Cena 15,00 zł
(VAT 5%)

Indeks 371262
ISSN 0033-2372

GŁÓWNY URZĄD STATYSTYCZNY
STATISTICS POLAND

PRZEGLĄD STATYSTYCZNY

STATISTICAL REVIEW

TOM 65

3

2018

100 lat GUS

Information for Authors

1. *Statistical Review* publishes scientific articles in the field of statistics, econometrics and other disciplines using quantitative methods to study economic phenomena. The submitted works should contain significant theoretical contributions or interesting empirical applications. Articles related to research conducted as part of research projects are especially required. Also, book reviews, reports on scientific activity of statisticians and econometricians, as well as studies containing original proposals in the field of didactics of statistics and econometrics are published in *Statistical Review*.
2. Articles are published in Polish and English. In the latter case, the author should send the English text carefully prepared and revised in terms of language.
3. A typescript with a maximum of 20 pages (including tables and diagrams, written using the Times New Roman 12-point font, with 1.5 line spacing, with margins of 2.5 cm) should be edited through the journal's editorial platform: http://www.editorialsystem.com/pst/
4. The literature that is cited should be listed in alphabetical order (in titles in English the first letter of the word should be capitalised), for example:
   Bauwens L., Laurent S., Rombouts J. V. K., (2006), Multivariate GARCH Models: A Survey, *Journal of Applied Econometrics*, 21 (1), 79–110.
   Brockwell P. J., Davis R. A., (1996), *Introduction to Time Series and Forecasting*, Springer-Verlag, New York.
   The text should use the Harvard referencing style, for example: This problem was discussed in the works of Granger (1969) and Brockwell and Davis (1996). This issue has been presented in many works (see *e.g.*, Granger, 1969; Brockwell and Davis, 1996; Bauwens *et al*., 2006).
5. If the work is divided into sections, the sections should be numbered with Arabic numerals. It is necessary to use continuous numbering for tables and figures (marked as Table 1, Table 2, *etc*., Figure 1, Figure 2, *etc.*).
6. Submitted papers should include (at the end of the paper) the title, a summary of the work (not exceeding 1/2 page of the typescript) and key words, all in Polish and English.
7. Please, also provide the information on the affiliation of the authors (name of the institution, faculty, department, postal address) and the e-mail address of the author responsible for correspondence in the footnote referring to the name. If the paper is an effect of a research project, this fact should be mentioned in the footnote referring to the title of the study, giving the number and title of the project.
8. Please be advised that in the process of reviewing the submitted papers, double anonymity will be maintained. Therefore, the paper should be sent in an anonymous version, and all identification information should be deleted.
9. Authors of papers accepted for publication are required to send scanned statements about the originality of the paper and the contribution of individual authors to the preparation of the publication as well as the transfer of the author's property rights.
10. Submission of an paper to *Statistical Review* and the fact of its acceptance for publication also means the author's consent for its publication on the website of the journal and in the databases of journals in which *Statistical Review* is included.
11. Papers that do not meet the requirements given will not be considered.

GUS

GŁÓWNY URZĄD STATYSTYCZNY
STATISTICS POLAND

PRZEGLĄD STATYSTYCZNY

STATISTICAL REVIEW

TOM 65

3

2018

WARSZAWA 2018

Strona www „Przegląd Statystyczny”:
http://www.przegladstatystyczny.pan.pl

Informacje w sprawie sprzedaży czasopisma tel.: 22 608 32 10, 22 608 38 10

ISSN 0033-2372

Indeks 371262

100 lat GUS Zakład Wydawnictw Statystycznych

ZAKŁAD WYDAWNICTW STATYSTYCZNYCH
al. Niepodległości 208, 00-925 Warszawa, tel. 22 608 31 45.
Zbigniew Karpiński (redaktor techniczny), Katarzyna Szymańska (skład i łamanie)

Zem. 516/2018 — nekłed 220 egz.

CONTENTS

DOI: 10.5604/01.3001.0014.0538

PRZEGLĄD STATYSTYCZNY
TOM LXV – ZESZYT 3 – 2018

Marta KORNAFEL[1]
Ivan TELEGA[2]

Natural capital in economic models[3]

1. INTRODUCTION

The problem of economic sustainability primarily covers issues of intergenerational equity (i.e. the concern for the well-being of future generations), the preservation of the capacity of natural capital to provide benefits important for social welfare, as well as the possibility of substituting natural capital with other forms of capital (Toman at al., 1995, p. 140). The role of natural capital, as the factor of growth and economic development, unfortunately is not emphasized strong enough in the mainstream economics. Sustainability concerns make the role of natural capital in the process of creating the social welfare more vivid. Many of interdisciplinary research indicate the crucial role of biodiversity for the ability of ecosystems to provide the ecosystem services (Cardinale at al., 2012, p. 59–67). There exists a fundamental discrepancy on the theoretical and methodological level between the mainstream and ecological economics. The most important issue is the substitutability of natural capital by another forms of capital (e.g. manufactured, human, etc.)

In the area of sustainability economics a lot of research was done in both, theoretical and empirical aspects (Pezzey, Toman, 2002). However, still the main approach is to use the perspective of neoclassical economics. It seems reasonable to introduce some aspects of ecological economics to the mainstream economics, in particular – to the analysis of the growth in the long run.

The goal of our paper is to make a critical analysis of selected growth models that use the notion of natural capital and construct the alternative model. In particular we treat the natural capital as a renewable resource and we use CES production function, weakening the assumption of substitutability of natural capital with other forms of capital. Therefore, our approach follows the main postulate of ecological economics, i.e. limited substitutability of natural resources.

---

[1] Cracow University of Economics, Faculty of Finance and Law, Department of Mathematics, 27 Rakowicka St., 31–510 Cracow, Poland, corresponding author – e-mail: marta.kornafel@uek.krakow.pl.

[2] Cracow University of Economics, Faculty of Finance and Law, Department of Mathematics, 27 Rakowicka St., 31–510 Cracow, Poland.

[3] Publication was financed from the funds granted to the Faculty of Finance and Law at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

2. THE BASIC ASPECTS OF ECOLOGICAL ECONOMICS – CONCLUSIONS FOR MODELLING

The natural capital is a quite new concept, being developed from the beginning of 90s. According to Constanza, Daly (1992) the natural capital is the extension of economic concept of capital as "a stock that yields a flow of valuable goods or services into the future" on environmental goods and services. The natural capital is understood differently in the mainstream economics and ecological economics. The mainstream economics focuses the attention on the role of natural resources (in particular – fossil fuels), while the ecological economics emphasizes those elements of natural capital, which creates ecosystems. The natural capital in the form of ecosystems provides many diverse ecological services for both, production and consumption, as well as for the maintenance of the life on the Earth. It may be said that the ecological services represent the stream of benefits, gained by humans from natural capital[4].

One of the first complete classifications of ecological services was the one proposed by Constanza at al. (1998, p. 253–260). The classification, which is most often referred to in the recent literature, is the one presented in the Millenium Ecosystem Assessment – thirty one ecological services were identified and grouped into four categories: supporting, provisioning, regulational and cultural (see Millenium Ecosystem Assessment, 2005, p. 40).

According to England (1998, p. 8) the starting point for defining the natural capital should be the theory of production by Georgescu-Roegen, which recognizes two main elements of production: funds elements, which represent the agents of production process, and flow elements, which are used and transformed by agents.

One of the most important problems that are considered in ecological economics is the issue of substitutability of natural capital by another form of capital (material production factors, knowledge, etc.). Ecological economics follows the rule of limited substitutability resulting in the idea of strong sustainability (Hediger, 2006).

Moreover, ecological economics postulates the existence of some limiting boundaries for usage of nature. Passing them makes a serious danger for ecosystems sustainability in the local and global scales. The attempt of estimation of those values was undertaken in the international research project (Rockström at al., 2009). According to critics setting the limits of growth is quite arbitrary if one takes into account that six of the mentioned limits, i.e. changes in the land use, loss of biodiversity, nitrogen level, consumption of the drinking water, chemical and aerosol pollutions, have local character, not global. Therefore there is no

[4] The links between biodiversity and the provision of services by ecosystems are important arguments for the protection of biodiversity and against the disappearance of species. However, as noted by Wilson (2002): "*The loss of the ivory-billed woodpecker has had no discernible effect on American prosperity. A rare flower or moss could vanish from the Catskill forest without diminishing the region's filtration capacity. But so what? To evaluate individual species by their known practical value at the present time is business accounting in the service of barbarism.*"

limit, that after passing it those processes start functioning in a fundamentally different way. There is too little evidence to state that breaking the limits in any of the mentioned areas would have negative influence on the social welfare (Nordhaus at al., 2012). The notion of critical natural capital is also developed within the ecological economics (Ekins, 2003; Chiesura, De Groot, 2003).

Taking into account the discussed issues we claim that the ecological growth models should meet the following assumptions:

1. The natural capital and other forms of capital have limited substitutability. This fact may be modelled in two nonexcluding ways – choosing a proper production function (CES or Leontieff) or taking some additional assumptions.
2. The natural capital has significant influence on the social welfare via economic sphere (resource function) as well as via direct influence on the well-being (regulatory and cultural functions). This should be captured by utility function, in particular in those models, where analysis is done from the social planner point of view.
3. The natural capital in the form of ecosystems is characterized by its internal dynamics and regeneration ability. At the same time the economic activity has negative effect on the resources usage and deterioration of natural environment, what influences the rate of regeneration and the availability of ecological services. Therefore the direction of evolution of natural capital is the result of several opposing factors.
4. The existence of the critical thresholds is postulated, so that the evolution of natural capital should be limited from below.

Ultimately, the adoption of a ”capital” definition of sustainability leads to a change in the research perspective. Instead of models describing the growth process of aggregate output (standard growth models), one should rather focus on modeling the process of shaping the resources of all types of capital, which are considered crucial for social well-being. Apart from the assumptions concerning the dynamics of particular types of capital, the assumptions as to the nature of mutual relations between them are also important.

3. SELECTED APPROACHES TO MODELLING NATURAL CAPITAL IN ECONOMIC PROCESSES

Literature on modelling the natural resource usage is abundant. However, the works in which authors directly use the term *natural capital* and try to describe its dynamics are less frequent. Another criterion for the selection of quoted results was that they present some characteristic and interesting ways of capturing the idea of natural capital in economic models - either using a particular form of production function (Kraev, 2002), considering a two-sector model (Commoli, 2006), a combination with the idea of material consumption (Rodrigues at al., 2005) or directly modeling the interaction between the four kinds of capitals (Roseta-Palma at al., 2010). Undoubtedly, in no way does our selection exhaust the rich literature of the subject.

The basic principles of ecological economics and the issues of formal description are discussed by Kraev (2002). Assuming strong complementarity between anthropogenic (H) and natural (N) capitals (what was expressed by the choice of Leontieff production function $Y = \min(AH, CN)$ [5]), an unbounded economic growth is impossible. However, the economy reacts differently in the long run depending whether the natural capital is treated as the stock or the flow (denoted by n) in the production function. The consequences of depletion of the flow are much more drastic – they lead to the zero production almost immediately (see Kraev, 2002, p. 281). As an attempt to weaken the assumption about strong complementarity Kraev considers a particular case of CES function:

$$Y = ((AH)^{-p} + (Cn)^{-p})^{-\frac{1}{p}}, \quad 0 < p < \infty, \tag{1}$$

where p is a parameter, characterizing the admissible rate of substitution[6].

The production function is characterized by weak complementarity, and simultaneously the main conclusion remains the same. If $AH \gg Cn$, then $Y \approx Cn$. Analogically to the production sector, the author extends the rules of ecological economics to consumption system. He assumes that the social welfare depends also on the ecological services (like clean air, water, the landscape), which are complementarities for the produced goods (market goods), usually considered in microeconomics. The utility function may have different form, nevertheless it should take into account the fact of limited substitutability.

Commoli (2006) considers two-sector economy, in which the produced capital K may be used for production of either, intermediate or final goods, i.e. $K = K_x + K_y$. The intermediate goods[7] are produced with use of capital K_x and natural resources X according to the production technology $M = H(K_x, X)$, in which the substitution is allowed. The final goods are produced with use of M and capital K_y, which are complementarities. Therefore:

$$M = H(K_x, X) = X^{\alpha} K_x^{1-\alpha} \;\; and \;\; F(M, K_y) = \min\left(M, \frac{K_y}{c}\right), \tag{2}$$

where $0 < \alpha < 1$ and $c > 0$ (parameters of model). The natural capital is considered in two forms: as the stock (X) and as the flow (M). Moreover, natural and man-made capitals are substitutes in intermediate goods sector and complementarities in the production of final goods.

[5] A, C are efficiency factors.

[6] Leontieff function is the limiting case for Y when $p \to \infty$.

[7] The intermediate goods may be understood as extracted resources.

Commoli assumes that the stock is renewable, i.e. $X = g(X) - M$, where g is strictly concave biological recruitment function. Most common assumption is that $g(X) = \gamma\, X\left(1 - \frac{X}{S}\right)$ with some parameter $\gamma > 0$ being intrinsic growth rate of the renewable resources and $S > 0$ being a parameter describing the environmental carrying capacity.

Under the standard assumptions on accumulation of produced capital, the author proves that there exists a stable stationary point satisfying $K > 0$ and $X > 0$ iff the condition: $\gamma > \left(\frac{s}{\delta} - c\right)^{\frac{1-\alpha}{\alpha}}$ is fulfilled (see Commoli, 2006, p. 159). If the economy is in the stationary state (i.e. $\dot{K} = 0$ and $\dot{X} = 0$), then the condition above takes form: $\gamma > \left(\frac{K_x}{X}\right)^{1-\alpha}$. As the right-hand side of the last inequality is not greater than $\frac{K_x}{X}$, the sufficient condition for existence of stationary point can be interpreted as follows: *the intrinsic growth rate of the renewable resource be at least as great as the long-run equilibrium ratio of manufactured to natural capital in the raw materials sector of economy* (Commoli, 2006, p. 159).

Rodrigues at al. (2005) propose a combination of neoclassical growth theory with the concept of allocation of natural capital and economy's dematerialization (the concept developed among others by Bringezu, 2003). They assume that anthropogenic impact depends on the degree of material intensity of the economy. Natural capital usage influences negatively the endogenous dynamics of ecosystems, reducing the volume of available ecological services. The authors show that under some conditions an unbounded growth is possible, keeping the natural capital at some constant level. In this paper the natural capital is divided between a production (fraction u) and "free" natural capital (fraction $1 - u$), directly influencing the social welfare. The natural capital has a character of renewable stock, but it also depends on carrying capacity CC), which evolves[8]. The dynamics of natural capital is (r is the growth parameter of N):

$$\frac{dN}{dt} = rN(CC - N) - P. \tag{3}$$

An increase of free natural capital increases CC, which grows with the growth of the free part of natural capital. This is why the dynamics of CC is described by (see Rodriguez at al., 2005, p. 385):

$$\frac{dCC}{dt} = \frac{l}{(1-u)N}\,\frac{d(1-u)N}{dt} = \frac{l}{N}\frac{dN}{dt} - \frac{l}{1-u}\frac{du}{dt}. \tag{4}$$

where $l > 0$ is a constant parameter determining the growth of CC. The dynamics of natural capital is affected by structural influences (via dependence on u)

[8] Assuming a constant CC implies that despite the damage in ecosystem it may always renew to the equilibrium value defined by CC. Notice that CC is the upper bound of natural capital stock.

and is characterized by endogenous dynamics of CC. Anthropogenic pressure P depends on amount of materials used in production and consumption. If material intensity is denoted by m, then $P = mY$. Notice that m is a quantity dependent on the technology A and the production Y. The authors assume that the structure of economy changes in result of the growth of production, what influences the quantity m. Therefore:

$$P = m_0 A^{-a} Y^n, \qquad m = m_0 A^{-a} Y^{n-1}, \tag{5}$$

where a, n, m_0 are constant parameters[9].

The production and capital accumulation are assumed in form:

$$\dot{K} = Y - C - \delta K, \tag{6}$$

$$Y = AK^{\alpha} (uN)^{1-\alpha}. \tag{7}$$

In addition, the dependence between growth rates of technology and capital are assumed:

$$\frac{\dot{A}}{A} = g\left(\frac{\dot{K}}{K}\right), \tag{8}$$

with some increasing, concave and bounded function g, for which $g(\cdot) \equiv 0$ for negative arguments. The utility function accounts the benefits from "free" natural capital part:

$$U = \ln C + \phi \ln\big((1-u)N\big). \tag{9}$$

An important characteristic of this function is that consumption is independent of environmental conditions. In biophysical steady-state $\dot{N} = \dot{CC} = 0$, what implies constant natural capital stock. The authors show that the economic growth is possible with constant level of natural capital, if the ratio of parameters a/n is bounded[10], i.e.:

$$1 + \frac{a}{g_0'} \leq \frac{a}{n} \leq \frac{1}{1-\alpha}. \tag{10}$$

Moreover, if the constraints (dependent on N^* and CC^*) on initial values of A and K are met, then it is possible to maintain increasing consumption (see Rodriguez at al., 2005, p.393).

[9] m_0 is scale parameter. If $n < 1$, then structural changes lead to a decrease in m. For example, the increase in share of service sector leads to smaller material intensity of the economy.

[10] Here $g_0' = g'(0)$ is maximal value of g'.

In many works, the role of human and social capital in creating prosperity is highlighted. Attempts to combine four capitals (i.e. produced, human, social and natural) into one model were undertaken by Roseta-Palma at al. (2010), thus referring to the ”capital” definition of sustainable development proposed by Pearce (see Pearce, Atkinson, 1998, p. 9).

It is also worth noting that attempts were also made to empirically evaluate the stock of individual capital types (World Bank, 2006, 2011), one of the conclusions of which was to indicate the specific role of human and social capital for the development of countries. It seems that the advantage of the proposed model is an attempt to take account of the relationships between different kinds of capitals. By denoting K_P, K_H, K_S, K_N respectively manufactured, human, social and natural capital, the evolution of each and the interactions are defined as follows (Roseta-Palma at al., 2010, p. 604):

$$\dot{K}_P = Y - C - \delta_P K_P. \tag{11}$$

Human capital can be used in the production, education, accumulation of social capital and environmental protection (research and development, emission reduction, etc.). Thus, respectively, $K_H = H_Y + H_H + H_S + H_N$ and:

$$\dot{K}_H = \xi H_H + \alpha K_S - \delta_H K_H, \tag{12}$$

where $\xi > 0, \alpha \geq 0$ are efficiency parameters.

Notice that human and social capitals are substitutable in the creation of human capital. Social capital is ”produced” with the use of human capital (via the creation of appropriate institutions and regulations), but at the same time its dynamics at all times depends on the size of the social capital, i.e.:

$$\dot{K}_S = \omega\, H_S + \Omega\, K_S, \tag{13}$$

where ω is again the efficiency parameter, and Ω may be positive or negative. The natural capital is again renewable resource, i.e.:

$$\dot{K}_N = R(K_N) - N_Y + P, \tag{14}$$

where $R(K_N)$ is natural regeneration rate, defined similarly to Rodrigues at al. (2005). N_Y denotes the stream of natural resources used in production, and P represents the positive effect of environmental protection. P depends positively on social capital K_S and the human capital H_N engaged in environmental protection, while negatively on manufactured capital K_P. This dependence is described in the form:

$$P = m_0 \frac{H_N^{\epsilon} K_S^{\kappa}}{K_P^{\varphi}}, \tag{15}$$

where m_0 is the scale parameter, and $\epsilon, \kappa, \varphi$ are respective elasticities. Production is aggregated via Cobb-Douglas technology, i.e. the substitutability between capitals is allowed:

$$Y = K_P^{\beta} N_Y^{\nu} H_Y^{\eta} K_S^{\sigma}, \qquad \beta + \nu + \eta = 1. \tag{16}$$

The essential feature of the model is the broader definition of social welfare, which depends not only on consumption, but also on the state of the environment (natural capital stock) and the level of trust and cooperation in society (social capital):

$$U(C, K_N, K_S) = \frac{\tau}{\tau - 1} \int_0^{\infty} (C K_N^{\Phi} K_S^{\Psi})^{\frac{\tau - 1}{\tau}} e^{-\rho t} \, dt, \tag{17}$$

where τ is the coefficient of elasticity of intertemporal substitution, whereas Φ, Ψ are the parameters of the preferences of natural capital (nature) and social capital respectively. Solving the problem of dynamic optimization, the authors derive the constant growth rates of Y and K_H in the steady-state.

4. CRITICAL REMARKS AND ALTERNATIVE MODEL

These approaches, despite a number of simplifying assumptions, allow for a holistic consideration of natural capital in the processes of wealth creation. Natural capital occurs both as a renewable resource (being an argument in utility function) and as a flow (a resource used in production). Understanding the anthropogenic pressure in line with the material flow concept has the advantage that it does not reduce the problem only to the stream of pollutants emitted or used energy carriers. It seems that the indicators of material requirements are the best measures of the consumption of natural capital by individual countries, while the amount of materials consumed by the economy is currently being estimated by Eurostat. The significant disadvantage of the majority of models is the use of Cobb-Douglas production functions, i.e. allowing for substitution between individual capitals. In the light of the theory of ecological economy this is very unlikely. Creating a realistic model requires limited factor substitutions, but without establishing strong complementarity. Using CES can be a reasonable compromise.

The most difficult issue is to consider the role of technology. There are many approaches in the literature that model the process of technology development emphasizing category of knowledge accumulation (Romer, 2012), innovation process (Howitt, Aghion, 1999) or human capital (Lucas, 1988). However, they are substantially different. In the context of this work, it is legitimate to use the category of human capital. Taking into account the chosen assumptions of the models discussed in part 3 and the differences in human capital approach, we propose the following specifications for the growth model. Let K, H, N denote manufactured, human and natural capital, respectively. Human capital H is interpreted in the sense of Mankiw at al. (1992), i.e. it is generated by investing part of the economic output Y. We exclude social capital in this case in order to simplify the model. Aggregated product is produced using individual capitals, but aggregation is made using the CES function[11] (cf. Kraev, 2002) with $p \in (0, \infty)$:

$$Y = f(K, H, N) = (\alpha\, K^{-p} + \beta\, H^{-p} + \gamma\, N^{-p})^{-\frac{1}{p}}. \tag{18}$$

The production is divided between consumption, investment in produced capital and investment in human capital (education):

$$Y = C + I + E. \tag{19}$$

The dynamics of produced and human capital is:

$$\dot{K} = s_K Y - \delta K = Y - C - E - \delta K, \tag{20}$$

$$\dot{H} = s_H Y = E, \tag{21}$$

where s_K and s_H are respectively the investment rates in produced and human capitals[12]. It is possible to set them constant or use dynamic programming approach, when C and E are control variables. In what follows we consider the optimal control problem. The natural capital is a renewable resource, diminished by environmental pressure[13]:

$$\dot{N} = rN - P. \tag{22}$$

where the pressure and material intensity are defined as (Rodrigues at al., 2005):

---

[11] CES function assumes a limited substitution between factors, including human and produced capital. Note, however, that we do not prejudge the degree of substitution, which depends on the parameter p. Taking into account the different degree of substitution between the factors requires the use of a nested CES function. The use of the nested production function seems to be a more realistic description of the economy, but we have abandoned this to simplify the model.

[12] Unlike Mankiw at al. (1992), we have abandoned the depreciation of human capital in order to simplify the model.

[13] Chapter 2 discusses the approach in which CC is variable, as does K_N is allocated between economic use and "free" part, according to u. This leads to the dynamics in the form $\dot{N} = rN(CC - N) - P$. However, the assumption of fixed natural capital at stationary state implies constant CC and u, therefore we simplify the model, choosing the form of eq. (22).

$$m = m_0 H^{-a} Y^{n-1}, \qquad P = mY \equiv m_0 H^{-a} Y^n. \tag{23}$$

The social welfare is defined as the value of functional:

$$W = \int_0^\infty e^{-\rho t}\, U(C,N)\, dt, \tag{24}$$

with the utility function $U(C,N) = \ln C + \ln N$[14]. Our goal is to maximize it and study the quantitative properties of solutions.

In our considerations we assume all the functions to be differentiable in their domains. It implies immediately that all the differential equations of the model have unique solutions. The control variables are consumption C and investment in human capital E, while the state variables are all types of capital: K, H, N. Notice that U is concave and due to economic meaning of controls we may assume them to be uniformly bounded. Therefore the functional W attains maximum for some values $C = C^*$ and $E = E^*$. The corresponding capital paths are denoted by K^*, H^* and N^* respectively. The current value hamiltonian is:

$$\mathcal{H}(C,E;\ K,H,N) = U(C,N) + \lambda_K(Y - C - E - \delta K) + \lambda_H E + \lambda_N(rN - P). \tag{25}$$

4.1. Conclusions from the structure of the model

The standard way to start the analysis of the model is to assume that there exists a steady-state, when all the variables have constants rates of growth (we follow the standard notation for rate: g_X for variable X). However, as in (Rodrigues at al., 2005), we have to make additional assumption that our economy is in the biophysical equilibrium, i.e. $\dot{N} = 0$, because it is impossible for natural capital to grow without bounds. Then $rN = P$ and $P = m_0 H^{-a} Y^n$ are constant. We derive from here:

$$g_Y = \frac{a}{n} g_H. \tag{26}$$

Additionally we have $mY = const$, what gives:

$$g_m = -g_Y. \tag{27}$$

[14] The utility function has been simplified to make the model and later calculations more legible. In general, you must enter the parameters differentiating the marginal utility of consumption and natural capital. For the needs of the model, we can assume that the units of measure of both goods are chosen to make unit utilities of both goods equal.

We have to consider two possibilities. If N is constant and $K \gg N$, $H \gg N$, it is impossible for Y to grow with constant rate. The main obstacle is the production function with limited sustainability. To see this, recall Kraev (2002, p. 283). With constant N, and K and H approaching infinity we have:

$$\lim_{K,H\to\infty} Y = \lim_{K,H\to\infty} (\alpha\, K^{-p} + \beta\, H^{-p} + \gamma\, N^{-p})^{-\frac{1}{p}} = \gamma^{-\frac{1}{p}} N, \tag{28}$$

so g_Y approaches zero, as K and H grows with constant N. Taking into account (26), we obtain that $g_Y = g_H = 0$. For the proposed production function:

$$g_Y = \alpha \left(\frac{K}{Y}\right)^{-p} g_K + \beta \left(\frac{H}{Y}\right)^{-p} g_H + \gamma \left(\frac{N}{Y}\right)^{-p} g_N, \tag{29}$$

Ultimately, we have zero growth rates for all variables in the long-run. In other words, with the assumption of limited substitutability and constant natural capital stock, unbounded steady state economic growth is not possible. Obviously, some kind of technological progress either in the form of an increase in the efficiency of natural capital use in the production function, or in the form of increasing total factor productivity is the only way to overcome this obstacle. The question about the nature of this progress, i.e. exogenous or endogenous, is still under consideration. It appears that this conclusion is in line with the concept of steady state economy by Daly (1980) with constant capital stock and a constant population size. We hope to take this into consideration in the further work.

If N is abundant, i.e. $N \gg H$ and $N \gg K$, we can assume the possibility of steady-state growth. In this case biophysical equilibrium boils down to:

$$g_Y = \alpha \left(\frac{K}{Y}\right)^{-p} g_K + \beta \left(\frac{H}{Y}\right)^{-p} g_H. \tag{30}$$

Taking into account (26), we obtain:

$$\frac{a}{n} g_H = \alpha \left(\frac{K}{Y}\right)^{-p} g_K + \beta \left(\frac{H}{Y}\right)^{-p} g_H, \tag{31}$$

and finally:

$$\left(\frac{a}{n} - \beta \left(\frac{H}{Y}\right)^{-p}\right) g_H = \alpha \left(\frac{K}{Y}\right)^{-p} g_K. \tag{32}$$

The rates g_H and g_K are constant by assumption, so differentiating the last identity with respect to time t we get:

$$\left(\frac{K}{H}\right)^p = \frac{\alpha g_K(g_K - g_Y)}{\beta g_H(g_Y - g_H)}. \tag{33}$$

We conclude that the ratio $\frac{K}{H}$ is constant and:

$$g_K = g_H = \frac{n}{a} g_Y. \tag{34}$$

Equation (21) implies that in the steady-state $\frac{E}{H} = const$, so:

$$g_E = g_H. \tag{35}$$

Therefore by (20) we get:

$$g_K = \frac{Y - C}{K} - \frac{E}{K} - \delta, \tag{36}$$

which leads to the observation that the ratio $\frac{Y-C}{K}$ is constant. This may happen if and only if:

$$g_Y = g_C = g_K. \tag{37}$$

Combining (34) and (37) we get the necessary condition of the steady state: $a = n$, i.e. P elasticities of technology and output are equal. In particular, the conclusion about equal growth rates boils down (30) to:

$$\alpha\left(\frac{Y}{K}\right)^p + \beta\left(\frac{Y}{H}\right)^p = 1. \tag{38}$$

However, it is worth noting that the condition $a = n$ is very unlikely, so in the given model the steady state occurs with a probability close to zero. At the same time, if $a = n$ then by (34) and (37) we have $g_Y = g_C = g_K = 0$. This leads to conclusions about the inability of long-term growth under steady state assuming constant natural capital.

4.2. Optimization conditions

Despite the impossibility of unlimited economic growth in the long-run, we still can analyze the conditions maximizing social welfare. Pontriagin Maximum Principle provides the following necessary conditions for optimal controls and paths:

$$\begin{cases} \frac{\partial \mathcal{H}}{\partial C} = 0 \\ \frac{\partial \mathcal{H}}{\partial E} = 0 \\ \rho\lambda_K - \dot{\lambda}_K = \frac{\partial \mathcal{H}}{\partial K} \\ \rho\lambda_H - \dot{\lambda}_H = \frac{\partial \mathcal{H}}{\partial H} \\ \rho\lambda_N - \dot{\lambda}_N = \frac{\partial \mathcal{H}}{\partial N} \end{cases} \Leftrightarrow \begin{cases} \lambda_K = \frac{1}{C}, & (39.1) \\ \lambda_K = \lambda_H, & (39.2) \\ \rho\lambda_K - \dot{\lambda}_K = \lambda_K\left(\frac{\partial Y}{\partial K} - \delta\right) - \lambda_N \frac{\partial P}{\partial Y}\frac{\partial Y}{\partial K}, & (39.3) \\ \rho\lambda_H - \dot{\lambda}_H = \lambda_K \frac{\partial Y}{\partial H} - \lambda_N\left[\frac{\partial P}{\partial H} + \frac{\partial P}{\partial Y}\frac{\partial Y}{\partial H}\right], & (39.4) \\ \rho\lambda_N - \dot{\lambda}_N = \frac{\partial U}{\partial N} + \lambda_K\frac{\partial Y}{\partial N} + \lambda_N r - \lambda_N\frac{\partial P}{\partial Y}\frac{\partial Y}{\partial N}, & (39.5) \end{cases}$$

Transversality conditions take the form:

$$\lim_{t\to+\infty} e^{-\rho t}\lambda_K(t)K(t) = 0, \quad \lim_{t\to+\infty} e^{-\rho t}\lambda_H(t)H(t) = 0, \quad \lim_{t\to+\infty} e^{-\rho t}\lambda_N(t)N(t) = 0. \quad (40)$$

Basing on (18) and (23) we derive the formulas for partial derivatives:

$$\frac{\partial Y}{\partial H} = \beta\left(\frac{Y}{H}\right)^{p+1}, \qquad \frac{\partial Y}{\partial K} = \alpha\left(\frac{Y}{K}\right)^{p+1}, \qquad \frac{\partial Y}{\partial N} = \gamma\left(\frac{Y}{N}\right)^{p+1}, \quad (41)$$

$$\frac{\partial P}{\partial Y} = \frac{nP}{Y}, \qquad \frac{\partial P}{\partial H} = -\frac{aP}{H}. \quad (42)$$

From $(39.1) - (39.2)$ we immediately have $\lambda_{K^*} = \lambda_{H^*} = \frac{1}{C^*}$ and consequently:

$$g_{\lambda_K^*} = g_{\lambda_H^*} = -g_{C^*}. \quad (43)$$

In the standard way (see Kamien, Schwartz, 2012, p. 138) we may calculate the shadow price of natural capital:

$$\lambda_{N^*}(t) = \frac{\partial W}{\partial N}(t) = \int_t^{+\infty} \frac{\partial U}{\partial N} e^{-r(s-t)} ds = \frac{1}{rN^*}. \quad (44)$$

Therefore $g_{\lambda_{N^*}} = -g_{N^*}$. Now we turn our attention to conditions (39.3) and (39.5). We divide the first equation by λ_K, the second one by λ_N. Using (43) and (44) we conclude:

$$g_{N^*} = 2r - \rho + \gamma\,\frac{nP^*}{Y^*}\left(\frac{Y^*}{N^*}\right)^{p+1} \frac{\rho + \delta + g_{C^*}}{\alpha\left(\frac{Y^*}{K^*}\right)^{p+1} - (\rho + \delta + g_{C^*})}. \quad (45)$$

$$N^* = C^*\,\frac{1}{r}\,\frac{nP^*}{Y^*}\,\frac{\alpha\left(\frac{Y^*}{K^*}\right)^{p+1}}{\alpha\left(\frac{Y^*}{K^*}\right)^{p+1} - (\rho + \delta + g_{C^*})}. \quad (46)$$

Next, from equations (39.3) and (39.4), we derive the dependencies:

$$g_{C^*} = \alpha \left(\frac{Y^*}{K^*}\right)^{p+1} \left(1 - \frac{\lambda_{N^*}}{\lambda_{K^*}} \frac{nP^*}{Y^*}\right) - \delta - \rho, \tag{47}$$

$$g_{C^*} = \beta \left(\frac{Y^*}{H^*}\right)^{p+1} \left(1 - \frac{\lambda_{N^*}}{\lambda_{K^*}} \frac{nP^*}{Y^*}\right) + \frac{\lambda_{N^*}}{\lambda_{K^*}} \frac{aP^*}{H^*} - \rho. \tag{48}$$

Eliminating the ratio $\frac{\lambda_N}{\lambda_K}$ from those equations, we express the growth rate of consumption in terms of average productivities of manufactured and human capitals:

$$g_{C^*} = \left[1 + \frac{\rho}{\frac{Y^*}{H^*}}\right] \frac{\alpha \left(\frac{Y^*}{K^*}\right)^{p+1}}{\beta \left(\frac{Y^*}{H^*}\right)^{p} - 1} - (\delta + \rho). \tag{49}$$

The conclusions that we conduct from (45) and (49) are:

a. $\frac{\partial g_{C^*}}{\partial (Y/K)} > 0$, i.e. raising the average productivity of manufactured capital leads to greater consumption rate. Similar result for the growth rate of natural capital holds under the following condition on g_{C^*}:

$$\frac{\partial g_{N^*}}{\partial (Y/K)} > 0 \quad \Leftrightarrow \quad g_{C^*} < \frac{\alpha}{\beta} \frac{\frac{Y^*}{H^*} + \rho}{\frac{Y^*}{H^*}\left[\left(\frac{Y^*}{H^*}\right)^{p} - 1\right]} \left(\frac{Y^*}{K^*}\right)^{p+1} - \rho - \delta. \tag{50}$$

b. $\frac{\partial g_{C^*}}{\partial (Y/H)} < 0$, i.e. (surprisingly) raising the average productivity of human capital leads to smaller consumption rate. Positive impact of this raise on the rate of growth for natural capital is possible, if the consumption rate satisfies:

$$\frac{\partial g_{N^*}}{\partial (Y/H)} > 0 \quad \Leftrightarrow \quad g_{C^*} > \frac{\alpha}{2} \left(\frac{Y^*}{K^*}\right)^{p+1} - \rho - \delta. \tag{51}$$

c. $\frac{\partial g_{C^*}}{\partial (Y/N)} = 0$, while

$$\frac{\partial g_{N^*}}{\partial (Y/N)} > 0 \quad \Leftrightarrow \quad g_{C^*} < \alpha \left(\frac{Y^*}{K^*}\right)^{p+1} - \rho - \delta. \tag{52}$$

d. Greater depreciation rate δ results in smaller consumption rate, while $\frac{\partial g_{N^*}}{\partial \delta} = 0$.

e. Increase in p (so equivalently: decrease in substitutability) causes decrease in g_{C^*} provided:

$$\begin{cases} \ln\frac{Y^*}{K^*} < \frac{\beta\left(\frac{Y^*}{H^*}\right)^p}{\beta\left(\frac{Y^*}{H^*}\right)^p - 1}\ln\frac{Y^*}{H^*}, & if \quad \left(\frac{Y^*}{H^*}\right)^p > \frac{1}{\beta}, \\ \ln\frac{Y^*}{K^*} > \frac{\beta\left(\frac{Y^*}{H^*}\right)^p}{\beta\left(\frac{Y^*}{H^*}\right)^p - 1}\ln\frac{Y^*}{H^*}, & if \quad \left(\frac{Y^*}{H^*}\right)^p < \frac{1}{\beta}. \end{cases} \tag{53}$$

Increase in p has negative influence on g_{N^*} if:

$$\ln\frac{Y^*}{N^*} < \ln\frac{Y^*}{H^*}\,\frac{\beta\left(\frac{Y^*}{H^*}\right)^{p+1}}{\beta\left(\frac{Y^*}{H^*}\right)^{p+1} - 2\frac{Y^*}{H^*} - \rho}. \tag{54}$$

Assuming constant rate of consumption on optimal path $g_{C^*} = const$, we may have additional conclusion from (49) about influence of average productivity of human capital on the rates of production, manufactured and human capitals:

$$\frac{g_{Y^*} - g_{H^*}}{g_{Y^*} - g_{K^*}} = \frac{(p+1)\left[\left(\frac{Y^*}{H^*}\right)^2 + \rho\right]}{p\left[\left(\frac{Y^*}{H^*}\right)^2 + \rho\right]\frac{\beta\left(\frac{Y^*}{H^*}\right)^{p+1}}{\beta\left(\frac{Y^*}{H^*}\right)^p - 1} - \rho\left(\frac{Y^*}{H^*}\right)^2}. \tag{55}$$

4.3. Coexistence of the steady-state and optimal growth path

In this section we are going to investigate the properties of steady state, being simultaneously the optimal growth path. Combining condition (38) (constant combination of productivities of manufactured and human capitals) with the formula on optimal consumption growth rate (49):

$$g_{C^*} = -\left[1 + \frac{\rho}{\frac{Y^*}{H^*}}\right] - (\delta + \rho) < 0. \tag{56}$$

In view of equal growth rates given by $(34), (35)$ and (37), we immediately get, that the economy collapses. On the other side, by differentiation of (49) with respect to time, we conclude that $g_{Y^*} = 0$, and therefore $g_{C^*} = 0$, what gives contradiction. Therefore in our model it is impossible to have steady-state, which simultaneously realizes welfare maximum.

5. CONCLUSIONS

The main conclusion we can derive from the discussed model in the impossibility of the economic growth in the long run if the limited substitutability and constant natural capital are assumed. Technological progress either in the form of the increase of natural capital efficiency or in the form of increasing total factor productivity seems to be the only way to overcome this limit. If the stock of natural capital is abundant, we can assume steady state growth (again, we must note that steady-state is very unlikely due to necessity of $a = n$), although it is still impossible to have steady-state growth, which simultaneously realizes welfare maximum.

Other conclusions resulting from the model are:

1. Raising the average productivity of manufactured capital leads to greater consumption rate. Similar result for the growth rate of natural capital holds if growth rate of consumption is limited from above.
2. Raising the average productivity of human capital leads surprisingly to smaller consumption rate. Positive impact of this raise on the rate of growth for natural capital is possible, if the consumption growth rate is limited from below. It could be interpreted as follows: faster consumption growth inhibits economic growth by reducing investment opportunities, thus slowing down the growth of natural capital usage P, which is positively dependent on Y. As a result, a faster growth of natural capital is possible.
3. Greater depreciation rate δ results in smaller consumption rate.
4. Increase in p (so equivalently: decrease in substitutability) could lead to decrease in g_{C^*}. Increase in p could also have negative influence on g_{N^*}. In both cases decrease in substitutability means lower growth rates.

Obviously, more in-depth reflection on the results is needed as well as an attempt to modify the model to enable sustainable growth, i.e. constant positive growth rate of production and consumption with constant natural capital stock. The model presented in the paper is just one of the many possibilities for describing the behaviour of the economy, thus another specification of assumptions, in particular the assumption on substitutability of natural capital with other kinds of capital, would potentially allow the long-term growth with preserved natural capital.

