekov model, test predznaka, pristup normalizovanog trgovinskog bilansa, svet*

Original Scientific Paper

POPULATION AGEING AND WELFARE DISSIMILARITIES WITHIN THE EUROPEAN UNION: A NEW APPROACH BASED ON CLUSTER ANALYSIS

UDC 330.342(4-672EU)

Mirela Cristea[1], Graţiela Georgiana Noja[2], Yannis Thalassinos[3]

[1]University of Craiova, Faculty of Economics and Business Administration,
Craiova, Romania
[2]West University of Timişoara, Faculty of Economics and Business Administration,
East European Center for Research in Economics and Business, Timişoara, Romania,
[3]Gulf University in Science and Technology, Kuwait, Kuwait

Abstract. *The research aims to identify several dissimilarities between the European Union Member States (EU-27 MS) in terms of welfare and labour market dimensions under the sheer implications of the ageing phenomenon. The quantitative research methodology emphasizes the cluster analysis based on the Ward method, performed for the year 2018. Main results denote that only two countries (Denmark and Sweden) registered soaring performances, especially for the labour market credentials (particularly the employment rate and active policies). Other 10 EU-27 MS accounted medium performances in terms of well-being, but also the lowest achievements as regards the old dependency rate, the employment rate of persons aged 55-64 and the birth rate. This paper brings to the fore the keen need to redesign specific policies and tailored strategies by the responsible authorities and business representatives across the EU, in order to enhance achievements and new solutions for the difficulties brought by population ageing, with spillover effects on the labour market integration of older employees and overall economic welfare. The study stands out through the new integrative approach based on cluster analysis that underlines the dissimilarities between the EU member states, and the features of each group of countries, in a pre-settled framework, thus grasping the difficulties, but also the opportunities faced in terms of ageing and economic welfare.*

Key words: *labour market, economic development, demographic economics, cluster analysis*

1. Introduction

Population ageing represents a topical subject on the European and global agenda, being placed in the open conversation on the transition to old societies that is already well under way, since the proportion of youg people has diminished compared to a few decades ago, while the share of early retirees has significantly increased (Sidlo et al., 2020; Wolf, 2020). These challenges are amplified today in the face of the Covid-19 pandemic crisis (Kashnitsky & Aburto, 2020; McKibbin & Fernando, 2020), while the globalization process weakens, and also the labour force shrinks since the growth of the working-age population is severely slowing in the core of the world economy (Skufina et al., 2020), and, particularly, in the European economies (Calvo-Sotomayor et al., 2019; Cristea et al, 2020a; Jivraj et al., 2020).

The European Union-wide reference study on grouping the Member States (MS) according to the representative indicators of population ageing measurement is the one deployed by the United Nations Economic Commission for Europe (UNECE) and European Commission (EC) (UNECE/EC, 2019), considering the Active Ageing Index (AAI) at the level of 2018. The AAI represents a composite indicator, including 22 indicators, grouped in four domains, namely, "employment, participation in society, independent healthy and secure living, and capacity and enabling environment for active ageing" (UNECE/EC, 2019, p. 17). As stated by in this report, the European Union (EU) MS were grouped according to the similar characteristics in terms of AAI, resulting in 4 groups of countries relative to the EU average (35.8), with scores between 28.1 - in the case of Greece, and 46.9 - in the case of Sweden. Moreover, in order to measure the existing inequalities between the EU MS, in this report, the AAI was connected for each country "to the GDP per capita, Gini index and life satisfaction of older people in the EU28" (UNECE/EC, 2019, p. 54), resulting in significant differences among them.

In this frame of reference, current paper aims to advance the UNECE/CE (2019) study by deploying a cluster analysis across the EU-27 MS, at the level of 2018, in order to assess the interlinkages between population ageing, labour market outcomes and welfare dissimilarities, and to propose specific policies and strategies for each group of countries.

After a concise introduction on the research topic, the paper presents a synthesis of the main results obtained in the literature, focused on the implications of demographic ageing on the labour market and economic development, as well as differences and similarities between states, based on these coordinates. Then follows the description of the data and the research methodology used, respectively the cluster analysis by the Ward method. The obtained results and the discussions are related to those obtained in the specialized literature and are supported by the data from the Annex. The conclusions summarize the rendering of the research hypothesis and the main strategies and policies proposed for each group of states.

2. Brief Literature Review

The connections between demographic ageing, the labour market and the degree of economic development of countries have been intensely debated in the literature. A number of studies have highlighted the unfavorable implications of the ageing population on labour productivity and, implicitly, on economic growth in all regions of the world.

Thereby, in the United States of America (USA), Maestas et al. (2016) showed that an increase in the stock of population aged over 60 years by 10% generates a reduction in the Gross Domestic Product (GDP) growth rate per capita by 5.5%, due to the lower increase of labour

productivity by age groups, concomitant with the easy-going increase of the working age population. At the EU-28 level, Cristea et al. (2020a) showed that the unfavorable implications of demographic ageing (measured by the workforce group aged between 55-64 years) on labour productivity were manifested only in countries that registered the AAI below the EU average or slightly above it. Similar conclusions were drawn by Calvo-Sotomayor et al. (2019) for 24 EU countries, namely the workforce ageing has negatively influenced the labour productivity for the analysed period 1983-2014. The confined analysis only for the new EU countries (EU-13) revealed that labour productivity in these countries "is shaped under the complex implications of older employment and ageing dimensions, further connected to health and well-being", with unfavourable implications if proper policies and strategies would not be implemented (Cristea et al., 2020b, p. 74).

Conde-Sala et al. (2016, p. 1059), studying the perceived life quality of people aged over 65 years at the European level, based on the results of the Survey of Health, Aging and Retirement in Europe (SHARE) in 2013, revealed that the percieved life is much better felt in the cluster of the Nordic countries (with a social democratic model), but also at the level of Continental countries, which have a "corporate model" (characterized by better socio-economic indicators), compared with clusters comprising the Eastern European countries and from the Mediterranean area (which have lower socio-economic indicators). Due to heteregeneity among the European countries, Suchecka and Urbaniak (2016, p. 173) proved that "it is possible to divide European countries on the basis of similarities of the status of people aged 65+", related to the economic and social conditions, with common strategies and policies for each group of countries (clusters). Authors (Suchecka & Urbaniak, 2016) found similarities at the year of 2012, by applying cluster analysis, as regards health and ageing dimensions among Switzerland, Sweden and Norway, on the one hand, and Central and Eastern European countries, on the other hand. Carrying out an analysis at the level of EU countries regarding the way in which active ageing (measured by AAI) correlates with the degree of economic development (measured by the GDP per capita and labour productivity), the dimensions of the labour market and other representative indicators of demographic ageing (old dependency ratio) at the level of 2018, Thalassinos et al. (2019, p. 598) concluded that "the lowest performance in terms of ageing, welfare and labour market policies was registered by the EU–13 countries (including also Greece and Portugal), and the highest, by countries from the old EU–15, especially by the Nordic States (Sweden, Denmark, Finland)".

The analysis of the Slovenian regions grouping, made by Rovan and Sambt (2003) by applying the cluster methodology, based on demographic factors (population ageing, population growth rate and migration) and certain socio-economic variables available at the territorial level (such as per capita income tax base, unemployment rate, population employed in agriculture, tertiary education and the car park), revealed that the most developed regions were characterized by a demographic ageing and a higher rate of lower unemployment, while in the less developed regions the situation was reversed. These results point out the same unfavorable dependence between population ageing and the degree of economic development of the country.

As such, the main results in the literature show significant differences between countries or regions in terms of the economy's response to the demographic ageing, for which there are different measurement units. In addition, the researchers show that the countries' homogeneity in terms of demographic ageing and well-being is changing from one period to another (Suchecka & Urbaniak, 2016).

<div align="center">3. DATA AND METHODOLOGY</div>

By reviewing the relevant underpinnings of the literature related to our own research subject, we have grouped the data into three categories of indicators, namely: *welfare variable, ageing credentials* and *labour market specific indicators*, as follows:

(*i*) *Welfare indicator*: Gross Domestic Product (GDP) per capita (*GDP_C*) (measured in constant 2010 United States Dollars – USD), for which the data were extracted from the World Development Indicators (WDI) (The World Bank, 2020);

(ii) *Ageing representative indicators*: Old age dependency ratio (*OD_65*) as share of population over 65 years aged to active population (15-64 years) (%); Life expectancy at birth total population (*LE*) (years); Crude birth rate (*BR*) (measured in "number of live births per 1,000 persons"), with data extracted from the Eurostat (European Commission, 2020a);

(*iii*) *Labour market specific indicators*: Employment rate, 55-64 years aged group or workforce ageing (*ER_55_64*) (% of total population), data extracted from the Eurostat (European Commission, 2020a); Active labour market policies (*ALMP*) (% of GDP) and Passive labour market policies (*PLMP*) (% of GDP), with data extracted from the Employment, Social Affairs & Inclusion (European Commission, 2020b).

The analysis is performed at the level of EU-27 for 2018.
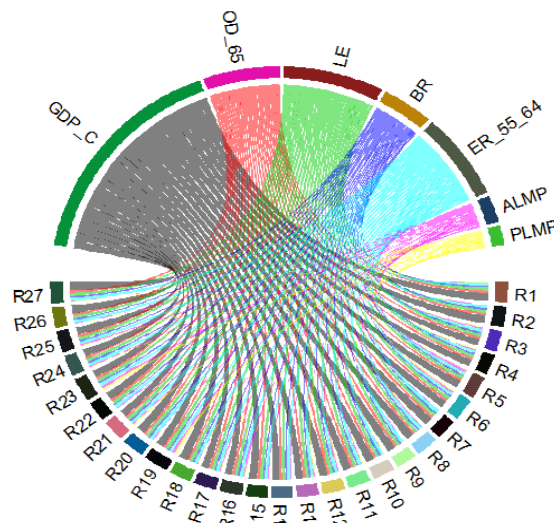
Fig. 1 Chord diagrams of the indicators used in the cluster analysis, EU-27, 2018
Legend: R1-Austria, R2-Belgium, R3-Bulgaria, R4-Croatia, R5-Cyprus, R6-Czech Republic, R7-Denmark, R8-Estonia, R9-Finland, R10-France, R11-Germany, R12-Greece, R13-Hungary, R14-Ireland, R15-Italy, R16-Latvia, R17-Lithuania, R18-Luxembourg, R19-Malta, R20-Netherlands, R21-Poland, R22-Portugal, R23-Romania, R24-Slovak Republic, R25-Slovenia, R26-Spain, R27-Sweden.
Source: own configuration in R version 3.6.3, based on the compiled dataset from Eurostat
(European Commission, 2020a)

Taking a closer look at the data (Fig. 1), at a glance, we could observe that there are important welfare dissimilarities across the EU, particularly as regards the ageing population and labour market indicators.

In terms of the economic development, measured by the GDP per capita (Fig. 2), the gap across the EU-27 countries between the poorest country (Bulgaria) and the richest one (Luxembourg), in 2018, is over 100 thousand USD. We can see that the highest welfare levels were obtained by the old EU countries and the lowest by the new EU countries (that adhered to the EU after 2004).
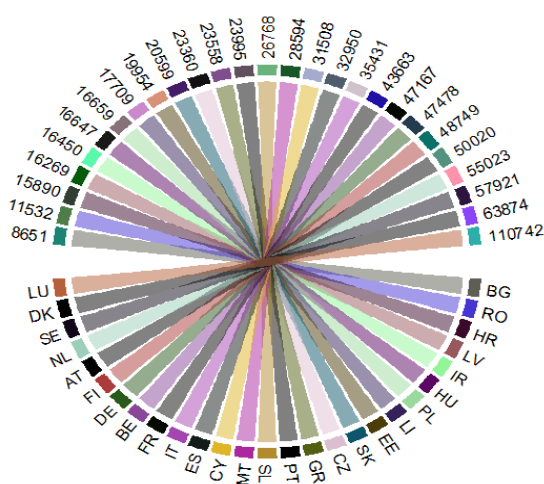
Fig. 2 GDP per capita (constant 2010 USD), EU-27, 2018
Source: own configuration in R version 3.6.3, based on the compiled
dataset from Eurostat (European Commission, 2020a)

At the same time, the old-age-dependency ratio per 100 persons (Fig. 3), namely "the ratio between the number of persons aged 65 and over (age when they are generally economically inactive) and the number of persons aged between 15 and 64" (European Commission, 2020a), is significantly high in Italy, Finland, Greece, Portugal, Germany, approaching 35% in 2018, while in Luxembourg, Ireland this ration is about 20%.

The old dependency ratio in the EU-27 (Fig. 3(a)) shows a share of 21% between people over 65 years old and the working population group, aged 15-64 years, registered in Luxembourg, and 35% in Italy. Increased old dependency ratios within the EU induce important long-run economic consequences, in terms of the savings and investment, housing markets or consumption patterns. The highest life expectancy (Fig. 3(b)) is in Spain (83.5 years), while the lowest is in Bulgaria (75 years), and the birth rate (Fig. 3(c)) is more pronounced in Ireland, Sweden and France (over 11 live births per 1,000 persons), while the less is in Italy Spain and Greece (between 7.3 - 8.1 live births per 1,000 persons).

The long-term problematic issues caused by a decrease in the labour force stock, due to decreasing of the birth rate, and an increase in the age dependency ratio could be alleviated by targeted policies designed to increase productivity and the labour market participation of the elderly.

(a)

(b)

(c)
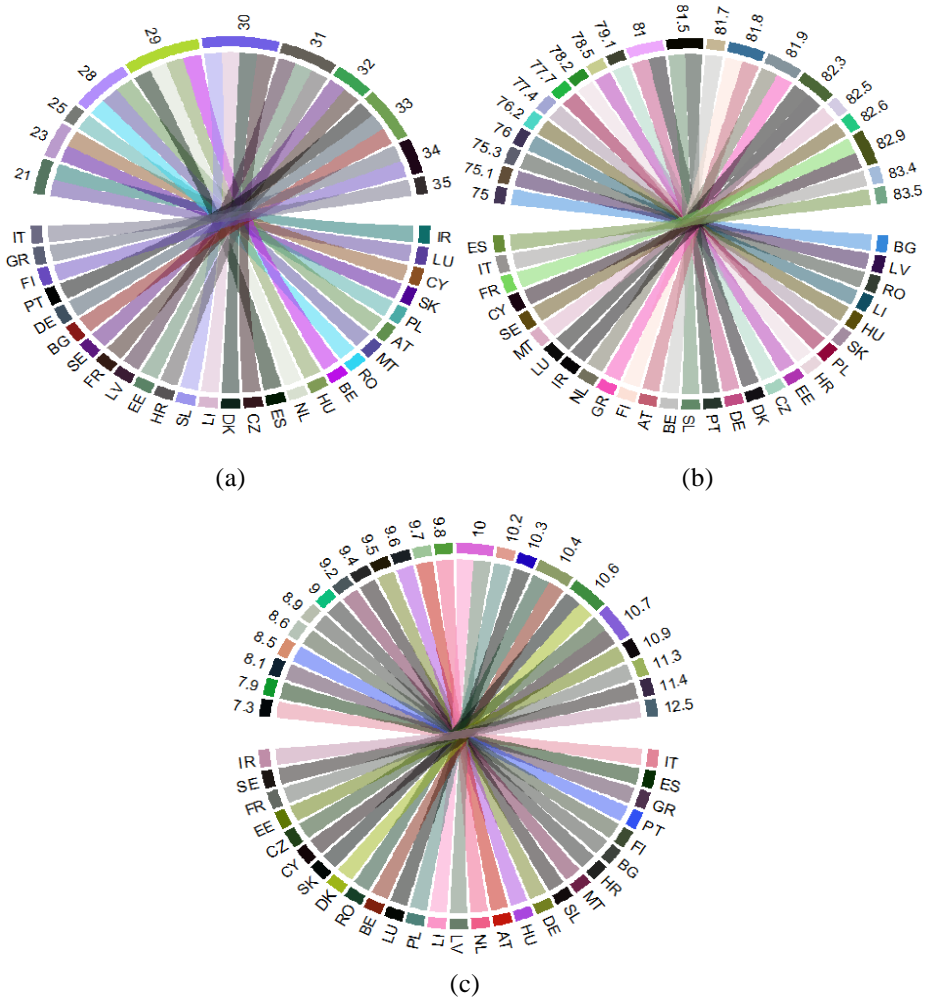
Fig. 3 Old-age-dependency ratio (*OD_65*) (a), Life expectancy (*LE*) (b) and Birth rate
(*BR*) (c), EU-27, 2018

Source: own configuration in R version 3.6.3, based on the compiled dataset
from Eurostat (European Commission, 2020a)

In the same perspective, in 2018, the employment rate of the population aged between
55-64 years (Fig. 4(a)) reached 78% in Sweden, while at the opposite end across the EU
MS we find Greece, Luxembourg, Croatia and Romania, where this rate is only slightly
around 40%. Progresses still therefore need to be made in this respect by most of the EU
countries, particularly from Central and Eastern Europe. Labour market policies allocation in
GDP (active and passive) is at the highest in Denmark and Sweden (Fig. 4(b)), as regards
ALMP, and in the Netherlands, Austria, Finland and Belgium (Fig. 4(c)), regarding PLMP. The

lowest participation of labour market policies in GDP is in Romania and Cyprus (Fig. 4(b)), considering ALMP, and Malta and Romania (Fig. 4(c)), as regards PLMP.
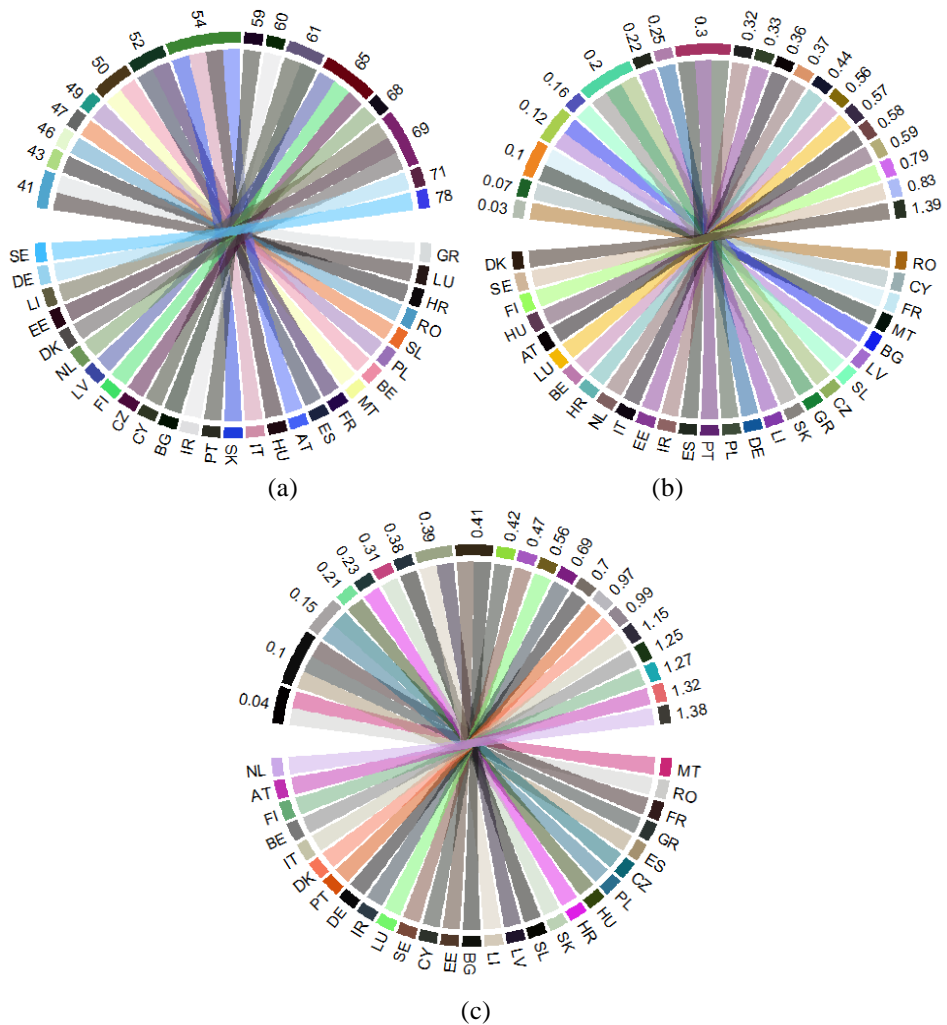
Fig. 4 Employment rate of the population aged from 55 to 64 years (*ER_55_64*) (a), *ALMP* (b), *PLMP* (c), EU-27, 2018

Source: own configuration in R version 3.6.3, based on the compiled dataset from Eurostat (European Commission, 2020a)

We have therefore focused our *research methodology*, based on *cluster analysis*, on these credentials, processed through the Stata program, in order to bring new evidence on the interlinkages between population ageing, labour market indicators and economic welfare dissimilarities across the EU member states, at the level of 2018. The cluster forming and analysis is based on *the Ward method inset on hierarchical clusters* (Härdle

& Simar, 2019), which set out that "the distance between two clusters A and B is shown by how much the sum of squares will increase when they are cumulated" (Cornish, 2007, p. 3). The data was adjusted by standardisation and logarithm in order to ensure a proper comparability between variables.

Accordingly, based on the literature review and own methodological endeavour, we pursued to verify the following research hypothesis (H): "*There are significant dissimilarities between the EU-27 MS in terms of welfare, ageing phenomenon and labour market implications, the developing EU countries confronting with low performances compared to the developed ones*".

4. RESULTS AND DISCUSSIONS

We have verified our hypothesis, *H*, by applying the cluster modelling, processed through the Ward method. Thus, we aim to group the EU-27 countries corresponding to the economic welfare (measured by GDP per capita), the ageing phenomenon dimensions (*OD_65*, *LE* and *BR*) and labour market indicators (*ER_55_64*, *ALMP* and *PLMP*) at the level of 2018.

The correlation matrix of the considered variables and dendrogram of cluster forming for EU-27 MS, at the level of 2018, are presented in Fig. 5(a), respectively Fig. 5(b).

| (a) | (b) |
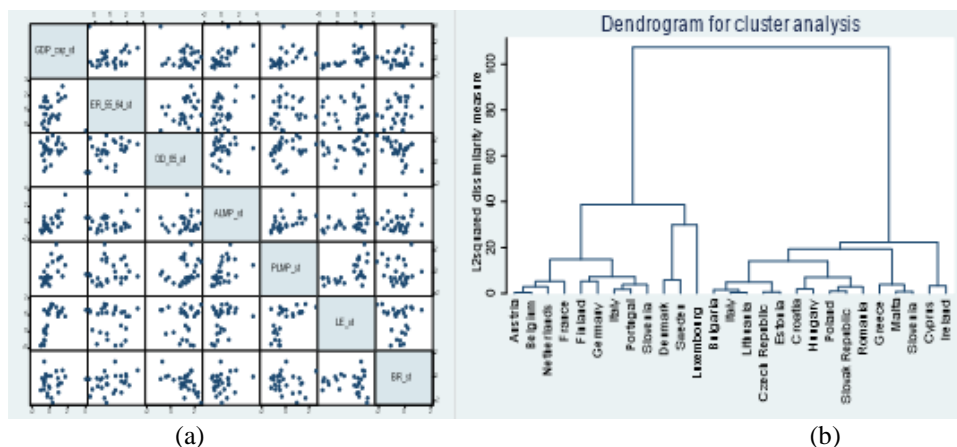
Fig. 5 Correlation matrix (a) and dendrogram of cluster modelling in terms of working ageing (b), EU-27, 2018
Source: own process of panel data in Stata

The performances achieved by each country, as revealed the *GDP_C*, ageing phenomenon dimensions and labour market indicators (Table 1, Appendix, Table A1) have denoted that in few countries from the developed EU MS (grouped into cluster C2), namely Denmark and Sweden, there are very high performances, especially for *ER_55_64* (the mean is 2.2866) and *ALMP* (the mean is 2.4618). These results are similar to those obtained by Thalassinos et al. (2019) (based on AAI), and Suchecka and Urbaniak (2016) (applied for the year 2012).

For the 10 countries of the EU-27 MS (grouped into clusters C1 and C3), there were registered medium performances in terms of well-being, ageing process and labour market

policies, the lowest *OD_65, ER_55_64* (cluster C3) and *BR* (cluster C1), and the highest *GDP_C* (cluster C3) and *PLMP* (cluster C1). More specifically, there were low performances accounted by 15 countries enclosed into C4 (mainly, countries of the new EU-13, aside from Greece, which registered the lowest economic welfare, *GDP_C*, from the EU-27 MS in relation with negative implications of *PLMP, BR* and *LE*, although the *ER_55_64* and *OD_65* was medium), and C5 clusters (despite the highest level of *LE* and *BR)*.

Table 1 Clusters associated with the economic welfare (GDP per capita), ageing phenomenon and labour market indicators, EU-27, 2018

Clusters (C)	EU-27 Member States	Cluster Modelling – Ward Method Performance
C1	Italy, France, Spain, the Netherlands, Finland, Germany, Belgium, Portugal, Austria	*Medium* (particularly in terms of *GDP_C, ER_55_64, ALMP, PLMP* and *LE*, but low in terms of *BR* and high *OD_65*)
C2	Denmark, Sweden	*Very High* (extremely significant in terms of *ER_55_65* and *ALMP*, but also *GDP_C, LE, PLMP* and *BR*)
C3	Luxembourg	*Medium to high* (particularly through high levels of *GDP_C*, *ALMP* and *LE*, and the lowest *OD_65*, despite reduced *ER_55_64*, *BR* and *PLMP*)
C4	Lithuania, the Czech Republic, Hungary, Greece, Bulgaria, Poland, Romania, Slovenia, the Slovak Republic, Latvia, Estonia, Malta, Croatia	*Low* (in terms of the lowest *GDP_C, PLMP, BR* and *LE*, and also reduced *ALMP,* despite medium *ER_55_64*, and *OD_65*)
C5	Ireland, Cyprus	*Low to medium* (particularly through medium levels of *ER_55_64, LE,* despite the highest level of *LE* and *BR*, and low level of *GDP_C, ALMP, PLMP* and *OD_65*)

Source: own process of panel data in Stata

Based on these findings, we can assert that there are significant dissimilarities between the EU-27 MS in terms of economic welfare connected to ageing phenomenon and labour market implications, the developing countries facing low performances compared to the developed ones, our hypothesis, H, being validated.

5. CONCLUSIONS

By applying the cluster modelling, for testing our *research hypothesis*, we have pursued for dissimilarities grouping between the EU-27 MS, thus: the lowest performance related to ageing, economic welfare and labour market was registered for the new EU-13 MS (including also Greece from the old EU-14 countries), and the highest, for the Nordic States (Sweden and Denmark); overall, the EU-13 countries are confronting with low performances compared to the EU-14 MS.

These findings are based on the fact that the Nordic States (Denmark and Sweden) "are well known for their labour market models and best practices at European level (namely the 'Danish model' of flexicurity) focused on balancing flexibility and security for a proper professional and personal development of employees and an adequate labour market insertion and integration (inclusion)" (Noja & Cristea, 2018, p. 735). The Danish labour market policies are justified by

the fact that „each person should, when possible, be active and participate in the social life by working or by attending activation initiatives" (Bazzani, 2017, p. 138). As main policies and strategies as regards ageing and welfare, the good practices of the Nordic countries consist in „activity centres" dedicated for the older people, with entailing them in governance of these centers, making them „empowered to define meaningfulness in activities, but only as an active member of a community" (Evans et al., 2018, p. 1).

Main limitations of the research conducted in this paper may infer the cross-sectional analysis performed at the level of one year that needs to be complemented with a panel-data analysis in order to be further advanced through accurate policies, including by accounting the implications upon the business environment (Pirtea et al., 2014), this type of analysis being targeted as main direction in our future research.

REFERENCES

Bazzani, T. (2017). Italy, Denmark and Germany: A comparative analysis in active and passive labour market policies. *European Labour Law Journal*, *8* (2), 133–153. https://doi.org/10.1177/2031952517712124

Calvo-Sotomayor, I., Laka, J. P., & Aguado, R. (2019). Workforce ageing and labour productivity in Europe. *Sustainability*, *11* (20), 5851. https://doi.org/10.3390/su11205851

Conde-Sala, J. L., Portellano-Ortiz, C., Calvó-Perxas, L., & Garre-Olmo, J. (2017). Quality of life in people aged 65+ in Europe: Associated factors and models of social welfare—analysis of data from the SHARE project (Wave 5). *Quality of Life Research, 26*, 1059–1070. https://doi.org/10.1007/s11136-016-1436-x

Cornish, R. (2007). Statistics: cluster analysis. *Mathematics Learning Support Centre*, 1-5.

Cristea, M., Noja, G. G., Dănăcică, D. E., & Ştefea, P. (2020a). Population ageing, labour productivity and economic welfare in the European Union. *Economic Research - Ekonomska Istraživanja*, *33* (1), 1354-1376. https://doi.org/10.1080/1331677X.2020.1748507

Cristea, M., Noja, G. G., Ponea, S., & Banaduc, I. (2020b). Labor productivity in the complex interplay between health and well-being of older employees: A focus on the New European Union Member States under the Covid-19 pandemic crisis. *Economic and Social Development: Book of Proceedings*, 69-78.

