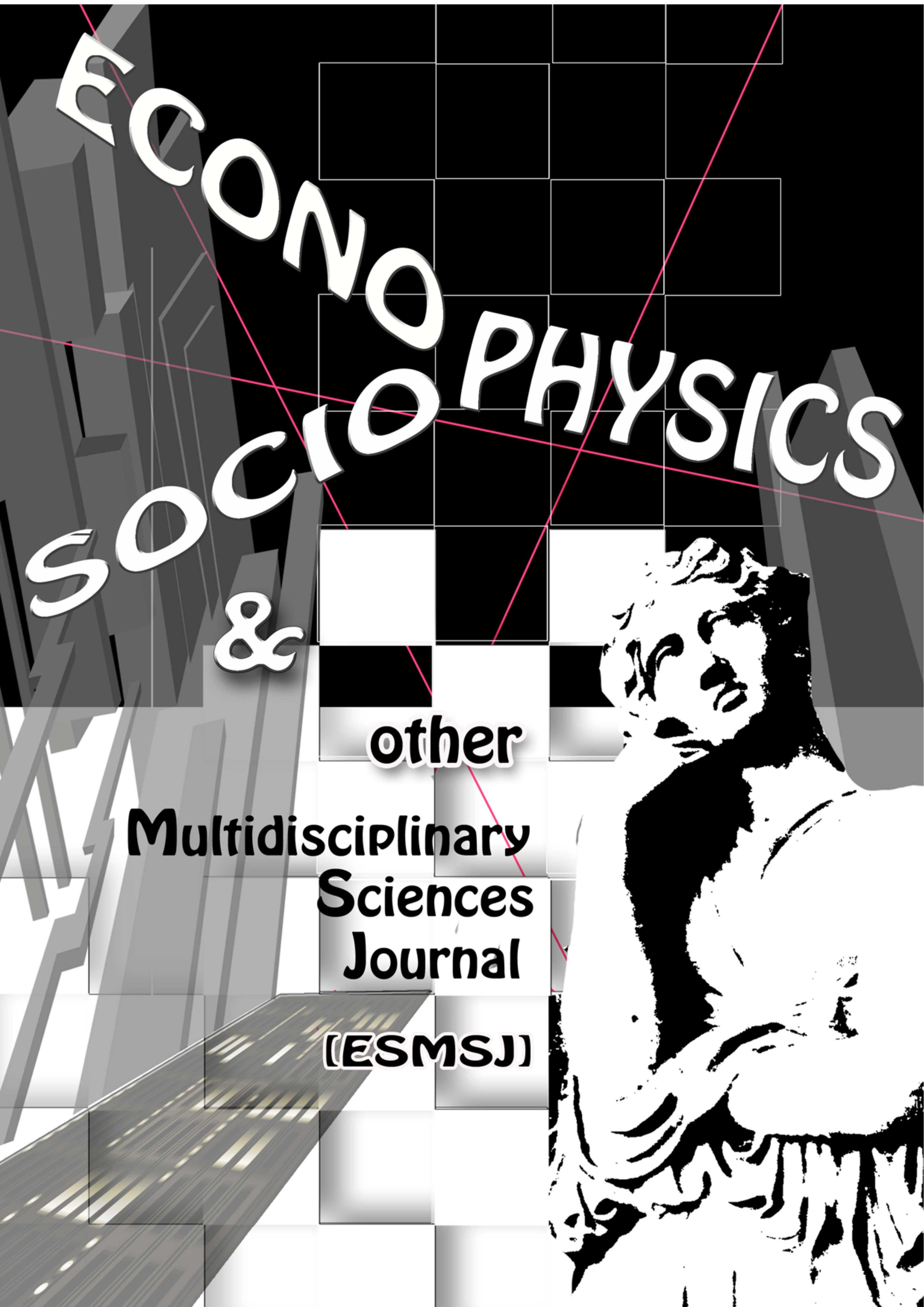

ECONOPHYSICS
SOCIO
&
other
Multidisciplinary
Sciences
Journal
[ESMSJ]

About

Econophysics, Sociophysics & Other Multidisciplinary Sciences Journal (ESMSJ) provides a resource of the most important developments in the rapidly evolving area of Econophysics, Sociophysics & other new multidisciplinary sciences. The journal contains articles from Physics, Econophysics, Sociophysics, Demographysics, Socioeconomics, Quantum Economics, Econooperations Research, or many other transdisciplinary, multidisciplinary and modern sciences and related fundamental methods and concepts.

CONTENT

A DEMOGRAPHIC, ECONOMIC AND STATISTICAL APPROACH TO RELIGION AND WELFARE

Gheorghe_Săvoiu[1], Ion_Iorga-Simăn[2],
[1]University of Piteşti, e-mail: gsavoiu@yahoo.com, [2]University of Piteşti, e-mail: ioniorgasiman@yahoo.com

Abstract: According to Michael J. Vlach, the world's great religions are defined by five significant factors: the number of adherents or followers, territorial dispersion or spreading (meaning the number of countries in which these great religions have found adherents), independence from any other religion of the world, the existence of a material body of their doctrine (usually, a book considered holy by the adherents), and the current practice of religion. In this paper a number of general criteria are also identified for grouping the world's populations, i.e. some synthetic demographic factors are detailed, as well, such as life expectancy and demographic aging. The contribution capitalizes on the data in the annual report of the CIA, available on the Internet, which includes the key indicators on world population, offered to the public by the U.S. institution, detailed and complete for a number of more than 200 countries in the world, where religious faith is present mainly as a firm option of the inhabitants; demographic evolution and economic growth are radically different, and so per capita GDP becomes a polarization axis. Also, the category of those having no religious creed, and that of the atheists are equally important contemporary landmarks in the initial segmentation of the population. Starting from the high degree of determination of religion as a factor of wealth, potential statistical associations areare quantified. Some conclusions naturally arise from the general approach of the statistical investigation.

Keywords: major world religion, religious demography, religion-calendar connections, religion-welfare associations, demography and demographic indicators, Yule association coefficient.

1. INTRODUCTION

Is there a special and statistically quantifiable relationship between religion and some essential elements of contemporary existence, welfare, calendar and demography?

One can quickly detail various general associations for the prolific conjecture questioningly suggested above. Both religion, and any other major concepts remain images of reciprocity in the sense of equidistant reflection, and also bipolar mutual extrapolation.

If the idea is agreed that welfare may be a Weberian reflection of religion consciously assumed, then a regular calendar or schedule is only a reflection in terms of astronomy, or will it also include religious or deep welfare elements, and will demography be influenced by both religion and welfare? Immediate echoes are also the numerous potential statistical associations, concering a certain distribution of wealth in relation to territorially dominant religion, regarding the birth rate, mortality and nuptiality, according to religion and welfare, etc. Statistically speaking, even a multiple, much deeper correlation can be developed, starting from religion as a factor of demography and welfare, and moving to migration flows as factors of welfare convergence, or segmentation or arbitrary setting of zero time along a time scale, symbolically called a calendar ...

If demography increasingly defined itself as a form of statistical and mathematical knowledge of human populations, and religion as a unified system of beliefs and practices of things considered sacred, or well isolated and preserved, (i.e. esoteric), the calendar, as a concept of Greek origin (kalendae), meaning an instrument through which human population is called, convened, welfare is delimited by special economic statistical concepts, such as GDP per capita, or the available average net income. Methodical induction, as used in demography, harmoniously blends with the seduction of religion, rigorous statistical information substitutes for the astronomical precision of the calendar; do the real and natural or earthly worlds, continually and naturally repopulated, have the astral support of the calendars of earthly religions, and also the welfare and effects of the natural or migratory demographic movements? Of course, the present paper cannot answer so many questions, still it can investigate some possible statistical associations centered on religion, focusing especially on religion and welfare.

2. HOW IS A RELIGION OF MAJOR IMPACT DEFINED?

One can define, through the simplicity of demography, and also with the characteristic calm of statistical data, a major religious preponderance of the world population, although there are nations dominated by atheism, lack of faith or irreligiosity. The first states, in which atheists and unbelievers, or people of no religious creed have the primacy in relation to their specific demographic weight, are presented by the American sociologist Phil Zuckerman, who focused mainly on their official data. The first-ranking states of the world, in proportion to the minimum and maximum estimates of the number and weight of unbelievers, atheists or people of no religious creed, in the first decade of the 20th century is presented in the following table

Table no.1

Country	Total population -number of inhabitants-	Atheists, unbelievers of people of no religious creed - in % -	Atheists, unbelievers of people of no religious creed - number of inhabitants -
Sweden	8 986 000	46 - 85	4 133 560 – 7 638 100
Vietnam	82 690 000	81	66 978 900
Denmark	5 413 000	43 - 80	2 327 590 – 4 330 400
Norway	4 575 000	31 - 72	1 418 250 – 3 294 000
Japan	127 333 000	64 - 65	81 493 120 – 82 766 450
Czech Republic	10 246 100	54 - 61	5 328 940 – 6 250 121
Finland	5 215 000	28 - 60	1 460 200 – 3 129 000

France	60 424 000	43 - 54	25 982 320- 31 420 480
South Korea	48 598 000	30 - 52	14 579 400 -25 270 960
Estonia	1 342 000	49	657 580
Germany	82 425 000	41 - 49	33 794 250 – 40 388 250
Russia	143 782 000	24 - 48	34 507 680 – 69 015 360
Hungary	10 032 000	32 - 46	3 210 240 – 4 614 720
Holland	16 318 000	39 - 44	6 364 020 – 7 179 920
England	60 271 000	31 - 44	18 684 010 – 26 519 240

Source: Martin, M., (2005), *The Cambridge Companion to Atheism,* Cambridge UK.: Cambridge University Press.

Atheists claim that God does not exist. They argue that there is no God in the world (Pantheists), or beyond the world (Deists), or, implicitly, do not accept the affirmative idea that God exists both in the world and beyond the world (Theists). This is a positioning simultaneously outside the inner and outer world, of a negativistic type, or affirmative as against a denial of faith, most atheists variously defining themselves as non-theistic, unbelieving, unreligious (or unchurched), etc.

Atheists in our contemporary world are either traditional (in their opinion there has never been, there is not, and will never be a God, and the certainty of there not being a God is, to them, total) or mythological (their God was once alive, that is, a model in which people used to believe, and according to which they lived, but they consider this myth dead and no longer valid, generating an intemediate positioning along a course ranging somewhere between the certainty of mythological existence and the incomplete uncertainty of God's disappearance), or of a semantic type (in this semantic context, any discussion about God is dead, because religious language has no cognitive sense, this concept does not imply a complete denial of the existence of God, but different degrees of uncertainty). Globally, one can define, in keeping with the specific range of arguments and manner of arguing, several atheistic arguments, which form bodies of a number of significant impact theories:

Old and new theories of contemporary atheism

Table no. 2

Theism versus humani-tarianism	A theory of Albert Camus, presented in one of the books of this existentialist philosopher, titled The Plague (La Peste), according to which we must join either the doctor and fight against the "plague" brought by the rats sent by God to a sinful city, or the priest and refuse to fight against the "plague", and so refrain from fighting against a theistic God. Refusing to fight is antihumanitary, and fighting means to face with God, who sent the "plague" as a punishment, and so God does not coincide with the idea of perfection (perfect goodness).
Suffering of the innocent	Contrary to some theists who claim that this is the best of all possible worlds, it seems an undeniable that life in this world could be improved. For example, not all evils are deserved, cruelty, cancer, rape hurt innocent victims. But an all-wise, almighty and perfectly good God would not allow innocents to suffer. Even a single act of injustice in the world (and it is sure that there are many such acts of injustice) would argue against the existence of a God about whom one can say is right(ful), in an absolute manner.
Suffering cannot be justified	Theists claim that some evils are a necessary condition, or a means to accomplish a greater good. For example, pain is sometimes used for endurance, but some atheists believe that this argument has a boomerang effect, and becomes a challenge to the existence of God, for if suffering could be justified, it would be completely wrong to make efforts to eliminate them. Therefore suffering cannot be justified, nor can the theistic God be said to exist, that new theoretical construct claims.

Another typology of contemporary atheism is generated by the means of presenting the fundamental antinomy within the theoretical body of unbelief. According to some authors, there are four major, clearly delimited and specific antinomies (Geisler, 1993):

a. the antinomy of omnipotence, which argues that a omnipotent God is a contradiction in terms (if there were an almighty God, he could make anything, including a stone so heavy that he were unable to lift, therefore there is no omnipotent God as theists claim).

b. the antinomy of perfection, which argues that God has every perfection, which is mutually exclusive (How could a single person possess both love and hate? Similarly, God cannot be both omniscient and all-loving).

c. the antinomy of creation favours the idea embraces the idea that God is a necessary being, and His will is one with His essence, but based on this statement some argue that what God wants in His will, He must necessarily want; on the other hand, theists claim that God was free not to create, as it is virtually impossible for a creation to necessarily arise from God, and also simultaneously with a void time, and therefore they state: either God is not necessary, or creation is not necessary, in both cases traditional theistic God can not exist.

d. the antinomy of time describes the world as having a beginning in time, an original point, not eternal, except for God, who is eternal; (but if the world had a beginning in time, it must have been a time before the beginning, however it is impossible there to be time before time started, and therefore there cannot be a theist God who created the world in time).

However, those who beleive, globally, for all their diversity and heterogeneity of belief, are far more numerous than those who do not, or refuse to believe in God. During the Renaissance three categories of essential religious populations were defined, Christians, Jews and pagans. Islam, Hinduism, and Buddhism joined, after 1800, Christianity and Judaism, and so the list of "five" has become, from that moment on, the list of the most amply shared religious beliefs in the world. Gradually, five more religions have been included on this list, namely Confucianism, Daoism, Jainism, Zoroastrianism and Shintoism.

Today, Christians are, despite their diversity, the largest religious community in the world (a little more than 33.0%), of which Roman Catholics are the majority, with more than 17.3%, while Protestants represent almost 5.8%, the Greek Orthodox have less than 3.4%, and the Anglicans represent 1.2%. Islamists are actually the second largest religious community worldwide, representing over 20.1% of the world population, Hinduism is placed in third position with 13.3%, and Buddhism is following with 5.9%; other religions represent 12.6%, 12.0% are unbelievers, and atheists represent 2.4%, according to the estimates published on http//:www. adherents.com at the end of 2005. Two strongly different limits are presented in assessing the number of atheists and unbelievers, infidels or people not having a religious creed. David Barrett provides the lower one, placed at about 918.3 million people, when assessing, in both the Encyclopaedia Britannica, and the World Christian Encyclopedia, 2001, according to recent censuses and specialized surveys, own reports and their journeys and investigations "on the place", a total number of people declared atheists over 150.1 million (2.5% of the world population), and about 768.2 million people unbelievers, infidels or having o religious creed (12.7% of the world population). It is the American sociologist Phil Zuckerman who tentatively established the upper limit, when, in one of his analyses published in "The Cambridge Companion to Atheism", edited by Michael Martin, at Cambridge University Press, in 2005, estimated the number of those who "do not believe" at over 1,500 million people...

In order to define or characterize a religion with global (world) incidence, a set of at least five significant factors of global or world impact are considered:

1. the large number of adherents, as most of the main world religions have the main magnitude of their followers into the millions of people;

2. the territorial spreading or dispersion in territorial areas, in terms of a significant number of countries where the major religions of the world have found adherents;

3. the independence from any other religion in the world;

4. the existence of a material body of their doctrine, which comprises, in the vast majority of the cases, a book considered sacred by the adherents;

5. the current practice of the religion, in order for it to be considered a truly major religion (Vlach, 2008).

The main elements that are the object of religions, fashioning their typology into groups or classes, are in turn rather limited, becoming in practice of mere criteria of structuring and classification:

a. belief in a supreme form (monotheism), or several superior beings (polytheism), or a superior form, worshiped among other deities, as if it were unique (henotheism);

b. the number of teachings about the supreme being or superior beings, and the believers' moral and religious duties;

c. the aggregating material nature, or the available expression of the religion represented:

d. the given form of internal worship (fear, love, hope), and external worship (sacrifice, prayers, songs, dances, etc.).

One can estimate that there are now over 10,000 religions in the world. Among the most popular religions of the world in terms of a single founder, of the adherents or followers, and their historical impact remain Zoroastrianism, Buddhism, Judaism, Shintoism, Christianity, Islam, Taoism, Confucianism, Hinduism, the most common of which is Christianity.

3. METHODOLOGY OF THE ASSOCIATIVE STATISTICAL PROCESS

The associative statistical approach follows the conceptualization of a scientific thinking and research, according to which statistics is a thinking system focusing on observation, treatment and research, the cycle involving five practical stages: problem definition → programme of investigation → resulting data → analysis and interpretation of data → conclusions. This view was formulated by Maxine Pfannkuch and Chris Wild from Auckland University, New Zealand, redefining the content of contemporary statistical research in the new millennium.

The main source used in the investigation is the CIA report is the year 2005/2006, entitled The World Factbook (the report has been available since 1981, initially referring to only 165 states, in a volume of only 225 pages, since 1996 it appeared on CD-ROM, and since 2005 it has comprised more than 250 countries and 700 pages of indicators, for which the data were used from the middle of the first decade of this new century, accessible on: http//www.cia.gov/library/publications/the-world-factbook / index.html or http://education.yahoo.com.reference/factbook /pk), as well as other specialized site like U.S. Census Bureau (http://www.census.gov/cgi-bin/ipc/idbagg and on http://www .adherents.com). The methods used in the statistical investigation are the correlation or identification and evaluation of the significant factors by their statistical association with a resultative variable of the GDP type.

The option for a Yule association coefficient, to the detriment of the contingency coefficient is related to the simplicity of calculations. Association, defined by G.U. Yule and M.G. Kendall as a situation where one variable cannot occur without the presence of another explanatory variable, while the other one may exist in an absolutely independent manner, is measurable through two classic ways derived from a table of double entry, maximally simplified (based on average values, defining, in a balanced manner, the populations into two groups, or proeeding from two opposite statuses "non alpha and alpha"):

The classic calculation procedure for a tetrachoric "r", or association coefficient

Table no. 3.

x_i	y_i		Total
	y_1	non y_1	
x_1	a	b	a+b
non x_1	c	d	c+d
Total	a+c	b+d	a+b+c+d

Three constructions of Yule interdependence coefficients dominate the practice of association in the area of the alternative variables and the variables converted from numerical into binary, with the average (variants below average becoming alpha, and the equal and above average variants, non alpha):

$$Q_1 = (ad-bc):(ad+bc) \quad (1)$$

$$Q_2 = \left[1-\sqrt{(bc):(ad)}\right]:\left[1+\sqrt{(bc):(ad)}\right] \quad (2)$$

$$Q_3 = \left[\sqrt{(ad)}-\sqrt{(bc)}\right]:\left[\sqrt{(ad)}+\sqrt{(bc)}\right] \quad (3)$$

although none of them mkes a structural causal distinction between the forms of complete association of a positive and negative nature, detailed for the three specific situations (either in positive form b = c, c = 0, or b = c = 0, or in a negative form a = 0, d = 0, or a = d = 0). However, the first method of calculation has remained in place as the fastest and easiest way to identify the an association and the intensity of an association.

The contingency coefficient proposed by Karl Pearson and determined by the relationship:

$$Qc = \sqrt{\phi^2} = [(ad)-(bc)] : \left[\sqrt{(a+d)(a+c)(b+d)(c+d)}\right] \quad (4)$$

or the Ciuprov coefficient, usually noted by T, whose method of calculation is:

$$T^2 = \phi^2 : \left[\sqrt{(k-1)(h-1)}\right] \quad (5)$$

although belonging to the same range of values, [-1,1], are less used, compared with the Yule coefficient, in the measurements aimed at identifying and ranking correlations through simple associations. The criterion of complete lack of association, or of independence, in all the variants of calculation used, is limited, after processing, to equal proportions ad = bc, and the association is a special case of contingency, when k = h = 2 şi $T^2 = Ø^2$, for square contingency. In this paper, the image of the method of Yule association coefficient in its Q1= (ad-bc):(ad+bc) variant was considered significant, being selected as the principal evaluation of the associative statistical approach. After an overview of the associations, using as a measure of explanatory and explained variables, the umber of states, it was found that the structural influences are better reflected, and have much more relevant final values if the number of inhabitants is turned to account.

4. GENERAL RESULTS OF ASSOCIATION QUANTIFICATION

An overview of world population, in relation to the major religions and adherents to these religions, according to the data in the CIA report entitled The World Factbook, presents the following detailed significant features:

- from a total of 251 states of the world, 226 were provided, in the report, with relevant data, and 221 religious countries of the world have been identified, out of which in 205 one religion is dominant, beloning to the broad scope of the four major groups, selected according to level of adhesion, i.e. at least 5% of all religious population (Christianity in 141, Islam in 51, Hinduism in 3, and Buddhism in 10), and 16 countries where the representative weight is held by indigenous religious groups, or is shared, fairly evenly, between far more religions;
- out of the 221de states dominated by religion, globally (203 in the end, according to data of CIA Report, where are 208 countries with complete and comparable data in point of time reference universe), a total of 194 countries of the world are mono-religious (Christianity in 125, Islam in 41, Buddhism in 8, Hinduism in 3, though there are comparable data only for 189 states), and 44 multi-religious nations (Christianity is relatively dominant in 16 countries, Islam in other 10, Buddhism in 2, while the remaining 16 states belonging to other indigenous religions, or do not have a dominant religion among those considered major, and out of the total number, only 14 provide complete data);
- five states are excessively dominated by atheism, and the lack of religiosity with the great majority of their population, viz. over 1.4 billion inhabitants worldwide;

The main religious structures in world population

Table no.4.

	Population with GDP≥ 10.200$*	Population with GDP < 10.200$*	Total number of inhabitants
Number of religious inhabitants, from which:	1339071039	3598688984	4937760023
Christianity, from which:	1058139076	804662384	1862801460
- Catholicism	491961713	582681939	1074643652
- Protestantism	319392447	45153337	364545784
- Orthodoxy	18073425	47008019	65081444
Islam – Total number	67393066	1380118235	1447511301
Hinduism - Total number	1240827	1123639142	1124879969
Buddhism – Total number	132408886	157598323	290007209
Atheists and unbelievers	153271889	1421574101	1574845990
Population – Total number	1492200034	5020178682	6512378716

**Note-GDP is expressed in $ at purchasing power parity (PPP)*

- a weight assessment, with shifts per cent, through the concentration of world population into just five major religions, gathering only 194 states prevalently defined by a single religion (including atheists and the category of declared non-religious people) and not exhaustive, as in the previous table, reveals an interesting structural hierarchy as a general tendency to concentrate in three major directions: Christianity 29%, atheists and nonreligious 27%, and 23% Islam, covering the Pareto optimum, 20/80;

Major religious groups, concentrated by state, according to the criterion of the majority of their members, with additional information on GDP and life expectancy

Table no. 5

Major religious groups	States	Number of adherents	Structure (%) in the five groups[a]	Structure (%) mini mal structure of adherents	GDP ($PPP / loc)	Life expectancy - years -
Christianity	126	1535281390	28.9	41.3	16180	71.99
Atheists and unbelievers	5	1415851205	26.7	59	7757	72.15
Islam	50	1206708912	22.7	47	3748	61.34
Hinduism	3	905153393	17.1	48	3749	64.60
Buddhism	10	246625108	4.6	42.5	17227	73.26
Total [a]	194	5309620008	100.0	-	10200[b]	64.77[b]

[a] Note - the structure was determined from the aggregate population of the group of the major religions (including the atheists and non-religious people, but excluding the states without an obvious religious dominant)

[b] Note - the final average data refer to the aggregate population, and reflects a polarization of the rest of the world's population not included the table, populations benefiting from a GDP between $500 and $3,000 / capita and average life expectancy between 35 and 50 years (belonging to populations where the majority are religions with fewer followers).

- the heterogeneity of populations of adherents to the great religions is different, being relatively highlighted by the magnitude of some of the variables, stressing the higher relative homogeneity of atheists and non-religious people, as well as Hinduism, in contrast with the maximal heterogeneity of Christianity;

The maximum and minimum GDP and life expectancy limits, in the major religious groups

Table no. 6

Major religious groups	GDP* ($PPP / capita) Minimum	GDP* ($PPP / capita) Maximum	Life expectancy - years - Minimum	Life expectancy - years - Maximum
Christianity	600	71400	32.62	83.51
Atheists and				

unbelievers	3100	21900	67.08	76.22
Islam	600	49700	40.22	78.4
Hinduism	1500	13700	60.18	72.63
Buddhism	1400	33100	54.78	82.19
Total value	600	71400	32.62	83.51

**Note-GDP is expressed in $ at purchasing power parity (PPP)*

- the economic and demographic consequences of the great religions are different, and their economic impact is polarizing (the average index of religious cohesion, determined as the ratio of the maximum GDP / capita and the minimum GDP / capita, is 459.6%, against a background of concentration of the major religious groups in only five categories, and the average life expectancy gap is circa 12 years, according to table no. 4), with maximum amplitude accents within the groups (the internal coefficient of cohesion in the groups ranging from 7.07 to 119.0, being determined similarly, and the internal difference of life expectancy ranges between 9.14 and 50.89 years, in keeping with table no. 6);

- econimic growth, although homogeneous at the level of the major religious groups, is heterogeneous in most groups of maximal amplitude, belonging to Islam, and the spectrum magnitude, from negative to positive values, pertains to Christianity;-

Maximum and minimum economic growth (rate index real GDP), per denomination

Table no. 7

Major religious groups	Economic growth rate index -%- ($I^{REAL\ GDP}$ – 100)	Economic growth (%)	
		Minimum	Maximum
Christianity	4.65	-4.4	18.6
Atheists and unbelievers	10.2	1.8	10.7
Islam	6.06	6.4	34.5
Hinduism	9.01	1.9	9.2
Buddhism	3.75	2.2	8.8

- the age structure as a whole reveals an aggregate population, which is still embraced by demographical appellation of young population, with only 7.3% of the population "5 years and above" (the theoretical threshold of aging is considered even 7%), with 64.9% population between 15 and 64, and 27.8% of the population below 14%, but distributed differently in relation to GDP per capita, in as many as 208 countries of the world, see table. 7;

Structure by age groups and in relation to the average GDP of 208 countries worldwide, dominated by religious faith, and their inhabitant

Table no. 8

Group-years-	GDP≥ 10.200 ($PPP/ capita)		GDP*<10.200 ($PPP/ capita)	
Group 0 - 14	*≥ media 2780%*	*< media 2780%*	*≥ media 2780%*	*< media 2780%*
Inhabitants	70949298	224807028	1082298806	435916188
States'number	14	65	97	32
Group 15 - 64	*≥ media 6490%*	*< media 6490%*	*≥ media 6490%*	*< media 6490%*
Inhabitants	833846919	166154688	1624738481	1599554205
States'number	66	13	37	92
Group ≥ 65 ani	*≥ media 730%*	*< media 730%*	*≥ media 730%*	*< media 730%*
Inhabitants	185515378	11069857	124238883	124001715
States'number	58	21	19	110

**Note-GDP is expressed in $ at purchasing power parity (PPP)*

- as can be seen, there occurs a significant structural differentiation in relation to macroeconomic outcome (GDP), countries with a GDP below the average of $ 10.200 PPP / capita form a more homogeneous population under 15 years, presents an important extension of the share of population covered by the group under 15, while those with a GDP above average constitute a more homogeneous population over 65 years, and reflects an increased aging process.