REFERENCES

Bringezu S., Schutz H., Moll S., (2003), Rationale for and Interpretation of Economy-Wide Materials Flow Analysis and Derived Indicators, *Journal of Industrial Ecology*, 7 (2), 43–64.

Cardinale B. J., Duffy J. E., Gonzalez A., Hooper D. U., Perrings C., Venail P., Narwani A., Mace G. M., Tilman D., Wardle D., Kinzig A. P., Daily G. C., Loreau M., Grace J. B., Larigauderie A., Srivastava D. S., Naeem S., (2012), Biodiversity Loss and Its Impact on Humanity, *Nature*, 486, 59–
–67.

Chiang, A. C., (2002), *Elementy dynamicznej optymalizacji*, WSHiFM, Warszawa.

Chiesura A., De Groot R., (2003), Critical Natural Capital: a Socio-cultural Perspective, *Ecological Economics*, 44 (2–3), 219–231.

Comolli P., (2006), Sustainability and Growth when Manufactured Capital and Natural Capital are Not Substitutable, *Ecological Economics*, 60, 157–167.

Constanza R., Daly H. E., (1992), Natural Capital and Sustainable Development, *Conservation Biology*, 6 (1), 37–46.

Costanza R., d'Arge R., de Groot R., Farber S., Grasso M., Hannon B., Limburg K., Naeem S., O'Neill R.V., Paruelo J., Raskin R.G., Sutton P., van den Belt M. (1997), The Value of the World's Ecosystem Services and Natural Capital, *Nature*, 387, 253–260.

Daly H., (1980), *Economics, Ecology, Ethics. Essays Towards a Steady-State Economy*, 2nd ed., W.H. Freeman and Company, San Francisco.

Ekins P., (2003), Identifying Critical Natural Capital: Conclusions about Critical Natural Capital, *Ecological Economics*, 44 (2–3), 277–292.

England R. W., (1998), Should We Pursue Measurement of the Natural Capital Stock?, *Ecological Economics*, 27 (3), 257–266.

Fischer-Kowalski M., Krausmann F., Giljum S., Lutter S., Mayer A., Bringezu S., Moriguchi Y., Schütz H., Schandl H., Weisz H., (2011), Methodology and Indicators of Economy-wide Material Flow Accounting. State of the Art and Reliability Across Sources, *Journal of Industrial Ecology*, 15 (6), 855–876.

Hediger W., (2006), Weak and Strong Sustainability, Environmental Conservation and Economic Growth, *Natural Resource Modeling*, 19 (3), 359–394.

Howitt P., Aghion P., (1999), *Endogenous Growth Theory*, The MIT Pres, Cambridge.

Kamien M. I., Schwartz N. L., (2012), *Dynamic Optimization. The Calculus of Variations and Optimal Control in Economics and Management*, Dover Publications Inc.

Kraev E., (2002) Stocks, Flows and Complementarity: Formalizing a Basic Insight of Ecological Economics, *Ecological Economics*, 43, 277–286.

Lucas R. E. Jr., (1988), On the Mechanics of Economic Development, *Journal of Monetary Economics*, 22 (1), 3–42.

Mankiw, N. G., Romer D., Weil D. N., (1992), A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics*, 107 (2), 407–437.

Nordhaus, T., Shellenberger M., Blomqvist L., (2012), Planetary Boundaries. A Review of the Evidence, Breakthrough Institute, http://thebreakthrough.org.

Pearce D., Atkinson D., (1998), The Concept of Sustainable Development: An Evaluation of Its Usefulness Ten Years After Brundtland, *Swiss Journal of Economics and Statistics*, 134 (3), 251–269.

Pezzey J., Toman M. A., (2002), The Economics of Sustainability: A Review of Journal Articles, Resources for the Future, www.rff.org

Rockström J., Steffen W., Noone K., Persson A., Stuart III Chapin F., Lambin E., Lenton T. M., Scheffer M., Folke C., Schellnhuber H.J., Nykvist B., de Wit C. A., Hughes T., van der Leeuw S., Rodhe H., Sörlin S., K. Snyder P. K., Costanza R., Svedin U., Falkenmark M., Karlberg L., Corell R. W., Fabry V.J., Hansen J., Walker B., Liverman D., Richardson K., Crutzen P., Foley J., (2009), Planetary Boundaries: Exploring the Safe Operating Space for Humanity, *Ecology and Society*, 14 (2), 32.

Rodrigues, J., Domingos T., Conceição P., Belbute J., (2005), Constraints on Dematerialisation and Allocation of Natural Capital Along a Sustainable Growth Path, *Ecological Economics*, 54, 382–396.

Romer D., (2012), *Advanced Macroeconomics*, 4th ed., The McGraw-Hill Series in Economics, New York.

Roseta-Palma C., Ferreira-Lopes A., Neves Sequeira T., (2010), Externalities in an Endogenous Growth Model with Social and Natural Capital, *Ecological Economics*, 69, 603–612.

The Millennium Ecosystem Assessment, (2005), *Ecosystems and Human Well-being: Synthesis*, Island Press, Washington, DC.

Toman M. A., Pezzey J., Krautkraemer J., (1995), Neoclassical Growth Theory and Sustainability, in: Bromley D. W., (ed.), *The Handbook of Environmental Economics*, Blackwell, Oxford.

Wilson E. O., (2002), *The Future of Life*, Knopf Doubleday Publishing Group.

World Bank, (2006), Where is the Wealth of Nations? Measuring Capital for the 21st Century, http://siteresources.worldbank.org.

World Bank, (2011), The Changing Wealth of Nations: Measuring Sustainable Development for the New Millennium, https://siteresources.worldbank.org.

KAPITAŁ NATURALNY W MODELACH EKONOMICZNYCH

Streszczenie

Celem artykułu jest dokonanie krytycznej analizy wybranych modeli wzrostu proponowanych w ramach szkoły ekonomii ekologicznej oraz odwołujących się do kategorii kapitału naturalnego, jak również próba konstrukcji alternatywnego modelu. W szczególności traktujemy kapitał naturalny jako odnawialny zasób i używamy funkcji produkcji CES, tym samym ograniczając możliwości substytucji kapitału naturalnego innymi formami kapitału. Analizowane są optymalne (tj. maksymalizujące dobrobyt społeczny) ścieżki kapitału i konsumpcji. Artykuł kończy się wnioskami formułowanymi na podstawie modelu.

Słowa kluczowe: ekonomia ekologiczna, trwałość, wzrost gospodarczy, kapitał naturalny, funkcja produkcji CES

NATURAL CAPITAL IN ECONOMIC MODELS

Abstract

The goal of our paper is to make a critical analysis of selected growth models that use the notion of natural capital and to construct the alternative model. In particular we treat the natural capital as a renewable resource and we use CES production function, weakening the assumption of substitutability of natural capital with other forms of capital. We investigate the optimal paths for capital and consumption, giving their characterization in the dependence on the parameters of the model. The paper ends with conclusions derived from the model.

Keywords: ecological economics, sustainability, growth model, natural capital, CES production function

DOI: 10.5604/01.3001.0014.0540 PRZEGLĄD STATYSTYCZNY
TOM LXV – ZESZYT 3 – 2018

Tadeusz KUFEL[1]
Sławomir PLASKACZ[2]
Joanna ZWIERZCHOWSKA[3]

Strong and safe Nash equilibrium in some repeated 3-player games[4,5]

1. INTRODUCTION

The Prisoner's Dilemma with its generalizations are very important as an example of conflicts and social dilemmas. As we can find in Dawes (1980), social dilemmas are real life problems which have two properties: ”1. each individual receives a higher payoff for a socially defection choice than for a socially cooperative choice, no matter what the other individuals in society do; 2. all individuals are better off if all cooperate than if all defect.” An example of such situation in real life is a problem of soldiers who fight in a battle. They are personally better off taking no chances, yet if no one fight against the enemy, then the result will be worst for all of soldiers. Such dilemmas can be found among resource depletion, pollution and overpopulation.

Social dilemmas are games in which there is a conflict between individual rationality and optimality of the equilibrium payoff. Since it is observable that people cooperate with each other in the real situations, game theorists have faced the obstacle, how to construct simple tools to encourage players in such games to cooperate with each other. The model need to approximate the real situation and strategies should be likely to use.

A natural approach is to consider the infinitely repeated game. Usually, all players observe the whole history of action profiles used in previous stages of the repeated game. Such situation is called the game with complete information.

[1] Nicolaus Copernicus University in Toruń, Faculty of Economic Sciences and Management, Department of Econometrics and Statistics, 13a Gagarina St., 87–100 Toruń, Poland.

[2] Nicolaus. Copernicus University in Toruń, Faculty of Mathematics and Computer Science, Department of Nonlinear Mathematical Analysis and Topology, 12/18Chopina St., 87–100 Toruń, Poland, corresponding author – e-mail: plaskacz@mat.umk.pl.

[3] Nicolaus Copernicus University in Toruń, Faculty of Mathematics and Computer Science, 12/18 Chopina St., 87–100 Toruń, Poland.

[4] This research was financially supported by the project ”Wzmocnienie potencjału dydaktycznego UMK w Toruniu w dziedzinach matematyczno-przyrodniczych” (Poddziałanie 4.1.1 POKL).

[5] We would like to thank the referees for helpful comments and corrections.

The strategies are functions from the set of the histories into the set of actions. Payoffs in the repeated game are either the discounted sum of stage payoffs or the limit of average payoffs. The aim of this approach is to obtain the Nash equilibrium in the repeated game with the pair of payoffs which is close to the cooperation payoffs in the stage game. Since the fifties of the last century there appeared various folk theorems which was not explicitly published and, in many cases, the original author is unknown.

The classic Prisoner's Dilemma is a 2-player game, in which each player has two actions, usually denoted as C (cooperation) and D (defection). The game has a unique Nash equilibrium – a pair of actions such that the action of each of the players optimize this player's payoff given the action of the opponent. The Nash equilibrium is the action profile (D, D) which is the pair of strictly dominant actions i.e. playing D is better than C whatever the other player does. What is more, both players benefit changing (D, D) into (C, C). So, the mechanism of individual rationality fails in the Prisoner's Dilemma and it leads to a loss of both players. It means that the Nash equilibrium is not Pareto-optimal in this case.

One of solutions for lack of cooperation of the Nash equilibrium in the stage game is an idea of good strategies introduced in Smale (1980) for the repeated Prisoner's Dilemma. Every pair of good strategies is a Nash equilibrium in the repeated game with Pareto-optimal payoffs corresponding to the payoff of (C, C) in the stage game. The second advantage of the good strategies equilibrium is the warranted minimal payoff for the non-deviating player. The minimal payoff is equal to the Nash payoff in the stage game. Good strategies have yet another advantage that has not been pointed in Smale (1980). Choosing a good strategy appropriately, the player controls the second player's payoff. For every $\varepsilon > 0$ there exists the ε-good strategy of the first player such that for an arbitrary second player's strategy, the first player's payoff will be at most ε smaller than the second player's payoff. The Prisoner's Dilemma is symmetric, so the second player also can choose the ε-good strategy which provides him no worse payoffs than the first player's one minus ε.

In fact, good strategies have properties that was postulated in Axelrod (1984). In 80's he studied the evolution of cooperation. It refers to how cooperation can emerge and persist as elucidated by application of game theory. He organized a tournament in which game theory experts submitted their strategies and each strategy was paired with each other for 200 iterations of Prisoner's Dilemma. Accumulated payoffs through the tournament was treated as a score. The winner was the strategy submitted by Arnold Rappaport – Tit for Tat. The additional advantage of this tournament was detecting what properties strategies should satisfy to encourage players to cooperate. They should be: nice, forgiving, retaliatory and are founded on simple rules. Good strategies have these properties and, what is more, player cooperates until the other player's average payoff is greater than his average payoff plus ε. By choosing ε, the player determines the level of his tolerance for the defection.

In this paper we shall consider the generalization of the idea of good strategies onto the Prisoner's Dilemma type repeated game for three players. We consider the repeated game with a partial monitoring. We assume that, after each stage, all players can only observe an aggregated history – the arithmetic mean of the payoffs from previous stages. The stage game is a symmetric 3-player game where each player has an action set consisting of two actions: I and NI[6]. We assume that the action profile (NI, NI, NI) is the only Nash equilibrium, and the sum of the players payoffs is minimal for this profile. The sum of players payoffs is maximal for the profile (I, I, I). The strategy profile in the repeated game is a function $s: S \to \{I, NI\}^3$, where $S \subset R^3$ is a convex hull of the set of the payoffs in the stage game.

An example of a game considered in the paper is given in Example 2 (section 3). The strategy profile $(\mathrm{NI}, \mathrm{NI}, \mathrm{NI})$ is the only Nash equilibrium. The common payoff corresponding to the equilibrium profile ($\sum = 60$) is the lowest possible one. The strategy NI dominates the strategy I, i.e. the action NI gives higher payoff then the action I despite of the action of other players. So, the example has properties typical for real-life situations called tragedy of commons. The rational player should choose the action NI that dominates I but in real-life situations the cooperation is often observed (comp. Axelrod, 1984). So, it appears a question "How to explain theoretically a player inclination to cooperation that is observed practically?" It is known that one of the strongest factors that motivate cooperative behavior is the repetition of the game. In the paper we assume that the game is repeated infinitely times. Infinite time horizon well approximate real life situations of finite (≥ 20) but unknown time horizons. Our aim is to construct an equilibrium strategy profile in the repeated game that motivates every player to cooperation. We assume that after every repetition players know the average payoff of every player from previous stages. Briefly speaking, an equilibrium strategy of player i bases on the comparison of her average payoff $\mathrm{x_i}$ with average payoffs $\mathrm{x_j}$, $\mathrm{x_k}$ of remaining players. Player i cooperates (chooses I) if $\mathrm{x_j} < \mathrm{x_i} + \varepsilon$ and $\mathrm{x_k} < \mathrm{x_i} + \varepsilon$. If one of the remaining players' average payoff is greater to $\mathrm{x_i} + \varepsilon$ then she stops cooperation and chooses NI. Precise definition of an ε-good strategy is given in (19). The positive constant ε is a measure of player's tolerance for others players defection.

Our aim is to construct a strategy profile s^* which is an approximated strong Nash equilibrium in the repeated game under consideration. The constructed equilibrium is safe in the meaning that the payoff of a player choosing strategy $\mathrm{s_i^*}$ is not less then the equilibrium payoff in the stage game. This payoff is assured even if the other two players choose an arbitrary strategy. Furthermore, the ε-good strategy guarantees that, in long time horizon, other player's average payoff will not exceed the good strategy player's average payoff by more than ε.

[6] From now on, we choose to name strategies with I and NI, where I means invest and it corresponds to strategy C and NI corresponds to strategy D.

In the framework of the repeated game, a ε-good strategy is an individually rational strategy. It theoretically explains the players' inclination to cooperation in the repeated three players Prisoner's Dilemma.

The notion of the strong equilibrium in the framework of repeated games was introduced by Aumann (1959, 1961, 1967), who showed that every payoff that belongs to the β-core of the stage game is a strong equilibrium payoff in the corresponding repeated game (comp. Sorin, 1992, Thm. 6.2.2). Despite the fact that the payoff corresponding to the profile (I, I, I) belongs to the β-core, our result is not exactly a case of the Aumann results. We have dropped the assumption of the full monitoring. Players do not observe the full history, i.e. the sequence of actions selected by all players in the previous periods. Instead, we assume that they observe the aggregate history, i.e. the arithmetic mean of the previous payoffs of all the players. It is worth noting that the results on the existence of strong equilibria (comp. Konishi, 1997 and Nessah, 2014) do not apply to the repeated game considered in the present paper.

The repeated Prisoner's Dilemma for more than two players has been considered in Behrstock (2015). The ε-good strategies constructed in the paper have some additional properties to the strategies in Behrstock (2015), in which the authors base on similar approchability results as we do in this paper. The difference is that authors consider N-players Prisoner's Dilemma Game in which strategies are stochastic processes. In our approach all strategies are deterministic.

The paper is organized as follows. In section 2 we present the basic information about sequences related to a map of a convex set. We adopt Blackwell's approachability method (comp. Blackwell, 1956) which was originally used in the framework of 2-player repeated games with vector payoffs. We show that the Blackwell condition is sufficient to obtain the convergence of the sequence of arithmetic means to a set called a weak attractor. The weak attractors introduced in subsection 2.1 have different properties in comparison with approachable sets in the sense of Blackwell. We provide an example of a singleton being a weak attractor that does not satisfy the Blackwell condition. Such a situation is not possible for approachable sets (comp. Shani, 2014, Thm. 8). In repeated games, there is considered a sequence of vector payoffs. Each payoff corresponds to one repetition of the state game. Subsection 2.1 provides us necessary results to analyze the directions in which the trajectory shifts and to examine the convergence of such sequence. This is crucial for defining the payoff in the repeated game. Subsection 2.2 provides basic properties of the Banach limit which shall be used to prove that ε-good strategies are ε Nash equilibria. In some of our arguments we not only require that the sequence of mean payoffs converges to a set, but that almost all its entries belong to the set. A similar problem named strong approachability was considered in Shani (2014). In section 2.3 we adopt a Lyapunov function method for discrete and discontinuous dynamical systems to obtain a deterministic strong approachability result.

In section 3 we consider a repeated 3-player symmetric game. Every player has two actions: invest (I) or not invest (NI). The vector payoff $\mathrm{B}=(\mathrm{p}_3,\mathrm{p}_3,\mathrm{p}_3)$ corresponding to the strategy profile $(\mathrm{I},\mathrm{I},\mathrm{I})$ is Pareto optimal and the strategy profile $(\mathrm{NI},\mathrm{NI},\mathrm{NI})$ is a Nash equilibrium in the stage game with the payoff vector $(\mathrm{r}_0,\mathrm{r}_0,\mathrm{r}_0)$). We assume that, in the repeated game, every player knows the average vector payoff from the previous stages of the game. The strategy $\mathrm{s}_\mathrm{i}\colon \mathrm{S}\to\{\mathrm{I},\mathrm{NI}\}$, $\mathrm{i}\in\{1,2,3\}$, is a function from the convex hull S of vector payoffs in the stage game to the set of his actions $\{\mathrm{I},\mathrm{NI}\}$. The strategy profile $\mathrm{s}=(\mathrm{s}_1,\mathrm{s}_2,\mathrm{s}_3)$ and the vector payoff function $\mathrm{G}\colon\{\mathrm{I},\mathrm{NI}\}^3\to\mathrm{R}^3$ determine the function $\varphi=\mathrm{G}\circ \mathrm{s}\colon\mathrm{S}\to\mathrm{S}$. The strategy profile s and the initial point $\mathrm{x}_1\in\mathrm{S}$ determine the trajectory $\bar{\mathrm{x}}_\mathrm{n}(\mathrm{s},\mathrm{x}_1)$ of a dynamic system given by

$$\bar{\mathrm{x}}_{\mathrm{n}+1}=\frac{\mathrm{n}\bar{\mathrm{x}}_\mathrm{n}+\varphi(\bar{\mathrm{x}}_\mathrm{n})}{\mathrm{n}+1}.$$

Our aim is to construct a strategy profile $\mathrm{s}_\varepsilon^*=(\mathrm{s}_1^*,\mathrm{s}_2^*,\mathrm{s}_3^*)$ such that for every $\mathrm{x}_1\in\mathrm{S}$

$$\lim_{\mathrm{n}\to\infty}\bar{\mathrm{x}}_\mathrm{n}=\mathrm{B}, \tag{1}$$

where $\bar{\mathrm{x}}_\mathrm{n}=\bar{\mathrm{x}}_\mathrm{n}(\mathrm{s}_\varepsilon^*,\mathrm{x}_1)$. If one player (for example player 3) deviates then

$$\limsup_{\mathrm{n}\to\infty}\bar{\mathrm{x}}_\mathrm{n}^3\le\mathrm{p}_3+\varepsilon, \tag{2}$$

where $\bar{\mathrm{x}}_\mathrm{n}=\bar{\mathrm{x}}_\mathrm{n}((\mathrm{s}_1^*,\mathrm{s}_2^*,\mathrm{s}_3),\mathrm{x}_1)$ and $\mathrm{s}_3\colon\mathrm{S}\to\{\mathrm{I},\mathrm{NI}\}$ is an arbitrary strategy of player 3. If two players deviate (for example players 2 and 3) then

$$\limsup_{\mathrm{n}\to\infty}(\bar{\mathrm{x}}_\mathrm{n}^2+\bar{\mathrm{x}}_\mathrm{n}^3)\le 2\mathrm{p}_3, \tag{3}$$

$$\liminf_{\mathrm{n}\to\infty}\bar{\mathrm{x}}_\mathrm{n}^1\ge\mathrm{r}_0, \tag{4}$$

$$\lim_{\mathrm{n}\to\infty}\mathrm{dist}\,(\bar{\mathrm{x}}_\mathrm{n},\{\mathrm{x}\in\mathrm{S};\ \mathrm{x}_2\le\mathrm{x}_1+\varepsilon,\ \mathrm{x}_3\le\mathrm{x}_1+\varepsilon\})=0, \tag{5}$$

where $\bar{\mathrm{x}}_\mathrm{n}=\bar{\mathrm{x}}_\mathrm{n}((\mathrm{s}_1^*,\mathrm{s}_2,\mathrm{s}_3),\mathrm{x}_1)$ and $\mathrm{s}_2,\mathrm{s}_3\colon\mathrm{S}\to\{\mathrm{I},\mathrm{NI}\}$ are the arbitrary strategies of players 2 and 3, respectively. By $\mathrm{dist}(\mathrm{x},\mathrm{A})$ we denote the distance from the point x to the set A, i.e. $\mathrm{dist}(\mathrm{x},\mathrm{A})=\inf\{|\mathrm{x}-\mathrm{a}|\colon \mathrm{a}\in\mathrm{A}\}$.

If the payoff is a Banach limit (comp. Conway, 1985) of the sequence of average payoffs then the strategy profile s_ε^* is a strong ε-Nash equilibrium in the repeated game as a consequence of (1–3). Property (4) implies that the non-deviating player's payoff is no smaller than the payoff corresponding to the Nash equilibrium in the stage game. Property (5) guarantees that the deviating player's payoff will not exceed the good strategy player's payoff by more than ε. The results presented in Theorems 3.1, 3.2, 3.3 give a partial answer to the question asked by Smale in the last Remark in section 1 of Smale (1980, p. 1623).

Section 4 contains concluding remarks. In Appendix we present the proofs of theorems from subsection 2.3.

2. PRELIMINARIES

2.1. Approachability results

Let H be a finite dimensional vector space and $\langle\cdot,\cdot\rangle$, $|\cdot|$ denote an inner product and a norm in H, respectively. We assume that S is a nonempty convex closed subset of H. By $\mathrm{N}^{\varepsilon}(\mathrm{B})$ we denote an ε-neighbourhood of the set B in S, i.e. $\mathrm{N}^{\varepsilon}(\mathrm{B}) = \{\mathrm{x} \in \mathrm{S}: \mathrm{dist}(\mathrm{x},\mathrm{B}) < \varepsilon\}$. The closure (the convex hull) of the set A we denote by $\mathrm{cl}(\mathrm{A})$ $(\mathrm{co}(\mathrm{A}))$.

We study limit properties of sequences $(\bar{\mathrm{x}}_{\mathrm{n}})_{\mathrm{n}=1}^{\infty}$ defined by a map $\varphi: \mathrm{S} \to \mathrm{S}$ and an initial point $\mathrm{x}_1 \in \mathrm{S}$ by

$$\bar{\mathrm{x}}_{\mathrm{n}+1} = \frac{\mathrm{n}\bar{\mathrm{x}}_{\mathrm{n}} + \varphi(\bar{\mathrm{x}}_{\mathrm{n}})}{\mathrm{n}+1}, \quad \bar{\mathrm{x}}_1 = \mathrm{x}_1, \tag{6}$$

The sequence $(\bar{\mathrm{x}}_{\mathrm{n}})$ can be interpreted as a sequence of arithmetic means $\bar{\mathrm{x}}_{\mathrm{n}} = \frac{1}{\mathrm{n}}(\mathrm{x}_1 + \cdots + \mathrm{x}_{\mathrm{n}})$, where $\mathrm{x}_{\mathrm{k}+1} = \varphi(\bar{\mathrm{x}}_{\mathrm{k}})$. The map φ defines a dynamical system $\beta_{\mathrm{n}}: \mathrm{S} \to \mathrm{S}$ by

$$\beta_{\mathrm{n}}(\mathrm{x}) = \frac{\mathrm{nx} + \varphi(\mathrm{x})}{\mathrm{n}+1}, \qquad \mathrm{n} = 1, 2, \ldots .$$

We denote by $\bar{\mathrm{x}}_{\mathrm{n}}(\varphi, \mathrm{x}_1)$ a trajectory determined by (6).

We say that a closed set $\mathrm{A} \subset \mathrm{S}$ is a weak attractor for a dynamic system determined by the map φ if for every $\mathrm{x}_1 \in \mathrm{S}$ we have

$$\lim_{\mathrm{n}\to\infty} \mathrm{dist}(\bar{\mathrm{x}}_{\mathrm{n}}(\varphi, \mathrm{x}_1), \mathrm{A}) = 0,$$

where $\mathrm{dist}(\cdot, \mathrm{A})$ denotes the distance to the set A. We provide some sufficient conditions for being a weak attractor.

First we formulate Blackwell approachability type theorem that originally was presented in Blackwell (1956) in the framework of repeated games with vector payoffs. We say that a map $\varphi: \mathrm{S} \to \mathrm{S}$ satisfies the Blackwell condition for a set $\mathrm{A} \subset \mathrm{S}$ in the domain $\mathrm{D} \subset \mathrm{S}$ if

$$\forall \mathrm{x} \in \mathrm{D}, \exists \mathrm{y} \in \Pi_{\mathrm{A}}(\mathrm{x}), \quad \langle \mathrm{x} - \mathrm{y}, \varphi(\mathrm{x}) - \mathrm{y} \rangle \leq 0, \tag{7}$$

where $\Pi_{\mathrm{A}}(\mathrm{x})$ denote the set of points in A that are proximal to x, i.e. $\Pi_{\mathrm{A}}(\mathrm{x}) = \{\mathrm{a} \in \mathrm{A}: |\mathrm{a} - \mathrm{x}| = \mathrm{dist}(\mathrm{x}, \mathrm{A})\}$.

The deterministic version of the Blackwell approachability result can be formulated in the following way.

Proposition 2.1 Suppose that the map $\varphi: S \to S$ satisfies the Blackwell condition for a closed set $A \subset S$ in the domain $D \subset S$. If almost all elements of the bounded sequence $\bar{x}_n(\varphi, x_1)$ belong do the set D then

$$\lim_{n\to\infty} \mathrm{dist}(\bar{x}_n, A) = 0.$$

We provide the proof of Proposition 2.1 for the reader convenience.

Proof: For a sufficiently large n we choose $y \in \Pi_K(\bar{x}_n)$ and then

$$\mathrm{dist}(\bar{x}_{n+1}, A)^2 \leq |\bar{x}_{n+1} - y|^2 = \left|\frac{n}{n+1}(\bar{x}_n - y) + \frac{1}{n+1}(\varphi(x_n) - y)\right|^2 =$$

$$= \left(\frac{n}{n+1}\right)^2 |\bar{x}_n - y|^2 + \left(\frac{1}{n+1}\right)^2 |\varphi(x_n) - y|^2 + 2\frac{n}{(n+1)^2}\langle \bar{x}_n - y, \varphi(x_n) - y\rangle \leq$$

$$\leq \left(\tfrac{n}{n+1}\right)^2 \mathrm{dist}(\bar{x}_n, A)^2 + \left(\tfrac{1}{n+1}\right)^2 C,$$

where C is an upper bound of $\mathrm{dist}(x_n, A)$. Setting $d_n = n^2 \mathrm{dist}(\bar{x}_n, A)^2$ we have $d_{n+1} \leq d_n + C$ for $n \geq n_0$. Thus $d_n \leq d_{n_0} + (n - n_0)C$. So

$$\mathrm{dist}(\bar{x}_n, A)^2 \leq \tfrac{1}{n}\left(d_{n_0} + \tfrac{n-n_0}{n} C\right). \qquad \text{QED}$$

Corollary 2.2 If the map $\varphi: S \to S$ satisfies the Blackwell condition for a closed set $A \subset S$ in the domain S, then the set A is a weak attractor for φ. If the set $A \subset S$ is convex and the map $\varphi: S \to A$ maps into the set A then the set A is a weak attractor for φ.

Taking $A = (-\infty, c]$ in Proposition 2.1 we obtain the following property of real sequences.

Corollary 2.3 Suppose that $(a_n)_{n=1}^{\infty}$ is a bounded sequence in $\mathbb{R}$ and $(\bar{a}_n)_{n=1}^{\infty}$ is the sequence of arithmetic means, i.e. $\bar{a}_n = \frac{1}{n}\sum_{k=1}^{n} a_k$. If we have

$$(\bar{a}_n > c \quad \Rightarrow \quad a_{n+1} \leq c)$$

for almost all n and a fixed constant $c \in \mathbb{R}$, then

$$\limsup_{n\to\infty} \bar{a}_n \leq c.$$

In many cases the set A is a weak attractor despite that the Blackwell condition is not satisfied. Such a situation occurs in repeated games that we study in section 3. Below we present two properties of weak attractors which are necessary for our reasoning.

Proposition 2.4 Suppose that the sets $\mathrm{A}, \mathrm{B} \subset \mathbb{R}^{\mathrm{d}}$ are nonempty closed and B is bounded. If a sequence $\mathrm{x_n}$ satisfies

$$\lim_{n\to\infty} \mathrm{dist}(\mathrm{x_n}, \mathrm{A}) = \lim_{n\to\infty} \mathrm{dist}(\mathrm{x_n}, \mathrm{B}) = 0,$$

then

$$\lim_{n\to\infty} \mathrm{dist}(\mathrm{x_n}, \mathrm{A} \cap \mathrm{B}) = 0.$$

Proof: We choose $\mathrm{a_n} \in \mathrm{A}$, $\mathrm{b_n} \in \mathrm{B}$ such that

$$|\mathrm{x_n} - \mathrm{a_n}| = \mathrm{dist}(\mathrm{x_n}, \mathrm{A}), \quad |\mathrm{x_n} - \mathrm{b_n}| = \mathrm{dist}(\mathrm{x_n}, \mathrm{B})$$

Since the set B is compact, we obtain that the sequences $(\mathrm{a_n})$, $(\mathrm{b_n})$, $(\mathrm{x_n})$ are bounded and they have the same nonempty set C of accumulating points. Thus $\lim_{n\to\infty} \mathrm{dist}(\mathrm{x_n}, \mathrm{C}) = 0$ and $\mathrm{C} \subset \mathrm{A} \cap \mathrm{B}$.

QED

Proposition 2.5 We suppose that a closed set $\mathrm{A} \subset \mathrm{S}$ is a weak attractor for the map $\varphi\colon \mathrm{S} \to \mathrm{S}$ and a closed subset $\mathrm{B} \subset \mathrm{A}$ satisfies

$$\forall \varepsilon > 0, \exists \delta > 0, \quad \begin{array}{l} \varphi \text{ satisfies the Blackwell condition} \\ \text{for the set } cl\big(N^{\varepsilon}(B)\big) \cap \ A \text{ in the domain } N^{\delta}(A). \end{array} \tag{8}$$

Then the set B is a weak attractor for φ.

Proof: Fix $\mathrm{x_1} \in \mathrm{S}$ and $\varepsilon > 0$. By (8), we choose $\delta > 0$ such that almost all elements of the trajectory $\bar{\mathrm{x}}_\mathrm{n}(\varphi, \mathrm{x_1})$ belongs to $\mathrm{N}^{\delta}(\mathrm{A})$. By Proposition 2.1, we obtain

$$\lim_{n\to\infty} \mathrm{dist}(\bar{\mathrm{x}}_\mathrm{n}, \mathrm{cl}(\mathrm{N}^{\varepsilon}(\mathrm{B})) \cap \mathrm{A}) = 0.$$

Thus

$$\limsup_{n\to\infty} \mathrm{dist}(\bar{\mathrm{x}}_\mathrm{n}, \mathrm{B}) \le \varepsilon.$$

QED

The method illustrated in Proposition 2.4 and Proposition 2.5 bases on the scheme that we explain in the following example.

Example 1 Let $S = \mathbb{R}^2$, $a, b \in \mathbb{R}^2$, $a_2 < 0$, $b_2 > 0$, $a_1 \neq b_1$ and

$$\varphi(x, y) = \begin{cases} a & \text{if } y > 0, \\ b & \text{if } y \leq 0. \end{cases}$$

We show that $\lim_{n\to\infty} \bar{x}_n = d$ for every $\bar{x}_1 \in \mathbb{R}^2$, where the limit d is the point of intersection of the interval $\overline{ab}$ with the line $p = \{(x, y): y = 0\}$. The set $D = \{d\}$ does not satisfy condition (7). Indeed, if $a_1 < b_1$ and $x > d_1$ then $\varphi(x, 0) = b$ and $\langle (x, 0) - (d_1, d_2), \varphi(x, 0) - (d_1, d_2) \rangle > 0$. To show that the set D is a weak attractor we point out weak attractors A, B such that $D = A \cap B$. We set $A = p$ and $B = \overline{ab}$. The sets A, B satisfy the Blackwell condition (7). By Theorem 2.1, we have

$$\lim_{n\to\infty} \operatorname{dist}(\bar{x}_n, A) = \lim_{n\to\infty} \operatorname{dist}(\bar{x}_n, B) = 0.$$

Applying Proposition 2.4 we obtain that $\lim_{n\to\infty} \bar{x}_n = d$.

Finally, we shall formulate a property of the dynamical system.

Proposition 2.6 If the set S is bounded then for every $\xi > 0$ there exists $N \in \mathbb{N}$ such that for all $n > N$ and for all $x \in S$

$$|\beta_n(x) - x| < \xi,$$

where the map $\varphi: S \to S$ determining β_n is arbitrary.

2.2. Payoff in the repeated game

Considering a sequence of payoffs in the repeated games we always receive a bounded sequence. As we presented in (6), the dynamic is the vector of the arithmetic mean of the payoffs received in the previous repetitions. To analyze such sequence, the following proposition shall be useful.

Proposition 2.7 Suppose that $a_0, a_1, \dots, a_k \in \mathbb{R}^d$ and let $T \in \mathbb{N}$. Then for all $\varepsilon > 0$ and for all $n_1, \dots, n_k \geq 0$ such that $n_1 + \dots + n_k = n$, where n is sufficiently large, we have

$$\frac{T}{T+n} a_0 + \frac{n_1}{T+n} a_1 + \dots + \frac{n_k}{T+n} a_k \in N_\varepsilon(\mathrm{co}\{a_1, \dots, a_k\}).$$

Proposition 2.7 is a consequence of the fact that $\frac{T}{T+n} a_1 + \frac{n_1}{T+n} a_1 + \dots + \frac{n_k}{T+n} a_k \in \mathrm{co}\{a_1, \dots, a_k\}$ and $\frac{T}{T+n} |a_0 - a_1|$ is small where n is sufficiently large.

To define the payoff in repeated games we shall use the Banach limit (comp. Conway, 1985). The Banach limit L is a continuous linear functional definite on the space l^∞ of bounded scalar sequences. If (x_n) is a bounded sequence of points in R^d then $\mathrm{Lim}\,(x_n) := (\mathrm{Lim}\,(x_{n1}), \mathrm{Lim}\,(x_{n_2}), \dots, \mathrm{Lim}\,(x_{nd}))$, where $x_n = (x_{n1}, x_{n2}, \dots, x_{nd})$. So Banach Limit can be extended onto the space of bounded sequences of points in R^d. If $\varphi: R^d \to R$ is a linear functional then $\varphi(\operatorname{Lim}(x_n)) = \operatorname{Lim}(\varphi(x_n))$.

Proposition 2.8 If A is a compact convex subset of R^d and a sequence $(x_n) \subset R^d$ satisfies $\lim_{n\to\infty} \operatorname{dist}(x_n, A) = 0$, then $\operatorname{Lim}(x_n) \in A$.

Proof. Suppose to the contrary that $\operatorname{Lim}(x_n) \notin A$. Then there exists a functional $\varphi: R^d \to R$ such that $\varphi(\operatorname{Lim}(x_n)) > \sup_{a\in A}\varphi(a)$. We have $\limsup_{n\to\infty}\varphi(x_n) \leq \sup_{a\in A}\varphi(a)$. Thus

$$\varphi(\operatorname{Lim}(x_n)) = \operatorname{Lim}(\varphi(x_n)) \leq \limsup_{n\to\infty}\varphi(x_n) \leq \sup_{a\in A}\varphi(a)$$

which gives the contradiction. QED

2.3. A lapunov type results

The Lapunov function method is typically used to study stability of equilibrium points for dynamical systems. Using the Lapunov function method we obtain a strong approachability result for a dynamical system determined by a multivalued map.

Let H be a Hilbert space and $p_1, \dots, p_k \in H$ be unit vectors, i.e. $|p_i| = 1$. We define a function $V: H \to \mathbb{R}$ by

$$V(x) = \max_{i\in\{1,\dots,k\}} V_i(x) \qquad \text{where } V_i(x) = \langle p_i, x\rangle. \tag{9}$$

The function V is a support function of the set $\{v_1, \dots, v_k\}$. So, the function V is convex, positively homogeneous and lipschitz continuous with the constant $L = 1$ (see [11]).

Set

$$\Delta_c = \bigcap_{i=1}^{k} \{x \in H: V_i(x) < c\} = \{x \in H: V(x) < c\}.$$

Let us denote by $\varphi: S \to S$ a multivalued map of a subset $S \subset H$.

Definition 2.9 We say that V is the Lapunov type function for the multivalued map φ with the constant $c > 0$ if

$$\exists 0 < \delta < c, \forall x \in S\backslash\Delta_c, \forall i = 1, \dots, k, \forall \omega \in \varphi(x),\ (V_i(x) \geq V(x) - \delta \ \Rightarrow V_i(\omega) \leq 0). \tag{10}$$

If the function V satisfies

$$\forall x \in S\backslash\Delta_c, \forall \omega \in \varphi(x), \forall i \in \{1, \dots, k\} \quad V_i(x) > 0 \ \Rightarrow \ V_i(\omega) \leq 0, \tag{11}$$

then V is the Lapunov type function for φ with the constant c.

If V is the Lapunov type function for φ with the constant c and $c_1 > c$ then V is the Lapunov type function for φ with the constant c_1.

To explain why we say that V is the Lapunov type function observe that if $V_i(x) = V(x)$ then $p_i \in \partial V(x)$, where $\partial V(x)$ is the subdifferential of a convex function. The condition (10) implies the following inequality

$$\langle p_i, \omega - x\rangle \leq \langle p_i, \omega\rangle - \langle p_i, x\rangle \leq 0 - V(x) + \delta < \delta - c < 0 \quad \text{for } x \in S\backslash\Delta_c,$$

which means that V is the Lapunov function for the vector field $f(x) = \omega - x$.

Proposition 2.10 Let S be a nonempty bounded convex subset of H and the function $V: H \to \mathbb{R}$ given by (9) be the Lapunov type function for the multivalued map $\varphi: S \to S$ with the constant $c > 0$. If a sequence $(\bar{x}_n)_{n=1}^{\infty}$ satisfies

$$\bar{x}_1 = x_1 \in S, \quad \bar{x}_{n+1} = \frac{n\bar{x}_n + x_{n+1}}{n+1}, \quad x_{n+1} \in \varphi(\bar{x}_n), \tag{12}$$

then

$$\forall c_1 > c, \exists N, \; \forall n \geq N, \quad \bar{x}_n \in \Delta_{c_1}.$$

The proof of Proposition 2.10 is technical and it is presented in Appendix.

3. THE MODEL AND MAIN RESULTS

Let G be a 3-player symmetric game and every player has two pure actions: ”invest” (I) or ”not invest” (NI). By P_I (P_{NI}) we denote the payoff for an investing (not investing) player. All payoffs depend on the total number of investing players. If $n \in \{0, 1, 2, 3\}$ is the total number of investing players, then

n	$P_I(n)$	$P_{NI}(n)$
0	$-$	r_0
1	p_1	r_1
2	p_2	r_2
3	p_3	$-$

The game G in the normal form is given by the matrix:

I	(p_2, r_2, p_2)	(p_3, p_3, p_3)
NI	(r_1, r_1, p_1)	(r_2, p_2, p_2)
	NI	I

when the third player invests, and by the matrix

I	(p_1, r_1, r_1)	(p_2, p_2, r_2)
NI	(r_0, r_0, r_0)	(r_1, p_1, r_1)
	NI	I

when the third player does not invest.