European Commission, (2020a). *Eurostat Database*. Retrieved from: http://ec.europa.eu/eurostat/data/database, Accessed on: 16 September 2020.

European Commission, (2020b). *Employment, Social Affairs & Inclusion. Statistical data.* Retrieved from: https://ec.europa.eu/social/main.jsp?catId=1143&intPageId=3227&langId=en Accessed on: 16 September 2020.

Evans, A. B., Nistrup, A., & Pfister, G. (2018). Active ageing in Denmark; Shifting institutional landscapes and the intersection of national and local priorities. *Journal of aging studies*, 46, 1-9. https://doi.org/10.1016/j.jaging.2018.05.001.

Härdle, W. K., & Simar, L. (2019). *Applied multivariate statistical analysis* 5th edition, Switzerland: Springer, Cham. https://doi.org/10.1007/978-3-030-26006-4_13

Jivraj, S., Goodman, A., Pongiglione, B., & Ploubidis, G. B. (2020). Living longer but not necessarily healthier: The joint progress of health and mortality in the working-age population of England. *Population Studies*, *74* (3), 399-414. https://doi.org/10.1080/00324728.2020.1767297

Kashnitsky, I., & Aburto, J. M. (2020). COVID-19 in unequally ageing European regions. *World development*, *136*, 105170. https://doi.org/10.1016/j.worlddev.2020.105170

Maestas, N., Mullen, K. J., & Powell, D. (2016). *The effect of population aging on economic growth, the labour force and productivity*. National Bureau of Economic Research, Inc., Working Paper Series No. 22452, Retrieved from: https://www.nber.org/papers/w22452 Accessed on: 02 October 2020.

McKibbin, W., & Fernando, R. (2020). The economic impact of COVID-19. In: Baldwin, R. & Weder di Mauro, B. (Eds.), *Economics in the Time of COVID-19* (pp. 45-52). London: Centre for Economic Policy Research (CEPR).

Noja, G. G., & Cristea, M. (2018). Working conditions and flexicurity measures as key drivers of economic growth: Empirical evidence for Europe. *Ekonomicky Casopis (Journal of Economics)*, *66* (7), 719-749.

Pirtea, M., Nicolescu, C., & Botoc, C. (2014). Do Romanian Companies Follow Pecking Order Financing?. *Economic Computation and Economic Cybernetics Studies and Research 48* (1), 1-15.

Rovan, J., & Sambt, J. (2003). Socio-economic differences among Slovenian municipalities: A cluster analysis approach. *Development in Applied Statistics*, *19*, 265-278.

Sidlo, L., Sprocha, B., & Durcek, P. (2020). A retrospective and prospective view of current and future population ageing in the European Union 28 countries. *Moravian Geographical Reports*, *28* (3), 187-207. https://doi.org/10.2478/mgr-2020-0014

Skufina, T., Baranov, S., & Samarina, V. (2020). Modeling and forecasting GDP production in russia, taking into account changes in the number of working-age population caused by the retirement age increasing. In: Solovev D., Savaley V., Bekker A. & Petukhov V. (Eds.), *Proceeding of the International Science and Technology Conference "FarEastCon 2019". Smart Innovation, Systems and Technologies,* 172 (pp. 201-209). Singapore: Springer. https://doi.org/10.1007/978-981-15-2244-4_17

UNECE/European Commission, (2019). 2018 Active Ageing Index: Analytical report. Report prepared by Giovanni Lamura and Andrea Principi, Geneva, June. Retrieved from: https://www.unece.org/fileadmin/ DAM/pau/age/Active_Ageing_Index/Stakeholder_Meeting/ACTIVE_AGEING_INDEX_TRENDS_2008-2016_web_cover_reduced.pdf Accessed on: 02 September 2020.

Thalassinos, E., Cristea, M., & Noja, G. G. (2019). Measuring active ageing within the European Union: Implications on economic development. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 14 (4), 591-609. https://doi.org/10.24136/eq.2019.028

The World Bank, (2020). *World Development Indicators*, Retrieved from: https://databank.worldbank.org/ reports.aspx?source=world-development-indicators Accessed on: 12 September 2020.

Wolf, M. (2020). Why inflation could be on the way back, *Financial Review,* Retrieved from: https://www.afr. com/policy/economy/why-inflation-could-be-on-the-way-back-20201118-p56fle Accessed on: 10 September 2020.

RAZLIKE U STARENJU POPULACIJE I BLAGOSTANJU UNITAR EVROPSKE UNIJE: NOVI PRISTUP ZASNOVAN NA KLASTER ANALIZI

Ovo istraživanje ima za cilj da identifikuje nekoliko razlika među zemljama članicama Evropske unije (EU-27 MS) u pogledu dimenzija blagostanja i tržišta rada pod implikacijama fenomena starenja. Metodologija kvantitativnog istraživanja naglašava klaster analizu zasnovanu na Vard-ovoj metodi za 2018. godinu. Glavni rezultati ukazuju da su samo dve zemlje (Danska i Švedska) registrovale visoke preformanse, naročito za akreditive na tržištu rada (konkretno stopa zaposlenosti i aktivne politike). Ostalih deset EU-27 zemalja su pokazale srednje performanse u smislu blagostanja, ali i najniža postignuća u smislu stope zavisnosti starih, stope zaposlenosti osoba između 55 i 64 godina starosti i nataliteta. Ovaj rad iznosi u prvi plan jaku potrenu da se ponovo osmisle konkretne politike i prilagođene strategije od strane nadležnih institucija i biznis predstavnika širom EU, da bi se naglasila postignuća i nova rešenja za izazove nastale starenjem populacije, sa efektima prelivanja na integraciju starijih zaposlenih u tržište rada i sveukupno ekonomsko blagostanje. Rad se ističe novim integrativnim pristupom baziranim na klaster analizi koja naglašava razlike među zemljama članicama EU, i karektereistike svake grupe zemalja, unutar unapred utvrđenog okvira, ističući tako poteškoće, ali i mogućnosti koje se pružaju u pogledu starenja i ekonomskog blagostanja.

Ključne reči: *tržište rada, ekonomski razvoj, demografsa ekonomija, klaster analiza*

APPENDIX

Table A1 Clusters results for the implications of the aging phenomenon and labour market policies upon the GDP per capita, EU-27, 2018

Indicators	Cluster 1 (C1)			Cluster 2 (C2)			Cluster 3 (C3)			Cluster 4 (C4)			Cluster 5 (C5)			F	R-sq
	N	mean	sd	N	mean	sd	N	mean	sd	N	mean	sd	N	mean	sd		
GDP_C	9	0.6719	0.4892	2	1.5704	0.2080	1	4.0341 *		13	-0.5037	0.2812	2	-0.2496	0.5194	47.24045***	0.8957
ER_55_64	9	1.0724	0.6265	2	2.2866	0.4993	1	-0.3690 *		13	0.7850	0.7891	2	1.2476	0.0283	3.011615**	0.3538
OD_65	9	1.6459	0.6118	2	1.5355	0.2615	1	-0.8447 *		13	1.1569	0.6995	2	-0.4519	0.3595	7.052035**	0.5618
ALMP	9	0.3614	0.5981	2	2.4618	1.3343	1	0.6421 *		13	-0.5035	0.5138	2	-0.6214	0.5957	11.94874***	0.6848
PLMP	9	1.2150	0.6932	2	0.1321	0.7225	1	-0.2020 *		13	-0.7567	0.2967	2	-0.2118	0.3751	20.63566***	0.7896
LE	9	1.2770	0.2590	2	1.1691	0.3328	1	1.3162 *		13	0.0601	0.7672	2	1.4045	0.1248	7.2119**	0.5673
BR	9	-0.8760	0.8314	2	0.2973	0.3733	1	-0.1647 *		13	-0.5150	0.5438	2	0.6932	0.8399	3.040925**	0.3560

Source: own process of panel data in Stata

Original Scientific Paper

CREDIT SCORING WITH AN ENSEMBLE DEEP LEARNING CLASSIFICATION METHODS – COMPARISON WITH TRADITIONAL METHODS

UDC 336.77:519.2

Ognjen Radović, Srđan Marinković, Jelena Radojičić

University of Niš, Faculty of Economics, Serbia

Abstract. *Credit scoring attracts special attention of financial institutions. In recent years, deep learning methods have been particularly interesting. In this paper, we compare the performance of ensemble deep learning methods based on decision trees with the best traditional method, logistic regression, and the machine learning method benchmark, support vector machines. Each method tests several different algorithms. We use different performance indicators. The research focuses on standard datasets relevant for this type of classification, the Australian and German datasets. The best method, according to the MCC indicator, proves to be the ensemble method with boosted decision trees. Also, on average, ensemble methods prove to be more successful than SVM.*

Key words: *credit scoring; classifier ensemble, deep learning, support vector machine*

JEL Classification: C38, C45, C55, G17, G24

1. INTRODUCTION

Credit scoring is a quantitative method for assessing the credit risk involved in granting a loan to a borrower. Credit scoring is one of the key stages in credit analysis. It is a method of assessing the creditworthiness of a client applying for a loan. The goal of creditworthiness assessment is to classify credit applications into acceptable and unacceptable, but also to provide the necessary inputs for the next phases of credit analysis, such as determining the credit volume, interest rates, collateral, restrictive clauses and the like. Traditional credit analysis relies on historical and verifiable information or accounting data. The most general credit analysis framework many traditional lenders use to assess credit users is the 5C approach. Credit analysts make decisions based on the following criteria: debtors' character, capacity (ability to repay), capital, collateral and market conditions. The main advantage of this method is that it can be

used to make credit decisions in various types of business and consumer loans without significant adjustments. Traditionally, creditworthiness assessment of a corporate borrower is based on financial indicators that indicate possible problems in loan repayment and play an important role in calculating credit risk levels. These indicators are characterized by objectivity because they are calculated on the basis of data in borrowers' financial statements. The analyst's experience and impressions can play a significant role in creditworthiness assessment of the entity applying for a loan. Prior to the credit scoring model, the decision to grant a loan was made based on the credit analyst's assessment. One of the disadvantages of such an approach is the inability to process a large number of applications per day, which has given rise to various credit scoring models to quantify the credit risk (Dastile, Celik, & Potsane, 2020). Banks used information received within a loan application (e.g., number of dependents, period in the current job, etc.) to calculate the borrowers' numerical score (Lewis, 1992). Credit scoring, as a precise and automatic creditworthiness assessment tool, is a particularly important factor in the expansion of consumer lending (Thomas, Crook, & Edelman, 2002). Credit scoring is mostly applied in consumer loans, credit cards and mortgage loans (Einav, Jenkins, & Levin, 2013). Advances in information technology facilitate further development of credit scoring models for making objective and quick decisions (Thomas, Crook, & Edelman, 2002). Credit scoring is a technique that financial organizations use when making a decision to approve a loan or reject their clients' loan application (application scoring). Credit scoring models can be applied to analyze the behavior of already existing clients, and then the score represents a numerical summary of the bank's experience with the client (behavior scoring) (Hui, Li, & Zongfang, 2017). The Basel II agreement of 2004 expanded the field for developing more sophisticated credit scoring models, thus allowing banks to assess the probability of default on their own under the internal ratings-based (IRB) approach (Goh & Lee, 2019).

2. TRADITIONAL STATISTICAL METHODS USED IN CREDIT SCORING

Credit scoring is a multi-stage process (Bequé & Lessmann, 2017), and it basically compares the borrower's characteristics and the characteristics of other clients from the previous period. The statistical model relies on historical data. Its goal is to predict future behavior in loan repayment based on previous experience with loan users with similar characteristics. If the borrower is similar to "bad" clients (who did not repay the approved loan properly), the application is rejected, and if the borrower is similar to "good" clients (who returned the approved loan properly), the loan application is approved. The loan applicant's score is compared with the established cut-off score. If the obtained score is higher than the cut-off score, credit is approved, and if the score is lower than the cut-off, application is rejected. The cut-off score is crucial for the usefulness of the credit scoring model and mainly relies on the credit decision makers' attitudes towards risk. So, there is no optimal cut-off value. "It varies from one environment to another and from one bank to another inside the same country" (Abdou & Pointon, 2011). The final step is to measure the accuracy of the credit scoring model and monitor business performance indicators. "The choice of a statistical model is crucial because it affects all subsequent activities and credit scoring performance" (Bequé & Lessmann, 2017).

The models that financial institutions use help them decide whether or not to grant a loan. As the final decision on approval is binary, there is a problem of binary classification. Credit scoring involves "formal statistical methods to classify loan applicants into "good"

and "bad" risk classes" (Hand & Henley, 1997). Credit scoring algorithms are basically statistical in nature: they use empirical evidence to formulate predictions about the future. Prediction models assess a continuous variable while classification models predict class membership. In credit scoring, the dependent variable is actually binary, so most algorithms can be considered classification algorithms (Abdou & Pointon, 2011).

Statistical methods such as linear regression and discriminant analysis require the assumption of a linear relationship between variables. In credit scoring, linear regression is used as a binary classification problem. Discriminant analysis is a simple parametric statistical technique for classifying loans into good and bad. Fischer (1936) suggests the application of discriminant analysis as a classification technique, and in 1941 Durand used discriminant analysis to classify "good" and "bad" car loans, thus beginning the trend of applying statistical models in credit scoring. Later, Altman (1968) developed a Z-score model with financial indicators, using variables from corporate financial statements as input variables for a discriminant analysis model to predict company bankruptcies. "Discriminant analysis is one of the frequently used techniques in credit scoring" (Abdou & Pointon, 2011). Ogrler (1971) applies regression analysis for credit scoring of consumer loans in banks, after using this method somewhat earlier to evaluate the already existing commercial loans (Orgler, 1971). He concludes that "information not included in the loan application form has a greater predictive ability" to assess loan quality in the future than the information included in the form.

In 1980, Ohlson proposed the use of logistic regression (LOGIT) as a creditworthiness assessment method in companies (Ohlson, 1980). The outcome variable in logistic regression is dichotomous (outcome 0/1), so this method is suitable for modeling binary outcomes and is widely used in creditworthiness assessment due to its simplicity and transparency (Dastile, Celik, & Potsane, 2020; Abdou & Pointon, 2011). Within traditional methods, logistic regression has become the standard credit scoring model due to its compliance with the Basel II standard (Goh & Lee, 2019).

3. APPLICATION OF MACHINE LEARNING IN CREDIT SCORING

More sophisticated methods of credit scoring that literature has offered in recent years are machine learning algorithms and data mining methods. The point of using sophisticated techniques is their ability to model extremely complex functions (Abdou & Pointon, 2011). Machine learning models learn from available data, thus allowing the calculation of predictive value. A machine learning algorithm learns a set of rules based on the information available in a training set of examples. Machine learning models have the potential to replace the logistic regression model in credit scoring because they show great prediction accuracy. However, the impossibility of certain models to explain the predictions, i.e. lack of transparency, limits their application in regulated financial institutions. The application of machine learning algorithms may include (Dastile, Celik, & Potsane, 2020): „k-*Nearest Neighbor* (k-NN), *Decision Trees* (DTs), *Support Vector Machines* (SVMs), *Artificial Neural Networks* (ANNs), *Random Forests* (RFs), *Boosting*, *Extreme Gradient Boost* (XGBoost), *Bagging*, *Restricted Boltzmann Machines* (RBMs), *Deep Multi-Layer Perceptron*, (DMLP), *Convolutional Neural Networks* (CNNs) and *Deep Belief Neural Networks* (DBNs)".

The Decision Tree creates a discriminant function in the form of a tree, from the root to the leaves. Each node represents a logical test of the attribute value, and the leaves denote

classes. Input observations are recursively split into subbranches, i.e. until the final result (class designation). Mathematical formulas such as the Gini index (CART) or entropy (in ID3, C4.5, C5, J4.8 decision tree algorithms) are used to determine the splitting threshold (Patil, Aghav, & Sareen, 2016; Bequé & Lessmann, 2017). The tree learns by asking questions that will solve the problem in the fastest and most accurate way. Each time the algorithm is repeated, the attribute value is compared to the threshold. Thresholds determine which attribute should be tested and when tree growth should be stopped (Bequé & Lessmann, 2017). After the training, the tree can predict the outcome if applied to data of the same type and format. The low decision tree accuracy may be affected by low depth and presence of noise (overfitting). This classification technique is widely used in credit scoring models (Dastile, Celik, & Potsane, 2020).

Since its introduction into the theory of statistical learning in 1998 (Vapnik, 1998), the support vector machines method (Support Vector Machines – SVM) has been used as a binary classifier of machine learning. SVM algorithms in the field of credit scoring were introduced by Baesens et al. (2003) and through comparison with other classification algorithms (logistic regression, discriminant analysis, k-nearest neighbor, neural networks, decision tree) indicated its good performance. SVM, as a discriminant model, is based on the margin for classification between classes, i.e. it focuses on finding the boundary that separates two classes with the smallest error. In the SVM model, data is viewed as vectors in n-dimensional space. The focus is on the maximum margin between classes, i.e. the space between two hyperplanes that separate data from different classes, i.e. which belong to one or more support vectors. The separating hyperplane is located farthest from the data and is determined by the position of the data of both classes closest to it. If the classes are denoted by y=+1 and y=-1, in the version of the linear SVM classifier, the margin $\|\alpha\| = \sqrt{\sum_{i=1}^{m} \alpha_i^2}$ between the negative and positive hyperplanes is maximized. The following equation is used to assign a class (Dastile, Celik, & Potsane, 2020):

$$y = \begin{cases} +1, if\ b + \alpha^T x \geq +1 \\ -1, if\ b + \alpha^T x \leq -1 \end{cases} \qquad (1)$$

where *b* is bias.

The SVM method can also be used for data classes that are not linearly separable. "For nonlinear classification, a kernel trick is used to modify the SVM formulation" (Dastile, Celik, & Potsane, 2020). Using the appropriate kernel function, the example is mapped to a space with a larger number of dimensions and the nonlinear problem is transformed into a linear one.

Neural network consists of several neurons that work in parallel, without centralized control. An artificial neural network mimics the way a biological neural network processes information. Neurons are usually complex in layers. The neural network usually consists of three layers: input, hidden, and output parameters. There are several types of neural networks, but their common components are a set of nodes and connections between nodes. Nodes represent computer units whose task is to receive inputs, process inputs and produce and output (Bequé & Lessmann, 2017). First, the input characteristics are processed to the hidden parameters, and then the hidden parameters calculate the adequate weight before forwarding the information to the output parameters. Each layer of the neural network "consists of several elements, i.e. neurons. The number of input neurons depends on the number of predictors, the number of hidden neurons is a setting parameter determined by the analyst, and the number of output neurons is determined by the modeling task itself,

e.g. for binary classification it is one" (Bequé & Lessmann, 2017). Artificial neural networks were first applied by Odom and Sharda (1990) in credit scoring.

For a given vector of the input attribute x, the three-layer neural network calculates the output value $\hat{y}$ as follows (Dastile, Celik, & Potsane, 2020):

$$\hat{y} = a_2\big(a_1\left(\alpha^{(1)}x + \alpha_0^{(1)}\right)\alpha^{(2)}x + \alpha_0^{(2)}\big) \tag{2}$$

where $(\alpha_0^{(1)}, \alpha^{(1)})$, $(\alpha_0^{(2)}, \alpha^{(2)})$ are weights, and a_2 and a_2 are activation functions between the input and hidden layer.

Neural networks are trained through a training set, and the final decision is made by applying the decision function to $\hat{y}$.

4. Ensemble of Classifiers – Literature Review

Elementary classifiers represent a unique set of statistical relationships. Ensemble of classifiers consist of a set of individually trained base classifiers whose decisions are combined in a certain way (weighted voting or unweighted voting) when new examples are classified. Ensemble of classifiers is a "combination of classifiers so that their fusion achieves better performance than stand-alone classifiers" (Nanni & Lumini, 2009). Combining different machine learning algorithms can improve the accuracy of results (Dastile, Celik, & Potsane, 2020). Ensemble algorithm techniques are used to aggregate the results of "unstable" algorithms in which small changes in the training set lead to large changes in the learned set of rules (e.g. neural networks, decision trees). The application of ensemble learning techniques requires the simultaneous fulfillment of the following assumptions (Pławiak, Abdar, & Acharya, 2019): "a) quality, b) statistical independence (diversity) and c) efficiency (speed)".

Ensemble of classifiers is used to achieve better performance in various research areas such as computer intelligence, statistics, and machine learning (Ren, Zhang, & Suganthan, 2016). Different ensembles of classifiers are used in literature (Pławiak, Abdar, & Acharya, 2019): a) Boosting (AdaBoost), b) Bagging (Bootstrap aggregation), c) Random Forest, d) Stacking (Stacked Generalization) and e) Mixtures of Experts.

Boosting is the most commonly used method in the ensemble of classifiers (Dastile, Celik, & Potsane, 2020). It starts with a weak model (for example, a shallow decision tree), and then the models are iteratively evaluated and amplified (Freund & Schapire, 1997). Boosting produces a series of classifiers (Bequé & Lessmann, 2017). Each subsequent classifier focuses on examples that have been misclassified by the previous classifier. Each example in the training set is assigned a weight in accordance with the significance of the example in the set. Examples misclassified by the previous model are assigned a higher weight. After individual classifier learning, the weights are updated on the test set. The accuracy of individual classifiers on a test set is determined by the weight of that classifier in the classification of new examples by applying an ensemble of classifiers.

One of the most well-known boosting techniques is Adaptive Boosting (AdaBoost) (Freund & Schapire, 1997). AdaBoost algorithm has one parameter T – the number of generated classifiers (iteration) and is characterized by simplicity and efficiency. AdaBoost assigns a class to an input attribute vector that is classified (x) as follows:

$$\hat{y} = sign\left(\sum_{t=1}^{T} \alpha_t \phi_t\left(x\right)\right) \tag{3}$$

where α_t is the weight of the classifier $\phi_t(x)$.

Bagging generates multiple versions of classifiers that are used as an aggregate predictor through a voting mechanism (Breiman, 1996). Bootstrap Aggregation is used to generate classifiers, with no iterative division into a training set and a test set, but random selection with return. The training set is formed by successive sampling (with repetition) of data from the initial set. Data never selected form a test set, while the rest is used for training. The process is repeated several times, and the overall score is obtained as the average score on all thus formed sets for verification (Breiman, 1996). Bootstrapping generates K training sets, and then one basic classifier is trained on each of them. The class rating is awarded by a majority vote of K classifiers, as follows (Dastile, Celik, & Potsane, 2020):

$$y = \underset{y \in \{+1, -1\}}{argmax} \sum_{i=1}^{K} 1(y = \phi_i(x)) \tag{4}$$

where

$$1(y = \phi_i(x)) = \begin{cases} 1, & if \ y = \phi_i(x); \\ 0, & if \ y \neq \phi_i(x). \end{cases} \tag{5}$$

The Random Forest algorithm consists of several decision trees. The new examples are classified by the voting method based on the decisions of individual trees. Not all samples and attributes are taken for training individual trees, but a certain number of randomly selected attributes and samples from the training dataset. Each decision tree develops on a subset of randomly selected attributes. The best attribute is chosen for the decision tree node. Selecting the right attributes (questions) to be tested in a particular node reduces the entropy, or provides additional information about the sample. In a random forest algorithm, multiple decision trees learn from randomly selected data leading to greater tree diversity and depth. The greater depth of the trees makes the random forest algorithm more resistant to underfitting (insufficiently good interpretation of the relationships between variables within the dataset) and overfitting (noise in the data) compared to individual decision trees.

Deep Learning, as one of the machine learning fields, is based on a hierarchical architecture that includes multiple layers of nonlinear operations and steps in information processing. Some of the deep machine learning techniques used in credit scoring are (Pławiak, Abdar, & Acharya, 2019): „(a) deep discriminant models such as: Deep Neural Networks (DNNs), Recurrent Neural Networks (RNNs), as well as Convolutional Neural Networks (CNNs), (b) unsupervised learning (generative models) such as: Restricted Boltzmann Machines (RBMs), Deep Belief Networks (DBNs), Deep Boltzmann Machines (DBMs) as well as Regularized Autoencoders." Deep learning classifiers are not widely used in credit scoring (Dastile, Celik, & Potsane, 2020). Disadvantages of deep learning are (Pławiak, Abdar, & Acharya, 2019): "(1) computationally complex training, (2) long and inefficient training and (3) the overfitting effect, which prevents its effective practical use".

Predictive accuracy is a basic measure of classification success. Prediction accuracy is the percentage of success in classifying new examples using learned rules. Even small errors in creditworthiness assessment can lead to large losses, so increasing the accuracy of credit scoring is of great importance for the profitability of banks (Pławiak, Abdar, & Acharya, 2019). In this regard, more sophisticated credit scoring models have significant potential (Dastile, Celik, & Potsane, 2020). In one of the earliest reviews of statistical methods and data mining methods applied in credit scoring, Hand and Henley (1997) conclude that further development should go towards more complex DM models, which later literature reviews confirmed (Sadatrasoul, et. al. 2013). A comparison of different approaches to credit scoring shows that advanced machine learning-based techniques may

have better predictive ability than conventional techniques such as logistic regression and discriminant analysis (Abdou & Pointon, 2011). Nanni and Lumini (2009) investigate several ensemble of classifier systems used in credit scoring and bankruptcy prediction and improve the performance obtained using the stand-alone classifiers. Wang et al. (2011) carry out a comparative performance evaluation of "three ensemble methods (Bagging, Boosting, Stacking) of basic classifiers: logistic regression, decision tree, artificial neural network, and support vector machine (SVM)". The results show that ensemble method improves performance and that bagging gives better results than boosting. Lou et al. (2017) compare classification success of deep learning algorithms in credit scoring and widely used models such as logistic regression and SVM, to find that deep learning models have better performance. Li et al. (2017) develop a "model for credit risk assessment using deep neural networks" and show that the proposed algorithm has greater accuracy in credit risk assessment. The results of the simulation by Zhou et al. (2012) show that Extreme learning machines (ELM) is a more suitable approach for credit risk assessment than SVM. Bequé and Lessmann (2017) investigate the potential application of ELM for credit scoring and compare it with other classifiers (neural networks, k-nearest neighbor, SVM, classification and regression decision tree, logistic regression) within three dimensions: "ease of use, computational complexity and predictive performance". They conclude that "ELM shows competitive or better results in each dimension of comparison and especially proves high discriminant power, both in isolation and within the ensemble and, therefore, represents a competitive alternative to already established classifiers in the field of credit scoring". Neagoe et al. (2018) design a credit scoring model using a neural network classifier, namely: The Multilayer Perceptron (MLP) approach and the thirteen-layer DCNN variant. The obtained results confirm the efficiency of the proposed approach, indicating a significant advantage of DCNN over MLP. Proceeding from the idea to imitate the work of the human brain in terms of fusion and information flow, Plawiak et al. (2019) create the "Deep genetic cascade ensembles of classifiers (DGCEC) based on the fusion of stratified 10-fold CV method, ensemble learning, deep learning, layered learning and supervised training. The applied model combines three machine learning techniques: evolutionary, ensemble and deep learning." The solution the authors propose provides a fast and efficient approach to training, which increases the accuracy of creditworthiness assessment. In the DGCEC method, each first-layer classifier is trained to increase the recognition performance of accepted or rejected borrowers based on the pre-processed data on borrowers. In other layers, based on the pre-processed user data and the classifier response from the first and previous layers using deep learning techniques and selection of genetic characteristics, a knowledge extraction process takes place that leads to the final result. The results show better performance of this approach compared to previously applied approaches in terms of the accuracy of creditworthiness assessment of borrowers in Australia. The highest accuracy of creditworthiness assessment in previously conducted studies is 91.97%, while the method proposed by the authors allows higher prediction accuracy, i.e. of 97.39%. (Pławiak, Abdar, & Acharya, 2019). Recent literature research shows that ensemble models have better performance than individual classifiers and that deep learning models give better results compared to statistical and traditional machine learning models (Dastile, Celik, & Potsane, 2020).

5. RESEARCH METHODOLOGY

Empirical research includes checking the performance of deep learning algorithms over known credit scoring datasets. In this paper, we use two datasets, the Australian Credit and the German Credit (UCI Machine Learning repository, Asuncion & Newman, 2010).

Datasets include a different number of independent variables. Australian credit data consists of 307 cases of creditworthy candidates and 383 cases of candidates to whom credit should not be granted. The German dataset is somewhat more asymmetric, with many more creditworthy examples (700) than those that should not be granted credit (300). The Australian dataset has 14 attributes, while the German has 24 attributes. Both sets have two classes {approved, rejected} and are a good mix of different types of attributes: continuous and nominal. Variables can be grouped into several categories (Beque and Lessmann, 2017): financial (assets, monthly income, etc.), socio-demographic (age, place of residence, etc.), others (possession of a credit card or a mobile phone).