The associations originally made between the numbers of states in keeping with religious dominance are relatively inconsistent. Changing the measuring unit from number of states to population, the adherents to one of the major religious groups, increased the accuracy of the associations, combine homogeneously and provide a thorough measurement, integrating almost completely the specific structural elements. The new comparable combinations, entirely based on the data from the 208 resulting countries, display high levels of the Yule coefficients, and are converted into some interesting oints of view, statistically substantiated in Table 9, which focuses the coefficients determined in the context of the new unit, the number of adherents being expressed in millions.

Table no. 9

The association between religiousity and economic level (GDP)			
	Religious	Non religious	Total
GDP≥10200$*	1130.8	96.1	1226.9
GDP<10200$*	3061.2	1319.8	4381.0
Total	4192.0	1415.9	5607.9
Q_1 = (ad-bc):(ad+bc) = - 0.671			
The association between Christianity and economic level (GDP)			
	Christian	Non Christian	Total
GDP≥10200$*	876.5	254.3	1130.8
GDP<10200$*	665.8	2395.4	3061.2
Total	1542.3	2649.7	4192.0
Q_1 = (ad-bc):(ad+bc) = 0.851			
The association between Orthodoxy and economic level (GDP)			
	Orthodox	Non Orthodox	Total
GDP≥10200$*	18.1	1040.1	1058.1
GDP<10200$*	47.0	757.7	804.7
Total	65.1	1797.7	1862.8
Q_1 = (ad-bc):(ad+bc) = - 0.562			
The association between mono/ multireligious and economic level (GDP)			
	Mono religious	Multi religious	Total
GDP≥10200$*	1052.0	78.8	1130.8
GDP<10200$*	2785.5	275.7	3061.2
Total	3837.5	354.5	4192.0
Q_1 = (ad-bc):(ad+bc) = 0.138			
The association between Catholicism and economic level (GDP)			
	Catholic	Non Catholic	Total
GDP≥10200$*	491.9	566.2	1058.1
GDP<10200$*	582.7	222.0	804.7
Total	1074.6	788.2	1862.8
Q_1 = (ad-bc):(ad+bc) = - 0.503			
The association between Protestantism and economic level (GDP)			
	Protestant	Non Protestant	Total
GDP≥10200$*	319.4	738.7	1058.1

GDP<10200$*	45.2	759.5	804.7
Total	364.6	1498.2	1862.8

Q_1 = (ad-bc):(ad+bc) = 0.758

The association between Islam and economic level (GDP)

	Islam	Non Islam	Total
GDP≥10200$*	65.7	1065.1	1130.8
GDP<10200$*	1077.6	1983.6	3061.2
Total	1143.3	3048.7	4192.0

Q_1 = (ad-bc):(ad+bc) = - 0.796

The association between Buddhism and economic level (GDP)

	Buddhist	Non Buddhist	Total
GDP≥10200$*	109.2	1021.6	1130.8
GDP<10200$*	137.4	2923.8	3061.2
Total	246.6	3945.4	4192.0

Q_1 = (ad-bc):(ad+bc) = 0.389

The association between age group 15-64 years (average value is 64.9%) and economic level (GDP)

	% ≥ 64.9	% <64.9	Total
GDP≥10200$*	833.8	166.2	1000.0
GDP<10200$*	1624.8	1599.5	3224.3
Total	2458.6	1765.7	4224.3

Q_1 = (ad-bc):(ad+bc) = 0.663

The association between Hinduism and economic level (GDP)

	Hindu	Non hindu	Total
GDP≥10200$*	595.6	1130,2	1130,8
GDP<10200$*	904.7	2156,5	3061,2
Total	905.3	3286,7	4192,0

Q_1 = (ad-bc):(ad+bc) = -0 997

The association between age group 0 -14 years (average value is 27.8%) and economic level (GDP)

	%≥ 27.8	%<27.8	Total
GDP≥10200$*	71.0	224.8	295.8
GDP<10200$*	1082.3	435.9	1518.2
Total	1153.3	660.7	1814.0

Q_1 = (ad-bc):(ad+bc) = - 0.774

The association between age group ≥ 65 years (average value is 7.3%) and economic level (GDP)

	% ≥7.3	%< 7.3	Total
GDP≥10200$*	185.5	11.1	196.6
GDP<10200$*	124.0	153.5	277.5
Total	309.5	164.6	474.1

Q_1 = (ad-bc):(ad+bc) = 0.908

The association between life expectancy and economic level (GDP) is, as is otherwise recognized in calculating the human development index (HDI), almost complete;

Table no. 10

The association between life expectancy (LE) and economic level (GDP)

	LE≥64,77	LE<64,77	Total
GDP≥10200$*	1445.8	46.4	1492.2
GDP<10200$*	2711.5	2308.7	5020.2
Total	4157.3	2355.1	6512.4

Q_1 = (ad-bc):(ad+bc) = 0.873

The synthesis of the statistical associations between religion and wealth with a comparability ensured methodologically, resulting from the use of two concepts embodied in different measuring units (number of states and people), tested and validated with *z* test, looks like this:

Table no.11

The associations' type	Association's coefficient based on the number of states	Association's coefficient based on the number of inhabitants
Religious/non religious variable and economic level (GDP)	Q_1 = *-0.043*	Q_1 = *0.671*
Monoreligious/multireligious variable and economic level (GDP)	Q_1 = *0.405*	Q_1 = *0.138*
Christian / non- Christian variable and economic level (GDP)	Q_1 = *0.552*	Q_1 = *0.851*
Orthodox /non-Orthodox variable and economic level (GDP)	Q_1 = *-0.487*	Q_1 = *- 0.562*
Catholic/ non-Catholic variable and economic level (GDP)	Q_1 = *0.317*	Q_1 = *-0.503*
Protestant/ non-Protestant variable and economic level (GDP)	Q_1 = *0.317*	Q_1 = *0.758*
Islam/non- Islam variable and economic level (GDP)	Q_1 = *-0.544*	Q_1 = *-0.796*
Hindu/non- Hindu variable and economic level (GDP)	Q_1 = *-0.299*	Q_1 = *-0.997*
Buddhist/non- Buddhist variable and economic level (GDP)	Q_1 = *-0.184*	Q_1 = *0.389*
The group of age 0 -14 variable and economic level (GDP)	Q_1 = *-0.867*	Q_1 = *- 0.774*
The group of age 15 -64 variable and economic level (GDP)	Q_1 = *0.853*	Q_1 = *0.663*
The group of age ≥ 65 variable and economic level (GDP)	Q_1 = *0.882*	Q_1 = *0.796*
The life expectancz (LE) variable and economic level (GDP)	Q_1 = *0.844*	Q_1 = *0.873*

The majority religion in the world remains the Christian type, which encourages marriage and child bearing. It is against contraception and divorce. These issues should be relevant as far as the rate of marriage, birth, contraception and divorce are concerned. The comparative data reveal the following aspects:

Marriage rate per 1,000 inhabitants, for the first typologies of religions

Table no.12

Religion	Demographic indicator
Atheists and unbelievers	6,8
Buddhism	5,7
Christianity	5,5
Judaism	6
Hinduism	8,7
Islam	8,9
Other different religions	*

**lack of homogeneous data*

At first glance, it is at least curious as far as Christians are concerned, because their religious precepts encourage le marriage, considering that it is not good for man to live alone. Let us not however forget that most Christians are in Europe, whose trend towards demographic implosion and accelerated aging is unique worldwide, among all the continents...

Birth rate per 1,000 inhabitants, for the first typologies of religions

Table no. 13.

Religion	Average value	Minimum value	Maximum value
Atheists and unbelievers	12.36	9	16.6
Buddhism	16.47	7.3	35

Christianity	19.64	8.2	48.1
Judaism	17.7	-	-
Hinduism	22.83	15.3	30.5
Islam	29.53	9	50.2
Other different religions	37.4	36.8	38.6

Note: the sign "-" indicates that it is not the case to make the determination (i.e. there is a single value)

We could exaggerate and say that this category includes people who want to be successful professionally and materially first, considering childcare a barrier to personal fulfillment. Christianity is placed behind Islam, even if the interval including the minimum and maximum values records them quite close. The other religions have the largest number of potential children

Fertility rate per 1,000 people, for the first typologies of religions

Table no. 14

Religion	Demographic indicator
Atheists and unbelievers	1.7
Buddhism	1.9
Christianity	2.7
Judaism	2.8
Hinduism	2.7
Islam	3.6
Other different religions	4.3

The birth rate indicator closely follows the normal population fertility, religion and faith being expressions of certain demographic policies, no less than contraception, which has become the result of atheism or clear lack of religiousness...

Contraceptive methods in the first-ranking religious types

Table no. 15

Religion	Number of cases of applied contraceptive methods (%)
Atheists and unbelievers	87.67
Buddhism	59.84
Christianity	60.06
Judaism	68.00
Hinduism	56.12
Islam	45.89
Other different religions	27.53

Atheists use the most contraceptive methods, and also, surprisingly enough, Judaism. The rates are relatively high for Christianity. Certainly the degree of economic and cultural level influence, in developed countries, the application of these methods, which eventually have higher and closer-ranging values.

Divorce rate for the first religion typologies

Table no. 16

Religion	Divorce rate to 1000 inhabitants	Divorce rate to 100 marriages
Atheists and unbelievers	2.77	44.43*
Buddhism	1.24	1.9
Christianity	1.74	24.42
Judaism	1.56	30.1
Hinduism	0.87	**
Islam	1.53	9.06
Other different religions	**	**

**China is not included in the calculation due to lack of specified data. **Missing data*

Through natural association, the highest rate of divorce is noted with atheists (unbelievers and unchurched or non-religious people). The following places are held by Judaism and Christianity.

Life expectancy in the first typologies of religion

Table no. 17

Religion	Life expectancy (LE)		
	Average value	Minimum value	Maximum value
Atheists and unbelievers	72.15	67.08	76.22
Buddhism	73.26	54.78	82.19
Christianity	71.99	32.62	83.51
Judaism	64.60	60.18	72.63
Hinduism	61.34	40.22	78.40
Islam	79.78	-	-
Other different religions	64.77	32.62	83.51

Note: the sign "-" indicates that it is not the case to make the determination (i.e. there is a single value)

Finally, it can be noticed that religion influences and stratifies the world's population according to the demographic indicator of life expectancy in a significant manner. The way demography is reflected in the world of religions defines a normal aspect of reciprocity. In conclusion, we many religious typologies are revealed, as well as clusters and discontinuities of adherence and classes of beliefs, according to the various and unsuspected religious traditions and customs of the peoples.

6. SOME USEFUL CONCLUSIONS AND SOME SUSTAINABLE DEMOGRAPHIC IMPLICATIONS

The two categories of associative evaluations are based on the idea that there exists a state policy, supported by the Constitution, as far as religious freedom and expression are expressed (the measure being the state), with immediate economic consequences, and also on the substance and homogeneous geographical distribution of the major religious groups (the unit being, in this latter case, the inhabitant adhering to such group).

I. One can state with certainty that there is a connection of a statistical type, identified as a simple association between religiosity and economic welfare, through the level of development (GDP). From the general picture of the confrontations there also occur abnormal situations between the two types of measurements, generated by geography of the spreading pattern of certain religions (especially Buddhism, where also occurs the alternating sign of multi-religious people, without however an obvious structural dominant).

II. One can certify a direct link, of average intensity in keeping with the number of states, and a link of strong intensity, if the value of adherents is capitalized on, as number of inhabitants, between the religious status of Christianity (as an alternative explanatory variable, with the states of being Christian of being non-Christian) and the economic result (materialized in GDP, and turned into an alternative resultative variant, in relation to the average \$10,200 PPP/capita). By working on, and refining the details, some statistical associations reveal unexpected variability of the intensities and signs or directions of association: in Christianity, according to the second, larger and more homogeneous approach, two different trends can be perceived, namely the first one: inverse association between Catholicism and the economic level (Q1= -0.503), and between Orthodoxy and the GDP level (Q1 = -0.562), and a second one: direct association between Protestantism and welfare (Q1 = 0.758).

III. An indirect association of particular complexity and almost maximal intensity occurs between Islam and Hinduism as religious status and wealth (through macroeconomic outcome) between Islam and the GDP, the inverse associationof high intensity and confirmed by both approaches (Q1 = -0.796) and betweenHinduism and welfare surprise occurs in the functional maximal inverse association (Q1 = -0.997).

IV. Between Buddhism and the GDP, due to its heterogeneity and specific geographical spreading, totally opposite signs appear (Q1 = -0.184 and Q1 = 0.389), as the adherents of this religion see the strongest polarization, in point of both life expectancy and GDP.

V. Between life expectancy and the economic level (GDP), as was expected, the association is virtually complete (Q1 = 0.873).

VI. Also, natural relations can be identified, of direct association between the degree of aging of the population and the economic level (GDP), with a very high coefficient of association, Q1= -0.796, and differentiated as indirect association between the high structural level of rejuvenation of the population and GDP, namely Q = - 0.774.

VII. Potential statistical and demographic associations can be identified and quantified, under the influence of religious phenomena such as the rate of marriage, fertility, birth, divorce, etc.

VIII. Several models can be proposed for estimating the GDP of a territorial aggregate where the great world religions are found.

A. A simple modelling, described in probabilistic terms exclusively by membership to one of the major religious groups, could lead to some improvements on GDP estimates, which are quite interesting in economic and statistical practice.

$$L=\begin{pmatrix} \text{christian} & \text{atheist} & \text{islam} & \text{hindu} & \text{buddhist} \\ 28.9 & 26.7 & 22.7 & 17.1 & 4.6 \end{pmatrix}, \text{ where}$$

the sum of the probabilities is: $\sum_{i=1}^{n} p_i = 1$.

B. Another modelling, described in terms of a simple lottery model, can assess the probability of achieving a certain level of GDP, based on its relative modification, described in terms of probabilities:

Δfinal GDP = $\sum\left[\Delta_I \times Q_I \times PIB_I\right]$ = (ΔChristians 0,851 GDP / Christian people) + [Δ Islam (-0.796) GDP / Islamic people] + [Δ Hindu (-0.997) GDP/Hindu people] + (ΔBuddhists 0.389GDP/Buddhist people) + (Δatheists 0.671 GDP /atheistic people).

C. A practical integrative model for estimating the GDP can be obtained through a type of modelling based on the composite Lq lottery solution, including other associated factors (aging or rejuvenation of the population, life expectancy etc., together with religious beliefs or denomination); it can be formalized as follows: $Lq = \alpha_1 p_1 + \alpha_2 p_2 \ldots + \alpha_n p_n$, and can be exemplified by a double lottery, according to the probabilities of religion and age group (here, with the specific variants limited only to Christianity), etc.

$$Lq = \begin{pmatrix} christian, & christian, & christian & \\ 28.9 & 28.9 & 28.9 & \cdots \\ 0.20 & 0.67 & 0.13 & \cdots \end{pmatrix}, \text{ its development}$$

capitalizing on the probabilities given by the following structure, refined and detailed by age and level of GDP per capita.

Table no. 18. Part I

	The group of age	Catholics	Orthodocs	Protestants	Christians
PIB>10200	0-14 years	20.85	14.28	20.43	20.35
	15-64 years	66.22	67.58	66.94	66.44
	>64 years	12.93	18.15	12.63	13.21
	Total	100.00	100.00	100.00	100.00
PIB<10200	0-14 years	33.57	16.85	40.58	32.90
	15-64 years	61.67	69.23	56.52	61.42
	>64 years	4.77	13.92	2.90	5.68
	Total	100.00	100.00	100.00	100.00

Table no. 18. Part II

Islam	Hindu	Buddhists	Atheists	Total
34.90	24.40	14.37	14.59	19.81
61.67	69.10	66.56	71.21	67.02
3.43	6.50	19.07	14.20	13.17
100.00	100.00	100.00	100.00	100.00
35.52	31.40	27.15	21.82	30.24
60.32	63.73	66.82	70.68	64.23
4.16	4.87	6.03	7.50	5.53
100,00	100,00	100,00	100,00	100,00

The model focused on such a compound lottery can easily become a model of a triple, quadruple, etc. nature (by continuously multiplying the number of factors), as other associations, statistically measurable are added, revealing new factors selected with their probabilities of occurrence (the value of the association coefficient being the factorial criterion ranking). Demographically, any religion also has specific influences. Certainly, in terms of marriage rates, divorce, fertility or the number of contraceptive methods, there are many factors that influence whose intensity is ever greater at the expense of religious morality and teachings. Christianity, the most widely spread of the religions in the world, does not have the overwhelming influence of Islam, for instance...

Marriages are made at increasingly advanced ages, and in a decreasing number. Bringing up children is considered an obstacle to the individual's self-assertion. People are becoming less tolerant, which results in more easily dissolving marriages. Contraceptive methods are increasingly used to prevent child birth. Consequently, even if religion has relatively diminished its influence on society, it is still an important factor, both demographically and in point of morality and peersonal identity, allowing meditation and self-retrieving, in a world of sheer materialism, more efficient economically to the detriment of the spiritual aspects.

Based on the data in the Demographic Yearbook of Romania published in 2001, covering a period of nearly 50 years, and also in an attempt to to discern associations and correlations with a lag, beyond the discussions, commented on in the introduction, of young people about the difficulty of booking a restaurant for an ordinary wedding party, or the rather difficult scheduling a religious marriage service in one of our Greek Orthodox churches, correlated with all the other events pertaining to an ordinary marriage, we tried to answer, demographically and statistically – while not ignoring the obviously economic, or marketing, or demographic policy facets – the natural question about a possible link between the seasonality of marriage and the type of religion where it is materialized, as well as a certain distribution of births in relation to the territorially dominant religion.

To do that, we will have recourse to the three long fasting periods in the Greek Orthodox calendar, placed in three significant periods of the year (astronomical spring, the hot weather period in August, and Santa Claus's winter). Our hypothesis is that marriages that fall outside the limits of traditional fasting periods, and also the births, follow the classical trajectory of a lag correlation, theoretically shaped for the first time by (Morice and Chartier, 1964):

$$r_h = \frac{\sum_{h=1}^{n}\left(xi - \overline{x}\right)\left(y_{t-h} - \overline{y}\right)}{\left(n-h\right)\sigma_x\sigma_y},$$

, where t is the number of terms of the original series, and h the number of terms which differentiate, as lags, the second series.

The statistical calculation shows a value that is practically identical with nine, as the number of years or terms of the series increases, so a gap of nine months after marriage night ritual or custom... The capacity that statistical diagrams have to reveal hidden phenomena is also present in this beautiful and unexpected story of the relationship between religion, calendar and demographics. Who has eyes to see, let them look carefully at the following historiogramme as a special chronological graph:

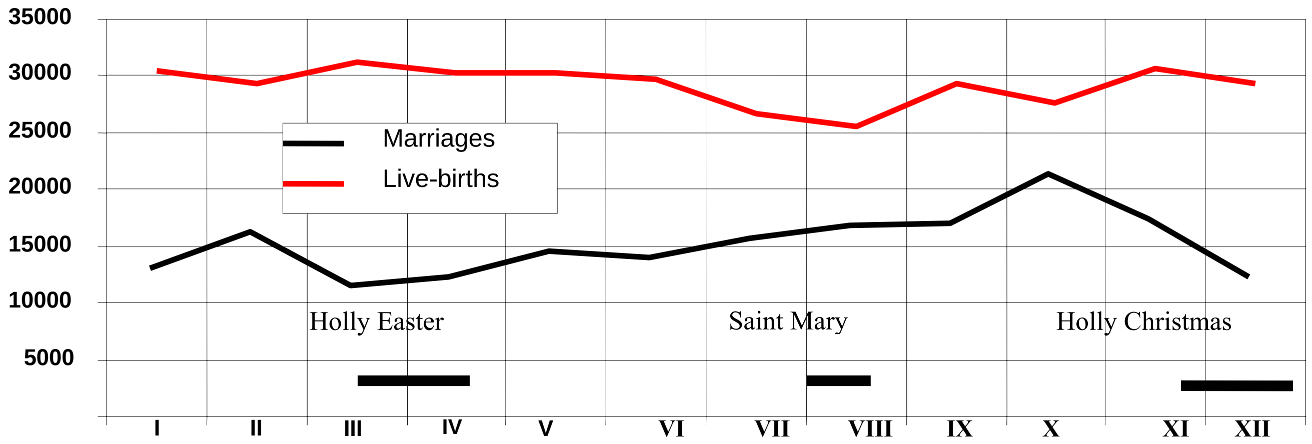


Figure no. 1. Marriages and Live-births related to religious major lents to Orthodocs

Three conclusions are clearly distinguished in their association, which are highlighted quite clearly in graph form:

A. The bottom winding line represents the number of marriages during an average year, determined according to demographic data from the period 1957-2000, available in the Demographic Yearbook of Romania, published in 2001 by INS (National Institute of Statistics), Bucharest. During the major fasting periods in spring and winter, the line goes down to the lowest levels, while in midsummer the upward slope is milder... So, Romanians are mostly Orthodox, observing the restrictions imposed by fasting in the processes related to marriage.

B. At about nine months from the graph peak of marriages, or their maximum number, three significant peaks, like three needle-stitches in the chart, stand out on the top line, that of the number of children born live: in September, 9 months after Advent, in November to December, again 9 months, this time after Easter fasting, and in March-April, evidently after nine months, but after the fasting period of Saint Mary's celebration. Romanians still observe the custom of the first night of their marriage, as well as the restrictions of the religious calendar..

C. The distance between the two lines is relative, still it is placed at about 1/3, or at most half the value recorded by the level of marriage, in the average value of the measurements. In other words, Romanian couples who bring children into the world before marriage or outside it is 1/3, or at most half of the total number of Romanians who meet the precepts of their Orthodox religion.

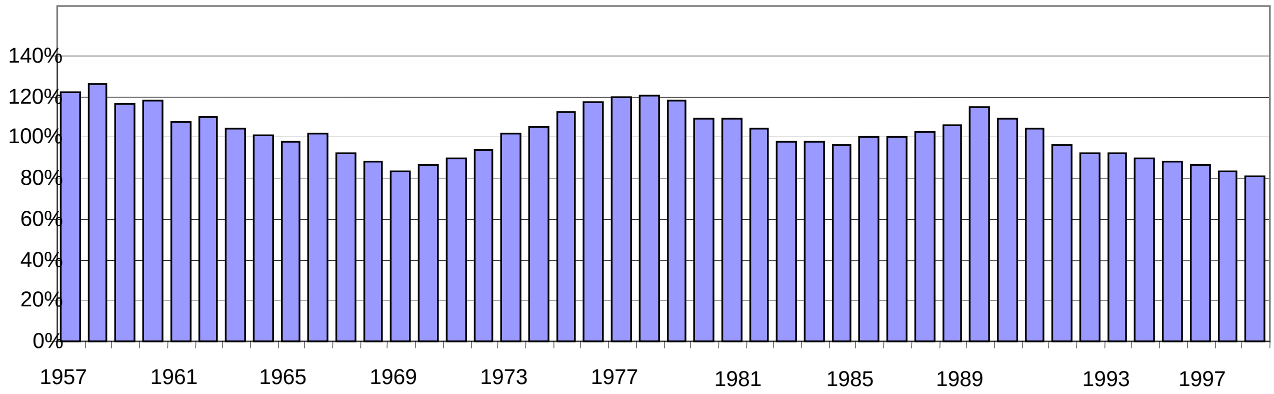


Figure no. 2. Marriage 's Statistical Index between 1957 and 2000, in Romania

Who could guess how many other influences have been maintained in relation to religion in the demographic evolution in the last half of the twentieth century, by the nuptial (or matrimonial) behaviour, and defined by this religious birth calendar? At the height of the atheist period, that was however a purely declarative one, under the influences of Marxism, the Orthodox Christians of the Eastern / Greek type, in our case Romania, remained the same unwavering traditionalists.

But does that not raise the issue of modernity and adaptability in the European Union, requiring the reconsideration of religion as an essential factor in multi- and trans-disciplinary approaches? Success, through tradition, in preserving national, and thus religious identity, especially the adequacy of other components of the human behaviour to modernity, reshapes human behaviour, making it similar to that of the rest of the population of the old, and also Christian Europe, of the inhabitants of planet Earth in general.

On a long term and very long, in the same global plane, there are two terrifying demographic projections in full contrast. These demographic projections, apparently absurd if considered with respect to time horizon and accuracy, belonging to François Héran, Director of the French National Institute of Demographic Studies (INED), focus on a diminished upward, or a steep downward trend; the former describes a positive rate of demographic surplus, which is however in decline, leading, in the year 2300, to a world population, still explosively booming, of about 36.4 billion people, and the latter, with a pessimistic level, almost stationary around 2075, a year that would correspond to a historic threshold of 9.2 billion people, a projection likely to be characterized by trends of early decline after the year 2100, and a decline severely installed beyond 2300, with a population of the Earth estimated at 2.3 billion inhabitants. A demographic century polarized and potentially conflicting, which is looming as a serious demographic reality, in the absence of an effort of the world religions to join in a desperate attempt to change, in a global and harmonized manner, the people's welfare, an alarming and uncertain century in the evolution of the human species...

7. A JUSTIFIED FINAL NOTE

Religion is still a "unified system of beliefs and practices relating to things considered sacred or well isolated and kept, as well as prohibited beliefs together with practices that finally cause a united, homogeneous community to become an entity from a moral standpoint". This classic definition given by Emil Durkheim brings about the conclusion that all religions also shape a specific attitude towards work and towards welfare. The diversity and specific tolerance of Christianity, evident in the maximal amplitude of its variability, dominates economically and demographically, and gradually decrease through the translation to other religions like Islam, Hinduism, Buddhism, becoming minimum in the area dominated by atheism. Within Christianity there are two contrary associations, one direct and strong, i.e. Protestantism, and another one, indirect and average, i.e. Catholicism and Orthodoxy... Although the importance of old age and its specific wisdom are indisputable, an old proverb like Who does not have an old man ought to buy one becomes an economic impossibility as in the areas where aging of the population is high, the economic level is also on the increase, in other words who has welfare has also secured the so much necessary elderly people.

The paradox outlined by the rapprochement between Buddhism and Christianity, respectively atheism, as variants that command the high level of life expectancy, and the relative level o welfare (GDP), emphasizes the fact that between the three attitudes there is still a special bond, which is not clearly revealed in an exclusively religious, economic or demographic manner. Demographically, religion associates different statistical influences. Certainly, in terms of marriage, divorce, and fertility rates, or the number of contraceptive methods, the influence factors are numerous that have an ever greater intensity at the expense of morality and religious teachings. Christianity, the most widely spread of religions, does not have the overwhelming influence that Islam has, for example...

Christian marriages are contracted at an increasingly advanced age, and in ever decreasing numbers. Bringing up children is considered an obstacle to individual self-assertion, for today's Christian scientist's career. The people of the new millennium, who have become increasingly less religious and tolerant, contribute in excess, and with increasing ease to the dissolution of marriages, while contraceptive methods are increasingly used... Religiosity finally enhances, in a spectacular manner, welfare, too.

A final remark could be that, although religion seems to have declined in its relative influence on society, it will remain the most important demographic factor, both morally and identity-wise, for every individual, an essential landmark in man's spiritual recuperation, in a world of an excessive materialism and declaratively more and more efficient economically.

8. REFERENCES

[1] Borovcnik M., Peard R., (1996), *Probability*, A. Bishop (eds.), *International Handbook of Mathematics Education*, Dordrecht: Kluwer Academic Publishers, pp. 239-288.