We shall assume that the functions $P_I(\cdot)$, $P_{NI}(\cdot)$ are increasing:

$$0 < r_0 < r_1 < r_2 \text{ and } \quad 0 < p_1 < p_2 < p_3. \tag{13}$$

We assume that

$$p_1 < r_0. \tag{14}$$

By (14), the outcome (NI, NI, NI) is a Nash equilibrium. We assume that the more players invest, the greater the sum of all players payoffs is, i.e.

$$3r_0 < p_1 + 2r_1 < 2p_2 + r_2 < 3p_3. \tag{15}$$

By (15), the vector payoff (p_3, p_3, p_3) is Pareto optimal. In fact, the condition (15) means even more – the vector payoff (p_3, p_3, p_3) maximize the sum of payoffs. To obtain a strong equilibrium in the repeated game we assume that:

$$p_1 + r_1 < 2p_3. \tag{16}$$

We additionally assume that:

$$p_2 < r_2. \tag{17}$$

Observe that from the opposite inequality $r_2 \leq p_2$ implies that (p_3, p_3, p_3) is a Nash equilibrium payoff, what we wanted to avoid.

We introduce the following notations

$$\begin{aligned}
&A = (r_0, r_0, r_0),\\
&B = (p_3, p_3, p_3),\\
&C_1^1 := (p_1, r_1, r_1),\\
&C_2^1 := (r_1, p_1, r_1),\\
&C_3^1 := (r_1, r_1, p_1),\\
&C_1^2 := (r_2, p_2, p_2),\\
&C_2^2 := (p_2, r_2, p_2),\\
&C_3^2 := (p_2, p_2, r_2).
\end{aligned}$$

If i players invest $(i \in \{1, 2\})$ then C_j^i denotes the vector payoff in the game G. If $i = 1$ then j shows which one invests, while if $i = 2$ then j tells which player does not invest.

The strategy profile in the iterated game is given by a map $s: S \to \{I, NI\}^3$, where S is the convex hull of vector payoffs set, i.e.

$$S = co\{A, B, C_1^1, C_2^1, C_3^1, C_1^2, C_2^2, C_3^2\}.$$

The strategy profile s determines a dynamical process $\beta_n: S \to S$

$$\beta_n(x) = \frac{nx+\varphi(x)}{n+1}, \qquad \text{for } x \in S, \quad n \in \mathbb{N}, \tag{18}$$

where $\varphi: S \to S$ is given by the formula $\varphi = G \circ s$. Observe that a pair (s, x_1), where s is a strategy profile and $x_1 \in S$, uniquely determines a sequence $(\bar{x}_n)_{n=1}^{\infty}$ by:

$$\bar{x}_1 = x_1, \quad \bar{x}_{n+1} = \beta_n(\bar{x}_n).$$

We denote the obtained sequence by $\bar{x}_n(s, x_1)$. A similar construction of a sequence was considered in section 2. The strategy profile s and the initial point $x_1 \in S$ uniquely determine a play path. The action profile in the next stage $s(\bar{x}_n)$ depends on the average vector payoff $\bar{x}_n$. The element x_{n+1} is the vector payoff in $n + 1$ stage. We do not assume that the players observe the full history of the game. Instead, they observe aggregated history – the arithmetic mean of vector payoffs.

Motivated by the Smale construction in Smale (1980) we define an ε-good strategy for the i-th player $s_i^{\varepsilon}: S \to \{I, NI\}$ by

$$s_i^{\varepsilon}(x) = \begin{cases} I & \text{if } x \in V_i, \\ NI & \text{if } x \in S \setminus V_i, \end{cases} \tag{19}$$

where

$$\begin{aligned} V_i &= \Omega_i^{\varepsilon} \setminus W_i, \\ \Omega_i^{\varepsilon} &= \{x \in S: \ x_i > x_j - \varepsilon \text{ and } x_i > x_k - \varepsilon\}, \\ W_i &= \{x \in S: \ x_i < r_0 \text{ or } x_j + x_k > 2p_3\}, \end{aligned}$$

where i, j, k are pairwise different elements of the set of players $\{1, 2, 3\}$. The player invests if his average payoff is greater than the every other players' average payoff minus ε. The player stops investing if his playing I has been exploited by his opponents, that is either the average payoff of the player is lower than the payoff guaranteed by Nash equilibrium ($x_i < r_0$) or the sum of the other players' average payoffs is greater then the sum of their payoffs corresponding to the Pareto optimal profile (I, I, I) ($x_j + x_k > 2p_3$).

First we consider the case when all players choose good strategies. Then the average payoffs vector tends to the point B corresponding to the Pareto optimal profile (I, I, I).

Theorem 3.1 Suppose that $s_i^{\varepsilon}: S \to \{I, NI\}$ are the ε-good strategies for $i = 1,2,3$. Then

$$\lim_{T\to\infty} \bar{x}_T = B,$$

where $\bar{x}_T = \bar{x}_T((s_1^\varepsilon, s_2^\varepsilon, s_3^\varepsilon), x_1)$ and x_1 is an arbitrary element of S.

Now, we consider the case when two players play good strategies and the third one deviates and chooses an arbitrary strategy. The deviating player does not improve their payoff more then $\frac{2}{3}\varepsilon$, where the positive constant ε can be chosen arbitrarily small by the two non-deviating players.

Theorem 3.2 Suppose that the first and the second player choose the ε-good strategies s_1^ε, s_2^ε and the third player plays an arbitrary strategy $s_3: S \to \{I, NI\}$. Then

$$\limsup_{T\to\infty} \bar{x}_T^3 \leq p_3 + \varepsilon \frac{2}{3}, (20) \tag{20}$$

where $\bar{x}_T = \bar{x}_T((s_1^\varepsilon, s_2^\varepsilon, s_3), x_1)$ and x_1 is an arbitrary element of S.

At the end of the section we show an example of the third player strategy, such that the upper limit of his average payoffs is strictly grater than p_3.

Now, we consider the case when two players deviate.

Theorem 3.3 Suppose s_1^ε is the ε-good strategy for the first player and s_2, s_3 are arbitrary strategies. Then

$$\liminf_{T\to\infty} \bar{x}_T^1 \geq r_0, \tag{21}$$

$$\limsup_{T\to\infty} (\bar{x}_T^2 + \bar{x}_T^3) \leq 2p_3, \tag{22}$$

$$\lim_{T\to\infty} \text{dist}\,(\bar{x}_T, V_1) = 0, \tag{23}$$

where $\bar{x}_T = \bar{x}_T((s_1^\varepsilon, s_2, s_3), x_1)$ and x_1 is an arbitrary element of S.

Suppose that the payoff in the repeated game is defined as the Banach limit of average payoffs. The inequality (22) provides that if two players deviate then at least one of them will not improve his payoff. Conclusions (21) and (23) mean that the good strategy is safe, i.e. the non-deviating player's payoff is not smaller than the Nash equilibrium payoff in the stage game and, moreover, the deviating player's payoff is not greater than the non-deviating player's payoff plus ε (comp. Proposition 2.8).

By Theorems 3.1 – 3.3, we obtain

Corollary 3.4 The strategy profile $s^\varepsilon = (s_1^\varepsilon, s_2^\varepsilon, s_3^\varepsilon)$ satisfies (1-5). If we define the payoff in the repeated game as a Banach limit of average payoffs, i.e $\text{Lim}\bar{x}_T$ then the strategy profile s^ε is a safe and strong ε Nash equilibrium.

Below we provide some elementary properties of sets V_i that are used in the definition of good strategies. We assume that i, j, k are pairwise different elements of the set of players $\{1, 2, 3\}$. We shall use the following notations

$$\begin{aligned} V^3 &= \bigcap_{i=1}^{3} V_i \\ V_i^2 &= (S\backslash V_i) \cap V_j \cap V_k \\ V_i^1 &= V_i \cap (S\backslash V_j) \cap (S\backslash V_k) \end{aligned}.$$

If each player plays good strategy then

$$\varphi(x) = \begin{cases} B & \text{if } x \in V^3 \\ C_i^1 & \text{if } x \in V_i^1 \text{ for } i \in \{1,2,3\}. \\ C_i^2 & \text{if } x \in V_i^2 \text{ for } i \in \{1,2,3\} \end{cases} \tag{24}$$

Proposition 3.5 Suppose that player i plays ε_i-good strategy for $i = 1, 2, 3$. Then

$$\Omega_i \cap W_i = \emptyset, \tag{25}$$

$$V_i^1 \subset \Omega_i, \tag{26}$$

where

$$\Omega_i = \{x \in S: x_i = \max\{x_1, x_2, x_3\}\}.$$

If we assume that $\varepsilon_i = \varepsilon_j$ $(=: \varepsilon)$ then

$$V_i \cap (S\backslash V_j) \subset \Omega_i \cup \Phi_j. \tag{27}$$

If we assume that $\varepsilon_i = \varepsilon_j = \varepsilon_k$ $(=: \varepsilon)$ then for every $i \in \{1, 2, 3\}$ we have

$$V_i^2 \subset \Phi_i, \tag{28}$$

where

$$\Phi_i = \{x \in S: x_i = \min\{x_1, x_2, x_3\}\}.$$

Proof: If $x_i < r_0$ and $x_i = \max\{x_1, x_2, x_3\}$ then $x_1 + x_2 + x_3 < 3r_0$. If $x_j + x_k > 2p_3$ and $x_i = \max\{x_1, x_2, x_3\}$ then $x_1 + x_2 + x_3 > 3p_3$. Since $3r_0 \leq x_1 + x_2 + x_3 \leq 3p_3$ for $x \in S$, we obtain (25).

As $(S\backslash V_j) \cap \Omega_j = \emptyset$ and $(S\backslash V_k) \cap \Omega_k = \emptyset$ we have $(S\backslash V_j) \cap (S\backslash V_k) \subset S\backslash(\Omega_j \cup \Omega_k) \subset \Omega_i$, and consequently we obtain (26).

To prove conclusion (27) we take $i = 1,\ j = 2$. As $(W_2 \backslash W_1) \cap \Phi_1 = \emptyset$ and $(\Omega_3 \backslash \Phi_1) \subset \Phi_2$ we obtain $W_2 \backslash W_1 \subset (\Omega_1 \cup \Omega_3) \backslash \Phi_1 \subset \Omega_1 \cup \Phi_2$. If $x \in \Omega_1^\varepsilon \backslash \Omega_2^\varepsilon$ then either

$$x_2 \le x_1 - \varepsilon$$

or

$$x_2 \le x_3 - \varepsilon \text{ and } x_1 > x_3 - \varepsilon \ \ (x \in \Omega_1^\varepsilon).$$

In both cases we obtain $x_2 < x_1$ and thus $x \in \Omega_1 \cup \Phi_2$. Since $V_1 \cap (S \backslash V_2) \subset (\Omega_1^\varepsilon \backslash \Omega_2^\varepsilon) \cup (W_2 \backslash W_1)$, we conclude that

$$V_1 \cap (S \backslash V_2) \subset \Omega_1 \cup \Phi_2.$$

If $x_i \le x_j - \varepsilon$ $(x \notin \Omega_i^\varepsilon)$ and $x_k > x_j - \varepsilon$ $(x \in \Omega_k^\varepsilon)$ then $x \in \Phi_i$. If $x_i \le x_k - \varepsilon$ $(x \notin \Omega_i^\varepsilon)$ and $x_j > x_k - \varepsilon$ $(x \in \Omega_j^\varepsilon)$ then $x \in \Phi_i$. Thus $(S \backslash \Omega_i^\varepsilon) \cap \Omega_j^\varepsilon \cap \Omega_k^\varepsilon \subset \Phi_i$.

If $x \in V_j$ then $x \notin W_j$ and hence $x_j \ge r_0$ and $x_i + x_k \le 2p_3$. If $x \in W_i \cap V_j \cap V_k$ then either $x_i < r_0,\ x_j \ge r_0,\ x_k \ge r_0$ or $x_j + x_k > 2p_3,\ x_i + x_k \le 2p_3,\ x_i + x_j \le 2p_3$. In both cases we deduce that $x \in \Phi_i$. So $V_i^2 \subset \Phi_i$. QED

First we prove Theorem 3.3.

Proof: The strategy s_1^* is the ε-good strategy, so if $\bar{x}_T^1 < r_0$ then $\bar{x}_T = (\bar{x}_T^1, \bar{x}_T^2, \bar{x}_T^3) \in W_1$ and $s_1^*(\bar{x}_T) = NI$. It means that the next vector payoff x_{T+1} belongs to the set $\{A, C_2^1, C_3^1, C_1^2\}$, so $x_{T+1}^1 \in \{r_0, r_1, r_2\}$, i.e. $x_{T+1}^1 \ge r_0$ (see (13)). By Corollary 2.3 we obtain that $\limsup_{T\to\infty} - \bar{x}_T^1 \le -r_0$, so $\liminf_{T\to\infty} \bar{x}_T^1 \ge r_0$.

Similarly, if $\bar{x}_T^2 + \bar{x}_T^3 > 2p_3$ then $s_1^*(\bar{x}_T) = NI$. Thus the sum $x_{T+1}^2 + x_{T+1}^3$ is one of the numbers: $2r_0, p_1 + r_1, 2p_2$. From the assumptions (13), (15) and (16), it follows that $x_{T+1}^2 + x_{T+1}^3 \le 2p_3$. By Corollary 2.3, we get

$$\limsup_{T\to\infty} \bar{x}_T^2 + \bar{x}_T^3 \le 2p_3.$$

If $x \in S \backslash V_1$ then $s_1^*(\bar{x}_T) = NI$. So, $\varphi(x) = G((s_1^*, s_2, s_3)(x)) \in \{A, C_2^1, C_3^1, C_1^2\} \subset V_1$. By Corollary 2.2, the set V_1 is a weak attractor for φ. QED

Let $\pi_u : \mathbb{R}^3 \to u$ be the orthogonal projection onto the line $u = \{x \in \mathbb{R}^3 : x_1 = x_2 = x_3\}$ and $\pi_P : \mathbb{R}^3 \to P$ be the orthogonal projection onto the plane $P = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$. Obviously $\pi_u(x) = \left(\frac{x_1+x_2+x_3}{3}, \frac{x_1+x_2+x_3}{3}, \frac{x_1+x_2+x_3}{3}\right)$ and $\pi_P(x) = x - \pi_u(x)$. In the remainder of the section we denote the projection of a point (a set) A onto the plane P by $\widetilde{A}$, i.e. $\widetilde{A} = \pi_P(A)$. The projection of the set S onto the plane P:

$$\tilde{S} := \pi_P(S)$$

is the convex hull of the hexagon with successive vertexes $\tilde{C}_1^1,\ \tilde{C}_2^2,\ \tilde{C}_3^1,\ \tilde{C}_1^2,\ \tilde{C}_2^1,\ \tilde{C}_3^2$.

Set

$$v_1 = \frac{1}{\sqrt{2}}(0,-1,1), \qquad v_4 = -v_1,$$
$$v_2 = \frac{1}{\sqrt{2}}(-1,0,1), \qquad v_5 = -v_2,$$
$$v_3 = \frac{1}{\sqrt{2}}(-1,1,0), \qquad v_6 = -v_3,$$

and

$$\Delta_c(K) = \bigcap_{i\in K} \{y \in \tilde{S}: \langle v_i, y\rangle < c\},$$

where $K \subset \{1, \dots, 6\}$ and $c > 0$. One can easy check that

$$x \in \Omega_1^\varepsilon \Leftrightarrow \pi_P(x) \in \Delta_c(\{2,3\}),$$
$$x \in \Omega_2^\varepsilon \Leftrightarrow \pi_P(x) \in \Delta_c(\{1,6\}),$$
$$x \in \Omega_3^\varepsilon \Leftrightarrow \pi_P(x) \in \Delta_c(\{4,5\}),$$

where $c = \frac{\varepsilon}{\sqrt{2}}$. Setting $\Omega^\varepsilon = \bigcap_{i=1}^3 \Omega_i^\varepsilon$ and $\Delta_c = \Delta_c(\{1, \dots, 6\})$ we obtain

$$x \in \Omega^\varepsilon \Leftrightarrow \pi_P(x) \in \Delta_c. \tag{29}$$

Now, we are able to prove Theorem 3.1.

Proof: Fix $x_1 \in S$. It is sufficient to show that in the sequence $\bar{x}_T = \bar{x}_T(s^*, x_1)$ there exists an element $\bar{x}_N$ belonging to V^3, where $s^* = (s_1^\varepsilon, s_2^\varepsilon, s_3^\varepsilon)$. Indeed, if $\bar{x}_N \in V^3$ then $\bar{x}_{N+k} = \frac{N}{N+k}\bar{x}_N + \frac{k}{N+k}B$, so $\lim_{T\to\infty}\bar{x}_T = B$.

First we show that almost all elements of the sequence $\bar{x}_T$ belong to $\Omega^\eta = \bigcap_{i=1}^3 \Omega_i^\eta$, for every $\eta > 0$.

The map φ given by (24) is determined by the strategy profile s^*, i.e. $\varphi = G \circ s^*$. Consider $\tilde{\varphi}: \tilde{S} \to \tilde{S}$ and $V: P \to \mathbb{R}$ given by

$$\tilde{\varphi}(x) = \{\pi_P(\varphi(y)):\ \pi_P(y) = x\},$$
$$V(x) = \max\{\langle v_i, x\rangle: i = 1, \dots, 6\}.$$

We verify that V is a Lapunov type function for $\tilde{\varphi}$ with the constant c, for an arbitrary $c > 0$. Let us fix $x \in \tilde{S}$ such that $\langle v_1, x\rangle > 0$. If $y \in S$ and $\pi_P(y) = x$ then $\langle v_1, y\rangle = \langle v_1, x\rangle$.Thus $y_3 - y_2 > 0$ and therefore $y \notin \Omega_2 \cup \Phi_3$. By (26), (28), we have $y \notin V_2^1 \cup V_3^2$. Since $\varphi(y) \in \{C_1^1, C_3^1, C_1^2, C_2^2, B\}$, we obtain $\langle v_1, \omega\rangle \le 0$ for $\omega \in \tilde{\varphi}(x))$. We use similar arguments to show that if $\langle v_i, x\rangle > 0$ and $\omega \in \tilde{\varphi}(x)$ then $\langle v_i, \omega\rangle \le 0$, for $i = 2, \dots, 6$.

Fix $\eta < \min\{\varepsilon, p_3 - \frac{2p_2+r_2}{3}, \frac{p_1+2r_1}{3} - r_0\}$. By Proposition 2.10 and (29), there exists N such that $\bar{x}_n \in \Omega^\eta$ for $n > N$. We claim that there exists $M > N$ such that $\bar{x}_M \in V^3$. Suppose to the contrary that $\bar{x}_M \notin V^3$ for every $M > N$. Then $\varphi(\bar{x}_M) \in \{C_1^1, C_2^1, C_3^1, C_1^2, C_2^2, C_3^2\}$ for $M > N$. By Proposition 2.7, we obtain that $z_M \in (\frac{p_1+2r_1-\eta}{3}, \frac{2p_2+r_2+\eta}{3})$ for M sufficiently large, where the point (z_M, z_M, z_M) is the projection of $\bar{x}_M$ onto u.

But, if $x \in \Omega^\eta \backslash V^3$ then $\frac{x_1+x_2+x_3}{3} \notin (\frac{p_1+2r_1-\eta}{3}, \frac{2p_2+r_2+\eta}{3})$. Indeed, if $x_i + x_j > 2p_3$ and $x \in \Omega^\eta$ then $\frac{x_1+x_2+x_3}{3} > p_3 - \frac{\eta}{3}$. If $x_i < r_0$ and $x \in \Omega^\eta$ then $\frac{x_1+x_2+x_3}{3} < r_0 + \frac{2}{3}\eta$.

QED

Now, we are in a position to prove Theorem 3.2.

Proof: Let $x_1 \in S$ and $\eta > 0$. Our aim is to prove that almost all elements of the sequence $\bar{x}_T = \bar{x}_T((s_1^\varepsilon, s_2^\varepsilon, s_3),\ x_1)$ belongs to $\Omega^{\varepsilon+\eta} := \Omega_1^{\varepsilon+\eta} \cap \Omega_2^{\varepsilon+\eta}$. We have

$$x \in \Omega^{\varepsilon+\eta} \Leftrightarrow \pi_P(x) \in \Delta_c(\{1,2,3,6\}), \tag{30}$$

where $c := \frac{\varepsilon+\eta}{\sqrt{2}}$. We show that the function $V^*: P \to \mathbb{R}$ given by

$$V^*(x) = \max\{\langle v_i, x\rangle : i = 1,2,3,6\}$$

is the Lapunov type function for $\tilde{\varphi}^*: \tilde{S} \to \tilde{S}$ with the constant c, where

$$\tilde{\varphi}^* = \{\pi_P(z);\ \ z \in \varphi^*(y), \qquad \pi_P(y) = x\}$$

and

$$\varphi^*(y) = \begin{cases} \{C_1^1, C_2^2\} & \text{if } y \in V_1 \cap (S\backslash V_2), \\ \{C_2^1, C_1^2\} & \text{if } y \in (S\backslash V_1) \cap V_2, \\ \{B, C_3^2\} & \text{if } y \in V_1 \cap V_2, \\ \{A, C_3^1\} & \text{if } y \in (S\backslash V_1) \cap (S\backslash V_2). \end{cases}$$

The map $\varphi: S \to S$ induced by the profile $(s_1^\varepsilon, s_2^\varepsilon, s_3)$ is a selection of φ^*.

If $\langle v_6, x\rangle > 0$ $(x \in \tilde{S})$ and $\pi_P(y) = x$ $(y \in S)$ then $y_1 > y_2$ and thus $y \notin \Omega_2 \cup \Phi_1$. By (27), we have $V_2 \cap (S\backslash V_1) \subset \Omega_2 \cup \Phi_1$. Thus $\varphi^*(y) \cap \{C_2^1, C_1^2\} = \emptyset$. So, we have $\langle v_6, \omega\rangle < 0$ for $\omega \in \tilde{\varphi}^*(x)$.

Using similar arguments we show that if $\langle v_3, x\rangle > 0$ then $\langle v_3, \omega\rangle \leq 0$ for $\omega \in \tilde{\varphi}^*(x)$.

Suppose that $\langle v_1, x\rangle \geq V^*(x) - \delta$ $(x \in \tilde{S})$ and $\pi(y) = x$ $(y \in S)$, where $\delta < \frac{\eta}{\sqrt{2}}$. Then $\langle v_1, x\rangle = \langle v_1, y\rangle > \frac{\varepsilon}{\sqrt{2}}$. If $z \in \Omega_2^\varepsilon$ then $\langle v_1, z\rangle \geq \frac{1}{\sqrt{2}}(z_3 - z_2) > \frac{\varepsilon}{\sqrt{2}}$. Thus, we have $y \notin \Omega_2^\varepsilon \supset V_2$ and therefore $\varphi^*(y) \subset \{C_1^1, C_2^2, C_3^1, A\}$. So, $\langle v_1, \omega\rangle \leq 0$ for $\omega \in \tilde{\varphi}^*|(x)$.

In the similar way we prove that if $\langle v_2, x\rangle \geq V^*(x) - \delta$ and $\omega \in \tilde{\varphi}^*(x)$ then $\langle v_2, \omega\rangle \leq 0$.

By Proposition 2.10, we obtain that almost all elements of the sequence $(\pi_P(\bar{x}_T))$ belongs to $\Delta_c(\{1,2,3,6\})$. By (30), we have that almost all elements of the sequence $(\bar{x}_T)$ belongs to $\Omega^{\varepsilon+\eta}$. If $x \in \Omega^{\varepsilon+\eta}$ then $x_1 > x_3 - (\varepsilon + \eta)$ and $x_2 > x_3 - (\varepsilon + \eta)$ and so $x_3 < p_3 + \frac{2}{3}(\varepsilon + \eta)$ $(x_1 + x_2 + x_3 \leq 3p_3$ for $x \in S)$.

QED

Remark. Reasoning as in the proofs of Theorem 3.2 and Theorem 3.3, we can conclude that good strategies are safe and strong Nash equilibria not only in the class of Smale's strategies, but also if ”loyal” players adopt good strategies, then ”disloyal” players can even play the random choice in each repetition. It does not change the properties (20), (21), (22) and (23).

Example 2 Let the stage game G be given by:

n	$P_I(n)$	$P_{NI}(n)$
0	−	20
1	10	28
2	18	36
3	26	−

This game satisfies conditions (13) – (17).

Let $s_i^*: S \to \{I, NI\}$ be the ε-good strategy for the i-th player, $i = 1,2$, and $0 < \varepsilon < \frac{1}{2}$. Let $Z = \text{conv}\{A, B, C_3^1, C_3^2\} = \{x \in S: x_1 = x_2\}$ and $D = (26 - \frac{\varepsilon}{2}, 26 - \frac{\varepsilon}{2}, 26 + \frac{\varepsilon}{2})$. We present the construction of the third player strategy $s_3^*: S \to \{I, NI\}$ such that

$$\lim_{T\to\infty} x_T(s^*, x_1) = D \text{ for every } x_1 \in Z.$$

We have $V_1 \cap Z = V_2 \cap Z$. We set

$$s_3^*(x) = \begin{cases} NI & \text{if } x \in V_1 \cap Z \cap \text{co}\{B, D, C_3^1\} \\ I & \text{elsewhere.} \end{cases}$$

The map φ induced by the strategy profile $s^* = (s_1^*, s_2^*, s_3^*)$ is given by

$$\varphi(x) = \begin{cases} B & \text{if } x \in V_1 \cap Z \backslash \mathrm{co}\{B, D, C_3^1\}, \\ C_3^2 & \text{if } x \in V_1 \cap Z \cap \mathrm{co}\{B, D, C_3^1\}, \\ C_3^1 & \text{if } x \in Z \backslash V_1. \end{cases}$$

The values of the map φ outside the set Z have no influence onto the trajectory $\bar{x}_T(s^*, x_1)$ if $x_1 \in Z$. The map $\varphi: Z \to Z$ satisfies the Blackwell condition for the triangle $\mathrm{co}\{C_3^1, C_3^2, D\}$ in the domain Z. The map $\varphi: Z \to Z$ satisfies the Blackwell condition for the sum of intervals $\overline{BD} \cup \overline{DC_3^1}$ in the domain Z. By Proposition 2.1, the sets $\mathrm{co}\{C_3^1, C_3^2, D\}$ and $\overline{BD} \cup \overline{DC_3^1}$ are week attractors. To conclude that the interval $\overline{BD}$ is a weak attractor we apply Proposition 2.5 taking $A = \overline{BD} \cup \overline{DC_3^1}$ and $B = \overline{BD}$. By Proposition 2.4, the intersection of weak attractors $\mathrm{co}\{C_3^1, C_3^2, D\}$ and $A = \overline{BD}$ is a weak attractor. The intersection equals to the set $\{D\}$.

4. CONCLUSIONS

This paper is concerned with the specific model of social dilemmas. Such models have a very special place in game theory as they describe real social problems of modern world: resources depletion, pollution and overpopulation. The main characteristic of such models is that each player gain more by not cooperating when opponents fix their choices and all individuals are better off if all cooperate. The lack of optimality of Nash equilibrium is the most interesting problem, because as we can observe in the real world, people are keen to cooperate with each other on the certain conditions. As we can find in Axelrod (1984), strategies that effectively encourage people to cooperate are: nice, forgiving, retaliatory and are found on simple rules.

The key idea in our approach is to apply Smale's idea for 3-payer extension of Prisoner's Dilemma. Our strategies are deterministic and satisfy conditions that are postulated in Axelrod (1984). What is more, ε-good strategies satisfy condition (5) which guarantee that using this strategy our payoff shall not be different than our opponents payoffs for more than ε. This constant ε is totally controlled by the player who choose it. This property is not received by any other author.

Our future aim is to extend the idea presented in Plaskacz (2018) onto the type of games considered in the paper - three players repeated social dilemmas. The idea is as follows. We would like to analyze the repeated three players game by evolutionary games methods. To achieve this goal, we threat the repeated game as a new game in which a player action is a point in the β-core of the original game. Using methods presented in the paper each point from the β-core should determine ε-good strategy. The main difficulty is to obtain the payoff in the case when players choose different points in the β-core. The payoffs in the new game are determined by the payoff in the repeated game.

REFERENCES

Aumann R. J., (1959), Acceptable Points in General Cooperative n-Person Games, in: Tucker A. W., Luce R. D., (eds.), *Contributions to the Theory of Games, IV*, Princeton University Press, Princeton, 287–324.

Aumann R. J., (1961), The Core of a Cooperative Game Without Side Payments, *Transactions of the American Mathematical Society*, 98, 539–552.

Aumann R. J., (1967), A Survey of Cooperative Games Without Side Payments, in: Shubik M., (ed.), *Essays in Mathematical Economics in Honour of Oscar Morgenstern*, 3–27.

Axelrod R., (1984), *The Evolution of Cooperation*, Basic Books, Inc., Publishers, New York.

Behrstock K., Benaim M., Hirsch M. W., (2015), Smale Strategies for Network Prisoner's Dilemma Games, *Journal of Dynamics and Games*, 2, 141–155.

Blackwell D., (1956), An Analog of The Minimax Theorem for Vector Payoffs, *Pacific Journal of Mathematics*, 6, 1–8.

Conway J. B., (1985), *A Course in Functional Analysis*, Springer Verlag, New York.

Dawes R. M., (1980), Social Dilemmas, *Annual Review of Psychology*, 31, 169–193.

Konishi H., Le Breton M., Weber S., (1997), Equivalence of Strong And Coalition-Proof Nash Equilibria in Games Without Spillovers, *Economic Theory*, 9, 97–113.

Nessah R., Tian G., (2014), On the Existence of Strong Nash Equilibrium, *Journal of Mathematical Analysis and Applications*, 414, 871–885.

Plaskacz S., Zwierzchowska J., (2018), Dynamical Systems Associated to the β-core in the Repeated Prisoner's Dilemma, *Dynamic Games and Applications*, https://doi.org/10.1007/s13235––018-0262-x.

Shani B., Solan E., (2014), Strong Approachability, *Journal of Dynamics and Games*, 1, 507–535.

Smale S., (1980), The Prisoner's Dilemma and Dynamical Systems Associated to Non-Cooperative Games, *Econometrica*, 48, 1617–1634.

Sorin S., (1992), Repeated Games with Complete Information, in: Aumann R. J., Hart S., (eds.), *Handbook of Game Theory*, 72–107.

APPENDIX

In this Appendix we shall present the proof of Proposition 2.10. We start with the necessary theorem.

Theorem 5.1 If S is a bounded convex subset of H and $V: H \to \mathbb{R}$ given by (9) is the Lapunov type function for the mulivalued map $\varphi: S \to S$ with the constant $c > 0$, then

$$\exists\gamma > 0, \quad \exists\alpha_0 > 0, \quad \forall\alpha \in [0, \alpha_0], \quad \forall x \in S\backslash\Delta_c, \quad \forall\omega \in \varphi(x)$$
$$V(\alpha\omega + (1-\alpha)x) \leq V(x) - \alpha\gamma.$$

Proof: By (10) we choose $\delta \in (0, c)$. Let $M = \sup\{|x|: x \in S\}$. For $x \in S\backslash\Delta_c$ we define a set of indexes $I(x)$ by

$$I(x) = \{j \in \{1, \dots, k\}: V_j(x) \geq V(x) - \delta\}$$

and a subset O_i of S, related to the fixed index i:

$$O_i := \{x \in S\backslash\Delta_c : V_i(x) \geq V(x) - \delta\}. \quad (31)$$

If $x \in O_i$ then $V_i(x) > 0$ and $\langle p_i, \omega\rangle \leq 0$ for all $\omega \in \varphi(x)$. Obviously, $i \in I(x)$ is equivalent to $x \in O_i$ for $x \in S\backslash\Delta_c$.

We fix positive constants: r, γ and α_0 such that

$$r < \frac{\delta}{2}, \qquad \gamma < c - \delta, \qquad \alpha_0 < min\left\{\frac{c-\delta-\gamma}{c-\delta+M}, \frac{r}{2M}, 1\right\}$$

and take an arbitrary $x \in S\backslash\Delta_c$. The following condition holds true

$$\forall y \in B(x,r) = \{y \in S\backslash\Delta_c : ||x-y|| < r\} \quad \exists i \in I(x) \quad V(y) = V_i(y). \quad (32)$$

Indeed, if $j \notin I(x)$, then $V_j(x) < V(x) - \delta$. Since V_j and V are lipschitz continuous with the constant L=1, we get $V_j(y) < V(y)$. Therefore, there exists $i \in I(x)$ such that $V(y) = V_i(y)$.

If $i \in I(x)$ then $V_i(x) \geq V(x) - \delta \geq c - \delta$ and $\langle p_i, \omega\rangle \leq 0$ for $\omega \in \varphi(x)$. Let $\alpha \in [0, \alpha_0]$ then $x_\alpha := \alpha\omega + (1-\alpha)x \in B(x,r)$ and $V_i(x_{\alpha_0}) \leq V_i(x_\alpha)$. Moreover,

$$V_i(x_\alpha) \geq V_i(x_{\alpha_0}) \geq -\alpha_0||\omega|| + (1-\alpha_0)(c-\delta) \geq c - \delta - \alpha_0(c-\delta+M) \geq \gamma$$

and

$$V_i(x_\alpha) \leq (1-\alpha)V_i(x) \leq V_i(x) - \alpha\gamma.$$

Thus we have obtained that

$$\forall i \in I(x), \quad \forall \alpha \in [0, \alpha_0], \quad \forall \omega \in \varphi(x), \quad V_i(x_\alpha) \leq V_i(x) - \alpha\gamma. \quad (33)$$

The function V has the following property: if $V(a) = V_i(a)$ and $V(b) = V_i(b)$ then $V(\lambda a + (1-\lambda)b) = V_i(\lambda a + (1-\lambda)b)$ for $\lambda \in [0,1]$ so the set

$$\{\alpha \in [0, \alpha_0] : V_i(\alpha\omega + (1-\alpha)x) = V(\alpha\omega + (1-\alpha)x)\}$$

is a closed segment. By (32) there exists $s \leq k$ and a partition $0 = \beta_0 < \beta_1 <\ldots< \beta_s = \alpha_0$ such that

$$\forall j \in \{0, \ldots, s-1\}, \exists i = i(j) \in I(x), \forall \alpha \in [\beta_j, \beta_{j+1}], \quad V(x_\alpha) = V_i(x_\alpha). \quad (34)$$

Let $\alpha \in [\beta_0, \beta_1]$. In view of (34) there exists $i = i(0) \in I(x)$ such that $V(x_\alpha) = V_i(x_\alpha)$ and by (33):

$$V(x_\alpha) = V_i(x_\alpha) \leq V_i(x) - \alpha\gamma = V(x) - \alpha\gamma.$$

Suppose that

$$V(x_\alpha) \le V(x) - \alpha\gamma, \quad \forall k = 1, \dots, j-1 \quad \forall \alpha \in [\beta_k, \beta_{k+1}]$$

and take $\alpha \in [\beta_j, \beta_{j+1}]$. By (34) there exists $i = i(j)$ such that $V(x_\alpha) = V_i(x_\alpha)$. Since $\alpha \in [\beta_j, \beta_{j+1}]$, there exists $\xi \in [0,1]$ such that $x_\alpha = \xi\omega + (1-\xi)x_{\beta_j}$. Therefore,

$$x_\alpha = \xi\omega + (1-\xi)(\beta_j\omega + (1-\beta_j)x) = (\xi + (1-\xi)\beta_j)\omega + (1-\xi)(1-\beta_j)x,$$

so

$$\alpha = \xi + (1-\xi)\beta_j \le \xi + \beta_j.$$

It is obvious that

$$V(x_\alpha) = V_i(x_\alpha) = V_i(\xi\omega + (1-\xi)x_{\beta_j}) = \xi < \omega, p_i > +(1-\xi) < x_{\beta_j}, p_i >$$
$$\le (1-\xi)V_i(x_{\beta_j}) = V_i(x_{\beta_j}) - \xi V_i(x_{\beta_j}) \le V_i(x_{\beta_j}) - \xi\gamma.$$

Then we get

$$V_i(x_{\beta_j}) - \xi\gamma \le V(x) - \beta_j\gamma - \xi\gamma = V(x) - \gamma(\beta_j + \xi) \le V(x) - \alpha\gamma,$$

hence,

$$V(\alpha\omega + (1-\alpha)x) \le V(x) - \alpha\gamma. \qquad \text{QED}$$

The proof of Proposition 2.10.

Proof: Fix $(\bar{x}_n)_{n=1}^{\infty}$ satisfying (12). First, we prove that

$$\forall M, \exists N \ge M, \quad \bar{x}_N \in \Delta_c. \tag{35}$$

Suppose, contrary to our claim, that $\bar{x}_n \notin \Delta_c$ for $n \ge m$. We choose $k \ge m$ such that $\frac{1}{k} < \alpha_0$, where α_0 and γ are given by Theorem 5.1. Thus

$$V(\bar{x}_{k+l+1}) = V\left(\frac{1}{k+l+1}x_{k+l+1} + \frac{k+l}{k+l+1}\bar{x}_{k+l}\right) \le V(\bar{x}_{k+l}) - \gamma\frac{1}{k+l+1} \le$$
$$\dots \le V(\bar{x}_k) - \gamma\left(\frac{1}{k+l+1} + \dots + \frac{1}{k+1}\right) \xrightarrow[l\to\infty]{} -\infty$$

which contradicts to the assumption that $V(\bar{x}_n) \ge c$ for $n \ge m$.

Fix $c_1 > c$. By Proposition 2.6, we choose M such that $|\bar{x}_{l+1} - \bar{x}_l| < c_1 - c$ for $l \geq M$. By (35), there exists $N \geq M$ such that $\bar{x}_N \in \Delta_c$. If $\bar{x}_{N+l} \in \Delta_c$ then $V(\bar{x}_{N+l+1}) \leq V(\bar{x}_{N+l}) + |\bar{x}_{N+l+1} - \bar{x}_{N+l}| < c_1$. If $\bar{x}_{N+l} \in \Delta_{c_1} \backslash \Delta_c$ then $V(\bar{x}_{N+l+1}) \leq V(\bar{x}_{N+l}) < c_1$. QED

SILNE I BEZPIECZNE RÓWNOWAGI NASHA W PEWNYCH GRACH POWTARZANYCH 3 GRACZY

Streszczenie

W pracy analizujemy grę nieskończenie powtarzaną 3-graczy będącą rozszerzeniem gry typu Dylemat Więźnia. Rozważamy grę 3-graczy w postaci normalnej z pełną informacją, w której każdy gracz ma dwa działania. Zakładamy, że gra jest symetryczna i powtarzana nieskończenie wiele razy. Strategią gracza w grze powtarzanej jest funkcja zdefinowana na uwypukleniu zbioru wypłat. Naszym celem jest skonstruowanie mocnej równowagi Nasha w grze powtarzanej, to znaczy profilu strategii, który jest odporny na odstępstwa od strategii równowagi przez koalicję graczy. Skonstruowane strategie równowagi są bezpieczne, to znaczy wypłata gracza, który nie odstępuje od strategii równowagi jest niemniejsza od wypłaty odpowiadającej równowadze w grze etapowej, oraz wypłata gracza odstępujacego od równowagi może być większa od wypłaty gracza nieodstępującego od strategii równowagi, ale nie więcej niż o pewną stałą dodatnią, która może być wybrana dowolnie mała przez gracza nieodstępującego od równowagi. Nasza konstrukcja jest inspirowana koncepcją dobrych strategii Smale'a opisaną w jego pracy z 1980 roku, gdzie rozważany był powtarzany Dylemat Więźnia. W dowodach wykorzystujemy wyniki o zbliżaniu oraz silnym zbliżaniu.