K-fold cross-validation is used to assess the classification model, which works as follows (Dietterich, 1998):

1. Dividing a training dataset into k randomly selected non-overlapping data subsets of approximately equal size;
2. One subset is used to validate the model of trained over the remaining data subsets;
3. This procedure is repeated k-times so that each subset is used exactly once for model validation;
4. Performance is assessed for each partition and the average error on all k-partitions is reported.

This is one of the most popular techniques for cross-validation and is good at assessing the predictive accuracy of the classification model. In our study, 5-fold cross-validation is used for both credit rating datasets. Also, in order to reduce the dimensionality of the predictor space, principal component analysis (PCA) is used. PCA linearly transforms predictors to remove redundant dimensions and prevent overfitting.

To assess the performance of classification models, several standard indicators are used, for the calculation of which the values of correct and incorrect predictions are used: the number of borrowers correctly classified as {approved} (defaults) (True Positives -TP), the number of borrowers incorrectly classified as {approved (defaults) (False Positives - FP), the number of borrowers correctly classified as {rejected} (non-defaults) (True Negatives - TN), and the number of borrowers incorrectly classified as {rejected} (non-defaults) (False Negatives - FN). The total number of examples is N=TP+FP+FN+TN. The false positive rate (FP) is defined as the share of misclassified loan approval cases. In contrast, the rate of false negative results (FN) is defined as the share of misclassified cases refused to be given a loan (qualifying for a loan).

The set of indicators consists of: PCC (Percentage Correctly Classified), AUC (area under the curve), sensitivity, specificity, precision, G-mean, F-measure and Matthews correlation coefficient - MCC. The selection of indicators is based on previous research (Oztekin, Al-Ebbini, Sevkli, & Delen, 2018; Kim et al, 2020).

Percentage Correctly Classified (PCC) is the ratio of correct predictions of a case classification model in two categories {approved, rejected}. It is calculated as PCC = (TP+TN)/(TP+TN+FP+FN). PCC is the average percentage of correctly classified cases and is a measure of correct classification over sets unused for learning (subset for validation in 5-fold cross-validation). Sensitivity/Recall is the ratio of correctly classified cases in the class {approved} to the total number of examples in the class {approved} and is calculated

as SEN = TP/(TP+FN). Specificity is the ratio of correctly classified cases in the class {rejected} to the total number of examples in the class {rejected} and is obtained as TN/(TN+FP). Sensitivity and specificity show the accuracy of class-level classifiers.

Geometric-Mean (G-mean) is obtained as follows:

$$G\text{-}mean = \sqrt{\frac{TN}{(FP + TN)} \, x \, \frac{TP}{(TP + FN)}} \qquad (6)$$

F-measure is calculated as:

$$F\text{-}measure = \frac{2 \times Sensitivity \times Precision}{Precision + Sensitivity} \qquad (7)$$

where Precision = TP/(TP+FP). G-mean and F-measure indicate imbalance between classes.

AUC or Area Under the Receiver Operating Characteristic Curve is one of the indicators that illustrates the performance of the classification model. The larger this area, the better the model. AUC can be seen as the ability to distinguish positive from negative classification.

Matthews correlation coefficient (MCC) is another performance indicator. The MCC value is -1 to 1. Perfect prediction has a value of 1, completely incorrect prediction is -1, while random prediction has a value of 0. The MCC generates a high score only if the model can correctly predict most positive credits and most correctly rejected credits. It is considered one of the best indicators of accuracy in the evaluation of machine learning algorithms. The formula for calculating the MCC is as follows (Matthews, 1975; Jurman et al, 2012):

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP) \cdot (TP+FN) \cdot (TN+FP) \cdot (TN+FN)}} \qquad (8)$$

The following classifiers are used to compare the performance of ensemble models: Logistic regression and Support vectors. The SVM models used to compare classification performance use different learning algorithms (Linear, Quadratic, Cubic, Fine Gaussian, Medium Gaussian, Coarse Gaussian). The tested ensemble models use three learning algorithms (Boosted Trees (AdaBoost), Bagged Trees and RUSBoosted Trees) over decision trees (Random Forest techniques). The research is conducted in the Classification Learner application module using the Matlab 2019b software package. All calculations are performed in a Windows 10 environment (AMD Ryzen 7 3700U with 12GB RAM).

6. RESULTS AND DISCUSSION

Tables 1 and 2 show the performance of predictive models for the Australian and German datasets, respectively. The tables show assessment of categorical prediction accuracy and the discriminant ability of the included classifiers. Performance assessment is average accuracy on test sets. For each model, AUC (Area Under the Receiver Operating Characteristic Curve), True Positive Classification Rate (TP), False Positive Classification Rate (FP), True Negative Classification Rate (TN), False Negative Classification Rate (FN), and PCC are shown (Percent Correctly Classified). These results are the average values determined for each of the 5 independent non-overlapping partitions of the dataset used in the 5-fold cross-validation.

For both datasets, the approved case rate is higher than the rejection rate for all models tested, meaning that false negative results are less common than false positive results relative to the overall prediction (with the exception of the Fine Gaussian SVM). It is

debatable whether false negative predictions are more serious than false positive ones from the point of view of credit risk. If someone is predicted to be able to repay the loan and it turns out that they are not able to, that can lead to certain losses. Conversely, if someone is rejected, and they are able to repay the loan, it leads to a loss of earnings on the loan.

At the Australian dataset, according to the AUC criterion, the best classifier is the Ensemble Bagged Trees model (Table 1) with a value of 0.92 (87.41%). The Gaussian SVM model has the lowest false positive case rate (FP), while the linear Gaussian SVM models have the lowest false negative case rate (FN). According to the criterion of false positive cases (FP), ensemble models show the best results. This means that these models prove to be less risky in terms of incorrectly granted credit. According to the PCC criteria, the Ensemble Bagged Trees model proves to be the best classifier (Table 1). However, some SVM methods (Gaussian SVM) show better results than other ensemble methods. However, in general, ensemble methods show better results than SVM methods. On average, according to the PCC indicator, ensemble models are better than other tested models. Our results confirm the Beque and Lessmann (2017), but it should be noted that our research relies on the Matlab package while Beque and Lessmann (2017) use the R programming environment. Our research on the Australian dataset shows that logistic regression is inferior to SVM and ensemble methods.

Table 1 Summary of results of individual classifiers obtained in predicting the Australian credit scoring dataset

Classification Learner	AUC	TP	FP	TN	FN	PCC
Logistic regression						
Logistic regression	0.90	266	69	314	41	84.06%
	(0.13,0.82)	(87%)	(18%)	(82%)	(13%)	
Support Vector Machines (Box constraint level, Manual kernel scale)						
Linear	0.91	**285**	78	305	**22**	85.51%
(1,-)	(0.07,0.80)	**(93%)**	(20%)	(80%)	**(7%)**	
Quadratic	0.90	257	62	321	50	83.77%
(1,-)	(0.16,0.84)	(84%)	(16%)	(84%)	(16%)	
Cubic	0.90	253	60	323	54	83.48%
(1,-)	(0.18,0.84)	(82%)	(16%)	(84%)	(18%)	
Fine Gaussian	0.89	233	**43**	**340**	74	83.04%
(1,0.94)	(0.24,0.89)	(76%)	**(11%)**	**(89%)**	(24%)	
Medium Gaussian	0.92	**285**	77	306	**22**	85.65%
(1,3.7)	(0.07,0.80)	**(93%)**	(20%)	(80%)	**(7%)**	
Coarse Gaussian	0.92	**285**	77	306	**22**	85.65%
(1,15)	(0.07,0.80)	**(93%)**	(20%)	(80%)	**(7%)**	
Ensemble (Method, Maximum number of splits, Number of learners, Learning rate)						
Boosted Trees	0.92	246	49	334	61	84.06%
(AdaBoost, 20,30,0.1))	(0.20,0.87)	(80%)	(13%)	(87%)	(20%)	
Bagged Trees	**0.92**	269	49	334	38	87.39%
(Bag, 689,30,-)	**(0.12,0.87)**	(88%)	(13%)	(87%)	(12%)	
RUSBoosted Trees	0.91	259	52	331	48	85.51%
(RUSBoost, 20,30,0.1)	(0.16,0.86)	(84%)	(14%)	(86%)	(16%)	

Source: Data processed by the author

In the German dataset, according to the AUC criterion, the best classifier is the ensemble RUSBoosted Trees with a value of 0.76 (70.38%), and logistic regression, linear and medium Gaussian SVM are close to it with 0.77 (about 65%) (Table 2). The RUSBoosted Trees model has the lowest rate of false positive cases (FP), while the Gaussian SVM models have the lowest rate of false negative cases (FN). According to the criterion of false positive cases (FP), ensemble models give the best results. According to the PCC criteria, the Medium Gaussian SVM model (Table 2) proves to be the best classifier with accuracy of 74.10%. On average, according to the PCC indicator, ensemble models are equal to SVM models. In the German dataset, logistic regression yields results on a par with SVM and ensemble methods.

Table 2 Summary of results of individual classifiers obtained in predicting the German credit scoring dataset

Classification Learner	AUC	TP	FP	TN	FN	PCC
Logistic regression						
Logistic regression	**0.77**	602	163	137	98	73.90%
	(0.54,0.86)	(86.0%)	(54.3%)	(45.7%)	(14.0%)	
Support Vector Machines (Box constraint level, Manual kernel scale)						
Linear	**0.77**	616	179	121	84	73.70%
(1,-)	**(0.60,088)**	(88.0%)	(59.7%)	(40.3%)	(12.0%)	
Quadratic	0.75	592	171	108	129	70.00%
(1,-)	(0.57,0.85)	(84.6%)	(57.0%)	(43.0%)	(15.4%)	
Cubic	0.70	547	158	142	153	68.90%
(1,-)	(0.53,0.78)	(78.1%)	(52.7%)	(47.3%)	(21.9%)	
Fine Gaussian	0.71	699	296	4	**1**	70.30%
(1,0.94)	(0.99,1.0)	(99.9%)	(98.7)	(1.3%)	**(0.1%)**	
Medium Gaussian	**0.77**	**645**	204	96	55	74.10%
(1,3.7)	**(0.68,0.92)**	**(92.1%)**	(68.0%)	(32.0%)	(7.9%)	
Ensemble (Method, Maximum number of splits, Number of learners, Learning rate)						
Boosted Trees	0.76	608	182	118	92	72.60%
(AdaBoost, 20,30,0.1))	(0.61,0.87)	(86.9%)	(60.7%)	(39.3%)	(13.1%)	
Bagged Trees	0.75	605	173	127	95	73.20%
(Bag, 689,30,-)	(0.58,0.86)	(86.4%)	(57.7%)	(42.3%)	(13.6%)	
RUSBoosted Trees	0.76	458	**74**	**226**	242	68.40%
(RUSBoost, 20,30,0.1)	(0.25,0.65)	(65.4%)	**(24.7%)**	**(75.3%)**	(34.6%)	

Source: Data processed by the author

Tables 3 and 4 show the performance indicators of different classifiers for Australian and German datasets, respectively. The models with the best performance indicators are in bold. All classifiers in both tested datasets record a high value of sensitivity, with large differences observed in terms of specificity indicator in German dataset. In both datasets, the Fine Gaussian SVM has the highest sensitivity value. In the Australian dataset, SVM algorithms have the highest specificity. However, the best ensemble in the German dataset is the Ensemble RUSBoosted Trees. The sensitivity of SVM methods is on average higher than ensemble methods in both datasets. However, the specificity of ensemble methods is on average higher than SVM in both datasets. Similarly, in terms of G-average and F-measure, which show a balance between sensitivity and specificity, ensemble methods, on average, show slightly better results than SVM in terms of G-average but lower in F-measure. In terms of MCC indicators, as the most relevant for the assessment of binary

classification techniques of machine learning, in both datasets, the best results are recorded with ensemble techniques. In the Australian dataset, the MCC correlation of 74.60% Ensemble Bagged Trees shows that the predicted class and the correct class are highly correlated. However, the MCC/Ensemble RUSBoosted Trees correlation of 37.44% shows that the predicted class and the correct class are not as highly correlated (as with the Australian dataset).

Taking into account all indicators, the analysis of classifiers and their predictive abilities do not identify an individual classifier with high predictive power for all, or at least most performance indicators.

Table 3 Performance comparison of individual classifiers for the Australian credit scoring dataset

Method	Sensitivity/Recall	Specificity	G-mean	F-measure	MCC
Logistic regression	81.98%	86.64%	84.28%	83.01%	68.24%
Linear SVM	79.63%	**92.83%**	85.98%	82.47%	72.13%
Quadratic SVM	83.81%	83.71%	83.76%	83.79%	67.31%
Cubic SVM	84.33%	82.41%	83.37%	83.90%	66.63%
Fine Gaussian SVM	**88.77%**	75.90%	82.08%	85.81%	65.60%
Medium Gaussian SVM	79.90%	**92.83%**	86.12%	82.67%	72.37%
Coarse Gaussian SVM	79.90%	**92.83%**	86.12%	82.67%	72.37%
Ensemble Boosted Trees	87.21%	80.13%	83.59%	85.60%	67.64%
Ensemble Bagged Trees	87.21%	87.62%	**87.41%**	**87.30%**	**74.60%**
Ensemble RUSBoosted Trees	86.42%	84.36%	85.39%	85.96%	70.70%

Source: Data processed by the author

Table 4 Performance comparison of individual classifiers for the German credit scoring dataset

Method	Sensitivity/Recall	Specificity	G-mean	F-measure	MCC
Logistic regression	86.00%	45.67%	62.67%	79.49%	34.23%
Linear SVM	88.00%	40.33%	59.58%	80.22%	32.16%
Quadratic SVM	82.11%	38.71%	56.38%	75.57%	21.96%
Cubic SVM	78.14%	47.33%	60.82%	73.23%	25.60%
Fine Gaussian SVM	**99.86%**	1.33%	11.54%	**82.51%**	7.73%
Medium Gaussian SVM	92.14%	32.00%	54.30%	82.14%	30.90%
Ensemble Boosted Trees	86.86%	39.33%	58.45%	79.09%	29.47%
Ensemble Bagged Trees	86.43%	42.33%	60.49%	79.27%	31.71%
Ensemble RUSBoosted Trees	65.43%	**75.33%**	**70.21%**	66.88%	**37.44%**

Source: Data processed by the author

7. Conclusion

Credit scoring is a widely used technique that helps banks decide when granting loans to applicants. In addition to using standard statistical decision-making techniques, such as logistic regression or decision tree, credit scoring is a very interesting task for machine learning and artificial intelligence methods. In recent years, machine learning technologies have been developing rapidly and ensemble learning is being studied more and more.

Several papers have shown the advantages of deep learning over traditional credit scoring methods. In this paper, we investigated the predictive capabilities of ensemble algorithms over credit scoring decision trees and compared them with traditional methods – logistic regression and SVM support vectors.

Ensemble methods are promising classifier and predictive techniques and represent an alternative to classical artificial neural networks (ANN). A large number of studies have shown that in problems of credit scoring classification, ensemble techniques are better than SVM techniques, as well as than traditional techniques such as logistic regression and decision tree. According to each comparison criterion, ensemble methods show better results or at least competitive results with tested predictive techniques.

Based on Australian and German credit scoring data, performance of different classification models is compared. The performance of logistic regression (LR), support vector machine (SVM), and deep learning based on decision tree ensembles is analyzed. Several different performance indicators are used. According to the MCC indicator, which is considered to be the most adequate for classification problems, on average, deep learning methods prove to be the best. Individually, ensembles with boosted trees work best in the Australian dataset, and ensembles with RUSBoosted trees in the German dataset.

Logistic regression performance is relatively poor compared to the ensemble method and SVM. Logistic regression, as the best of the traditional methods, fails to match the tools of machine learning.

The results of this paper confirm previous research on the advantages of machine learning methods over traditional models. Also, the division regarding the obvious advantage of deep learning over SVM methods is confirmed. According to certain criteria, SVM shows better characteristics compared to ensembles. Nevertheless, ensemble methods are promising tools and provide potential for future research.

References

Abdou, H., & Pointon, J. (2011). Credit Scoring, Statistical Techniques and Evaluation Criteria: A Review of the Literature. *Intelligent Systems in Accounting Finance & Management*, *18*, 59–88.

Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, *23*(4), 589-609.

Asuncion, A., & Newman, D. J. (2010). *UCI machine learning repository*. School of information and computer science, Retrieved from: http://archive.ics.uci.edu/ml/

Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., & Vanthienen, J. (2003). Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring. *The Journal of the Operational Research Society*, *54*(6), 627-635.

Bequé, A., & Lessmann, S. (2017). Extreme learning machines for credit scoring: An empirical evaluation. *Expert Systems With Applications*, *86*, 42-53.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*, 123-140.

Dastile, X., Celik, T., & Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing Journal*, *91*, 1-21. https://doi.org/10.1016/j.asoc.2020.106263

Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning. *Neural Computation*, *10*(7), 1895–1923.

Durand, D. (1941). *Risk Elements in Consumer Instalment Financing*. New York: National Bureau of Economy Research.

Einav, L., Jenkins, M., & Levin, J. (2013). The impact of credit scoring on consumer lending. *The RAND Journal of Economics*, *44*(2), 249–274.

Fisher, R. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, *7*(2), 179-188.

Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, *55*(1), 119-139.

Goh, R., & Lee, L. (2019). Credit Scoring: A Review on Support Vector Machines and Metaheuristic Approaches. *Advances in Operations Research*, 1-30. https://doi.org/10.1155/2019/1974794

Hand, D. J., & Henley, W. E. (1997). Statistical classifcation methods in consumer credit scoring: a review. *Journal of the Royal Statistical Society*, *160*(3), 523-541.

Hui, L., Li, S., & Zongfang, Z. (2017). The Model and Empirical Research of Application Scoring Based on Data Mining Methods. *Procedia Computer Science*, *17*, 911-918.

Jurman, G. R. S. (2012). A comparison of MCC and CEN error measures in multi-class prediction. *PLoS ONE*, *7*(8), 41882.

Kim, A. Y.-C. (2020). Can deep learning predict risky retail investors? A case study in financial risk behavior forecasting. *European Journal of Operational Research*, *283*(1), 217-234.

Lewis, E. (1992). *An Introduction to Credit Scoring*. San Rafael: Fair, Isaac and Co., Inc.

Li, Y., Lin, X., Wang, X., Shen, F., & Gong, Z. (2017). Credit Risk Assessment Algorithm Using Deep Neural Networks with Clustering and Merging. 13th International Conference on Computational Intelligence and Security (CIS) (pp. 173-176). Hong Kong: IEEE.

Lou, C., Wu, D., & Wu., D. (2017). A deep learning approach for credit scoring using credit default swaps. *Engineering Applications of Artificial Intelligence*, *65*, 465-470.

Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure*, *405*(2), 442–451.

Nanni, L., & Lumini, A. (2009). AAn experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring. Expert Systems with Applications, *36*, 3028–3033.

Neagoe, V., Ciotec, A., & Cucu, G. (2018). Deep Convolutional Neural Networks Versus Multilayer Perceptron for Financial Prediction. *2018 International Conference on Communications (COMM)* (pp. 201-206). Bucharest: IEEE

Odom, M., & Sharda, R. (1990). A neural network model for bankruptcy prediction. *1990 IJCNN International Joint Conference on Neural Networks*, 2, pp. 163-168.

Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, *18*(1), 109-131.

Orgler, Y. (1971). Evaluation of bank consumer loans with credit scoring models. *Journal of Bank Research*, *2*(1), 31-37.

Oztekin, A., Al-Ebbini, L., Sevkli, Z., & Delen, D. (2018). A decision analytic approach to predicting quality of life for lung transplant recipients: A hybrid genetic algorithms-based methodology. *European Journal of Operational Research*, *266*(2), 639–665.

Patil, P. S., Aghav, J. V., & Sareen, V. (2016). An Overview of Classification Algorithms and Ensemble Methods in Personal Credit Scoring. *International Journal of Computer Science and Technology, 7*(2), 183-188.

Pławiak, P., Abdar, U., & Acharya, R. (2019). Application of new deep genetic cascade ensemble of SVM classifiers to predict the Australian credit scoring. *Applied Soft Computing Journal*, *84*, 105740.

Ren, Y., Zhang, L., & Suganthan, P. (2016). Ensemble Classification and Regression-Recent Developments, Applications and Future Directions. *IEEE Computational Intelligence Magazine*, *11*(1), 41-53.

Sadatrasoul, S. M., Gholamian, M. R., Siami, M., & Hajimohammadi, Z. (2013). Credit scoring in banks and fnancial institutions via data mining techniques: a literature review. *Journal of AI and Data Mining*, *1*(2), 119-129.

Thomas, L., Crook, J., & Edelman, D. (2002). *Credit Scoring and Its Applications, Second Edition*. Philadelphia: Society for Industrial and. and Applied Mathematics.

Vapnik, V. N. (1998). *Statistical Learning Theory*. New York: Wiley- Interscience.

Wang, G., Hao, J., Ma, J., & Jiang, H. (2011). A comparative assessment of ensemble learning for credit scoring. *Expert Systems with Applications*, *38*(1), 223-230.

Zhou, H., Lan, Y., Soh, Y., Huang, G., & Zhang, R. (2012). Credit risk evaluation with extreme learning machine. *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, (pp. 1064-1069). Seoul.

KREDITNO BODOVANJE POMOĆU ANSAMBLERSKIH METODA DUBOKOG UČENJA ZA KLASIFIKACIJU – POREĐENJE SA TRADICIONALNIM METODAMA

Kreditni skoring (kreditno bodovanje) privlači posebnu pažnju finansijskih institucija. Poslednjih godina, posebno su interesantni metodi zasnovani na dubokom učenju. U ovom radu, upoređujemo performanse ansamblerskim metodama dubokog učenja zasnovanih na stablima odlučivanja sa najboljom klasičnom metodom, logističkom regresijom i već postavljenim reperom za metode mašinskog učenja, mašinama sa vektorskom podrškom. Za svaku metodu testirano je više različitih algoritama. Takođe, korišćeni su različiti indikatori performansi. Istraživanje je izvršeno nad standardnim bazama za ovu vrstu klasifikacije, Australijskim i Nemačkim skupom podataka. Kao najbolja metoda, prema MCC indikatoru, pokazala se ansamblerska metoda sa boosted stablima odlučivnja. Takođe, u proseku, ansamblerske metode su se pokazele uspešnijim od SVM.

Ključne reči: *kreditno bodovanje, ansambli za klasifikaciju, duboko učenje, mašine za vektorsku podršku*

Original Scientific Paper

SUSTAINABILITY RISK MANAGEMENT IN THE DIGITAL ECONOMY

UDC 005.334:501.131.1

Jelena Z. Stanković, Evica Petrović, Jovica Stanković

University of Niš, Faculty of Economics, Serbia

Abstract. *The problem of sustainable development has become an imperative of globalization, which resolutely sets the request for companies to operate socially responsibly, i.e., to create value in a manner that is sustainable in the future by achieving economic, environmental and social goals. The wave of change, conditioned by digital transformation, is considered an opportunity, but also a challenge for the realization of the concept of sustainable development. Therefore, the aim of this paper is to consider the risks of sustainable business emerging from the implementation of ICT in the business process, with focus on the companies in the Republic of Serbia.*

Key words: *sustainability risks, digitalization, risk management, sustainable development*

JEL Classification: Q01, Q56, L86, Q55, G32

1. INTRODUCTION

The increase in the number of risks associated with climate change and social disturbances, as well as the intensity of their effects, is estimated to reduce the gross domestic product of the global economy by up to 20% in the future (Stern & Stern, 2007). In such business conditions, the problem of sustainable development is becoming increasingly important in corporate governance, due to the fact that companies are required to operate in a socially responsible way by achieving economic, environmental and social goals (Dangelico & Pujari, 2010). Based on data on corporate social responsibility, it can be concluded that the number of companies that apply the principles of sustainable development in their business has increased. Thus, in 2011, only 20% of the total number of companies

---

included in the S&P 500 index basket published reports on sustainable operations, while in 2016 their number increased to 81% (Coppola, 2016).

The application of information and communication technologies (ICT) can provide solutions to many existing business challenges, such as shortening innovation cycles, increasing volatility and competitiveness in a dynamic business environment (Kagermann, 2015). However, if the importance of the digitization process for human life and work is taken into account, it can be concluded that this process will be crucial for achieving a long-term balance between social development and the environment (Souter et al., 2010). As the concept of sustainable development cannot be achieved without global communication and knowledge exchange (MacLean et al., 2007), business digitalization can be an opportunity to overcome development constraints, but also a serious risk factor for sustainable development, because the effects of Industry 4.0 on sustainability are still uncertain (Birkel et al., 2019). Due to the fact that 8 of the 10 most important business risks are directly related to the environmental and social problems (Schulte & Hallstedt, 2017), risk management is a key factor for integrating the principles of sustainable development into business operations (Schulte & Hallstedt, 2018). Thus, the aim of this paper is to consider the risks of sustainable business, arising from the implementation of ICT in the business process. Moreover, the application of ICT in companies in the Republic of Serbia for achieving the goals of sustainable business is assessed. According to this, the paper is structured in the following way: the second part describes the impact of ICT on socially responsible business, the third part explains the risks of sustainable development in the digital economy and opportunities for their integration into existing risk management frameworks. In order to access the risks of business digital transformation in the Republic of Serbia, a SWOT analysis of the application of ICT in companies in Serbia is presented in the fourth part of the paper. In the conclusion, the potential effects of the synergy of the digital economy and socially responsible business are presented, as well as the limitations of the existing application of ICT, especially in companies in the Republic of Serbia.

2. SOCIALLY RESPONSIBLE BUSINESS IN CONDITIONS OF DIGITALIZATION

The technological development of modern civilization has reached the stage of meeting different development goals, and the tendency to maintain the traditional model of economic growth leads to a conflict between economic and environmental goals. The combination of multiple pressures, such as continuous population growth, changes in consumption patterns and climate change, leads to a reduction of limited resources such as food, water, energy and materials, below the critical limit. Therefore, the concept of sustainable development has been gradually expanding by connecting sociology, economy and ecology. The focus of interest is on a complex, multidimensional ecosocial system, which provides basic services to society, such as food, drinking water, energy, and whose key components are environmental and social subsystems.

The complex and dynamic process of economic development has initiated a number of structural changes in the economy, both locally and globally. These changes include changes in the structure of demand and production, changes in the professional and qualification structure of the workforce, foreign trade and finance. The main characteristic of these changes is that they are not realized immediately, but in a long period of time through economic crises, transformations and improvements. As a result, there is an enormous increase in the impact of

human activities on the ecological subsystem and the ecosocial system as a whole, but also the emergence of risks. Digitalization, as one of the biggest drivers of change in the 21st century, has completely changed the concept of development of the global economy and society, putting in the forefront the potential of ICT and the challenges of environmental sustainability (Ciocoiu, 2011).

The process of digitalization, as a process of connecting people and things and continuous convergence of the real and virtual world based on ICT, has had such a great impact on the ecosocial system that it has caused a re-examination of the basic structures of this system. Information and communication technologies have a dominant role in redefining products, processes, organizational structure and business models in many economic activities. Digital transformation can be defined as the expanded use of modern information technologies, such as analytics, mobile devices, social networks or smart integrated devices, as well as the advanced use of traditional technologies, such as ERP, to improve business (Westerman et al., 2014). This great wave of change is considered an opportunity, but also a challenge to achieve the United Nations Sustainable Development Goals (SDG), especially when it comes to reducing various forms of inequality, ensuring decent work and responsible production and consumption. Implementation in various spheres and the development of digital technologies create unique opportunities for improving social and environmental well-being, and further enhancing global living standards while preserving and improving the environment for future generations (Linkov et al., 2018). However, its impact on sustainable development is still not possible to assess reliably, although digitalization is expected to transform the economy, as well as society, in terms of economic, environmental and social performance (Muller et al., 2018). The classification of possible impacts of ICT on sustainable development depends on the level of development and application of information technologies in one community and is shown in Table 1.

Table 1 Classification of the impact of ICT on sustainable development

Order of effects	Impact	Direction of action
First order effects: production and application of ICT	Increased level of pollution and energy consumption necessary for the production and application of ICT, as well as for its disposal after use.	Negative
Second order effects: use of ICT in order to reduce energy / water consumption, reduce pollution	Optimization of production processes and unsustainable consumption due to the application of ICT (e.g. reduction of energy consumption using various applications).	Mostly positive
Third order effects: change of life habits	The cumulative medium- to long-term impact that the use of ICT can have by large numbers of people (e.g., teleworking instead of traveling).	Positive
Fourth order effects	Improving overall social decision-making capacities and measuring their impact in the field of real-time sustainable development policy implementation.	Positive

Source: Forge et al. (2009)

Current research studies are devoted mainly to the technical challenges of implementing the Industry 4.0 concept (Liao et al., 2017), while the analysis of the risks associated with it has not yet been systematically described in the scientific literature. Isolated risk observation

does not contribute to the detection of interdependencies that exist between the risks of traditional and digital business, which can cause a decrease in profit or reallocation of business due to loss of comparative advantages of companies that do not achieve adequate ICT implementation in the existing business model (Ghanbari et al., 2017). The transformation of the organization of the company can be further slowed down by social risks, such as requirements for retraining of employees or termination of jobs, especially for simple and routine operations, which are replaced by automated processes (Kazancoglu & Ozkan-Ozen, 2018). Therefore, in the continuation of the paper, the changed effect of basic and additional business risks conditioned by the application of ICT in business will be explained.