[2] Buchet C., (1998), *Religie şi putere în relaţiile internaţionale contemporane*, Bucharest: Editura Didactică şi Pedogogică.

[3] Culianu, I.P., (1996). *Religie şi putere,* Bucharest: Editura Nemira.

[4] Eisenberg J., (1993, *O istorie a evreilor*, Bucharest: Ed. Humanitas.

[5] Eliade, M., Culianu I P., (1996), *Dicţionar al religiilor*, Bucharest: Ed. Humanitas.

[6] Gaudin, P., (2005), *Marile Religii,* Bucharest: Editura Orizonturi.

[7] Geisler, N., (1993), *Apologetică Creştină*, Ilinois, SUA: Ed. Societatea Misionară.

[8] Golopenţia, A., (2002), *Opere complete, vol.II, Statistică, demografie şi geopolitică*, Bucharest: Editura Enciclopedică.

[9] Gheţău, V., (1979), *Perspective demografice,* Bucharest: Editura Ştiinţifică şi enciclopedică.

[10] Hoerl, R.W., Snee, R.D., (2002), *Statistical Thinking: Improving Statistical Performance*, Pacific Grove.

[11] Hexham, I., **(**1999**),** *Concise dictionary of religion*, Vancouver**:** Regent College Press**.**

[12] Kritz. M., Lim, L, Zlotnik, H. (1992), *International migration systems–a global approach*, Oxford: Clarendon Press.

[13] Martin, M.,(2005), *The Cambridge Companion to Atheism*, Cambridge, UK: Cambridge University Press.

[14] Mihălcescu, I., (1946), *Istoria religiunilor lumii*, Bucharest: Editura Cugetarea.

[15] Morice, E., Chartier, F., (1964), *Methode statistique*, Paris: Imprimerie Nationale, pp. 501-508.

[16] Richards, E.G., (1999), *Cronologie şi Civilizaţie*, Bucharest: Editura Tehnică.

[17] Săvoiu, G., Cruceru, G., *Religia şi demografia - factori de risc şi incertitudine ai deciziei şi ai dezvoltării economice, Progrese in teoria deciziilor economice in condiţii de risc şi incertitudine, vol III,* Ed. Performantica Iaşi, pp. 47-61

[18] Săvoiu, G., (2006), *Populaţia lumii între explozie şi implozie demografică,* Bucharest: International University Press.

[19] Trebici V., (1975), *Mică enciclopedie de demografie*, Bucharest: Editura Ştiinţifică şi Enciclopedică.

[20] Ţarcă, M., (1996), *Demografie*, Bucharest: Editura Economică.

[21] Vlach, M. J., (2008), *Dispensationalism: Essential Beliefs and Common Myths*. Los Angeles: Theological Studies Press.

[22] Wild, C. J. and Pfannkuch, M. (1999), *Statistical Thinking in Empirical Enquiry*. International Statistical Review, 67(1), pp. 223–248.

A GENERALIZED CLASS OF REGRESSION TYPE ESTIMATORS IN TWO PHASE SAMPLING

Subhash_Kumar_Yadav[1], Sant_Sharan_Mishra[2] and Alok_Kumar_Shukla[3]

[1,2] Department of Mathematics & Statistics, R M L Avadh University, Faizabad, India, [3] Department of Statistics, D. A. V. College, Kanpur, India
[1]e-mail: drskystats@gmail.com, [2] e-mail: sant_x2003@yahoo.co.in, [3] e-mail: alokshukladav@gmail.com

Abstract. *In the present investigation, a generalized class of linear regression model of estimators has been proposed for estimating the population mean and population total when auxiliary information is not available in survey sampling. The proposed model has lesser mean squared error as compared to ordinary regression method of estimation under two phase sampling scheme. The improvement of the previous estimator has been validated with the help of an empirical data under the aforesaid sampling scheme.*

Keywords: *Auxiliary variable, Mean Squared Error, Regression type estimators, two phase sampling.*

1. INTRODUCTION

The auxiliary information has been in use in sampling theory since the development of the theory and application of modern sample surveys. It is well established that the intelligent use of auxiliary information improves the efficiency of the sampling design by increasing the precision of the estimates. The auxiliary information has been used for the purposes of stratification in stratified sampling, measures of sizes in PPS (Probability Proportional to Size) sampling. In regression method of estimation, one uses auxiliary information so as to improve precision of the estimate of population parameters like population mean and population total etc.

Let y and x be the study and the auxiliary variables respectively. When the variable y under study and the auxiliary variable x is highly correlated and the line of regression of y on x passes through origin, the ratio and product type estimators are used to estimate the parameter under study. When y and x are highly positively correlated and the line of regression of y on x passes through origin, ratio estimator is find better to estimate the parameter under study while the product estimator is used when y and x are negatively correlated. When the regression line does not passes through origin or its neighborhoods, regression estimator is better one to use. In practice it has been seen that the regression line hardly passes through the neighborhoods of the origin, in this situation ratio estimator is not as good as regression estimator. So it is better to seek for some other improved regression type estimator having lesser mean squared error.

Let $U = (U_1, U_2, ..., U_N)$ be the finite population of size N out of which a sample of size n is drawn with simple random sampling without replacement technique. Let y and x be the variable under study and the auxiliary variables respectively.Let $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ & $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$ be the population means of study and the auxiliary variables and $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ & $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ be the respective sample means. When $\bar{X}$ is not known, double sampling or two phase sampling is used to estimate the population mean of the study variable y. Under This sampling technique a large sample of size n' in first phase is taken with simple random sampling without replacement (srswor), without increasing the cost of the survey to estimate $\bar{X}$ and to estimate $\bar{Y}$, a sample of required size n in second phase is drawn.

The linear regression estimate of population mean of variable Y is defined as:

$$\bar{y}_{lr} = \bar{y} + b(\bar{X} - \bar{x}) \tag{1.1}$$

Where $\bar{y}$ *and* $\bar{x}$ are sample mean of variable Y and X respectively. $\bar{X}$ is population mean of auxiliary variable X and is supposed to be known. $\bar{y}_{lr}$ is linear regression estimate of population mean and b is a constant. Sukhatme *et.al.* (1984) have described in detail, the procedures for deriving estimates of population parameters along with their biases, mean square error etc.

But when $\bar{X}$ is unknown, the double sampling version of estimator (1.1) is defined as

$$\bar{y}_{lrd} = \bar{y} - b(\bar{x} - \bar{x}') \tag{1.2}$$

Where $\bar{x}' = \frac{1}{n'}\sum_{i=1}^{n'} x_i$ is the sample mean of auxiliary variable x based on sample of size n'.

The mean square error (MSE) of the estimator (1.2) is

$$MSE(\bar{y}_{lrd}) = \lambda S_y^2 + \hat{\beta}_1^2 \lambda^* S_x^2 - 2\hat{\beta}_1 \lambda^* S_{xy} \tag{1.3}$$

2. A PROPOSED MODEL

Following Misra et.al (2010), we are proposing a generalized class of linear regression type estimator under two phase sampling as

$$\bar{y}_{gld} = \bar{y} - \hat{\beta}_1(\bar{x} - \bar{x}') - \hat{\beta}_2(\bar{z} - \bar{z}') \tag{2.1}$$

Where $\hat{\beta}_1$ and $\hat{\beta}_2$ are the estimates of β_1 and β_2 respectively with $\bar{z}'$ as mean of Z based on sample of size n'. $Z = f(X)$ is a function of auxiliary variable X. When $Z = X^2$, it assumes the double sampling version of relationship considered by Ekpenyong et.al (2008), If $Z = \frac{1}{X}$, it takes the form of Misra et.al. (2009) and many more functions may be considered. It has been shown by Misra et.al (2009) that their estimators of population mean and total are more efficient as compared to estimators of Ekpenyong et.al (2008) and ordinary linear regression estimator. U is independently and identically distributed random variable with mean zero and fixed variance σ^2.

The estimator of population mean based on (2.1) is given by:

$$\bar{y}_{gld} = \bar{y} - \hat{\beta}_1(\bar{x} - \bar{x}') - \hat{\beta}_2(\bar{z} - \bar{z}') + u \qquad (2.2)$$

Where $\bar{y}_{gld}$ is a general regression type estimator of population mean based on proposed model defined in relation (2.1), $\bar{z}$ and $\bar{z}'$ are sample means of the variable Z based on samples of size n and n' respectively.

3. BIAS AND MEAN SQUARED ERROR OF $\bar{y}_{gld}$

To calculate the bias and MSE of $\bar{y}_{gld}$ estimator of population mean, we define.

Let $\bar{y} = \bar{Y}(1+e_0)$,

$\bar{x} = \bar{X}(1+e_1)$,

$\bar{x}' = \bar{X}(1+e_1')$

$\hat{\beta}_1 = \beta_1(1+e_2)$,

$\hat{\beta}_2 = \beta_2(1+e_3)$,

$\bar{z} = \bar{Z}(1+e_4)$ and $\bar{z}' = \bar{Z}(1+e_4')$

Such that $E(e_i) = E(e_i') = 0 \quad \forall\ i = 0,1,2,3,4$ and

$E(e_0^2) = \lambda C_y^2$, $E(e_1^2) = \lambda C_x^2$, $E(e_1'^2) = \lambda' C_x^2$

$E(e_4^2) = \lambda C_z^2$, $E(e_4'^2) = \lambda' C_z^2$ $E(e_0 e_1) = \lambda C_{xy}$,

$E(e_0 e_1') = \lambda' C_{xy}$, $E(e_1 e_1') = \lambda' C_x^2$, $E(e_0 e_4) = \lambda C_{yz}$,

$E(e_0 e_4') = \lambda' C_{yz}$, $E(e_4 e_4') = \lambda' C_z^2$, $E(e_1 e_4) = \lambda C_{zx}$,

$E(e_1 e_4') = \lambda' C_{zx}$, $E(e_1' e_4) = \lambda' C_{zx}$, $E(e_1' e_4') = \lambda' C_{zx}$

Where $\lambda = \frac{1}{n} - \frac{1}{N}$ and $\lambda' = \frac{1}{n'} - \frac{1}{N}$

Putting these values in equation (2.1), we get

$$\bar{y}_{gld} = \bar{Y}(1+e_0) - \beta_1(1+e_2)[\bar{X}(1+e_1) - \bar{X}(1+e_1')] -$$
$$\beta_2(1+e_3)[\bar{Z}(1+e_4) - \bar{Z}(1+e_4')]$$

$$\bar{y}_{gld} = \bar{Y}(1+e_0) - \beta_1(1+e_2)[\bar{X}(e_1 - e_1')] -$$
$$\beta_2(1+e_3)[\bar{Z}(e_4 - e_4')]$$

$$\bar{y}_{gld} = \bar{Y} + \bar{Y}e_0 - \beta_1\bar{X}(e_1 - e_1' + e_1 e_2 - e_1' e_2) -$$
$$-\beta_2\bar{Z}(e_4 - e_4' + e_3 e_4 - e_3 e_4')$$

Taking expectation both sides, we get

$$E(\bar{y}_{gld} - \bar{Y}) = \bar{Y}(e_0) - \beta_1\bar{X}[E(e_1 e_2) - E(e_1' e_2)] -$$
$$\beta_2\bar{Z}[E(e_3 e_4) - E(e_3 e_4')]$$
$$= -\beta_1\bar{X}[E(e_1 e_2) - E(e_1' e_2)] - \beta_2\bar{Z}[E(e_3 e_4) - E(e_3 e_4')]$$

$$Bias(\bar{y}_{gld}) = -[Cov(\bar{x}, \hat{\beta}_1) - Cov(\bar{x}', \hat{\beta}_1)] -$$
$$[Cov(\bar{z}, \hat{\beta}_2) - Cov(\bar{z}', \hat{\beta}_2)] \qquad (3.1)$$

which is negligible for large sample size. For large samples usually $Cov(\bar{x}, \hat{\beta}_1)$ and $Cov(\bar{x}', \hat{\beta}_1)$ decreases and it becomes zero if the joint distribution of y and x is bivariate normal.

Similarly $Cov(\bar{z}, \hat{\beta}_2)$ and $Cov(\bar{z}', \hat{\beta}_2)$ vanishes if the joint distribution of y and z follows bivariate normal distribution. In this case the proposed regression estimator is exactly unbiased.

To the first order of approximation, we have

$$\bar{y}_{gld} - \bar{Y} = e_0\bar{Y} - (e_1 - e_1')\beta_1\bar{X} - (e_4 - e_4')\beta_2\bar{Z}$$

Therefore

$$MSE(\bar{y}_{gld}) = E[e_0\bar{Y} - (e_1 - e_1')\beta_1\bar{X} - (e_4 - e_4')\beta_2\bar{Z}]^2$$
$$= E[e_0^2\bar{Y}^2 + \beta_1^2\bar{X}^2(e_1 - e_1')^2 + \beta_2^2\bar{Z}^2(e_4 - e_4')^2 -$$
$$2\beta_1\overline{XY}(e_0 e_1 - e_0 e_1') + 2\beta_2\overline{YZ}(e_0 e_4 - e_0 e_4') -$$
$$2\beta_1\beta_2\overline{ZX}(e_1 e_4 - e_1' e_4 - e_1 e_4' + e_1' e_4')]$$

Applying expectation, putting different values of expectations and simplifying, we get

$$MSE(\bar{y}_{gld}) = [\lambda s_y^2 - 2\beta_1\lambda^* s_{xy} + \beta_1^2\lambda^* s_x^2 -$$
$$2\beta_2\lambda^* s_{yz} + \beta_2^2\lambda^* s_z^2 + 2\beta_1\beta_2\lambda^* s_{xz}] \qquad (3.2)$$

And we know that

$$MSE(\bar{y}_{lrd}) = \lambda S_y^2 + \hat{\beta}_1^2\lambda^* S_x^2 - 2\hat{\beta}_1\lambda^* S_{xy}$$

Where $\lambda^* = \frac{1}{n} - \frac{1}{n'} = \lambda - \lambda'$ and s_{xy}, s_{yz} & s_{zx} are estimators of the population covariances, S_{XY}, S_{YZ} and S_{ZX} respectively, while the variances s_x^2, s_y^2 and s_z^2 are unbiased estimators of population variances S_X^2, S_Y^2 and S_Z^2 respectively.

Now it is required to estimate β_1 and β_2 in such a way that $MSE(\bar{y}_{gld})$ is a minimum. Using the method of ordinary least square, we differentiate partially (3.2) with respect to $\hat{\beta}_1$ and $\hat{\beta}_2$ and obtain following normal equations.

$$\hat{\beta}_2 s_{zx} + \hat{\beta}_1 s_x^2 = s_{xy} \qquad (3.3)$$

$$\hat{\beta}_2 s_x^2 + \hat{\beta}_1 s_{zx} = s_{yz} \qquad (3.4)$$

Solving (3.3) and (3.4) simultaneously, we get

$$\hat{\beta}_1 = \frac{(s_{yz}s_{zx} - s_{yx}s_z^2)}{(s_{xz}^2 - s_x^2 s_z^2)} \qquad (3.5)$$

$$\hat{\beta}_2 = \frac{(s_{yx}s_{zx} - s_{yz}s_x^2)}{(s_{xz}^2 - s_x^2 s_z^2)} \qquad (3.6)$$

Thus for these values of $\hat{\beta}_1$ and $\hat{\beta}_2$, the $MSE(\bar{y}_{gld})$ will be minimum.

The estimate of population total (y_{gld}) and its variance using proposed estimator $\bar{y}_{gld}$, are as follows:

$$y_{gld} = N\,\bar{y}_{gld}$$

$$MSE(y_{gld}) = N^2 MSE(\bar{y}_{gld})$$

4. EFFICIENCY COMPARISON:

In view of (1.3) and (3.2), we have

$$MSE(\bar{y}_{lrd}) - MSE(\bar{y}_{gld}) = \lambda^*\left[2\hat{\beta}_2 s_{yz} - \hat{\beta}_2^2 s_z^2 - 2\hat{\beta}_1\hat{\beta}_2 s_{zx}\right]$$

Now

$$2\hat{\beta}_2 s_{yz} - \hat{\beta}_2^2 s_z^2 - 2\hat{\beta}_1\hat{\beta}_2 s_{zx} = \hat{\beta}_2(2s_{yz} - \hat{\beta}_2 s_z^2 - 2\hat{\beta}_1 s_{zx})$$

$$= \frac{(s_{yx}s_{zx} - s_{yz}s_x^2)}{(s_{xz}^2 - s_x^2 s_z^2)^2}\left[2s_{yz} - \frac{(s_{yx}s_{zx} - s_{yz}s_x^2)}{(s_{xz}^2 - s_x^2 s_z^2)}s_z^2 - 2\frac{(s_{yz}s_{zx} - s_{yx}s_z^2)}{(s_{xz}^2 - s_x^2 s_z^2)}s_{zx}\right]$$

$$= \frac{(s_{yx}s_{zx} - s_{yz}s_x^2)}{(s_{xz}^2 - s_x^2 s_z^2)^2}\left[2s_{yz}(s_{xz}^2 - s_x^2 s_z^2) - (s_{yx}s_{zx} - s_{yz}s_x^2)s_z^2 - 2(s_{yz}s_{zx} - s_{yx}s_z^2)s_{zx}\right]$$

$$= \frac{(s_{yx}s_{zx} - s_{yz}s_x^2)}{(s_{xz}^2 - s_x^2 s_z^2)^2}\left[2s_{yz}s_{xz}^2 - 2s_{yz}s_x^2 s_z^2 - s_{yx}s_{zx}s_z^2 + s_{yz}s_x^2 s_z^2 - 2s_{yz}s_{zx}s_{zx} + 2s_{yx}s_z^2 s_{zx}\right]$$

$$= \frac{(s_{yx}s_{zx} - s_{yz}s_x^2)}{(s_{xz}^2 - s_x^2 s_z^2)^2}\left[-s_{yz}s_x^2 s_z^2 + s_{yx}s_{zx}s_z^2\right]$$

$$= \frac{(s_{yx}s_{zx} - s_{yz}s_x^2)}{(s_{xz}^2 - s_x^2 s_z^2)^2}\left[s_{yx}s_{zx} - s_{yz}s_x^2\right]s_z^2$$

$$= \frac{(s_{yx}s_{zx} - s_{yz}s_x^2)^2 s_z^2}{(s_{xz}^2 - s_x^2 s_z^2)^2} > 0$$

$$\Rightarrow \quad Var(\bar{y}_{lrd}) > Var(\bar{y}_{gld})$$

which shows that the estimator $\bar{y}_{gld}$ is more efficient than the estimator $\bar{y}_{lrd}$ as it has lesser mean squared error. This result has also been verified with the help of an empirical example.

5. NUMERICAL VALIDATION

For numerical validation, we have considered the data given in Des Raj (1972). The size of the population has been considered as 100 and a random sample without replacement of size 30 on first phase has been drawn from it and on second phase a sample of size 8 has been drawn from the sample of first phase.

Table no. 1

Estimates / Methods	Mean ($\bar{y}$)	$V(\bar{y})$	Total (y)	$V(y)$
Model ($z = 1/x^2$)	4.0353	0.1312	403.53	1312.00
Model ($z = 1/x$)	4.3075	0.1396	430.75	1396.00
Model ($z = \sqrt{x}$)	4.0367	0.1571	403.67	1571.00
Model ($z = x^2$)	4.3242	0.1682	432.42	1682.00
Linear Regression	4.2158	0.1858	421.58	1858.00

It is observed that estimates of population mean and population total obtained from $\bar{y}_{gld}$ are more efficient as compared to estimates of population mean and population total obtained from $\bar{y}_{lrd}$.

6. CONCLUSION

The precision of the estimates of population parameters such as population mean and population total etc are improved by including a linear term in ordinary linear regression estimator even in two phase sampling. The proposed model is in general form in which it includes the double sampling versions of models of Ekpenyong et al (2008) and also Misra et al (2009) as particular cases. It has been shown with the help of an example that the proposed model provides précised estimates of population mean and population total as compared to ordinary linear regression estimator of population mean and population total under double sampling.

The additional linear term i.e., z , which is a function of auxiliary variable x has been considered as $\frac{1}{x^2}$, $\frac{1}{x}$, $\sqrt{x}$, x^2 in the present study and many more functions may be formed and theoretically we have shown that it is improved over ordinary linear regression estimator.

7. REFERENCES

[1] Cochran, W.G. (1999), *Sampling Technques*, John Wiley & Sons.

[2] Des Raj (1972), *The design of sampling surveys,* New York*:* McGraw-Hill.

[3] Ekpenyong, E. J., Okonnah, M.I. and John, E.D. (2008), *Polynomial (Nonlinear) Regression Method for Improved Estimation Based on Sampling*, Journal of Applied Sciences, vol. 8(8), pp.1597-1599.

[4] Misra, G.C., Shukla, A.K. and Yadav, S.K. (2009), *A Comparison of Regression Methods for Improved Estimation in Sampling*, Journal of Reliability and Statistical Studies, Vol. 2, Issue 2, pp. 85-90.

[5] Misra, G.C., Shukla, A.K. and Yadav, S.K. (2010), A class of regression type Estimators in survey sampling. Communicated to "*Statistics in transition-new Series*" Journal. BaNoCoSS conference papers.

[6] Sampath, S. (2005), *Sampling Theory and Methods*, Narosa Publishing House, India.

[7] Sukhatme,P.V., Sukhatme, B. V., Sukhatme S. and Asok, C. (1984), *Sampling Theory of Surveys with Applications*, Indian Society of Agricultural Statistics.

PHISICS OF URBANISM: THE FRACTAL GROWTH AND DISTRIBUTION OF THE ROMANIAN CITIES AND TOWNS

Mircea_Gligor[1]

[1]National College "Roman Voda", Roman-5550, Neamt, e-mail: mrgligor@yahoo.com

Abstract. *In some previous works (M. Gligor and L. Gligor, 2004, 2008) we considered the fractal distribution of cities in Romania by population and area of the urban perimeter. The dataset was taken according to the 2002 census, referring to 265 urban settlements. Subsequently, there were officially declared an additional 55 towns (Wikipedia.org). Today (September, 2011) in Romania there are 320 towns. In the present paper, we demonstrate that using the updated dataset, the basic features of distributions remain essentially the same. In the second part of the paper, the Central Places Theory, the diffusion-limited aggregation and the self-organized criticality mechanisms are investigated by means of some numerical simulations and the last two are found to fit better the urban perimeter growth.*

Keywords: Zipf law, master equation, diffusion-limited aggregation, self-organization.

1. HISTORICAL FRAMEWORK

More than seven decades ago, Englewood Cliffs from New Jersey published a pioneering book by Christaller (1933), where several key questions were for the first time posed: "What type of dynamics describe the growing of the urban locations?" and, further: "Are there laws which determine the number, size and distribution of towns?" The Christaller's theory – the so-called *central places theory (CPT)*, later developed by Beckman (1968) among others, describes the urban morphology in the terms of the Euclidean geometry considering that the urban development is structured around a *central business district.* From this he deduced that settlements would tend to form in a triangular/hexagonal lattice, this being the most efficient pattern for travel between settlements.

In fact, Christaller's theory is an important brick in a larger theoretical edifice, namely *the location theory*. Before him, Alfred Weber (1909) formulated a least cost theory of industrial location which tried to explain and predict the location pattern of the industry at a macro-scale. It emphasized that firms seek a site of minimum transport and labor cost, by taking into account several economic factors as the point of optimal transportation based on the costs of distance to the "material index", the labor distortion, the agglomeration and de-agglomeration.

A complementary approach of the CPT was formulated by August Lösch (1940). While Christaller was starting, in effect, with the largest market area and then turned to commodities with ever smaller market areas, Lösch considered first the commodity with the smallest market area and then introduced other commodities with successively larger market areas. Thus while Christaller's approach is an inductive one, Lösch's model is essentially a deductive one. In Lösch's theory the deviations from optimal spatial layouts for individual commodities are relatively small, and the more flexible distribution of functions between centers permits smaller centers to provide goods and services to larger centers (that is not allowed in Christaller's approach).

Essentially, in order to avoid the inconstancies, both Christaller and Lösch theories had to make some basic simplifying assumptions such as:

a) The framework of the model consists of an isotropic (all flat), homogeneous, unbounded limitless surface (abstract space), an evenly distributed population, and evenly distributed resources;

b) All consumers have a similar purchasing power and demand for goods and services;

c) There is only one type of transport and this would be equally easy in all directions; transport cost is proportional to distance traveled;

d) There are no external economies or diseconomies permitted, in shopping or in production which could distort the systems of hexagons;

e) No statements about the sizes of central places are possible (except, in the case of Christaller, that each higher order central place is at least as large as all lower order central places);

f) There can be no Thünen-type ring formation because of the need to have an even distribution of demand;

g) Industrial and service production cannot consume any space; otherwise factor prices for land would be different in different-size centers.

All the above assumptions can be hardly considered realistic, especially for the urban systems in the developing countries. Here the configuration is strongly influenced by local factors, such as climate, topography, history of development, technological improvement and personal preference of consumers and suppliers. Economic status of consumers in an area is also crucial in a developing country. Consumers of higher economic status tend to be more mobile and therefore bypass centers providing only lower order goods. On the other hand, the purchasing power and density affect the spacing of centers and hierarchical arrangements.

Obviously, the applicability of the CPT is drastically limited by two factors the exclusive using of Euclidean varieties as lines and surfaces. The modeling of the urban perimeter in Euclidean terms leads to results in strong discrepancy with the empirical evidence, both for large cities and for small towns configurations. That is why some modern approaches consider that the random cluster models and the fractal properties derived from them constitute necessary ingredients in modeling urban development.

The roots of these approaches are placed in the second half of the XX century, when B. Mandelbrot (1975) opened the door towards a more realistic description of many natural and social phenomena by introducing the mathematical varieties with fractional dimensions, usually called *fractals*. In the fractal theories, the dimension has a higher degree of generality than in Euclidean it has. For instance, considering an object M which can be decomposed in N parts, each one in the ratio r (r<1) with the whole object, the (self-similarity) dimension is defined as:

$$D = \frac{\log N}{\log(1/r)} \tag{1}$$

Thus, the dimension associated to the object M may appear as a fractional number.

It is well-known that the distinctive feature of the fractal object is *self-similarity i.e.* it displays the same aspect regardless the scale at which it is viewed. This geometrical property is mathematically expressed in the *scaling* of the characteristic functions $F(x)$ describing the fractal object: $F(cx) = c^{\alpha}F(x)$. The parameter α completely describes the object geometric features; this is like saying that a simple number is able to characterize a very complicated shape. Moreover, the fractal character of the system may be reflected not only in its geometric properties, but also in its characteristic distribution functions. In this way, numerous natural systems display a fractal structure: the mountain ranges, river networks, coastlines, etc. The word "fractal" here means that some correlation functions show non-trivial power law behavior.

In their classical paper on *self-organized criticality (S.O.C.)*, Bak, Tang and Wiesenfield (1987) argued that the dynamics which give rise to the robust power-law correlations seen in the non-equilibrium steady states in nature must not involve any fine-tuning of parameters. It must be such that the systems under their natural evolution are driven to a state at the boundary between the stable and unstable states. Such a state then shows long-range temporal-temporal fluctuations similar to those in equilibrium critical phenomena. Bak et al. (1987) proposed a simple example of a system whose natural dynamics drives it towards, and then maintains it, at the edge of stability: a sand pile. When the average slope of the sand pile is larger than a certain value θ_c, addition of a small amount of sand often results in an avalanche whose size is of the order of the system size, while in a pile where the average slope is θ_c, the response to addition of sand is less predictable. In the steady state of this process the sand may be added to the system at a constant small rate, but it leaves the system in a very irregular manner, with long periods of apparent inactivity interspersed by events which may vary in size and which occur at unpredictable intervals. A long power law tail characterizes the typical distribution of the frequency and size of the avalanches, with an eventual cut-off determined by the system size.