Słowa kluczowe: gra powtarzana, silna równowaga Nasha, metoda Blackwell'a w problemie zbliżania, metoda funkcji Lapunowa

STRONG AND SAFE NASH EQUILIBRIUM IN SOME REPEATED 3-PLAYER GAMES

Abstract

The paper examines an infinitely repeated 3-player extension of the Prisoner's Dilemma game. We consider a 3-player game in the normal form with incomplete information, in which each player has two actions. We assume that the game is symmetric and repeated infinitely many times. At each stage, players make their choices knowing only the average payoffs from previous stages of all the players. A strategy of a player in the repeated game is a function defined on the convex hull of the set of payoffs. Our aim is to *construct a strong Nash equilibrium in the repeated game, i.e. a strategy profile being resistant to deviations by coalitions. Constructed equilibrium strategies are safe, i.e. the non-deviating player payoff is not smaller than the equilibrium payoff in the stage game, and deviating players' payoffs do not exceed the non-deviating player payoff more than by a positive constant which can be arbitrary small and chosen by the non-deviating player. Our construction is inspired by Smale's good strategies described in Smale's paper (1980), where the repeated Prisoner's Dilemma was considered. In proofs we use arguments based on approachability and strong approachability type results.*

Keywords: repeated game, strong Nash equilibrium, Blackwell's approachability, Lapunov function method

DOI: 10.5604/01.3001.0014.0541

PRZEGLĄD STATYSTYCZNY
TOM LXV – ZESZYT 3 – 2018

Agnieszka LACH[1]
Łukasz SMAGA[2]

Comparison of the goodness-of-fit tests for truncated distributions

1. INTRODUCTION

The role of statistical testing has been the subject of discussion for years. An overview of this topic was given decades ago for instance in Cox et al. (1977). Nevertheless, the matter is still important, which confirms for example the paper on statistical testing in finance written by Kim, Ji (2015). As was mentioned in that paper, many statistical tests are used in practice with little consideration of their key characteristics as size and power. These characteristics should be intensively studied at least in simulations as, for example, in Pavia (2015), Górecki, Smaga (2015) or Orzeszko (2014). In this paper, we investigate the finite sample behavior of some goodness-of-fit testing procedures for truncated distributions known in the literature. Such behavior was not considered in the original paper introducing these tests. The paper is an extension of the results obtained in the bachelor thesis by Lach (2017).

The shape of a distribution in the tails is very important in many areas of science. Chernobai et al. (2015) adapted the standard goodness-of-fit tests for left-truncated distributions. The modifications of standard procedures help to take the decision, whether the tail belongs to a specified distribution or not. The tests were implemented in the R package truncgof (R Core Team, 2017; Wolter, 2012). The detailed description of their seven tests is given in section 2. Five of them are the commonly used standard tests with the modified null hypothesis cumulative distribution function. Following the original notation, these tests will be referred to as the AD^* (supremum Anderson-Darling), AD^{2*} (quadratic Anderson-Darling), KS^* (Kolmogorov-Smirnov), V^* (Kuiper) and W^{2*} (Cramér-von Mises) tests, respectively, in the remainder of the article. The other two tests are specially designed for the upper tails. They use the modified null hypothesis cumulative distribution function and the new weighing function. These are modified Anderson-Darling tests, which will be referred to as AD_{up}^* and AD_{up}^{2*} tests, respectively.

---

[1] Poznan University of Economics and Business, Faculty of Informatics and Electronic Economy, Operations Research Department, 10 Al. Niepodległości St., 61–875 Poznań, Poland, corresponding author – e-mail: agnlach1@gmail.com.

[2] Adam Mickiewicz University, Faculty of Mathematics and Computer Science, Department of Probability and Mathematical Statistics, 87 Umultowska St., 61–614 Poznań, Poland.

The tests by Chernobai et al. (2015) are often used in the literature, especially in the field of the operational risk calculation. Here, the choice of appropriate severity distribution is of crucial importance. In the process of calculating the aforementioned risk, Fischer, Jakob (2016) used a compound severity distribution, which involves dividing it into the body and the tail by a threshold. The Authors conclude that positive tempered α-stable distribution better fits empirical data in the tail than lognormal, Weibull, gamma and generalized gamma distributions. To assess goodness-of-fit of the distributions in the tail they used among others the AD_{up}^{2*} test for truncated distributions. Chernobai et al. (2006) considered the following severity distributions: exponential, lognormal, Weibull, Burr, generalized Pareto (GPD) and log α-stable. The null hypothesis that the cumulative distribution function belongs to truncated versions of the families of these distributions was verified by using the procedure described in Chernobai et al. (2015). The tests for truncated distributions were also used by Chernobai et al. (2010), who analyzed the effects of model misspecifications on Value-at-Risk and Conditional Value-at-Risk figures.

Examples of applications of the tests by Chernobai et al. (2015) can also be found in hydrology and social sciences. To estimate flood peaks, Brunner et al. (2017) used among others modified Anderson-Darling test for the upper tail to verify fitting of GPD and generalized extreme value distribution (GEV) to observed flood hydrographs. As was stated in the study, the test confirmed that the GPD fits well to the peak discharges and the GEV distribution fits well to the flood volumes. In the field of social sciences, Fagiolo et al. (2010) studied distributional properties of Italian household consumption expenditures. To study the tails of the distributions, they truncated distributions in several points and then they used standard truncated goodness-of-fit normality tests. Clementi et al. (2012) proposed a new model for income distribution: the κ-generalized distribution. As the fit in the right tail was of greater importance here, they decided to compare it with Singh-Maddala or Dagum type I distributions using upper tail goodness-of-fit tests.

In the majority of the studies listed above, the truncated tests were not the only ones, upon which the decisions were taken. However, it is clear that they had impact on the researchers' final decisions and that the range of possible applications of them is wide. Until now no studies concerning the size and the power of these tests for the left truncated distributions were published. The aim of this paper is to fill this gap.

The research of this paper is similar to that conducted by Pavia (2015). The main difference is that Pavia concentrated on complete distributions, while this paper refers to truncated ones. Pavia conducted the research for different sample sizes (10, 20, 50, 100, 200, 500). In this paper, the research is conducted for the sample of size 1000. Pavia verified the empirical sizes and the empirical powers of several goodness-of-fit tests available in the R packages, including five tests from the truncgof package (AD^*, AD^{2*}, KS^*, V^* and W^{2*}). As the Author was interested only in complete distributions, he omitted the tests from the truncgof package designed for the upper tail. When the truncation point is not set, the tests from this package can be used also in the case of complete samples. According to the analysis of the sizes of the tests, the main conclusion was that most of the tests from the truncgof package are giving unacceptable results, especially for bounded distributions. Only the AD^* test from this package achieved acceptable rejecting rates for all examined distributions except for the uniform one. In case of the exponential distribution, the V^* test also gave reasonable results. When analyzing the power of the tests, in most of the examples the tests implemented in the truncgof package showed superior power over the rest of the tests taken into comparison. The results for the bounded distribution were again unacceptable.

The study in this paper is based on the artificial data generated from the distributions that are used to describe the tails of asset returns. The shape of the tails has great importance in the assessment of the risk. The origins of the studies on the distribution of asset returns dates back to the year 1900. At that time Louis Bachelier noticed, that according to the Central Limit Theorem the distribution of the asset returns in long term should be Gaussian (Haas, Pigorsch, 2009). That implies that the tails of the distributions should be thin and tend to zero faster than exponentially (Feller, 1950). This conception was prevailing until 1963, when Mandelbrot (1963) noticed fat tails of distributions of the cotton prices logarithms. One of the first distributions proposed to replace the normal distribution was the t-distribution with power decaying tails (Haas, Pigorsch, 2009). However, recent studies show that most of the asset returns have semi-heavy tails (Echaust, 2014; Piasecki, Tomasik, 2013). The power-exponential distribution could be proposed here as alternative. Depending on parameters, its tails can change from thinner tails than those of normal distribution to fat ones. Another example might be Weibull distribution, whose tails vary from thin to fat. The distributions mentioned in this paragraph were chosen for the research due to their historical meaning or the possibilities they offer. For the details of the Central Limit Theorem and these distributions the reader is referred to Krzyśko (2000) and Magiera (2005).

The remainder of this paper is organized as follows. In section 2, the tests for truncated distributions introduced in Chernobai et al. (2015) are presented. Section 3 contains the results of the simulation studies. Finally, section 4 draws some conclusions.

2. TESTS FOR TRUNCATED DISTRIBUTION

This section contains description of seven goodness-of-fit tests for truncated distributions, which were introduced in Chernobai et al. (2015). Five of these tests are modifications of standard goodness-of-fit tests. The remaining testing procedures are specifically constructed for upper tails of distributions. Before the description of the tests, short information about upper tails in finance is given.

Upper tails in finance are defined as $\overline{F(x)} = P(X > x)$, where x is sufficiently high, which means that $x \to \infty$ (Haas, Pigorsch, 2009). In case of the asset returns, even 5% is enough high to be set as the truncation point (Haas, Pigorsch, 2009). However, banks and other financial institutions may wish to define the tails in terms of the quantiles of the distribution. When calculating risk measures, like VaR or CVaR, usually quantiles of level 0.95, 0.975 or 0.99 are taken into account, although higher quantiles also appear (Haas, Pigorsch, 2009). On the other hand, when choosing an investment strategy, investors might be interested in much lower quantiles of distributions.

Chernobai et al. (2015) adapted the Kolmogorov-Smirnov, Kuiper, Cramér-von Mises and Anderson-Darling tests, which are standard goodness-of-fit tests, for truncated distributions. Anderson, Darling (1952) enabled giving different weights to specific parts of a distribution function, multiplying classical Kolmogorov and Cramér-von Mises statistics by the weight function $\Psi(x)$ (where $\Psi(x) \geq 0$ for $x \in [0,1]$). Anderson and Darling considered two weight functions: $\Psi(x) = 1$ and $\Psi(x) = 1/[x(1-x)]$. While for the first function test statistics reduce to the standard Kolmogorov and Cramér-von Mises statistics, the second function gives greater importance to the tails of the distribution function.

Let us assume that we have a sample $\mathbf{X} = (X_1, \dots, X_{n_c})'$ of i.i.d. variables with an unknown distribution function F. To formulate a goodness-of-fit problem for truncated distributions, Chernobai et al. (2015) used the appropriate distribution function for the truncated sample. Let F_0 denote distribution function for the complete sample and let H be the truncation point. The modified distribution function for the truncated sample is then defined by the following formula:

$$F_0^*(x) = \begin{cases} \dfrac{F_0(x) - F_0(H)}{1 - F_0(H)} & , \text{for } x \geq H, \\ 0 & , \text{for } x < H. \end{cases} \qquad (1)$$

The complete sample of observations consists of n_c items. The ordered sample of observations $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n_c)}$ has the empirical distribution function (Krzyśko, 2004):

$$F_{n_c}(x; \mathbf{X}) = \frac{\#\{1 \leq j \leq n_c : X_j \leq x\}}{n_c}, \qquad x \in R, \mathbf{X} \in R^{n_c}. \qquad (2)$$

The difference between the values of the empirical distribution function for two neighboring points is equal to $1/n_c$. In case of left truncated distribution, the complete sample is the sum $n_c = m + n$, where m denotes the number of unknown observations below the truncation point and n is the number of observations equal to or greater than the truncation point. The empirical distribution function for the truncated sample is the same as for the complete sample, but the difference between the values of the empirical distribution function for two neighboring points is equal to $1/n$. The empirical distribution function for the observed part of the whole population is then (Chernobai et al., 2015):

$$F_n(x;\mathbf{X})(1-F_0(H))+F_0(H) = \begin{cases} F_0(H) & x < x_{(1)}, \\ \frac{j}{n}(1-F_0(H))+F_0(H) & x_{(j)} \le x < x_{(j+1)}, j=1,\dots,n-1, \\ 1 & x \ge x_{(n)}. \end{cases} \quad (3)$$

Thus, the null and alternative hypothesis can be formulated as follows:

$$\begin{aligned} H_0 &: F = F_0^*, \\ H_1 &: F \neq F_0^*. \end{aligned} \quad (4)$$

To test the null hypothesis against the alternative one, the null distributions of the test statistics (described below) are approximated by the Monte Carlo method. The detailed procedure for computing the corresponding p-values is as follows:

1. Compute the test statistic T_{obs} for the original data.
2. Generate a sample of n observations from the theoretical distribution function F_0^*. Each observation has to be greater than or equal to H.
3. Compute the test statistic T for the data generated in step 2.
4. Repeat steps 2 and 3 N times. Let $T_1, \dots, T_N$ denote the obtained values of the test statistic.
5. Compute the p-value according to the formula $(1/N)\sum_{i=1}^{N} I\,(T_i \ge T_{obs})$, where $I(S)$ denotes the indicator function of a set S.

The null hypothesis is rejected, when the p-value is less than or equal to the nominal significance level α. Otherwise, we do not have any evidence to reject the null hypothesis. The asymptotic distributions of the test statistics considered in this paper are not known, which is one of the reasons of using the above procedure.

Following Chernobai et al. (2015), the test statistics applied to verify the null hypothesis are divided into three groups: (1) the supremum class, (2) the quadratic class, (3) the test statistics specifically designed to test goodness-of-fit in the upper tail.

The first group is made up of three modified statistics: Kolmogorov-Smirnov (KS^*), Kuiper (V^*) and Anderson-Darling (AD^*) in supremum version. The Kolmogorov-Smirnov test is based on the statistic called Kolmogorov distance, which measures the distance between empirical distribution function and given distribution function. The modified version of the Kolmogorov-Smirnov statistic is as follows:

$$KS^* = \sqrt{n}\sup_x |F_n(x;\mathbf{X}) - F_0^*(x)|. \quad (5)$$

The Kuiper test statistic derives from the Kolmogorov distance. It is the sum of the greatest positive and negative difference between the empirical distribution function and the given distribution function. The modified version of the Kuiper test statistic is given by the following formula:

$$V^* = \sqrt{n}\left(\sup_x\{F_n(x;\mathbf{X}) - F_0^*(x)\} + \sup_x\{F_0^*(x) - F_n(x;\mathbf{X})\}\right). \tag{6}$$

The Anderson-Darling test statistic in the supremum version is also based on the distance between two distribution functions, but it put more emphasis on the tails:

$$AD^* = \sqrt{n}\sup_x \frac{|F_n(x;\mathbf{X}) - F_0^*(x)|}{\sqrt{F_0^*(x)(1 - F_0^*(x))}}. \tag{7}$$

The second group of statistics consists of two statistics: Cramér-von Mises (W^{2*}) and Anderson-Darling (AD^{2*}) in quadratic version. Both statistics measure the area between empirical distribution function and given distribution function, but they assign different weights to observations. Cramér-von Mises statistic has the weight function equal to one, and its customized version is of the form:

$$W^{2*} = n\int_H^\infty (F_n(x;\mathbf{X}) - F_0^*(x))^2 dF_0^*(x). \tag{8}$$

The Anderson-Darling statistic in quadratic version again puts more weight in the tails:

$$AD^{2*} = n\int_H^\infty \frac{(F_n(x;\mathbf{X}) - F_0^*(x))^2}{F_0^*(x)(1 - F_0^*(x))} dF_0^*(x). \tag{9}$$

New statistics proposed in Chernobai et al. (2015) are based on the Anderson-Darling statistics and give more importance to the upper tail of the distribution. The Authors introduced a new weight function, namely $\Psi(x) = 1/(1-x)$. After substituting this function, the Anderson-Darling statistics for the truncated samples in supremum (AD_{up}^*) and quadratic (AD_{up}^{2*}) version are respectively as follows:

$$AD_{up}^* = \sqrt{n}\sup_x \frac{|F_n(x;\mathbf{X}) - F_0^*(x)|}{1 - F_0^*(x)}, \tag{10}$$

$$AD_{up}^{2*} = n\int_H^\infty \frac{(F_n(x;\mathbf{X}) - F_0^*(x))^2}{(1 - F_0^*(x))^2} dF_0^*(x). \tag{11}$$

The computational formulas for the test statistics for truncated distributions (for quadratic versions of the statistics and for the new statistics) can be found in Chernobai et al. (2015).

3. SIMULATION STUDIES

This section contains the results of the simulation studies, conducted for seven modified goodness-of-fit statistics presented in section 2 and for the selected distributions described in section 1. The aim of the studies was to evaluate the size and the power of the goodness-of-fit tests for truncated distributions on the basis of artificial data. The simulation studies were conducted for different tail thickness and truncation points. This section is organized as follows: first part describes the methodology of the studies, next the results of the evaluation of the size and the power are presented, finally some details of implementation in R program are given.

3.1. Description of simulation experiments

To compute the empirical sizes of the analyzed tests, the following procedure was applied:

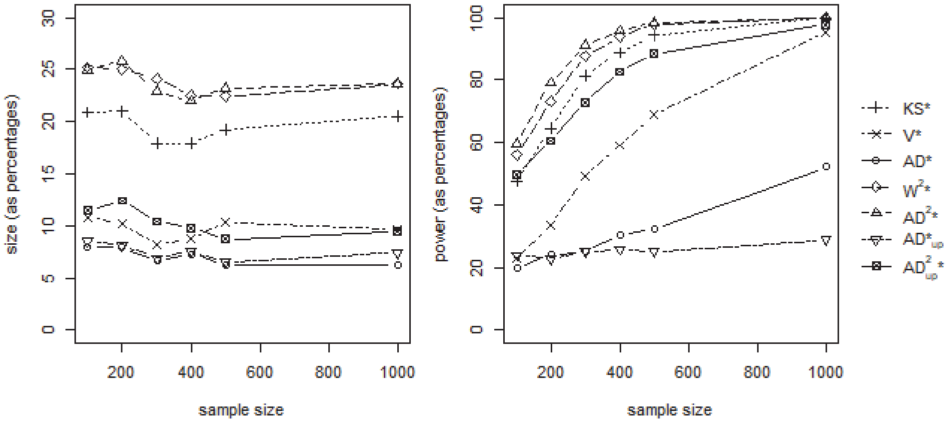
1. Generate n observations from the theoretical distribution that appears in the null hypothesis.
2. Apply all the analyzed tests to the data generated in point 1. Note the p-values of the tests.
3. Repeat the steps described in points 1 and 2 M times, were M is sufficiently large number.
4. Compute the empirical size of each test as the mean of a number of rejections of the null hypothesis.

To compute the empirical power of the tests, in point 1 of the above procedure, the data were generated from a different distribution than it was stated in the null hypothesis. The steps from 2 to 4 remained the same.

To determine the p-values, the testing procedures described in section 2 were carried out. The p-values were calculated on the basis of $N = 100$ Monte Carlo samples, which is the default value of N in the truncgof package. Within each simulation a sample of $n = 1000$ observations was generated. The number of simulation replicates was $M = 1000$. The studies were conducted for 12 distributions described in the next paragraph and for 5 truncation points $H = 2, 4, 6, 8, 10$. Altogether 60 experiments were conducted to evaluate both the empirical size and power. The results were verified on the significance level $\alpha = 5\%$.

The sample size $n = 1000$, was determined on the basis of the simulation studies conducted for the selected cases. Namely, figure 1 presents the empirical size and power of the analyzed tests for some cases under t-distribution. The power of all analyzed tests improved with the increase of the sample size to $n = 1000$. In many cases also improvement in the size is visible.

Figure 1. The size ($t(5)$ distribution, H=6) and the power ($t(5)$ vs $t(6)$ distributions, H=6) of the analyzed tests with respect to the sample size



Source: own calculation.

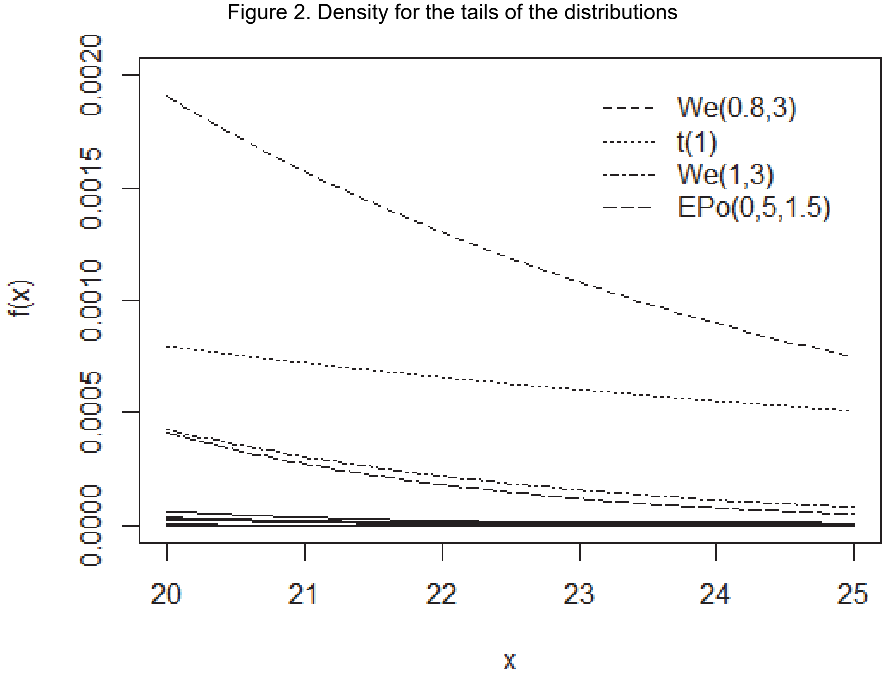
Values of the cumulative distribution functions for the chosen truncation points range from 0.4592 to 0.9999. In insurance data studies even lower levels are considered. Respective values in Chernobai et al. (2006) amounted to 0.0387 and 0.8212 for the conditional distributions. On the other hand, in finance, some of the risk measures like VaR and CVaR, are based on as high quantiles of distributions as 0.95, 0.975 or 0.99, as was mentioned at the beginning of section 2.

The actual distributions in experiments are: normal, t-distribution, power-exponential and Weibull. The notation used for these distributions in the paper is as follows: $N(\mu,\sigma^2)$ for the normal distribution, where $\mu \in \mathbb{R}$ is the location parameter and $\sigma > 0$ is the scale parameter; $t(n)$ for the t-distribution, where $n \in \mathbb{N}$ denotes the degrees of freedom; $EPo(\mu,\sigma_p,p)$ for the power-exponential distribution, where $\mu \in \mathbb{R}$, $\sigma_p > 0$ and $p > 1$ are the location, scale and shape parameters, respectively; $We(\alpha,\beta)$ for the Weibull distribution with the shape parameter $\alpha > 0$ and the scale parameter $\beta > 0$. Normal distribution, t-distribution and power-exponential distribution were used among other distributions by Piasecki, Tomasik (2013) to verify the shapes of the log asset returns on the polish market. The values of the estimators of these distributions' parameters, e.g. for the WIG index for the chosen period labeled as "h3", were as follows: $N(0.1956, 1.8623)$, $t(2.1561)$, $EPo(0.1982, 1.5881, 1.3618)$. The t-distribution used by Piasecki, Tomasik (2013) is the generalized t-distribution with $n \in \mathbb{R}^+$, while EPo is referred to as generalized error distribution (GED). Burnecki et al. (2015) studied the tails of the asset returns and considered among others: normal distribution $N(0, 2)$ and t-distribution $t(4)$. Weibull distribution was used for instance to assess the operational risk by Guegan, Hassani (2018). The estimated parameters of the distributions for the whole analyzed period were as follows $We(0.5896, 182.9008)$. However, the standard version of the Weibull distribution is rarely used. If X has Weibull distribution, then $-X$ has extreme value distribution of type III and is used in extreme value theory (Magiera, 2005). For the detailed overview of domestic and external researches in the field of asset returns distributions please refer to Piasecki, Tomasik (2013).

For each distribution, three sets of parameters were considered. As average return rates on the stock exchange are not significantly different from zero, all the location parameters of the considered distributions (if they exist) were set to zero. To obtain different thickness of tails, the remaining parameters were changed. In case of the normal distribution, the thickness of the exponential tail was controlled by the standard deviation, that was set to $\sigma \in \{3,4,5\}$. In case of the t-distribution, the thickness of the power tail was controlled by the number of degrees of freedom. They were set to $n \in \{1,3,5\}$. For the power-exponential distribution, the parameter p was set to 1.5, so the distribution has the tail thicker than exponential. Here the thickness of the semi-heavy tail was controlled by $\sigma_p \in \{3,4,5\}$. In case of the Weibull distribution, $\alpha \in \{0.8,1,1.2\}$, which results in the power, exponential and faster than exponential decaying tail.

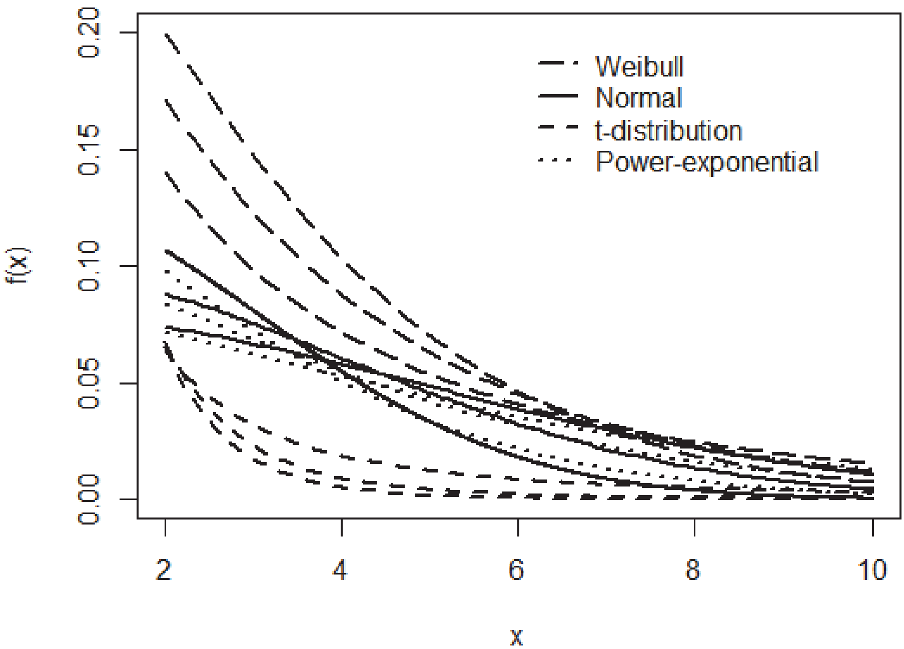
The tails of the distributions considered in the simulation studies are visualized in figure 2. The greatest probability mass in the tails appear in the $We(0.8,3)$, $t(1)$, $We(1,3)$ and $EPo(0,5,1.5)$ distributions, respectively. The tails of the remaining distributions practically disappeared.

Figure 2. Density for the tails of the distributions



Source: own calculation.

Figure 3. Probability distributions used in power study



Source: own calculation.

3.2. Discussion of simulation results

Empirical sizes and powers of a test are identified with the number of rejected null hypothesis. The empirical size of a test should be close to a determined significance level. The empirical power should be as large as possible.

The empirical sizes of the tests obtained in the simulation studies are presented in table 1. The results suggest dividing the tests into three groups. First group contains the KS^*, W^{2*} and AD^{2*} tests. In this group, the empirical sizes were on average 8-times higher than the determined significance level. The tests achieved visibly better results for the t-distribution, but the average rate of rejection was here still 4-times higher than the determined significance level. The second group includes the following tests: V^* and AD^{2*}_{up}. For these tests, the average rate of rejection was 3-times higher than the significance level. These tests also noted better results for the t-distribution. Here probability of rejection of the true null hypothesis was twice higher than the significance level. The third group consists of the remaining two tests: AD^* and AD^*_{up}. In this group, the rate of rejection of the null hypothesis was on average 1.5-times higher than the determined significance level. No clear differences among the distributions were detected.

On the basis of figure 1 and the above results it can be stated, that the considered tests require large number of observations to control the type I error. Unfortunately, the tests in the first and second group may not keep the preassigned type I error level even for samples of 1000 observations. In case of the fat tailed distribution like $t(1)$, the Authors made additional calculations to verify the size of the tests, when threshold levels are very high, $H = 50, 100, 150, 200, 250$. The size of the tests remained similar to the ones given in table 1 or figure 1. The tests from the first group are the most liberal, and the results of the powers of these tests will not be further analyzed. The results for the tests from the second group will be presented only for the illustrative purposes.

Table 1. EMPIRICAL SIZES OF THE TESTS (AS PERCENTAGES, $\alpha = 5\%$)

Actual distribution	H	KS^*	V^*	AD^*	W^{2*}	AD^{2*}	AD_{up}^*	AD_{up}^{2*}
$N(0,3)$	2	48.1	16.3	7.0	51.8	50.6	7.5	14.9
	4	44.9	15.0	7.0	49.6	48.3	7.9	16.7
	6	42.3	13.8	6.9	46.9	45.0	8.2	17.5
	8	40.4	12.5	7.2	44.3	42.1	8.3	16.6
	10	37.0	12.2	7.4	41.8	39.6	7.9	15.8
$N(0,4)$	2	49.3	17.0	6.7	52.9	51.5	7.5	14.6
	4	46.6	15.5	7.0	50.8	48.9	7.6	15.6
	6	44.2	14.7	6.9	49.2	47.5	8.0	17.2
	8	43.0	13.9	7.0	46.8	44.8	8.3	17.9
	10	41.2	13.5	7.2	45.4	43.7	8.3	17.6
$N(0,5)$	2	49.5	16.8	6.7	53.0	52.2	7.4	14.2
	4	47.4	16.2	7.0	51.6	50.2	7.6	15.2
	6	45.5	15.2	7.0	50.2	48.4	7.8	16.5
	8	44.0	14.3	6.9	48.4	46.7	8.0	16.5
	10	42.4	13.3	7.1	46.7	44.4	8.4	17.0
$t(1)$	2	18.8	9.7	6.4	22.9	22.8	7.5	9.4
	4	20.9	9.9	6.3	24.3	24.2	7.5	9.3
	6	21.1	9.9	6.3	24.6	24.3	7.5	9.3
	8	21.2	9.9	6.3	24.6	24.4	7.5	9.3
	10	21.2	9.9	6.3	24.6	24.4	7.5	9.3
$t(3)$	2	16.0	9.3	6.4	20.4	20.2	7.4	9.8
	4	18.8	9.8	6.4	23.1	23.2	7.4	9.6
	6	20.7	9.8	6.3	24.0	24.0	7.5	9.4
	8	21.1	9.9	6.3	24.3	24.2	7.5	9.3
	10	21.1	9.9	6.3	24.5	24.3	7.5	9.3
$t(5)$	2	15.5	9.2	6.5	18.8	19.6	7.7	10.0
	4	18.3	9.7	6.4	22.8	22.8	7.4	9.6
	6	20.5	9.7	6.3	23.6	23.6	7.4	9.5
	8	20.9	9.9	6.3	24.3	24.2	7.5	9.4
	10	21.1	9.9	6.3	24.4	24.2	7.5	9.3
$EPo(0,3,1.5)$	2	52.5	24.5	7.6	61.0	61.6	9.7	24.7
	4	49.0	19.9	7.9	55.6	55.8	10.7	26.8
	6	45.9	17.2	8.0	51.9	51.1	10.7	25.4
	8	43.3	16.0	7.8	48.6	47.8	9.9	23.7
	10	41.7	14.7	7.8	46.1	45.9	9.5	22.2
$EPo(0,4,1.5)$	2	53.9	25.4	7.6	62.5	62.7	9.2	24.2
	4	51.3	21.7	7.8	58.2	58.7	10.4	26.2
	6	49.2	19.1	7.9	54.8	55.4	10.9	27.1
	8	45.6	17.2	8.0	51.3	50.4	10.5	24.9
	10	43.4	16.1	8.0	49.1	47.8	9.9	22.9

Table 1. EMPIRICAL SIZES OF THE TESTS (AS PERCENTAGES, $\alpha = 5\%$) (cont.)

Actual distribution	H	KS^*	V^*	AD^*	W^{2*}	AD^{2*}	AD_{up}^*	AD_{up}^{2*}
$EPo(0,5,1.5)$	2	54.6	25.6	7.6	62.6	64.2	9.3	23.1
	4	51.2	23.3	7.8	60.6	59.8	10.2	25.7
	6	50.0	20.2	7.9	56.5	56.0	10.7	27.0
	8	47.7	18.4	7.8	53.1	53.1	10.4	25.6
	10	45.1	16.7	7.6	50.7	50.0	10.4	23.9
$We(0.8,3)$	2	47.4	15.6	7.0	51.2	49.8	7.6	15.8
	4	44.7	14.8	6.9	49.5	48.5	7.8	17.0
	6	44.0	14.3	7.0	48.4	46.2	7.9	17.9
	8	43.1	14.0	7.1	47.3	45.5	8.2	18.3
	10	42.7	13.8	7.2	46.3	45.0	8.2	18.5
$We(1,3)$	2	48.0	15.6	7.0	51.5	50.1	7.6	15.7
	4	44.3	14.5	6.9	49.2	47.9	7.8	17.2
	6	43.3	14.3	7.0	47.7	45.7	8.2	18.0
	8	42.5	13.7	7.2	46.4	45.1	8.2	18.5
	10	41.7	13.2	7.3	45.3	44.7	8.2	19.0
$We(1.2,3)$	2	48.3	16.0	7.0	51.8	50.6	7.6	15.4
	4	44.3	14.5	6.9	49.1	47.5	7.9	17.3
	6	43.0	14.0	7.1	47.2	45.6	8.2	18.3
	8	41.8	13.5	7.3	45.6	44.7	8.3	18.7
	10	40.5	13.2	7.4	44.4	44.1	8.3	19.2

Source: own calculation.

Empirical powers of the tests are presented in table 2. The visualization of the probability distributions used in the power study is presented in figure 3. The tests from the second group, that is the V^* and AD_{up}^{2*} tests, have high empirical powers, that on average amount to 97% for the first test and 98% for the second test. However, it has to be reminded, that these tests are too liberal. The AD^* and AD_{up}^* tests from the third group are much more realistic. While the average rate of rejection of the false null hypothesis for the first test is 64%, it is only 30% for the second test. The results for the AD_{up}^* are very irregular. For the distributions with the fast decaying tails, that is for the normal and power-exponential ones, the average powers are lower than 3%.

The powers of the V^*, AD_{up}^{2*} and AD^* tests show common behavior with respect to the decaying rates of the tails. With regard to the distributions with the fast decaying tails, that is the normal and power-exponential distributions, the empirical powers of the tests decrease with the growing thickness of the tail. In case of the distributions with thicker tails, that is the t-distribution and Weibull distribution, the relation is opposite, the powers of the tests increase with the growing thickness of the tail. Summarizing, the powers of the tests are higher for extreme tails, that are decaying exponentially or powerly. In case of the distributions with semi-heavy tails, the considered tests had more problems with recognizing the actual distribution. It is also worth noting that the powers of the V^*, AD_{up}^{2*} and AD^* tests were increasing with the growth of the truncation point.

Table 2. EMPIRICAL POWERS OF THE TESTS (AS PERCENTAGES, $\alpha = 5\%$)

Null hypothesis	Actual distribution	H	V^*	AD^*	AD_{up}^*	AD_{up}^{2*}
$N(0,3.5)$	$N(0,3)$	2	100.0	71.3	0.4	99.9
		4	100.0	81.0	0.5	100.0
		6	100.0	87.0	0.5	100.0
		8	100.0	89.5	0.4	100.0
		10	100.0	90.5	0.5	100.0
$N(0,4.5)$	$N(0,4)$	2	99.2	35.1	1.2	98.6
		4	99.7	45.5	1.2	98.8
		6	99.7	53.2	1.2	99.5
		8	99.8	59.8	1.2	99.6
		10	99.9	64.7	1.2	99.7
$N(0,5.5)$	$N(0,5)$	2	93.4	15.6	1.8	91.4
		4	95.8	19.8	1.7	94.8
		6	97.1	23.6	1.8	96.7
		8	98.3	27.9	1.8	97.4
		10	98.5	32.7	1.8	97.6
$t(2)$	$t(1)$	2	100.0	100.0	100.0	100.0
		4	100.0	100.0	100.0	100.0
		6	100.0	100.0	100.0	100.0
		8	100.0	100.0	100.0	100.0
		10	100.0	100.0	100.0	100.0
$t(4)$	$t(3)$	2	99.3	83.6	48.8	99.6
		4	100.0	93.4	53.1	100.0
		6	100.0	95.3	54.2	100.0
		8	100.0	96.0	55.7	100.0
		10	100.0	96.2	55.9	100.0
$t(6)$	$t(5)$	2	56.0	28.4	24.3	83.9
		4	90.4	44.9	27.5	95.1
		6	95.4	52.2	28.9	97.8
		8	97.4	56.1	29.5	98.4
		10	98.0	57.7	29.9	98.5
$EPo(0,3.5;1.5)$	$EPo(0,3,1.5)$	2	99.9	63.4	1.4	100.0
		4	100.0	69.7	1.4	100.0
		6	100.0	73.8	1.5	100.0
		8	100.0	74.0	1.3	100.0
		10	100.0	74.4	1.3	99.9
$EPo(0,4.5,1.5)$	$EPo(0,4,1.5)$	2	97.7	26.4	2.1	98.3
		4	98.1	32.7	2.2	98.9
		6	98.2	36.4	2.1	99.3
		8	98.3	38.4	2.0	99.2
		10	98.2	38.9	2.0	99.3
$EPo(0,5.5,1.5)$	$EPo(0,5,1.5)$	2	91.3	14.4	2.8	91.0
		4	91.9	16.8	2.6	94.1
		6	92.0	19.0	2.8	95.6
		8	91.5	19.9	2.7	96.1
		10	91.5	20.0	2.4	95.9

Table 2. EMPIRICAL POWERS OF THE TESTS (AS PERCENTAGES, $\alpha = 5\%$) (cont.)

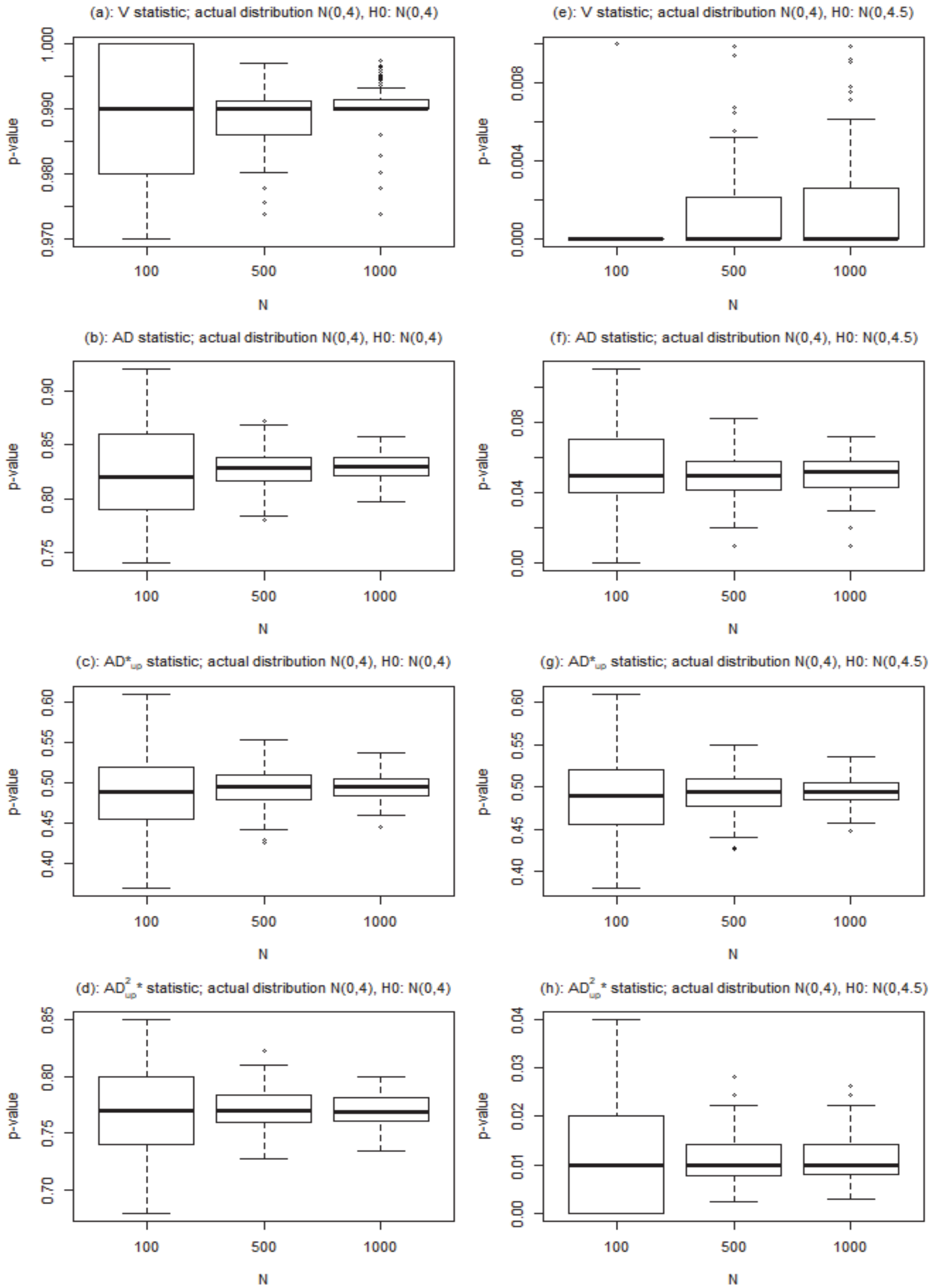
Null hypothesis	Actual distribution	H	V^*	AD^*	AD^*_{up}	AD^{2*}_{up}
$We(0.9{,}3)$	$We(0.8{,}3)$	2	95.8	80.3	58.5	99.7
		4	99.9	90.4	64.8	100.0
		6	99.9	94.9	69.3	100.0
		8	100.0	97.2	72.5	100.0
		10	100.0	98.3	75.9	100.0
$We(1.1{,}3)$	$We(1{,}3)$	2	82.9	58.4	43.0	98.2
		4	96.9	74.2	49.3	99.8
		6	99.8	84.7	55.3	100.0
		8	99.8	92.0	59.3	100.0
		10	99.9	95.2	62.3	100.0
$We(1.3{,}3)$	$We(1.2{,}3)$	2	67.4	43.6	35.2	93.1
		4	90.7	60.1	39.8	99.2
		6	97.8	73.7	43.9	99.9
		8	99.8	82.9	49.0	100.0
		10	99.8	89.5	53.4	100.0

Source: own calculation.