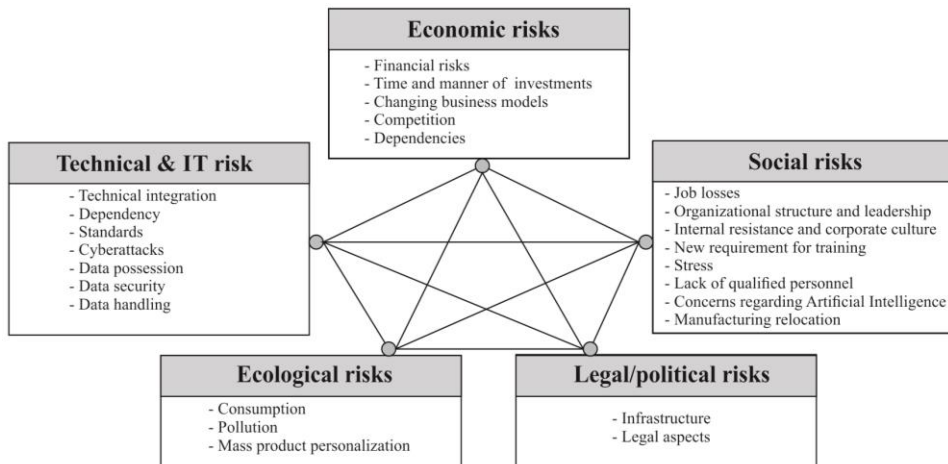## 3. RISKS OF SUSTAINABLE BUSINESS IN THE DIGITAL ECONOMY

Digitalization of management process, supported by comprehensive data analysis, advanced decision-making algorithms, and artificial intelligence, affects all three dimensions of sustainability (Kiel et al., 2017). However, in addition, the technical aspects of the implementation of the Industry 4.0 concept have opened up a number of questions regarding technical and regulatory requirements (Müller & Voigt, 2018), which, if not adequately regulated, can become risk factors (Fig. 1).

Economic risks of ICT implementation in business operations affect the achievement of business results. Considering the complexity of information technology, the high value of investments, both in equipment and in employees, can be identified as one of the basic economic risks (Birkel et al., 2019). Due to the inability to accurately assess the economic life of the implemented technology, as well as the uncertainty regarding cash flow planning in the long run, decisions on the realization of these business ventures are very often classified in the group of those that require a careful assessment of the beginning of the investment period (Raj et al., 2020). Actually, investing in ICT in the mature phase of development, may cause smaller investments, but it does not provide the company with the expected comparative advantage. On the other hand, investments in underdeveloped technologies carry an additional risk of choosing the wrong technological solutions, which will not provide the company with the desired competitive position in the market. As an additional aspect of the impact of ICT on economic risks, the emergence of new business models can be pointed out, such as business cooperation through platforms (Kagerman, 2015), which, in addition to being an additional instrument for increasing business value, also creates the so-called risk of cannibalization of existing business models.

The implementation of digital technologies in business practice leads to the emergence of numerous environmental risks. The first effects of the use of ICT are mostly negative and relate to the excessive use of energy and a large quantity of raw materials for the production of equipment (Stock et al., 2018). Excessive use of energy is the result of the complex work process, especially the usage of additional services in data processing. Environmental pollution can be caused by inadequate disposal of equipment after use since for the most of the technology materials recycling processes are not available or are unprofitable to conduct (Behrendt, 2018). Despite significant efforts to reuse or adapt existing plants and equipment, this is not possible in most cases when more advanced technical and technological solutions have appeared in order to reduce costs and resources. The complexity of new technologies and the required capacities for their implementation

limit the possibilities for using the existing capacities of the company. In addition, there may be an increase in the amount of waste due to mass personalization of products (Wang et al., 2017), which highlights the negative effects of the application of ICT in business and in the mature stages of its implementation.

The integration of digital technologies into the business process inevitably imposes the need for new competencies of employees, especially in terms of using ICT, reducing the number of business activities, but also increasing the flexibility of work organization, which can create additional pressures on the existing social tension (Schwab, 2017). In other words, employees are expected to have qualifications and skills to use advanced tools and technologies, as well as to be able to solve problems and perform multiple work tasks simultaneously in a complex environment (Bonekamp & Sure, 2015). Such requirements often create a stressful work environment in which employees do not understand the role of advanced technologies in work, and often do not accept the changes that business modernization requires, despite additional training and education. In extreme situations, companies decide to relocate their business facilities, which increases unemployment and deepens the problem of social inequality.

Fig. 1 ICT risks of sustainable development and their interdependence
Source: Birkel et al. (2019)

A specific group of sustainability risks in the digital economy is represented by the risks of technical integration of the concept of digitalization into the business of the company (Kagermann, 2015). An integrated approach to business can have great potential for the growth and development of companies, but it is conditioned by a high level of complexity in implementation. This way of doing business becomes too dependent on technology, so the failure of software or systems can lead to a break in the entire value chain. Vertical and horizontal integration of business through ICT opens the business system to various external influences, and makes it vulnerable to cyber-attacks, but also calls into question the protection of data and the way they are used. Therefore, at the national level, it is necessary to have an adequate infrastructure, which will support the implementation of advanced ICT in business (Raj et al., 2020), and the globalization of business implies a high level of standardization

(Schwab, 2017). Mass digitalization of business opens numerous problems, such as data protection, determining working hours, but also the protection and safety of employees at work, which is why the impact of advanced technologies on sustainability needs to be considered in terms of regulatory risks.

In accordance with these risks, there is a change in the role and importance of enterprise risk management in modern business conditions in a way that the long-term sustainability of business is closely linked to enterprise risk management (ERM). In such circumstances, the triple bottom line becomes a mandatory framework for measuring the results of socially responsible business. Companies that focus on socially responsible business are primarily engaged in activities that avoid future business problems, but also society, as a whole, is protected from potentially harmful effects of business operations (Godfrey et al., 2009). The fact is that these risks may have an impact on firm performance in the long run, but the effects of their manifestation cannot yet be adequately assessed, nor the link between long-term and short-term effects of sustainability risks and business risks (Bromiley et al., 2015). This situation is a consequence of the inconsistency of the way of integrating the concept of sustainable development into the existing framework of corporate governance and ERM. In other words, the requirement for corporate social responsibility can be seen as a specific business goal of the company (Saardchom, 2013) or existing business goals can be seen through the perspective of corporate social responsibility (Faris et al., 2013). The chosen aspect of the integration of the principles of sustainable development into the business will also affect the way of integrating sustainability risk into the ERM framework. ERM frameworks should be adapted to the new requirements for sustainability by including the concept of sustainability in all phases of risk management (Saardchom, 2013). Defining an adequate risk management framework enables companies to comprehensively and systematically assess, control, use, finance and monitor financial, strategic and operational risks, which in the short and long term may affect the interests of stakeholders.

In order to incorporate the principles of corporate social responsibility into the strategic management process, it is necessary to be aware of the risks of sustainability when performing all business activities - from daily, operational business to investment decision-making. The process of business digitalization goes beyond the digitalization of resources and involves the transformation of key business activities, products and processes, leading to modified or completely new business models, so the implementation of this process requires an appropriate strategy (Bharadwaj et al., 2013). In order to formulate and implement a digital transformation strategy, companies must consider: the use of technology, changes in value creation, structural changes and financial aspects of this process. When it comes to the external environment, market conditions, competition, technological trends, regulatory requirements, climate change, country risks and political risks can be considered the key factors of business success and sustainability, but also the conditions and method of financing, needs and requirements of external stakeholders and all other external influences and risks associated with them. The complexity of the company in terms of its size, number of organizational units, level of vertical integration, internal regulations, as well as the existing business strategy can be considered the internal factors that affect ERM in terms of digitalization and sustainability of the company. Therefore, in the following part of the paper the main risks of digitalization in the Republic of Serbia will be assessed using SWOT analysis.

4. RISKS OF DIGITALIZATION IN THE REPUBLIC OF SERBIA

The ICT sector in Serbia offers potential for sustainable development. In 2017, the ICT sector participated with 4.42% in the creation of GDP (Eurostat Database), while in the same period, EUR 768 million (NBS Database) was generated from the export of ICT services. However, the application of ICT is still not in the function of sustainable development and increasing competitiveness. In the Republic of Serbia, 100% of companies have Internet connection, 84.4% of companies have a website, but only 27.9% of companies sell products or services online (SORS, 2020). Moreover, the largest number of companies (over 50%) made less than 25% of profit through the online sales. This situation, as well as the fact that only 23% of companies use CRM to manage customer information, indicates that companies in Serbia do not sufficiently use the potential of the ICT sector. The impact of digitalization on sustainable development in the Republic of Serbia will be assessed by analyzing two indices: The Network Readiness Index and the Sustainable Development Index.

The Network Readiness Index (NRI) is a composite index that comprehensively assesses how society and the economy influence the development and application of information and communication technology. In order to adequately assess the impact of ICT on the development of society and the economy, NRI has changed its structure since 2002, when it was first presented to the public. Thus, in 2019 the index was significantly revised and improved. Data for the Republic of Serbia in comparison with the surrounding countries are higher, and a growing trend can be observed during the observed period (2009 - 2015), but not as high as in the EU Member States (Table 2).

Table 2 Descriptive statistics of NRI and SDI indicators in the period 2009 – 2015 in the Republic of Serbia and neighboring countries

	Serbia		North Macedonia		Bulgaria		Croatia		Bosnia and Herzegovina		Albania	
	NRI	SDI	NRI	SDI	NRI	SDI	NRI	SDI	NRI	SDI	NRI	SDI
Min	3,51	0,40	3,64	0,71	3,66	0,76	3,91	0,70	3,07	0,72	3,23	0,78
Max	4,00	0,68	4,40	0,74	4,00	0,78	4,34	0,75	3,99	0,74	3,89	0,81
Mean	3,70	0,52	3,93	0,72	3,85	0,77	4,13	0,73	3,50	0,73	3,58	0,80
St. Dev.	0,18	0,10	0,28	0,01	0,11	0,01	0,17	0,02	0,37	0,01	0,25	0,01

Source: Authors' calculation

On the other hand, if the level of Sustainable Development Index (SDI) is observed, it can be concluded that in the observed period (2009 – 2015) the Republic of Serbia was the least ecologically efficient in the process of human development (Table 2). Contrary to expectations, the value of SDI decreased during the period, so that the value of SDI fell from 0.69 in 2009 to 0.40 in 2015 (Fig. 2).

Spearman's rank correlation coefficient was calculated and its significance was tested in order to examine the interdependence between variations in NRI and SDI values in the Republic of Serbia. This indicator belongs to the category of nonparametric indicators and represents a nonparametric analogy of the simple linear correlation coefficient. Since the calculation of Spearman's rank correlation coefficient does not require the fulfillment of assumptions related to data distribution, its application is considered adequate in this case.
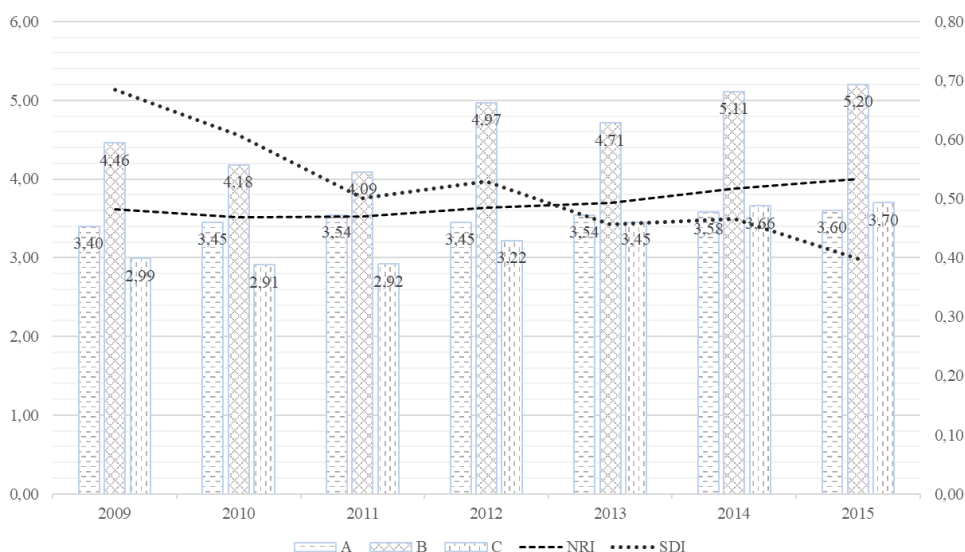
Fig. 2 NRI and its sub-indices scores and SDI scores for the Republic of Serbia
Source: Authors' presentation

Table 3 Spearman's correlation coefficient values by indices and their sub-indices

	NRI	A: Environment	B: Readiness	C: Usage	SDI
NRI	1.00	0.73	**0.93***	**0.99***	-0.79*
A: Environment	0.73	1.00	0.58	0.73	**-0.93***
B: Readiness	**0.93***	0.58	1.00	**0.93***	-0.61
C: Usage	**0.99***	0.73	**0.93***	1.00	**-0.79***
SDI	**-0.79***	**-0.93***	-0.61	**-0.79***	1.00

*statistically significant correlation with *p*-value less than 0.05
Source: Authors' calculation

The calculated values of Spearman's rank correlation coefficient between variations in the estimates of NRI and its sub-indices and SDI, shown in Table 3, indicate the existence of a strong negative correlation, which is at the same time statistically significant. The value of the correlation coefficient between NRI and SDI is -0.79 and is statistically significant. The high values of Spearman's correlation coefficient between SDI and the NRI sub-indices Environment and Usage indicate a strong degree of indirect agreement in variations in the capacity of the environment and the use of ICT with the assessment of sustainable development in the Republic of Serbia. Specifically, the constant increase in the value of NRI scores is not sufficiently in accordance with the changes in SDI values.

Due to the adoption of Agenda 2030 and the Sustainable Development Goals in 2015, the method of calculating SDI has changed significantly, and during 2019 and 2020, the NRI index included for the first time the impact of ICT on sustainable development. Having in mind the mentioned risks of digitalization for sustainable development, the potential impacts of ICT on sustainable development in the Republic of Serbia were assessed using SWOT analysis. In this case, the assessments of the most important

constituents of the NRI for the Republic of Serbia and their rank in relation to the score of the related sub-pillar, as well as in relation to the average score for the EU and countries of the same level of development were discussed. In accordance with the chosen method of assessment, the main weaknesses of Industry 4.0 and threats that could potentially slow down the achievement of the Sustainable Development Goals of the economy of the Republic of Serbia have been identified (Table 4).

If economic risks are observed, it can be concluded that insufficient investment is one of the basic weaknesses of companies in the Republic of Serbia. Digital technologies in the form of e-services, robotics or process management solutions can help companies achieve their business goals in the SDG context. However, investments of Serbian companies in the development of the ICT are very modest - only 2.06% of total investments in research and development relate to programming, consulting and related activities (Eurostat Database). Specifically, only 18.6% of companies pay for the Cloud services (RZS, 2020).

Table 4 SWOT analysis of the ICT risks and opportunities in the Republic of Serbia

Strengths	Weaknesses
▪ Large number of firms having own website ▪ Secured access to website ▪ Quality of e-Government services ▪ Accessibility, quality and relevance of government websites	▪ Insufficient investment in and adoption of AI, robotics, applications and websites for e-commerce, big data analytics and Cloud computing ▪ Low value of purchased or leased computer software ▪ Inadequate computer literacy skills of employees ▪ Low level of mobile and e-business usage
Opportunities	**Threats**
▪ High R&D expenditure in ICT by governments and higher education ▪ Internet access and usage ▪ Fixed-broadband access ▪ Wide area coverage for Mobile network access (4G) ▪ Regulatory and legal framework for ICT ▪ Satisfactory legal protection of Internet privacy	▪ Low availability of government published open data ▪ Cybersecurity risks ▪ Low level of usage of e-Bank and m-Bank services, especially in rural areas ▪ Inadequate policy and regulatory environment for private sector development ▪ Slow policy and regulatory adaptability to disruptive technologies ▪ Unsatisfactory legal framework for e-commerce ▪ Internet usage gender inequality ▪ Low level of environment monitoring

Source: Authors' elaboration according to Dutta and Lanvin (2020)

In regards to social risks, companies are burdened by the lack of basic ICT skills of employees, but also by the lack of permanently employed ICT experts. Only 19.3% of companies in Serbia permanently employ the ICT experts, while 14.6% provide ICT training for their employees and 56.6% outsource ICT services (SORS, 2020). Low level of investments in employee training prevents better exploitation of available ICT capacities and slows down their integration into business models of companies. However, demand for qualified ICT experts in companies is high, since 43.7% of companies have unsuccessfully tried to employ ICT experts. The main obstacles for ICT experts' engagement are the lack of relevant experience (71.4%), the lack for adequate qualifications (65.6%), as well as high expectations regarding salaries (28.6%) (SORS, 2020). Considering the low level of robots' usage – only 3% of companies use some type of robots in their business

(Eurostat Database), it can be concluded that there is still no significant risk of job losses nor organization structure changes.

ICT infrastructure and legal regulations in the field of ICT provide a solid basis for the development of Industry 4.0 in the Republic of Serbia. However, the legislation does not sufficiently follow the specific areas of ICT application and does not adapt quickly enough to the changes that have occurred. Specifically, the weakest component of ICT legislative environment is the e-commerce legislation (Serbia's rank 77/134) and the adaptability of current legislative to emerging technologies (Serbia's rank 72/134) (Dutta & Lanvin, 2020). This can be an obstacle to the development of e-business and the use of e-banking and m-banking services, as well as significant change of existing business models.

The technical risks that companies in the Republic of Serbia face are mainly related to the availability of open data, data security and the implementation of Cloud solutions. Low availability of government published open data places Serbia to 65th place out of 134, while cybersecurity risks rank is 60/134 (Dutta & Lanvin, 2020). Thus, very small number of companies uses Big Data analysis - only 2% of companies in the Republic of Serbia comparing to the 13% of companies in the EU (Eurostat Database).

On the other hand, the contribution of ICT to the environment in the Republic of Serbia cannot be reliably assessed, because despite the fact that the application of ICT should contribute to improving environmental planning, natural disaster management and increased energy efficiency, environmental monitoring in the Republic of Serbia is very low and this information is usually not available. Instead, the most important contribution of ICT to SDG is observed for sustainability of cities and communities (41/134), quality education (45/134) and gender equality (53/134), while the contribution to achieving goals of affordable and clean energy is rather low (109/134), as well as good health and well-being (88/134) (Dutta & Lanvin, 2020).

5. CONCLUSION

The company's business has changed significantly in recent decades. Companies operating in modern business conditions are directed to socially responsible business, which implies progress in terms of economic prosperity, environmental quality and social justice. Technical and technological changes, which have influenced knowledge and networks to play a greater role in economic development than capital, have led to the emergence of the concept of a sustainable digital economy as a solution to environmental and social problems. The synergetic effects of the digital economy and sustainable development have been recognized in developed economies. Besides this, based on experience, it can be concluded that ICT can support sustainable development in the following ways (IISD, 2010):

▪ reducing the direct effects of production, distribution, use and disposal of equipment on the environment by improving energy efficiency, using energy from renewable sources, reducing the use of toxic materials and improving recycling and waste disposal. ICTs are used to optimize the operations, thus increasing their efficiency while reducing pollution and resource consumption.

▪ increasing the effects of the ICT usage on sustainable development through improving the efficiency of production, distribution and consumption of goods and services. Complete or partial dematerialization of products and services enables the creation of their virtual

substitutes, the production and use of which can significantly reduce emissions and energy use.

▪ supporting systemic effects resulting from changes in the behavior, affinities and value systems of individuals as citizens and consumers; changes in economic and social structures and political processes. ICT is used in the visualization and communication of information relevant to sustainable development, such as the use of energy, water or CO2 emissions, which is a prerequisite for the application of the principles of sustainable development.

Although countries have recognized the potential of combining the digital economy and the principles of sustainable development, a large number of companies still do not consider the integration of sustainability principles to be essential for their business (Ernst & Young, 2017) and view these requirements exclusively from a risk management perspective. Analysis of the NRI and SDI index imply that Serbian companies do not apply ICT in the context of achieving SDGs. Companies confront many ICT risks, mainly regarding ICT skilled employees and financial means for investment in ICT. Moreover, legislative does not regulate the specific areas of ICT application and does not adapt quickly enough to the dynamics of digital economy decreasing the potential of ICT for sustainable development.

However, effective sustainability risk management implies adequate integration of these risks into existing risk management frameworks, but there are some limitations in practical application (Schulte & Hallstedt, 2017). Unlike traditional risks, these risks have certain properties, such as the dynamics of change, qualitative assessment of the effects, pronounced uncertainty of manifestation. According to the available reports, the effects of most ICT risks on the economy are not available. In addition, there is no consensus on the definition of sustainable development, which prevents the creation of a clear risk management strategy, obscures accountability and prevents the establishment of an interdependence between costs and benefits. All this points to the insufficient maturity of the ERM to manage the risks of sustainable development.

REFERENCES

Behrendt, S. (2018). „Recycling von Technologiemetallen "Transformationsfeldanalyse im Rahmen des Projektes „Evolution2Green–Transformationspfade zu einer Green Economy "Institut für Zukunftsstudien und Technologiebewertung.

Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. (2013). Digital business strategy: toward a next generation of insights. *MIS quarterly*, 471-482.

Birkel, H. S., Veile, J. W., Müller, J. M., Hartmann, E., & Voigt, K. I. (2019). Development of a risk framework for Industry 4.0 in the context of sustainability for established manufacturers. *Sustainability*, *11*(2), 384.

Bonekamp, L., & Sure, M. (2015). Consequences of Industry 4.0 on human labour and work organisation. *Journal of Business and Media Psychology*, *6*(1), 33-40.

Bromiley, P., McShane, M., Nair, A., & Rustambekov, E. (2015). Enterprise risk management: Review, critique, and research directions. *Long range planning*, *48*(4), 265-276.

Ciocoiu, C. N. (2011). Integrating digital economy and green economy: opportunities for sustainable development. *Theoretical and Empirical Researches in Urban Management*, *6*(1), 33-43.

Coppola, L. (2016). Eighty one percent of the S&P 500 index companies published corporate sustainability reports in 2015. Flash report. *Governance & Accountability Institute*.

Dangelico, R. M., & Pujari, D. (2010). Mainstreaming green product innovation: Why and how companies integrate environmental sustainability. *Journal of business ethics*, *95*(3), 471-486.

Dutta, S., & Lanvin, B. (2019). The Network Readiness Index 2020: Accelerating Digital Transformation in a post-COVID Global Economy. *Washington DC: Portulans Institute, WITSA. Retrieved November*, *14*, 2020.

Ernst & Young. (2017). Is your nonfinancial performance revealing the true value of your business to investors? Retrieved from: https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/purpose/purpose-pdfs/ey-nonfinancial-performance-may-influence-investors.pdf, Accessed on: 14 December 2020.

Eurostat Database available at: https://ec.europa.eu/eurostat/web/main/data/database, Accessed on: 1 December 2020.

Faris, C., Gilbert, B., LeBlanc, B., Ballou, B., & Heitger, D.L. (2013), *Demystifying Sustainability Risk: Integrating the Triple Bottom Line into an Enterprise Risk Management Program*, Committee of Sponsoring Organizations of the Treadway Commission.

Forge, S., Blackman, C., Bohlin, E., & Cave, M. (2009). *A Green Knowledge Society*. An ICT policy agenda to 2015 for Europe's future knowledge society, A study for the Ministry of Enterprise, Energy and Communications, Government Offices of Sweden, published by SCF Associates Ltd, September 2009.

Ghanbari, A., Laya, A., Alonso-Zarate, J., & Markendahl, J. (2017). Business development in the Internet of Things: A matter of vertical cooperation. *IEEE Communications Magazine*, *55*(2), 135-141.

Godfrey, P.C., Merrill, G.B., & Hansen, J.M. (2009). The relationship between corporate social responsibility and shareholder value: an empirical test of the risk management hypothesis. *Strategic Management Journal, 30(4)*, 425– 445.

IISD (2010). *The Digital Economy and the Green Economy: Opportunities for strategic synergies*. International Institute for Sustainable Development.

Kagermann, H. (2015). Change through digitization—Value creation in the age of Industry 4.0. In *Management of permanent change* (pp. 23-45). Springer Gabler, Wiesbaden.

Kazancoglu, Y., & Ozkan-Ozen, Y. D. (2018). Analyzing Workforce 4.0 in the Fourth Industrial Revolution and proposing a road map from operations management perspective with fuzzy DEMATEL. *Journal of Enterprise Information Management*.

Kiel, D., Müller, J. M., Arnold, C., & Voigt, K. I. (2017). Sustainable industrial value creation: Benefits and challenges of industry 4.0. *International Journal of Innovation Management*, *21*(08), 1740015.

Liao, Y., Deschamps, F., Loures, E. D. F. R., & Ramos, L. F. P. (2017). Past, present and future of Industry 4.0-a systematic literature review and research agenda proposal. *International journal of production research*, *55*(12), 3609-3629.

Linkov, I, Trump, B. D., Anklam, E., Berube, D., Boisseasu, P., Cummings, C., ... & Vermeire, T. (2018). Comparative, collaborative, and integrative risk governance for emerging technologies. *Environment Systems and Decisions*, *38*(2), 170-176.

MacLean, D., Andjelkovic, M., & Vetter, T. (2007). *Internet governance and sustainable development: towards a common agenda*.

Müller, J. M., & Voigt, K. I. (2018). Sustainable industrial value creation in SMEs: A comparison between industry 4.0 and made in China 2025. *International Journal of Precision Engineering and Manufacturing-Green Technology*, *5*(5), 659-670.

Müller, J. M., Kiel, D., & Voigt, K. I. (2018). What drives the implementation of Industry 4.0? The role of opportunities and challenges in the context of sustainability. *Sustainability*, *10*(1), 247.

NBS Database available at: https://nbs.rs/sr_RS/drugi-nivo-navigacije/statistika/, Accessed on: 14 December 2020.

Raj, A., Dwivedi, G., Sharma, A., de Sousa Jabbour, A. B. L., & Rajak, S. (2020). Barriers to the adoption of industry 4.0 technologies in the manufacturing sector: An inter-country comparative perspective. *International Journal of Production Economics*, *224*, 107546.

SORS (2020). *Usage of information and communication technologies in the Republic of Serbia*. Statistical Office of the Republic of Serbia.

Saardchom, N. (2013). Enterprise Risk Management under Sustainability Platform. *Journal of Business and Economics, 4 (1)*, 32-41.

Schulte, J., & Hallstedt, S. (2017). Challenges for integrating sustainability in risk management-current state of research. In *21st International Conference on Engineering Design, ICED, Vancouver, Canada, 21 August 2017 through 25 August 2017* (No. DS87-2, pp. 327-336). The Design Society.

Schulte, J., & Hallstedt, S. I. (2018). Company risk management in light of the sustainability transition. *Sustainability*, *10*(11), 4137.

Schwab, K. (2017). *The fourth industrial revolution*. Currency.

Souter, D., MacLean, D., Okoh, B., & Creech, H. (2010). ICTs, the Internet and Sustainable Development: Towards a new paradigm. *International Institute for Sustainable Development*, *39*.

Stern, N., & Stern, N. H. (2007). *The economics of climate change: the Stern review*. Cambridge University press.

Stulz, R. M. (2008). Risk management failures: What are they and when do they happen?. *Journal of Applied Corporate Finance, 20(4)*, 39-48.

Stock, T., Obenaus, M., Kunz, S., & Kohl, H. (2018). Industry 4.0 as enabler for a sustainable development: A qualitative assessment of its ecological and social potential. *Process Safety and Environmental Protection*, *118*, 254-267.

Wang, Y., Ma, H. S., Yang, J. H., & Wang, K. S. (2017). Industry 4.0: A way from mass customization to mass personalization production. *Advances in Manufacturing*, *5*(4), 311-320.

Westerman, G., Bonnet, D., & McAfee, A. (2014). The nine elements of digital transformation. *MIT Sloan Management Review*, *55*(3), 1-6.

UPRAVLJANJE RIZICIMA ODRŽIVOG RAZVOJA U DIGITALNOJ EKONOMIJI

Problem održivog razvoja je postao imperativ globalizacije, koji pred preduzeća sve odlučnije postavlja zahtev da posluju društveno odgovorno, odnosno stvaraju vrednost na način koji je održiv u budućem periodu ostvarujući ekonomske, ekološke i društvene ciljeve. Talas promena uslovljen digitalnom transformacijom se smatra mogućnošću, ali i izazovom za ostvarenje koncepta održivog razvoja. Stoga je cilj ovog rada sagledavanje rizika održivog poslovanja, koji proizilaze iz implementacije IKT u proces poslovanja, sa posebnim osvrtom na preduzeća koja posluju u Republici Srbiji.