The "random cluster" models (Batty & Xie, 1996) consider the cities growth as the growth of two-dimensional aggregates of particles – problem of particular interest in physics of disordered media. In particular, the model of *diffusion-limited aggregation (DLA)* (Witten Jr. & Sander, 1981) has been applied to describe urban growth, and results in tree-like dendritic structures, which have a core or "central business district". The DLA model predicts that there exists only one large fractal cluster that is screened from incoming "development units" (people, capital, resources, etc).

The DLA model predicts that the urban population density decreases from the centre to the periphery as a power law:

$$\rho \sim r^{D-2} \tag{2}$$

where r is the radial distance from the core and D is the fractal dimension of the model.

Alternatively, an exponential decay is considered in the so-called *correlated percolation model (CPM)* (Makse, Havlin & Stanley, 1995; 1998), where the spatial correlations in urban settlements are also embodied. A modified version of the DLA model in which the cluster density decays as a complementary error function was recently elaborated (Pica Ciamarra & Coniglio, 2006).

Clearly, the CPM offers the best description of the great urban agglomerations (e.g. London, Berlin). Nonetheless, it is a geographical stylized fact that the small villages, at least the ones situated in a plane environment, are roughly concentrically structured, so they seem to be better fitted by the DLA model. It might be of interest to study what happens when the cities are formed by merging some independent developed villages. Bucharest, the analyzed city in the present paper, is one example.

On the other hand, the fractal behavior may be found not in the spatial structure itself, but is manifested through the power-law dependence between some physical observables. Examples include the earthquakes, the fluid turbulence, etc. Note that a lot of time-series from the social sciences, *e.g.* stock market price variations both for the greatest stock markets (Mantegna & Stanley, 2000) and for some emerging ones (Gligor, 2004) display power-law tails in their power spectra (the so-called "1/f noise").

In this new meaning of fractality, let us recall that almost six decades ago, an important result for the urbanism was pointed out: the population and area distributions of cities and towns follows power law behavior (Zipf, 1949). Indeed, if n_s is the number of cities having the population s, then:

$$R(s) = \int_s^{\infty} n_x dx \tag{3}$$

defines the *rank* of the city in a hierarchy: the largest city has R = 1; the second largest R = 2, etc. Zipf found that R is a function of s, which can be inverted as:

$$s(R) \sim R^{-\gamma} \tag{4}$$

with $\gamma \approx 1$.

The first striking property of the above result (known today as "the Zipf's law for cities" is the scale invariance, reflecting an underlying fractal structure; the second consists in universality: the statistical datasets shows that the law is valid for many different societies and during various time periods.

The universality of the power law behavior suggests the possibility of study the urban system by tools that do not depend in an explicit way on the concrete nature of the interactions between its elementary constituents. Gabaix (1999) suggested that the behavior described by Eq. (4) can be explained by assuming an auto-catalytic process characterized by the rule that the growth of each individual entity is proportional to its present size. Nonetheless, few years later, Blank and Solomon (2000) showed that a growing system with a fixed number of components and a fixed smallest component size cannot converge to a power law. Instead, by fixing the minimal population to a certain fraction of the average, they defined the so called "generalized Lotka-Volterra process" with variable number of components, which converges to a power law for a very wide range of parameters. Moreover, in a very large subset of this range, they obtained for the power law exponent the special value 1 specific for the cities population distribution.

The power law distribution is dramatically cut at the upper end, i.e. in the region of the small size towns. Taking into account this aspect, Malacarne, Mendes and Lenzi (2002) suggested the q-exponential distribution (derived from the generalized nonextensive statistical mechanics) as a possible alternative to the power law. However, the cut off can be simply due to some marginal effects related to the finiteness of sampling. At present, the Zipf's law is generally accepted as an empirical fact describing quite different societies (Mulianta, Situngkir, & Surya, 2004; Newman, 2005, and references therein; Moura Jr. & Ribeiro, 2006).

In the following section we show that the Zipf's law can be easily derived by supposing that the development process is Markovian. It is well-known that Markovian stochastic processes can be described by a master equation. In the last decades Weidlich and Haag have successfully introduced this formalism for the description of social processes (Weidlich & Haag, 1983; Weidlich, 1991) like opinion formation (Haag, 1989), migration (Haag & Weidlich, 1984), agglomeration (Weidlich & Haag, 1987) and settlement processes (Weidlich, 1997), and have shown how well-known outcomes might well arise in a dynamic context. Particularly important are the evolution of the mean value and quasi-mean value equations that can depict expected outcomes over time (Weidlich, 2002). These equations operate in a stochastic framework, which is supposed to represent the actions of the individuals or other lower-level units in the system. Following synergetics, trends in order parameters usually determine the overall outcomes.

Some inherent difficulties with the above approaches should be noted. One that has been hinted at is the relative lack of empirical work related to or based on socio-dynamics. In general it is not very easy to empirically estimate the many of the transition probabilities that are crucial to many of the models. On the other hand, the above approaches seem to leave out the possibility that the trends of the order parameters themselves may be altered by changes in individual behavior. In any case, all models have their limits, and socio-dynamic approaches only deal with aggregated phenomena rather than individual outcomes (Weidlich, 2002).

The main points of our research are outlined below. In Section 2 we argue that the Zipf's distributions of cities populations and areas can be derived from the basic assumption that the development process is Markovian, without other additional constraints. This result, well known for the largest cities in all over the world, is empirically tested using data referring to Romania as an example of developing country. Section 3 brings into discussion the shape of the urban perimeter by means of some numerical simulations of DLA and SOC models, versus the empirical founded structure of the largest urban settlement in Romania, namely the capital Bucharest. We find that the basic assumptions of CPT simply do not work. No hexagonal structure can be found in the central places disposition. Instead, the real structure is found to be well fitted by DLA and SOC simulations, and may be entirely explained by particular historical and economic facts of evolution. The last section summarizes the findings and draws some conclusions.

2. THE POPULATION AND AREA DISTRIBUTIONS OF CITIES AND TOWNS

2.1 The stochastic model

Let us consider N towns, and let be s_i the size of the i-th city (expressed as number of citizens as well as units of urban area). The model is built using the general framework of the master equation. We assign the transition rates for the growth $\Psi_+(s_i)$ or decrease $\Psi_-(s_i)$ of the size s_i. In other words, $\Psi_+(s_i)$ is the probability that a new citizen arrives (or a new economic/residential unit-area location is created) in the city i in the time interval $(t, t + dt)$, so that $s_i \to s_i + 1$. Analogous, $\Psi_-(s_i)$ is the probability that one of the s citizens departs (or a unit-area location is left) in the same time interval, so that $s_i \to s_i - 1$.

We introduce now the average number $n(s, t)$ of cities of size s at time t, for a *given N*. The quantity $n(s, t)$ satisfies the master equation:

$$\frac{\partial n(s,t)}{\partial t} = \psi_-(s+1)n(s+1,t) - \psi_-(s)n(s,t) + \psi_+(s-1)n(s-1,t) - \psi_+(s)n(s,t) \tag{5}$$

where $\partial n(s, t)/\partial t$ is the variation of n. The parameters of the model are the transition rates $\Psi_\pm(s_i)$.

If the total number of cities N is considered not be constant, at least an additional parameter must be introduced (e.g. Marsili & Zhang, 1998; Blank and Solomon, 2000), describing the probability that a citizen leaves the system (or a unit-area location is created outside of the system). However, a new city formation implies an *allocation* process, strongly depending on external conditions which cannot be simply included in this approach (but might be an interesting challenge for further works). Moreover, the time scale of this process certainly exceeds the time scale of the statistical data recording.

Thus, as in the study of the most interacting-agent systems, we are firstly interested in finding the stationary solution of the master equation, for which s and N are constant on average and:

$$\frac{\partial n(s,t)}{\partial t} = 0 \tag{6}$$

In this case $n(s, t) \equiv n(s)$, *i.e.* the quantities n and Ψ do not depend explicitly on the time. Equation (5) becomes:

$$\psi_-(s+1)n(s+1) - \psi_-(s)n(s) + \psi_+(s-1)n(s-1) - \psi_+(s)n(s) = 0 \tag{7}$$

The simplest way to take into account the interactions among agents is assuming these interactions pair-wise type, so that: $\Psi \sim s^2$. This assumption simply means that all the s city units are in interaction each other, displaying a fully connected social network. In the simplest way, choosing $\Psi_-(s) = k_1 \cdot s^2$ and respectively $\Psi_+(s) = k_2 \cdot s^2$, a straightforward calculus leads to:

$$n(s) = C / s^2 \tag{2}$$

where $C = N \big/ \sum_{i=1}^{N} s_i^{-2}$. Using the rank relation (3) one finds: $R(s) \sim s^{-1}$, or, inverting: $s(R) \sim 1/R$, that is the usual form of the Zipf's law.

2.2 The empirical dataset

The empirical data referring to the urban area and population of the Romanian cities and towns were supplied by the most recent census performed by Romanian Institute of Statistics (2003). A number of 320 large and medium size towns are ordered by decreasing the urban area, and by decreasing the urban population.

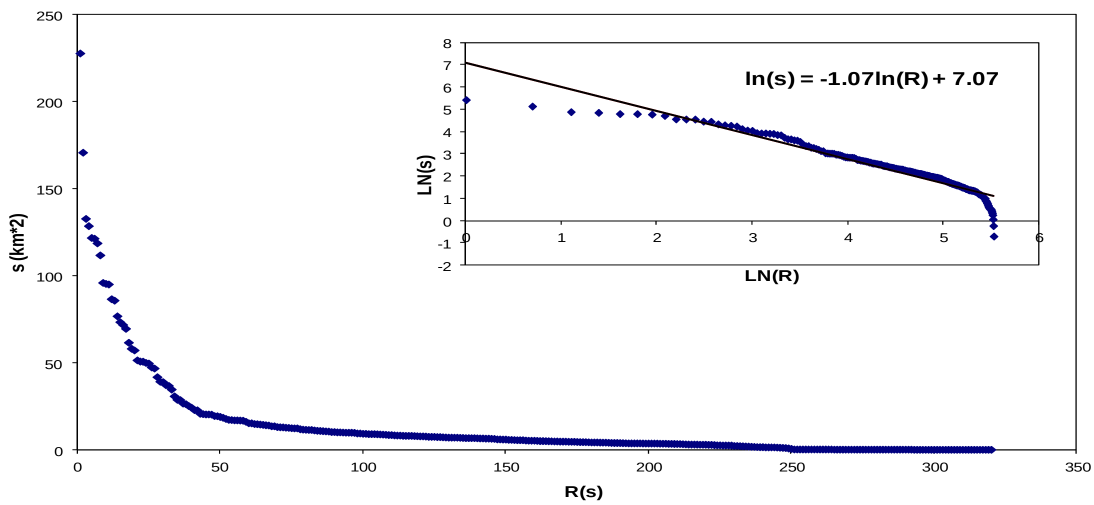


Fig. 1 The urban area distribution for 320 Romanian cities and towns. Inset: the Zipf plot.

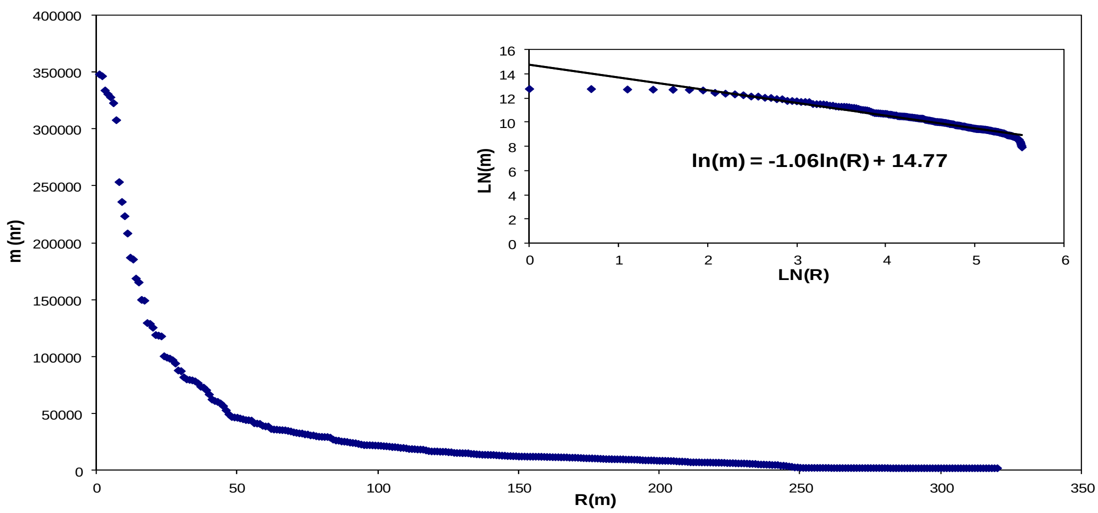


Fig. 2 The urban population distribution for 320 Romanian cities and towns. Inset: the Zipf plot.

The data are well fitted by power laws, leading to the scaling exponents: μ = 1,07 (for the urban area – Fig. 1) and γ = 1,06 (for the urban population – Fig. 2). These results are in well agreement with the similar ones reported in the literature cited in Introduction.

One can see a good agreement of fitting as regards to the Pearson product moment correlation coefficient through the given data points $(R^2)_1$ = 0.92 for the area distribution, and $(R^2)_2$ = 0.94 for the population distribution. The standard errors are σ_1 = 1.07 and σ_2 = 1.05 respectively. According with the Chebyshev's theorem, removing the points situated at more than $\pm 4\sigma$ from the fitted curve, i.e. the "outliers" (9 points in Fig. 1 and 7 points in Fig. 2), one gets $(R^2)_1$ = 0.98 and $(R^2)_2$ = 0.99 respectively. The χ^2 statistic test applied to the remainder points gives more than 98% confidence interval in both cases.

Taking into account the above results we can conclude that the both distributions are well fitted by the Zipf's law. The fact that in any sampling a threshold minimal value is a priori chosen imposes a biased consideration of the values situated *around* the threshold and can explain the distribution cut-off.

3. THE URBAN PERIMETER MODELING

3.1 Numerical simulations of the growth process

The simulations are performed model on a lattice, which we take for simplicity to be the two dimensional square lattice. The system evolves in discrete time. In the first version (Fig. 3), the simulation follows the slightly modified DLA model: we start from a central site containing a large number of particles. Now we introduce a new particle at a large distance from the seed, and let it perform a random walk. Ultimately, that second particle will either escape to infinity or contact the seed, to which it will stick irreversibly.

Now introduce a third particle into the system and allow it to walk randomly until it either sticks to the two-particle cluster or escapes to infinity. In addition, at each time step, some particles are moved from the central place to the neighboring sites with a power-law decay probability.

In the second version (Fig. 4), one starts from the sand pile model with a uniform distribution of heights (Bak et al. 1987). There is a positive integer variable at each site of the lattice, called the height of the sand pile at that site. At each time step a site is picked randomly, and its height z_i is increased by unity. If the site height is larger than a critical value z_c, the site relaxes by toppling whereby z_c grains leave the site, and each of the four neighboring sites gets $z_c/4$ grains. In case of toppling at a site at the boundary of the lattice, grains falling "outside" the lattice are *not* removed from the system, but they are added randomly to the highest ones. This process continues until all sites are stable.The spatial distribution of "avalanches"/"relaxation" processes are followed at two different times of simulation

3.2 Comparison with the real data

As shown, in the DLA model, only a large central place or large cluster is generated. The cluster generated by this process are both highly branched and fractal (Fig. 3). The cluster's fractal structure arises because the faster growing parts of the cluster shield the other parts, which therefore become less accessible to incoming particles. A new arriving random particle is far more likely to attach to one of the tips of the cluster shown in figure 1a than to penetrate deeply into one of the cluster's "fjords" without first contacting any surface site.

However, a real urban area is rather composed of central places that are spatially distributed following a certain hierarchy, thus the sand pile model of evolution (Fig. 4a and 4b) offer a more realistic description of the real urban perimeter shown in Fig. 5.

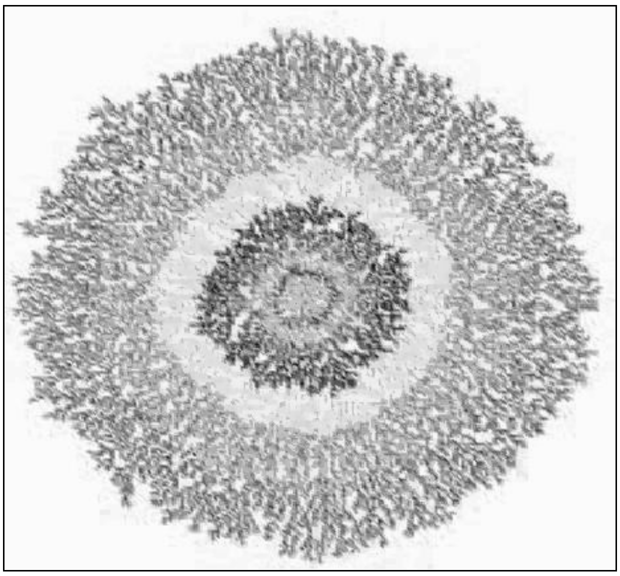


Fig. 3 A numerical simulation of a growth process in a dendritic-like structure, from the DLA model (Eq. 2 with *D* = 1.7). The growth begins in the centre and extends to the periphery.

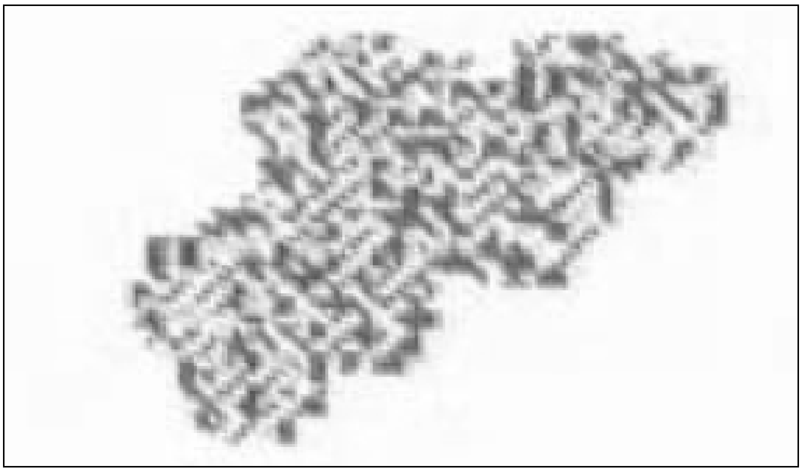


Fig. 4a

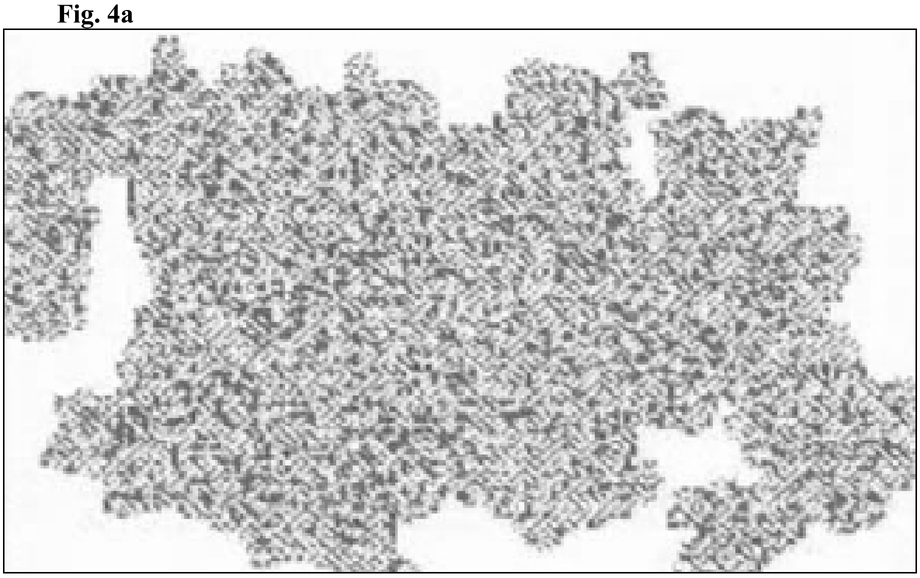


Fig. 4b

Fig. 4 The result of the numerical simulation of a growth process in the sand pile model after **(a)** $n = 10^2$; **(b)** $n = 10^3$ simulation time steps.

Fig.5a **Fig.5b**

Fig. 5 The structure of Bucharest. **a.** Map. In square: The Central Business District; **b.** Satellite view of the fractal structure.

One can also see in Fig. 5 that the basic assumptions of CPT simply do not work. No hexagonal structure can be found in the central places disposition. Instead, we found one first order and two second order central places, in a quasi linear disposition. This structure can be exclusively explained by historical facts, taking into account that the city was formed by fusion of three independent old villages that must be seen as independent centers of growth.

Each independent center has developed a dendritic-like structure around itself, as predicted by DLA model. Moreover, the density of commercial units decreases following a power law dependence on the distance from the local center.

On the other hand, the actual shape of the urban perimeter is well described by a SOC process, namely the spatial extending of the avalanches processes in the sand pile model. This result indicates that the first steps of any urban development are governed by “trial and error” principle rather than economic efficiency reasons. The randomness of the avalanches locations means here that the firms’ growth or failure result from the interaction of numerous social, political, economic factors, too many to be considered separately as explanatory facts, as well as from the agents idiosyncratic behavior.

4. CONCLUSIONS

In the present paper some basic ideas of the fractal city theory have been briefly reviewed. Particularly the questions of urban perimeter growth and towns’ distribution were pointed out. While an increasing amount of literature is devoted to the large cities structure and distribution, a relatively low interest has been so far given in the study of medium and small-size urban locations, especially those situated in the developing countries. Generally here are not mega-polis-like cities and the most urban centers are formed by merging some small units (villages).

In this case, we found that:

(i) The Zipf’s law can be directly derived from the assumption of the Markovian process of development without other auxiliary hypothesis.

(ii) The power law distribution with exponent roughly unit was found to be valid for the urban settlements distribution in a developing country as well as it was found in the previous studies referring to the cities distribution of the developed countries;

(iii) The basic assumptions of the Central Places Theory are not fulfilled in the case of cities formed by merging old village-like settlements. Particularly, the regular hexagonal structure predicted by CPT cannot be found.

(iv) The Diffusion-Limited Aggregation model fits very the growth process around the local (second-order) poles.

(v) The self-organized sand pile model seems to fit very well the urban perimeter shape. This fact can bring into discussion the relative importance of various economic, political and geographical particular factors, as well as the intrinsic competition between order and hazard.

5. REFERENCES

[1] Bak, P., Tang, C., & Wiesenfeld, K. (1987), Self-organized criticality: An explanation of 1/f noise. *Physical Review Letters, 59*, 381-384.

[2] Batty, M. & Xie, Y. (1996), Preliminary evidence for a theory of the fractal city. *Environment and Planning A*, *28*, 1745-1762.

[3] Beckmann, M. J. (1968), *Location Theory*. New York: Random House.

[4] Blank, A., & Solomon, S. (2000), Power laws in cities population, financial markets and internet sites (scaling in systems with a variable number of components). *Physica A, 287*, 279-288.

[5] Christaller, W. (1933), *Central Places in Southern Germany*. New Jersey, Prentice-Hall: Englewood Cliffs.

[6] Frankhauser, P. (1994), *La Fractalité des Structures Urbaines*. Paris: Collection Villes, Anthropos.

[7] Gabaix, X. (1999), Zipf’s Law for cities: An explanation. *Quarterly Journal of Economics, 114*, 739-767.

[8] Gligor, M. & Gligor L. (2004), Fractal properties of the structure and distribution of urban settlements. In *Actual topics in statistics*. Junimea, Iasi. (in Romanian)

[9] Gligor, M. (2004). Statistical properties of Romanian emerging stock market RASDAQ. *International Journal of Theoretical and Applied Finance*, *7*, 723-739.

[10] Gligor, L. & Gligor M. (2008), The fractal city theory revisited: new empirical evidences from the distribution of the Romanian cities and towns. *Nonlinear Dynamics, Psychology and Life Sciences, 12 (1),* 15-28.

[11] Haag, G. (1989), *Dynamic Decision Theory: Application to Urban and Regional Topics*. Dordrecht: Kluwer Academic Publishers.

[12] Haag, G. & Weidlich W. (1984), A Stochastic Theory of Interregional Migration. *Geographical Analysis 16*, pp. 331-335.

[13] Lösch, A. (1940), *Die Raumliche Ordnung der Wirtschaft* (Jena: Gustav Fischer Verlag), translated by W. H. Woglom, *The Economics of Location* (New Haven: Yale University Press, 1954)

[14] Machedon, L. & Scoffham, E. (1999), *Romanian Modernism: The Architecture of Bucharest, 1920-1940*. Cambridge, M.A.: MIT Press.

[15] Makse, H. A., Havlin, S. & Stanley, H. E. (1995), Modelling urban growth patterns. *Nature*, *377*, 608-612.

[16] Makse, H. A., Havlin, S. & Stanley, H. E. (1998), Modelling urban growth patterns with correlated percolation. *Physical Review E, 58*, 7054-7062.

[17] Malacarne, L. C., Mendes, R. S. & Lenzi, E. K. (2002), q-Exponential distribution in urban agglomeration. *Physical Review E, 65*, 017106.

[19] Malescio, G., Dokholyan, N. V., Buldyrev, S.V., & Stanley, H. E. (2000), Hierarchical Organization of Cities and Nations. URL:http://xxx.lanl.gov/abs/cond-mat/ 0005178.

[20] Mandelbrot, B. (1975), *Les objets fractals: forme, hasard et dimension*. Paris: Flammarion.

[21] Mantegna, R. N., & Stanley, H. E. (2000), *An introduction to econophysics: Correlations and complexity in finance*. Cambridge: Cambridge University Press.

[22] Marsili, M. & Zhang, Y. C. (1998), Interacting individuals leading to Zipf’s law". *Physical Review Letters, 80*, 2741-2744.

[23] Moura Jr., N. J. & Ribeiro, M. B. (2006), Zipf Law for Brazilian cities. *Physica A, 367*, 441-448.

[24] Mulianta, I., Situngkir, H. & Surya, Y. (2004), Power law signature in Indonesian population. URL: http://xxx.lanl.gov/abs/nlin.PS/0409006.

[25] Newman, M. E. J. (2005), Power laws, Pareto

distributions and Zipf's law. *Contemporary Physics, 46*, 323-351.

[26] Pica Ciamarra, M., & Coniglio, A. (2006), Random walk, cluster growth, and the morphology of urban conglomerations. *Physica A, 363*, 551–557.

[27] Romanian Institute of Statistics (2003), Romanian Statistical Yearbook, 2003. On-line edition: http://www.insse.ro/anuar_2004/aseng2004.htm

[28] Weber, A. (1909), *Über den Standort der Industrien* (Tübingen), translated by C. J. Friedrich, *Alfred Weber's Theory of the Location of Industries* (Chicago: University of Chicago Press, 1929)

[29] Weidlich, W. (1991), Physics and social science - The approach of synergetics. *Physics Reports 204*, 1–163.