Due to the time-consuming procedures, the *p*-values were calculated on the basis of $N = 100$ Monte Carlo samples, the default value of N in the truncgof package. To justify the obtained results, the randomness of the *p*-values was studied for the selected cases (similar analysis was considered in Smaga, 2017). The tests chosen to the power study were applied 100 times to a single data set, with different values of N. The study was performed for the actual distribution $N(0{,}4)$, under the true and false null hypothesis. Figure 4 presents the results. The median of each analysed case does not vary considerably between different numbers of N. The variance of *p*-values decreases with the increase of N, therefore, in the unconvincing cases, it is recommended to repeat the tests with a higher value of N. This may at least slightly improve the results, for example, the empirical sizes of the tests AD^*, AD^{2*}, KS^*, V^*, W^{2*}, AD^*_{up} and AD^{2*}_{up} were equal to 7.5, 42.8, 39.4, 13.6, 44.8, 7.6 and 13 respectively for $N = 1000$, actual distribution $N(0{,}4)$ and $H = 6$, while for $N = 100$, they were equal to 6.9, 47.5, 44.2, 14.7, 49.2, 8 and 17.2 respectively.

All the calculations were done in R statistical environment (R Core Team, 2017). Except the R package truncgof (Wolter, 2012), the R packages normalp (Mineo, 2014) and doParallel (Calaway et al., 2017) were also used, since they deal with power-exponential distributions and parallel computing, respectively. When writing the code in R program, many tips and hints were drawn from the handbook written by Górecki (2011).

Figure 4. Boxplots for the randomness of the *p*-values



For each boxplot the testing procedure was repeated 100 times. For a-d the null, and for e-h the alternative hypothesis was true. The p-values determined on the basis of $N = 100000$ for a-h were as follows: 0.99010, 0.80513, 0.49084, 0.76880, 0.00053, 0.04464, 0.48980, 0.00680.

Source: own calculation.

4. CONCLUSIONS

The aim of this paper was to present the results of the simulation studies, that were evaluating the finite sample behavior of the tests for truncated distributions introduced in Chernobai et al. (2015) and implemented in the R package truncgof (Wolter, 2012). The tests were performed with default values of parameters used in the truncgof package. The research was based on artificial data generated from the distributions that are often describing the tails of asset returns. The study was conducted for different tail thickness and for changing truncation point. In the cases considered in the article, the KS^*, W^{2*}, AD^{2*}, V^* and AD_{up}^{2*} tests did not maintain the preassigned type I error level. The remaining two tests, AD^* and AD_{up}^*, obtained reasonable rejection rates for the true null hypothesis. The power of the AD^* test was much higher than the power of the AD_{up}^* test. While the average rate of rejection of the false null hypothesis for the first test is 64%, it is only 30% for the second test. It was also noticed, that the power of AD^* and AD_{up}^* tests is higher for extreme tails and it grows with the truncation point. On the basis of the obtained results, it is recommended to assess the behavior of the tests analyzed in this article, in terms of the sample size, theoretical distribution and truncation point, before every application. In the unconvincing cases (e.g., when the *p*-value is close to the significance level), it is suggested to use greater number N of Monte Carlo samples to estimate the *p*-values of the tests than $N = 100$, which is the default value of N in the truncgof package.

REFERENCES

Anderson T. W., Darling D. A., (1952), Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes, *The Annals of Mathematical Statistics*, 23 (2), 193–212.

Brunner M. I., Viviroli D., Sikorska A. E., Vannier O., Favre A.-C., Seibert J., (2017), Flood Type Specific Construction of Synthetic Design Hydrographs, *Water Resources Research*, 53 (2), 1390–1406.

Burnecki K., Chechkin A., Wylomanska A., (2015), Discriminating Between Light- and Heavy-Tailed Distributions with Limit Theorem, *PLoS ONE*, 10 (12), 1–23.

Chernobai A., Burnecki K., Rachev S., Trück S., Weron R., (2006), Modelling Catastrophe Claims with Left-Truncated Severity Distributions, *Computational Statistics*, 21 (3), 537–555.

Chernobai A., Menn C., Rachev S. T., Trück S., (2010), Estimation of Operational Value-at-Risk in the Presence of Minimum Collection Threshold: An Empirical Study, Working Paper Series in Economics, 4, Karlsruhe Institute of Technology (KIT), Department of Economics and Business Engineering.

Chernobai A., Rachev S. T., Fabozzi F. J., (2015), Composite Goodness-of-Fit Tests for Left-Truncated Loss Samples, in: Lee C. F., Lee J., (eds.), *Handbook of Financial Econometrics and Statistics,* Springer, New York, NY.

Clementi F., Gallegati M., Kaniadakis G., (2012), A New Model of Income Distribution: The κ – Generalized Distribution, *Journal of Economics*, 105 (1), 63–91.

Cox D. R., Spjøtvoll E., Johansen S., Van Zwet W. R., Bithell J. F., Barndorff-Nielsen O., (1977), The Role of Significance Tests, *Scandinavian Journal of Statistics*, 4 (2), 49–70.

Echaust K., (2014), *Ryzyko zdarzeń ekstremalnych na rynku kontraktów futures w Polsce*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań.
Fagiolo G., Alessi L., Barigozzi M., Capasso M., (2010), On the Distributional Properties of Household Consumption Expenditures: The Case of Italy, *Empirical Economics*, 38 (3), 717–741.
Feller W., (1950), *An Introduction to Probability Theory and Its Applications*, Wiley, New York.
Fischer M., Jakob K., (2016), pTAS Distributions with Application to Risk Management, *Journal of Statistical Distributions and Applications*, 3 (1), 1–18.
Górecki T., (2011), *Podstawy statystyki z przykładami w R*, Wydawnictwo BTC, Legionowo.
Górecki T., Smaga Ł., (2015), A Comparison of Tests for the One-Way ANOVA Problem for Functional Data, *Computational Statistics*, 30, 987–1010.
Guegan D., Hassani B.K., (2018), More Accurate Measurement for Enhanced Controls: VaR vs ES?, *Journal of International Financial Markets, Institutions & Money*, 54, 152-165.
Haas M., Pigorsch C., (2009), Financial Economics, Fat-Tailed Distributions, *Encyclopedia of Complexity and Systems Science*, 4, 3404–3435.
Kim J. H., Ji P. I., (2015), Significance Testing in Empirical Finance: A Critical Review and Assessment, *Journal of Empirical Finance*, 34, 1–14.
Krzyśko M., (2000), *Wykłady z teorii prawdopodobieństwa*, Wydawnictwa Naukowo-Techniczne, Warszawa.
Krzyśko M., (2004), *Statystyka matematyczna*, Wydawnictwo Naukowe UAM, Poznań.
Lach A., (2017), Testy zgodności z rozkładami uciętymi, *Praca licencjacka*, Uniwersytet im. Adama Mickiewicza, Poznań.
Magiera R., (2005), *Modele i metody statystyki matematycznej. Cz. 1: Rozkłady i symulacja stochastyczna*, Oficyna Wydawnicza GiS, Wrocław.
Mandelbrot B., (1963), The Variation of Certain Speculative Prices, *The Journal of Business*, 36, 394-419.
Mineo A.M., (2014), normalp: Routines for Exponential Power Distribution, R package version 0.7.0, https://CRAN.R-project.org/package=normalp.
Orzeszko W., (2014), Symulacyjna ocena rozmiaru testu BDS, *Przegląd Statystyczny*, 4 (61), 335–361.
Pavia J. M., (2015), Testing Goodness-of-Fit with the Kernel Density Estimator: GoFKernel, *Journal of Statistical Software*, 66, 1–27.
Piasecki K., Tomasik E., (2013), *Rozkłady stóp zwrotu z instrumentów polskiego rynku kapitałowego*, Wydawnictwo edu-Libri, Kraków, Warszawa.
R Core Team, (2017), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/.
Revolution Analytics, Weston S., (2015), *doParallel: Foreach Parallel Adaptor for the 'parallel' Package*, R package version 1.0.10, http://CRAN.R-project.org/package=doParallel.
Smaga Ł., (2017), A Two-Sample Test Based on Cluster Subspaces for Equality of Mean Vectors in High Dimension, *Discussiones Mathematicae Probability and Statistics*, 37, 147–156.
Wolter T., (2012). Truncgof: GoF Tests Allowing for Left Truncated Data, R package version 0.6-0, https://CRAN.R-project.org/package=truncgof.

PORÓWNANIE TESTÓW ZGODNOŚCI DLA ROZKŁADÓW UCIĘTYCH

Streszczenie

Celem artykułu jest empiryczne zbadanie mocy i rozmiaru siedmiu testów zgodności, zaprezentowanych w pracy Chernobai i inni (2015), przeznaczonych dla rozkładów lewostronnie uciętych. Badania symulacyjne oparto na danych, wygenerowanych z rozkładów, które były w przeszłości lub są obecnie sto-

sowane do opisu ogonów rozkładów stóp zwrotu. Badania przeprowadzono dla różnych grubości ogonów rozkładów oraz zmieniających się poziomów ucięcia. Wyniki symulacji wskazują na istnienie znacznych różnic pomiędzy poszczególnymi procedurami testowymi. Ponadto otrzymanie zadowalających wyników w przypadku niektórych procedur wymaga dość dużej liczby obserwacji.

Słowa kluczowe: moc testu, program R, rozkłady ucięte, rozmiar testu, testy zgodności

COMPARISON OF THE GOODNESS-OF-FIT TESTS FOR TRUNCATED DISTRIBUTIONS

Abstract

The aim of this paper is to investigate the finite sample behavior of seven goodness-of-fit tests for left truncated distributions of Chernobai et al. (2015) in terms of size and power. Simulation experiments are based on artificial data generated from the distributions that were used in the past or are used nowadays to describe the tails of asset returns. The study was conducted for different tail thickness and for changing truncation point. Simulation results indicate that the testing procedures do not work equally well under finite samples, and some of them require quite large number of observations to perform satisfactorily.

Keywords: goodness of fit tests, power of test, R program, size of test, truncated distributions

DOI: 10.5604/01.3001.0014.0543

PRZEGLĄD STATYSTYCZNY
TOM LXV – ZESZYT 3 – 2018

Piotr SULEWSKI[1]

Nonparametric versus parametric reasoning based on two-way and three-way contingency tables

1. INTRODUCTION

The analysis of contingency tables (CTs) is one of the most common tasks performed by statisticians. CTs display the frequency distribution of two (two-way CTs), three (three-way CTs) or more (multi-way CTs) categorical variables. The information about categorical data can be found e.g. in Bishop et al. (1975), Agresti (2002), Van Belle et al. (2014). Information presented as CTs features in a wide variety of areas such as the social sciences (Wickens, 1969), genetics (El Galta et al., 2008; Dickhaus et al., 2012), demography (Cung, 2013) and psychology (Iossifova et al., 2013). Basic methods of testing for dependency in CTs in details is described e.g. in Steinle et al. (2006), Bock (2003), Kaski et al. (2005), Allison, Liker (1982). Other examples of applications may be found in Ilyas et al. (2004), Oates, Cohen (1996), Schrepp (2003), Haas et al. (2007).

One can recognize two general cases in which CTs can be useful. This distinction between the cases is made with respect to the tasks which CTs are used for.

Case A. Dependency is unwanted. The general population is sought to be in its normal state or be under control when levels of feature X are independent of levels of feature Y. Revealing dependency means revealing abnormality of members of the general population. If so, a large scale and very costly actions have to be obligatorily initiated. That is why a decision-maker tries to avoid false alarm. This case is typical, for instance, in security guarding. A classic statistical way of reasoning is tailored to case A. Please notice that the main hypothesis, commonly denoted by H_0, states that: X and Y are independent. Moreover, H_0 is guarded against rejection by setting significance level at 5% or less.

Case B. Dependency is wanted. The state of the general population is assessed upon feature X. Unfortunately, levels of feature X are difficult to be determined e.g. determination is risky, costly or time consuming. In contrast, levels of another feature Y are easy to be determined. Assessors are concerned with finding out whether there is a tie between X and Y. In other words, whether X

[1] The Pomeranian University, Institute of Mathematics, 6–7 Kozietulskiego St., 76–200 Słupsk, Poland, e-mail: piotr.sulewski@apsl.edu.pl.

and Y are dependent or independent. Assessors use $Y_1,\dots,Y_k$ levels as sensible indicators of $X_1,\dots,X_w$ levels. Case B is typical in diagnostics, both medical and technical. In case B another way of statistical reasoning is needed, different from the classic way.

Conservativeness of the classic statistical way of reasoning often obstructs progress in numerous situations where rejecting H_0 means making a step ahead. This is a strong motivation for making a turnaround in statistical reasoning. In this new statistical reasoning there is no null hypothesis. In contrast to the classic way, there is a set of competing hypotheses. Moreover, the testing procedure warrants equality of all the alternatives when the test begins. The former null hypothesis is no longer the main one, but exists among the other ones of equal importance. Particular hypotheses relate to scenarios under which particular CTs are created. Details are presented in section 7. There are two reasons for which this likelihood based reasoning is developed and put forward:

a) Undoubtedly, CT-based classic statistical reasoning is the nonparametric reasoning. It is commonly known that parametric statistical reasoning, if applicable, is much more sensitive to untruthfulness of H_0 than nonparametric reasoning. In this paper we propose a parametric reasoning. Particular scenarios are parameterized with the probability flow parameter (PFP).
b) Let us again retrace a way of the classic thinking. A value of the test statistics is smaller than the appropriate critical value results in failing to reject H_0. In case A the decision maker is comfortable about independence. A value of the test statistics is no smaller than the appropriately determined critical value results in rejecting H_0. In case B the decision maker is comfortable about dependence because there is no word said what the reason of rejecting H_0 is. The most likely scenario is selected whereas reasons to reject H_0 or not are embedded in scenarios. One can say that the method put forward in this paper offers a transition from ”unfathomable” to ”fathomable” reasons.

Nonparametric and parametric reasoning based on 2×2 CT is presented in (Sulewski, 2018b), therefore this paper is devoted to bigger tables, e.g. 2×3, 2×4, 3×3, 3×4, 4×4, $2 \times 2 \times 2$, $3 \times 2 \times 2$ ones.

This paper is organized as follows. Variants of presentation of CT are described in section 2. CT coming into being are presented in section 3. Statistic tests including the power divergence tests and the $|\chi|$ test are defined in section 4. Section 5 is devoted to measures of untruthfulness of H_0 including the measure that is defined by means of an absolute value. In section 6 maximum likelihood method is applied to estimate the PFP. Section 7 is devoted to instructions how to generate two-way and three-way CTs. Section 8 presents numerical examples and section 9 presents closing remarks.

Monte Carlo simulation is performed in Visual Basic for Applications embedded in Microsoft Excel 2016.

2. VARIANTS OF PRESENTATION OF CT

This section is devoted to the $w \times k \times p$ CT. If $p = 1$, then we obviously have $w \times k$ CT.

Let X, Y, Z be three features of the same object, respectively, have levels $X_1, \dots, X_w$, $Y_1, \dots, Y_k$, $Z_1, \dots, Z_p$. Testing these three features for independence with an appropriately arranged CT is probably one of the most common statisticians' tasks. At the moment one can distinct between four variants of presentation of CTs. It is because each variant is intended for a different purpose. Below details of particular variants are given:

— TP Variant (theoretical probabilities). Cells contain probabilities p_{ijt} intrinsic to the phenomenon being investigated (see table 1). The exact values of these probabilities are unknown to the investigator. This variant is introduced a little bit in advance since CTs will be simulated with the Monte Carlo method in further sections of this paper. And just then CT variant filled with probabilities arbitrarily set by Monte Carlo experimenter will be applied.

— TC Variant (theoretical counts). Cells contain theoretical expected counts $n_{ijt} = np_{ijt}$. These counts are theoretical in this sense that they result from TP variant.

— EP Variant (empirical probabilities). Cells that result from EC variant and contain estimates $p^*_{ijt} = n^*_{ijt}/n$ of the unknown content of TP.

— EC Variant (experimental counts). Cells contain n^*_{ijt} counts observed on a sample drawn from general population subjected to the investigation.

Table 1. TP VARIANT OF THREE-WAY CT

Z	Z_1			...	Z_p			Total
X / Y	Y_1	...	Y_k	...	Y_1	...	Y_k	
Y_1	P_{111}	...	P_{1k1}	...	P_{11p}	...	P_{1kp}	$P_{1\bullet\bullet}$
...	...	...	...	...	...	...	...	...
X_w	P_{w11}	...	P_{wk1}	...	P_{w1p}	...	P_{wkp}	$P_{w\bullet\bullet}$
Total	$P_{\bullet 11}$	...	$P_{\bullet k1}$	...	$P_{\bullet 1p}$	...	$P_{\bullet kp}$	1

3. ON HOW CT COMES INTO BEING

One can treat CTs as a mathematical expression of a certain phenomenon we deal with. This formulation suggests that there is an internal mechanism in this phenomenon that determines probabilities of particular X, Y or X, Y, Z combinations and ascribes these probabilities to the cells of the table. Below are "progenitors" of all the $w \times k$ CTs

$$T_p = \begin{bmatrix} 1/(wk) & \cdots & 1/(wk) \\ \vdots & \cdots & \vdots \\ 1/(wk) & \cdots & 1/(wk) \end{bmatrix} \tag{1}$$

and all the $w \times k \times p$ CTs

$$W_p = \begin{bmatrix} 1/(wkp) & \cdots & 1/(wkp) & \cdots & 1/(wkp) & \cdots & 1/(wkp) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1/(wkp) & \cdots & 1/(wkp) & \cdots & 1/(wkp) & \cdots & 1/(wkp) \end{bmatrix} \tag{2}$$

A variety of tables may be generated when portions of PFP a flow from ”maternal” cells of (1) or (2) to other cells. Obviously, the total probability always equals 1. In this paper twenty eight scenarios that seem fundamental are developed (tables 2–3). These scenarios are created based on scenarios for the 2×2 CT that seem fundamental and describe different levels of dependence (Sulewski, 2018b).

Table 2. THE CONTENTS OF $w \times k$ CTs RESULTING FROM SCENARIOS IN QUESTION

Table	Scenario	Content
2×3	I	$p_{11} = 1/6 - a, p_{12} = p_{13} = p_{21} = p_{22} = 1/6, p_{23} = 1/6 + a$
	II	$p_{11} = p_{12} = 1/6 - a, p_{13} = p_{21} = 1/6, p_{22} = p_{23} = 1/6 + a$
	III	$p_{11} = p_{12} = 1/6 - a, p_{13} = p_{23} = 1/6, p_{21} = p_{22} = 1/6 + a$
	IV	$p_{11} = p_{23} = 1/6 - a, p_{12} = p_{22} = 1/6, p_{13} = p_{31} = 1/6 + a$
2×4	V	$p_{11} = 1/8 - a, p_{12} = p_{13} = p_{14} = p_{21} = p_{22} = p_{23} = 1/8, p_{24} = 1/8 + a$
	VI	$p_{11} = p_{12} = p_{13} = 1/8 - a, p_{14} = p_{21} = 1/8, p_{22} = p_{23} = p_{24} = 1/8 + a$
	VII	$p_{11} = 1/8 - a, p_{12} = p_{13} = p_{14} = p_{22} = p_{23} = p_{24} = 1/8, p_{21} = 1/8 + a$
	VIII	$p_{11} = p_{12} = 1/8 - a, p_{13} = p_{14} = p_{23} = p_{24} = 1/8, p_{21} = p_{22} = 1/8 + a$
3×3	IX	$p_{11} = 1/9 - a, p_{12} = p_{13} = p_{21} = p_{22} = p_{23} = p_{31} = p_{32} = 1/9, p_{33} = 1/9 + a$
	X	$p_{11} = p_{12} = 1/9 - a, p_{13} = p_{21} = p_{22} = p_{23} = p_{33} = 1/9, p_{31} = p_{32} = 1/9 + a$
	XI	$p_{11} = p_{33} = 1/9 - a, p_{12} = p_{21} = p_{22} = p_{23} = p_{32} = 1/9, p_{13} = p_{31} = 1/9 + a$
	XII	$p_{11} = p_{21} = p_{33} = 1/9 - a, p_{12} = p_{22} = p_{32} = 1/9, p_{13} = p_{23} = p_{31} = 1/9 + a$
3×4	XIII	$p_{11} = 1/12 - a, p_{12} = p_{13} = p_{14} = p_{21} = p_{22} = p_{23} = p_{24} = p_{31} = p_{32} = p_{33} =$ $= 1/12, p_{34} = 1/12 + a$
	XIV	$p_{11} = p_{12} = 1/12 - a, p_{13} = p_{14} = p_{21} = p_{22} = p_{23} = p_{24} = p_{33} = p_{34} =$ $= 1/12, p_{31} = p_{32} = 1/12 + a$
	XV	$p_{11} = p_{12} = 1/12 - a, p_{21} = p_{22} = p_{23} = p_{24} = p_{31} = p_{32} = p_{33} = p_{34} =$ $= 1/12, p_{13} = p_{14} = 1/12 + a$
	XVI	$p_{11} = p_{12} = p_{33} = p_{34} = 1/12 - a, p_{21} = p_{22} = p_{23} = p_{24} = 1/12, p_{13} =$ $= p_{14} = p_{31} = p_{32} = 1/12 + a$

Table 2. THE CONTENTS OF $w \times k$ CTs RESULTING FROM SCENARIOS IN QUESTION (cont.)

Table	Scenario	Content
4×4	XVII	$p_{11} = p_{12} = p_{13} = 1/16 - a, p_{14} = p_{21} = p_{22} = p_{23} = p_{24} = p_{31} = p_{32} = p_{33} = = p_{34} = p_{44} = 1/16, p_{41} = p_{42} = p_{43} = 1/16 + a$
	XVIII	$p_{11} = p_{12} = 1/16 - a, p_{13} = p_{14} = p_{21} = p_{22} = p_{23} = p_{24} = p_{31} = p_{32} = p_{33} = = p_{34} = p_{41} = p_{42} = 1/16, p_{43} = p_{44} = 1/16 + a$
	XIX	$p_{11} = p_{12} = p_{13} = p_{21} = p_{22} = p_{23} = 1/16 - a, p_{14} = p_{24} = p_{21} = = p_{31} = p_{41} == 1/16, p_{32} = p_{33} = p_{34} = p_{42} = p_{43} = p_{44} = 1/16 + a$
	XX	$p_{11} = p_{12} = p_{43} = p_{44} = 1/16 - a, p_{21} = p_{22} = p_{23} = p_{24} = p_{31} = p_{32} = p_{33} = p_{34} = 1/16, p_{13} = p_{14} = p_{41} = p_{42} = 1/16 + a$

Source: own elaboration.

Table 3. THE CONTENTS OF $w \times k \times p$ CTs RESULTING FROM SCENARIOS IN QUESTION

Table	Scenario	Content
$2 \times 2 \times 2$	XXI	$p_{111} = p_{112} = 1/8 - a, p_{121} = p_{122} = p_{211} = p_{212} = 1/8, p_{221} = p_{222} = 1/8 + a$
	XXII	$p_{111} = p_{121} = p_{112} = 1/8 - a, p_{122} = p_{211} = 1/8, p_{221} = p_{212} = p_{222} = 1/8 + a$
	XXIII	$p_{111} = 1/8 - a, p_{121} = p_{112} = p_{122} = p_{221} = p_{212} = p_{222} = 1/8, p_{211} = 1/8 + a$
	XXIV	$p_{111} = p_{112} = 1/8 - a, p_{121} = p_{122} = p_{221} = p_{222} = 1/8, p_{211} = p_{212} = 1/8 + a$
$3 \times 2 \times 2$	XXV	$p_{111} = p_{112} = 1/12 - a, p_{121} = p_{122} = p_{211} = p_{221} = p_{212} = p_{222} = p_{311} = p_{312} = = 1/12, p_{321} = p_{322} = 1/12 + a$
	XXVI	$p_{111} = p_{112} = 1/12 - a, p_{121} = p_{211} = p_{221} = p_{212} = p_{222} = p_{311} = = p_{312} = p_{322} = 1/12, p_{122} = p_{321} = 1/12 + a$
	XXVII	$p_{111} = p_{112} = p_{211} = 1/12 - a, p_{212} = p_{222} = p_{311} = p_{321} = = p_{312} = p_{322} == 1/12, p_{121} = p_{122} = p_{221} = 1/12 + a$
	XXVIII	$p_{111} = p_{112} = p_{321} = p_{322} = 1/12 - a, p_{211} = p_{221} = p_{212} = = p_{222} = 1/12, p_{121} = p_{122} = p_{311} = p_{312} = 1/12 + a$

Source: own elaboration.

In all the above scenarios the PFP a takes values in $\left[0, \frac{1}{wp}\right]$ or $\left[0, \frac{1}{wpk}\right]$. The scenarios are selected in such a way that they correspond to different levels of dependence expressed by means of an appropriate measure of untruthfulness of H_0 (MoU). The MoU takes values on interval $[0,1]$. A simulation study is carried out for MoU values no bigger than $2/3$. It is obvious that the detection of a strong dependence is very simple. You can find more information about the MoU in section 5.

Obviously, scenarios do not cover all the cases. They may be locally mutated by reversing rows or columns to better fit the analyzed data. These are simple equal-portion scenarios. In the scenarios you can use a part of PFP, e.g. $a/2$, $a/3, \ldots$. Surely, real scenarios can be more or less similar to these above. This is typical in relations between theory and real life. With the current availability of computers, the statistician can afford situations that interest him and instantly repeat such simulations. All examples presented here have a very precise algorithmic description in a form of a step list.

The researches can be generalized by introducing several PFPs. This, however, causes a significant deterioration in the properties of the parameter estimators. The Weibull distribution has a simple analytical form. For its generalization, the Generalized Gamma Distribution (URG) can be considered. Due to big problems with estimating URG parameters the author does not know any practical applications of URG to describe the reliability results of technical objects. You can always add more parameters to the model, however, this might worsen their estimation.

4. *INDEPENDENCE TESTS*

4.1. Two-way contingency table

Features X, Y are independent what means that H_0 is true, if $p_{ij} = p_{i\bullet} \cdot p_{\bullet j} (i, j = 1,2)$ for each pair of i, j. The alternative hypothesis denoted H_1 is such one that negates H_0. Let e_{ij} be the expected counts

$$e_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n} = n p_{i\bullet} p_{\bullet j} (i = 1, \dots, w; j = 1, \dots, k). \tag{3}$$

The expected counts e_{ij} have the same one-way marginal values as the observed table n_{ij} (Gokhale, Kullback, 1978).

Statistical science has been enriched with many other statistics intended for research on test independency. Cressie, Read (1984) propose the power divergence statistics (PDS). The PDS for $w \times k$ CTs is given by

$$\begin{aligned} P^2 &= \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^{w} \sum_{j=1}^{k} n_{ij}^* \left[\left(\frac{n_{ij}^*}{e_{ij}^*} \right)^{\lambda} - 1 \right] = \\ &= \frac{2n}{\lambda(\lambda+1)} \sum_{i=1}^{w} \sum_{j=1}^{k} p_{ij}^* \left[\left(\frac{p_{ij}^*}{p_{i\bullet}^* p_{\bullet j}^*} \right)^{\lambda} - 1 \right] - \infty < \lambda < \infty, \end{aligned} \tag{4}$$

where e_{ij} values are given by (3). Equation (4) always takes positive values and is defined as a limit of P^2 at -1 and 0. P^2 contains a very rich class of test statistics, for example: the χ^2 statistics ($\lambda = 1$), the G^2 statistics (the limit as λ goes to 0), the Freeman-Tukey statistics ($\lambda = -0.5$), the modified G^2 statistics (the limit as λ goes to -1), the Neyman modified χ^2 statistics ($\lambda = -2$) and the Cressie-Read statistics ($\lambda = 2/3$). If H_0 is true, statistics (4), for large n (i.e. asymptotically), follows the chi-square distribution with $(w-1)(k-1)$ degrees of freedom.

The following PDS are selected to Monte Carlo study: the χ^2 statistics (Pearson, 1900), the Freeman-Tukey *FT* statistics (Freeman, Tukey, 1950), the Cressie-Read *CR* statistics (Cressie, Read, 1984):

$$\chi^2 = \sum_{i=1}^{w}\sum_{j=1}^{k}\frac{\left(n_{ij}^{*} - e_{ij}\right)^2}{e_{ij}}, \tag{5}$$

$$FT = 4\sum_{i=1}^{w}\sum_{j=1}^{k}\left(\sqrt{n_{ij}^{*} - \sqrt{e_{ij}^{*}}}\right)^2, \tag{6}$$

$$CR = \frac{9}{5}\sum_{i=1}^{w}\sum_{j=1}^{k}n_{ij}^{*}\left[\left(\frac{n_{ij}^{*}}{e_{ij}^{*}}\right)^{2/3} - 1\right]. \tag{7}$$

The G^2 statistics (Sokal, Rohlf, 2012), the modified G^2 statistics (Kullback, 1959) and the Neyman modified χ^2 statistics (Neyman, 1949) have not been subjects in the Monte Carlo study because they are applicable only in a case where all n_{ij} $(i = 1, \dots, w; j = 1, \dots, k)$ counts are not equal to zero.

The square used in the numerator of χ^2 statistics (5) makes that large differences between expected and theoretical counts even bigger and the small differences even smaller. Another aim of the use of the square is to avoid that the differences are mutually exclusive. For this purpose one can use their absolute value instead of squared deviations. The $|\chi|$ statistics is selected to Monte Carlo study, too. It is an authorial modification of χ^2 statistics and it has the form (Sulewski, 2013)

$$|\chi| = \sum_{i=1}^{w}\sum_{j=1}^{k}\frac{\left|n_{ij}^{*} - e_{ij}^{*}\right|}{e_{ij}^{*}} = \sum_{i=1}^{w}\sum_{j=1}^{k}\frac{\left|p_{ij}^{*} - p_{i\bullet}^{*}p_{\bullet j}^{*}\right|}{p_{i\bullet}^{*}p_{\bullet j}^{*}}, \tag{8}$$

where n_{ij}^{*} are experimental counts, e_{ij}^{*} are expected counts and p_{ij}^{*} are empirical probabilities. It is shown in (Sulewski, 2016) that $|\chi|$ test is more powerful than tests (5)–(7).

4.2. Three-way contingency table

In this paper the research has been limited only to complete independence. Features X, Y, Z are completely independent from one another, and H_0 is true, if

$$P_{ijt} = P_{i\bullet\bullet} P_{\bullet j\bullet} P_{\bullet\bullet t} \tag{9}$$

for each $i = 1, \dots, w; j = 1, \dots, k; t = 1, \dots, p$. The alternative hypothesis H_1 negates H_0.

Let e_{ijt} be the expected counts under complete independence of i, j, t

$$e_{ijt} = \frac{n_{i\bullet\bullet} n_{\bullet j\bullet} n_{\bullet\bullet t}}{n^2} = np_{i\bullet\bullet} p_{\bullet j\bullet} p_{\bullet\bullet t} (i = 1, \dots, w; j = 1, \dots, k; t = 1, \dots, p). \tag{10}$$

The expected counts e_{ijt} under complete independence of i, j, t have the same one-way

marginal values as the observed table n_{ijt} (Gokhale, Kullback, 1978)

To study the complete independence of the features X, Y, Z we use the statistics that are extensions of those for two-way CTs (Pardo, 1996)

$$\chi_3^2 = \sum_{i=1}^{w} \sum_{j=1}^{k} \sum_{t=1}^{p} \frac{\left(n_{ijt}^* - e_{ijt}^*\right)^2}{e_{ijt}^*}, \tag{11}$$

$$FT_3 = 4 \sum_{i=1}^{w} \sum_{j=1}^{k} \sum_{t=1}^{p} \left(\sqrt{n_{ijt}^*} - \sqrt{e_{ijt}^*}\right)^2, \tag{12}$$

$$CR_3 = \frac{9}{5} \sum_{i=1}^{w} \sum_{j=1}^{k} \sum_{t=1}^{p} n_{ijt}^* \left[\left(\frac{n_{ijt}^*}{e_{ijt}^*}\right)^{2/3} - 1\right]. \tag{13}$$

Statistics (11)–(13) for CT, when H_0 is true, asymptotically follow the chi-square distribution with $wkp - (w + k + p) + 2$ degrees of freedom. Statistics

$$G_3^2 = 2 \sum_{i=1}^{w} \sum_{j=1}^{k} \sum_{t=1}^{p} n_{ijt}^* \ln\left(\frac{n_{ijt}^*}{e_{ijt}^*}\right), N_3 = \sum_{i=1}^{w} \sum_{j=1}^{k} \sum_{t=1}^{p} \frac{\left(n_{ijt}^* - e_{ijt}^*\right)^2}{n_{ijt}^*},$$

$$KL_3 = 2 \sum_{i=1}^{w} \sum_{j=1}^{k} \sum_{t=1}^{p} e_{ijt}^* \ln\left(\frac{e_{ijt}^*}{n_{ijt}^*}\right)$$

also belong to the PDS. However, these statistics have not been applied in the Monte Carlo study, because they do not take into account the condition $n_{ijt}^* = 0$.

The $|\chi_3|$ statistics is selected to Monte Carlo study, too. It is an authorial modification of χ_3^2 statistics and it has the form (Sulewski, 2018a)

$$|\chi_3| = \sum_{i=1}^{w}\sum_{j=1}^{k}\sum_{t=1}^{p}\frac{|n_{ijt}^* - e_{ijt}^*|}{e_{ijt}^*}, \tag{14}$$

where n_{ijt}^* are experimental counts and e_{ijt}^* are expected counts. It is shown in (Sulewski, 2018a) that $|\chi_3|$ test is more powerful than tests (11)–(13).

5. MEASURES OF UNTRUTHFULNESS OF H₀

5.1. Two-way contingency table

When the equality $p_{ij} = p_{i\bullet}p_{\bullet j}$ is not fulfilled, H_0 is not true and an appropriate measure of untruthfulness of H_0 (MoUH) is needed. There are many different measures in literature, e.g.: the Pearson's φ, the Tschuprow's T, the Cramer's V, the corrected contingency c, the Goodman and Kruskal's τ.

In this paper we use a MoUH which is given by (Sulewski, 2016):

$$MoU = \frac{1}{n}\sum_{i=1}^{w}\sum_{j=1}^{k}\left|n_{ij}^* - \frac{n_{i\bullet}^* \cdot n_{\bullet j}^*}{n}\right| = \sum_{i=1}^{w}\sum_{j=1}^{k}\left|p_{ij}^* - p_{i\bullet}^* \cdot p_{\bullet j}^*\right|. \tag{15}$$

The MoU takes values in interval $\langle 0,1\rangle$. This measure, doubtlessly, springs from the essence of H_0 and has a very simple form. The MoU formulas and the maximal MoU values (the minimal MoU values are equal to zero) under scenarios I-XX are presented in table 4. The MoU is a function of the PFP a. Owing to this the MoU values are very easy to calculate.

Table 4. THE MoU UNDER SCENARIOS I-XX FOR A TWO-WAY CT

Table	Scenario	MoU	MoU_{max}	Table	Scenario	MoU	MoU_{max}
2×3	I	$4a/3$	0.2222	2×4	V	$2a$	0.25
	II	*	0.3333		VI	**	0.3125
	III	$8a/3$	0.4444		VII	$3a$	0.375
	IV	$4a$	0.6667		VIII	$4a$	0.5
3×3	IX	$2a^2 + 2a$	0.2469	3×4	XIII	$8a/3$	0.2222
	X	$8a/3$	0.2963		XIV	$4a$	0.3333
	XI	$4a$	0.4444		XV	$16a/3$	0.4444
	XII	$16a/3$	0.5926		XVI	$8a$	0.6667

Table 4. THE MoU UNDER SCENARIOS I-XX FOR A TWO-WAY CT (cont.)

Table	Scenario	MoU	MoU_{max}	Table	Scenario	MoU	MoU_{max}
4×4	XVII	$3a$	0.1875				
	XVIII	$4a$	0.25				
	XIX	***	0.3125				
	XX	$8a$	0.5				

* $4a/3\, I(0 \le a < 0.1), 8a^2 + 2a/3\, I(0.1 < a \le 1/6)$
** $2a\, I(0 \le a \le 0.075), 12a^2 + a\, I(0.075 < a \le 0.125)$
*** $4a\, I(0 \le a \le 0.0375), 48a^2 + 2a\, I(0.0375 < a \le 1/16)$

Source: own elaboration.

5.2. Three-way contingency table

The theory devoted to MoUH for TT is not as rich as for the two-way contingency table, where the Goodman—Kruskal index plays an important role (Goodman, Kruskal, 1954). Numeric extensions of this index for three way CT are: the Marcotorchino index τ_M (Marcotorchino, 1984), the delta index τ_L (Lombardo, 2011) and the Gray—Williams index τ_{GW} (Gray, Williams, 1975). Information about other less popular indices can be found in (Beh, Davy, 1998; Harshman, 1970; Lombardo, Beh, 2010; Trucker, 1963).

Based on the classical definition of independence of X, Y, Z, the MoU_3 in the form

$$MoU_3 = \frac{1}{n}\sum_{i=1}^{w}\sum_{j=1}^{k}\sum_{t=1}^{p}\left|n_{ijt}^{*} - \frac{n_{i\bullet\bullet}^{*} n_{\bullet j\bullet}^{*} n_{\bullet\bullet t}^{*}}{n^2}\right| = \sum_{i=1}^{w}\sum_{j=1}^{k}\sum_{t=1}^{p}\left|p_{ijt} - p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet t}\right|, \quad (16)$$

is put forward in (Sulewski, 2018a). The measure (16) takes the value 0 when H_0 is true. The higher the MoU_3 value, the greater the possibility of H_0 falsity. More information about the measures MoU_3, τ_M, τ_L and τ_{GW} defined under some scenarios for three-way CT you can find in (Sulewski, 2018a).

The MoU_3 as a natural measure, resulting from the definition of independence, is used in the Monte Carlo simulation. The MoU_3 formulas and the maximal MoU_3 values (the minimal MoU_3 values are equal to zero) under scenarios XXI–XXVIII are presented in table 5. The MoU_3 is a function of the PFP *a*. Owing to this the MoU_3 values are very easy to calculate.

Table 5. The MoU_3 under scenarios XXI-XXVIII for three-way contingency tables

Table	Scenario	MoU_3	$MoU_{3\,max}$	Table	Scenario	MoU_3	$MoU_{3\,max}$
$2 \times 2 \times 2$	XXI	$16a^2$	0.25	$3 \times 2 \times 2$	XXV	$8a/3$	0.2222
	XXII	*	0.3359		XXVI	**	0.3889
	XXIII	$3a$	0.375		XXVII	$6a$	0.5
	XXIV	$4a$	0.5		XXVIII	$8a$	0.6667

*

a	0	0.0125	0.025	0.0375	0.05	0.0625	0.075	0.0875	0.1	0.1125	0.125
MoU_3	0	0.025	0.0504	0.0763	0.103	0.1309	0.1601	0.1969	0.238	0.283	0.3359

**

a	0	0.0083	0.0167	0.025	0.0333	0.0417	0.05	0.0583	0.0667	0.075	0.0833
MoU_3	0	0.0364	0.0733	0.1108	0.1489	0.1875	0.2267	0.2664	0.3067	0.3475	0.3889

Source: own elaboration.