Ključne reči: *rizici održivosti, digitalizacija, upravljanje rizicima, održivi razvoj.*

Original Scientific Paper

THE NIGERIAN INSURANCE INDUSTRY COMPETITIVE ADVANTAGE: THE ROLE OF INNOVATION CAPABILITY

UDC 368(669)

Sunday Abayomi Adebisi, Joyce M. Odiachi, Abdul-Hammed A. Sulaimon

University of Lagos, Lagos, Nigeria

Abstract. *Emerging technologies have given rise to greater opportunities within the insurance sector with innovations upturning the traditional business model. This study examined competitive advantage and the relationship with innovation capability in Nigerian insurance sector. Employing a cross sectional research design, the study utilised primary data obtained from selected insurance companies through the use of a structured questionnaire. Multistage sampling technique was used and data collected was analysed using regression analysis. The findings revealed innovation capability had a significant influence on competitive advantage. The study concluded that there was a need to pay attention to this capability with a capacity to boost the activities of the industry and recommendations were made.*

Key words: *Competitive advantage, insurance, innovation capability*

1. INTRODUCTION

Organisations that aspire to effectively sustain competitive advantage may require an edge to contest on the boundaries of competence and technology. Increased force in terms of product development, ease of imitation and shorter product life cycles amid global market competition, have created immense pressure on organisations' activities geared towards innovation and their bid to sustain competitive advantage (Chadha, 2009). Porter (1985) observed technological change as one of the major catalysts of competition and is

not important for its own sake, but for the effect it has on competitive advantage and industry structure.

According to Iddris (2016), latest changes in technology, market conditions as well as variations in consumer preferences, have shown the importance of innovation capability to an organisation and its survival.

Innovation activity is the ability of an organisation to replicate innovation success and has been recognised as an essential source of competitive advantage (Aas & Breunig, 2017; Calantone, Cavusgil & Zhao, 2002). Innovation helps insurance firms reduce costs, create efficiencies in service delivery, improved flexibility and demand for its products and services (Ernst & Young, 2015). An organisation's ability to change the base of its resource by crafting, incorporating, recombining and discharging same is referred to as capability (Eisenhardt and Martin, 2000).

Breakthrough technologies such as cloud computing, data analytics, telematics, and so on have spurred fundamental transformation of the insurance industry worldwide. These same innovations may perhaps assist the Nigerian insurance industry enhance its product offerings, create better efficiencies in customer experience, minimise the claims processing time amongst others. These technologies are enabling new product creation and business models (IIF, 2016). Whilst these emerging technologies offer opportunities for growth, practitioners with the existing old models, are forced to respond to these changes. The last few decades have seen the rise in globalisation which has shrunk the barriers in trade across continents thus enabling cross continent business and has inadvertently given rise to the need for insurance on a global scale. It is thus expected that companies who want to play within this global space, have capabilities that will enable them offer customers a robust platform to deal with regarding insurance.

The resource-based view theory (RBV) has been repeatedly employed as the base theory in terms of organisational resources usage to attain competitive advantage. When referring to the RBV, it does not relate only to just resources, but also attributes, required by these resources in order to gain competitive advantage (Barney & Clark, 2007). For organisations to gain competitive advantage, they require competitive positioning further than their marginal competitor (Peteraf & Barney, 2003). Organisations' leaders devise strategies and allocate resources to meet goals. To achieve this, technology is required for value creation, seamless integration and improved competitive advantage. The need to pay attention to value creation has risen due to an increased demand for insurance. The focus for organisations may perhaps be delivering products and services to customers that are appropriate and align with their needs as against selling to them existing products, with an obligation to meet these demands through improved capacity while still keeping an eye on making a difference.

Insurance is more or less a promise that cannot be seen, as such, a certain level of guarantee is required and this guarantee is very important to the customer. Mathur and Tripathi (2014) note that the reinforcement of this comfort is further guaranteed when the customers believe and experience differentiated processes, speed and efficiency. Challenges within the insurance sector cut across a plethora of issues from political, regulatory to technological, which all exert pressure thus leading to increased competition with little or no effort to diversify (Rajapathirana & Hui, 2017). Developments and changes in technology, growth within the economy, regulatory frameworks have all started calling for new techniques in the way things are done. While some of these frameworks like recapitalisation, will help create excess liquidity and opportunities for the companies to improve their level

of innovative offering, others such as product guidelines may unwittingly suffocate innovation (Abongo, Mutinda and Otieno, 2019).

In recent times, there have been quite a number of innovations globally in the insurance sector relating to telematics, customer relations management, big data, product innovation, amongst others (Klapkiv & Klapkiv, 2017). The use of big data is to mine information in profiling customers, identifying specific needs, innovating new products to suit these needs and enhancing product offerings. Better efficiencies are also created in customer experience with the use of customer relationship management tools that assist to store information, recognize sales opportunities, record possible service issues and give opportunity for improvement. Another benefit to the insurance industry is the use of telematics in the tracking of vehicles which has become quite common and assist in the recovery of vehicles when stolen. These and many more, may assist the industry create more interconnections, improve risks management and enhance service delivery.

According to Czerniak and Klapkiv (2018), the process of underwriting is a major stage in an insurance contract. Gnatzy and Moser (2012) in a study note that there is a need to have an innovation business model that is designed for insurance, while on the other hand, Maina (2016) observe that there existed very poor implementation within insurance companies of innovation strategies and sound management programs designed towards innovation which could help create competitive advantage. The Nigerian financial sector has quite a number of players within the insurance sub-sector which is the 5th largest insurance market in sub-Saharan Africa with reports however, that contribution to the GDP and penetration rate is still lower than most other African countries (PWC, 2015). The industry Gross Premium income grew from N300bn in 2017 to N413bn in 2018 (NIA, 2018). Despite this improvement, Nigeria still is one of the few countries with such a huge market base to invest less in innovation capability.

According to Porter (1996), organisations who focus more on strategic positioning by executing different tasks than their contenders or comparable jobs in a unique way, rather than operational efficiency as a source of competitive positioning stand a better chance of attaining competitive advantage. Application of strong innovation capability may improve the competitive advantage of insurance companies in Nigeria.

Several studies have shown that innovation capability facilitates competitive advantage, most of these studies have focused on developed economies with a few on the developing economies, some on other sectors of the economy, while others on marketing or claims management (Alsamydai, Alnawas & Yousif, 2010; Bell and Figueiro, 2012; Rajapathirana & Hui, 2017; Anand & Monin, 2013). However, to the knowledge of these researchers, there seems to be very little highlighted on innovation within the underwriting process in the Nigerian insurance industry. To address this gap, this study explores companies within the sector and seeks to examine the relationship between innovation capability and competitive advantage. In line with the above objective, research question was asked and hypothesised thus:

Innovation capability will not influence competitive advantage in the Nigerian insurance industry.

1.1. Theoretical Underpinning

The resource-based view theory underpins this study. It acknowledges organisational resource as a mixture of its assets, information, knowledge, processes and capabilities

within its control and assists in the conceptualisation and implementation of various strategies that have positive impact on the organisation (Kimani & Juma, 2015). The RBV theory as popularised by Wernerfelt (1984) and Barney (1991) is a well-known viewpoint in strategic management which states that every organisation requires specific competences and resources to sustain its competitive position.

Capabilities are complex patterns of skills which the RBV theory explains as a combination of a firm's skills and knowledge accumulated over time (Zhou, Pan & Urban, 2018). Innovation capability may be categorised as an intangible resource. According to Rua, Franca and Ortiz, (2017), intangible assets are considered to be strategic variables and sources of sustainable value.

In line with the RBV theory, it is the conception, ownership, management and distribution of intangibles that expound variations in performance. Resources, where they assist an organisation in improving its performance, may be designated strategic instruments (Massey, 2016). According to Wu, Gu and Zhang (2008) innovation capabilities are the abilities of an organisation to create and transmit to financial practice, new technological potentials that have the ability to respond to the changing environment. RBV theory supposes that a relationship exists between the resources and capabilities being deployed and competitive advantage (Shan, Luo, Zhou & Wei, 2019). It is thus not farfetched, in the context of this study, to say that innovation capability being an intangible asset and acting as an organisational resource, may perhaps be critical to competitive advantage of insurance firms.

1.2. Innovation Capability (IC)

Innovation is a tactical tool for organisations in their quest for competitive advantage and sustainable performance and its significance has repeatedly improved and supported the growth of businesses globally (Chaudhry & Verma, 2016). Zawislak, Fracasso and Tello-Gamarra (2018) describe innovation as a means of competitive advantage and an organisation's ability to absorb, alter and adapt into particular managerial and operational routines, a certain technology that has the capacity to lead the organisation to immense profitability and thus preserve itself. The OECD (2005) describe innovation as the implementation of a process, new method in business practice or marketing, or implementation of a new product or an improved product, goods or service that impacts the organisation or its external relations. Studies on innovation have established that innovation not only is the opening up of new markets, but also the offering of new ways to serve existing and established markets (Lambert & Davidson, 2013). Innovation as a basis of a business's advantage is a consolidated subject in literature, and is viewed as a process dependent on technological capability (Katz, 1984; Desai, 1984; Lall, 1992).

Capability as a concept is not a parameter for performance, but a guide as to the level to which an organisation is prepared in terms of development of innovation forces (Borjesson & Elmquist, 2011). Innovation capability is, therefore, defined as the capacity of an organisation to identify new ideas and transform the same into new products, improved services and processes that are beneficial to the organisation. Rajapathirana and Hui (2018) observed that innovation is more a matter of survival than of competitive advantage. Innovation capability can be described as the capacity of an organisation to identify new ideas and transform the same into new products, improved services and processes that are beneficial to the organisation. It is entrenched with all the strategies and structure that are in support of innovation in an organisation (Gloet & Samson,

2016). An organisation's innovation capability according to Lerro, Linzalone and Schiuma (2014), is the skill to transform continuously ideas and knowledge and convert to new products, processes and systems that are beneficial to the organisation and other stakeholders. The study viewed IC from the perspectives of process, product, market and organizational innovation.

Existing literature has focused on different aspects of innovation and revealed a positive relationship between innovation activities implementation and organisational performance (Slater, Mohr & Sengupta, 2014; Vicente, Abrantes & Teixeira, 2015; Hill, Brandeau, Truelove & Lineback, 2015). IC has been measured as a single dimension, a dichotomous construct and multi-dimensional (Cavusgil, Calantone & Zhao, 2003; Cassiman & Golovko, 2011). This study, while adopting the definition of Lerro et al. (2009), will view innovative capability as a uni-dimensional construct.

1.3. Competitive Advantage (CA)

Competitive advantage is exhibited as positional advantage greater than that of competitors in either marketing or technological know-how which translates into hard-to-imitate innovative products (Hwang, Choi & Shin, 2020). According to Barney and Hesterly (2009), competitive advantage is the ability of a company to produce a higher amount of financial worth than their competitors. The art of creating value for an organisation's client in a more improved manner than the competition is competitive advantage which can be achieved through differentiation, focus and cost strategies (Porter, 2008). Peteraf (1993) noted that an organisation that wishes to have competitive advantage, needs to ensure that resources are properly matched to environmental opportunities.

The RBV premises that an organisation's specific capabilities are warehoused within a single firm and the search for competitive advantage lies therein (Wernerfelt, 1984). With the increase in global rivalry for supremacy by organisations, achieving and sustaining competitive advantage may have become important. Stevenson (2009) noted that the achievement of an organisation in the deployment of its resources in meeting requests of its consumers when compared to their competitors is the measure of competitive advantage. For an organisation to effectively sustain competitive advantage, they need to look inwards to their unique resources and capabilities that they have control over and ensure that the foundation of its advantage is not easily duplicated by competitors (Chen, Zhou, Zhou & Xue, 2017; Mahdi, Nassar & Almsafir, 2018). Several extant literatures have defined and measured competitive advantage in various ways (Christensen & Fahey, 1984; Stevenson, 2009), this study while reviewing as a uni-dimensional construct, will adopt the definition of Li, Ragu-Nathan, Ragu-Nathan and Rao (2006).

1.4. Innovation Capability and the Insurance Industry Competitive Advantage

Insurance the world over has grown over time, Africa and Nigeria not excluded. Underlying drivers of change within the insurance industry across the world and great moves by insurtechs have seen the rise of innovations towards customer experience (IIF, 2016). In recent times the insurance sector globally with the evolution of digital platforms, has focused quite heavily on technology solutions to enhance their services ranging from automated processes, use of telematics in motor insurance, building systems, and more recently, applying artificial intelligence and application programming interface in their activities.

In an increasing competitive global market, organisations within this sector have come to understand innovation as an important component in the drive to meet up with the constantly changing demographic and technological competition. Empirical evidence suggests that innovation thrives in a setting where frontrunners are able to see possible connections, spot chances and take advantage of them. And that organisations who imbibe innovation along with the application of better productive process, tend to have a more positive customer perception (Gundaya, Ulousoy, Kilic & Alpkan, 2011). Economists concur that innovation is responsible for a considerable proportion of progress in the profitability of organisations and that innovation thrives in a setting where frontrunners are able to see possible connections, spot chances and take advantage of them (Chatterji, Glaeser & Kerr, 2013; Abongo, Mutinda & Otieno, 2019). According to AIO (2018) the insurance industry across the globe has grown tremendously and in Africa, this growth is led by South Africa at an outstanding 14.2%, Kenya at 3% and Ghana at 1%. While the sector is plagued with a few challenges as earlier highlighted, some repeatedly identified ones are low penetration, poor awareness, low trust level, and regulatory stiffness amongst others. To drive the continued growth of the sector, and help address the perennial problems, there is a need to embrace innovation.

According to Lambert and Davidson (2013), a business model is a value creation design by which an insurance company creates value to its clients and is embedded with delivery mechanism. Most companies use a model by which they are well identified and this signifies the type of offerings they provide to their clients The business models are often very rigid in nature, constraining and prolonged in terms of response times (McGrath, 2010). Traditional models of agency, brokerage and bancassurance in the distribution and marketing of their products and services have been the norm (Gera, Costonis, Sandquist, Bramblet & Secchi, 2018) while the Nigerian model is commonly a revenue generation focused one where sale of policy is the major indicator (Mudaly, 2017). However, of recent, the enforcement of a customer-based approach has been included in the market conduct and business practice guidelines for insurance institutions issued by the National Insurance Commission (NAICOM).

Though these still constitute the major channels in most insurance markets, the recent entry of insurtechs start-ups providing insurance services are turning the tide of these old models (Kottmann & Dordrechter, 2018). The market is a maze of forever changing scenarios across all fronts and this requires companies to implement very high-level expertise in its structure and service delivery (Ernst and Young, 2015). Though there have been some great signs of improvement within the Nigerian insurance sector, in comparison to other emerging markets, there is still room for more with a current penetration level of 0.3% in comparison to other countries both overseas and in Africa (AIO, 2018).

Insurance as a risk transfer mechanism is a socially germane to the economy, Innovation is an indispensable driver of change. The sector globally, is gradually embracing recent technological developments (Schanz, & Sommerrock, 2016). The penetration rate and GDP contribution is quite discouraging and in need for a shift, with discussions around technology and innovation. The AIO (2018) noted that increasing levels of literacy and wealth have created a new middle class across the continent, thus increasing the pool of insurable assets and lives of an untapped potential market. They further noted that to increase penetration, strategies need to move from the previous inward looking to a more outer focused view through implementation of technology such as data analytics, internet of things and so on. Innovative technology will further aid ease of interaction and access to customers as well as client profiling, product design and quality service (Lehman, 2017).

1.5. Empirical Review

Rajapathirana and Hui (2018) explored innovation capability and innovation type and their relationship to firm performance in Sri Lanka. The results observed that IC and innovation efforts had significant relationship with firm performance and concluded that effective innovation capability had the ability to assist insurance companies deliver more effective innovation outcomes. The study recommended that insurance companies apply this as it would help improve firm performance.

Czerniak and Klapkiv (2018) explored characterization of innovations of underwriting patents in the insurance market. The study results showed that there was no significant correlation between premium growth and the number of patents from the industry. The study concluded that innovation activity was a strong propellant for the insurance industry development. Hwang, Choi and Shin (2020) investigated the mediating role of IC in the relationship between individual level entrepreneurial competencies (EC) and firm level CA in Korean SMEs. Results showed that a stronger relationship between the indirect effects of EC through organisational innovation capabilities and CA than the direct effects. The study concluded that improving IC was beneficial and increased competitiveness. It recommended that firms should inculcate innovation capability to sustain their superior position. Abd Aziz and Samad (2016) investigated the moderating effect of firm age on the relationship between innovation and competitive advantage. Using SMEs in the food sector of Malaysia, the study employed structural equation modelling to test the hypotheses. The results revealed that innovation had a significant impact on competitive advantage and that firm age moderated the relationship. The study recommended that SMEs should invest in innovation to gain competitive advantage.

Waseem, Loo-See, Adeel and Riaz (2018) examined innovation capability and organisational performance. The study observed that intellectual capital dimensions portrayed substantial affirmative direct and indirect effects on organisational performance unlike structural capital which was found to be insignificant. Abongo et al., (2019) investigated the application of knowledge and skills in marketing of insurance in Kenya as an exchange driver and found that innovation capabilities were largely directed towards external rather than internal developments. The study also noted that these components had no impact on performance and that the companies had failed to sustain the innovative capabilities for the future. Anand and Monin (2013) reviewed innovation processes in emerging markets. The study identified seven processes and postulated that these processes responded to some basic issues in respect of values, affordability and accessibility. The study concluded that while these issues may also be existent in developed markets, they have higher prevalence in emerging markets and recommended that innovation processes within the service space aim to trigger evolution and development. Klapkiv and Klapkiv (2017) investigated technological innovations in the insurance industry. The study viewed the use of technologies such as analytics, algorithms and sensors within the industry and its effect on the value chain. Results showed that most of the processes within the value chain had been impacted by technology and concluded that whilst it had a positive impact, there might be increased dependence of the industry on other sectors.

Lerro, Linzalone and Schiuma (2014) who viewed described innovation capability from the perspectives of process, product, market and organisational innovation, investigated the role of innovation, intellectual capital, performance improvement and competitive advantage on an organisation's value creation. The study found certain assumptions and concluded that

there was a need for organisations to have dynamic capabilities and also engage the passion and imagination of the public through innovation.

Hagen, Hales, Reifel, Pei and Miller (2009) examined technology as the pivot point in the insurance industry and noted five main trends currently and that technology could play a role in transforming the insurance industry. The findings of the study suggested that social networking, telematics and service-oriented architectures are essential to the growth of a competitive advantage. Chaudhry and Verma (2016) reviewed innovation capability evaluation factors and observed that innovation capability had a strong influence on an organization's competency and its capacity to sustain competitive advantage.

2. METHODOLOGY OF RESEARCH

The study employed a cross-sectional survey design. The study populace consisted of employees of 55 registered insurance companies in Nigeria across license types (composite business, general business and life business) (NAICOM, 2019). The target population consisted of employees of selected companies.

The sample size for this study was total enumeration of professional staff of the selected insurance companies. Stratified random sampling was used in the first stage to select thirty companies ranked into three groups comprising top ten from each license type (composite business, general business and life business). The justification for selection was based on organisations with operational location in Lagos State and shareholders' funds. These criteria are assumed to provide good liquidity base and act as a possible edge in terms of innovation. Six (6) companies were randomly selected in the second stage, with two companies from each of the group to ensure sufficient variance within the sample. Sample size of 150 was derived employing convenience sampling of 25 employees of each organisation. Data collection instrument was a questionnaire.

Measures for the study variables were adapted from extant literature. Measurement for innovation capability which was defined as the skill to continuously transform ideas and knowledge and convert the same to new products and systems beneficial to all stakeholders (Lerro et al., 2014), was done with ten items; competitive advantage defined as an organisation's advantage to outperform its competitors (Kay, 1994), was measured using nine items. The Cronbach Alpha for all scales were above the accepted range ($\alpha = 0.70$). Responses were elicited by means of a 5-point Likert scale of strongly agree at "5" to strongly disagree at "1". In line with the study of Wu & Chen, (2011) the study included size of the company and age as control variables in order to account for the effects of extraneous variables. The organisation size was calculated using the employee size while the age was calculated using number of years of operation. The study employed regression analysis to test the relationship between the study variables.

3. RESULTS

Table 1 report the descriptive statistics, the mean and standard deviation of the study constructs.

Table 1 shows the result of the descriptive statistics of the study constructs. The results show that innovation capability items had a mean score range of between 3.50 and 4.17 with a reasonable distribution about this central tendency measure, while the standard

deviation had a range between 0.722 and 1.007. The study results further show that competitive advantage measured with 9 items had a mean score range between 3.52 and 3.95 while the standard deviation had a range of between 0.754 and 0.954.

Table 1 Descriptive Statistics

Constructs	Items	Questionnaire description	Mean	SD
INNOVATION CAPABILITY				
	1	In my organisation, there is adequate funding for innovation	4.01	.995
	2	There is a culture of risk taking in my organisation	3.50	1.007
	3	There is a process for innovation in my organisation	3.75	1.018
	4	In my organisation, there is recognition for innovation	3.70	.938
	5	There is development of new product, services and solutions in my organization	3.91	1.001
	6	My organisation, successfully introduced and implemented innovative products/services	3.72	.933
	7	Our market ranking has increased due to innovation in my organization	4.17	.760
	8	There is a target for innovation in my organisation	3.75	1.007
	9	We have savings due to our innovation strategy in my organization	3.59	.912
	10	In my organisation, our research and innovation management is excellent	3.66	.722
COMPETITIVE ADVANTAGE				
	1	The prices of our products are competitive in my organization	3.59	.954
	2	In my organisation, our products/services have unique benefits than those of our competitors	3.52	.906
	3	Our products/services are of superior quality than those of our competitors / We are able to compete based on quality	3.57	.800
	4	Our services are more advanced than those of competitors in the same market in my organisation	3.89	.851
	5	In my organisation, our products/services are customized to meet clients' need	3.94	.765
	6	We respond to clients' request for new features in my organization	3.62	.887
	7	In my organisation, our products/services are delivered on time and to specification	3.74	.754
	8	In my organisation, our products/services are the first to hit the market	3.53	.861
	9	Our product development is fast in my organisation	3.95	.809

Note: N=103; The summary statistics are reported.

To effectively measure the relationship between the study variables in response to our hypothesis, the study employed regression analysis to test the relationship and the result is as shown in Table 2 and 3 below.

Table 2 shows the result of the study when the control variables of firm size and firm age are treated. The result indicates that firm size and firm age when controlled, have no significant relationship with competitive advantage (F=2.921; p>.05). The results further show that (R^2=0.055), thus indicating that the control variables account for only 1% variation in competitive advantage.

Table 2 Model Summary

Model	R	R^2	Adjusted R^2	Std Error of the Estimate
	0.235	0.055	0.036	4.972
ANOVA[a]				
Model	Sum of Squares	Df	Mean Square	F	Sig.
Regression	144.393	2	72.197	2.921	.058
Residual	2471.588	100	24.716		
Total	2615.981	102			
Coefficients[a]					
	Unstandardized Coefficients		Std Coeff		
Model	B	Std Error	Beta	t	Sig.
(Constant)	38.703	2.290		16.900	.000
Firm size	-.200	.469	-.048	-.426	.671
Firm age	-1.428	.782	-.207	-1.827	.071

Predictors: Firm size, Firm age
Dependent Variable: Competitive Advantage
Source: Field Survey 2020

Table 3 shows the regression results when the study variables firm size, firm age and innovation capability are treated.

Table 3 Model Summary

Model	R	R^2	Adjusted R^2	Std Error of the Estimate
	0.628	0.395	0.377	3.998
ANOVA[a]				
Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	1033.272	3	344.424	21.544	.000[b]
Residual	1582.708	99	15.987		
Total	2615.981	102			
Coefficients[a]					
	Unstandardized Coefficients		Std Coeff		
Model	B	Std Error	Beta	t	Sig.
Firm size	-.303	.378	-.073	-.801	.425
Firm age	-.742	.635	-.107	-1.168	.246
Innovation Capability	.515	.069	.590	7.457	.000

Dependent Variable: Competitive Advantage
Source: Field Survey 2020

The results in the model summary show correlation amongst the study variables with R= 0.628. The result also shows that $R^2 = 0.395$, suggesting that 39.5% variation in competitive advantage is due to innovation capability, firm size and age with an unexplained 61.5% outside the scope of this study. The Anova results reflect a positive significant relationship between the study variables in the model (F=21.544; p<.05). Furthermore, the results show that innovation capability makes a significant contribution to the overall model ($\beta = 0.590$; t = 7.457, p < 0.05). Therefore, the null hypothesis cannot be accepted, meaning that innovation capability does have an influence on competitive advantage after controlling for firm size and firm age.

4. DISCUSSION AND CONCLUSION

Using selected companies with emphasis on the underwriting process within the companies, this study examined the relationship between innovation capability and competitive advantage in Nigerian insurance companies. The study results showed that innovation capability had a significant relationship with competitive advantage. The results align with the study of Chaudhry and Verma (2016) who notes that innovative organisations tend to do better than organisations that are not very innovative. The study findings in relation to the contributions of innovation capability on competitive advantage corroborate existing studies (Hagen, Hales, Reifel, Pei & Miller, 2009; Czerniak & Klapkiv, 2018) who observe that technology played a pivotal role in the required transformation and development of the insurance industry market. The results also agree with the study of Klapkiv and Klapkiv (2017) that processes within the insurance value chain had been impacted by technology.

However, the result of this study extends the literature on innovation capability further, with emphasis on the underwriting process within the context of developing countries like Nigeria. A very critical component of the insurance contract is the underwriting process. And one of the important areas within this process is the risk assessment and premium rating. Having correct estimates from these two activities assists the organisation derive accurate results in terms of premiums charged, which translates to income which in turn aids financial stability. This study results in general highlight the importance of innovation capability and the need for the insurance industry in Nigeria to take cognizance of this capability and its potential to assist achieve competitive advantage. The study concludes that insurance companies maximizing this potential focus on improving and harnessing the same across all operational functions rather than engage in a rate cutting war.

The study results however, should be considered in light of some limitations which include the number of participants and the use of cross-sectional survey. Expansion in the participants and a longitudinal research is suggested for future studies. This study looked at underwriting process, deviating from the usual focus on product and customer. Future studies should look at other units within the organisation. The population of the study was limited. It is thus suggested that the population be expanded for future research.

The study contributes to existing literature on innovation capability and competitive advantage by considering the impact on underwriting process within the insurance market in Nigeria. The policy implication is that global market and customers are aligning more with industry players who have the capability to meet the requirements, immediate interactions and play internationally, which are only enabled through innovative techniques. Thus, insurance firms should harness available growth in the technology space as this is a key factor to their ability to sustain competitive advantage. The study further recommended that insurance companies in Nigeria understand and exploit intangibles as valuables, especially in terms of innovation capability, identify and harness innovation capability through adequate investments and align with market shifts to help improve competitive advantage.

REFERENCES

Aas, T. H., & Breunig, K. J. (Eds.). (2017). *Innovation Capabilities: Affirming an Oxymoron?*. Fundation for the Dissemination of Knowledge and Science" Cognitione".

Abongo, B., Mutinda, R., & Otieno, G. (2019). Innovation capabilities and process design for business model transformation in Kenyan insurance companies: A service dominant logic paradigm. *Journal of Information*

and Technology, *3*(1), 15-45. Retrieved from http://stratfordjournals.org/journals/index.php/Journal-of-Information-and-Techn/article/view/297

AIO. (2018). Africa insurance barometer: Market survey. *African Insurance Organisation*, 1-49. Retrieved from: http://turnkeyafrica.com/africa-insurance-barometer-2018/

Alsamydai, M. J., Alnawas, I. A. M., & Yousif, R. A. (2010). The impact of marketing innovation on creating sustainable competitive advantage: The case of private commercial banks in Jordan. *Asian Journal of Marketing, 4,* 113-130.

Anand, M., & Monin, P. (2013). Innovation processes in emerging markets: Empirical evidence from the Indian insurance industry. *Management International, 18*(1), 40-57.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of management*, *17*(1), 99-120.

Barney, J. B., & Clark, D. N. (2007). *Resource-based theory: Creating and sustaining competitive advantage* (1st ed.) Oxford University Press on Demand

Barney, J. B., & Hesterly, W. S. (2009). *Strategic management and competitive advantage.* Upper Saddle River, NJ: Pearson Education.

Bell, M., & Figueiredo, P. N. (2012). Innovation capability building and learning mechanisms in latecomer firms: Recent empirical contributions and implications for research. *Canadian Journal of Development Studies/Revue canadienne d'études du développement*, *33*(1), 14-40.

Calantone, R. J., Cavusgil, S. T., & Zhao, Y. (2002). Learning orientation, firm innovation capability, and firm performance. *Industrial marketing management*, *31*(6), 515-524.

Cassiman, B., & Golovko, E. (2011). Innovation and internalization through exports. *Journal of International Business Studies, 42*(1), 56-75.