[30]Weidlich, W. (1997), From fast to slow processes in the evolution of urban and regional settlements structures. In *Self Organization of Complex Structures – From Individual to Collective Dynamics.* Ed. F Schweitzer (Amsterdam: Gordon and Breach Science Publisher) pp 475-488.

[31] Weidlich, W. (2002), *Sociodynamics: A Systematic Approach to Mathematical Modelling in the Social Sciences*. London: Taylor & Francis.

[32] Weidlich, W. & Haag, G. (1983), *Concepts and models of a quantitative sociology*: *The dynamics of Interacting Populations* (Springer series in synergetics). Berlin: Springer.

[33] Weidlich, W. & Haag, G. (1987), A dynamic phase transition model for spatial agglomeration processes. *Journal of Regional Science 27(4)*, 529–569.

[34] Wikipedia.org: http://en.wikipedia.org/wiki/List_of_cities_and_towns_in_Romania

[35] Witten Jr., T.A., & Sander, L. M. (1983), Diffuzion-limited aggregation. *Physical Review B, 27*, 5686-5697

[36] Zipf, G. K. (1949), *Human Behavior and the Principle of Least Effort*. Cambridge MA: Cambridge University Press.

THE OCCURRENCE OF POWER LAW FOR COMPOSITE INDICATOR OF ICT ADOPTION IN SERBIANS ORGANIZATIONS

Mladen Čudanov[1], Jelena Minović[2], Gheorghe Săvoiu[3]

[1]University of Belgrade, Serbia [2]Institute of Economic Sciences, Belgrade, Serbia, [3]University of Pitesti, Romania
[1]e-mail:mladenc@fon.rs, [2]e-mail:jelena.minovic@gmail.com,[3]e-mail: gsavoiu@yahoo.com,

Abstract. *This paper presents analysis of power law occurrence in the Serbian organizations. Its occurrence in physics, economics and finance, computer science, biology, earth and planetary sciences, demography, and the social sciences has been covered by research, but this study researches its occurrence in business application of ICT, where published research is still scarce. Parameter used in measurement is composite indicator of ICT adoption, already successfully used as a tool to measure ICT adoption in organizations in three studies of different context. It is described in detail in first part of the article, and compared with other approaches in measurement of ICT adoption in organizations. We analyzed the occurrence of power law in three cases: distribution of values of ICT adoption indicator in 67 selected enterprises, distribution between values of ICT adoption and size of the enterprises measured by number of employees, and distribution between values of ICT adoption and profit per employee. Thus, we found that ICT adoption in the Serbian organization is distributed according to Pareto's power law distribution, in all cases, with the tail exponent (or parameter* α *) smaller than 1, which means that ICT adoption in the Serbian organizations is undeveloped. This is the first paper that shows the distribution of ICT adoption possesses properties of Pareto's law in the Serbian organizations.*

Keywords: *Power law, ICT adoption, organization size, profit per employee*

1.INTRODUCTION

Power-law states that appearance of small values is quite common, while large values appearances are rear. This law is literately regarded as Zipf or Pareto's law[1]. When the probability of measuring a particular value of some quantity varies inversely as a power of that value, the quantity is said to follow a power law, also known the Pareto's distribution [2].

Atkinson, Piketty, and Saez wrote that the Pareto's law for top ICT index is given by the following (cumulative) distribution function for ICT indicator *z* [3]:

$$1 - F\left(z\right) = \left(\frac{k}{z}\right)^{\alpha}, \left(k > 0, \alpha > 1\right), \qquad (1)$$

where *k* and α are given parameters, α is called the Pareto parameter. The corresponding density function is given by

$$f\left(z\right) = \alpha k^{\alpha} z^{-1-\alpha} \qquad (2)$$

The value of the exponent depends on the value of the lower ICT indicator bound. Indeed, empirical studies show that the value of changes across different countries, and is typically in the range [4].

Further details on the mathematics of power laws can be found in Newman research [2]. Power laws are common in the systems that are constituted out of pieces that have no specific size and in the systems that are constituted out of self-catalyzing elements. They represent the link between simple microscopic basic laws on individual level and macroscopic phenomena that occur collectively [5].

2. LITERATURE REVIEW

Power laws appear widely in physics, economics and finance, computer science, biology, earth and planetary sciences, demography, and the social sciences [2]. Pareto's law has also been proposed as a model of word frequency [6], fluctuation in finance [7], firm sizes [8,9], citations of scientific papers [10,11], web hits [12,13], the cumulative distribution of the number of telephone calls received on a single day [14,15], frequencies of family names [16], turnover [17], income [18,19,20]), wealth [4,21] [22,23] ,), and so on. Clauset, Shalizi, and Newman [24] present power law distributions in empirical data from a range of different disciplines. Gabaix [25] surveys empirical power laws regarding income and wealth, the size of cities and firms, stock market returns, trading volume, international trade, and executive pay. He surveyed power laws in finance and economics, and he showed that Pareto laws have also been applied in several areas outside of income and wealth distribution. Atkinson, Piketty, and Saez [3] surveyed some of these theoretical models.

3. COMPOSITE INDICATOR OF ICT ADOPTION

In order to analyze power law occurrence we first need to present objective measurement of ICT adoption in the company, which can also be useful for comparison with other characteristics of the organization, as a guideline for changing processes in organization, measuring infrastructure requirements high technology projects or organizational changes that include information and communication technologies, or as a benchmarking tool among different companies. A tool for objective assessment of ICT adoption has long been needed by the consulting and scientific community in order to enable comparison with other characteristics in organizations[26], benchmarking [27,28] or measurement of existing state of infrastructure before implementation of Enterprise information system [29]. Presenting objective measurement of ICT adoption level can ease problems of managing ICT in organization. It could lead to certain level of formalization and as all control systems could be misrepresented and distorted to fit subjective, illegitimate goals, but would nevertheless present benefits of ordered measurement system in organization. In mentioned and other studies in the field problems of defining adoption emerge, and different approaches are applied in different studies. For the purpose of this study, we will use term "adoption" as establishing prerequisite ICT infrastructure in organization, both software and hardware. In such definition, adoption is different from actual usage by employees of the organization, which has other, behavioural dimensions, and introduces humanware requirements. It describes potential for usage, and is necessary, but not sufficient condition for actual usage of ICT by employees, that depends on other, less tangible factors.

Formula of composite indicator used for the purpose of this research was the result of several years of experimentation during consulting experience in companies, where authors searched for different tools to objectively quantify adoption of ICT in the organization. It was developed as a tool, basically for the purpose of scientific research, to the point subjective assessment of business consultants was that it describes adoption of ICT accordingly to the qualitative assessment observed in organizations that were included in the research. In practice, organizations with low level of index were estimated with low level of adoption of ICT and organizations with good or excellent adoption of ICT had higher levels of this indicator. For example, if indicator was low in value, computers were rarely or even not at all used, core processes in organization were driven without ICT and at best ICT was used below its potential – e.g. computers were used as typewriters and calculators. In contrast, companies with advanced practices of ICT usage, like ERP systems application, active internet presence, TEL programs or electronic office collaboration had high levels of this indicator. In this form, formula has been applied successfully used in studies that compared levels of ICT adoption in organizations with autocratic, democratic and liberal management styles [30], organization size [31] and dominant management orientation [32]. Currently, efforts are given into further improvement of the formula, which is in its original formula given below.

For the purpose of calculation of this formula data from wider research of 71 organizations in transitional countries in Southern and Eastern Europe was used. During mentioned wider research each organization was analyzed, resulting in reports covering wide range of data extracted into more than 40 indicators that served as a basis for several different research papers. Problems with difference of estimated values by different personnel or misunderstanding of questions by interviewed personnel were solved by asking for clarification and second estimation, common to simplified application of Delphi method used in articles of De Icaza et al. [33] and Radauceanu et al. [34]. Written reports from companies were used to detect inconsistencies in data, comparing qualitative with quantitative descriptions.

There are other approaches in measuring ICT adoption, besides by development of composite indicator as used in this paper:

- Subjective assessment by users on a questionnaire scale, e.g. "yes/no" or "1-10" scale (Thong and Yap [35]; Thong et al. [36]; Yang et al. [37])
- Subjective categorization by users or external party into defined qualitative categories, usually by intensity of adoption and usage (Howard [37]; Redoli et al. [38])
- Calculation of ICT adoption by single factor, like number of computers per employee, (Moore and Benbasat [39]; Golubeva and Merkuryeva [40]; El-Mashaleh [41])
- Quantitative assessment of ICT adoption, similar to composite indicator of ICT adoption is more usual in studies that measure trends in national economies, like Kauffman and Kumar [42] and Hanafizadeh et al. [43], who mainly utilize statistics on population usage of ICT to develop quantitative assessment of ICT adoption. However, there has not been much research with measuring similar indicator on national level.

$$CIICT = \frac{NoC}{NoE} + \frac{NoCC}{NoE} + \sum_{i=1}^{i=8} Cfi + \prod_{i=1}^{i=8} \left(\sqrt{\frac{NoCC}{NoE}} + Cf_i \right) + CDB + DBA \quad (3)$$

Where mentioned factors represent: CIICT = composite indicator of ICT adoption in company in its original form; NoC = Number of computers in the company; NoE = Number of employees in the company; NoCC = Number of computers connected to internal network in the company; Cf_i = Coverage of enterprise function by ICT, where for different values of i functions are: 1 – human resources, 2 – accountancy, 3 – financial, , 4 – technical, 5 commercial, 6 – administrative, 7 – legal, 8 – protection; Coverage of business function was estimated by IT staff, functional staff and top manage-ments as percentage of usual job in that function supported by ICT existing in the organization; DB = Existence of integrated company database (0=no, 1=yes); DBA = Database administrator present (0=no, 1=yes).

First addend in the formula was given as ratio between number of computers and number of employees. By using the ratio this factor becomes relative to organization's size. Second addend was introduced as extension of the first to emphasize importance of computer networks in a company. This factor will probably make first factor of the formula obsolete in several years, and replace it completely. First factor will not be as fit to measure ICT adoption in future companies as the second, because only computers that have ultimate security demands or otherwise special status will not be in network, even in small enterprises. Today, however, especially in small and medium enterprises we have a large number of unconnected computers so both addends were used. Third addend is a simple sum of coverage of business functions by information system. It has more impact on the final value, because practice and literature review like in Scheer and Habermann [44] or Scheer and Schneider [45] connects coverage of business functions by ICT and ICT adoption in organization. Fourth addend emphasizes importance of synergetic effect of ICT appliance in the company. Since business can be seen as chain with several interconnected functions, its strength is determined by the weakest link, and when all major business functions are fully covered, ICT adoption is much better than when one function is omitted, leaving weak link that will slow down other functions. In other words, adoption increase does not follow coverage of business function linearly, company with half of main business function covered by ICT compared to company with all main functions covered by ICT should be valued as much less than half successful as the second example in ICT adoption, because of the lack of synergy. Coverage of all functions in the organization by ICT, usually realized by implementation of all or most necessary ERP modules gives synergetic benefits. Fifth addend was treated as binary in original formula, representing existence of integrated database (0 if it does exist and 1 if it exists). The problem of integrating separate databases in organization is important in business adoption of ICT, and this factor values contribution to its solution. Sixth factor in used formula represented employment of worker with database administrator duties, again as binary value, 0 if none exists, and 1 if there are one or more employees. Tarafdar and Vaidya [46] compared companies with "Pioneer", "Advanced", "Late" and "Laggard" roles of ICT adoption, and indicative is that in Laggard case roles of database administrators were performed by staff from other functions, subsequently trained for those tasks. Importance of database administrator role for ICT adoption was also observed by Vogel et al. [47] and Van den Hoven [48].

Main goal of this indicator is to satisfy following requirements:

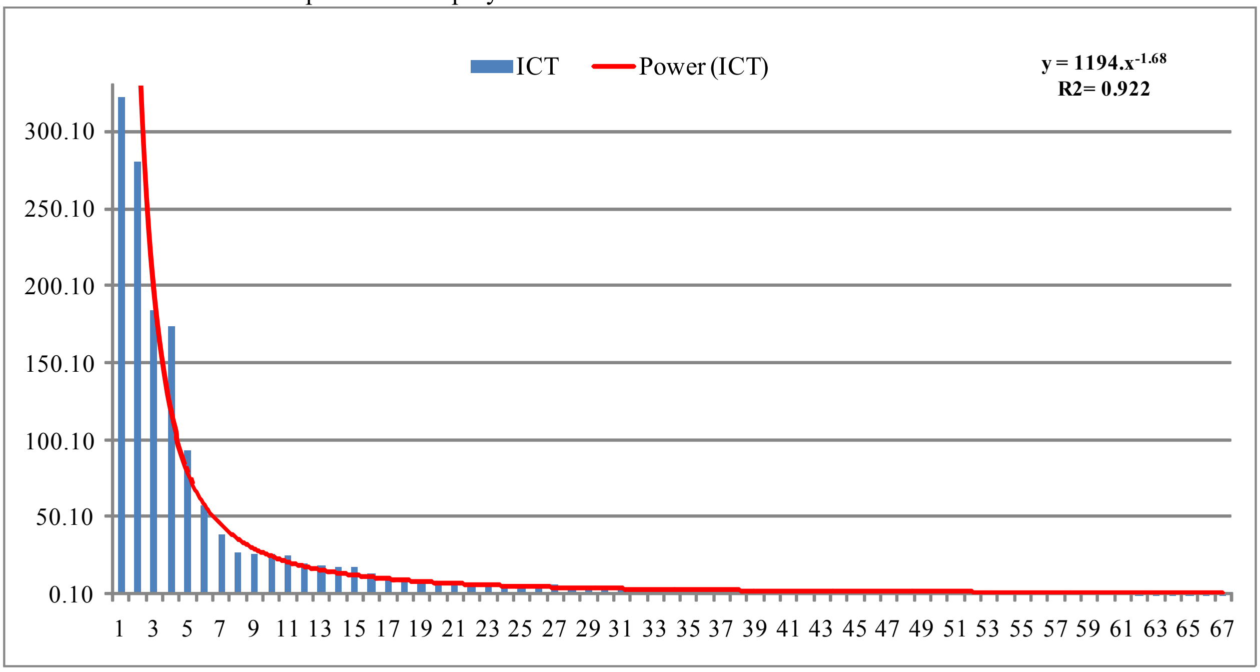
a) To take into consideration various different factors of ICT adoption in organization described in literature and observed during our consulting experience.

b) To maximize practical usability of the formula for other consultants and researchers, it should consist of data which can usually be easily collected during interview in a company, and be as simple as possible, but not simpler, according to famous Einstein recommendation – still retaining power to accurately describe level of ICT adoption;

c) To be as objective as possible, because of perceived tendency that in different organizations individuals have different criteria of subjective assessment of ICT adoption.

4. RESULTS

The occurrence of power law was measured in comparison with size of enterprises measured by number of employees, and profit per employee.



Source: Data from primary source, Authors' estimation of Pareto's parameters.

Fig.1. Power law distribution of values of ICT adoption indicator in 67 selected enterprises.

The corresponding density function, from Figure 1, for ICT indicator is:

$$f\left(z\right) = 1194 \cdot z^{-1-0.68}. \tag{4}$$

An estimate of the expected statistical error on the exponent α is given by http://www-personal.umich.edu/~mejn/courses/2006/cmplxsys899/

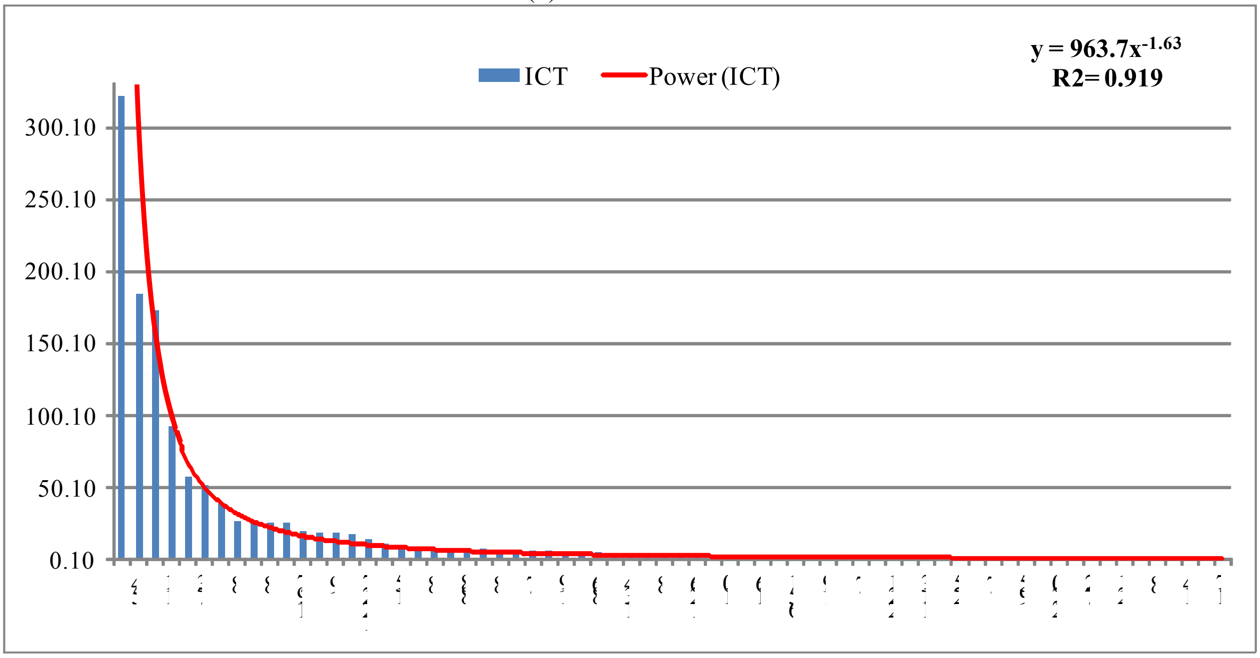
$$\sigma = \frac{\alpha}{\sqrt{n}} = 0.08 \quad (5)$$

where n is number of observation.

The Pareto parameter is:

$$\alpha = 0.68 \pm 0.08. \quad (6)$$

Power law has been checked between ICT adoption and size of the enterprises measured by the number of the employees.



Source: Data from primary source, Authors' estimation of Pareto's parameters.

Fig. 2. Power law distribution between ICT adoption indicator and the size of enterprises measured by number of employees.

The corresponding density function, from Figure 2, for ICT indicator is:

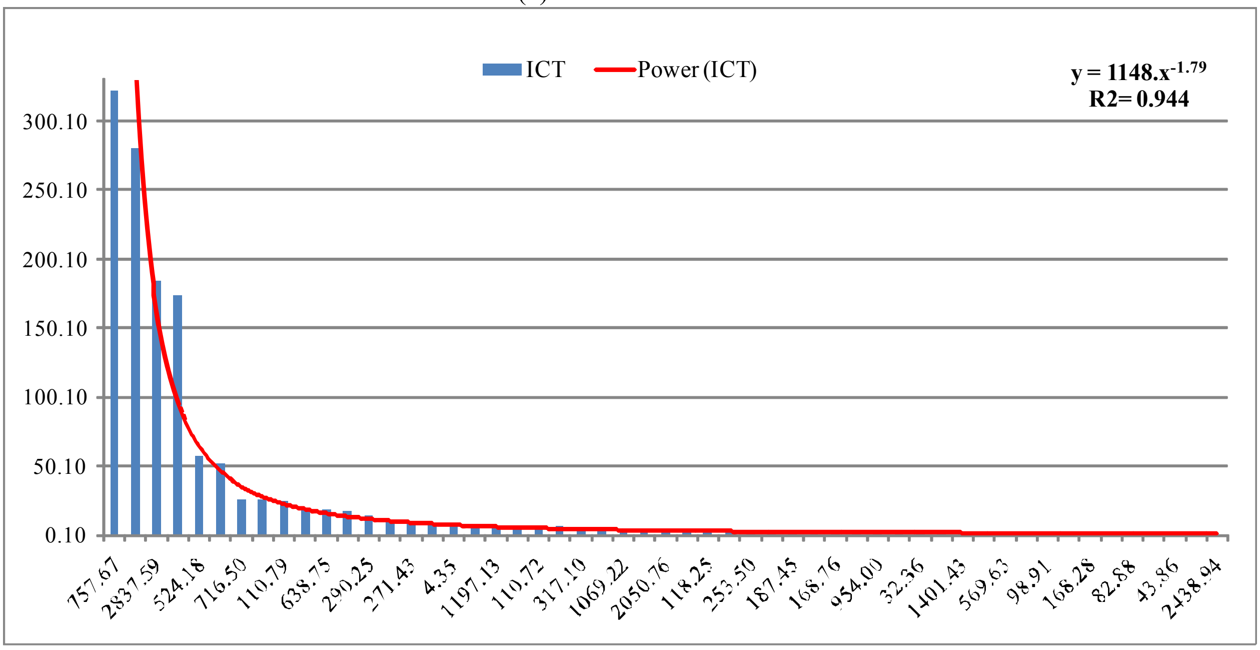
$$f(z) = 963.7 \cdot z^{-1-0.63} \quad (7)$$

The Pareto parameter with the expected statistical error is:

$$\alpha = 0.63 \pm 0.07 \quad (8)$$

where the expected statistical error on the exponent α is calculated by equation (5).

Power law between ICT adoption measured by composite indicator and profit per employee has also been analyzed. Results are given in Figure 3.



Source: Data from primary source, Authors' estimation of Pareto's parameters.

Fig.3. Power law distribution between ICT adoption indicator and the size of enterprises measured by profit per employee.

The corresponding density function, from Figure 3, for ICT indicator is:

$$f\left(z\right) = 1148 \cdot z^{-1-0.79} \tag{9}$$

The Pareto parameter is:

$$\alpha = 0.79 \pm 0.10. \tag{10}$$

From equations (6), (8), and (10), we can observe that the tail exponents of Pareto's distributions are smaller than 1, in all cases. Generally, empirical studies show that the value of α changes across different countries, and is typically in the range $1 < \alpha < 2$. In our three cases $\alpha < 1$, which means that ICT adoption in the Serbian organizations is undeveloped. In all cases the coefficient of determination (R^2) is pretty high (92% and 94%). These values indicate that the interpolation line very well fits the real ICT data.

5. CONCLUSIONS

This paper has two distinctive parts. After introduction remarks and explanation of power law and its occurrences outside physics, mostly in economy, sociology and business sciences, indicator which properties are tested is described. Formula proposed as composite indicator of ICT adoption is proposed as advancement over usual, mostly subjective methods of measuring ICT adoption in organization. It has already been successfully used in three studies, and in this study power law properties have been tested. The paper presents the occurrence of power law distribution of values of ICT adoption indicator in 67 selected enterprises. Described ICT indicator has been compared with firm size measured in number of employees, and profit per employee. In both cases, we found that the distribution of ICT adoption possesses properties of Pareto's law in the Serbian organization. Therefore, in all cases, ICT adoption is distributed according to Pareto's distributions with the tail exponent smaller than 1, which means that ICT adoption in the Serbian organizations is undeveloped. In all cases the coefficient of determination (R^2) is pretty high (above 90%), which indicates that the interpolation line very well fits the real ICT data. This is the first paper that shows the distribution of ICT adoption possesses properties of Pareto's law in the Serbian organizations.

6. ACKNOWLEDGEMENTS

This paper is a part of research projects numbers 47009 (*European integrations and social and economic changes in Serbian economy on the way to the EU*), 179015 (*Challenges and prospects of structural changes in Serbia: Strategic directions for economic development and harmonization with EU requirements*), and 47003 *"Infrastructure for electonically supported learning in Serbia"*, financed by the Ministry of Science and Technological Development of the Republic of Serbia. All remaining errors are the authors' responsibility.

7. REFERENCES

[1] Adamic, L. A. (2007), *Zipf, Power-laws, and Pareto-a ranking tutorial*, Xerox Palo Alto Researches Center, Palo Alto, CA 94304, http://www.hpl.hp.com /research/idl/papers/ranking/ranking.html

[2] Newman, M.E.J. (2005), *Power laws, Pareto distributions and Zipf's law. Contemporary Physics*, Vol. 46, No. 5, pp. 323 – 351.

[3] Atkinson, A. B, Piketty T., Saez E. (2009), *Top incomes in the long run of history*, Working Paper 15408, National Bureau of Economic Research.

[4] Klass, O. S., Biham O., Levy M., Malcai O., Solomon S. (2004), *The Forbes 400, the Pareto Wealth Distribution and Efficient Markets*, The European Physical Journal B, 55, pp. 143-147.

[5] Solomon, S. (1999), *Generalized Lotka-Volterra Models, in "Aplications of Simulations to Social Sciences"*, eds. G. Balot and G. Weisbuch; Hemes Science Publications. http://xxx.lanl.gov/PS_cache/cond- mat/pdf/ 9901/9901250v1.pdf

[6] Zipf, G. K. (1932), *Selected studies of the principle of relative frequency in language*. Harvard University Pree. Cambridge. MA

[7] Mandelbrot, B. B. (1997), *Fractals and Scaling* in Finance in France, Springer, Berlin.

[8] Axtell, R. L. (2001), *Zipf distribution of U.S. firm sizes*. Science, 293 (5536), pp. 1818-1820.

[9] Stanley, M. H. R. Amaral L. A. N., Buldyrev S. V., Havlin S., Leschhron H., Maass P., Salinger M. A., and Stanley H. E. (1996), *Scaling behavior in the growth of companies*, Nature, 379, pp. 804-806.

[10] Price, D.J. de S. (1965), *Networks of Scientific Papers*, Science,149 (3683), pp. 510-515.

[11] Redner, S. 1998. *How popular is your paper? An empirical study of the citation distribution*. European Physical Journal B, 4, pp. 131-134.

[12] Adamic, L. A., Huberman B. A. (2000), *Power-Law Distribution of the World Wide Web*, Science, Vol. 287, no. 5461, p. 2115.

[13] Barabási A. L., Albert R. (1999), *Emergence of scaling in random networks*. Science, 286, pp. 509-512.

[14] Abello, J., Buchsbaum A., and Westbrook J. (1998), *A Functional Approach to External Graph Algorithms,* in Proceedings of the 6th European Symposium on Algorithms (Springer, Berlin), pp. 332-343.

[15] Aiello, W., Chung F., and Lu L. (2000), *A random graph model for massive graphs*, in Proceedings of the 32nd Annual ACM Symposium on Theory of Computing (Association for Computing Machinery, New York), pp. 171–180.

[16] Miyazima, S. L. Y, Nagamine T., Miyajima H. (2000), *Power-law distribution of family names in Japanese societies*, Physica A, 278, pp. 282-288.

[17] Bentley, A., Ormerod P., and Batty M. (2008), *Evolution and turnover in scaling systems, Working paper series*, paper 135, UCL Centre for Advanced Spatial Analysis University College London.

[18] Atkinson, A.B, Piketty,T. (2007), *Top Incomes over the Twentieth Century*. Oxford, UK: Oxford Univ. Press

[19] Yamamoto, K., Miyazima S., Koshal K. R., Yamada Y. (2002), *A model for distribution of high-tax payers*, Physica A, 314, pp. 743-748.

[20] Aoyama, H., Souma W., Nagahara Y., Okazaki M. P., Takayasu H., Takayasu M. (2000), *Pareto's law for income of individuals and debt of bankrupt companies*. Fractals, 8, pp. 293-300.