6. APPLYING THE MAXIMUM LIKELIHOOD METHOD TO ESTIMATE THE PROBABILITY FLOW PARAMETER

This section is a simply attempt of replacing a nonparametric statistical inference method by the parametric one. Maximum likelihood method is applied to estimate the PFP.

Let us remember that in cells of two-way CT are values n_{ij} $(i = 1, \dots, w; j = 1, \dots, k)$, in cells of three-way CT are values n_{ijt} $(i = 1, \dots, w; j = 1, \dots, k; t = 1, \dots, p)$. These values are components of the multinomial distribution. Thus the multinomial distribution was taken as a groundwork of the likelihood functions. A family of these likelihood functions is given below. Every function from this family has an index. Indices assign functions to particular scenarios presented in section 3 of the paper.

6.1. Two-way contingency table

Let n^*_{ij} be the value of (i, j) cell and a is the probability flow parameter. Then likelihood functions under CT $w \times k$ have the form

a) table 2×3

$$L_I(a) = C(1/6 - a)^{n^*_{11}}(1/6 + a)^{n^*_{23}}(1/6)^{n^*_{12}+n^*_{13}+n^*_{21}+n^*_{22}}, \qquad (17)$$

$$L_{II}(a) = C(1/6 - a)^{n^*_{11}+n^*_{12}}(1/6 + a)^{n^*_{22}+n^*_{23}}(1/6)^{n^*_{13}+n^*_{21}}, \qquad (18)$$

$$L_{III}(a) = C(1/6 - a)^{n^*_{11}+n^*_{12}}(1/6 + a)^{n^*_{22}+n^*_{22}}(1/6)^{n^*_{13}+n^*_{23}}, \qquad (19)$$

$$L_{IV}(a) = C(1/6 - a)^{n^*_{11}+n^*_{23}}(1/6 + a)^{n^*_{13}+n^*_{21}}(1/6)^{n^*_{12}+n^*_{22}}, \qquad (20)$$

b) table 2×4

$$L_V(a) = C(1/8 - a)^{n^*_{11}}(1/8 + a)^{n^*_{24}}(1/8)^{n^*_{12}+n^*_{13}+n^*_{14}+n^*_{21}+n^*_{22}+n^*_{23}}, \qquad (21)$$

$$L_{VI}(a) = C(1/8 - a)^{n^*_{11}+n^*_{12}+n^*_{13}}(1/8 + a)^{n^*_{22}+n^*_{23}+n^*_{24}}(1/8)^{n^*_{14}+n^*_{21}}, \qquad (22)$$

$$L_{VII}(a) = C(1/8 - a)^{n^*_{12}}(1/8 + a)^{n^*_{21}}(1/8)^{n-n^*_{11}-n^*_{21}}, \qquad (23)$$

$$L_{VIII}(a) = C(1/8 - a)^{n^*_{11}+n^*_{12}}(1/8 + a)^{n^*_{21}+n^*_{22}}(1/8)^{n^*_{13}+n^*_{14}+n^*_{23}+n^*_{24}}, \qquad (24)$$

c) table 3×3

$$L_{IX}(a) = C(1/9 - a)^{n_{11}^*}(1/9 + a)^{n_{33}^*}(1/9)^{n - n_{11}^* - n_{33}^*}, \tag{25}$$

$$L_{X}(a) = C(1/9 - a)^{n_{11}^* + n_{12}^*}(1/9 + a)^{n_{31}^* + n_{31}^*}(1/9)^{n_{13}^* + n_{21}^* + n_{22}^* + n_{23}^* + n_{33}^*}, \tag{26}$$

$$L_{XI}(a) = C(1/9 - a)^{n_{11}^* + n_{33}^*}(1/9 + a)^{n_{11}^* + n_{33}^*}(1/9)^{n_{12}^* + n_{21}^* + n_{22}^* + n_{23}^* + n_{32}^*}, \tag{27}$$

$$L_{XII}(a) = C(1/9 - a)^{n_{11}^* + n_{21}^* + n_{33}^*}(1/9 + a)^{n_{11}^* + n_{23}^* + n_{33}^*}(1/9)^{n_{12}^* + n_{22}^* + n_{32}^*}, \tag{28}$$

d) table 3×4

$$L_{XIII}(a) = C(1/12 - a)^{n_{11}^*}(1/12 + a)^{n_{34}^*}(1/12)^{n - n_{11}^* - n_{34}^*}, \tag{29}$$

$$L_{XIV}(a) = C(1/12 - a)^{n_{11}^* + n_{12}^*}(1/12 + a)^{n_{31}^* + n_{32}^*}(1/12)^{n - n_{11}^* - n_{12}^* - n_{31}^* - n_{32}^*}, \tag{30}$$

$$L_{XV}(a) = C(1/12 - a)^{n_{11}^* + n_{12}^*}(1/12 + a)^{n_{13}^* + n_{14}^*}(1/12)^{n - n_{11}^* - n_{12}^* - n_{13}^* - n_{14}^*}, \tag{31}$$

$$\begin{gathered} L_{XVI}(a) = C(1/12 - a)^{n_{11}^* + n_{33}^* + n_{34}^*}(1/12 + a)^{n_{13}^* + n_{14}^* + n_{31}^* + n_{32}^*} \\ (1/12)^{n_{21}^* + n_{22}^* + n_{23}^* + n_{24}^*}, \end{gathered} \tag{32}$$

e) table 4×4

$$\begin{gathered} L_{XVII}(a) = C(1/16 - a)^{n_{11}^* + n_{12}^* + n_{13}^*}(1/16 + a)^{n_{41}^* + n_{42}^* + n_{43}^*} \\ (1/16)^{n - n_{11}^* - n_{12}^* - n_{13}^* - n_{41}^* - n_{42}^* - n_{43}^*}, \end{gathered} \tag{33}$$

$$L_{XVIII}(a) = C(1/16 - a)^{n_{11}^* + n_{12}^*}(1/16 + a)^{n_{43}^* + n_{44}^*}(1/16)^{n - n_{11}^* - n_{12}^* - n_{43}^* - n_{44}^*}, \tag{34}$$

$$\begin{gathered} L_{XIX}(a) = C(1/16 - a)^{n_{11}^* + n_{12}^* + n_{13}^* + n_{21}^* + n_{22}^* + n_{23}^*} \\ (1/16 + a)^{n_{32}^* + n_{33}^* + n_{34}^* + n_{42}^* + n_{43}^* + n_{44}^*}(1/16)^{n - n_{14}^* - n_{24}^* - n_{31}^* - n_{41}^*}, \end{gathered} \tag{35}$$

$$\begin{gathered} L_{XX}(a) = C(1/16 - a)^{n_{11}^* + n_{12}^* + n_{43}^* + n_{44}^*} \\ (1/16 + a)^{n_{13}^* + n_{14}^* + n_{41}^* + n_{42}^*}(1/16)^{n_{21}^* + n_{22}^* + n_{23}^* + n_{24}^* + n_{31}^* + n_{32}^* + n_{33}^* + n_{34}^*}, \end{gathered} \tag{36}$$

In formulas (17)–(36) $C = n!/\prod_{i=1}^{w}\prod_{j=1}^{k} n_{ij}^*!$.

The logarithmic likelihood function under scenario I in CT 2×3 is given by

$$l_I(a) = 1\text{n}L_I(a) = 1\text{n}(\text{C}) + n_{11}^* 1\text{n}(1/6 + a) + (n_{12}^* + n_{13}^* n_{21}^* n_{22}^*)1\text{n}(1/6),$$

then

$$\frac{\partial l_I(a)}{\partial a^2} = \frac{n_{23}^*}{1/6 + a} - \frac{\partial l_I(a)}{\partial a} = 0 \Rightarrow \hat{a} = \frac{n_{23}^* - n_{11}^*}{6(n_{23}^* + n_{11}^*)},$$

Let us check what kind of extremum we can found. As a result of a simple transformation we have

$$\frac{\partial l_I(a)}{\partial a^2}=\frac{\partial}{\partial_a}\left[\frac{n_{23}^*}{1/6+a}-\frac{n_{11}^*}{1/6-a}\right]=-\frac{n_{23}^*}{(1/6+a)}-\frac{n_{11}^*}{(1/6-a)^2}<0, \quad (37)$$

for $a<1/6$. It means that the logarithmic likelihood function has always a maximum at $a=\hat{a}$. So, $\hat{a}$ is the maximum likelihood estimator of a which is the probability flow parameter. It may be proven that inequality (37) holds for all scenarios considered in this paper.

Formulas for the maximum likelihood estimator of a under scenarios I-XX for two-way CT are given by:

a) table 2×3

$$\hat{a}_I=\frac{n_{23}^*-n_{11}^*}{6(n_{23}^*+n_{11}^*)},\hat{a}_{II}=\frac{n_{22}^*+n_{23}^*-(n_{11}^*+n_{12}^*)}{6(n_{11}^*+n_{12}^*+n_{22}^*+n_{23}^*)},$$
$$\hat{a}_{III}=\frac{n_{21}^*+n_{22}^*-(n_{11}^*+n_{12}^*)}{6(n_{11}^*+n_{12}^*+n_{21}^*+n_{22}^*)},\hat{a}_{IV}=\frac{n_{13}^*+n_{21}^*-(n_{11}^*+n_{23}^*)}{6(n_{11}^*+n_{12}^*+n_{21}^*+n_{23}^*)}, \quad (38)$$

b) table 2×4

$$\hat{a}_V=\frac{n_{24}^*-n_{11}^*}{8(n_{24}^*+n_{11}^*)},\hat{a}_{VI}=\frac{n_{22}^*+n_{23}^*+n_{24}^*-(n_{11}^*+n_{12}^*+n_{13}^*)}{8(n-n_{14}^*-n_{21}^*)},$$
$$\hat{a}_{VII}=\frac{n_{21}^*-n_{11}^*}{8(n_{11}^*+n_{21}^*)},\hat{a}_{VIII}=\frac{n_{21}^*+n_{22}^*-(n_{11}^*+n_{12}^*)}{8(n_{11}^*+n_{12}^*+n_{21}^*+n_{22}^*)}, \quad (39)$$

c) table 3×3

$$\hat{a}_{XI}=\frac{n_{33}^*-n_{11}^*}{9(n_{11}^*+n_{33}^*)},\hat{a}_X=\frac{n_{31}^*+n_{32}^*-(n_{11}^*+n_{12}^*)}{9(n_{11}^*+n_{12}^*+n_{31}^*+n_{32}^*)},$$
$$\hat{a}_{XI}=\frac{n_{13}^*+n_{31}^*-(n_{11}^*+n_{33}^*)}{9(n_{11}^*+n_{13}^*+n_{31}^*+n_{33}^*)},\hat{a}_{XII}=\frac{n_{13}^*+n_{23}^*+n_{31}^*-(n_{11}^*+n_{21}^*+n_{33}^*)}{9(n-n_{12}^*-n_{22}^*-n_{32}^*)}, \quad (40)$$

d) table 3×4

$$\hat{a}_{XIII}=\frac{n_{34}^*-n_{11}^*}{12(n_{11}^*+n_{34}^*)},\hat{a}_{XIV}=\frac{n_{31}^*+n_{32}^*-(n_{11}^*+n_{12}^*)}{12(n_{11}^*+n_{12}^*+n_{31}^*+n_{32}^*)},$$
$$\hat{a}_{XV}=\frac{n_{13}^*+n_{14}^*-(n_{11}^*+n_{12}^*)}{12(n_{11}^*+n_{12}^*+n_{13}^*+n_{14}^*)}, \quad (41)$$
$$\hat{a}_{XVI}=\frac{n_{13}^*+n_{14}^*+n_{31}^*+n_{32}^*-(n_{11}^*+n_{12}^*+n_{33}^*+n_{34}^*)}{12(n-n_{21}^*-n_{22}^*-n_{23}^*-n_{24}^*)},$$

e) table 4×4

$$\hat{a}_{XVII} = \frac{n_{41}^* + n_{42}^* + n_{43}^* - (n_{11}^* + n_{12}^* + n_{13}^*)}{16(n_{11}^* + n_{12}^* + n_{13}^* + n_{41}^* + n_{42}^* + n_{43}^*)},$$

$$\hat{a}_{XVIII} = \frac{n_{43}^* + n_{44}^* - (n_{11}^* + n_{12}^*)}{16(n_{11}^* + n_{12}^* + n_{43}^* + n_{44}^*)},$$

$$\hat{a}_{XIX} = \frac{n_{32}^* + n_{33}^* + n_{34}^* + n_{42}^* + n_{43}^* + n_{44}^* - (n_{11}^* + n_{12}^* + n_{13}^* + n_{21}^* + n_{22}^* + n_{23}^*)}{16(n - n_{14}^* + n_{24}^* + n_{31}^* + n_{41}^*)}, \quad (42)$$

$$\hat{a}_{XX} = \frac{n_{13}^* + n_{14}^* + n_{41}^* + n_{42}^* - (n_{11}^* + n_{12}^* + n_{31}^* + n_{44}^*)}{16(n_{11}^* + n_{12}^* + n_{13}^* + n_{14}^* + n_{41}^* + n_{42}^* + n_{43}^* + n_{44}^*)}.$$

To decide which of the defined scenarios takes place, you use the following algorithm:

1. Find dimension of CT in question and read out related q according to the rule: $2 \times 3\ (q = 1), 2 \times 4\ (q = 2), 3 \times 3\ (q = 3), 3 \times 4\ (q = 4), 4 \times 4 (q = 5)$
2. Set a set of scenario indices $\{4q - 3, 4q - 2, 4q - 1, 4q\}$.
3. Calculate a^*, which is an estimate of parameter *a* for each scenario from step 2 by means of (38)–(42).
4. Calculate corresponding values of the maximum likelihood functions $L_{4q-3}(a^*), L_{4q-2}(a^*), L_{4q-1}(a^*), L_{4q}(a^*)$ by means of (17)–(36).
5. Choose a scenario for which the value $L(a^*)$ is the greatest.

This algorithm will be used in section 8.1, Example 1.

6.2. Three-way contingency table

Let n_{ijt}^* be the value of (i, j, t) cell and a is the PFP. Then likelihood functions for selected three-way CT have the form:

a) table $2 \times 2 \times 2$

$$L_{XXI}(a) = D(1/8 - a)^{n_{111}^* + n_{112}^*}(1/8 + a)^{n_{221}^* + n_{222}^*}(1/8)^{n_{221}^* + n_{211}^* n_{122}^* + n_{212}^*}, \quad (43)$$

$$L_{XXII}(a) = D(1/8 - a)^{n_{111}^* + n_{121}^* n_{122}^*}(1/8 + a)^{n_{221}^* + n_{212}^* + n_{222}^*}(1/8)^{n_{122}^* + n_{211}^*}, \quad (44)$$

$$L_{XXIII}(a) = D(1/8 - a)^{n_{111}^*}(1/8 + a)^{n_{211}^*}(1/8)^{n - n_{111}^* + n_{211}^*}, \quad (45)$$

$$L_{XXIV}(a) = D(1/8 - a)^{n_{111}^* + n_{112}^*}(1/8 + a)^{n_{211}^* + n_{212}^*}(1/8)^{n_{121}^* + n_{221}^* + n_{122}^* + n_{222}^*}, \quad (46)$$

b) table $3 \times 2 \times 2$

$$L_{XXV}(a) = D(1/8 - a)^{n_{111}^* + n_{112}^*}(1/8 + a)^{n_{321}^* + n_{322}^*}(1/8)^{n - n_{111}^* - n_{112}^* - n_{321}^* - n_{322}^*}, \quad (47)$$

$$L_{XXVI}(a) = D(1/8 - a)^{n_{111}^* + n_{112}^*}(1/8 + a)^{n_{122}^* + n_{321}^*}(1/8)^{n - n_{111}^* - n_{112}^* - n_{112}^* - 321}, \quad (48)$$

$$L_{XXVII}(a) = D(1/8 - a)^{n_{111}^* + n_{112}^* + n_{221}^*}(1/8 + a)^{n_{121}^* + n_{122}^* + n_{211}^*} (1/8)^{n_{212}^* + n_{222}^* + n_{311}^* + n_{321}^* + n_{312}^* + n_{222}^* + n_{322}^*}, \quad (49)$$

$$L_{XXVIII}(a) = D(1/8 - a)^{n^*_{111}+n^*_{112}+n^*_{321}+n^*_{322}}(1/8 + a)^{n^*_{121}+n^*_{122}+n^*_{311}+n^*_{312}} (1/8)^{n^*_{211}+n^*_{221}+n^*_{212}+n^*_{221}}. \quad (50)$$

In formulas (43)–(50) $D = n!/\prod_{i=1}^{w}\prod_{j=1}^{k}\prod_{t=1}^{p} n^*_{ijt}!$.

Formulas for the maximum likelihood estimator of a under scenarios XXIV–XXVIII for CT $w \times k \times p$ are given by

a) table $2 \times 2 \times 2$

$$\hat{a}_{XXI} = \frac{n^*_{221} + n^*_{222} - (n^*_{111} + n^*_{112})}{8(n^*_{111} + n^*_{112} + n^*_{221} + n^*_{222})},$$
$$\hat{a}_{XXII} = \frac{n^*_{221} + n^*_{212} + n^*_{222} - (n^*_{111} + n^*_{121} + n^*_{122})}{8(n - n^*_{122} - n^*_{211})}, \quad (51)$$
$$\hat{a}_{XXIII} = \frac{n^*_{211} - n^*_{111}}{8(n^*_{111} + n^*_{211})}, \hat{a}_{XXIV} = \frac{n^*_{211} + n^*_{212} - (n^*_{111} + n^*_{112})}{8(n^*_{111} + n^*_{112} + n^*_{211} + n^*_{212})}.$$

a) table $3 \times 2 \times 2$

$$\hat{a}_{XXV} = \frac{n^*_{321} + n^*_{322} - (n^*_{111} + n^*_{112})}{12(n^*_{111} + n^*_{112} + n^*_{321} + n^*_{322})},$$
$$\hat{a}_{XXVI} = \frac{n^*_{122} + n^*_{321} - (n^*_{111} + n^*_{112})}{12(n^*_{111} + n^*_{112} + n^*_{122} + n^*_{321})},$$
$$\hat{a}_{XXVII} = \frac{n^*_{121} + n^*_{122} + n^*_{211} - (n^*_{111} + n^*_{112} + n^*_{221})}{8(n^*_{111} + n^*_{112} + n^*_{211} + n^*_{212})} \quad (52)$$
$$\hat{a}_{XXVIII} = \frac{n^*_{121} + n^*_{112} + n^*_{311} + n^*_{312} - (n^*_{111} + n^*_{112} + n^*_{321} + n^*_{322})}{12(n - n^*_{211} - n^*_{221} - n^*_{212} - n^*_{222})}.$$

To decide which of the defined scenarios takes place, you use the following algorithm:

1. Find dimension of CT in question and read out related q according to the rule:

$$2 \times 2 \times 2\ (q = 6), 3 \times 2 \times 2\ (q = 7)$$

2. Set a set of scenario indices$\{4q - 3, 4q - 2, 4q - 1, 4q\}$.
3. Calculate a^*, which is an estimate of parameter a for each scenario from step 2 according to (51)–(52)
4. Calculate corresponding values of the maximum likelihood functions $L_{4q-3}(a^*), L_{4q-2}(a^*), L_{4q-1}(a^*), L_{4q}(a^*)$ by means of (43)–(50).
5. Choose a scenario for which the value $L(a^*)$ is the greatest.

This algorithm will be used in section 8.2, Example 4.

7. GENERATING CONTINGENCY TABLE

Generating CT is very important in the simulation study. The approach in the literature for the generating two-way CTs is the Markov Chain Monte Carlo (Diaconis, Sturmfels, 1998; Cryan, Dyer, 2003; Cryan et al., 2006; Chen et al., 2005; Fishman, 2012), the Sequential Importance Sampling (Chen et al., 2005; Chen et al., 2006; Blitzstein, Diaconis, 2011; Yoshida, 2011), the probabilistic divide-and-conquer technique (DeSalvo, Zhao, 2015), the Generalized Gamma Distribution (Sulewski, 2009), the bar method (Sulewski, Motyka, 2015). The bar method for the generating three-way CT is presented in (Sulewski, 2018a).

In this paper we use an algorithm for generating two-way and three-way CTs using the bar method. The bar method is identical to the method that generates random numbers that follow the multinomial distribution.

8. PARAMETRIC REASONING PUT INTO PRACTICE

8.1. Two-way contingency table

Example 1. This example compares decision making by means of classic statistic testing with likelihood based decision making. Tables 6-10 show a set of example two-way CTs filled one by one accordingly to the scenarios I–XX. The table is divided into two parts. The left hand side is related to likelihood based decisions. To decide which of the defined scenarios takes place, see algorithm in section 6.1. The right hand side is related to classic statistical testing and presents values of test statistics. The H_0 states that X and Y are independent. Critical values, indicated by underlining, are determined by Monte Carlo method based on 10^6 order statistics. Such a large number of repetitions guaranties very precise results. When reading rows of the table it turns out that all the decisions made in a classic way are wrong. It is because untrue H_0 hypotheses have not been rejected. But it does not reveal anything new. This is just one more confirmation of what is commonly known – the classic statistical test is very conservative. The PFP a is a maximal value of this parameter for which untrue H_0 is rejected in no scenarios.

Tables 6–10 show that all the decisions made in a classic way are wrong under the scenarios in question. It is because untrue H_0 hypotheses have not been rejected even in situations where the MoU does not have such small values, e.g. $MoU = 0.3$ in scenario XX. In turn, the parametric approach detected a dependency between features.

Example 2. The conservativeness of classic testing is a reason why we suggest making a turnaround in this domain. The new proposal is a method of statistical inference and not a classical parametric test. Now there will be no null hypothesis, but there will be a set of competing alternative hypotheses instead. The former H_0 is no longer the main one, but exists among the competitors of an equal rank. All the hypotheses state: “the considered two-way CT is generated accordingly to particular scenarios”. Each figure from 1 to 5 shows sets of four likelihood curves. Each curve has its four attributes, namely $\hat{a}$, n, table dimensions and, of course, the actual generation scenario that is specified in the figure's title. The legend lists all competing scenarios including the actual one.

Table 6. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE 2×3

Experiment settings: $n = 60, a = 0.05, \alpha = 0.05$

Scenario (MoU)	Content			X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	$\|\chi\|$	H_0 true	H_0 rejected
						6.041	6.422	6.083	1.767		
I(0.067)	7	10	10	Yes	I	0.324	0.324	0.324	0.418	No	No
	10	10	10								
II(0.067)	7	7	10	Yes	II	0.334	0.338	0.335	0.402	No	No
	10	13	13								
III(0.133)	7	7	10	Yes	III	1.250	1.237	1.246	0.833	No	No
	13	13	10								
IV(0.200)	7	10	13	Yes	IV	3.600	3.706	3.612	1.200	No	No
	13	10	7								

Source: own elaboration.

Table 7. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE 2×4

Experiment settings: $n = 80, a = 0.0625, \alpha = 0.05$

Scenario (MoU)	Content				X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	$\|\chi\|$	H_0 true	H_0 rejected
							<u>8.058</u>	<u>8.657</u>	<u>8.131</u>	<u>2.283</u>		
V(0.125)	5	10	10	10	Yes	V	1.439	1.462	1.443	1.084	No	No
	10	10	10	15								
VI(0.125)	5	5	5	10	Yes	VI	1.648	1.642	1.645	1.086	No	No
	10	15	15	15								
VII(0.188)	5	10	10	10	Yes	VII	3.810	4.104	3.855	1.524	No	No
	15	10	10	10								
VIII(0.250)	5	5	10	10	Yes	VIII	5.333	5.482	5.350	2.133	No	No
	15	15	10	10								

Source: own elaboration.

Table 8. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE 3×3

Experiment settings: $n = 90, a = 0.0444, \alpha = 0.05$

Scenario (MoU)	Content			X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	$\|\chi\|$	H_0 true	H_0 rejected
						9.416	10.067	9.458	2.544		
IX(0.093)	6	10	10	Yes	XI	1.138	1.160	1.142	0.870	No	No
	10	10	10								
	10	10	14								
X(0.119)	6	6	10	Yes	X	2.297	2.257	2.283	1.148	No	No
	10	10	10								
	14	14	10								
XI(0.178)	6	10	14	Yes	XI	6.400	6.750	6.440	1.600	No	No
	10	10	10								
	14	10	6								
XII(0.237)	6	10	14	Yes	XII	8.688	8.991	8.690	2.172	No	No
	6	10	14								
	14	10	6								

Source: own elaboration.

Table 9. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE 3×4

Experiment settings: $n = 80, a = 0.0625, \alpha = 0.05$

Scenario (MoU)	Content				X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	$\|\chi\|$	H_0 true	H_0 rejected
							12.565	13.463	12.638	3.325		
XIII(0.067)	7	10	10	10	Yes	XIII	0.770	0.781	0.772	0.813	No	No
	10	10	10	10								
	10	10	10	10								
XIV(0.100)	7	7	10	10	Yes	XIV	1.841	1.856	1.843	1.228	No	No
	10	10	10	10								
	13	13	10	10								
XV(0.133)	7	7	13	13	Yes	XV	2.424	2.474	2.433	1.616	No	No
	10	10	10	10								
	10	10	10	10								
XVI(0.200)	7	7	13	13	Yes	XVI	7.200	7.413	7.225	2.400	No	No
	10	10	10	10								
	13	13	7	7								

Source: own elaboration.

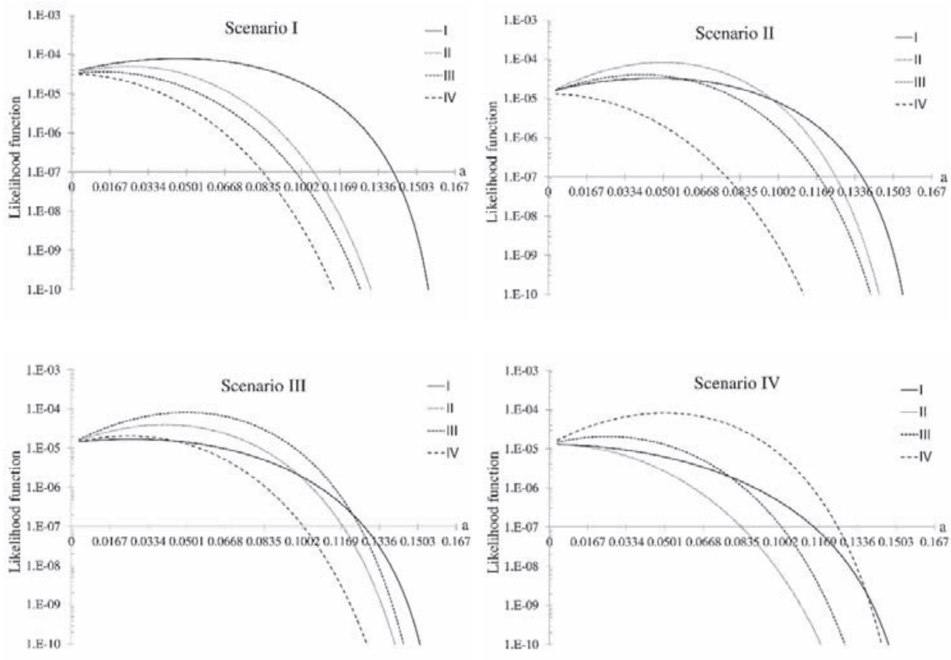
Table 10. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE 4×4

Experiment settings: $n = 80, a = 0.0375, \alpha = 0.05$

Scenario (MoU)	Content				X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	$\|\chi\|$	H_0 true	H_0 rejected
							16.798	22.675	16.965	6.370		
XVII(0.113)	2	2	2	5	Yes	XVII	3.386	3.105	3.297	2.257	No	No
	5	5	5	5								
	5	5	5	5								
	8	8	8	5								
XVIII(0.150)	2	2	5	5	Yes	XVIII	2.206	2.277	2.217	2.698	No	No
	5	5	5	5								
	5	5	5	5								
	5	5	8	8								
XIX(0.150)	2	2	2	5	Yes	XIX	2.704	2.677	2.690	2.712	No	No
	2	2	2	5								
	5	8	8	8								
	5	8	8	8								
XX(0.300)	2	2	8	8	Yes	XX	14.400	16.421	14.613	4.800	No	No
	5	5	5	5								
	5	5	5	5								
	8	8	2	2								

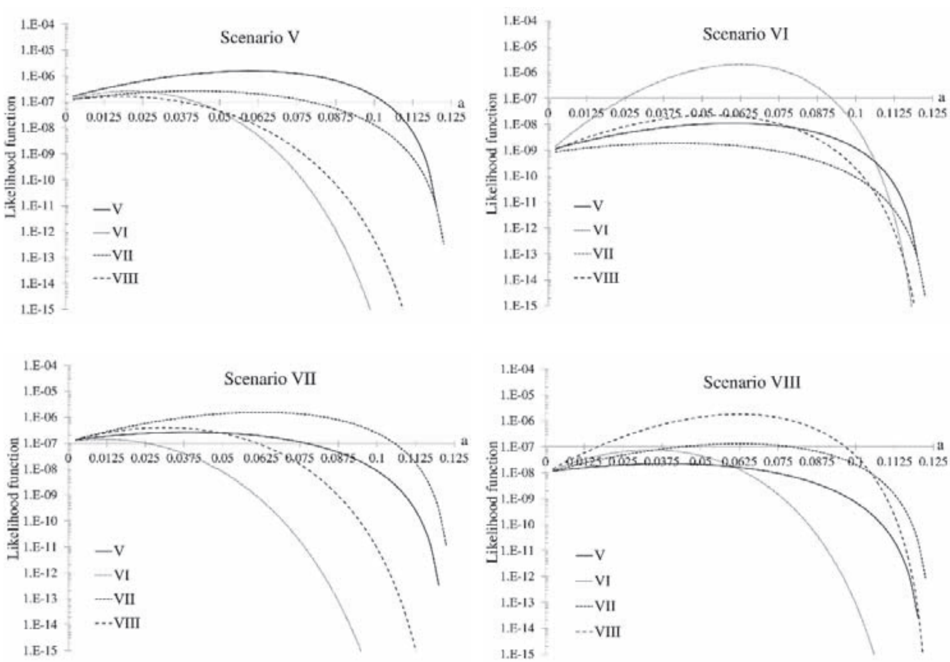
Source: own elaboration.

Figure 1. The likelihood functions for $\hat{\boldsymbol{a}} = \mathbf{0.05}, \boldsymbol{n} = \mathbf{60}$, table $\mathbf{2} \times \mathbf{3}$



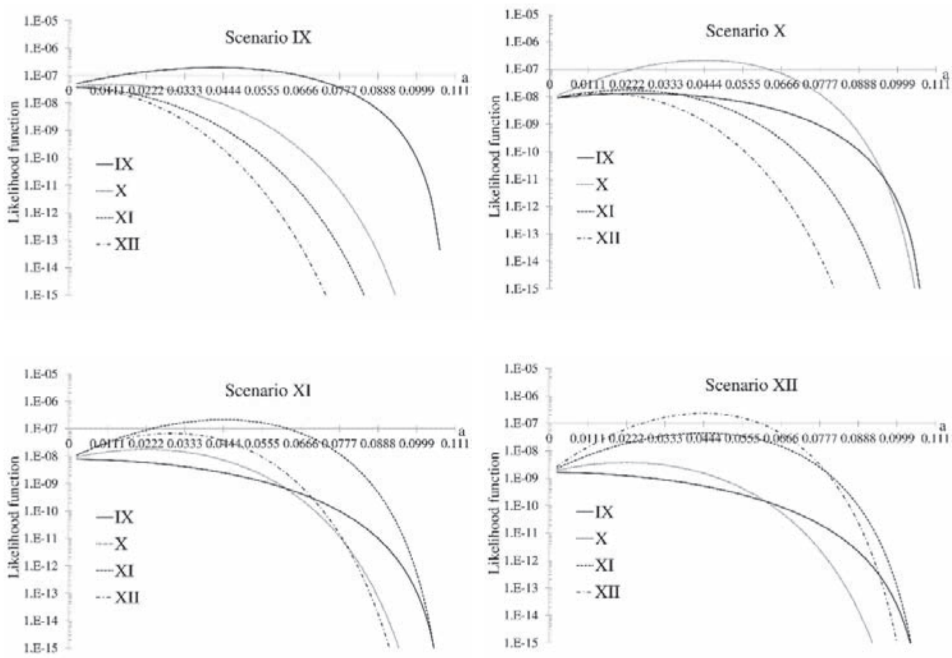
Source: own elaboration.

Figure 2. The likelihood functions for $\hat{a} = 0.0625, n = 80$, table 2×4



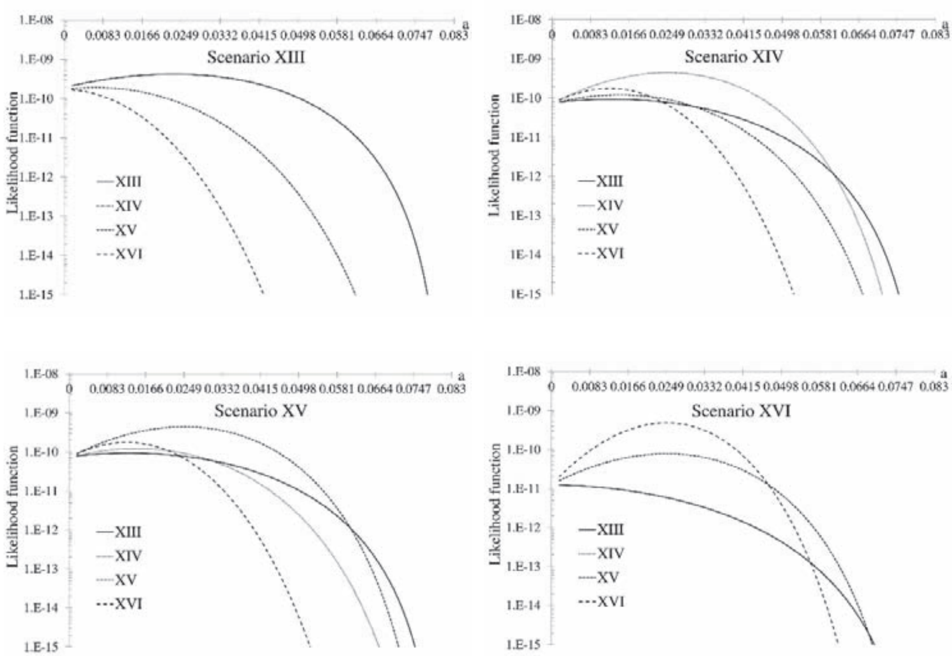
Source: own elaboration.

Figure 3. The likelihood functions for $\hat{a} = 0.0444, n = 90$, table 3×3



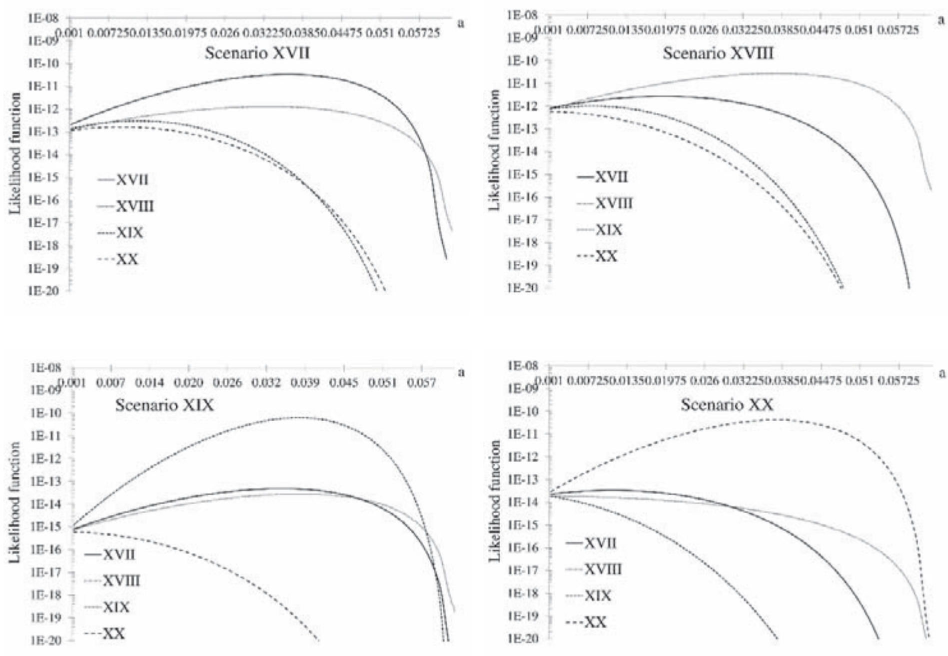
Source: own elaboration.

Figure 4. The likelihood functions for $\hat{a} = 0.0250, n = 120$, table 3×4



Source: own elaboration.

Figure 5. The likelihood functions for $\hat{a} = 0.0375, n = 80$, table 4×4



Source: own elaboration.

It is noteworthy (see figures 1–5) that in each set of likelihood curves, the curve related to the actual scenario predominates over the others. It is instructive to read out a value of a_{max} that maximizes likelihood function on a particular figure and notice that this value is close to assumed $\hat{a}$.

Example 3. This example is carried out in accordance with the following algorithm:

1. Find dimension of CT in question and read out related q according to the rule:

   $$2 \times 3\ (q = 1), 2 \times 4\ (q = 2), 3 \times 3\ (q = 3), 3 \times 4\ (q = 4), 4 \times 4 (q = 5)$$

2. Set a set of scenario indices $\{4q - 3, 4q - 2, 4q - 1, 4q\}$.
3. Calculate values of PFP by means of $a_i = 0.1i/(wk)$ $(i = 1,2, \dots ,10)$.
4. Set a sample size n.
5. Repeat the following steps $u = 10^4$ times:
   5.1. Let $Sc_{4q-3} = 0, Sc_{4q-2} = 0, Sc_{4q-1} = 0, Sc_{4q} = 0$
   5.2. Generate CT under the scenarios that you have chosen in Step 2.
   5.3. Calculate $L_{4q-3}(a), L_{4q-2}(a), L_{4q-1}(a), L_{4q}(a)$ by means of (17)–(36).
   5.4. If $MaxL = L_{4q-3}(a)$, then $Sc_{4q-3} = Sc_{4q-3} + 1$,
   If $MaxL = L_{4q-2}(a)$, then $Sc_{4q-2} = Sc_{4q-2} + 1$,

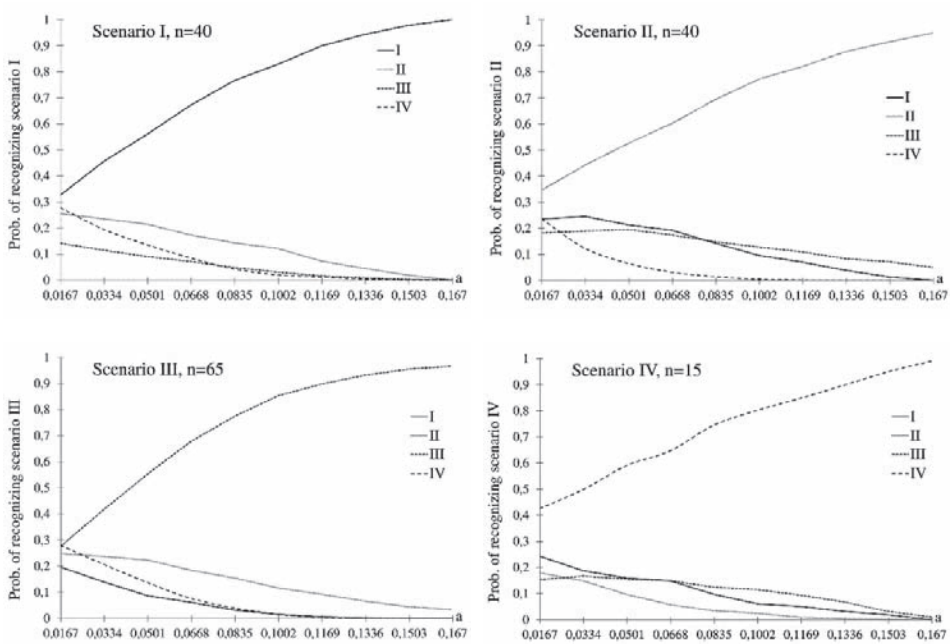
If $MaxL = L_{4q-1}(a)$, then $Sc_{4q-1} = Sc_{4q-1} + 1$,
If $MaxL = L_{4q}(a)$, then $Sc_{4q} = Sc_{4q} + 1$.