Cavusgil, S. T., Calantone, R. J., & Zhao, Y. (2003). Tacit knowledge transfer and firm innovation capability. *Journal of Business & Industrial Marketing, 18*(1), 6-21.

Chadha, A. (2009). Product cycles, innovation and exports: A study of Indian pharmaceuticals. *World Development, 37*(9), 1478-1483.

Chatterji, A., Glaeser, E., & Kerr, W. (2014). Clusters of entrepreneurship and innovation. *Innovation policy and the economy*, *14*(1), 129-166.

Chaudhry, B., & Verma, P. K. (2016). Technological innovation capabilities: a critical review. *International Journal of Latest Technology in Engineering, Management & Applied Science*, *5*(4), 95-101.

Chen, L., Zhou, Y., Zhou, D., & Xue, L. (2017). Clustering enterprises into eco-industrial parks: Can interfirm alliances help small and medium-sized enterprises?. *Journal of Cleaner Production, 168*, 1070-1079.

Christensen, K. & Fahey, L. (1984). Building distinctive competencies into competitive advantage. *Strategic Planning Management, 2*(1), 113-123.

Czerniak, J., & Klapkiv, L. (2018). Innovations on the insurance market: Patents in underwriting. *Handel Wewnetrzny, 6*, 151-163.

Desai, A. V. (1984). India's technological capability: An analysis of its achievements and limits. *Research Policy, 13*(5), 303-310. Retrieved from: https://ideas.repec.org/a/eee/respol/v13y1984i5p303-310.html

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they?. *Strategic Management Journal*, *21*(10-11), 1105-1121.

Ernst & Young (2015). *Sustaining growth and innovation in the insurance sector.* Insurance Governance Leadership Network Summit, November 2015. Retrieved from https://www.ey.com/Publication/vwLUAssets/EY-igln-summit-sustaining-growth-and-innovation/$FILE/EY-IGLN-Summit-Sustaining-growth-and-innovation.pdf

Gera, P., Costonis, M., Sandquist, E., Bramblet, J., & Secchi, A. (2018). Insurance as a living business: Explosive growth. *Accenture.* Retrieved from https://www.accenture.com/t00010101T000000Z__w__/ar-es/_acnmedia/PDF-72/Accenture-Insurance-Living-Business-POV.pdf

Gloet, M., & Samson, D. (2016). Knowledge management and systematic innovation capability. *International Journal of Knowledge Management, 12*(2), 54-72.

Gnatzy, T., & Moser, R. (2012). Scenario development for an evolving health insurance industry in rural India: INPUT for business model innovation. *Technological Forecasting and Social Change*, *79*(4), 688-699.

Gunday, G., Ulusoy, G., Kilic, K., & Alpkan, L. (2011). Effects of innovation types on firm performance. *International Journal of production economics*, *133*(2), 662-676.

Hagen, C., Hales, M., Reifel, J., Pei, A., & Miller, J. (2009). Technology: the insurance industry's pivot point. *A.T. KEARNEY, 1*-7.

Hill, L. A., Brandeau, G., Truelove, E., & Lineback, K. (2015). The capabilities your organization needs to sustain innovation. *Harvard Business Review*, 1-9.

Hwang, W., Choi, H., & Shin, J. (2020). A mediating role of innovation capability between entrepreneurial competencies and competitive advantage. *Technology Analysis & Strategic Management, 32*(1), 1-14.

Iddris, F. (2016). Innovation capability: A systematic review and research agenda. *Interdisciplinary Journal of Information, Knowledge, and Management*, *11*, 235-260.

IIF. (2016). *Innovation in insurance: How technology is changing the industry*. Institute of International Finance, pp. 1-12. Retrieved from https://www.iif.com/Publications/ID/1246/Innovation-in-Insurance-How-Technology-is-Changing-the-Industry

Katz, J. (1984). Domestic technological innovations and dynamic comparative advantage. *Journal of Development Economics, 16*(1/2), 13-37.

Kimani, M., & Juma, W. (2015). Assessment of strategies for gaining competitive advantage in insurance firms in Thika Town, Kenya. *International Journal of Economics, Commerce and Management*, *3*(2), 1305-1316.

Klapkiv, L., & Klapkiv, J. (2017). Technological innovations in the insurance industry. *Journal of Insurance, Financial Markets and Consumer Protection, 26*(4/2017), 67-78.

Kottmann, D., & Dordrechter, N. (2018). Insurtech caught on the radar hype or the next frontier? *Oliver Wyman/Policendirekt, Global Insurtech Report*, 1-52. Retrieved from https://www.oliverwyman.com/our-expertise/insights/2017/may/InsurTech-caught-on-the-radar.html

Lall, S. (1992). Technological capabilities and industrialization. *World development*, *20*(2), 165-186.

Lambert, S. C., & Davidson, R. A. (2013). Applications of the business model in studies of enterprise success, innovation and classification: An analysis of empirical research from 1996 to 2010. *European management journal*, *31*(6), 668-681.

Lehman, M. (2017). How will insurtech influence claims?. *Accenture Insurance Blog,* 1-4. Retrieved from insuranceblog.accenture.com

Lerro, A., Linzalone, R., & Schiuma, G. (2014). Managing intellectual capital dimensions for organizational value creation. *Journal of Intellectual Capital*, *15*(3), 350-361.

Li, S., Ragu-Nathan, B., Ragu-Nathan, T. S., & Rao, S. S. (2006). The impact of supply chain management practices on competitive advantage and organizational performance, *Omega*, *34*(2), 107-124.

Mahdi, O. R., Nassar, I. A., & Almsafir, M. K. (2019). Knowledge management processes and sustainable competitive advantage: An empirical examination in private universities. *Journal of Business Research, 94*, 320-334.

Maina, M. N. (2016). *Effect of innovation strategies on the performance of insurance firms in Kenya*. Unpublished Master of Science in Entrepreneurship and Innovations Management thesis, University of Nairobi. UON Repository.

Massey, B. L. (2016). Resource-based analysis of the survival of independent web-native news ventures. *Journalism & Mass Communication Quarterly*, *93*(4), 770-788.

Mathur, D., & Tripathi, A. S. H. I. S. H. (2014). Factors influencing customer's choice for insurance companies: A study of Ajmer City. *IOSR Journal of Business Management*, *16*(2), 35-43.

McGrath, R. G. (2010). Business models: A discovery driven approach. *Long range planning*, *43*(2-3), 247-261.

Mudaly, N. A. (2017). Business model innovation: the drivers, enablers and inhibitors of firms facing disruptive change (Doctoral dissertation, University of Pretoria). Retrieved from: URI: http://hdl.handle.net/2263/59891

National Insurance Commission (NAICOM) 2019. Retrieved from https://www.naicom.gov.ng

NIA (2019). 2018 Nigeria Insurers Digest

OECD. (2005). *OECD science, technology and industry scoreboard*, OECD, Paris. Retrieved from https://www.oecd.org/sti/scoreboard.htm

Peteraf, M. A. (1993). The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal*, *14*(3), 179-191.

Peteraf, M. A., & Barney, J. B. (2003). Unravelling the resource-based tangle. *Managerial and Decisions Economics, 24*(4), 309-323.

Porter, M. E. (1985). Technology and competitive advantage. *Journal of Business Strategy, 5*(3), 60-78.

Porter, M. (1996). What is strategy? *Harvard Business Review*, *74*(6), 61-78.

PwC (2015). *Africa insurance trends*. Retrieved from https://www.pwc.com/ng/en/assets/pdf/nigeria-insurance-survey.pdf

Rajapathirana, R. J., & Hui, Y. (2018). Relationship between innovation capability, innovation type, and firm performance. *Journal of Innovation & Knowledge*, *3*(1), 44-55.

Rua, O., França, A., & Ortiz, R. F. (2018). Key drivers of SMEs export performance: the mediating effect of competitive advantage. *Journal of Knowledge Management*.

Schanz, K. U., & Sommerrock, F. (2016). Harnessing technology to narrow the insurance protection gap. *Report/The Geneva Association: Zurich*. Retrieved from https://www.genevaassociation.org/sites/default/files/research-topics-document-type/pdf_public/harnessing-technology-to-narrow-the-insurance-protection-gap.pdf

Shan, S., Luo, Y., Zhou, Y., & Wei, Y. (2019). Big data analysis adaptation and enterprises' competitive advantages: the perspective of dynamic capability and resource-based theories. *Technology Analysis & Strategic Management, 31*(4), 406-420.

Slater, S. F., Mohr, J. J., & Sengupta, S. (2014). Radical product innovation capability: Literature review, synthesis, and illustrative research propositions. *Journal of Product Innovation Management*, *31*(3), 552-566.

Stevenson, W. J. (2009). *Operations management* (10th ed,). McGraw-Hill Irwin Higher Education.

Vicente, M., Abrantes, J. L., & Teixeira, M. S. (2015). Measuring innovation capability in exporting firms: The INNOVSCALE. *International Marketing Review*, *32*, 29-51.

Waseem, B., Loo-See, B., Adeel, A., & Riaz, A. (2018). Impact of intellectual capital on innovation capability and organizational performance: An empirical investigation. *Serbian Journal of Management*, *13*(2), 365-379.

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic management journal*, *5*(2), 171-180.

Wu, X., Gu, Z., & Zhang, W. (2008). The construction of innovation networks and the development of technological capabilities of industrial clusters in China. *International Journal of Innovation and Technology Management*, *5*(02), 179-199.

Wu, J., & Chen, X. (2011). Leaders social ties, knowledge acquisition capability and firm competitive advantage. *Asia Pacific Journal of Management, 29*(2), 331-350.

Zawislak, P. A., Fracasso, E. M., & Tello-Gamarra, J. (2018). Technological intensity and innovation capability in industrial firms. *Innovation & Management Review*, *15*(2), 189-208.

Zhou, Y., Pan, M., & Urban, F. (2018). Comparing the international knowledge flow of China's wind and solar photovoltaic (PV) industries: Patent analysis and implications for sustainable development. *Sustainability*, *10*(6), 1883.

KONKURENTNA PREDNOST NIGERIJSKE INDUSTRIJE OSIGURANJA: ULOGA INOVACIONE SPOSOBNOSTI

Nove tehnologije su omogućile veće mogućnosti u okviru sektora osiguranja, gde su inovacije preokrenule tradicionalni biznis model. Ovaj rad proučava konkurentnu prednost i njen odnos sa sposobnošću za inovacije nigerijskog sektora osiguranja. Uz pomoć međusektorskog dizajna istraživanja, studija koristi primarne podatke dobijene od odabranih osiguravajućih kompanija uz pomoć strukturisanog upitnika. Korišćena je tehnika višefaznog uzrokovanja, a dobijeni podaci su analizirani uz pomoć regresivne analize. Rezultati su pokazali da je sposobnost za inovacije imala značajnog uticaja na komparativnu prednost. Studija zaključuje da postoji potreba da se obrati pažnja na ovu sposobnost sa mogućnošću da se pojačaju aktivnosti industriuje i date su preporuke.

Ključne reči: *konkurentna prednost, osiguranje, inovaciona sposobnost*

Original Scientific Paper

ARE COSTS STICKY? EVIDENCE FROM SERBIA

UDC 657.471.1(497.11)

Jovana Jugović

University of Belgrade, Faculty of Economics, Belgrade, Serbia

Abstract. *This paper is focused on the theory of sticky costs, created out of researches which pointed to the fact that costs do not act symmetrically in the case of equivalent increase and decrease of the activity volume, as it is implied by the traditional cost theory. Deliberate business decisions, the ones made in order to increase company's value, as well as opportunistic decisions aimed at the realization of managers' personal goals are found as some of essential causes of cost stickiness. In order to examine the phenomenon of stickiness in the cost behavior of companies that operate in Serbia, we conducted a research on a sample of 917 medium and large companies from manufacturing sector for the period 2007 – 2016. The analysis of panel data pointed to the presence of stickiness in the behavior of operating costs - it showed that they grow by 0.847% as revenues grow by 1%, and they fall by 0.718 % due to 1% drop in revenues. We also found a lagged adjustment to operating costs for changes in operating revenues and partial reversal of stickiness in the period after a revenue decrease.*

Key words: *cost behavior, cost stickiness, adjustments costs, behavioral finance, agency theory*

JEL Classification: M41, D24, G40

INTRODUCTION

Cost behavior is the way costs respond to the changes in cost drivers. Efficient cost management assumes understanding of the cost functions because of cost planning, calculation and control. Cost management is imposed as a complex task, having in mind that different purposes of using information about the costs demand cost categorization to be done in different ways. Managers need to make many decisions based precisely on the information about the costs: whether to introduce a new product in the assortment or eliminate some of the existing ones; produce or buy at the external market; sell or process

further and the like. The analysis of cost behavior is supposed to be put in the context of how the resource consumption changes along with the change of activity volume.

This paper is organized as follows. The basic ideas of sticky cost theory are described in section 1 and compared with the traditional one. Section 2 identifies three groups of factors that determine business decisions whose realization causes cost stickiness: adjustment costs, managers' personal benefits and behavioral factors. The results of panel analysis, conducted with the aim to test cost stickiness in operating costs of manufacturing companies in Serbia, are presented in section 3, as well as author's interpretation of the results.

1. COST BEHAVIOR: TRADITIONAL AND STICKY COST THEORY

Traditional cost theory classifies costs in two groups, depending on the way costs, in the overall amount, react to the change of activity volume: fixed and variable costs. Unlike fixed, whose total amount does not change in the short term, the total variable costs change not only linearly, but also in proportion to the change in the activity volume. Reaction of variable costs to the change of activity level is symmetrical in both directions, which means that this category reacts uniformly both to the increase and equivalent decrease of the activity volume. Thus, direction and volume of the change of activity are irrelevant for the change of costs. The magnitude of a change in costs will be the same, regardless of whether the activity goes from lower to upper volume or vice-versa, as well as whether it is about a relatively small or relatively big change of the activity volume. Besides these two basic groups, there is also the third one – mixed costs which, due to the effects of economies of scale, learning curves and the like, show characteristics of both previous groups. Still, "these extensions continue to assume a mechanistic relation between cost driver quantities and costs, leaving no room for deliberate managerial decisions motivated by various economic constraints, incentives, and biases" (Banker et al., 2018, p. 192-193). Upon categorizing costs to fixed and variable ones it is necessary to bear in mind that this classification is conditioned by a relevant range of activities, and it was done in relation to the time period being observed. Curve of overall costs shows symmetrical behavior of costs. Traditional cost accounting theory, as opposed to economic theory, presents total costs as a linear function of the activity volume, for the relevant range of activities. The reason for this is that the cost curve is almost linear in any relevant range of activities, as shown in Figure 1, so the linearity assumption is used for the purposes of short-term planning, calculation and control of costs.
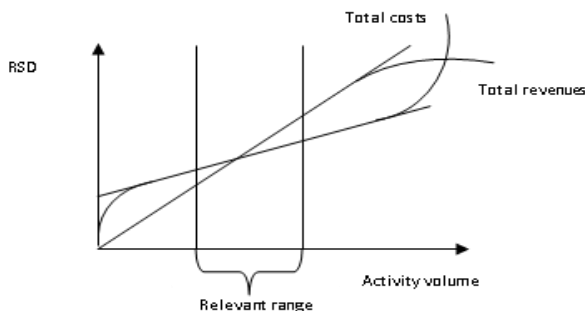
Fig. 1 Cost and revenue function
*Source*: Malinić et al., p. 324

Some empirical studies, however, showed that costs do not act symmetrically, as traditional theory claims. The results of these studies pointed out the fact that costs react more to the increase than to equivalent decrease of the activity volume. If demand exceeds the available capacity of existing resources, it will lead to the increase of capacity by acquiring additional resources, which is aimed at absorbing greater demand. However, in case unused capacities exist, "some committed resources will not be utilized unless managers make the deliberate decision to remove them" (Anderson et al., 2003, p. 49). Costs will be fixed only if the management does not take any action regarding the reduction of unused capacities. This can be an indicator of active cost management. The first empirical confirmation of this idea comes from the research of Noreen & Soderstrom (1997). The authors showed that costs adjust more difficultly to the fall than to the increase of the activity volume, pointing to a different pattern of cost behavior as determined by the direction of the change of the activity. Also, Hamerman and Pfann (1996) stressed there is no necessary reason why the marginal cost of increasing activity would be the same as that of an equal-size decrease.

However, this idea gains importance after the publishing of the research of Anderson et al. (2003) which opened 'the black box' about the cost behavior and motivated research in the field. Authors used the sample of 7,629 industrial companies, observing the period of 20 years (1979 – 1998) and showed that sales, general and administrative costs act asymmetrically. In other words, these costs grow by 0.55% on average upon the increase of sales by 1%, and on average they decrease by only 0.35% in case of 1% fall in sales. This kind of cost behavior is named „sticky" by the authors. "Costs are sticky if the magnitude of the increase in costs associated with an increase in volume is greater than the magnitude of the decrease in costs associated with an equivalent decrease in volume" (Anderson et al., 2003, 48). So that, according to sticky cost theory, costs are determined both by the change of volume and the direction of activities (Banker & Byzalov, 2014), unlike traditional cost theory claims, which can be presented:

$$\Delta C = f (\Delta q, \text{ direction of change q}), \qquad (1)$$

where:
C – amount of overall costs of a certain category,
q – activity volume.[2]
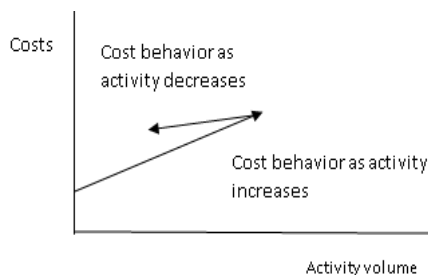
Fig. 2 Sticky costs
Source: Maher, Stickey, Weil (2008), p. 160

---

[2] It would be important to stress that the unlinearity in the behavior of costs is not an equivalent to sticky cost behavior. In case of nonlinear behavior, the same as with linear behavior, marginal costs are determined exclusively by the change of the activity volume, not by the change of direction, while in the case of sticky cost behavior marginal costs were determined both by the change of volume and change of the direction of activity. More to be seen in Banker and Bzyalov (2014).

Anderson et al. (2003) consider deliberate decision making on resource adjustments as the primary cause of stickiness in the cost behavior. Cost stickiness is just one of many ways how managerial decisions may affect cost behavior (Banker et al., 2018). Numerous researches also highlight the importance of the managers' characteristics and behavior for cost stickiness (Kama & Weiss, 2010; Chen et al. 2012; Dierynck et al. 2012; Qin et al. 2015; Salamah & Abulezz, 2017).

Taking into account time component, as numerous authors have shown (e.g. Andersen et al. 2003; Dalla Via & Perego 2014), delayed decision-making and contracting lags cause cost stickiness observed in one period to be reversed in subsequent periods. Changes in revenues may reflect current market conditions and may be an indicator of long-term orientation in terms of demand for the company's products. Therefore, if there is a drop in demand, managers will need some time to gather information on the basis of which they will conclude about the cause of the fall and make decisions about reducing resources. In addition, even after a decision has been made to reduce resources, it often takes some time for that decision to reflect on cost reduction since contractual arrangements take time to be undone. These time lags will result in cost stickiness bearing in mind that companies will maintain unused resources during the period between the reduction in volume and the adjustment decision. The consequence of this will be that the stickiness observed in one period might be reverted in subsequent periods.

2. DELIBERATE DECISION MAKING AS CAUSE OF COST STICKINESS

Decisions of the business entity may be a factor of cost variability (Malinić et al., 2019, p. 219). "Managers choose resource levels subject to various constraints (e.g., demand conditions, production technology, resource adjustment costs, strength of corporate governance, debt covenants, government regulation), incentives (e.g., performance compensation, earnings targets, ownership type, stakeholder activism), and biases (e.g., overconfidence)" (Banker et al., 2018, p. 195). Managers choose different resource levels for the same activity volume depending on whether it is the upward or downward resource adjustment (Anderson et al, 2003). On that occasion, managers make those decisions whose realization they expect to produce the increase of company's value, as well as the ones which will bring them the fruition of personal benefits, while neglecting the effects that the realization of those decisions could have for the company. We identify three groups of factors that determine business decisions whose realization causes cost stickiness: adjustment costs, managers' personal benefits and behavioral factors.

Companies change their demand for resources more slowly than shocks require (Hamermesh & Pfann,1996, p. 1264.) Explanation of this phenomenon lies in adjustment cost, concept which was first presented by Lucas (1967). In case of external shocks, caused by the change of demand, it is impossible to adapt resources to changed circumstances regarding the need for them, without following cost adjustments. Dismissal, as well as the employment of additional units of resources, entails certain adjustment costs. These can be explicit costs, expressed in monetary units, such as severance costs, costs of organizing vacancies, costs of training and education of new employees, disposal costs, equipment installation costs. However, implicit costs that are not covered by the accounting information system, such as the decline in employee morale due to dismissals, should not be overlooked. These costs may be reflected in the

financial statements of some future periods, for example, by increasing costs due to reduced efficiency and productivity.

Cost-benefit analysis is based on the comparison of costs of maintaining unused capacities and adjustment costs. In case of the drop in demand, the release of resources would happen to the point where managers would become indifferent to either releasing or keeping the additional resource unit. In this case keeping resources could be done, not only with the intention to avoid current, but also future cost adjustments, which would occur in case there was a need to acquire additional resource units. Contrary to drop, a more considerable increase of market demand can be absorbed only if necessary resources were acquired so that a larger volume of business activities could be performed, under the assumption of complete exploitation of the existing resources. "Managers' resource commitments depend not only on concurrent sales, but also on (1) prior period resource levels, which affect adjustment costs that have to be incurred in the current period, (2) expected future sales, which affect future adjustment costs, and (3) agency and behavioral factors, which drive a wedge between optimal decisions for the firm and managers' actual choices" (Banker & Byzalov, 2014, p. 47). Proving cost stickiness of overheads is quite challenging. Novák et al. (2017) showed that the variability of overheads is affected by the two factors - production in kg and total production in minutes - in the previous reporting period (month). "In contrast, a fixed/variable model of cost behavior asserts that the amount of costs incurred depends on the volume of realized activity in the current period only" (Balakrishnan & Gruca, 2008, p. 994).

Some resources, such as raw material, are purchased with the expectation that they will be consumed in the production process in a short period of time. Organizations commit, however, to making many other resources available whether or not the resources will be fully used for current and future activities (Cooper & Kaplan, 1992), such as buildings, machines, equipment. Managers will rather keep a larger volume of the resources, whose adjustment to current circumstances causes high costs, both explicit and implicit, than do what the existing activity level demands. Exploiting such resources creates a curve of conventional fixed costs. Accordingly, fixed and variable costs are two extreme cost categories determined by higher or lower cost adjustments. If cost adjustments are lower, we can expect that the change of costs will follow the change of the activity volume that caused it. Some forms of costs may be classified as variable or fixed costs with huge difficulties, as they behave differently in different situations, such as logistics costs, which is of special interest for cost management (see e.g. Novák et al., 2018). Adjustment costs may vary with institutional and contractual arrangements (Banker & Byzalov, 2014). This implies consideration of macro environment, characteristics of sector, capital market movements, labor legislation and social security, contractual obligations to suppliers and the like (see e.g. Guenther et al., 2014). Also, as observed by Li & Zheng (2017), management in financially stronger firms has more resources for investment expenditures in spite of a sales fall that can also cause cost stickiness.

Empirical research on the impact of adjustment costs on the level of cost stickiness has been hampered by difficulties in calculating adjustment costs. Banker et al. (2013) investigated the correlation between employment protection and cost stickiness for companies from OECD member countries. They concluded that stricter regulation, which implies higher adjustment costs, is associated with a higher level of cost stickiness, which is in line with theoretical assumptions.

Upon the analysis of factors that determine managers' decisions it is inevitable that in specific circumstances the decisions will be motivated, not by maximizing the value of

company, but by maximizing managers' personal benefits. Jensen's agency theory foresees that in the relationship between the principal and agent there are justified reasons to believe that the agent, obliged by the contract to perform a service for the principal, will not always work in the best interest of the principal when their interests are in conflict. Knowing about the asymmetry of information and agency's problems, establishing efficient board of directors, suitable compensation systems, acting on the market of corporate control and through other mechanisms of corporate management, rational investors limit managers' selfish, rational orientation (Todorović, 2011, p. 282). One of manifestations of this problem relates to manager's tendency to expand the limits of their company beyond optimal scope – empire expansion syndrome. According to Jensen (1986), growth increases manager's power through the increase of resources that they control. Insisting on unjustified growth is a reason for investment in projects with negative Net Present Value, overpayment of other companies' shares and the like. These actions can be interpreted as a need for gaining greater control.

Despite the assumption that managers' opportunistic behavior might result in the occurrence of cost stickiness, for many years no empirical research, which would confirm it, has been done. Also, empire expansion syndrome is dominantly analyzed in literature through the prism of complex M&A transactions (Motis 2007). However, Chen et al. (2012) found strong evidence that cost stickiness is positively associated with managers' empire building incentives due to the agency problem.

Contrary conclusions are reached if we consider managers' decisions that are made in order to accomplish defined profit goals. In order to realize personal gains such as a pay rise, bonus and other forms of compensation conditioned by achieving targeted financial performances, in case of a temporary drop in the activity volume managers may be motivated to adjust their resources to that lower level. Based on the concept of sticky costs, stimulus to meet earning targets will mostly result in the decrease of the level of cost stickiness (Kama & Weiss, 2010). Empirical research confirmed these assumptions (Kama & Weiss 2012). Dierynck et al. (2012) came to the conclusion that companies which tend to meet or beat zero earnings benchmark show greater readiness to reduce labor costs as the activity volume goes down, while labor costs in companies without this stimulus show stickiness in behavior.

However, it often happens that managers, in an attempt to maximize the value of the company, make decisions based on heuristics, which do not necessarily result in positive performance. Behavioral finance is based on the breakthrough of psychological principles into the finance theory. Tending to explain market anomalies as size effect, calendar effect, price bubble, market crash and the like (Todorović, 2011), supporters of this concept think that participants at the market are not usually rational, as it is envisaged by the conventional finance theory. Namely, it goes from the fact that individuals have limited abilities for solving complex problems, processing more data and instead of using sophisticated mathematical and statistical models, decisions are mostly based on intuition, former experience, emotions, as well as socially acceptable practices. Decisions on the level of resources are primarily based on expectations regarding the activity volume. In addition to the fact that these expectations are based on information on previous trends, the macroeconomic environment, they are also based on the personal judgment of managers about future trends. In general, people in a good mood are more optimistic in their choices and judgments than those in a bad mood. This also applies to managers. Among the heuristics recognized by behavioral finance theory, two stand out in the context of the effects they

have on cost behavior: overconfidence and optimism. It is often considered that these two characteristics are so closely linked that they usually appear together. An individual, who shows overconfidence, is also overly optimistic, and vice-versa. However, there is a need to make a difference between overconfidence and optimism (Malmedier & Tate 2005; Fabre & Heude-Francois 2009).

Overconfidence is generally defined as having inaccurate, overly positive perceptions of one's abilities or knowledge, which is conceptualized in two ways in the psychology literature: better-than-average effect and miscalibration (Qin et al., 2015). The "better than average" effect is a type of social comparison in which individuals, comparing their characteristics (or behaviors) to a group (peers, co-workers) believe to be superior to an average representative of that group in various fields (Skala, 2008). According to Biais et al. (2004) miscalibration involves overestimating the precision of someone's information. Individuals who are calibrated well are able to see if they are making a mistake, and in this context, a surplus of confidence appears as a form of wrong calibration. "Optimism, on the other hand, can be defined as the tendency or inclination to perceive an event or action as more likely to result in a favorable outcome, irrespective of the objective probability of that outcome actually occurring" (Fabre & Heude-Francois, 2009, p. 80). According to Li & Zhang (2017, p. 307) "managerial optimism about future demand may prompt managers in firms facing a competitive environment to make investment expenditures in spite of a current sales downturn, while pessimism may cause managers to be unwilling to infuse resources, even to the point where they face strong competition from rivals". Therefore, confidence is primarily based on personal characteristics, skills, while optimism is a perception of external phenomena (Malmendier & Tate, 2005, p. 2662).

Numerous authors researched behavioral factors as causes of stickiness in cost behavior. Chen et al. (2013), Qin et al. (2015) introduced managers' overconfidence into literature, as a behavioral factor of stickiness. Both studies confirmed that companies whose CEO shows a greater level of confidence experience higher level of cost stickiness. Salamah & Abulezz (2017) showed that company's costs that are managed by managers who have the aversion to risk show the reversed stickiness (anti stickiness) in behavior. Namely, with a drop in the volume of sales, these managers adjust resources to the drop, not wanting to be exposed to the risk of bearing the costs of unused capacity. Li & Zheng (2017, p. 311) found "evidence suggesting that cost stickiness arising from product market competition is increasing in managerial optimism, proxied by a prior period sales rise; whereas, managerial pessimism, proxied by a prior period sales fall, does not affect the relationship between competition and cost stickiness".