[21] Minović, J., Minović M., Milovanović M. (2009), *Modeling wealth distribution for 100 reachest people in South-Eastern Europe*, Proceedings of the 10th International Symposium on Operational Research SOR' 09, Slovenian Society Informatika, Nova Gorica, Slovenija, pp. 545-552.

[22] Nirei M., Souma W. (2007), *A two factor model of income distribution dynamics*. Rev. Income Wealth 53, pp. 440–59.

[23] Kleiber, C., Kotz S. (2003), *Statistical Size Distributions in Economics and Actuarial Sciences*. New York: Wiley.

[24] Clauset, A., Shalizi C. R., and Newman M. E. J. (2007), *Power-law distributions in empirical data.* http:// arxiv.org/abs/0706.1062v1.

[25] Gabaix, X. (2009), *Power Laws in Economics and Finance*, The Annual Review of Economics, 1, pp. 255–93.

[26] Gagnon, M.P., Legare, F., Labrecque, M., Fremont, P., Pluye, P., Gagnon, J., Car, J., Pagliari, C., Desmartis, M., Turcot, L. & Gravel, K., (2009), *Interventions for Promoting Information and Communication technologies Adoption in Healthcare Professionals*, The Cochrane Collaboration,Vol 2, pp.1-31.

[27] Cragg, P.B., (2002), *Benchmarking information technology practices in small firms*, European Journal of Information Systems, Vol. 11, pp. 267-282.

[28] Caldeira, M.M. and Ward, J.M., (2003), *Using resource-based theory to interpret the successful adoption and use of information systems and technology in manufacturing small and medium sized enterprises*, European Journal of Information Systems,Vol.12, pp.127-141.

[29] Ma, X., Lombardo, S., Sciaba, E.S., (2008), *Learning mechanisms in a technology based organizational change programme: an exploratory study between Chinese and Norwegian corporations*, International Journal of Technology Management, Vol.41 No. 1/2, pp.75-95.

[30] Čudanov,M., Jaško,O., Jevtic,M., (2009), *Influence of Information and Communication Technologies on Decentralization of Organizational Structure.* Computer Science and Information Systems Journal, 6, pp. 93-108.

[31] Čudanov, M, Jaško, O, Savoiu, G, (2010), *Interrelationships of Organization Size, Information and Communication Technology Adoption*. Journal of Applied Quantitative Methods, 5/1, pp 29-40

[32] Čudanov M, Jaško O, (2010), *Adoption of information and communication technologies and dominant management orientation in organizations*. Behaviour and Information Technologies, online first available on: http://www.informaworld.com/smpp/content ~db=all~content=a923093688~frm=titlelink

[33] De Icaza, Alfonso and Pick, J., (1998), *Quantitative Model of Size and Complexity of Prospective Geographic Information Systems for Regional Governments in Mexico*. Proceedings of Americas Conference on Information Systems (AMCIS) Paper No. 134. pp. 396-398, Available from: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1561&context=amcis1998 [last accessed on 12.10.2011.]

[34] Radauceanu, A., Chouanière, D., Wild, P., Héry M., and. Créau Y., (2008), *Assessment of asbestos exposure in asbestos removal operations: Proposal of a task exposure matrix.* Archives des Maladies Professionnelles et de l'Environnement, Vol. 69, Iss. 4, pp. 586-592

[35] Thong, J.Y.L., and Yap, C.S., (1995), *CEO characteristics, organizational characteristics and information technology adoption in small businesses. Omega* – The International Journal of Management Science. 23, pp. 429-442.

[36] Thong, J.Y.L., Yap, C.S. and Ramen, K., (1996), *Top management support, external expertise and information systems implementations in small businesses*, Information Systems Research, 7, pp. 248-267.

[37] Howard, K., (1997), *IT Means Business? A survey of attitudes in smaller businesses to information and technology*, London, UK: Institute of Management.

[38] Redoli, J., Mompó, R., García-Díez, R.J., López-Coronado, M., (2008), *A model for the assessment and development of Internet-based information and communication services in small and medium enterprises*, Technovation, Vol. 28 No. 7, pp. 424-435.

[39] Moore, G.C., and Benbasat, I., (1991), *Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation*. Information Systems Research, 2, pp. 192-222.

[40] Golubeva, A., Merkuryeva I., (2006), *Demand for Online Government Services: Case Studies from St. Petersburg.* Information Polity Vol.11,No.3–4, pp.241–54.

[41] El-Mashaleh, M.S., (2007), *Benchmarking information technology utilization in the construction industry in Jordan*. Journal of Information Technology in Construction, Vol. 12, No. 19, pp. 279-291.

[42] Kauffman, R. J., and Kumar, A., (2005). *A critical assessment of the capabilities of five measures for ICT development*. MIS Research paper, available from: http:// citeseerx.ist.psu.edu/viewdoc/download [last accessed 04 October 2011]

[43] Hanafizadeh, P., Hanafizadeh, M.R., Khodabakhshi, M., (2009), *Extracting Core ICT Indicators Using Entropy Method.* The Information Society, Vol. 25, pp. 236–247.

[44] Scheer, A.W., and Habermann, F., (2000), *Making ERP a Success*. Communications of the ACM,43,pp.57-61.

[45] Scheer, A.W, and Schneider, K., (2006), *ARIS — Architecture of Integrated Information Systems*, in Bernus, P., Mertins, K., and Schmidt, G., International Handbooks on Information Systems, 2006, Part Three - Handbook on Architectures of Information Systems, pp. 605-623.

[46] Tarafdar, M. and Vaidya S.D., (2006), *Challenges in the Adoption of E-Commerce Technologies in India: The Role of Organizational Factors*. International Journal of Information Management, vol. 26, pp. 428-441.

[47] Vogel, D.R., Orwig, R., Dean, D., Lee, J. and Arthur, C., (1993), *Reegineering with Enterprise Analyzer.* Proceedings of the 26th Annual Hawaiian International Conference on System Science (HICSS), Los Alamitos, California, USA, Vol 3, pp. 127 – 136.

[48] Van den Hoven, J., (1999), *Information resource management: Stewards of data.* Information Systems Management, Vol. 16 No. 188-91.

UNCERTAINTY IN EU STOCK MARKETS BEFORE AND DURING THE FINANCIAL CRISIS

[1]Daniel _Traian_ Pele and [2]Miruna_Mazurencu-Marinescu

[1,2] Academy of Economic Studies, Bucharest, Department of Statistics and Econometrics, Bucharest, Romania
[1]e-mail: danpele@ase.ro, [2]e-mail: miruna@ase.ro

Abstract. *The aim of this paper is to study the uncertainty behaviour of the EU stock markets before and during financial crisis by using daily data for main stock market indexes of EU countries. The results show a significant difference in the uncertainty pattern of EU stock markets before and during financial crisis. Also, a large variability among EU stock markets in uncertainty patterns is noticed and this fact could be explained trough inequalities in stock market performances and economic development.*

Keywords: *stable distributions, financial crisis, uncertainty, entropy*

1. INTRODUCTION

The financial crisis had a severe impact on stock markets across European Union time, following the pattern from US stock market. This impact could be revealed by using a key feature of financial crisis, which is a higher level of uncertainty.

The level of uncertainty of a stock market could be analyzed by using two complementary approaches: information theory and statistical modelling. From an information theory point of view, a robust measure of uncertainty is entropy; in classical form, the Shannon entropy is positively correlated with the level of uncertainty. Also, by using the statistical approach, the probability in the tails of the returns' distribution is an estimate of uncertainty level of a stock market.

As a measure of stock market uncertainty we are using the entropy of the distribution function of returns, while the tail behaviour of returns' distribution is captured by using alpha-stable distributions.

Given the probability distribution of returns is essential for any statistical inference concerning the stock market. In general, the major distribution characterizing the evolution of returns is the normal distribution (Gaussian) or its derivatives (e.g. log-normal distribution).

Recent papers (Rachev et al., 2000 and 2010) show that stable distributions are a much better approach than classical distributions in financial modeling. The fact that the observed distribution of returns is heavy-tailed cannot be explained by a normal distribution.

The relationship between stable distributions and financial crisis has been previously addressed by Barunik, Vacha and Vosvrda (2010). In this study, they are estimating parameters of stable distributions of US and Central Europe stock markets, using daily and intraday data. By analyzing the distribution of returns for 2005-2009, and separately for the periods 2005-2007 (before the financial crisis) and 2007-2009 (the crisis), the authors conclude that there is a significant difference between the probability distribution of returns before and during the financial crisis.

Thus, the pre-financial crisis period does not present large deviation from normal distribution, while the crisis period is characterized by a significant deviation from normality.

In this paper we are using daily data for main stock market indexes of EU-27 countries in order to study the uncertainty behaviour of these stock markets before and during financial crisis.

The study is structured as follows: a first section contains a theoretical presentation of measures of uncertainty, in the second section the main results are presented and the last section is dedicated for conclusions.

2. MEASURES OF UNCERTAINTY

2. 1. Stable distributions

Stable distributions is a class of distributions which have the property of being invariant under linear combinations; Gaussian distribution is a special case of stable distributions.

The difficulty that occurs for stable distributions is that in most cases, an explicit form of probability density function is not known, only the expression of its characteristic function. Thus, a random variable X follows a stable distribution with parameters $(\alpha,\beta,\gamma,\delta)$ (Nolan, 2011) if exists $\gamma > 0, \delta \in \Re$ such as X and $\gamma Z + \delta$ have the same distribution, where Z is a random variable with the following characteristic function:

$$\phi(t) = \mathbf{E}[e^{itZ}] = \begin{cases} \exp(-|t|^{\alpha}[1 - i\beta \tan(\frac{\pi\alpha}{2}) sign(t)]), & \alpha \neq 1 \\ \exp(-|t|[1 + i\beta t \frac{2}{\pi} sign(t)(\ln(|t|)]), & \alpha = 1 \end{cases}.$$

In the above notations $\alpha \in (0.2]$ is the stability index, controlling for probability in the tails (for Gaussian distribution $\alpha = 2$), $\beta \in [-1,1]$ is the skewness parameter, $\gamma \in (0,\infty)$ is the scale parameter and $\delta \in \mathbf{R}$ is the location parameter.

A random variable X follows a stable distribution $S(\alpha,\beta,\gamma,\delta;0)$ if its characteristic function has the following form:

$$\phi(t) = \mathbf{E}[e^{itX}] = \begin{cases} \exp(-\gamma^{\alpha}|t|^{\alpha}[1 + i\beta\tan(\frac{\pi\alpha}{2})sign(t)(|\gamma t|^{1-\alpha} - 1)] + i\delta t), \alpha \neq 1 \\ \exp(-\gamma|t|[1 + i\beta t\frac{2}{\pi}sign(t)(\ln(|\gamma t|)] + i\delta t), \alpha = 1 \end{cases} . \quad (1)$$

A random variable X follows a stable distribution $S(\alpha,\beta,\gamma,\delta;1)$ if its characteristic function has the following form

$$\phi(t) = \mathbf{E}[e^{itX}] = \begin{cases} \exp(-\gamma^{\alpha}|t|^{\alpha}[1 - i\beta\tan(\frac{\pi\alpha}{2})sign(t)] + i\delta t), \alpha \neq 1 \\ \exp(-\gamma|t|[1 + i\beta t\frac{2}{\pi}sign(t)(\ln(|t|)] + i\delta t), \alpha = 1 \end{cases} . \quad (2)$$

Parametrisation $S(\alpha,\beta,\gamma,\delta;1)$ has the advatage that is more suitable for algebric manipulations, altough its characteristic function is not continuous for all parameters.

Parametrisation $S(\alpha,\beta,\gamma,\delta;0)$ is suitable for numerical simulations and statistical inference, altough the expression of characteristic function is more difficult to utilise in algebric calculus. Nolan (2011) shows that the two parametrisations are equivalent; if $X \sim S(\alpha,\beta,\gamma,\delta_1;1)$ and $X \sim S(\alpha,\beta,\gamma,\delta_0;0)$, then

$$\delta_0 = \begin{cases} \delta_1 + \beta\gamma\tan\frac{\pi\alpha}{2}, \alpha \neq 1 \\ \delta_1 + \beta\frac{2}{\pi}\gamma\ln\gamma, \alpha = 1 \end{cases} \quad (3)$$

The behavior of stable distributions is driven by the values of stability index α : small values are associated to higher probabilities in the tails of the distribution.

2.2. Information entropy

Information entropy is the most widely used measure of uncertainty, its applications covering a wide range, from physics to economics and biology. The concept of entropy originates from physics in the 19th century; the second law of thermodynamics stating that the entropy of a system cannot decrease other way than by increasing the entropy of another system. As a consequence, the entropy of a system in isolation can only increase or remain constant over time.

If the stock market is regarded as a system, then it is not an isolated system: there is a constant transfer of information between the stock market and the real economy. Thus, when information arrives from (leaves to) the real economy, then we can expect to see an increase (decrease) in the entropy of the stock market, corresponding to situations of increased (decreased) randomness.

Most often, entropy is used in one of the two main approaches, either as Shannon Entropy – in the discrete case – or as Differential Entropy – in the continuous time case. Shannon Entropy quantifies the expected value of information contained in a realization of a discrete random variable. Also it is a measure of uncertainty, or unpredictability: for a uniform discrete distribution, when all the values of the distribution have the same probability, Shannon Entropy reaches his maximum. Minimum value of Shannon Entropy corresponds to perfect predictability, while higher values of Shannon Entropy correspond to lower degrees of predictability (the minimum value of Shannon Entropy is 0).

Differential Entropy is an extension of Shannon Entropy to the continuous case, but is not a good measure of uncertainty; as it can take negative values and in addition, it is not invariant to some linear transformations.

Dioniso et al. (2006) provide a review of the theoretical and empirical work about the entropy and the variance as measures of uncertainty. Several conclusions could be drawn from this review: first of all, the entropy is a more general measure of uncertainty than the variance or the standard deviation (Philippatos and Wilson, 1972), since the entropy depends on more characteristics of a distribution as compared to the variance and may be related to the higher moments of a distribution (Ebrahimi et al., 1999). Secondly, both the entropy and the variance reflect the degree of concentration for a particular distribution, but their metric is different. While the variance measures the concentration around the mean, the entropy measures the diffuseness of the density irrespective of the location parameter (Ebrahimi, Maasoumi and Soofi, 1999).

In this paper we use a recently developed concept, the entropy of a function (Lorentz, 2009) in order to estimate the entropy of a distribution function,by employing very general assumptions and in a non-parametric context.

Basically, our methodology involves the following steps *to estimate the entropy of a distribution function (Lazar et al.(2011)). For a sample $X_0,...,X_{n-1}$ of i.i.d. observations drawn form the distribution F:*

Step 1. Estimate the distribution function using a Kernel Estimator or Empirical Distribution Estimator, obtaining values $\hat{F}_n(X_i)$ for $i = 0,..,n-1$;

Step 2. Sampling from the distribution function, using the sampled function $S_n(\hat{F}_n)(i) = \hat{F}_n(X_i)$ for $i = 0,..,n-1$;

Step 3. Define a quantum $q > 0$; then

$Q_q S_n(\hat{F}_n)(j) = (i+1/2)q$, if $\hat{F}_n(X_j) \in [iq,(i+1)q)$;

Step 4. Compute the probabilities

$$p_n(i) = \frac{c_n(i)}{\sum_j c_n(j)} = \frac{c_n(i)}{n} = \frac{card\{\hat{F}_n(X_j) \in [iq,(i+1)q)\}}{n};$$

Step 5. The estimator of entropy of distribution function is then:

Step 6. $H_q(\hat{F}_n) = -\sum_i p_n(i)\log_2 p_n(i)$. (4)

In order to insure comparability among various distributions, one can define a normalized entropy,

as a ratio between the entropy and the entropy of uniform distribution:

$$NH_q(\hat{F}_n) = -\sum_i p_n(i)\log_2 p_n(i) / \log_2 n \in [0,1]. \quad (5)$$

In the following we will refer to entropy as the normalized entropy, taking values between 0 and 1.

Low values of entropy are associated with heavy-tailed distributions, while high values of entropy correspond to Gaussian distribution. In other words, as the tails probability is higher, the expected value of entropy is lower.

3. DATA AND EMPIRICAL RESULTS

In order to asses the impact of financial crisis on uncertainty behavior of stock market indexes for EU countries, we use a sample of daily observations for 26 European countries. Starting from observed index ? p_t, we compute the logreturns as $r_t = \log p_t - \log p_{t-1}$ and using the methodology described above, we estimate the entropy of the returns distribution function and also the stability index α for stable distribution (we use STABLE. EXE, available on http://academic2.american.edu/~jpnolan /stable/stable.html)

Stock market indexes

Table no. 1.

Country	Country code	Index	2005/2007 Entropy	2008/2010 Entropy	2005/2007 α	2008/2010 α
Malta	MT	MSE	0.613	0.640	1.283	1.373
Bulgaria	BG	SOFIX	0.693	0.622	1.414	1.371
Slovenia	SI	SBITOP	0.661	0.590	1.482	1.501
Cyprus	CY	CYSMMAPA	0.593	0.742	1.578	1.809
Lithuania	LT	VILSE	0.669	0.565	1.590	1.443
Portugal	PT	PSI20	0.589	0.528	1.629	1.592
Ireland	IR	ISEQ	0.642	0.645	1.641	1.682
Latvia	LV	RIGSE	0.597	0.675	1.663	1.695
Sweden	SE	OMX Stockholm	0.675	0.687	1.668	1.655
Denmark	DK	OMX Copenhagen	0.675	0.692	1.668	1.657
Romania	RO	BET	0.584	0.659	1.680	1.637
UK	UK	FTSE100	0.687	0.614	1.682	1.598
Belgium	BE	BEL20	0.751	0.654	1.706	1.714
Czech Republic	CZ	PX50	0.609	0.559	1.718	1.572
Austria	AT	ATX20	0.617	0.672	1.729	1.672
Greece	GR	ASE	0.633	0.719	1.741	1.808
Luxemburg	LU	LUXX	0.596	0.642	1.770	1.738
Netherlands	NL	AEX	0.744	0.631	1.773	1.566
Spain	ES	IBEX	0.729	0.601	1.776	1.684
Finland	FI	OMXH15	0.813	0.701	1.805	1.724
Poland	PL	WIG	0.694	0.706	1.854	1.667
France	FR	CAC40	0.787	0.634	1.857	1.648
Italy	IT	FTSEMIB	0.752	0.649	1.858	1.671
Hungary	HU	BUX	0.775	0.610	1.871	1.699
Germany	DE	DAX	0.809	0.630	1.881	1.620

The results obtained for the two subsamples show significant differences between EU countries from the point of view of stock market uncertainty.

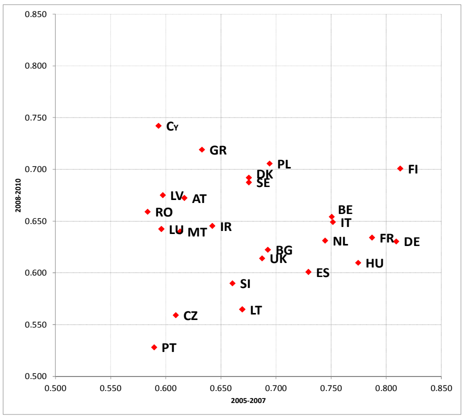


Fig. 1. Behavior of entropy : 2005-2007 vs 2008-2010

Thus, for 2005-2007 period, before the financial crisis, one can distinguish three clusters of countries, based on entropy behavior:

- first cluster – countries with low uncertainty (high entropy) is formed by the following countries: Finland, Germany, France, Hungary, Netherlands, Italy, Belgium, Spain;
- second cluster – countries with moderate uncertainty is formed by the following countries: Lithuania, UK, Bulgaria, Poland, Denmark, Sweden, Slovenia;
- third cluster – countries with high uncertainty is formed by the following countries: Portugal, Greece, Czech Republic, Malta, Luxembourg, Romania, Cyprus, Latvia, Ireland.

The situation is slighly different in period 2008-2008 (financial crisis): for the entire sample of countries, entropy is lower than in period 2005-2007, indicating an increased likelihood of extreme events on stock market and a higher degree of uncertainty.

Thus, for 2008-20010 period, during the financial crisis, one can distinguish three clusters of countries, based on entropy behavior:

- first cluster – countries with low uncertainty (high entropy) is formed by the following countries: Greece, Cyprus, Poland, Finland;
- second cluster – countries with moderate uncertainty is formed by the following countries: Germany, France, Hungary, Netherlands, Italy, Belgium, Spain, UK, Bulgaria, Denmark, Sweden, Greece, Malta, Luxembourg, Romania, Cyprus, Ireland;
- third cluster – countries with high uncertainty is formed by the following countries: Portugal, Czech Republic, Lithuania, Slovenia.

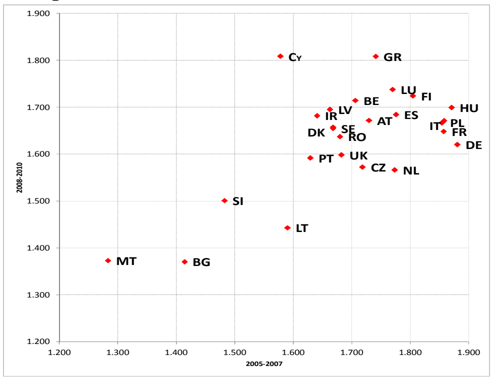


Fig.2. Behavior of α (stability index):2005-2007 vs 2008-2010

The estimation of stability index α of the stable distribution shows a different clustering behavior of EU stock markets: for most of the countries, there is a departure from normality induced by the financial crisis. Two extreme clusters could be identified based on this criterion:

- countries with extreme departure from normality: Malta, Bulgaria, Slovenia, Lithuania;
- countries for which stability index α shows a behavior close to Gaussian distribution: Greece and Cyprus.

Return and volatility

Table no. 2. - %-

Country	Country Code	Index	Return 2005/ 2007	Return 2008/ 2010	Volatility 2005/ 2007	Volatility 2008/ 2010
Malta	MT	MSE	0.068	-0.033	0.879	0.795
Bulgaria	BG	SOFIX	0.137	-0.212	1.035	1.841
Slovenia	SI	SBITOP	0.133	-0.144	0.878	1.492
Cyprus	CY	CYSMMAPA	0.206	-0.203	1.523	3.021
Lithuania	LT	VILSE	0.079	-0.031	0.986	1.654
Portugal	PT	PSI20	0.064	-0.091	0.678	1.859
Ireland	IR	ISEQ	0.015	-0.115	1.081	2.300
Latvia	LV	RIGSE	0.049	-0.055	0.975	1.895
Sweden	SE	OMX Stockholm	0.050	0.009	1.106	1.933
Denmark	DK	OMX Copenhagen	0.050	0.009	1.106	1.943
Romania	RO	BET	0.110	-0.083	1.709	2.434
UK	UK	FTSE100	0.039	-0.012	0.844	1.730
Belgium	BE	BEL20	0.045	-0.061	0.841	1.754
Czech Republic	CZ	PX50	0.075	-0.052	1.146	2.237
Austria	AT	ATX20	0.083	-0.059	1.169	2.365
Greece	GR	ASE	0.083	-0.174	1.022	2.257
Luxemburg	LU	LUXX	0.083	-0.059	0.941	1.949
Netherlands	NL	AEX	0.051	-0.049	0.840	1.979
Spain	ES	IBEX	0.067	-0.057	0.850	2.027
Finland	FI	OMXH15	0.033	-0.010	1.252	2.096
Poland	PL	WIG	0.098	-0.021	1.199	1.664
France	FR	CAC40	0.050	-0.051	0.909	1.952
Italy	IT	FTSEMIB	0.029	-0.085	0.822	2.010
Hungary	HU	BUX	0.077	-0.016	1.408	2.258
Germany	DE	DAX	0.084	-0.020	0.908	1.842

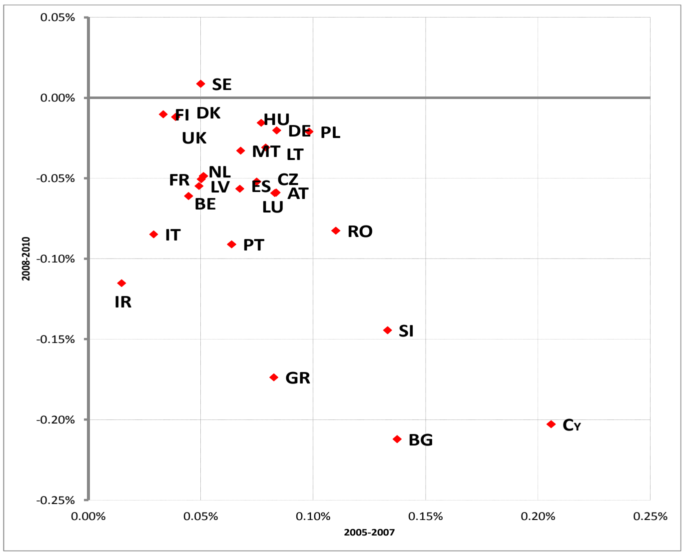


Fig. 3. Performance of stock market: 2005-2007 vs 2008-2010

The analysis of the European stock markets could be performed also in terms of performance (returns) and volatility; what financial crisis brings from the point of view of stock market performance is a significant drop in average daily returns and also a significant increase in volatility.

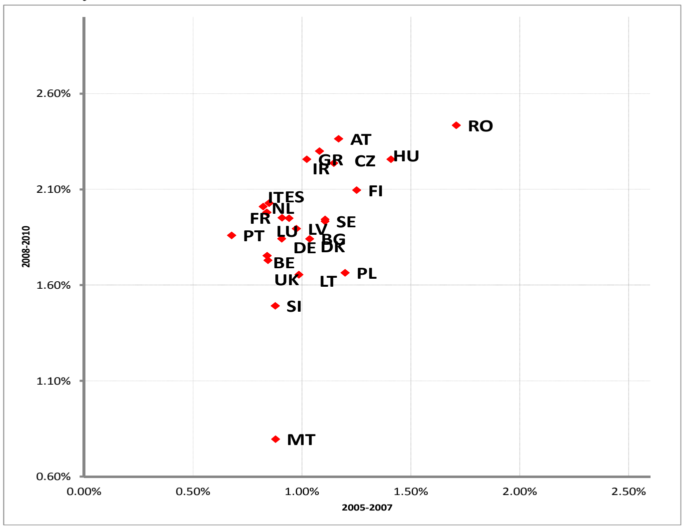


Fig. 4. Volatility of stock market: 2005-2007 vs 2008-2010

From the point of view of average daily returns before and during the financial crisis, most of European stock markets have a similar behavior, with few notable exceptions: Greece, Cyprus, Bulgaria and Slovenia. For these countries, the financial crisis had a severe impact on stock market performance; for example, Cyprus had a drop in average daily return from 0.2% to -0.2%, while Greece average daily return dropped from 0.08% to -1.7%.

Also in terms of volatility, both before and during financial crisis, we can observe a cluster of countries with the highest volatilities: Greece, Austria, Romania, Ireland, Hungary, Czech Republic, Finland, Hungary.

As the above results show, the impact of financial crisis on European stock markets was not uniform across countries; there are significant differences in terms of performance, volatility, uncertainty and behavior towards normality.

4. CONCLUSIONS

Using daily data for main stock market indexes of EU-27 countries, we have studied the behaviour of these stock markets before and during the financial crisis.

The analysis was conducted on two directions, looking for significant differences between properties of return distribution and also looking for homogenous groups of countries based on stock market indicators.