6. Calculate probabilities of recognizing (PoR) scenarios. These are probabilities for the actual scenario to be recognized as one of scenarios in question.

$$Pr_{4q-3} = Sc_{4q-3}/u,\ Pr_{4q-2} = Sc_{4q-2}/u,\ Pr_{4q-1} = Sc_{4q-1}/u,\ Pr_{4q} = Sc_{4q}/u.$$

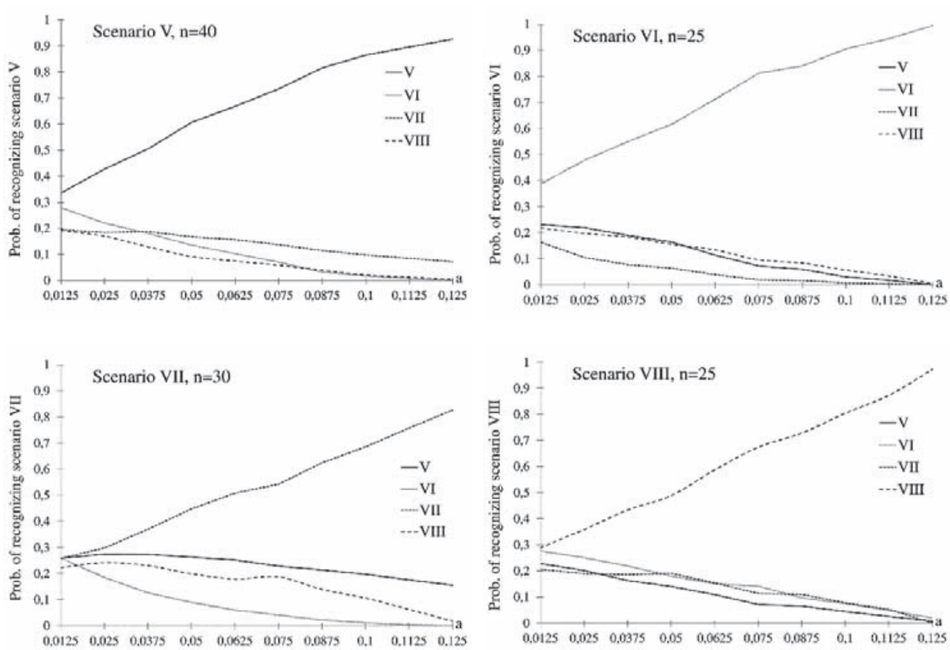
Figures 6-10 present PoR(a) functions for two-way CT in question. Sample sizes for a given CT are different because maximal MoU values under the scenarios are different (see table 4). The minimal sample sizes are chosen in such a way that probabilities of proper recognition (actual I as I, …, actual XX as XX) are greater than probabilities of improper recognitions for all the a values.

Figure 6. The PoR actual scenario as one of I–IV scenarios, table 2×3



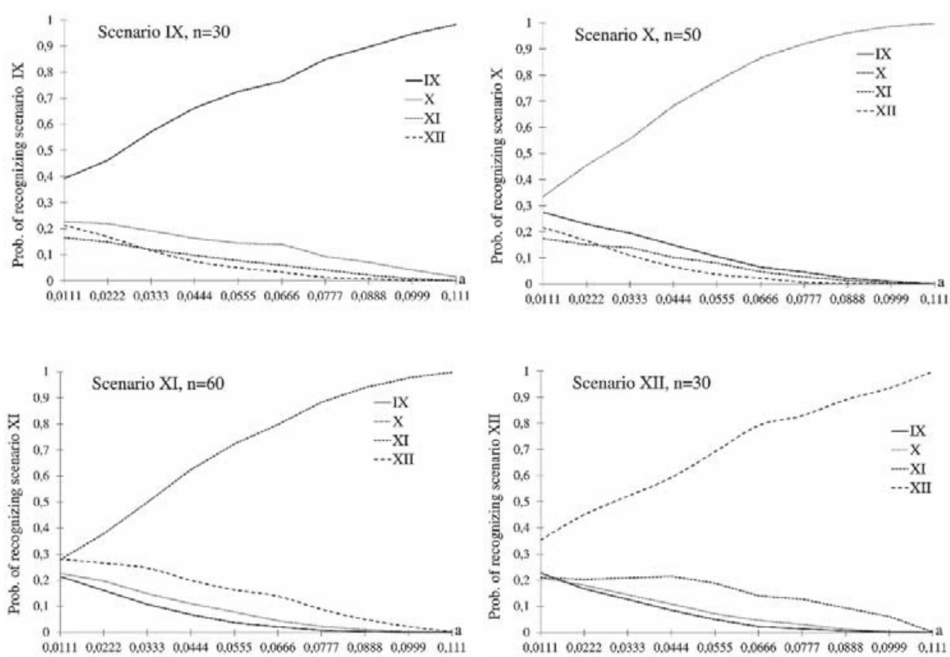
Source: own elaboration.

Figure 7. The PoR actual scenario as one of V–VIII scenarios, table 2×4



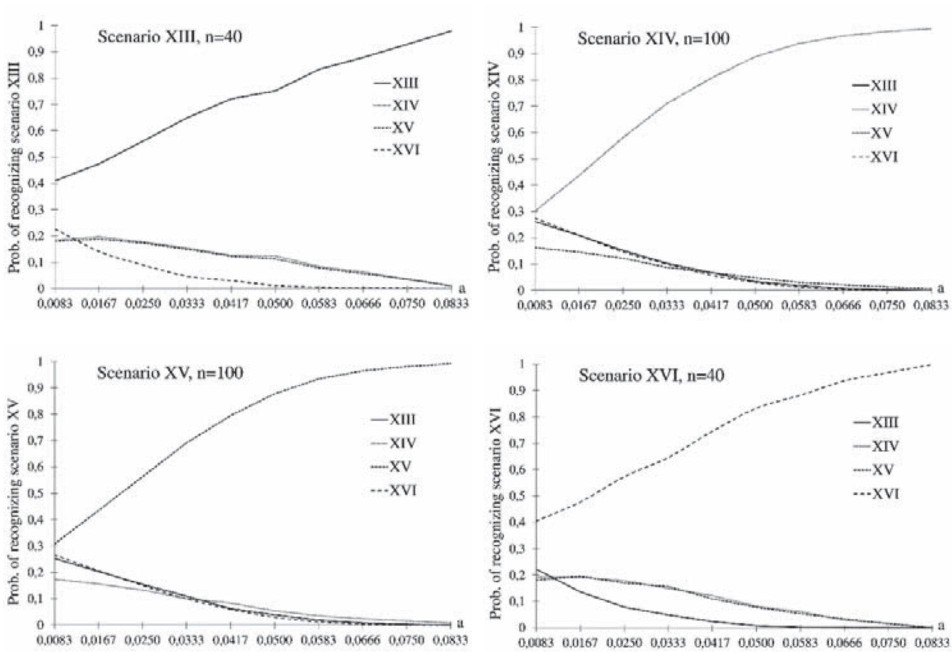
Source: own elaboration.

Figure 8. The PoR actual scenario as one of IX–XII scenarios, table 3×3



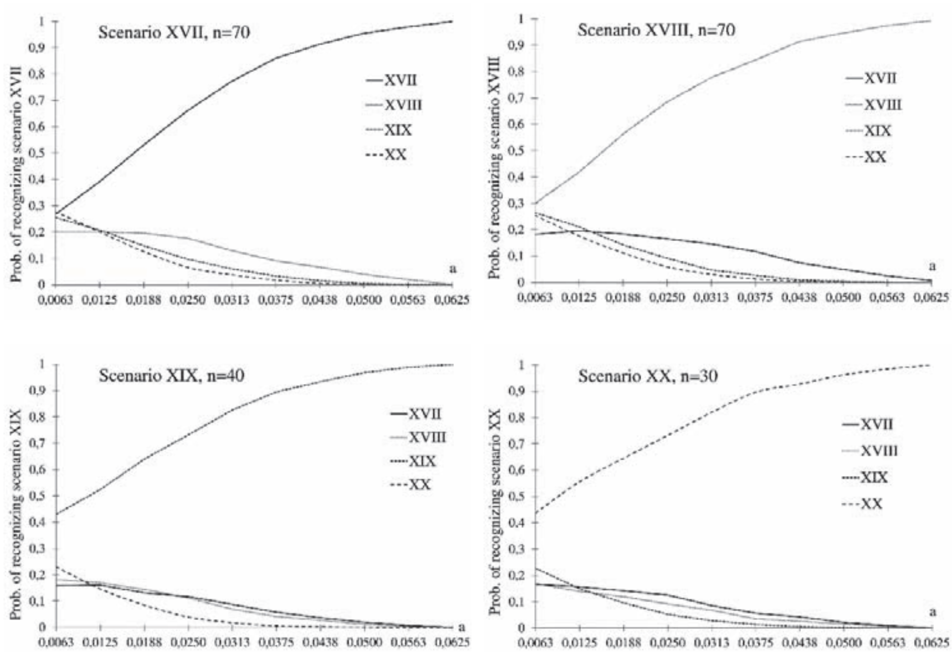
Source: own elaboration.

Figure 9. The PoR actual scenario as one of XIII–XVI scenarios, table 3×4



Source: own elaboration.

Figure 10. The PoR actual scenario as one of XVII–XX scenarios, table 4×4



Source: own elaboration.

Figures 6–10 show that even when samples are small (e.g. 15 items), probabilities of proper recognition are greater than probabilities of improper recognitions, regardless how small the PFP is. The bigger PFP a, the bigger PoR actual scenario. In the classic statistical testing related to 2×3 CT (see table 6) untrue H_0 has not been rejected even if $n = 60$, PFP $a = 0.05$ and $MoU = 0.2$. In a likelihood based decision dependence between features in 2×3 CT is visible already for $n = 15$ and PFP $a < 0.05$ (see figure 6). In the classic statistical testing related to 2×4 CT (see table 7) untrue H_0 has not been rejected even if $n = 80$, PFP $a = 0.0625$ and $MoU = 0.25$. In a likelihood based decision dependence between features in 2×4 CT is visible already for $n = 25$ and PFP $a = 0.0625$ (see figure 7). In the classic statistical testing related to 3×3 CT (see table 8) untrue H_0 has not been rejected even if $n = 90$, PFP $a = 0.0444$ and $MoU = 0.237$. In a likelihood based decision dependence between features in 3×3 CT is visible already for $n = 30$ and PFP $a < 0.0444$ (see figure 8). A similar situation occurs related to 3×4 and 4×4 CTs.

8.2. Three-way contingency table

Example 4. This example compares decision making by means of classic statistic testing with likelihood based decision making. Tables 11–12 show a set of example three-way CTs filled one by one accordingly to the scenarios XXI–XXVIII. The description of these tables has been presented in the example 1. To decide which of the defined scenarios takes place, see algorithm in section 6.2.

Tables 11–12 show that all the decisions made in a classic way are wrong under the scenarios in question. It is because untrue H_0 hypotheses have not been rejected even in situations where the MoU does not have such small values, e.g. $MoU = 0.267$ in scenario XXVIII. In turn, the parametric approach detected a dependency between features.

Example 5. This example is very similar to the Example 2. The new proposal is a method of statistical inference, not a classical parametric test. All the hypotheses state: “the considered three-way CT is generated accordingly to particular scenarios”. Each of figures from 11 to 12 shows a sets of four likelihood curves. Each curve has its four attributes, namely $\hat{a}$, n, table dimensions and, of course, the actual generation scenario that is specified in the figure's title. The legend lists all competing scenarios including the actual one.

Table 11. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE $2 \times 2 \times 2$

Experiment settings: $n = 80, a = 0.0625, \alpha = 0.05$

Scenario (MoU)	Content				X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	$\|\chi\|$	H_0 true	H_0 rejected
							9.477	10.161	9.521	2.402		
XXI(0.130)	5	9	5	14	Yes	XXI	1.505	1.469	1.496	0.979	No	No
	9	15	9	14								
XXII(0.168)	5	5	5	10	Yes	XXII	1.669	1.661	1.665	1.088	No	No
	10	15	15	15								
XXIII(0.188)	5	10	10	10	Yes	XXIII	3.810	4.104	3.855	1.524	No	No
	15	10	10	10								
XXIV(0.250)	5	10	5	10	Yes	XXIV	5.333	5.482	5.350	2.133	No	No
	15	10	15	10								

Source: own elaboration.

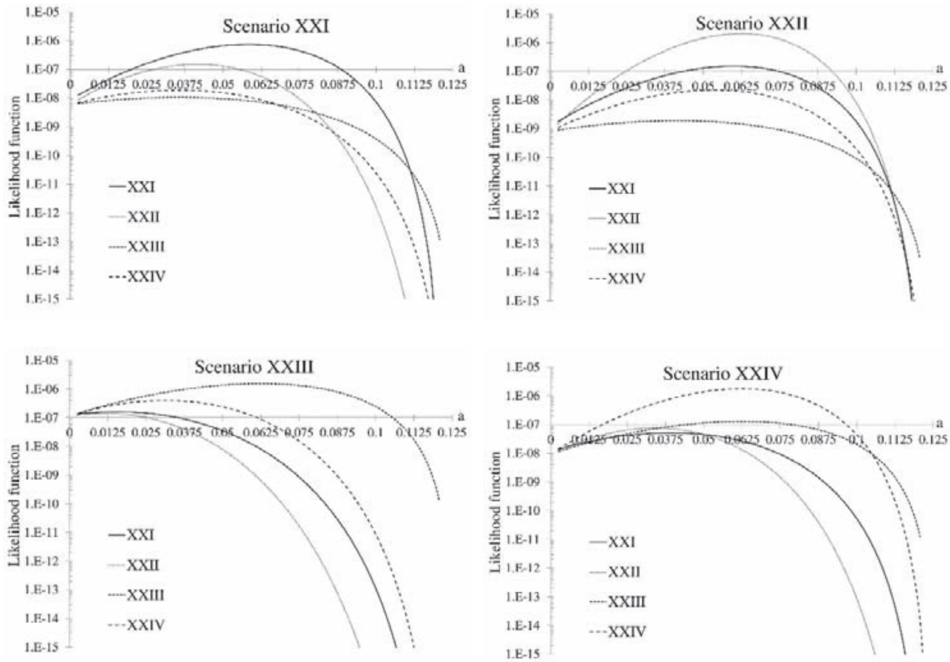
Table 12. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE $3 \times 2 \times 2$

Experiment settings: $n = 120, a = 0.0333, \alpha = 0.05$

Scenario (MoU)	Content				X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	$\lvert\chi\rvert$	H_0 true	H_0 rejected
							13.989	15.069	14.043	3.489		
XXV(0.089)	6	10	6	10	Yes	XXV	1.222	1.222	1.221	1.154	No	No
	10	10	10	10								
	10	14	10	14								
XXVI(0.149)	6	10	6	14	Yes	XXVI	3.697	3.629	3.674	1.815	No	No
	10	10	10	10								
	10	14	10	10								
XXVII(0.200)	6	14	6	14	Yes	XXVII	5.000	5.078	5.009	2.500	No	No
	6	14	10	10								
	10	10	10	10								
XXVIII(0.267)	6	14	6	14	Yes	XXVIII	12.800	13.500	12.879	3.200	No	No
	10	10	14	6								
	14	6	14	6								

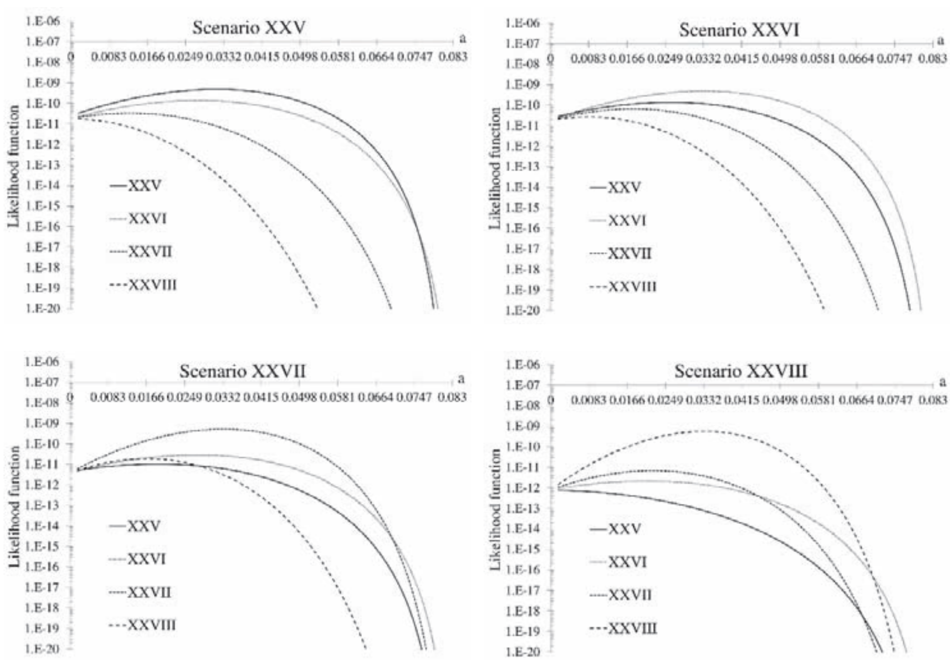
Source: own elaboration.

Figure 11. The likelihood functions for $\hat{a} = 0.0625, n = 80$, table $2 \times 2 \times 2$



Source: own elaboration.

Figure 12. The likelihood functions for $\hat{a} = 0.0333, n = 120$, table $3 \times 2 \times 2$



Source: own elaboration.

It is noteworthy (see figures 11–12) that in each set of likelihood curves, the curve related to the actual scenario predominates over the others. It is instructive to read out a value of a_{max} that maximizes likelihood function on a particular figure and notice that this value is close to assumed $\hat{a}$.

Example 6. This example is carried out in accordance with the following algorithm:

1. Find dimension of CT in question and read out related q according to the rule:

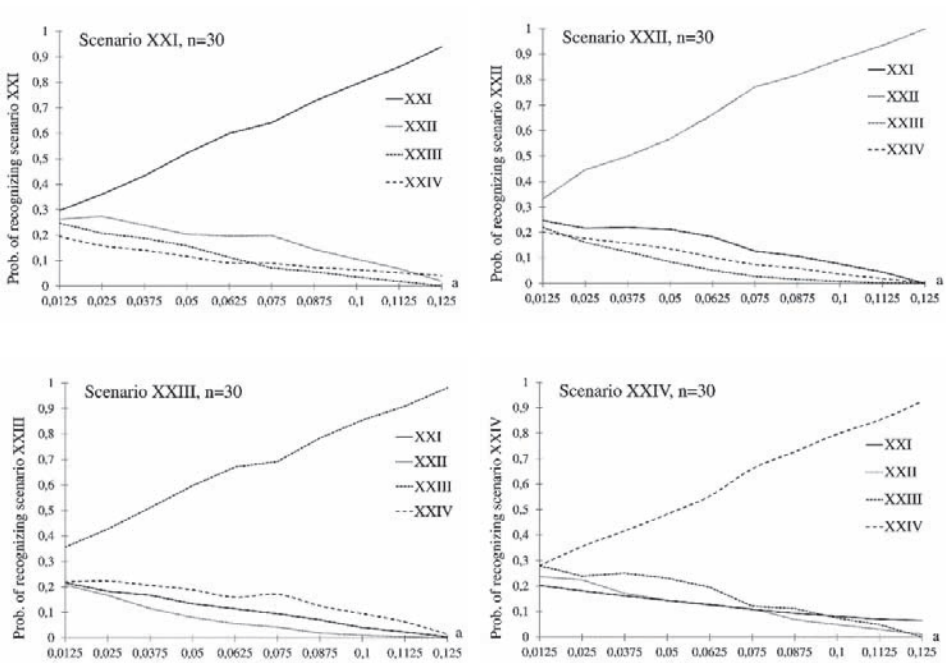
$$2 \times 2 \times 2\ (q = 6), 3 \times 2 \times 2\ (q = 7)$$

2. Set a set of scenario indices $\{4q-3, 4q-2, 4q-1, 4q\}$.
3. Calculate a value of PFP by means of $a_i = 01i/(wkp)\ (i = 1,2,\dots,10)$.

Steps 4–6 are the same as in the example 3.

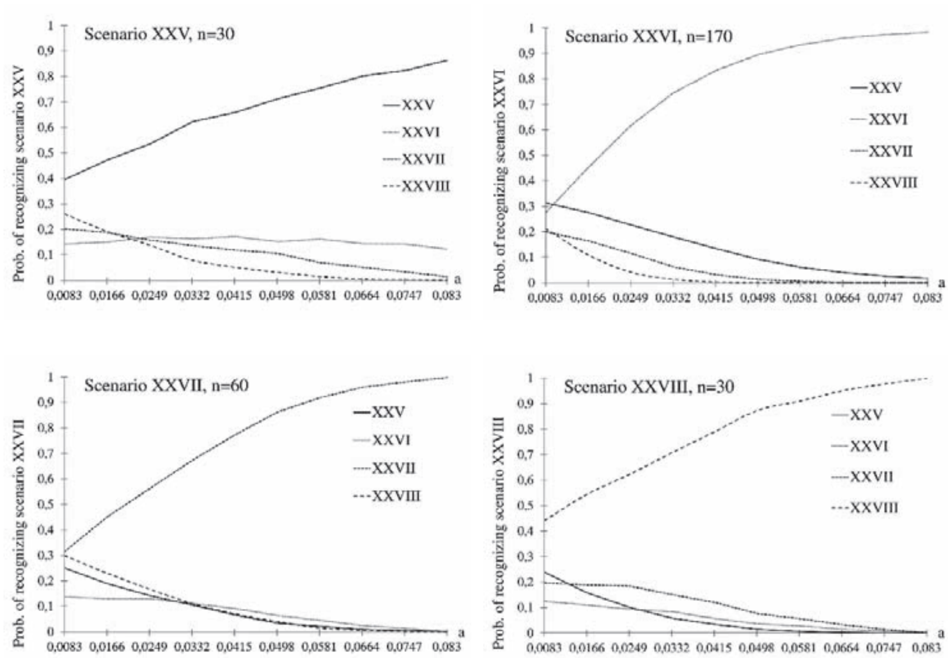
Figure 13–14 present PoR(a) functions for three-way CT in question. Sample sizes for a given CT are different because a maximal MoU values under the scenarios are different (see table 5). The minimal sample sizes are chosen in such a way that probabilities of proper recognition (actual XXI as XXI, …, actual XXVIII as XXVIII) are greater than probabilities of improper recognitions for all the a values.

Figure 13. The por actual scenario as one of XXI–XXIV scenarios, table $2 \times 2 \times 2$



Source: own elaboration.

Figure 14. The por actual scenario as one of XXV–XXVIII scenarios, table $3 \times 2 \times 2$



Source: own elaboration.

Figures 13–14 show that even when samples are small (e.g. 30 items), probabilities of proper recognition are greater than probabilities of improper recognitions, regardless how small the PFP is. The bigger PFP a, the bigger PoR actual scenario. In the classic statistical testing related to $2 \times 2 \times 2$ CT (see table 11) untrue H_0 has not been rejected even if $n = 80$, PFP $a = 0.0625$ and $MoU = 0.25$. In a likelihood based decision dependence between features in $2 \times 2 \times 2$ CT is visible already for $n = 30$ and PFP $a < 0.0625$ (see figure 13). In the classic statistical testing related to $3 \times 2 \times 2$ CT (see table 12) untrue H_0 has not been rejected even if $n = 120$, PFP $a = 0.0333$ and $MoU = 0.267$. In a likelihood based decision dependence between features in $3 \times 2 \times 2$ CT is visible already for $n = 30$ and PFP $a < 0.0333$ (see figure 14).

9. CONCLUSIONS

There are two new elements in the method of statistical reasoning from CTs presented in this paper. Firstly, CTs are parameterized with the probability flow parameters. Parametric reasoning turns out to be much more sensitive in revealing dependency between features than classic reasoning. Secondly, we suggest a scenario (i.e. internal mechanism) under which particular CT comes into being.

Figuring up more and more general scenarios does not seem very difficult. The researches can be generalized by introducing a part of flow parameter, e.g. $a/2, a/3, \dots$ also remembering about the condition of normalization. The researches can also be generalized by introducing several flow parameters. This, however, causes a significant deterioration in the properties of the parameter estimators. You can always add more parameters to the model, however, this might worsen their estimation. Hence, inflated generalizations should be avoided.

REFERENCES

Agresti A., (2002), *Categorical Data Analysis*, Wiley, New Jersey.

Allison P. D., Liker J. K., (1982), *Analyzing Sequential Categorical Data on Dyadic Interaction*: A Comment on Gottman.

Beh E. J., Davy P. J., (1998), Partitioning Pearson's Chi-squared Statistics for a Completely Ordered Three-Way Contingency Table, *The Australian and New Zealand Journal of Statistics,* 40, 465–477.

Bishop Y., Fienberg S., Holland, P., (1975), *Discrete Multivariate Analysis – Theory and Practice,* Cambridge, MA: MIT Press.

Blitzstein J., Diaconis P., (2011), A Sequential Importance Sampling Algorithm for Generating Random Graphs with Prescribed Degrees, *Internet mathematics*, 6 (4), 489–522.

Bock H. H., (2003), Two-Way Clustering for Contingency Tables: Maximizing A Dependence Measure, *Between data science and applied data analysis*, 143–154.

Chen Y., Diaconis P., Holmes S. P., Liu J.S., (2005), Sequential Monte Carlo Methods for Statistical Analysis of Tables, *Journal of the American Statistical Association*, 100 (469), 109–120.

Chen Y., Dinwoodie I. H., Sullivant S., (2006), Sequential Importance Sampling for Multiway Tables, *The Annals of Statistics*, 523–545.

Cressie N., Read T. R., (1984), Multinomial Goodness-Of-Fit Tests, *Journal of the Royal Statistical Society,* Series B (Methodological), 440–464.

Cryan M., Dyer M., (2003), A Polynomial-Time Algorithm to Approximately Count Contingency Tables When the Number of Rows Is Constant, *Journal of Computer and System Sciences*, 67 (2), 291–310.

Cryan M., Dyer M., Goldberg L. A., Jerrum M., Martin R., (2006), Rapidly Mixing Markov Chains for Sampling Contingency Tables with A Constant Number of Rows, *SIAM Journal on Computing*, 36 (1), 247–278.

Cung, B., (2013), *Crime and Demographics: An Analysis of LAPD Crime Data*.

DeSalvo S., Zhao J. Y., (2015), Random Sampling of Contingency Tables Via Probabilistic Divide-And-Conquer. arXiv preprint arXiv:1507.00070.

Diaconis P., Sturmfels B., (1998), Algebraic Algorithms for Sampling from Conditional Distributions, *The Annals of statistics*, 26 (1), 363–397.

Dickhaus T., Straßburger K., Schunk D., Morcillo-Suarez C., Illig T., Navarro A., (2012), How to Analyze Many Contingency Tables Simultaneously in Genetic Association Studies, *Statistical Applications in Genetics and Molecular Biology*, 11 (4),12.

El Galta R., Stijnen T., Houwing-Duistermaat J. J., (2008), Testing for Genetic Association: A Powerful Score Test, *Statistics in Medicine*, 27 (22), 4596–4609.

Fishman G. S., (2012), Counting Contingency Tables Via Multistage Markov Chain Monte Carlo. *Journal of Computational and Graphical Statistics*, 21(3), 713–738.

Freeman M. F., Tukey J. W., (1950), Transformations Related to the Angular and the Square Root, *Annals of Mathematical Statistics*, 21, 607–611.

Gokhale D., Kullback S., (1978), *The Information in Contingency Tables*, Marcel Dekker, Inc., New York.

Goodman L., Kruskal W., (1954), Measures of Association for Cross Classifications, *Journal of the American Statistical Association*, 49, 732–764.

Gray L. N., Williams J. S., (1975), Goodman and Kruskal's Tau b: Multiple and Partial Analogs, In: Proceedings of the Social Statistics Section, *Journal of the American Statistical Association,* 444–448.

Haas P. J., Hueske F., Markl V., (2007), Detecting Attribute Dependencies from Query Feedback, Proceedings of the 33rd International Conference on Very Large Data Bases, VLDB Endowment, 830–841.

Harshman R. A., (1970), Foundations of the PARAFAC Procedure: Models and Conditions for an Explanatory Multi-modal Factor Analysis, UCLA Working Papers in Phonetics, 16, 1–84.

Ilyas I. F., Markl V., Haas P., Brown P., Aboulnaga A., (2004), CORDS: Automatic Discovery of Correlations and Soft Functional Dependencies, Proceedings of the 2004 ACM SIGMOD international conference on Management of data, ACM, 647–658.

Iossifova R., Marmolejo-Ramos F., (2013), When The Body Is Time: Spatial and Temporal Deixis in Children with Visual Impairments and Sighted Children, *Research in Developmental Disabilities*, 34 (7), 2173–2184.

Kaski S., Nikkila J., Sinkkonen J., Lahti L., Knuuttila J. E., Roos C., (2005), Associative Clustering for Exploring Dependencies Between Functional Genomics Data Sets, I*EEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 2 (3), 203–216.

Kullback S., (1959), *Information Theory and Statistics*, Wiley, New York.

Lombardo R., (2011), Three-Way Association Measure Decompositions: The Delta index, *Journal of Statistical Planning and Inference*, 141, 1789–1799.

Lombardo R., Beh E. J., (2010), Simple and Multiple Correspondence Analysis for Ordinal-Scale Variables Using Orthogonal Polynomials, *Journal of Applied Statistics*, 37 (12), 2101–2116.

Marcotorchino F., (1984), *Utilisation des Comparaisons par Pairesen Statistique des Contingences. Partie (I),* Etude IBM F069, France.

Neyman J., (1949), Contribution to the Theory of the x2 Test, Proceedings of the First Berkeley Symposium on Mathematical Statistics and Probability, Univ. of Calif. Press, 239–273.

Oates T., Cohen P. R., (1996), Searching for Structure in Multiple Streams of Data, ICML 96.

Pardo M. C., (1996), An Empirical Investigation of Cressie and Read Tests for the Hypothesis of Independence in Three-way Contingency Tables, *Kybernetika* 32 (2), 175–183.

Pearson K., (1900), On the Criterion That a Given System of Deviations From the Probable in the Case of a Correlated System of Variables is Such That it Can be Reasonably Supposed to Have Arisen From Random Sampling, *Philosophy Magazine Series*, 50, 157–172.

Schrepp M., (2003), A Method for The Analysis of Hierarchical Dependencies Between Items of a Questionnaire, *Methods of Psychological Research Online*, 19, 43–79.

Sokal R. R., Rohlf F. J., (2012), *Biometry: The Principles and Practice of Statistics in Biological Research*, Freeman, New York.

Steinle M., Aberer. K, Girdzijauskas S., Lovis C., (2006), Mapping Moving Landscapes by Mining Mountains of Logs: Novel Techniques for Dependency Model Generation, *P*roceedings of the 32nd international conference on Very large data bases, VLDB Endowment, 1093–1102.

Sulewski P., (2009), Two-By-Two Contingency Table as a Goodness-Of-Fit Test, *Computational Methods in Science and Technology*, 15 (2), 203–211.

Sulewski P., (2013), Modyfikacja testu niezależności, *Wiadomości Statystyczne*, 10, 1–19.

Sulewski P., (2016), Moc testów niezależności w tablicy dwudzielczej większej niż 2x2, *Przegląd Statystyczny*, 63 (2), 191–209.

Sulewski P., (2018a), Power Analysis of Independence Testing for the Three-way Contingency Tables of Small Sizes, *Journal of Applied Statistics*, 45 (13), 2481–2498.

Sulewski P., (2018b), Nonparametric Versus Parametric Reasoning Based on 2×2 Contingency Tables, *Computational Methods in Science and Technology*, 24 (2), 143–153.

Sulewski P., Motyka R., (2015), Power Analysis of Independence Testing for Contingency Tables. *Scientific Journal of Polish Naval Academy*, 56, 37–46.

Tucker L. R., (1963), Implications of Factor Analysis of Three-Way Matrices for Measurement of Change. *In Problems in Measuring Change*, 122–137.

Van Belle G., Fisher L.D., Heagerty P. J., Lumley T., (2004), *Categorical Data: Contingency Tables*, John Wiley & Sons.

Yoshida R., Xi J., Wei S., Zhou F., Haws D., (2011), Semigroups and Sequential Importance Sampling for Multiway Tables. arXiv preprint arXiv:1111.6518.

WNIOSKOWANIE PARAMETRYCZNE I NIEPARAMETRYCZNE W TABLICACH DWUDZIELCZYCH I TRÓJDZIELCZYCH

Streszczenie

W artykule proponowane są scenariusze generowania tablic dwudzielczych (TD) z parametrem przepływu prawdopodobieństwa i zdefiniowane są miary nieprawdziwości H_0. W artykule wykorzystywane są statystyki z rodziny X^2 oraz statystyka modułowa $|X|$. Niniejsza praca jest prostą próbą zastąpienia nieparametrycznej metody wnioskowania statystycznego metodą parametryczną. Metoda największej wiarygodności jest wykorzystana do oszacowania parametru przepływu prawdopodobieństwa. W pracy opisane są także instrukcje generowania TD za pomocą metody słupkowej. Symulacje komputerowe przeprowadzono metodami Monte Carlo.

Słowa kluczowe: wnioskowanie statystyczne, funkcja największej wiarygodności, tablice kontyngencji, test parametryczny, parametr przepływu prawdopodobieństwa

NONPARAMETRIC VERSUS PARAMETRIC REASONING BASED ON TWO-WAY AND THREE-WAY CONTINGENCY TABLES

Abstract

This paper proposes scenarios of generating two-way and three way contingency tables (CTs). A concept of probability flow parameter (PFP) plays a crucial role in these scenarios. Additionally, measures of untruthfulness of H_0 are defined. The power divergence statistics and the $|X|$ statistics are used. This paper is a simple attempt to replace a nonparametric statistical inference from CTs by the parametric one. Maximum likelihood method is applied to estimate PFP and instructions of generating CTs according to scenarios in question are presented. The Monte Carlo method is used to carry out computer simulations.

Keywords: statistical inference, likelihood function, contingency tables, parametric test, probability flow parameter

DOI: 10.5604/01.3001.0014.0544

PRZEGLĄD STATYSTYCZNY
TOM LXV – ZESZYT 3 – 2018

Maciej RYCZKOWSKI[1]
Paweł STOPIŃSKI[2]

Labour market areas in Poland

1. INTRODUCTION

Administrative regions are becoming discrepant with naturally grown functional areas. The historical boundaries diverge often from the contours of territories consistent due to socioeconomic reasons.[3] An urgent need emerges to develop the concept of 'functional regions', often in the form of Labour Market Areas (LMAs). LMA is a functional, geographic region beyond the administrative boundaries. It is economically integrated, while residents may find jobs within a reasonable commuting distance and might change their job without changing their place of residence (compare Gołata, 2004). Therefore, LMAs allow to analyze the effects of commuting on the labour market centres and their hinterland. It should be beneficial for the design of employment, labour mobility and urban planning policies. LMAs could help to decide on the investment plans, road, train infrastructure, pre-schools, kindergartens or managing bus, train connections and other.

A close concept to the LMAs are the Larger Urban Zones (LUZ) (Carlquist, 2006). LUZ is a functional urban area which includes the central city (i.e. core) and an area of the core's main commuting flows from the neighbouring localities (i.e. commuting zone). Młodak (2012) presented a delineation of metropolises and LUZ for Poland and explained the differences between LUZ and Functional Urban Areas (FUA). In turn, FUA for Poland were delineated by taking into account a broader set of selected economic issues like: commuting, migration between the core and the commuting zone, share of the employment outside agriculture sector, number of companies per number of residents, housing market and population density (Śleszyński, 2013). Other works on the analyses of func-

---

[1] Nicolaus Copernicus University in Toruń, Faculty of Economic Sciences and Management, Economics Department, 13a Gagarina St., 87-100 Toruń, Poland, Statistical Office in Bydgoszcz, Surveys and Analysis of Labour Market Centre, 1–3 Stanisława Konarskiego St., 85–066 Bydgoszcz, Poland, corresponding author – e-mail: m_ryczkowski@umk.pl.

[2] Statistical Office in Bydgoszcz, Surveys and Analysis of Labour Market Centre, 1–3 Stanisława Konarskiego St., 85–066 Bydgoszcz, Poland.

[3] See: Młodak (2008) for the selected methods to cluster spatial areas in the surveys of population flows and Klapka, Halás (2016) for a typology of functional regions in geographical research.

tional urban areas for Poland include, for example: Sołtys, Golędzinowska (2017), Szafranek (2017), OECD (2016), Śleszyński (2014).

LUZ, similarly like LMAs and as opposed to FUA, are delineated by using mainly data on commuting flows, which are the main indicator of functional linkage (Rakowska, 2014). However, LMAs are sub-regional geographical areas where the majority of the labour force works and lives. As opposed to it, a commuting zones of LUZ are delineated by Eurostat if at least 15 percent of employed residents work in a city, while if 15 percent of the employed live in one city and work in another city, these cities are treated as a single city. In consequence, LUZ are larger than LMAs and LUZ should contain several LMAs to capture the impact zone of a city. Therefore, LMAs have the potential to play a key role in the evaluation and monitoring of policies at smaller, local areas.

The delineation of Polish local labour markets presumably for the first time in over several years was carried out by Gruchociak (2012, 2013). Author used variants of the taxonomic approach to obtain two LMA's delineations, which were next compared to each other. The algorithm was based on a methodology which was an early attempt of Eurostat (1992) for delineating employment zones at that time. The more recent article of Gruchociak (2015) compared the differences between local labour markets obtained by three different methods and at two points in time, namely for the years 2006 and 2011. The 2011 data came from the 2011 Polish Census of Population and Housing, whereas for the European variant author used the European Algorithm for Regionalisation (Eurostat, 1992). Similarly, Wdowicka (2016) using data from the 2011 Census delineated three different sets of local labour markets with the Eurostat (1992) methodology. Wdowicka (2016) discussed and applied also the selected criteria for evaluation of the results. In turn, Szczebiot-Knoblauch, Kisiel (2014) analysed the supply side of the labor market in rural areas, although authors used the administrative boundaries and they have not delineated any new areas.

The goal of the article is to delineate LMAs in Poland, to discuss LMA methodology and its problems. The novelty is that we use a more recent version of the Mike Coombes, Office for National Statistics (2015) algorithm. The algorithm has been simplified in comparison to previous versions of the TTWA (Travel to Work Areas) algorithm in order to remove the no longer needed set of initial stages to form 'proto-TTWAs' before the main LMA definition process is applied. Moreover, the final version of the algorithm was a result of the discussions during Eurostat Task Force's meetings, seminars and workshops. We add new insights into the literature as we select typical input parameters for Poland. We also propose the output of Polish taxonomy for this purpose. The relevance of the proposal rests on the fact that in literature exists no unambiguous solution to select the optimal values of parameters in the TTWA algorithm. Polish experiences could be beneficial – especially that most empirical papers on LMAs concern highly developed countries.

The policies, decisions and investments resting on regional, functional data (instead on historical boundaries) seem to be required in modern, knowledge-based economies (KBE). The reason is that information infrastructure is a key determinant of KBE (Madrak-Grochowska, 2016). Moreover, Ręklewski, Ryczkowski (2016) evidenced that regional labour market well-being improves the quality of public's life. In turn, Balcerzak, Pietrzak (2015) found a positive impact of the efficiency of the institutions in relation to the potential of the global KBE on the quality of life. The access to the proper spatial information may accelerate economic growth and in consequence may lead to a faster progress towards Europe 2020 targets (European Commission, 2010). Šnajder, Bobek (2014) confirmed that the concept of functional regions is relevant for the effective development. Authors argue that in Slovenia LMAs or functional regions could improve the quality of services, work against state centralization and provide additional stimulus for the regional development. Moreover, Stimson et al. (2016) point that administrative regions create the modifiable area unit problem, which makes it necessary to address spatial autocorrelation issues. Finally, labour mobility associated with functional regions is getting a lot of attention nowadays because it is a mean by which knowledge circulates at the regional scale and matches labour supply and demand (Huber, 2012).

The article is organized as follows. Second section presents literature review and international experiences with LMAs. In the third section we describe the methodology and briefly the EU-TTWA algorithm, followed by a fourth section with data description. In the fifth section, we present the results (with LMA maps) while, in the sixth section, we discuss key problems and possible solutions.