3. DESCRIPTION OF DATA AND SAMPLE

Presence of stickiness in cost behavior of companies that operate in Serbia will be tested on the sample of medium-sized and big companies belonging to manufacturing sector for the period 2007-2016. The size of a company is determined according to the level of sales revenue throughout the observed period, as a criterion. The research is based on the dataset based on data available in the official financial reports. As it is envisaged by appropriate rulebooks on the content and form of financial reports for companies, cooperatives and entrepreneurs that the income statement is to be made by using the method of total cost, we are not able to analyze the influence of the change of

operating revenues on the change of the sales, general and administrative costs. Having in mind analogous studies (Calleya et al., 2006; Pervan & Pervan, 2012; Kama & Weiss, 2012; Dalla Via & Perego, 2014; Pamplona et al., 2016), we decided to assess the influence of the change of sales revenue on the change of operating costs. In accordance with the requirements of econometric models in terms of data, we included into the sample companies having the data on operating revenues and operating costs for three years in a row, in a mentioned time interval, at which point the amount of operating revenues and operating costs is higher than 0. Afterwards we calculated the change of operating revenues and costs during two years in a row and performed data logarithmic which is in line with the conclusions of Banker & Byzalov (2014). Finally, we excluded from the sample 1% of the highest and lowest variable values that are included in the model in order to neutralize the extreme values in the sample.

The final sample consists of 917 companies. It is an unbalanced panel, with the number of observations per company ranging from 1 to 8, with an average of 7.125 observations, which makes that our sample has a total of 6,534 observations. Of the 917 companies in the sample, 197 are large and 720 are medium-sized.
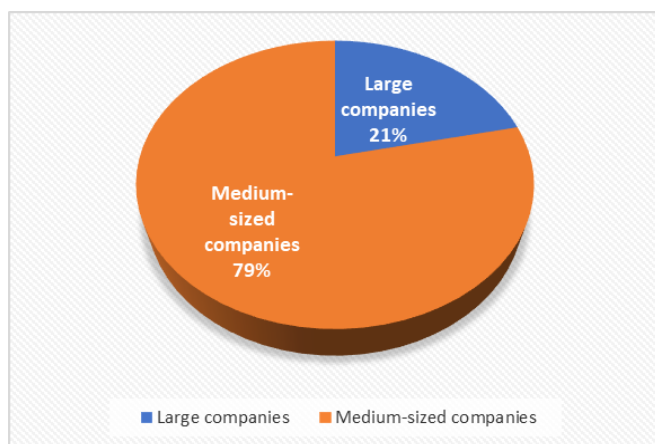
Fig. 3 Percentage of large and medium-sized companies in sample
Source: Author's calculation

An overview of the movement of mean operating revenues and operating costs, with appropriate standard deviations, by year, for the group of large and the group of medium-sized companies from sample is presented in table 1.

In 2009 both groups of companies mark a fall in average operating revenues, but also in costs. After that came a period of continuous growth of these indicators, with the exception that in in 2014, compared to last year, there was a slight fall in average operating revenues and costs in the group of large companies. Growing operating revenues may be an indicator of the end of recession and overall improvement of business environment since 2010.

Table 1 Mean values and standard deviations of operating revenues / operating costs by years, for the group of medium-sized and large companies in sample

Year	Medium sized companies				Large companies			
	Operating revenues (mill RSD)		Operating costs (mill RSD)		Operating revenues (mill RSD)		Operating costs (mill RSD)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
2007	577	638	541	600	6,286	21,094	6,055	20,460
2008	665	729	623	680	7,051	24,736	6,807	24,256
2009	636	649	590	584	5,810	12,771	5,519	12,474
2010	758	798	702	723	7,196	16,656	6,665	14,863
2011	911	1,011	848	932	7,776	18,243	7,292	16,166
2012	1,031	1,102	964	1,019	8,564	20,029	7,879	16,575
2013	1,044	1,144	987	1,068	9,693	24,666	9,032	21,667
2014	1,102	997	1,047	993	9,578	23,468	8,946	20,703
2015	1,209	1,008	1,148	996	10,239	21,464	9,542	19,281
2016	1,323	1,087	1,235	1,043	11,197	23,530	10,316	21,064

Source: Author's calculation

4. SPECIFICATION OF PANEL MODEL AND INTERPRETATION OF RESULTS

Methodology that Anderson et al. (2003) used was also used in later studies that dealt with this issue. Information about the sales, general and administrative costs are publicly available, while authors approximated the activity volume to the amount of operating revenues due to the data on physical sales volume being unavailable. Even though it is often stated as the limitation of the analysis, Banker and Byzalov (2014) emphasize that the use of operating revenues would be more purposeful even if this problem did not exist because the products, even of the companies doing business in the same branch, are somewhat differentiated, which makes physical units incomparable. Having in mind previous researches, as well as availability of data, we estimated the following model (1) in order to test cost stickiness:

$$\log \frac{C_{i,t}}{C_{i,t-1}} = \beta_0 + \beta_1 \log \frac{R_{i,t}}{R_{i,t-1}} + \beta_2 * Dummy_{i,t} * \log \frac{R_{i,t}}{R_{i,t-1}} + \varepsilon_{i,t} \tag{2}$$

where:

$C_{i,t}$ - amount of operating costs for company i in period t,

$R_{i,t}$ - amount of operating revenues[3] (sales revenue is increased by revenues from activating own goods and increase of value of finished goods and work in process, and it is decreased by the drop in the value of finished goods and work in process) for company i in period t,

Dummy variable$_{i,t}$ – takes on value 1 in case of $R_{i,t} < R_{i,t-1}$ or value 0 in case of $R_{i,t} > R_{i,t-1}$.

---

[3] Having in mind the discrepancy between the schemes of profit and loss account which were used till 2014 and the schemes that were prescribed to be used starting from 2014, certain adjustments were made in order to achieve comparability. When we say discrepancies, we dominantly think of previously prescribed use of production orientated concept method of overall costs in the scheme of profit and loss account, and then the rulebook which envisages the use of market oriented concept method of overall costs was enforced. In order to carry out the analysis, data on operating revenues and expenses for the period 2014-2016 were corrected in a way that costs classified by type remain in the framework, and operating revenues includes, apart from revenues gained from sales, also revenues from activating own goods, as well as appropriate correction through finished goods and work in process.

As dummy variable takes on value 0 upon the growth of revenues, coefficient β_1 shows proportional growth of operating costs with 1% growth of operating revenues, while in the case of fall in revenues dummy variable takes on value 1, which means that the addition of coefficients $\beta_1 + \beta_2$ shows proportional fall of operating costs upon 1% fall of operating revenues. Under the assumption that costs change, as it is envisaged by traditional theory, with a same amount in case of equivalent growth and fall of the activity volume, it is expected for a coefficient β_2 to take on value 0.

In order to take into account the effect of time on cost stickiness, based on previous research, we applied extended model (2) to test if cost stickiness is reversed in subsequent periods:

$$\log \frac{C_{i,t}}{C_{i,t-1}} = \beta_0 + \beta_1 \log \frac{R_{i,t}}{R_{i,t-1}} + \beta_2 * \text{Dummy}_{i,t} * \log \frac{R_{i,t}}{R_{i,t-1}}$$
$$+ \beta_3 \log \frac{R_{i,t-1}}{R_{i,t-2}} + \beta_4 * \text{Dummy}_{i,t-1} * \log \frac{R_{i,t-1}}{R_{i,t-2}} + \varepsilon_{i,t} \tag{3}$$

where:

$\text{Dummy}_{i,t-1}$ – takes on value 1 in case of $R_{i,t-1} < R_{i, t-2}$, or value 0 in case of $R_{i,t-1} > R_{i, t-2}$.

According to previous research, the β_3 coefficient, if positive, indicates a lagged adjustment to operating costs for changes in sales revenue. Further, the β_4 coefficient, if positive and lower than absolute value of β_2, indicates a partial reversal of the cost stickiness that follows a revenues decrease.

Choice of specification models was done in line with econometric procedure which entails the application of certain specification tests. The existence of individual effects was tested by the application of the F test for fixed specification and Breusch Pagan test for the random one. The existence of time effects is tested only by the application of the F test, having in mind that it is about a short-term time horizon and a large number of observation units.

Table 2 Specification tests

Test	Realized value of test statistics		p-value	
	Model 1	Model 2	Model 1	Model 2
F-test (fixed individual)	0.74	0.70	1.0000	1.0000
Breusch-Pagan (random individual)	0.00	0.00	1.0000	1.0000
F-test (time effects)	13.15	13.03	0.0000	0.0000

Source: Author's calculation

Conducted tests point to the fact that individual effects are not statistically important. On the other hand, based on realized value of F statistics, it is concluded that there are significant time effects. We can assume that such a result is the result of a dynamic macroeconomic environment conditioned by the crisis effect of 2008 and 2009. The inclusion of time effects in a regression models was done by introducing dummy variables in the model. In order to neutralize potential autocorrelation and heteroscedasticity problems, specifications with cluster-robust standard errors are used. Estimated coefficients are presented in table 3 for both models.

Table 3 Estimated coefficients

	Model 1	Model 2
β_1	0.8469394*	0.8485476*
	(0.0129047)	(0.0143084)
β_2	-0.1286175*	-0.140712*
	(0.0244261)	(0.026359)
β_3		0.0107644*
		(0.0048706)
β_4		0.0705936*
		(0.0211013)

Notes: Standard errors are given in parentheses; *-significance level of 5%.
Source: Author's calculation using Stata software

Estimated regression models are both significant as a whole, with the calculated F statistics of 865.21 and 788.04 respectively and appropriate p-values of 0. 0000.

Estimated value of the coefficients β_1 in model 1 points to the fact that operating costs increase on average 0.847 % per 1% growth of operating revenues, while, on average, they decrease by 0.718 % upon the fall of operating revenues by 1%, which points to the existence of stickiness in operating costs of observed companies. Estimated values of coefficients are statistically significant, at the value level of 5%. Results are comparable with the results gained from the research of Calleya et al. (2006), where the authors established that operational costs increase by 0.97% on average with 1% growth of revenues, while on the contrary they decrease by 0.91%. Dalla Via and Perego (2014) showed that operational costs increase by 0.94% on average with 1% growth of revenues in manufacturing industry, but they found coefficient β_2 to be positive. Bugeja et al. (2015) found that costs of Australian listed firms increase by 0.885% with 1% increase in operating revenues, but decrease by only 0.797% for 1% decrease in operating revenues.

We can find few potential causes of identified presence of cost stickiness. Companies from our sample belong to manufacturing sector which is characterized by greater level of fixed assets: machinery, plants and equipment than trading and service sectors. These committed resources cause high adjustment costs, which promotes the emergence of cost stickiness. This is in line with conclusions of Weidenmier and Subramaniam (2003). Having in mind that managers rely on information about past operating revenues in forming expectations about future demand, other potential cause of observed cost stickiness may be managers' optimism. Since there is a continuous growth of business activity, measured by the amount of operating revenues in the observed period, managers will be hesitant to reduce unused resources in the case of a slight decrease of demand in order to avoid adjustment costs. We can assume that the trend of growth is caused by the exit from recession and overall improvement of business environment since 2010. Besides, as it is emphasized in Pervan & Pervan (2012), the only similar research done in the field of sticky costs in the region, managers of state owned companies are surely less prone to dismiss workers in the conditions of fall in business activities due to the implementation of social and employment policies, which can be reflected in these results. However, the confirmation of these assumptions requires similar researches to be done at the level of labor costs, in companies with a different ownership structure. Another potential cause of identified cost stickiness may be managerial empire-building incentives. According to

Chen et al. (2012) cost stickiness is more pronounced under weak corporate governance. This is, mainly, a problem of state owned companies.

Estimated coefficients β_1 and β_2 in model 2 are similar to their counterparts in model 1, which confirms cost stickiness. The positive and significant estimated coefficient β_3 indicates a lagged adjustment to operating costs for changes in operating revenues. Additionally, the estimated coefficient β_4 is, also, positive and significant, and its value is less than absolute value of estimated coefficient β_2 ($0.0705936 < |-0.140712|$) implying a partial reversal of stickiness in the period after a revenue decrease. We can interpret this as manager's caution in deciding to reduce resources, which takes some time to obtain information about the persistence of changing demand. The other explanation could be that contractual arrangements related to committed resources take time to be undone. Besides, cost stickiness observed in one period is partly reverted in subsequent period, after a revenue decrease. "During longer time intervals, the managerial assessment about the permanence of a change in revenues becomes more precise and then adjustment costs become lower relative to the cost of keeping unused resources" (De Medeiros & Costa, 2004). This shows us that cost stickiness is less pronounced when observed in a longer time period.

Conclusion

Determining the pattern of cost behavior is not at all simple. Traditional cost theory classifies costs into two groups, depending on how they react in total to a change in the activity volume: fixed and variable costs. However, the costs show different features depending on the time horizon that is observed. In addition, the complexity is especially pronounced when we keep in mind that costs do not have a strictly defined character, but that their behavior is determined by numerous factors. Determining the way in which costs move with the change in the activity volume is essential for the needs of planning, cost control, as well as making individual business and financial decisions.

The theory of sticky costs indicates that the variable and fixed are the two extreme cost categories determined by low and high adjustment costs, respectively. Anderson et al. (2003) consider deliberate decision making on resource adjustment as the primary cause of stickiness in the cost behavior. Taking into account the direction of activity volume change, the authors concluded that managers make different decisions on resource adjustment in case of increase or decrease in activity volume.

Despite popularity and growing interest in this subject in the expert literature, some authors were critical of the theory of sticky cost behavior (e.g. Anderson & Lanen, 2007). In spite of this, it is certain that the results of the studies based on sticky cost theory provided researchers with enthusiasm regarding the analysis of cost behavior, placing it into new context which is compatible with contemporary ideas of the theory of finance, presented by behavioral finance. Framework of sticky costs presents a different way of thinking about the cost behavior in relation to the activity volume, not necessarily opposed to the conventional theory. A better understanding of cost behavior results in more precise and better cost planning. "At a shorter time horizon, cost stickiness potentially affects those corporate decisions which use some variant on a theme of average/standard costs and which typically do not acknowledge that costs, for managerial considerations, do not behave as anticipated" (Calleya et al., 2006, 140). The settings of this theory are also important for the

purpose of predicting financial performance, both for financial analysts and existing and potential investors. Ignoring this phenomenon in cost behavior could result in underestimation of budgeted costs in case of less desirable outcomes (decline in sales volume), and thus overestimation of profitability. Previous researches testify to minimal awareness of asymmetric cost behavior (e.g. Novák et al., 2018).

Starting from the methodology introduced by Anderson et al. (2003), we have shown that the costs of medium and large manufacturing companies that operate in Serbia show stickiness. Similar research for our country has not been conducted, which makes this one a basis for further studies. In addition, it would be possible to include certain elements of operating costs in the research – i.e. certain categories of costs, where, in the context of adjustment costs, the focus would be on salary costs. Future research should focus on determinants of cost stickiness of domestic companies, bearing in mind the complex macroeconomic environment to which companies are exposed in the domestic market, which exceeds the empirical part of this paper. This type of research would enable a better understanding of the business decision-making process, but would also make some progress in the development of cost accounting in our country

This research has certain limitations. First of all, we used the amount of sales revenue as a proxy variable for the scope of activities due to the lack of a more adequate publicly available variable. In addition, we must keep in mind that changes in costs and revenues, as proxy variables for activity volume, are conditioned not only by higher sales volumes, but also by changes in sales prices.

REFERENCES

Anderson, M. C., Banker, R. D., & Janakiraman, S. N. (2003). Are selling, general, and administrative costs "sticky". *Journal of Accounting Research*, *41*(1), 47–63.

Anderson, S., & Lanen, W. (2007). *Understanding cost management: What can we learn from the evidence on 'sticky costs'?* (Working Paper), Melbourne: University of Melbourne.

Balakrishnan, R., & Gruca, T. (2008). Cost Stickiness and Core Competency: A Note. *Contemporary Accounting Research*, *25*(4), 993-1006.

Banker, R., & Byzalov, D. (2014). Asymmetric Cost Behavior. *Journal of Management Accounting Reserch, 26*(2), 43-79.

Banker, R., Byzalov, D., & Chen, L. (2013). Employment Protection Legislation, Adjustment Costs and Cross-Country Differences in Cost Behavior. *Journal of Accounting & Economics, 55*,111-127. http://dx.doi.org/10.1016/j.jacceco.2012.08.003

Banker, R., Byzalov, D., Fang, S., & Liang, Y. (2018). Cost Management Research. *Journal of Management Accounting Research*, *30*(3), 187-209.

Banker, R., Ciftci, M., & Mashruwala, R. (2008). Managerial Optimism, Prior Period Sales Changes, and Sticky Cost Behavior. Retrieved from: https://dx.doi.org/10.2139/ssrn.1599284

Biais, B., Hilton, D., Mazurier, K., & Pouget, S. (2008). Judgemental Overconfidence, Self-Monitoring, and Trading Performance in an Experimental Financial Market. *Review of Economic Studies, 72*(2), 287-312.

Bugeja, M., Lu, M., & Shan, Y. (2015). Cost stickiness in Australia: Characteristics and determinants. *Australian Accounting Review*, *25*(3), 248-261.

Calleya, K., Steliaros, M., & Thomas, D. C. (2006). A note on cost stickiness: Some international comparisons. *Management Accounting Research*, *17*(2), 127-140.

Chen, C. X., Gores, T., & Nasev, J. (2013). *Managerial overconfidence and cost stickiness* (Working paper), University of Illinois at Urbana-Champaign and University of Cologne. https://doi.org/10.2139/ssrn.2208622

Chen, C. X., Lu, H., & Sougiannis, T. (2012). The Agency Problem, Corporate Governance, and the Asymmetrical Behavior of Selling, General, and Administrative Costs. *Contemporary Accounting Research*, *29*(1), 252-282.

Cooper, R., & Kaplan, R. (1992). Activity-based Systems: Measuring the Costs of Resource Usage. *Accounting Horizons*, *September 1992*, 1-13.

Dalla Via, N., & Perego, P. (2014). Sticky cost behaviour: evidence from small and medium sized companies. *Accounting and Finance*, *54*(3), 753–778.

De Medeiros, O. R., & Costa, P. D. S. (2004). Cost stickiness in Brazilian firms. *Available at SSRN 632365*: https://ssrn.com/abstract=632365.

Dierynck, B., Landsman, W. R., & Renders, A. (2012). Do managerial incentives drive cost behavior? Evidence about the roles of the zero earnings benchmark for labor cost behavior in Belgian private firms. *The Accounting Review, 87*(4), 1219-1246. https://doi.org/10.2308/accr-50153

Fabre, B., & Heude-Francois, A. (2009). Optimism and overconfidence investors' biases: a methodological note. *Finance, 30*(1), 79-119.

Guenther, T. W., Riehl, A., & Rößler, R. (2014). Cost stickiness: state of the art of research and implications. *Journal of Management Control, 24*(4), 301-318.

Hamermesh, D. S., & Pfann, G. A. (1996). Adjustment costs in factor demand. *Journal of Economic literature*, *34*(3), 1264-1292.

Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *The American economic review*, *76*(2), 323-329.

Kama, I., & Weiss, D. (2012). Do Ernings Targets and Managerial Incentives Affect Sticky Costs. *Journal of Accounting Research, 51*(1), 201-224. https://doi.org/10.1111/j.1475-679X.2012.00471.x

Kama, I., & Weiss, D. (2010). *Do managers' deliberate decisions induce sticky costs?*. Tel Aviv University, Faculty of Management, The Leon Recanati Graduate School of Business Administration.

Li, W. L., & Zheng, K. (2017). Product market competition and cost stickiness. *Review of Quantitative Finance and Accounting*, *49*(2), 283-313.

Lucas, R. (1967). Adjustment Costs and the Theory of Supply. *The Journal of Political Economy*, *75*(4), 321-334.

Maher, M., Stickey, C., & Weil, R. (2008). *Managerial Accounting: An Introduction to Concepts, Methods and Uses.* Thomson: South-Western.

Malinić, D., Milićević, V., & Stevanović, N. (2018). *Upravljačko računovodstvo [Management Accounting].* Ekonomski fakultet Beograd.

Malmendier, U., & Tate, G. (2005). CEO overconfidence and Corporate Investment. *The Journal of Finance, 60*(6), 2661-2700.

Motis, J. (2007). Mergers and acquisitions motives. *Toulouse School of Economics EHESS (GREMAQ) and University of Crete.* Retrieved from https://economics.soc.uoc.gr/wpa/docs/paper2mottis.pdf

Noreen, E., & Soderstrom, N. (1997). The Accuracy of Proportional Cost Models: Evidence from Hospital Service Departments. *Review of Accounting Studies*, *2*(1), 89–114.

Novák, P., Dvorský, J., Popesko, B., & Strouhal, J. (2017). Analysis of overhead cost behavior: case study on decision-making approach. *Journal of International Studies, 10*(1),74-91. https://doi.org/10.14254/2071-8330.2017/10-1/5

Novák, P., Hrušecká, D., & Macurová, L. (2018). Perception of cost behaviour in industrial firms with emphasis on logistics and its costs. *FME Transactions, 46*(4), 658-667. https://doi.org/10.5937/fmet1804658N

Pamplona, E., Fiirst, C., de Jesus Silva, T. B., & da Silva Zonatto, V. C. (2016). Sticky costs in cost behavior of the largest companies in Brazil, Chile and Mexico. *Contaduría y Administración*, *61*(4), 682-704.

Pervan, M., & Pervan, I. (2012). Sticky costs: evidence from Croatian food and beverage industry. *International Journal of Mathematical Models and Methods in Applied Sciences*, *8*(6), 963-970.

Qin, B., Mohan Warsha, A., & Kuang, F. Y. (2015). CEO Overconfidence and Cost Stickiness. *Management Control & Accounting, 2015*(2), 34-38.

Regulation on the classification of activities, https://www.paragraf.rs/propisi/uredba_o_klasifikaciji_delatnosti.html

Salamah, M. A., & Abulezz, E. M. (2017). Cost Stickiness: Does Manager's Preference Toward Risk Matter? An Empirical Study. *SSRN Electronic Journal*. https://dx.doi.org/10.2139/ssrn.3492541

Serbian Business Registers Agency. http://www.apr.gov.rs

Skala, D. (2008). Overconfidence in Psychology and Finance – an Interdisciplinary Literature Review. *Bank i Kredyt*, *4*, 33-50.

Todorović, M. (2011). Psihologija i finansijski menadžment: bihevioralne korporativne finansije [Psychology and Financial Management: Behavioral Corporate Finance]. *Ekonomika preduzeća*, 275-287.

Weidenmier, M., & Subramaniam, C. (2003). *Additional Evidence on the Sticky Behavior of Costs.* Working Paper, Texas Christian University.

DA LI SU TROŠKOVI ASIMETRIČNI – PRIMER SRBIJE

Rad je fokusiran na teoriju asimetričnih troškova, iznedrenu istraživanjima koja su ukazala na činjenicu da troškovi ne reaguju simetrično u slučaju ekvivalentnog povećanja i smanjenja obima aktivnosti, kao što to podrazumeva tradicionalna teorija troškova. Poslovne odluke donete u cilju povećanja vrednosti kompanije, kao i oportunističke odluke, usmerene na ostvarenje ličnih ciljeva menadžera, ističu se kao jedan od osnovnih uzroka asimetrije troškova. Da bismo ispitali postojanje asimetrije u kretanju troškova kompanija koje posluju u Srbiji, sproveli smo istraživanje na uzorku od 917 srednjih i velikih preduzeća iz proizvodnog sektora za period 2007 - 2016. Analiza panel podataka ukazala je na prisustvo asimetrije u ponašanju ukupnih troškova, odnosno, pokazala je da ukupni troškovi rastu za 0,847% u slučaju 1%-og rasta poslovnih prihoda, a padaju za 0,718% usled 1%-og pada poslovnih prihoda. Utvrdili smo i da se troškovi uz kašnjenje prilagođavaju poslovnim prihodima, kao i da postoji parcijalna reverzibilnost asimetrije troškova u periodu koji sledi period pada poslovnih prihoda.

Ključne reči: *kretanje troškova, asimetrija troškova, troškovi prilagođavanja, bihevioralne finansije, agencijska teorija*

Original Scientific Paper

THE IMPORTANCE OF MARKETING COMMUNICATION FOR ATTRACTING AND RETAINING INSURANCE SERVICE USERS

UDC 658.8
368

Slavica Popović, Mira Avramović

Academy of Vocational Studies Southern Serbia, Department of Business Studies, Leskovac, Serbia

Abstract: *The assumption of successful marketing communication within the service sector is the establishment of two-way interactive communication between service users and service companies. The aim of the strategy of marketing communication is attracting new users, but also retaining the existing users of services and establishing long-term relations with them. A successful achievement of the aim of marketing communication requires that the marketing message is created according to the characteristics and preferences of the target audience. The role of marketing communication differs depending on the stage of the selling process. During the pre-purchase stage, the aim of marketing communication is to inform potential users so that they could create preferences towards the company's offer. By applying the adequate forms of marketing communication during the stage of using the service, the company tends to explain to and introduce users to the process of servicing and to provide support in solving the occurring problems. The confirmation of users' positive impressions and the resolution of possible problems which can occur during the process of service provision represent the aim of marketing communication during the post-purchase stage.*

Key words: *marketing communication, service user, attracting, satisfaction, retaining service users*

INTRODUCTION

The marketing concept in insurance companies implies a complete orientation towards meeting the needs of insurance service users. Through the process of defining and implementing the research programme, the insurance company identifies the needs, desires and demands of service users. On the basis of research results, it creates an adequate offer. By selecting the form, marketing communication informs and creates the awareness of potential users about the offer, and stimulates interest and demand. The offer of insurance companies is standardised and the possible diversification is minimal. Therefore, the achievement of competitive advantage of insurance companies is based on the achievement of satisfaction and, thus, retaining insurance service users.

The strategy of marketing communication originates from the company's marketing strategy. Depending on the goals set by the company, the priorities and goals of the strategy of marketing communication are defined. The effective and efficient strategy of marketing communication is a prerequisite for business success of insurance companies. If that defined goal of the insurance company is to attract and retain service users, then the role of marketing communication is different in every stage of the purchase process. A logical time order of activities is the creation and announcement of realistic promises, fulfilment of the given promises and the establishment of communication after the completion of the purchase process. A successful implementation of a defined goal requires that the insurance company continually monitors the changes in the users' needs and demands and that it achieves an intense communication with users during the provision of services, as well as with those users who have stopped cooperating with the insurance company. These activities make service users perceive that the efficient fulfilment of their needs and demands is the focus of all the activities of the insurance company. In this way, marketing communication creates positive impressions which increase the satisfaction of users and contribute to the development of long-term relations with them.

The factors such as intense competitiveness, fast development and the application of new information-communication technology, fragmentation of the media and changes in users' behaviours have influenced changes in the strategy of marketing communication. Contemporary business conditions determine the establishment of an interactive dialogue between the company and users and it occurs in all stages of the purchase process. In the pre-purchase stage, the role of marketing communication is to inform potential users, to make them interested and to create preferences towards the company's offer. By applying the adequate forms of marketing communication during the stage of using the service, the company tends to explain the stages of the process of servicing to users and to provide them with support in solving possible problems. In the post-purchase stage, the aim of marketing communication is to increase the positive experience of service users. The strategy of marketing communication should contribute to the creation of realistic expectations and the fulfilment of promises given to service users. The unrealistic and exaggerated promises given in the pre-purchase stage can increase the level of users' expectations. If the company is not able to fulfil the given promises during the process of servicing, service users will feel betrayed after the completion of the process of providing services and they will perceive a low level of service quality. All this can affect their decision to end cooperation with the company.

The aim of the paper is to show the connection between marketing communication and attracting, satisfying and retaining insurance service users. The paper should point

out the role of marketing communication activities in the establishment, development and maintenance of long-term relations with insurance service users. The results of the empirical research, which are presented in the paper, are significant because they indicate that the insurance company can improve its business performance by implementing and investing into the activities of marketing communication. A better understanding of the role and the increased investment of insurance companies into the activities of marketing communication add to the fact that these types of research are becoming more important.

The paper consists of five parts. The first part presents a theoretical concept of marketing communication – the understanding of the role of marketing communication. The second part of the paper analyses the influence of marketing communication on attracting insurance service users. The third part analyses the role of marketing communication in achieving the satisfaction of insurance service users and retaining them. The fourth part of the paper refers to the research methodology, and the fifth part includes a discussion about the research results.

1. MARKETING COMMUNICATION – THEORETICAL CONCEPT

A well-defined and implemented strategy of marketing communication enables the achievement of excellent business results and recognition of the company on the market. Kotler & Keller (2006, p.536) states that by using marketing communication companies tend to inform, persuade and remind the target segment of consumers of the company's offer. Marketing communications represent the "voice" of the company and its brands and are a means by which it can establish a dialogue and build relationships with and among consumers (Keller, 2009, p.141). The efficient communication between the company and the target audience assumes that it occurs in both directions (Stanković, et.al., 2014, p.15).

The fact that, in the contemporary conditions, consumers are exposed to a large number of marketing messages has created the need for companies to start viewing marketing communication as an interactive dialogue which occurs during the process of purchase.