From a distributional point of view, most of the countries exhibit large departure from normality during financial crisis, values of stability index α being significantly lower than 2(the case of Gaussian distribution).

The same conclusion was revealed using entropy of distribution function of returns as an estimator of stock market uncertainty. For majority of countries form our sample, there is a clear movement towards high uncertainty levels during financial crisis.

Further research need to be conducted in order to explain this large variability among European stock markets in terms of uncertainty patterns, perhaps trough existing inequalities in stock market and economic development.

5. ACKNOWLEDGMENTS

This paper was cofinanced from the European Social Fund, through the Human Resources Development Operational Sectorial Program 2007-2013, project number POSDRU/89/1.5/S/59184, „Performance and excellence in the post-doctoral economic research in Romania".

6. REFERENCES

[1] Barunik, J., Vacha, L., Vosvrda, M.(2010), *Tail Behavior of the Central European Stock Markets during the Financial Crisis,* Working Papers IES 2010/04, Charles University Prague, Faculty of Social Sciences, Institute of Economic Studies.

[2] Dionisio, A., Menezes, R., and Mendes, D. A., (2006), *An econophysics approach to analyse uncertainty in financial markets:An application to the Portuguese stock market,* European Physics Journal B, vol 50, pp. 161–164.

[3] Ebrahimi, N., Maasoumi, E., and Soofi, E. S. (1999), *Ordering Univariate Distributions by Entropy and Variance*., Journal of Econometrics, vol 90, pp. 317-336.

[4] Ebrahimi, N., Massoumi, E., and Soofi, E. S, (1999), *Measuring Informativeness of Data by Entropy and Variance,* Advances in Econometrics: Income Distribution and Methodolgy of Science, Essays in Honor of Camilo Dagumi, D. Slottje (ed), Chapter 5, Springer Verlag.

[5] Kim, Y., Rachev, S., Bianchi, M., Mitov, I. and Fabozzi, F.(2010), *Time Series Analysis For Financial Market Meltdowns*, No 2, Working Paper Series In Economics, Karlsruhe Institute Of Technology (KIT), Department Of Economics And Business Engineering.

[6] Lazar, E., Dufour, A., Pele, D.T., Țepuş, A.M. (2011), *Information Entropy and Measures of Market Risk*, Unpublished manuscript.

[7] Lorentz, R. (2009), *On the entropy of a function,* Journal of Approximation Theory, vol 158, (2) (June 2009), pp. 145-150.

[8] Maasoumi, E. and J. Racine (2002), *Entropy and predictability of stock market returns,* Journal of Econometrics, Elsevier, vol. 107(1-2), pp. 291-312,

[9] Nolan, J. P. (2011), *Stable Distributions - Models for Heavy Tailed Data*, Boston: Birkhauser.Unfinished manuscript, Chapter 1 online at academic2.american. edu/_jpnolan.

[10] Nolan, J. P. (2011), *Stable Distributions - Models For Heavy Tailed Data*, Boston:

[11] Birkhauser, J. (2011), *Unfinished Manuscript*, Chapter 1 Online At Academic2.American.Edu/_Jpnolan.

[12] Philippatos, G. C. and Wilson, C.(1972). *Entropy, Market Risk and the Selection of Efficient Portfolios*, Applied Economics, 4, 209-220.

[13] Rachev, S. And Mitnik S. (2000). *Stable Paretian Models* In Finance, John Wiley, Series In Financial Economics And Quantitative Analysis, Chechester, New York.

INTERDISCIPLINARITY AND SPHERES OF MEANING - THE CASE OF INFORMATION SYSTEMS

Mila_Mitic
Belgrade University, Mihajlo Pupin Institute, Belgrade, e-mail: mila.mitic@pupin.rs

Abstract. Abstract. *Solving complex social problems requires interdisciplinary research. However, there are a lot of barriers to successful interdisciplinary studies. It is considered that philosophy could help in understanding of interdisciplinarity. It has been recognized a need for better understanding human life for improving inter-disciplinarity, as well as a need for understanding barriers impeding interdisciplinarity. In this paper it is considered how the philosophy of Dooyeweerd, who has interested in human everyday experience, recognizes barriers impeding interdisciplinarity and how his spheres of meaning could help in building understanding of human life needed for better interdisciplinary activities. That consideration, together with extant deliberation of definition and dignity of the IS phenomenon (in which problems are similar to ones of other interdisciplinary phenomena), suggests a way of thinking, a multi-aspectual one, towards better understanding of interdisciplinarity.*

Keywords: *Interdisciplinarity, Dooyeweerd, Spheres of meaning, Information systems*

1. INTRODUCTION

Interdisciplinarity has become synonymous with all things modern, creative and progressive about scientific research. The interdisciplinary imperative has arisen from complexity of existing problems ([9]). However, interdisciplinary researches are often unsuccessful because there are differences in values, differences in theories and differences in epistemologies in them ([11]). Besides that, extant social view on importance of different disciplines is often a barrier to success of interdisciplinarity. It has been recognized a need for better understanding human life for improving interdisciplinarity ([1], [15]), as well as a need for understanding barriers impeding interdisciplinarity ([11]).

Since theoretical thought always tends to narrow and distort human understanding of everyday experience, it is considered that philosophy, as an integrative discipline, could help in understanding interdisciplinarity itself ([2]).

Dooyeweerd was a philosopher who was interested in everyday experience. His interest was in the whole breadth and depth of reality. He differentiated two "sides" of reality: entities and laws. Entities are subject to law in their functioning and, in their structure. Functional laws are relating to functioning (activities) of entities as subjects or objects. They are called aspectual laws too. Dooyeweerd recognized the suite of aspects, which are irreducible and in harmony. An aspect is distinct sphere of meaning, distinct way of being, distinct rationality, distinct mode of being, distinct way of functioning, distinct sphere of law, distinct kind of normativity, distinct type of repercussion, and distinct way of knowing ([2]).

A brief description of interdisciplinarity is given in the section 2, and a brief presentation of Dooyeweerd's theory of modal aspects in given the section 3 of this paper. After that, in the section 4 it is considered how Dooyeweerd's philosophy recognizes barriers impeding interdisciplinarity and how his spheres of meaning could help in building understanding of human life needed for better interdisciplinary activities.

Dooyeweerd's philosophy has been applied in the IS field, which requires some interdisciplinary combinations of the social and technical spheres of organizational studies and computer science. The description of IS as an interdisciplinary field and the case of applying Dooyeweerd's suite of aspects for defining and dignifying the IS phenomenon ([3]) is presented in the section 5.

Concluding remarks, given in the section 6, are relating to possibilities of using the spheres of meaning as a way of thinking towards better understanding of inter-disciplinarity.

2. INTERDISCIPLINARITY

2.1 Definition

There are many different interpretations and definitions of interdisciplinarity. They disagree on details, but they do agree that the participating scientists work together on a common question by somehow exchanging concepts and tools in order to solve problem.

Nowadays interdisciplinarity is seen as communication and collaboration across academic disciplines. Interdisciplinarity is supposed to integrate knowledge and solve problems that individual disciplines cannot solve alone ([9]). Interdisciplinary studies may be defined as a process of answering a question, solving a problem, or addressing a topic that is too broad or complex to be dealt with adequately by a single discipline or profession ([15]). Some previous interdisciplinary fields have become disciplines. There is an attitude that when inter-disciplinarity is successful it becomes a discipline ([1]).

Nowadays the degree of mutual dependency of scientific fields is high. Intellectual fields generally exhibit (among other things) weakened boundaries, increased mobility of ideas and skills across those boundaries, and increased inter-field coordination of research objectives, strategies, and results ([9]).

2. 2 Terminological ambiguity

The literature is characterized by considerable terminological ambiguity. The different terms ("intradisciplinary", "crossdisciplinary", "multi-disciplinary", "interdisciplinary", and "transdisciplinary") are used to distinguish between low, moderate, and high levels of interconnectedness or intellectual integration.

Intradisciplinary analysis involves work within a single discipline. Crossdisciplinary activity views one discipline from the perspective of another. Multidisciplinary and interdisciplinary analyses draw on the knowledge of several disciplines. In multidisciplinary analysis each discipline provides a different perspective on a problem or issue and makes a contribution to the overall understanding of the issue, but in a primarily additive fashion. Interdisciplinary analysis, on the other side, requires integration of knowledge from the disciplines so that the resulting understanding is greater than the sum of its disciplinary parts. Transdisciplinary analysis is concerned with the unity of intellectual frameworks beyond the disciplinary perspectives; it may deal with philosophical questions about the nature of reality and the nature of knowledge systems that transcend disciplines ([15]).

However, everybody doesn't make such distinctions and uses the terms "interdisciplinary" and "interdisciplinarity" as general ones for describing interrelationships among academic disciplines. In this paper these terms are used in that general meaning.

2.3 Significance of interdisciplinarity

In recent years, interdisciplinarity has become synonymous with all things modern, creative and progressive about scientific research. Almost every new research effort calls itself interdisciplinary. Interdisciplinarity has become a kind of sales argument. It has become a label of good research ([9]).

The interdisciplinary imperative has arisen from complexity of existing problems. Thinking collectively about complex problems requires crossing boundaries both horizontally (across disciplines) and vertically (across experts, policymakers, practitioners, and the public) ([11]).

There are assumptions that interdisciplinary research has vast potential for societal good in the form of new kinds of knowledge. It is believed that more interdisciplinary research is better than less ([9]).

2.4 Barriers to interdisciplinarity

There are numerous barriers impeding inter-disciplinarity: the difference in values, the difference in theories and models, the difference in epistemology, the difference in the way in which society interacts and organizes academia; the relative absence of motivation ([11]).

Values are embedded in all inquiries and at all stages: in the choice of questions, theoretical positions, and style of research. However, scientists are loath to acknowledge that. Since they are expected to provide "objective" advice for problem solving, acknowledgment of value not-neutrality and importance of researchers' normative positions is even more difficult. Because of that unrecognized difference in values, the collective judgment required in interdisciplinary research is especially difficult.

The same phenomenon is often studied by different theories or models in different disciplines. The superiority of one theory over another in a particular case is difficult to prove. Commitment to one's school of thought often is so important that a need for exploring assumptions embedded in used theories and models and their suitability for considered context is neglected.

The epistemic barriers have been the most emphasized ones in the literature on interdisciplinarity. They involve incompatible styles of thought, research traditions, techniques, and different languages ([9]). Very often there is epistemological sovereignty in interdisciplinary research, i.e., the research ends up entitling a single discipline or epistemology, incorporating others only in a support or service role ([12]). Research questions for a project are formulated and then, because of the complexity of the system under study, a scientist from another discipline is invited to help in the investigation. But, the research usually remains framed by the theory of the researcher who formulated research questions, what limits the scope of other researcher's contribution.

Because of epistemic differences, a researcher in the interdisciplinary team often finds that other members of the team define the problem quite differently or seek different types of answers. That's why too much effort for communication and sharing knowledge within the team is required. Moreover, collaboration between scientists even within the same broad area can be difficult from the same reason ([11]).

The way in which society interacts with and organizes academia influences on interdisciplinary research. The importance of a certain discipline or a particular disciplinary crossing is often determined by society, i.e., outside academia. There are significant differences in the manner in which society treats the social and natural sciences. There is a deep-rooted belief of the superiority of the natural sciences over social ones. That's why, for example, social problems in which some importance of the technical dimension is recognized are often solved by teams of researchers from the natural sciences. Extant social viewpoint on different importance of sciences contributes to the relative absence of motivation of researchers from some disciplines to work together ([11]). Society doesn't value enough the problem driven knowledge because of its lack of abstraction ([9]).

The superficial success of the label "interdisciplinarity" goes together with an enormous resistance against interdisciplinarity by the most powerful groups in society ([1]). Whole human culture founded its success in fragmentation and specialization of labor. The idea of interdisciplinarity is, however, contrary to the idea of the labor division in the intellectual domain ([9]).

2.5. A need for changes in understanding human life

It has been recognized that some changes in understanding human life are needed in order to improve interdisciplinarity. There is a need for answering some basic questions such as "What is a human being? How does a human being function normally in a normal environment?" ([1]). There is a need for understanding and appreciating diversity of human life. There is a need to understand that drivers of human behavior seen from different views (for example, material benefits from view of mainstream economists, power from the view of certain schools within sociology, and cultural norms and value systems from certain schools within anthropology) aren't mutually incompatible ([11]). There is a need for epistemological pluralism which recognizes that, in any given research context, there may be several valuable ways of knowing, and that accommodating this plurality can lead to more successful integrated study ([12]).

Social science theories and their adherents have to take into account the constraints imposed by natural resources and processes on human actions ([11]). The dimensions of values, interests and power have to be acknowledged in any knowledge building ([16]). Participants in interdisciplinary research projects must overcome various biases and prejudices that accompany disciplinary training; they need to be self-reflective about the value judgments embedded in their choice of theories and models, willing to give respect to and also learn more about the "other," and able to work with new models and theories used by others ([11]).

Reflecting on how to think across academic disciplines is only a first step toward bridging the various divides involved in collectively addressing complex problems ([11]). It would be clear what means that fields across the natural and social sciences are well-connected in the web of intellectual ties. There is a need for system wide shift in structuring academic careers ([9]).

An understanding how people function in everyday life might be helpful in understanding interdisciplinarity. In order to make everyday decisions, interpret phenomena, and generally make sense of the world, people do informal interdisciplinary analysis drawing on and integrating diverse information ([15]). They have multiple personalities (for example, the economic one in the market, the political one in elections, and some other personality in the interaction with other people). They are driven by different factors in different activities.

4. THEORY OF MODAL ASPECTS

The diversity and the richness of human life require consideration of several aspects. Dooyeweerd's theory of modal aspects, presented in this paper according to [2] and [3], shows how aspects account for diversity and coherence, being and doing, normativity, etc.

Spheres of meaning

Aspects are spheres of meaning. Each aspect is some kind of origin, which enables being, doing, knowing, and the like.

Aspects cannot be directly observed, but only as they are expressed in things, events, situations, and so on, as ways these can be meaningful. Each entity has aspect(s) that determine its nature, its purpose. Each aspect is centered on its kernel meaning.

Dooyeweerd formed a list of 15 aspects of everyday experience: quantitative, spatial, kinematic, physical, biotic, sensitive, analytical, formative, lingual, social, economic, aesthetic, juridical, ethical, and pistic. The list of aspects with their kernel meaning is given in the table **Eroare! Fără sursă de referință.** (columns "norms" and "sciences" are used later).

Aspectual meaning is grasped by the intuition. Intuition isn't absolute and it is subject to cultural, experiential and pistic modification. From the analytic aspect onwards, the full meaning of aspects cannot be understood without reference to human living. So, for example, the lingual aspect is not about abstract notions of symbolic signification but about the human activities of recording, informing, and communicating.

Aspects are irreducibly distinct in respect of their meaning. Irreducibility, usually called sphere sovereignty, means that no aspect can be eliminated in favor of another. Aspectual irreducibility provides philosophical grounds for understanding diversity and helps us to avoid overlooking important factors. In everyday experience, every aspect is important: none can be dismissed as less meaningful, less interesting, or deserving of less of our attention.

Aspects are in harmony. Their laws are not in conflict. No aspect is absolute. No aspect can be the foundation for all the others. No aspect has its full meaning within itself. Each aspect refers to, or relates to each of the others.

Dooyeweerd's suite of aspects/spheres

Table no. 1

Sphere	Kernel meaning	Norms	Sciences
Quantitative	Amount	Sequence	Mathematics
Spatial	Continuous space	Simultaneity	Geometry, topology
Kinematic	Movement	Dynamism	Mechanics
Physical	Energy, mass, forces, material	Persistence	Physics, chemistry, materials and fluid sciences
Biotic/organic	Organism, life functions	Health: integrity of organism	Biology, ecology
Psychic/ sensitive	Sense, feeling, response	Sensitivity, responsiveness	Psychology
Analytical	Distinction, concepts, logic, pieces of data	Clarity, non-contradiction	Logic, analytical science
Formative	Structures, construction, processing, goals, technique, technology, history	Achievement	Design science, engineering
Lingual	Symbolic signification	Understandability	Linguistics, informatics
Social	Social relationships and institutions, roles	Respect	Social sciences
Economic	Management of scarce resources	Frugality	Economics, management science
Aesthetic	Harmony	Rich harmony	Aesthetics
Juridical	Due, rights, responsibilities	Justice: due, appropriateness	Legal science
Ethical	Self-giving love, generosity	Self-giving love	Ethics
Pistic	Vision, commitment, belief	Faithfulness	Theology

Source: (Basden, 2010), pp. 13-20

Different aspects, as different spheres of meaning, provide different rationalities. The aspects pertain, across all situations, all cultures, all times, whether we acknowledge or understand them, or not.

4.1.Aspects and norms

The earlier aspects (especially quantitative to physical) are determinative while the later aspects (especially from the analytic aspect onwards) allow more and more freedom. Freedom means that the future is opening (it is not determined). People have freedom to go against laws of non-determinative aspects, but they are never free from repercussions of their doing. Beneficial or positive repercussions come from functioning in line with the laws of aspects and detrimental or negative repercussions come from going against the laws of aspects.

Aspects are spheres of law that establish a variety of norms (see table **Eroare! Fără sursă de referință.**). Normativity distinguishes what is "right" or beneficial from what is "wrong" or detrimental. It yields a distinct type of good and evil for each aspect. In general, time-response of repercussion lengthens with the aspects, from almost immediate in the earliest aspects to centuries in the pistic aspect.

Human activities, in order to be sustainable, have to be in line with each aspect. Since each aspect is important, neglecting any of them threatens successfulness of human activities. That's why all aspects and inter-aspectual dependencies have to be considered.

4.2 Aspects and science

Since universality, which science seeks to know, is of the law side, the role of science is to study the laws of aspects. Different aspects are studied by different sciences (see table **Eroare! Fără sursă de referință.**). Integral scope of reality is reflected by philosophy. Philosophy concerns itself with the connections between aspects.

When human beings engage in a disciplinary activity, they focus on certain things that are meaningful to them. The things that are the most meaningful may be thought of as forming a sphere of meaning. Things that are less meaningful are in other spheres of meaning. That's why, though centered on a single sphere, the work and research of the discipline reach out to other spheres, which are progressively less meaningful ([3]).

In other words, each aspect defines the central interest of a scientific area, determining the types of entities, processes and laws about which each science concerns. Each discipline may, therefore, be seen as centered on one sphere of meaning, while also concerning itself with others. If a discipline has two aspects at its centre, one of them is usually primary.

Disciplines in the same aspect are sibling ones. Each of them has different links with other aspects.

4.4 Towards overcoming barriers to interdisciplinarity applying spheres of aspects

Since Dooyeweerd's philosophy could help in understanding human life, it could be used for building understanding of interdisciplinarity. In this section the barriers impeding interdisciplinarity are examined from the view of spheres of meaning and the need for improving understanding of human life is considered by applying Dooyeweerd's philosophy.

4.3 Spheres of meaning and barriers to interdisciplinarity

In order to understand the barriers impeding interdisciplinarity using Dooyeweerd's theory of modal aspects, we have to consider what it might tell us about them.

<u>The difference in values</u>

The difference in value exists. It could be explained by different perspectives of researchers and their life-and-world-views (LWVs). A LWV (Weltanschauungen) embodies deep assumptions, aspirations and quality criteria. It is often centered on one aspect and, as such, it is often reductionist one ([2]). Thus, researchers in an interdisciplinary team have different aspectual profiles. Since different aspects are spheres of different meaning with different norms and different rationalities, researchers in an interdisciplinary team do have different values.

Dooyeweerd has very explicitly stated that there is no truth in itself, i.e., there is no truth which is self-dependent and able to stand as truth without reference to anything else ([2]). In other words, there is no objective knowledge and no objective advice from an interdisciplinary research could be expected.

<u>Difference in theories and models</u>

Different aspects are different sphere of meanings and are studied by different methods ([2]). Since all aspects are important, no method form one aspect is superior over a method from some other aspect. The main sphere of meaning of a science and its links to other aspects (sciences) could help in understanding assumptions embedded in a method and possibilities of its linking to other method.

<u>Difference in epistemology</u>

Dooyeweerd has pointed out that each aspect requires distinct way of knowing. In other words, there is difference in epistemology in different aspects, i.e., in an interdisciplinary team.

Recognizing importance of inter-aspectual links in a problem which has to be solved in an interdisciplinary research could help in better understanding what styles of thinking and research approaches have to harmonize, as well as whether initial research question reflects well these links.

<u>Difference in the way in which society interacts and organizes academia (and relative absence of motivation)</u>

Understanding that all aspects are important could help in building understanding that no science from any aspect is superior over a science from some other aspect. Understanding main sphere of meaning for the problem under consideration, and how this sphere is connected to others, could help in forming good team and proper evaluation of its result. Causes for relative absence of motivation would be removed.

Intellectual labor division reinforces further division of research domains. Narrowing of domains makes more sibling disciplines and so more links in a web of domains of an interdisciplinary research. Since each sibling discipline is somehow distinct from others, the web is more complex, and there are more barriers to successful implementation of interdisciplinary research.

4.5 Dooyeweerd's philosophy and a need for changes in understanding human life

Dooyeweerd was a philosopher who was interested in everyday experience. His interest was in the whole breadth and depth of reality.

It is the philosopher who tried to understand diversity of human life and how human being functions in everyday life. He recognized the suite of aspects, which are irreducible and in harmony. An aspect is distinct sphere of meaning, distinct way of being, distinct rationality, distinct mode of being, distinct way of functioning, distinct sphere of law, distinct kind of normativity, distinct type of repercussion, and distinct way of knowing.

In other words, Dooyeweerd pointed out to a need for epistemological pluralism, for consideration constraints of natural resources and processes on human actions, and vice versa. He recognized that building knowledge is dependent on pistic, ethical, social, and other aspects.

In order to function well in an interdisciplinary research, all participants have to behave well in all aspects. They have to care about colleagues and their values, theories, methods, and assumptions, as well to try to harmonize research approaches of all participants.

Thus, Dooyeweerd's philosophy supports understanding a need for changes in understanding human life. Well functioning of all involved in an interdisciplinary work, in Dooyeweerd's sense, would allow better interdisciplinary work, and better organized academia and greater social benefit.Information systems – interdisciplinary field or science discipline?

The field of information systems is a relatively new research field. An information system is seen as an interdisciplinary combination of the social and technical spheres of organizational studies and computer science. There have been a lot of crises in the IS field, similar ones in other interdisciplinary fields.

Dooyeweerd's notion of spheres of meaning has been used for consideration a direction which might be taken if IS were to be a discipline. A proposal of the definition and dignity of the IS discipline was given ([3]).

5. INFORMATION SYSTEMS AS INTER - DISCIPLINARY FIELD

The field of information systems (IS) is a relatively new one. Its development is closely linked to the development of information technology (IT) which is very recent. The first computers were built in the 1940s and first business application supported by IT at 1950s.

The core concern of the IS field is taken to be orderly provision of data and information within an organization using IT, so that provided information are relevant to ever-changing activity of the organization and / or its members ([5]). The potential of IT has been so seductive that IS researchers and practitioners are often not interested in wider questions as what new IT really brings to their users and what the social implications of IT are.

The idea of nature of an IS has been changed. An IS was seen first as a technical one, then as a technical system with some social repercussions, and afterwards as a socio-technical system, i.e., as a social system with technical implementation ([8]). Thus, nowadays an information system is often seen as an interdisciplinary combination of the social and technical spheres of organizational studies and computer science ([3]). However, since information is a rich phenomenon and very important one for IS, it has been suggested that an IS could not be seen only as a technical system and as a social one, but also as a knowledge system ([10]).

As well as interdisciplinary studies in general, there are many barriers for success in the IS field. Very different assumptions have been made in the IS field, and they led to very different schools of thought in IS work ([5]). Those schools imply different values, different theories and methods, and different epistemologies. They effect IS departments position in university. IS departments are most often situated at technical faculties or at schools of business and management.

All IS developers approach to their task with a number of explicit and implicit assumptions about the nature of human organizations, the nature of their task, and expectations of them. Very important assumptions for IS development are ones associated with the way of systems developers' knowing (epistemological assumptions) and ones relating to their view of social and technical world (ontological assumptions). Different assumptions lead to different systems development approaches, different roles for IS developers and different system outcomes ([7]). About hundred theories, including the theory of modal aspects, are used in IS research ([14]). These theories are from different science disciplines.

Such state in the IS field leads to different, often unrealistic, expectations of IS users. For example, if a IS user sees IS developer as an expert, he will be expected to know how to develop IS that optimally supports organization. But, such IS developer could assume a rational organization, with discoverable objective cause-effect relationships. But, the nature of modern organization isn't such. On the other hand, if the IS developer has different assumptions, his values and assumptions will be in conflict with ones of their users. Even if IS users and developers have the same beliefs and assumptions, they, by a rule, have different education, and that's why they will use different theories and models, and probably they will be not able to understand well each others because of their different professional languages.

Besides that, if IS developers or researchers use social not-acceptable approaches, they could be evaluated as unprofessional ones ([13]). Because of extant way of scientists' evaluation, researchers are unwilling to do right, hard investigation, but prefer easier and better evaluated technical ones ([4]).

In other words, the barriers impeding interdisciplinarity has been recognized as barriers developing an IS. The barriers of different values, different theories and models, different epistemologies, as well as the way in which university is organized, have been recognized. That's why the fact that IS practice is very different from IS theory ([6]) is not very much surprising.

6. INFORMATION SYSTEMS AS A SCIENCE DISCIPLINE

In order to define the IS discipline, by finding its dignity and destiny and suggesting how the IS discipline can relate to other disciplines by way of responsible application, foundation and anticipation, Dooyeweerd's spheres of meaning has been used ([3]). In that research the focus was not on what is meaningful within IS discipline itself, but on what is meaningful to IS researchers and practitioners, to others and to the broader scheme of things. Respecting other disciplines allowed understanding what future for information systems is worthwhile and beneficial in the broader scheme of things. In that way, a proposal for the dignity, destiny and responsibility of the IS discipline was established.

It was determined that most meaningful sphere of meaning for the IS discipline is the lingual one. It is the sphere in which the human activities of crucial importance for an IS (i.e., recording, informing, and communicating by means of signification, by symbol-as-expressed-meaning) are expressed and guided by a norm of understandability.

Sibling disciplines to the IS discipline are languages, linguistics and semiotics, and media. The presupposition of advanced technology differentiates the IS discipline from languages. Treating IT users as full, multi-aspectual human beings rather than as mere language-users, signifiers or audience differentiates it from linguistics, semiotics and media.

IS links to other spheres of meaning was investigated too. It was considered how they affect, and are affected by, human informing, recording and communication.

The neighboring aspects are very important. The formative aspect, as the neighboring foundational aspect, offers important topics for the IS discipline: the structure and processing of information, the creative human activity of IS development, and technology, techniques and artifacts involved in an IS. The IS discipline is interested in social topics relating to recording, informing and communication from the informational angle, not the social per se.

The analytical aspect is exhibited in IS as individual pieces of data. Availability of data types determines what can be easily said or understood. The economic aspect is important because of widespread use of IS in business, as well as because of the topics of information resources management.

The IS discipline itself should not penetrate far into formative and analytic matters as such, and should rely on disciplines centered on those aspects to do so. Thus the IS discipline would relate to foundational disciplines with mutual respect. The IS discipline relates to anticipatory disciplines in a spirit of willing service.

Of course, all aspects have been considered in order to build a fuller picture. For example, the juridical aspect is relating to emancipatory IS, and the aesthetic aspect to IS architecture in that picture.