2. EXPERIENCES OF COUNTRIES

Numerous and alternative approaches to the definition of local labour market areas have been developed in recent decades. Casado-Diaz, Coombes (2011) review international scientific research on the delineation of local labour market areas. Official recognition of functional regions varies considerably between countries (OECD, 2002). For instance, in the United States already since 1980s, hierarchical cluster analysis was used by the Census Bureau to group counties into Commuting Zones and LMAs. LMAs were defined similarly to Commuting Zones, except that they were restricted by a minimum population of 100,000 persons. Nevertheless, LMAs were only estimated in 1980 and in 1990 (United States Department of Agriculture). The approach to LMAs differs among countries. In this respect, it was an important and challenging task ahead of Eurostat to define a methodology for creating harmonized LMAs throughout Europe. Ultimately, the algorithm is largely based on its British counterpart.

TTWAs became the official British definition of LMAs since 1960s, although their predecessors go further back in time. Following each national census and starting from the 1971 one, functional areas have been delineated using the data on commuting flows based on the place of residence of an employee and his/her workplace (UK Office for National Statistics, 2015). The method was revealed to work well in practice and has led Coombes (2000) to ascertain that: ‘the TTWA method had been shown to be the *best practice* for defining local labour market areas across Europe’.

In December 2014, Italian National Institute of Statistics (Istat) released LMAs (Istat, 2014; Franconi et al., 2016) that were based on the commuting data with the use of the TTWA method of Coombes, Bond (2008). Istat has been working on this method together with Eurostat and with members of the designated Task Force. The leading role in the Task Force was assigned to Istat, due to its long experience in defining functional regions. Istat released LMAs already in 1989 using the commuting data from the 1981 population Census. Next, similar exercises were repeated in 1991, 2001 and 2011.

For other developed countries the concept of LMAs has also been developed. Kropp, Schwengler (2016) with the use of a novel three-step method obtained 50 German labour market regions that were quite heterogeneous in terms of size. Stimson et al. (2016) implemented the Intramax procedure to the journey-to-work (JTW) commuting flows from the 2011 census data to derive functional, economic regions for Australia. In fact, many approaches could be used to identify functional regions. For instance, Kim et al. (2015) proposed a spatial optimization model with the p-functional regions problem, to solve a regionalization dilemma. Authors considered geographic flows and grouped areal units into smaller number of clusters to classify the areal units with similar properties.

3. METHODOLOGY

The description of the algorithm implemented in Poland (EU-TTWA) can be found in the Eurostat Task Force’s LMA Final Report (Eurostat, 2015). Therefore, we will present it only briefly. We measure self-containment for the supply and demand side. Supply side self-containment (SSC) is the number of people living and working in an area divided by the number of residents in the area. Demand side self-containment (DSC) is the number of people living and working in an area divided by the number of jobs (residents and not-residents employed) in the area. The EU (TTWA) method uses four parameters:

a) minimum self-containment (minSC) – a level of self-containment SC, where SC=min(SSC, DSC), at which clusters of large sizes are acceptable,
b) target self-containment (tarSC), a level of SC at which clusters of small sizes are acceptable,
c) minimum number of working residents (minSZ) for a cluster to be considered a valid LMA,
d) target number of working residents (tarSZ) – a value at which lower levels of self-containment are acceptable for an LMA.

We apply a function which assesses whether a grouping of gminas comprises a viable LMA. This function has the following properties:

a) Cluster of LAUs-2 (Local Administrative Unit: *gmina*) with self-containment (on supply and demand side) that exceeds tarSC and has at least minSZ workers living in the area should be accepted.
b) Cluster of LAUs-2 with self-containment (on supply and demand side) that exceeds minSC and has at least tarSZ workers living in the area should be accepted.
c) Cluster of LAUs-2 in which fewer than minSZ workers live should be rejected.
d) Cluster of LAUs-2 with self-containment (on either supply or demand side) that is less than minSC should be rejected.
e) For cluster of LAUs-2 where live between minSZ and tarSZ workers, the required self-containment (on both supply and demand side) should progressively decrease from tarSC for the smallest areas to minSC for the largest ones.

Therefore, we classify a cluster of LAU-2s to be an LMA if it is consistent with the a–e points, that is the following validity condition must be met:

$$\frac{\text{minSC}}{\text{tarSC}} \leq \left(1 - \left(1 - \frac{\text{minSC}}{\text{tarSC}}\right) MAX\left(\frac{\text{tarSZ} - SZ}{\text{tarSZ} - \text{minSZ}}, 0\right)\right)\left(\frac{MIN\big(SC, \text{tarSC}\big)}{\text{tarSC}}\right). \quad (1)$$

The right-hand side of the condition (1) represents a function that measures the trade-off between the size of a LAU2 unit (SZ) [in occupied persons] and the minimum of SSC and DSC (SC=min(SSC, DSC)). The validity condition (1) fulfils all the properties from ‘a’ to ‘e’. Coombes and Bond’s (2008) methodology allows for a certain degree of flexibility in defining the value of the four parameters. In line with Eurostat’s guidelines each country should define its individual values depending on the specificity of the economy. Nevertheless, the typical values for tarSC are between 0.75 and 0.8, for the minSC are between 0.6 and 0.7, while the size parameters depend on the data – usually they take values from 10 thousands and more.

The algorithm to delineate LMAs checks every LAU against the condition (1). If not all LAUs fulfil the condition, the LAU_A that gives the lowest value for the right-hand side of the condition (1) is selected. This LAU_A is then assessed against all other LAUs to find the LAU_B which has the most important commuting flows in line with the formula:

$$MAX\left(\frac{\text{CF}(LAU_A \rightarrow LAU_B)^2}{ER_{LAU_A} \times E_{LAU_B}} + \frac{CF(LAU_B \rightarrow LAU_A)^2}{ER_{LAU_B} \times E_{LAU_A}}\right). \quad (2)$$

Where, CF – commuting flows, ER – employed residents, E – employed residents and non-residents. LAU_A and LAU_B are grouped. The grouping of LAU_A and LAU_B (LAU_{AB}) is now considered as one entity and the joined commuting flows to the other LAUs are recalculated. Cases when LAUs need to be re-grouped or put on a reserve list (a list of LAUs that fail the validity condition during merging) are described in Eurostat (2015). The process stops when all LAUs or groupings of LAUs fulfil the condition (1).

The problem with the EU-TTWA algorithm is a lack of unambiguous method to select the final values for the minSC, tarSC, minSZ and tarSZ. Therefore, we compared basic data between Poland and countries, which had already used the TTWA algorithm (Italy and Great Britain). As a result of the comparisons and the discussions at the Eurostat LMA's seminars and workshops, we have chosen the following parameter values for further analyses:

- minSZ = {1,000; 2,000; 3,000; 3,500; 4,000; 5,000}
- tarSZ = {7,500; 10,000; 15,000; 16,000; 17,000; 18,000; 19,000; 20,000; 25,000; 30,000; 35,000}
- minSC = {0.5, 0.55, 0.6, 0.667, 0.7}
- tarSC = {0.667, 0.7, 0.75, 0.8, 0.85, 0.9}

We have taken into account the combinations of the above values. In consequence, we estimated several hundred LMAs and evaluated their properties using functions implemented in 'LabourMarketAreas' R-package available on CRAN[4]. In particular, we calculated the characteristics for the alleged LMAs: number of clusters consisting of only one *gmina* (undesirable), number of clusters which fulfill the validity condition (desirable), number of clusters with no central *gmina* (namely, no *gmina* with more people commuting into the *gmina* than commuting out of the *gmina* – undesirable), arithmetical means of the supply side and demand side self containment (the lower is the worse, as both values indicate then that LMA is too 'loose', namely the commuting flows interact strongly with other areas) and the Q-modularity index[5] of Newman, Girvan (2004). We resigned from the sets of input parameters if at least one measure had undesirable properties and we accepted the input parameters, which delivered expected and desirable properties for all the measures. Finally, LMAs with too many non-contiguous areas or with high number of *gminas* on a reserve list were treated as inappropriate. We refer to this part: a 'sensitivity analysis'.[6]

Next, to rank the sets of input parameters with desirable properties (and after abandonment of the sets with undesirable properties), we used the Hellwig's Method of a taxonomic measure of development (Hellwig, 1968; Nowak, 1990;

[4] The most recent version is: Ichim et al. (2017), https://cran.r-project.org/web/packages/LabourMarketAreas/index.html.

[5] Q-modularity index stands for a quality measure for each split of a network into communities. It verifies the correlation between the probability of having an edge joining two sites and the fact that the sites belong to the same community. It is verified for every split while moving down a dendrogram in order to detect local peaks. It equals unity in case of strong community structures – in reality strong community structures are represented by smaller values, typically from 0.3 to 0.7.

[6] The desired and undesired values for the measures were selected by comparisons with the average outcome for all the potential LMAs.

Suchecki, 2010). The method was recently applied, for instance, in Balcerzak (2016). The level of a complex phenomenon (quality of LMAs) is evaluated by a synthetic variable calculated for every LMA as a distance from the abstract, ideal solution (defined by multiple-criteria: stimulants and destimulants). The variables, for which an increase in their value lead to the improvement of the LMA's properties (stimulants) were: mean of the demand side self-containment of LMAs, mean of the supply side self-containment of LMAs, median of the percentage of internal flows (excluding flows having the same *gmina* as origin and destination) within LMAs to the total internal flows (Cohesion 1), median of the ratio of the number of links between communities inside LMA (excluding itself) to the maximum number of possible links (Cohesion 2), Q-modularity index. The variables, for which an increase in the value lead to the deterioration of LMA's properties (destimulants) were: number of LAUs in a reserve-list, percentage of LMAs with only one LAU, percentage of clusters with the right-hand side of the validity condition (1) smaller than unity, percentage of clusters with no LAU having a centrality measure[7] greater than unity. All these statistics were computed cross all LMAs constructed using a given set of input parameters.

In consequence, we have obtained 144 efficient sets of parameters. To unify the variables (stimulants and destimulants), we standardized them to obtain variables z_{ij} with a zero mean and a unitary variance for i=1,2,…,n j=1,2,…,m, where:

n – number of sets of expected input parameters, where n=144,

m – number of variables (stimulants and destimulants), where m=9.

Next, we changed all the variables into stimulants by multiplying the values of the standardized destimulants by minus one. The optimal object would be then:

$$z_j^W = \max_i z_{ij},$$

The distance of each LMA from the abstract, optimal benchmark was measured by the Euclidean formulae:

$$d_{i0} = \sqrt{\sum_{j=1}^{m} (z_{ij} - z_j^W)^2 w_j} \quad for\ i = 1,2,\dots,n. \tag{3}$$

We assumed all the weights w_j to be equal (in the basic scenario), where:

$$\sum_{j=1}^{m} w_j = 1 \text{ and } w_j \geq 0\ for\ j = 1,2,\dots,m.$$

[7] The centrality measure $C_k=\frac{f_{*k}-f_{kk}}{f_{k*}-f_{kk}}$, where f_{*k} denotes commuting flows from all gminas to gmina k, f_{k*} are the commuting flows from gmina *k* to all gminas and f_{kk} denotes commuting flows within gmina *k*.

The development measure (synthetic variable) is given by (the higher the values of m_i, the better LMAs' properties):

$$m_i = 1 - \frac{d_{i0}}{d_0}, \qquad (4)$$

where d_0 is the Euclidean distance between z_j^W and z_j^A, $m_i \in [0,1]$ and:

$$d_0 = \sqrt{\sum_{j=1}^{m}(z_j^W - z_j^A)^2 w_j}\,. \qquad (5)$$

Table 1. ALTERNATIVE WEIGHTS USED TO CONSTRUCT THE TAXONOMIC MEASURE OF DEVELOPMENT

Variables	Weights 1	Weights 2	Weights 3
Destimulants			
Number of LAUs in a reserve-list	0	0.1	0
LMAs with only one LAU	0.125	0.15	0.16
LMAs with expression (1) smaller than unity	0.125	0.17	0.18
LMAs with centrality measure greater than unity	0.125	0.13	0.14
Stimulants			
Mean SC demand side	0.125	0.07	0.1
Mean SC supply side	0.125	0.07	0.1
Cohesion 1	0.125	0.06	0.06
Cohesion 2	0.125	0.06	0.06
Q modularity index	0.125	0.19	0.2
Sum of weights	1	1	1

Source: own elaboration.

We experimented also with different structures of weights (table 1). Eventually, the final, chosen by us set of input parameters was evaluated on the basis of the knowledge of experts from the statistical offices. The acceptance of the delineated LMAs was based on comparisons of the LMAs with the official labour market data and spatial distribution of companies that could attract employment.

4. DATA

Data come from the Polish National Census of Population and Housing 2011. The Census was based on direct interviews and twenty-eight administrative sources. Data for creating LMAs are aggregated at the LAU–2 level. We used administrative part of the Census to create a matrix of commuting flows. Persons at the age of fifteen years old or more and having in the National Insurance System an insurance code of the employed were taken into account. Persons, who were not employed, working abroad or those for whom it was impossible to define a place of work from the registers, were excluded from the matrix. Next, for each person, two LAU-2 codes: living_code (place of residence of an employee) and working_code (place of work of an employee) were specified. Each farmer was considered as living and working in the same *gmina*. Persons who have not declared travelling to work in a tax registry, had their working_code and living_code set equal. The final matrix of commuting flows between living_code and working_code contained 277,686 links (the number stands for existing combinations of gminas, not for people commuting to work between gminas).

5. RESULTS

Since density of population in Poland is lower than the density of population in Great Britain and in Italy, it was decided to check the target self-containment between 0.667 and 0.85. Lower (higher) values of the target self-containment resulted in too many (too few) LMAs. Simple and proportional relationship between the density of population and the number of LMAs in Italy and Great Britain, suggests that the number of LMAs in Poland should fall between 110 and 378. According to Eurostat guidelines in each LMA the majority of persons is supposed to both live and work. Therefore, the values for minimal self containment below 0.5 were not considered. To maintain the international comparability of LMAs' definitions, the value of the minimal self-containment was decided to be analyzed between 0.6 and 0.7. The upper bounds for the minimal size of LMA and for the target size of LMA were chosen close to the values in Great Britain. Their lower bound was chosen equal to the values in Italy. The most important characteristics of labour market in Italy, Great Britain (UK) and Poland are summarized in table 2.

Table 2. BASIC DATA ON ITALY, GREAT BRITAIN AND POLAND IN 2011

Variable	Italy	Great Britain	Poland
Population (persons)	59,433,744	63,182,180	38,044,565
Population of 15 years & more (persons)	51,107,701	52,082,285	32,262,995
Economically active (persons)	25,985,295	32,442,335	17,576,246
Employed (persons)	23,017,840	30,008,635	15,443,421
Unemployed (persons)	2,967,455	2,433,705	2,132,825
Area (thousands km^2)	302,073	248,528	312,679
Density of population (persons/km^2)	197	254	122
Number of building blocks[1]	8,092	10,399	3,081
Minimal size of LMA (working residents)	1,000	3,500	4,000
Target size of LMA (working residents)	10,000	25,000	30,000
Minimal self-containment of LMA	0.6	0.667	0.667
Target self-containment of LMA	0.75	0.75	0.8

Table 2. BASIC DATA ON ITALY, GREAT BRITAIN AND POLAND IN 2011 (dok.)

Variable	Italy	Great Britain	Poland
Number of LMAs	611	228	339
Average population in a LMA (persons)	97,273	277,114	112,226
Average area of a LMA(thousands km^2)	494	1,090	922
Average number of building blocks in a LMA	13	46	9.1

[1]In Poland building blocks consist of gminas (LAU-2); in Italy LAU-2; in Great Britain: a) in England and Wales layer super output areas (LSOA); b) in Scotland data zones (DZ); c) in Northern Ireland super output areas (SOA)

Source: own elaboration on the basis of Eurostat, Europa.eu portal, Istat, INSEE, CSO Poland and the Population Census 2011 data.

Using different combinations of the selected values of the four parameters (p. 7), we estimated several hundred LMAs in line with the EU-TTWA method. Next, after the 'sensitivity analysis' (i.e. the procedure of acceptance of the results or their rejection due to LMAs' properties), we performed the rank analysis for the *n*=144 most reliable sets of parameter values (table 3).

Table 3. RANKING OF THE TOP TEN VALUES OF THE PARAMETERS ACCORDING TO THE TAXONOMIC MEASURE OF DEVELOPMENT FOR POLAND

Development measure	Rank	MinSZ	MinSC	TarSZ	TarSC	NbClusters
0.653	1	4000	0.667	30000	0.8	339
0.649	2	3000	0.667	20000	0.8	391
0.646	3	4000	0.667	25000	0.8	346
0.644	4	4000	0.667	20000	0.8	366
0.643	5	5000	0.667	25000	0.75	386
0.643	6	5000	0.667	20000	0.75	390
0.642	7	5000	0.667	15000	0.8	368
0.637	8	3000	0.667	25000	0.8	371
0.637	9	4000	0.667	15000	0.8	387
0.633	10	5000	0.667	25000	0.8	337

Source: own elaboration.

The chosen, final set of parameter values (table 2; table 3, row 1) allowed us to delineate 339 areas (Map 1) within which people commute. Different weights for the stimulants and destimulants in the ranking method have not affected the choice.

The smallest population density in Poland corresponds to the highest minimal size of LMA and to the highest target size of LMA. In consequence, both the average population and average area of an LMA is between values obtained for Italy and Great Britain. Moreover, Poland having the biggest average area of building blocks, has the smallest number of building blocks per LMA. Two LMAs consist of one *gmina* as they fulfill the condition of being a valid LMA (see table 4 for other characteristics).

Map 1. Labour Market Areas in Poland in 2011

Grey contours mark gminas (LAU2), while LMAs are reflected by different colors.
Source: own elaboration using R software and EU-TTWA algorithm on the Population Census 2011 data.

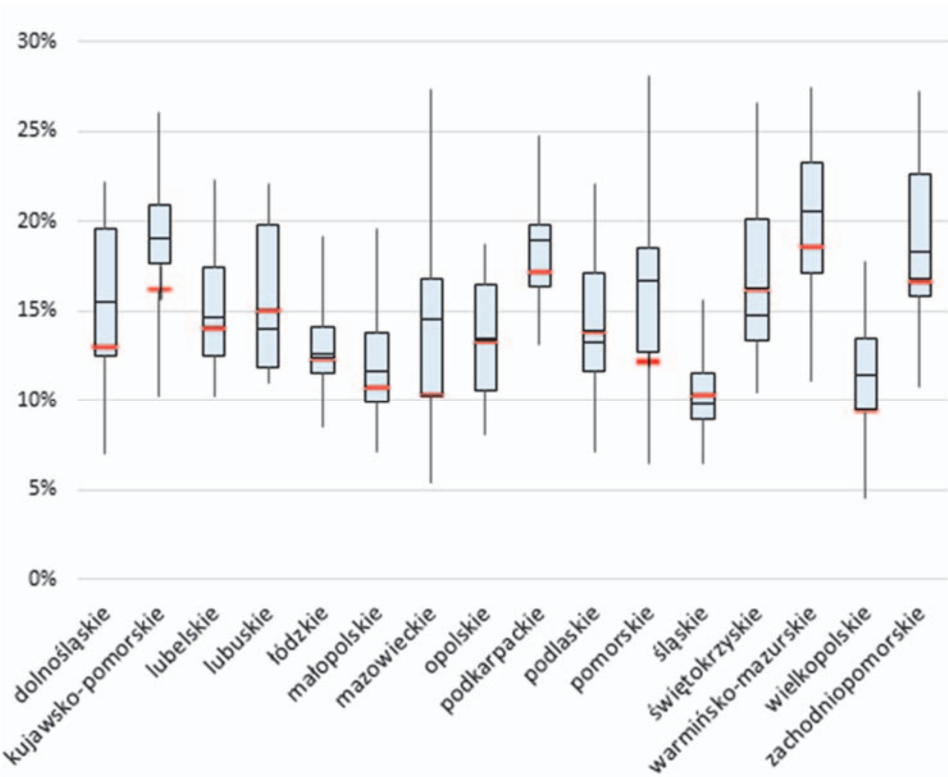
Table 4. POLISH LABOUR MARKET AREAS – SUMMARY

Characteristic	Value
Number of LMAs	339
Mean self-containment	0.816
Mean size (persons)	41,818
Mean number of gminas forming the LMA	9.1
Mean validity	1.12
Number of LMAs with validity < 1	1
Number of links between LMAs	46,167
Number of LMAs with no gminas having a centrality measure > 1	41
Mean SC (demand side)	0.902
Standard deviation of the DSC	0.050
Mean SC (supply side)	0.822
Standard deviation of the SSC	0.055

Source: own elaboration on the Population Census 2011 data.

LMAs should constitute economically integrated regions. Indeed, for the ratio of the unemployed to the population in the working age: an average of the ratio's absolute deviation between LMAs is 2.5%, whereas an average of the ratio's absolute deviation inside LMAs is 1.3%. This indicates that the differences are greater between LMAs than within LMAs.

Figure 1. Selected statistics on the LMA unemployment rate for voivodships, March 2011



Any LMA_x is assigned to the voivodship, where gmina with the highest number of working residents among all gminas belonging to this LMA_x.is located For every voivodship, the minimum, first quartile, median, third quartile and the maximum values of the unemployment rate were visualized. The red lines present the unemployment rate in the voivodship.

Source: own elaboration.

We find that Mazowieckie voivodship has the largest difference between its minimum and maximum values of unemployment rates between LMAs (figure 1). The registered unemployed to the working age population ratio was lower in LMA 1550 (containing Warsaw)[8] (4.08%) than in the neighbouring LMAs (4.96%-

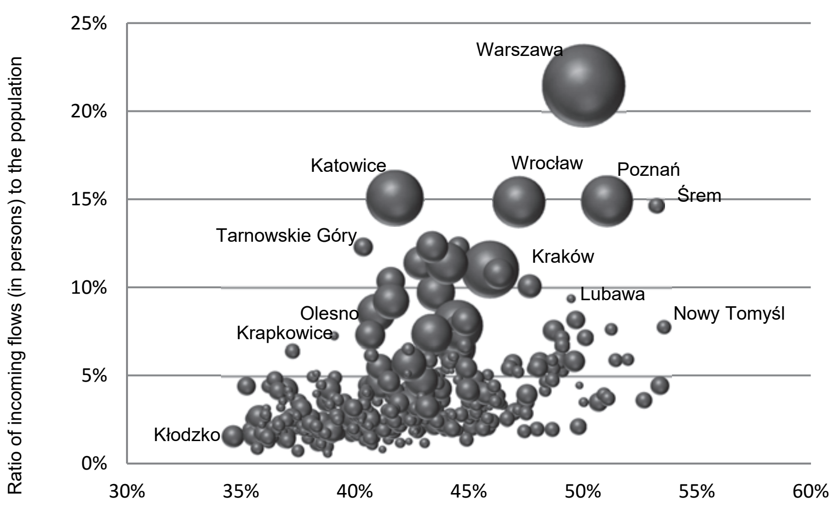
[8] The numbers next to LMA are numbers assigned by an algorithm to an LMA and thus they have no interpretation.

13.79%). LMA 1550 has also more sizeable ratio of the employed to the working age population in 2011 than the neighbouring LMAs (85.66% in comparison to 31.70%-47.46% in the neighbouring LMAs). This indicates relatively considerable differences between LMAs in voivodships.

LMA 1550 has the largest size of all LMAs in Mazowieckie voivodship with 979,478 inhabitants. Nevertheless, one could expect it to be even larger because of frequent commuting flows to Warsaw (including flows from distant gminas). The reason it does not happen is that commuting in both directions is crucial while defining LMAs. Thus when the neighbouring LMAs were analyzed, all of them turned out to have stronger[9] commuting links inside themselves than with the LMA 1550. The neighbouring LMAs fulfil the conditions to be valid LMAs on their own.

In turn, the ratio of incoming flows was highest in big cities. Nevertheless, in some LMAs in spite of low ratio of incoming flows to population aged over 14 years old and in spite of small population, the percentage of the employed persons exceeded 50% (figure 2).

Figure 2. Flows and employment ratios in 2011 – population aged 15 years or more



Source: own elaboration.

For example, below Nowy Tomyśl, there are two LMAs with high ratio of the employed persons to the population and low ratio of incoming flows to the population (figure 2). Both Grojec LMA and Grodzisk Wielkopolski LMA are neigh-

[9] Namely, more commuters travel to work inside particular LMAs (other than LMA 1550) than they travel to work between a neighbouring LMA and LMA 1550.

bouring to LMAs containing big cities (Warsaw and Poznan respectively). This fact causes a small number of incoming flows, relatively big number of outcoming flows (mainly to the neighbouring big cities) and high value of the ratio of the employed persons to the population. In general, LMAs are not homogeneous. For good policy making, one should consider not only their borders but also the specificity of the regions where employees commute.

6. PROBLEMS

Poland-specific problems appeared, which were not solved by the EU-TTWA algorithm. The knowledge about them may help to delineate LMAs in countries without specified commuting areas.

We found commuters between very distant gminas. The reason was that under certain conditions the algorithm assigned people to work in the headquarters instead of in the actual place of work due to the specificity of Polish registers. Even though in tax registers information on the fact of travelling to work is available (and it was used), it has not fully solved the problem, at least for those who travel to work[10]. The solution could be to analyze distance between *gminas* and to neglect insignificant links between *gminas* located unreasonably far from each other and with no fast transportation connections to commute. It is intended to be tested on the next Census data.

After running the algorithm, 26 non-contiguous LMAs appeared. According to Eurostat's requirements non-contiguous LMAs can be accepted only if they contain an administrative island (a non-contiguous part of a LAU-2). A contiguity for all the other LMAs in each country must be provided during the fine tuning process. Gminas which caused non-contiguity were ordered by size and each of them was assigned to the neighbouring LMA where it had the maximum value of expression (2). Moreover, we found 352 towns surrounded by a rural part of the gmina. In consequence, the town and the rural *gmina* may have been initially assigned to different LMAs. It resulted in 'holes' in LMAs. During the fine tuning process both parts were attached to the same LMA.

An example of non-contiguity was the LMA 1613 that consisted of two separate parts - part one with four *gminas* and 9,636 residents and part two with three *gminas* and 7,795 residents (Map 2). To provide contiguity we could split the LMA into two independent LMAs. However, self-containment of the 'part two' was lower than the required threshold. As for 'part one', the condition of being a valid LMA was not met. Both arguments were against creating two LMAs out of one. Therefore we calculated the number of 'attracting *gminas*' using the centrality index C_k (if C_k>1 then '*gmina* is attracting').

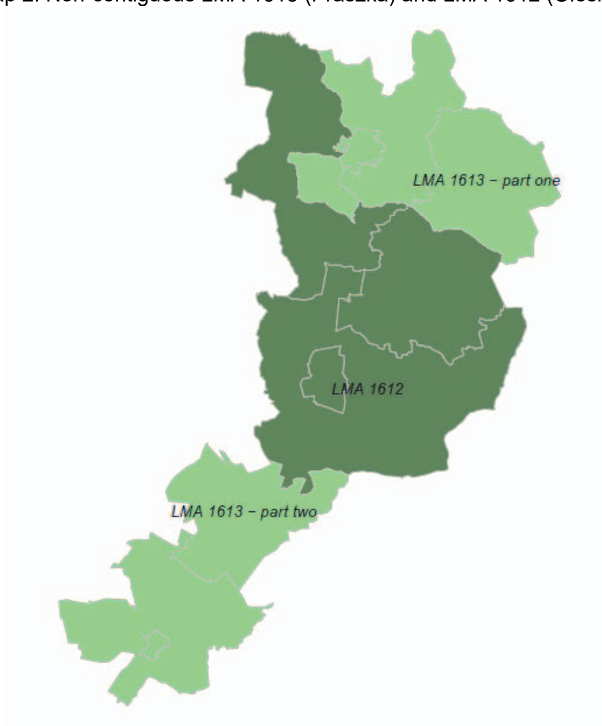
[10] The reason is that for all the enterprises in the Census at most three last registered places of business activity were available. Moreover, for some companies in the Census data only the headquarters' address was available.

Table 5. Centrality index values for gminas in LMA 1613

Territorial code of a gmina	Gmina	Part	Centrality index value
1608024	Gorzów Śląski	one	**1.39**
1608044	Praszka	one	**1.02**
1608045	Praszka	one	0.20
1608062	Rudniki	one	0.15
1608072	Zębowice	two	0.17
1609084	Ozimek	two	0.95
1609085	Ozimek	two	2.35

Source: own elaboration on the Population Census 2011 data.

Map 2. Non-contiguous LMA 1613 (Praszka) and LMA 1612 (Olesno)



Source: own elaboration using the *LabourMarketAreas R*-package.

Part one contained two attracting *gminas* and part two had only one (table 5). Therefore, we let part one to stay as an autonomous LMA[11] and we split the 'part two' into *gminas*. Each of them was assigned to one of the neighbouring LMAs with the largest commuting flows with it. After the new assignment of gminas to the neighbouring LMAs, the validity condition was still met.

[11] Eurostat allows for a small number of LMAs which do not fulfil the validity condition if there are required changes during the fine-tuning process. In consequence, the part one was accepted as an LMA.

7. SUMMARY

EU-TTWA algorithm allowed us to delineate 339 Labour Market Areas in Poland. The novelty is that we use a more recent version of the Mike Coombes, Office for National Statistics (2015) algorithm. Moreover, the final version of the algorithm was a result of the discussions during Eurostat Task Force's meetings, seminars and workshops. Therefore, we maintained international comparability with the results of European countries by following the recent EU guidelines. We add new insights into the literature as we select typical input parameters for Poland. We also propose the usage of Polish taxonomy as a way to select values of the input parameters. Labour Market Areas may be used for compiling and evaluating data related to the road, train, plane, or bus infrastructure, the spatial distribution of pre-schools or kindergartens, the timetables of public transportation, the premises for investments and many more. After all, right policies, decisions and investments resting on exact regional data (instead on historical boundaries) are required in modern knowledge-based economies. Labour Market Areas by delivering new spatial information may accelerate the growth of an economy. Nevertheless, for good policy making, one should consider not only their borders but also the specificity of the region where employees commute as we found large dissimilarities between LMAs. A future task would be to differentiate LMAs by occupational categories, employment by gender, mode of travel to work and other. The revised LMAs should be delineated after the Population Census 2021 to assess the changes in the Polish commuting patterns within recent decade. Finally, delimitation of industrial districts seems also to be a promising concept. Assessing the functional polycentricity may be applied to relatively small regions (Hanssens et al., 2014). Detailed analysis of functional regions covering, for instance, an urban area of Silesian Metropolis could be an interesting case study as well (see Sojka, 2013).

REFERENCES

Balcerzak A. P., (2016), Multiple-Criteria Evaluation of Quality of Human Capital in the European Union Countries, *Economics and Sociology*, 9 (2), 11–26.

Balcerzak A. P., Pietrzak M. B., (2015), Wpływ efektywności instytucji na jakość życia w Unii Europejskiej. Badanie panelowe dla lat 2004–2010, *Przegląd Statystyczny*, 62 (1), 71–91.

Carlquist T., (2006), Larger Urban Zones in the URBAN AUDIT Programme, in Dziechciarz J., (ed.), 25th SCORUS Conference on Regional and Urban Statistics and Research. Globalization Impact on Regional and Urban Statistics, Publishing House of the Wrocław University of Economics, 30, Wrocław.

Casado-Diaz J. M., Coombes M., (2011), The Delineation of 21st Century Local Labour Market Areas: A Critical Review and a Research Agenda, *Boletín de la Asociación de Geógrafos Españoles*, 57, 7–32.

Coombes M., (2000), Geographic Information Systems: A Challenge for Statistical Agencies*, Research in Official Statistics*, 3 (2), 77-87.

Coombes M., Bond S., (2008), *Travel-to-Work Areas: the 2007 Review*, Office for National Statistics, London.

Coombes M., Office for National Statistics, (2015), *Travel to Work Areas*,

www.ncl.ac.uk/media/wwwnclacuk/curds/files/RR2015-05.pdf.

European Commission, (2010), Europe 2020, *A Strategy for Smart, Sustainable and Inclusive Growth,Brussels,* http://eurlex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:52010C2020&from=EN.

Eurostat, (1992), *Study on Employment Zones*, E/LOC/20, Luxembourg.

Eurostat, (2015), *Task Force, Harmonised Labour Market Areas, Final Report*,

https://ec.europa.eu/eurostat/cros/system/files/Task%20Force%20on%20LMA%20Final%20Report.pdf_en.

Franconi L., D'alò M., Ichim D., (2016), *Istat Implementation of the Algorithm to Develop Labour Market Areas*, https://www.istat.it/en/files/2016/03/Description-of-the-LabourMarketAreas-algorithm.pdf.

Gołata E., (2004), *Estymacja bezpośrednia bezrobocia na lokalnym rynku pracy*, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań 2004.

Gruchociak H., (2012), Delimitacja lokalnych rynków pracy w Polsce, *Przegląd Statystyczny*, 2, 277–297.

Gruchociak H., (2013), Delimitacja lokalnych rynków pracy w Polsce na podstawie danych z badania przepływów ludności związanych z zatrudnieniem, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 278, 343–350.

Gruchociak H., (2015), Porównanie struktury lokalnych rynków pracy wyznaczonych przy wykorzystaniu różnych metod w Polsce w latach 2006 i 2011, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 385, 111–119.

Hanssens H., Derudder B., Van Aelst S., Witlox F., (2014), Assessing the Functional Polycentricity of the Mega-City-Region of Central Belgium Based on Advanced Producer Service Transaction Links, *Regional Studies*, 48 (12), 1939–1953.

Hellwig Z., (1968), Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr, *Przegląd Statystyczny*, 15 (4), 307–327.

Huber F., (2012), Do Clusters Really Matter for Innovation Practices in Information Technology? Questioning the Significance of Technological Knowledge Spillovers, *Journal Of Economic Geography*, 12 (1), 107-126.

Ichim D, Franconi L., D'Alo' M., Heuvel G., (2017), *LabourMarketAreas: Identification, Tuning, Visualisation and Analysis of Labour Market Areas*, https://cran.r-project.org/web/packages/LabourMarketAreas/index.html.

Istat, (2014), *Labour Market Areas, Italian National Institute of Statistics,* http://www.istat.it/en/files/2014/12/EN_Labourmarketareas_2011.pdf?title=Labour+Market+Areas+-+17+Dec+2014+-+Full+text.pdf

Kim H., Chun Y., Kim K., (2015), Delimitation of Functional Regions Using a p-Regions Problem Approach, *International Regional Science Review*, 38 (3), 235–263.

Klapka P., Halás M., (2016), Conceptualising Patterns of Spatial Flows: Five Decades of Advances in the Definition and Use of Functional Regions, *Moravian Geographical Reports*, 24 (2), 2–11.

Kropp P., Schwengler B., (2016), Three-Step Method for Delineating Functional Labour Market Regions, *Regional Studies*, 50 (3), 429-445.

Madrak-Grochowska M., (2016), The Information Infrastructure of Knowledge-Based Economies in the Years 1995–2010, *Ekonomia i Prawo*, 15 (4), 503-515.

Młodak A., (2008), Metody grupowania w badaniach przepływów ludności, *Wiadomości Statystyczne*, 53 (9), 28–40.

Młodak A., (2012), Statystyka metropolii polskich – problemy i perspektywy, *Studia Regionalne i Lokalne*, 2 (48), 20–38.

Newman M. E. J., Girvan, M., (2004), Finding and Evaluating Community Structure in Networks, *Physical Review E*, 69 (2), 026113.

Nowak E., (1990), *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*, PWE, Warszawa.

OECD, (2002), *Redefining Territories The Functional Regions: The Functional Regions*, http://www.urbanpro.co/wp-content/uploads/2017/04/Redefining-TerritoriesThefunctional-regions.pdf.

OECD, (2016), Functional Urban Areas In OECD Countries: Poland, https://www.oecd.org/cfe/regional-policy/functional-urban-areas-all-poland.pdf.

Rakowska J., (2014), Codzienne dojazdy do pracy jako ekonomiczne kryterium rządowych klasyfikacji i delimitacji obszarów (na przykładzie USA i Kanady), *Studia Regionalne i Lokalne*, 3 (57), 46–59.

Ręklewski M., Ryczkowski M., (2016), The Polish Regional Labour Market Welfare Indicator and Its Links to Other Well-Being Measures, *Comparative Economic Research*, 19 (3), 113–132.

Śleszyński P., (2013), Delimitacja miejskich obszarów funkcjonalnych stolic województw, *Przegląd Geograficzny*, 85 (2), 173–197.

Śleszyński P., (2014), Delimitation and Typology of Functional Urban Regions in Poland Based on Commuting, 2006, *Geographia Polonica*, 87 (2), 317–320.

Šnajder L., Bobek V., (2014), Regionalization of Slovenia by Establishing Functional Regions, *Our Economy (Nase Gospodarstvo)*, 60 (1/2), 26-36.

Sojka E., (2013), Analiza sytuacji na lokalnym rynku pracy z wykorzystaniem zmiennej syntetycznej, *Studia Ekonomiczne – Zeszyty Naukowe Wydziałowe Uniwersytetu Ekonomicznego w Katowicach*, 2 (160), 33-43.

Sołtys J., Golędzinowska A., (2017), Integrated Development Plans of the Functional Urban Areas in Pomeranian Region in Poland, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 476, 30–38.

Stimson R., Mitchell W., Flanagan M., Baum S., Tung-Kai S., (2016), Demarcating Functional Economic Regions Across Australia Differentiated By Work Participation Categories, *Australasian Journal of Regional Studies*, 22 (1), 27-57.

Suchecki B., (2010), *Ekonometria przestrzenna. Metody i modele analizy danych przestrzennych*, Wydawnictwo C.H. Beck, Warszawa.

Szafranek E., (2017), Miejskie obszary funkcjonalne a kształtowanie spójności terytorialnej, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 467, 113-129.

Szczebiot-Knoblauch L., Kisiel R., (2014), Podaż siły roboczej na rynku pracy na obszarach wiejskich w Polsce, *Oeconomia Copernicana*, 5 (1), 97-115.

UK Office for National Statistics, (2015), Changes in Travel to Work Areas from 2001 to 2011,webarchive.nationalarchives.gov.uk/20160106004208/http://www.ons.gov.uk/ons/dcp171766_426621.pdf.

United States Department of Agriculture, https://www.ers.usda.gov/data-products/commuting-zones-and-labor-market-areas/.

Wdowicka H., (2016), Analiza sytuacji na lokalnych rynkach pracy w Polsce, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 426, 235-244.

OBSZARY RYNKU PRACY W POLSCE

Streszczenie

Celem artykułu jest wyznaczenie obszarów rynku pracy w Polsce przy wykorzystaniu metody EU-TTWA opracowanej pod auspicjami Eurostatu. Wykorzystując wspomnianą metodę otrzymano ponad 300 obszarów rynku pracy składających się z gmin. W artykule opisano charakterystyczne dla Polski problemy i rozwiązania, które przyjęto. Przeprowadzono także porównanie liczby obszarów rynku pracy w państwach, w których wyznaczono za pomocą tego algorytmu wspomniane obszary uwzględniając gęstość zaludnienia, liczbę ludności oraz wielkość danego państwa. W pracy zaproponowaliśmy wykorzystanie metody taksonomicznej celem wyboru parametrów wejściowych w algorytmie EU-TTWA. Obszary rynku pracy mogą dostarczać wartościowej informacji przestrzennej. Niemniej jednak okazały się one niejednorodne pod kątem wybranych statystyk. Dlatego właściwa polityka gospodarcza nie powinna ograniczać się wyłącznie do interpretowania ich granic.

Słowa kluczowe: obszary funkcjonalne, obszary rynku pracy, dojazdy do pracy

LABOUR MARKET AREAS IN POLAND

Abstract

The aim of the article is to delineate Labour Market Areas (LMAs) in Poland with the use of the European version of the Travel to Work Areas (EU-TTWA) methodology that was developed under Eurostat auspices. We received over 300 areas that consist of LAU-2 units (gminas) – the smallest administrative regions in Poland. We discuss Poland-specific results and problems. We compare numbers of LMAs in countries with EU-TTWA-delineated LMAs in relation to population density, total population and area. We propose the taxonomic rank method to select the parameter values for the EU-TTWA algorithm. LMAs may deliver useful spatial information, although one needs to account for their heterogeneity.

Keywords: functional regions, labour market areas, travel to work, commuting