In the pre-purchase stage, by employing marketing messages, potential users are given a promise about future satisfaction with the services. In this stage, the role of marketing communication is to inform, draw the attention of potential users and create the preferences towards the company's offer. The behaviour of potential service users, which is characteristic for this stage, represents focusing a significant amount of attention to the comparison of alternatives with the aim of minimising the level of perceived risk. In this sense, each form of marketing communication will have different credibility with decision-makers. When deciding about the choice of a service or a service company, potential users collect information from different sources. Besides a previous experience of the user with the given service or service company, promotional activities of service companies are a significant source of information. However, due to the specificity of services, those potential users who do not have that personal experience will tend to collect information through direct communication, primarily with friends, relatives and acquaintances, but also with the employees of the service company. Contrary to classic mass communication of the producers of consumer goods, the employees of service companies, who use a direct contact with service users, have a significant position as the carriers of external marketing communication (Senić, et. al., 2008).

A marketing message should be created by taking into account the specifics and demands of the target segment of users. In this way, during the stage of providing the service,

the ratio of their cancellation of the service is minimised. The role of marketing communication during the stage of providing service is to introduce and explain to users the stages of the service process. However, during the service process, certain problems can occur. In those situations, marketing communication represents the key instrument for an efficient resolution of the problem, while the emphasis is on the role of the employees who establish direct communication with service users.

The sense of confirming or not confirming expectations and promises given to service users is characteristic for the post-purchase stage. In this stage, the role of marketing communication is to use an adequate approach and information in order to resolve potential misunderstandings with users or to strengthen their positive impressions.

Marketing communication is a mutual term for different types of planned messages used by a company – economic propaganda, sales improvement, direct marketing, personal selling, public relations, sponsorship (Duncan, 2005, p.7). Shimp states that marketing communication is an approach whose implementation is used by marketing managers in order to inform, persuade and remind users of the company's offer (Shimp, et.al., 2013). Informing and persuading users and brand differentiating, with the aim of forming their positive attitudes and completing the transaction, are basic functions of marketing communication (Stanković, et.al., 2014, p.16). By using the combination of instruments of the communication mix, marketing managers tend to achieve a maximum communication influence. The starting and the ending point in the development of the strategy of marketing communication is the consumer. The consumer does not differentiate between direct and indirect forms of marketing communication, but believes that those are all the ways of persuasion used by the company. The consumer finds it confusing and unconvincing if he/she is faced with a non-consistent message. Therefore, in order to efficiently achieve defined goals of marketing communication, it is important that the messages to the consumers are presented in an integrated and consistent manner (Avramović, 2008).

2. MARKETING COMMUNICATION AND ATTRACTION OF INSURANCE SERVICE USERS

The main aim of marketing activities of service companies is attracting and retaining service users. The leading factor which affects attracting of new and retaining the existing service users is satisfaction. Veljković (2009, p.101) believes that satisfaction is the reaction of consumers to the evaluation of the perceived difference between previous expectations (or a specific standard of performance) and the actual performances. For Ljubojević (1998), the concept of satisfaction is abstract, and it represents the function of perception (the probability of an event) and expectation (the evaluation of an event after the delivery of service) of the consumer.

The activities of marketing communication contribute to the attraction of service users. Potential service users form their expectations on the basis of marketing messages of service companies given through direct and indirect forms of marketing communication, and based on conversations with friends and relatives. They form their expectations during the pre-purchase stage, and the degree of achieved satisfaction is evaluated during the stage after the use of service. The expectations of users are, primarily, oriented towards the quality of service, the adjustment of the offer to their requests, desires and needs, the empathy of employees, and the fulfilment of promises given through marketing communication.

The ability of the company to fulfil and overcome the expectations of service users creates certain advantages. First of all, service users who have achieved a certain level of satisfaction will transfer a positive message to potential users. Users are often included in the process of service provision and they transfer their positive or negative experiences through informal direct communication to other potential users. When a potential user needs to make a decision regarding the choice of a service company or a service, an important source of information will include friends, relatives, and acquaintances. Then, their positive or negative experiences, i.e. the achieved degree of satisfaction, becomes significant. After making a decision regarding the purchase of a certain service and establishing interaction with the service company, the user forms an experience as a result of comparing previous expectations and achieved gain. They will share their experiences (positive or negative) with friends, acquaintances, relatives or through social networks, which speeds up the transfer of information.

When it comes to services, the greatest significance goes to personal selling as a direct form of marketing communication. Its goal is to inform potential users about the company's offer, to stimulate interest and to encourage them to purchase. Although the marketing message which is presented by employees to potential users is created in advance, it can be changed and adjusted, depending on a specific situation. The ultimate goal is to succeed in selling.

The most significant segment of personal selling includes the employees who establish direct communication with the potential users. Personal selling is a very effective factor in establishing contact between salespersons and buyers in two directions. The customers can inquire and ask the salespersons face-to-face questions. At the same time, they can discuss and reach a clear understanding and make a proper buying decision (Yousif, 2016, p.130). The contribution of those employees is reflected in understanding needs, collecting information about specific requests, expressing empathy and creating the offer according to the needs and desires of potential users. The role of personal selling in the pre-purchase stage is not emphasised, because the potential user is still collecting necessary information. During the stage of using the service, the user needs additional information which he/she receives through direct communication with employees with the aim of achieving complete satisfaction (Vračar, 2005). Also, the employees who know the needs of users well enough can represent the channel for offering new services.

Development and the increasing application of modern technologies has significantly influenced the expressed growth of applying direct marketing in the service sector. Direct marketing is based on databases. It enables personalised and interactive communication, receiving an immediate response from the user in the form of purchasing or asking for further information, and the effects can be measured. Marketing messages are most often sent by e-mail or SMS. The activities of direct marketing enable the identification of potential service users, but also the increased level of loyalty of the existing ones. By receiving a direct response from the users (by phone, mail, or e-mail), companies create a database of potential users. The increase in the level of loyalty is achieved by adjusting the offer created on the basis of data from the database about users' preferences or special interests.

When it comes to mass forms of marketing communication, economic propaganda is applied most often. It attempts to persuade potential customers to purchase or consume a product or service. It is designed in such a way that it creates and reinforces brand image and brand loyalty (Kumar, et.al., 2011, p.83). Its significance is expressed with attracting new users, i.e. during the pre-purchase stage. The goal is to inform potential users about

the company's offer and to minimise the perceived risk. The purpose of economic propaganda is to influence potential users so that they could think about or react to a service or a company in a certain way (Senić, et.al., 2008). Taking into consideration the characteristics of services, primarily the fact that they are not material, the company can "materialise" services by sending adequate marketing messages through economic propaganda. In this way, the potential service users are given the opportunity to compare alternatives and decision-making regarding the purchase is easier.

Considering the fact that the service sector does not have a large number of mediators, the improvement of sales as a mass form of marketing communication is less significant in relation to the previously mentioned forms. However, the growth of competition and the sensitivity of users to additional services, have conditioned a larger application of the improvement of sales. The activities of sales improvement imply the use of discounts, gifts, coupons, and organization of prizewinning games. The application of means for the improvement of sales enables a short-term increase in the sales volume, encourages users to respond quickly and attracts users of competitive companies, thus increasing the database of its own users, and the accomplished effects are easily measured.

3. MARKETING COMMUNICATION, SATISFACTION AND RETAINING SERVICE USERS

The key factor for accomplishing a long-term competitive advantage and success of a company on the market is the satisfaction of service users. The specificity of the service sector is the fact that it represents a significant source of dissatisfaction in users. The tendency of users who are not satisfied with a service or a service company is to make a decision to end their cooperation. Therefore, with the aim of keeping and developing long-term relations with the users, companies tend to provide them with complete satisfaction. That is why it is important to base the activities of the company on users' expectations and needs. The satisfaction of users depends on the quality of service and it is based on a specific transaction. Satisfied users perceive the company's offer as having quality, which increases the ratio of repeated purchases, the spread of positive information, and a higher level of loyalty and trust (Thaichon, et.al., 2013).

The significance of the quality of services is particularly expressed in the process of retaining the existing users. Traditionally, service quality has been explained as the gap between customer expectation about the service and perception of how the service has been performed (Kumar, 2017, p.160). The quality of service represents the key characteristic of the service offer because that is the dimension through which the user perceives the successfulness of a service company. In order for the company to provide a service of a satisfactory quality, it is necessary to include all employees and to have a good understanding of the needs, demands and desires of users. Besides the focus of the company on the creation of a quality offer, there also needs to have a system which will efficiently control the creation and delivery of an adequate quality service.

The users form their expectations on the basis of information from formal and informal sources. Although the influence of information from informal sources is expressed, the key significance belongs to marketing communication. Users' expectations should be realistic when it comes to the possibilities of the company to satisfy their needs. Unrealistic expectations of users can add to the increase in sales volume only short-term, but, long-term, they will result in their dissatisfaction and the end of cooperation.

The satisfaction of users is most often evaluated on the basis of their experiences with the company's offer, but also on the basis of their overall experience with consumption which is the result of comparing expected and delivered values (Stanković, 2002). The measurement of satisfaction is a complex and hardly noticed variable which is the result of the subjective evaluation of service users. The aim of measuring the satisfaction of service users is to quantitatively express their subjective perceptions in an objective manner (Veljković, 2009, p.122). The essence of the measuring of satisfaction is reflected in identifying factors which affect the level of loyalty in users and a repeated purchase. The measuring of users' satisfaction enables the company to acquire information about the cause of satisfaction and which elements of the offer have the strongest effect on the increase in the level of satisfaction.

Retaining users refers to focusing the marketing activities on the existing database of users. The company which wants to retain existing users, besides the fact that it tends to satisfy their expectations, is also oriented towards motivating them to remain at the company's offer by increasing their satisfaction (Lovreta, et.al., 2010). By delivering the service with additional value, traditional relations between the company and the user becomes a partnership. A satisfied user will use the company's service again, he/she will recommend the company and its services to other potential users and they will have no reason to consider an offer from competitors. This means that satisfaction is the key factor for retaining users.

By creating an offer which is in accordance with the users' demands and needs, the service company tends to make potential users into first-time users, and then into users who repeat the purchase. However, the users who use the service for the first time, as well as those who repeat their purchase, can simultaneously use competitors' services. In that case, a well-defined and implemented strategy of marketing communication enables the achievement of their loyalty and the development of long-term relations. Personal selling and sales improvement are the forms of marketing communication which have the greatest influence on retaining users, but the activities of economic propaganda, whose aim is reminding about the company's offer, are also significant (Kotler, et.al., 2017).

 Personal selling is directed towards monitoring users' reactions and the level of achieved satisfaction. The employees tend to collect information about users' satisfaction through direct communication, to retain them and to encourage them to spread positive information. Also, they contribute to the establishment of long-term relations which are based on the confirmation of promises given in the pre-purchase stage. The implementation of activities which are directed towards the achievement of users' satisfaction will have an influence on their decision to use the company's services in the future.

The role of economic propaganda in the post-purchase stage is crucial for overcoming users' potential doubts in the correctness of the decision (Stanković, et.al., 2014). In this stage, the message of the economic propaganda has the aim of confirming users' right choice and to encourage them to repeat their purchase. By reminding them of the company's offer and affirming positive attitudes which have been previously adopted by users, economic propaganda contributes to the increase in repeated purchases. Economic propaganda leads to the creation of loyal users who are less sensitive to competitors' prices and offers (Avramović, 2008).

4. RESEARCH METHODOLOGY

4.1. The conceptual research model of the significance of marketing communication for attracting and retaining insurance service users

The subject of research of this paper is the analysis of the significance of marketing communication for attracting and retaining insurance service users. The subject of research has been concretised by defining the research framework as it is shown in Figure 1.
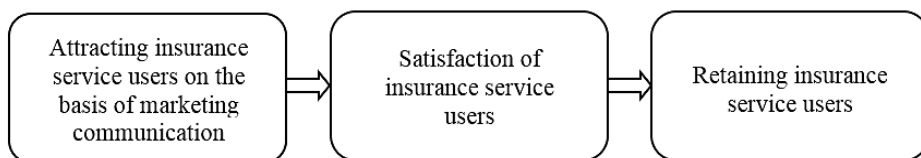
Fig. 1 The conceptual research model

Research questions for the definition of attracting insurance service users on the basis of marketing communication are:
1) What are the sources of information of insurance service users regarding the company's offer?
2) Do insurance service users notice the forms of marketing communication of insurance companies?
3) Do insurance service users notice marketing messages of insurance companies sent through the mass media?
4) Do insurance service users believe in marketing messages of insurance companies sent through the mass media?
5) Do insurance service users believe in marketing messages of insurance companies which they receive from a direct contact with the sellers of insurance services?

The research questions which refer to the satisfaction of insurance service users are:
6) How is the quality of the complete offer of the insurance company evaluated by insurance service users?
7) Are insurance service users satisfied with the services of the insurance company?

The research question for retaining insurance services users is:
8) What will be insurance service users' decision if the conditions of insurance are changed?

4.2. The characteristics of the sample and the method of research

In accordance with the defined subject and research questions, the research has been conducted by applying the research method. The structured questionnaire has been used as the main instrument for collecting primary data. The survey questionnaire mostly contains multiple choice questions, statements with which the respondents either agree or disagree, and the ranking of attitudes has been done by applying a Likert scale with scores from 1 to 5.

The primary data are collected from the group which consists of 145 respondents, insurance service users. According to the gender of respondents, the numbers are almost equal, 74 women and 71 men. Out of the total number of respondents, the largest number of them are 46-55 years old (48), followed by the category of respondents who are 36-45 (38)

and 26-35 (36), while the least represented include the categories of the youngest respondents who are 18-25 (4) and respondents who are over 60 years old (8). According to the level of education, the largest number of respondents has acquired higher education – a university degree (46), a master's degree (20) or a PhD (18), but the participation of respondents with a high-school degree (29) is also significant. Most respondents have their own sources of income. The largest number of respondents, 50 of them, state that they earn a monthly household income of 40,001-70,000 RSD, followed by the category of those who earn the income of 70,001-100,000 RSD (45).

5. RESEARCH RESULTS AND DISCUSSION

This part of the paper presents the research results and a discussion on the influence of marketing communication on attracting, satisfying and retaining insurance service users.

a) The analysis of attracting insurance service users.

The analysis of the answers of respondents to the first research question, regarding the sources of information while selecting the insurance company, shows that most insurance service users, 53 of them, have stated that they receive necessary information about insurance services through direct communication with the employees of the insurance companies, through personal selling. A significant source of information for insurance service users is also the informal communication with friends, relatives and acquaintances. Internet browsing and economic propaganda, as the mass form of marketing communication, are a less prevalent source on the basis of which users make a decision to establish cooperation with an insurance company. The research results are shown in Table 1.

Table 1 Sources of information of insurance service users

Atitudes of insurance service users	N	The number of users who provided an answer	% users who provided an answer
Economic propaganda	145	25	17,24
Internet browsing	145	16	11,03
Talking to friends, relatives, acquaintances	145	51	35,17
Personal selling	145	53	36,55

Source: Authors' calculations

In the further course of the research, the respondents were supposed to state the forms of marketing communication of insurance companies which they have noticed. The answers of the respondents to the second research question – whether service users notice the forms of marketing communication of insurance companies – are shown in Table 2. The research results indicate that one insurance service user can notice more than one form of marketing communication. Table 2 also shows the respondents' answers to the third research question – whether service users notice marketing messages of insurance companies sent through the mass media. The largest percentage of respondents has noticed marketing messages sent through television, radio, newspapers, magazines and billboards (44.25%). The percentage of noticed messages sent through the Internet (28.29%), as well as the activities of personal selling (17.14%), are also significant. The activities of sales improvement (discounts, prizewinning games, gifts) and direct marketing are noticed the least.

Table 2 The percentage of noticed forms of marketing communication of insurance companies

The forms of marketing communication	N	The number of users who provided an answer	% of users who provided an answer
Economic propaganda	145	111	44.25
Personal selling	145	43	17.14
Sales improvement	145	14	5.58
Direct marketing	145	12	4.78
Internet	145	71	28.29

Source: Authors' calculations

The answers of the respondents to the fourth research question – whether insurance service users believe in the marketing messages of insurance companies sent through the mass media – are shown in Table 3. Although insurance service users have largely noticed the activities of economic propaganda of insurance companies, most of them either partially believe in or are indifferent towards the marketing messages sent through the mass media. The attitude of a relatively small number of respondents is, on the one hand, to completely believe in, or, on the other hand, not to believe at all in the activities of economic propaganda.

Table 3 The trust of insurance service users in the activities of economic propaganda of insurance companies

The attitudes of insurance services users	Mark	N	The number of users who provided an answer	% of users who provided an answer
I completely believe	5	145	14	9.6
I partially believe	4	145	69	47.59
I neither believe nor disbelieve	3	145	36	24.83
I slightly believe	2	145	20	13.80
I don't believe at all	1	145	6	4.14

Source: Authors' calculations

Unlike the indirect forms of marketing communication, the largest number of insurance service users, 93 of them, partially believe in the information they receive through communication with the sellers of insurance companies. The attitude of 26 respondents is that they completely believe in direct communication with the employees, 10 users slightly believe, while there are no users who do not believe at all. The respondents' attitudes regarding the fifth research question – whether insurance service users believe in the marketing messages which they receive from a direct contact with the employees – are shown in Table 4.

Table 4 The trust of insurance service users in the activities of personal selling of insurance companies

The attitudes of insurance service users	Mark	N	The number of users who provided an answer	% of users who provided an answer
I completely believe	5	145	26	17.93
I partially believe	4	145	93	64.15
I neither believe nor disbelieve	3	145	16	11.00
I slightly believe	2	145	10	6.92
I don't believe at all	1	145	0	0

Source: Authors' calculations

b) The analysis of the level of achieved satisfaction of insurance service users.

The quality of the complete offer of an insurance company has been evaluated by the respondents with the application of a Likert scale, where 5 is the highest mark, and 1 is the lowest. The respondents' answers to the sixth research question – the evaluation of the quality of the complete offer of insurance companies – are shown in Table 5. The results of the research indicate that a significant number of insurance service users have evaluated the quality of the complete offer of insurance companies with high marks. There have been no lowest marks, while only one user has evaluated the quality of the offer with the mark 2.

Table 5 The evaluation of the quality of the complete offer of insurance companies

Mark	N	The number of users who provided an answer	% of users who provided an answer	Mean value
5	145	70	48.28	
4	145	62	42.76	
3	145	12	8.27	4.39
2	145	1	0.69	
1	145	0	0	

Source: Authors' calculations

In the further course of the research, the respondents were supposed to express their attitudes about the level of achieved satisfaction with insurance services. The respondents' answers to the seventh research question – whether insurance service users are satisfied with the services of insurance companies – are shown in Table 6. The attitude of most insurance service users is that they are very satisfied (58), or somewhat satisfied (63) with the services of insurance companies. The presented research results indicate that there are neither dissatisfied nor very dissatisfied users, while there are 24 users who are indifferent.

Table 6 The coefficient of the satisfaction of insurance service users

The attitudes of insurance service users	Mark	N	The number of users who provided an answer	Value	Mean value
Very satisfied	5	145	58	290	
Somewhat satisfied	4	145	63	252	
Neither satisfied nor dissatisfied	3	145	24	72	4.23
Dissatisfied	2	145	0		
Very dissatisfied	1	145	0		

Source: Authors' calculations

c) The analysis of retaining insurance service users

Depending on the level of achieved satisfaction and the fulfilment of promises given in the pre-purchase stage through marketing communication, insurance service users will make a decision regarding further cooperation with an insurance company. The respondents' answers to the eighth research question – what decision of insurance service users will be in case of changed conditions – are shown in Table 7. Retaining insurance service users has been measured as a percentage of users out of the whole sample who claim that they will use the services of the same insurance company. The attitude of the majority of insurance service users (66.20%) is that, in case of changed conditions, they would think through

before they make a decision. The decision to still use the services of the same insurance company, regardless of the changes of insurance conditions, will be made by 20.70% of users, while 13.10% of users will change their insurance company.

Table 7 The ratio of retaining insurance service users

The attitudes of insurance service users	Mark	N	The number of users who provided an answer	% of users who provided an answer
Still use the services of the same insurance company	3	145	30	20.70
Think through before making a decision	2	145	96	66.20
Change the insurance company	1	145	19	13.10

Source: Authors' calculations

CONCLUSION

The paper analysed the importance of marketing communication for the user in reaching a decision on using insurance services. The aim of the research was to determine the influence of marketing communication on attracting and retaining insurance beneficiaries.

The empirical research was conducted on the sample of 145 insurance beneficiaries. The results of the research showed that there is a positive correlation between marketing communication, and informing and creating the consciousness of potential users about the insurance company's offer. Also, the results indicated that there is a positive correlation between marketing communication and the user's decision to remain loyal to the insurance company. On the assumption that they have a choice, beneficiaries will remain loyal to the company as long as they are getting higher value in comparison with what they expect. This implies that a well-defined and implemented strategy of marketing communication considerably contributes to the attraction of new and retainment of current insurance beneficiaries. The realized research makes a significant contribution to the research of the role of marketing communication in establishing, developing and advancing long-term relations between beneficiaries and insurance companies.

The research has certain limitations. Above all, the results refer to the service sector, i.e. the insurance sector, and do not take into consideration details of all the organizations and sectors. Also, the research was carried out on the sample which cannot be considered representative. In accordance with the above-mentioned points, the authors believe that further research in this field is necessary in order to overcome the weaknesses.

REFERENCES

Avramović, M. (2008). Strategija privredne propaganda u sistemu integrisanih marketing komunikacija [Economic propaganda strategy in the system of integrated marketing communications]. Ph.D. Thesis. University in Niš.

Becić, S., Stojanović, M., & Nikolić, M. (2018). Role of marketing and social networks in improving business effectiveness. *Ekonomika, 64(*3), 77-88.

Boateng, I., & Aguey, S. (2020). Effect of Customer Satisfaction and Marketing Communication Mix on Customer's Loyality in the Ghanian Banking Industry. *European Journal of Management and Marketing Studies*, *5(*4), 22-39. Retrived from: https://www.researchgate.net/publication/344151751_Issue_4_ 2020_EFFECT_OF_CUSTOMER_SATISFACTION_AND_MARKETING_COMMUNICATION_MIX_ ON_CUSTOMERS'_LOYALTY_IN_THE_GHANAIAN_BANKING_INDUSTRY Accessed on: November 01, 2020.

Blaga, L. (2018). Integrated Marketing Communication and Its Impact on Consumer Behavior. *Studies in Business and Economics*, *13*(2), 92-102. Retrieved from: https://content.sciendo.com/view/journals/sbe/13/2/article-p92.xml?language=en Accessed on: November 01, 2020.

Ćurčić, N., Piljan, I., & Simonović, Z. (2019). Marketing concept in insurance companies. *Ekonomika, 65*(3), 21-33.

Duncan, T. (2005). *Principles of Advertising & IMC*. 2nd edition. New York: McGraw-Hill, Irwin.

Kaftandjiev, Ch., & Shustova, E. (2018). *Fear Appeals in marketing Communications of Banks and Insurance Companies in the Continuum of Dramatic-Tragic Situations – The Semiotic Perspective.* Faculty of Business Economics and Entrepreneurship, 1-2, 83-110. Retrived from: https://www.researchgate.net/publication/331427975_Fear_appeals_in_marketing_communicatioof_banks_and_insurance_companies_in_the_continuum_of_dramatictragic_situations_The_Semiotic_Perspective Acessed on: November 05, 2020.

Keller, K. (2009). Building strong brands in modern marketing communications environment. *Journal of Marketing Communications*, *15*(2-3), 139-155.

Kotler, Ph., & Keller, K. (2006). *Marketing menadžment 12.izdanje [Marketing management 12th edition]*. Beograd: Data Status.

Kotler, Ph., & Keller, K. (2017). *Marketing menadžment 15.izdanje [Marketing management 15th edition]*. Beograd: Data Status.

Kumar, N., Noida, G., Gangal, D. V. K., & Singh, K. (2011). Advertising and consumer buying behaviour: a study with special reference to Nestle Ltd. *International Journal of Research in Commerce & Management*, *2*(10), 83-87.

Kumar, A. (2017). Effect of service quality on customer loyalty and the mediating role of customer satisfaction: an empirical investigation for the telecom service industry. *Journal of Management Research and Analysis*, *4*(4), 159-166.

Lovreta, S., Berman, B., Petković, G., Veljković, S., Crnković, J., & Bogetić, Z. (2010). *Menadžment odnosa sa kupcima [Customer relationship management]*. Beograd: CID Ekonomski fakultet.

Ljubojević, Č. (1998). *Menadžment i marketing usluga [Service management and marketing]*. Novi Sad: Stylos.

Marinković, V., & Senić, V. (2012). Analiza elemenata kvaliteta usluga u korporativnom bankarstvu [Analysis of service quality elements in corporate banking]. *Ekonomski horizonti* (januar-april *1*), 13-22.

Rabasović, B. (2013). Očekivanja potrošača u turizmu [Consumer expectations in tourism]. *Anali ekonomskog fakulteta u Subotici*, *(29),* 93-104.

Robson, J. (2015). General insurance marketing: A review and future research agenda. *Jornal of Financial Service Marketing*, *20*(4), 282-291. Retrieved from: https://www.researchgate.net/publication/281972587_General_insurance_marketing_A_review_and_future_research_agenda Accessed on: November 12, 2020.

Senić, R., & Senić, V. (2008). *Menadžment i marketing usluga [Service management and marketing]*. Kragujevac: Prizma.

Senić, R., & Senić, V. (2013). Sprečavanje odlaska postojećih i pospešivanje povratka izgubljenih kupaca [Preventing the departure of existing and facilitating the return of lost customers]. *Marketing*, *44*(4), 333-342.

Shimp, T., & Andrews, C. (2013). Advertising, Promotion and other aspects of Integrated Marketing Communications, 9th edition, South Western, Cengage Learning. Retrived from: https://fandango.home.xs4all.nl/Advertising%20Promotion%20and%20Other%20Aspects%20of%20Integrated%20Mng%20Communications%20%20Terence%20A.%20Shimp%20&%20J.%20Craig%20Andrews.pdf   Accessed on: November 12, 2020.

Stanković, Lj. (2002). *Međuzavisni marketing [Interdependent marketing]*. Niš: Ekonomski fakultet.

Stanković, Lj. & Đukić, S. (2014). *Marketing komuniciranje [Marketing communication]*. Niš: Ekonomski fakultet.

Stanković, J. (2018). Zeleni marketing i zeleni proizvod [Green marketing and green product]. *Economics of Sustainable Development*, *2*(2). 37-46.

Thaichon, P., Quach, T.N. & Lobo A. (2013). Marketing communications: Factors influencing brand loyalty of Internet service provider. Conference: *Australian and New Zealand Marketing Academy Conference* December 2013. Retrieved from: https://www.researchgate.net/publication/279970569_MARKETING_COMMUNICATIONS_FACTORS_INFLUENCING_BRAND_LOYALTY_OF_INTERNET_SERVICE_PROVIDERAccessed on: November 10, 2020

Veljković, S. (2009). *Marketing usluga [Services Marketing]*. Beograd: CID Ekonomski fakultet.

Vračar, D. (2005). *Strategije tržišnog komuniciranja [Marketing communication strategies]*. Beograd: CID Ekonomski fakultet.

Yousif, R. (2016). The Impact of Personal Selling on the Purchasing Behavior towards Clothes: A case Study on the Youth Category. *International Journal of Marketing Studies, 8*(5), 128-135.

Zeithaml, V. & Bitner, M. (2003). *Services Marketing, integrating customer focus acros the firm*. International edition. New York: McGraw-Hill Higher Education.

ZNAČAJ MARKETING KOMUNICIRANJA ZA PRIVLAČENJE I ZADRŽAVANJE KORISNIKA USLUGA OSIGURANJA

Pretopstavka uspeha procesa marketing komuniciranja u sektoru usluga je ostvarivanje dvosmerne interaktivne komunikacije između korisnika usluga i uslužnih organizacija.Uspešna strategija marketing komuniciranja usmerena je na privlačenje novih korisnika, ali i na zadržavanje i izgradnju dugoročnih odnosa sa postojećim korisnicima usluga. Marketing poruka treba da bude kreirana u skladu sa definisanim ciljevima, karakteristikama i preferencijama ciljnog auditorijuma. Uloga marketing komunikacija je različita u zavisnosti od faze procesa kupovine. U fazi pre kupovine,cilj marketing komunikacija je da informišu potencijalne korisnike i da kreiraju preferencije ka ponudi organizacije. U fazi korišćenja usluge, primenom adekvatnih oblika marketing komunikacija, organizacija nastoji da objasni i upozna korisnike sa procesom usluživanja i da pruži podršku prilikom rešavanja eventualnih problema. Potvrđivanje pozitivnih impresija korisnika i rešavanje eventualnih problema nastalih tokom procesa pružanja usluga, cilj je marketing komuniciranja u fazi posle kupovine.

Ključne reči: *marketing komuniciranje, korisnik usluga, privlačenje, satisfakcija, zadržavanje korisnika usluga*