In other words, the IS discipline can be defined and suited by reference to the spheres of meaning that are most important to it. The dignity, destiny and responsibility of information systems as a discipline is that it opens up new potential of the lingual aspect in the service of other aspects in a way that no other discipline can.

7. CONCLUDING REMARKS: TOWARDS MORE SUCCESSFUL INTERDISCIPLINARY RESEARCH

It has been recognized a need for philosophy in a deliberation on interdisciplinarity. Needs for better understanding human life in order to improve interdisciplinarity ([1]) and for better understanding the barriers impeding interdisciplinarity ([11]) have also been recognized.

Dooyeweerd was a philosopher who was interested in everyday human experience. His interest was in the whole breadth and depth of reality. His theory of modal aspects shows how irreducibly distinct aspects account for diversity and coherence, being and doing, normativity, etc.

In this paper it was investigated how his philosophy recognizes barriers impeding interdisciplinarity and how his spheres of meaning could help in building understanding of human life needed for better interdisciplinary activities.

On the other side, investigations on applying spheres of meaning in the IS field, have shown that the spheres of meaning cast a different light on interdisciplinarity, that interdisciplinarity can be seen, not as bringing different processes or phenomena together, but as looking at different aspects of the same phenomenon ([2]). By looking at different aspects of the phenomenon of information systems, the frameworks for IS understanding and the proposal for the definition and dignity of the IS discipline have been given.

Since the spheres of meaning could be helpful in better understanding and improving interdisciplinarity and that they have been used for IS phenomenon, in which there are problems similar to ones of other interdisciplinary phenomena, a similar way could be helpful for other interdisciplinary phenomena. In other words, a central focus of interdisciplinary research phenomenon should be found, as well as its links to other aspects (and disciplines) by way of responsible application, foundation and anticipation.

It should be expected difficulties in disclosing the central sphere and links to others in some cases, but thinking about spheres of meaning of the phenomenon and their links should help in building more understanding of extant and possible barriers impeding that interdisciplinarity activity.

Defining and suiting the phenomenon by references to the spheres of meaning, like in the case of the IS phenomenon, might be an indication that it matures to be a discipline, with disclosed definition, dignity, destiny and responsibilities.

8. ACKNOWLEDGMENTS.

The presented work has been financially supported by the Ministry of Education and Science of Republic of Serbia under the project no. TR 36005.

9. REFERENCES

[1] Apostel, L., Vanlandschoot, J., (1994), ***Interdisciplinarity: The Construction of Worldviews and the Dissemination of Scientific Results***, Issues in Integrative Studies, Vol. 12, pp. 9-22

[2] Basden, A., (2008), ***Philosophical Frameworks for Understanding Information Systems***, IGI Publishing

[3] Basden, A., (2010), ***On using spheres of meaning to define and dignify the IS discipline***, International Journal of Information Management, Vol. 30, No. 1, pp. 13-20

[4] Carola, L., (2005), ***Development and Status of the Information Systems / Wirtschaftsinformatik Discipline - An Interpretive Evaluation of Interviews with Renowned Researchers: Part II - Results Information Systems Discipline***, ICB - Research Reports, No. 3, University Duisburg-Essen, Institute for Computer Science and Business Information Systems

[5] Checkland, P., Holwell, S., (1998), ***Information, Systems and Information Systems***, Wiley

[6] Ciborra, C. U., (1998), ***Crisis and foundations: an inquiry into the nature and limits of models and methods in the information systems discipline***, Journal of Strategic Information Systems, Vol. 7, No. 1, pp 5-16

[7] Hirschheim, R., Klein, H. K., (1989) , ***Four Paradigms of Information Systems Development***, Communication of the ACM, Vol. 32, No. 10, pp. 1199-1216

[8] Hirschheim, R., Klein, H.K., Lyytinen, K., (1995), ***Information Systems Development and Data Modeling: Conceptual and Philosophical Foundations***, Cambridge

[9] Jacobs, J.A., Frickel, S., (2009), ***Interdisciplinarity: A Critical Assessment***, Annual Review of Sociology, Vol. 35, pp. 43–65

[10] Lee, A. S., (2004), ***Thinking about social theory and philosophy in information systems***, in Mingers, J., Willcocks, L.P. (Eds.), Social theory and philosophy for information systems, Wiley, pp. 1–26

[11] Lélé, S., Norgaard, R.B., (2005), ***Practicing Interdisciplinarity***, BioScience, Vol. 55, No. 11, pp. 967-975

[12] Miller, T. R., Baird, T. D., Littlefield, C. M., Kofinas, G. , Chapin, III, F. , Redman, C. L., (2008), ***Epistemological pluralism: reorganizing interdisciplinary research***, Ecology and Society, Vol. 13, No. 2, 46 [online] URL: http://www.ecologyandsociety.org/vol13/iss2/art46/

[13] Mitic, M., (2010), ***IS development issues understanding – a condition and an object of social responsibility for IS in Serbia***, Management, Vol. 15, No. 54, pp. 69-79

[14] Schneberger, S. Wade, M.M., ***Theories Used in IS Research***, URL: http://istheory.byu.edu/wiki/Main_Page

[15] Seipel, M, (2007), ***Interdisciplinarity: An Introduction***, URL: http://www.uta.edu/faculty/repko/INTS_2301/SEIPEL.pdf

[16] Wesselink, A., (2008), ***Interdisciplinarity, Problem Focused Research and Normativity***, SRI Papers, No. 11, University of Leeds, Sustainability Research Institute

EXPLAINING IN MULTIDISCIPLINARY AND INTERDISCIPLINARY SCIENTIFIC INQUIRY SOME PHILOSOPHICAL EXPLORATIONS

Richard_ David-Rus
Romanian Academy, Iasi Branch, e-mail: rusdavid@gmail.com

Abstract. *The subject of scientific explanation is one of the major topics in philosophy of science. The explanations that were primarily studied were from the mature disciplines of science esp. from physics and biology. The context of interdisciplinary and multidisciplinary research was usually ignored by philosophical reflections. The recent modelistic orientation in philosophy of science makes plausible such an approach on scientific inquiry. In the present paper I will first look at the ways the classical accounts on explanation can accommodate the explanations that draw on more disciplines. I will than argue for the modelistic view as a proper way to foster such an approach. A general schema of explanation through modeling will be also introduced and discussed in regard to the advantages it brings and some related issues.*

Keywords: *scientific explanation, model and modelistic view, model-based explanation.*

1. INTRODUCTION

The subject of scientific explanation in philosophy of science is one of the major topics that marked the debates of the discipline in the last half of century. It dominated the working agenda of the field and concentrated the efforts of the philosophers and logicians starting with the 'classical' period of neopositivism. The philosopher C.G. Hempel initiated and raised the subject at its prime status and its deductive-nomological model concentrated the debate for more than three decades[1]. The fourth decade was marked by major new approaches and fresh insights that diversified investigation and revealed important novel aspects. Nonetheless minimal agreement was not reached and the enthusiasm for the topic gradually faded. Towards the end of the last century the debate was marked by a sort of fatigue and lack of fresh insights.

In recent years the new tendencies in the philosophical research brought with it the possibility of reframing the topic. The search for a general account of explanation valid for all varieties of explanation ranging over different scientific fields was dropped. The investigation concentrates nowadays on particular forms of explanation in a more local setting taking into account aspects that were previously ignored such as the pragmatic ones.

The above mentioned philosophical literature focuses mainly on scientific explanations articulated within the boundaries of the disciplines. Scientific areas that involve the appeal to more than one discipline, either multidisciplinary or interdisciplinary studies, were ignored as possible subjects of philosophical inquiry. This reflects the legacy of the 'received view' in the agenda of the topic and discloses the limitations of the classical view. The recent developments and tendencies in philosophy of science, especially the one that emphasizes the importance of models in the economy of scientific knowledge, open the possibility to approach in a plausible way the question of how explanation practices take place in multidisciplinary scientific research.

2. THE CLASSICAL APPROACHES ON EXPLANATION

Hempel's [10] deductive-nomological (DN) account construes explanation as a deductive inference in which the phenomenon to be explained (the explanadum) takes the position of the conclusion. The premises (the explanans) contain two sets of statements: laws of nature and statements describing the conditions under which the explanandum occurs. So to explain why the bottle filled with water broke over night we invoke the law of physics that state the increase in volume of frozen water and the fact that over night the temperature dropped below the freezing point. This is a more mundane interpretation of an explanation but we can say, as Batterman [1] also noticed, that Hempel had in mind primarily the explanation given in physical sciences by solving initial value problems for ordinary differential equations. As for example when we want to predict and explain this way the position of the Earth at some future point we will need Kepler's equations and the positions at some specific moment. Extending this type of explanation to other areas of science proved in the end to be problematic.

Now for Hempel the laws he has in mind are the fundamental laws of nature as they appear in physics. More particular physical laws as Kepler's laws of celestian motion, or Galileo's laws of moving bodies at the Earth surface are to be explained in the same way by deriving them from the fundamental laws of Newtonian mechanics. This might pose a first problem for the DN-model in regard to multidisciplinary explanations. In such studies the laws that are referred must belong to one or another of the disciplines involved; they are laws specific for the disciplines involved. So for example the explanation of a social phenomenon by appeal to a physical law would not count according to the DN model as a genuine explanation. In fact this situation can be described better as a modeling situation which for Hempel refers to an exploratory aspect of scientific inquiry and not an explanatory one. Hempel addresses the issue of a possibility of explanation through modeling and his illustration draws on mathematical modeling in social science. He denies such a possibility claiming any attempt to be only a sort of heuristic attempt to extend the theory in the realm of new phenomena.

One could say that Hempel was on the right track when addressing the issue of explanation through models but the

[1] For a detailed historical review of the debate in its first decade one can use Salmon's book *Four decades of scientific explanation.* Salmon was an important participant in the debate with major contribution in the analysis of statistical explanation and causal explanation.

frame of the official conception, the neopositivistic one, prohibited him to pursue any further such an investigation.

I will further look to the other existing approaches in order to see how they might address the issue of multidisciplinary explanation.

The approach that conceived explanation as unification was proposed by M. Friedman [6] in a seminal paper in 1974. His account didn't got too far due to a logical flaw but the idea got propagated. Its best articulation at the time, is to be found in Kitcher's account [11]. Friedman's account took derivation of laws or lawlike generalizations as the paradigm of a scientific explanation. Explanation and also understanding[2] is given by unifying various disparate phenomena expressed through more particular laws under a general law. So for example we unify specific laws as Kepler's laws, Galileo's laws, Hook law and many other under the fundamental laws of Newtonian mechanics by deductive derivation.

Kitcher's account is more elaborated. He develops a special notion of argument pattern as sort of a more adaptable schema for an inferential judgment. Such a pattern is meant to capture also the specificity of the reasoning in particular scientific fields. An argument pattern is a triple that contains a schematic sentence, a set of filling instructions (with instructions for each term in the schematic sentence) and a classification which describes the inferential characteristics of the schema. An argument pattern becomes explanatory if it belongs to the set of argument patterns that best unifies the corpus of knowledge ("the explanatory store"). The determinant process for the establishing the explanatory virtues of patterns takes place as the macrolevel were the sets of patterns are compared weighted and selected.

For the purpose of our investigation it seems that nothing prohibits such patterns to be instantiated in multidisciplinary inquiry, which will mean that they are part of the explanatory corpus. But Kitcher's intention is to see these patterns at the core of the grand scientific theories. He illustrates his ideas by reference to Newtonian mechanics or the explanatory patterns of Darwinian evolutionary theory. We might think of pattern as involved in interdisciplinary research as partially transferred from other theories and showing variation in some of their components. The most obvious case would be the variation in filling instruction as we would apply a schematic sentence to another sort of phenomena than those initially intended. Nevertheless we cannot tell if this pattern will belong to the explanatory store. The criteria for such a decision are too global and involve comparison of entire corpuses of knowledge. They are too uninformative to be used for a more local inquiry into how explanatory claims are built.

What about the other major accounts on explanation? The causalist approach proved to be one of the most successful one. It is not only in the classical period through mainly Salmon's work [14]. Recently the causalist approaches found different articulations in more general accounts but also in specialized domain-specific accounts, the last ones by analyzing causal explanatory forms in particular areas of science. The basic intuition of the approach is that an explanation tracks the causes of an event or fact, identifies causal influence and exposes causal mechanisms. The next step undertaken in such account is to spell out what this causal relation involves. Hempel considered he accounted for this sort of explanation through his DN model, being also faithful to the neopositivistic attitude towards causality and demystifying it from any metaphysical interpretation. The later causalists nevertheless offered consistent explications of causality.

In what regards the multidisciplinary and interdisciplinary investigation, the causalist approach seems to have some clear advantages in comparison to other previously discussed approaches. Tracking causes is independent of the theoretical structure appealed. So one can read out here a neutrality towards the sort of theoretical setting: it can fall into the boundaries of one discipline or of more disciplines without any problem. In fact one can find some important work on explanation through causal mechanisms with reference to such interdisciplinary domains as biophysics or biochemistry. I will mention here only the extended studies of William Bechtel on complex systems found at the intersection between biology, physics and chemistry and neurology.

The major problem with the causalist approach is quite well-known. Not all explanations are causal. This limitation was acknowledged by the adherents of the causal approach. The major example that is invoked and draws the limits for this approach (as Salmon recognizes) is the case of explanation in quantum mechanics. For an multidisciplinary point of view it will be unreasonable to restrict scientific explanation only to causal types of explanation and exclude as pseudo-explanation any involvement of other disciplines.

A last type of approach that I'm going to address is the one that draws on pragmatics.[3] It found one of its boldest expression in van Fraassen's account on explanation [16]. Putting it roughly a scientific explanation is an answer to a why-question given by making use of scientific information. Inspired by the developments of the erotetic logic[4] he tries to formalize the why-questions in a similar way. Why-questions are problematic in comparison to the other wh-question that Belnap & Steele [2] analyze, due to their open structure. For van Fraassen why-questions are contrastive: we ask "why P_k rather than P_j ?" where these Ps are propositions. A why-question Q is represented by a triple of the form $< P_k, X, R>$, where X is the contrast class i.e. a set of propositions: $\{P_1, P_2, .., P_n\}$, P_k is one of these propositions (called the topic) and R the relevance relation. A direct answer might be formalized as "P_k in contrast to (the rest of) X because A"; so more precise "B is a direct answer to $Q = < P_k, X, R>$ if there is some proposition A such that A bears relation R to $<P_k, X>$ and B is the proposition which is true exactly if (P_k; and for all i other than k, not P_i; and A) is true".

[2] Friedman argues for the fact that an approach on explanation should also give us insight into understanding.

[3] Similar to linguistic studies in logic and philosophy of science the syntactic, semantic and pragmatic aspects are separated , the last one concerning the use of the constructs.

[4] Belnap & Stele book *The logic of questions and answers* appeared at that time;

Letting apart the details of the account, the main emphasize falls on the relation of relevance R, which ultimately is determined contextually. Subsequently what might be taken as a satisfactory explanation is ultimately determined by contextual factors: such as the interest of the inquirer, the audience towards the explanation is directed etc. To see the main difference from a causal explanation I'll draw on a classical example from the literature. We can explain the length of a shadow casted by a flagpole by invoking the height of the flagpole and the laws of optics under both approaches; nevertheless only for the interrogative approach a derivation of the height of the flagpole from the length of the shadows will count as an explanation.

The interrogative approach is quite generous and can accommodate any sort of explanation. Its drawback seems to lie exactly in the fact that it is too generous – an almost "anything (that is an answer to a why-question) goes" sort of approach. That was what the critiques [12] pointed out from the beginnings; putting it a bit more precise the accusation goes to the fact that the relation of relevance remains unconstrained. The main point is indeed that there is a need for more than a totally open contextual determination. Nevertheless, taken the larger perspective of the debate on explanation van Fraassen's approach revealed the need to consider more seriously the contextual factors – a moral that was nowadays widely accepted.

An important consequence of the debate on explanation in its classical period is that a pluralistic view made its way as the most plausible one. Nowadays the existence of different irreducible explanatory formats is accepted by the philosophical community. Causal explanation is one of the most researched ones as it appears in various scientific areas. The investigation concentrates nowadays on particular forms of explanation as they are articulated in specific scientific domains.

There seems to be no problem with claiming the interrogative view on explanation as adequate to address the interdisciplinary explanation. The problem remains the same as identified by the general critique – the too unconstrained nature of the explanatory relevance relation. What we might need is some more consistent reference to scientific contexts and specific nature. I will further suggest a frame that can better harbor a philosophical inquiry into explanations in interdisciplinary settings.

3. TOWARDS A MODELISTIC VIEW

As we saw previously the philosophical accounts on explanation made reference to such scientific units as laws, theories or even larger types as field or entire corpuses of knowledge. Scientific models were neglected as potential bearers of explanations. This attitude can be attributed to the neopositivistic legacy and the fact that the topic of explanation was part of it. Interesting enough Hempel was the single major author that discussed the issue of explanation through scientific models; he rejected the possibility of any genuine explanation due the models limitations in range and purposes. Hempel, as one of the important figures of neopositivism, shared the "depreciative view" on models that the conception promoted (view that originates in P. Duhem's work [5]). According to this view models are only additional means that can be used for specific purposes in the production of scientific knowledge, usually for making more intelligible the abstract theoretical principles, but also for other purposes such as for example the search for new applications of a theory. Nevertheless they tell us little about the real structure of science. After we reach our goal they might be dropped; so they are dispensable means of little importance for scientific knowledge.

My suggestion, in the spirit of the recent modelistic orientation, points on the contrary to their importance in the scientific knowledge. In this sense, by taking seriously models as bearers of explanations we can better address and analyze the way explanations are build in multidisciplinary research. As I have argued in another place [4] taken into account the drawbacks of the classical approaches on explanation, a plausible direction of reframing it might be characterized as a local dynamic and non-theory driven sort of approach. Let me briefly clarify these characterizations. A local account is more sensitive to the contextual setting of the explanation in contrast to the global and generalizing modality of classical approaches. A dynamic view will take seriously the process of explanation rather than explanation as a final product of that process as it is conceived in the mentioned accounts. The non-theory driven sort of approach makes reference to the recent reactions against the sort of theory-dominated philosophical agenda as Cartwright called it, that characterized philosophy of science almost up to the end of the last century.

I have argued that such an approach could be developed in a modelistic frame - by investigating how the explanation builds up in the frame through modeling processes. There are some clear advantages that makes such a frame adequate for addressing explanation at a local level. Models are local units heavily influenced by the pragmatics of the investigation. They are constructs that represent specific aspects of reality or specific purposes. The scientists intentions play a major role in their building and functioning. The philosopher R. Giere [8] saw the relation between models and reality as a many-placed relationship and proposed the following definition for the model-based representation: "*S* uses *X* to represent *W* for purposes *P*."

The recent modelistic reorientation[5] in philosophy of science redefined also the working agenda that was previously marked by the theory-dominated sort of approach. By drawing on modeling practices in different scientific disciplines we can reset the general philosophical questions in more specific contexts. We make this way the philosophical investigation more relevant to scientific practice. In fact another recent trend stimulated and connected to the modelistic movement is the practice-oriented philosophy of science[6].

---

[5] The volume *Models as Mediators. Perspectives on Natural and Social Science* [13] edited by Morgan and Morrison gathers some important early works of this orientation.

[6] Developed around the Society of Philosophy of Science in Practice, it brings together philosophers that also have important contributions to the study of scientific modeling. A recent number from 2011 of the *European Journal of Philosophy of Science* gathers some of the works of this orientation.

Not least are the dynamical aspects of inquiry that a model oriented frame can accommodate. In comparison to other scientific units such as theories or laws, models could be much easier approached through a dynamic perspective involving their building, functioning or modification in the process of scientific inquiry. According to F. Weinert [17] models are ales constrained than theories and so it is easier to represent the alterations and modifications they are subjected to.[7] Besides we can much easirer draw on concrete examples from scientific practice in order to document these processes.

4. A GENERAL SCHEMA FOR MODEL-BASED EXPLANATION

I'll further present a general frame for approaching explanation through models proposed by Frigg and Hartmann (the LOOP account as they call it) [7]. The scheme is a sufficient general one but also enough articulated to provide a proper setting for an inquiry into explanation in multidisciplinary contexts.

The authors starting point are some scattered suggestions made by philosophers Hare and Cartwright. Hare [9] takes explanations to provide us pictures of the facts and these pictures are in fact presented through models. Cartwright [3] in her simulacrum account also states explicit that "to explain a phenomenon is to find a model that fits into the basic framework of the theory and that thus allows us to derive analogues for the messy and complicated phenomenological laws which are true of it." The LOOP account seeks to give a more detailed articulation of this idea by making explicit the dynamic in the model and between the model and the represented target system.

The explanandum is conceived to be a feature or propriety of the target system or an event or phenomenon within this system. They exclude other types of explanadum from their account. The explanans on the other side is the model itself. The problem is then: how does a model M explains an occurrence O exhibited by the target system T that is represented by M? Their account specifies four steps that make out the process of an explanation. The first two steps are called identification steps. In the first one we identify the occurrence in the target (OIT as they call it), i.e., the behavior of interest in the target system that has to be explained. Using their example, Boltzmann ideal gas model in which the gas is represented through an ensemble of a huge number of hard balls moving in a confined space under Newtonian classical laws, the OIT is the expansion of the gas in the entire volume of the container when a separating wall is removed. In the second identifying step, the occurrence in the model (OIM) is identified, i.e., the element in the model that corresponds to the occurrence in the target that we wish to explain. In our example it corresponds to the spreading of the balls in the entire volume.

The next two steps are called the explanatory steps. In the first one, called explanation_1 we have to reproduce the OIM in the model, meaning that the OIM has to follow from the basic assumptions of the model. 'Follow' is not made more explicit in any way but is not reducible to deduction as in the DN model. In the mentioned example of the ideal gas one has to show that the approach to equilibrium follows from the assumptions about the balls (the fact that are hard, that they collide elastically, etc). The fourth and last step, called explanation_2, involves the translation of knowledge obtained in the model (and about the model) to the target system. In our example, we know that the balls bounce around such that they reach the equilibrium distribution (what Boltzmann proved) and that the balls are idealizations of a certain kind of the molecules. This way what holds true in the model approximately carries over to the real system.

Having laid out the above sketchy account before proceeding further I want to emphasize two important points. The first one is linked to the precaution the authors take regarding the issue of explanation and truth. I completely agree with their position that an explanation is an explanation due to its "inner constitution' and not for how good, bad or fruitful it is. The classical accounts were hinging on the truth requirement as a fundamental one. That anchored the explanation problem in a dense philosophical context and made it dependent on such philosophical topics as the subject of laws or of causality and the solution to such issues. The second point I want to make is that the LOOP schema is an empty, quite unsubstantiated schema that has to be filled out. The authors recognize this and see the needed content to be delivered by the different types of representations corresponding to different explanatory strategies. The LOOP schema makes sense only in reference to a specific scientific context, articulate in a particular modeling episode from a specific area of science.

***

For the remaining part I will briefly point to some advantages the LOOP schema offers for the inquiry into multidisciplinary explanations. I will also suggest some immediate issues that might rise and should be addressed in applying the schema.

The LOOP schema offers clear advantages over the older approaches. One of the major is the fact that it conceives explanations a process (not as a final product as in older accounts) that is embedded into the more general dynamic of the scientific inquiry. It makes this way possible to account for the production and modifications of scientific knowledge when engaging into scientific inquiry.

The schema offers the possibility to tackle the way different elements are integrated in the flow of the inquiry processes. This has direct relevance for the study of integration of different disciplinary contributions in case of interdisciplinary or multidisciplinary inquiries. Distinct methodologies or techniques specific to the engaged disciplines are integrated in the run of the LOOP cycles. The different conceptual frames belonging to different disciplines might interact and combine or be used subsequently during the explanatory process.

The schema allows also to draw and articulate some distinctions that we need to take into consideration. It allows

[7] P. Duhem the French scientist and philosopher at the turn of the XXth century was warning against the dangers of these less rigourous constructs of science. His idea are at the origin of a depreciative view of models that got propagated by the neopositivists.

this way to reveal and identify the division of labor and the different contributions of the involved disciplines. So, to point only to some first sight rough distinctions the schema can separate the fact that the explanation in the model (the E1 step) might belong (or characterize this way) to one of the disciplines. To take an example we can see such a derivation in the frame of a model specific to statistical physics which can be applied to the study of some social phenomena. The OIM is the occurrence that properly falls into the physics realm described by statistical physics. Meanwhile the corresponding OIT is the real feature of a social system that is targeted by the model. The derivation inside the model is an uninterpreted inferential move one that draws on techniques for the physical science. Nonetheless this is not a totally physics bounded affair since many assumptions made in the run of the derivation might quite well make use of information specific to the social phenomena modeled. The second explanatory step is the moment when we judge the result in the model transferring the knowledge to the targeted system. At this step we are involved in interpretation of the results obtained in the model and have to take into account the knowledge and constraints from the other discipline (in our example the particular social science that claims the system as its domain of expertise). Of course the above observations are only the general sketchy distinctions that could be drawn at a first sight. The application on a specific scientific episode would reveal the finer configuration and provide consistency to the analysis.

Another important distinction is the one that regards the understanding that builds by going through such an explanatory loop. The division of labor between E1 and E2 makes possible to distinguish between understanding developed in the model which might draw primarily on one discipline (the one on which the model relies more heavily) from understanding claimed for the target system, that draws on the other discipline. A further step of inquiry will look at the different moments, how they develop during the modeling process and how they might contribute to the overall final understanding that the researcher can gain. In the end we would like to gain some novel insight on the targeted system and make the conceptual resources of the model as much as possible relevant to the nature of the modeled system. Conceptual novelties and improvements coming from the modeling discipline could suggest a further level of extending the multidisciplinary inquiry suggesting the possibility of a further more complex level of modeling.

The above discussion sketched only some general lines on which we might follow and develop the further inquiry. In order to flesh out the schema and better articulated the questions and the sort of answers they generate we need to make reference to a particular scientific context, a concrete piece of scientific inquiry. Working into such a specific context we can identify and characterize the modality explanations and understanding articulates in particular areas of research but also how these characterizations might be transferred to other domains.

Not at least one last important thing to be emphasized. The LOOP schema allows a Bayesian reading of the knowledge dynamics but also of the explanatory claims. By running through the loop we gain new information and update this way the degree of belief in our findings. Regarding explanation, as the authors also suggested, we can attach them values between 0 and 1 and represent this way the degree of confidence in an particular explanation. We do not have to drop an explanation entirely if we are not totally confident about it and we cannot validate it beyond doubt. We might hold on and work with explanations though we have some doubts about them (which we may try to clarify by using it) and continue to expand this way our corpus of knowledge. This makes possible to adopt a failibilistic perspective over the production of scientific knowledge.

5. FINAL REMARKS

The recent modelistic reorientation in philosophy of science could be seen as sort of sub-paradigm change in the philosophical research. As its working agenda expands it unveils unexpected opportunities to rethink some older philosophical topics. Such is the well-honored explanation topic. The changes also brought with them the possibility to address issues that were ignored by classical approaches. Such is study of multidisciplinary and interdisciplinary research and the explanatory claims that could be made in such contexts. We are engaging here in new areas of philosophical research that open the possibility to gain fresh insights into the study of science. The most important thing of the recent tendencies is the fact that they promote in a more determined and efficient way the close contact of philosopher with the science practician.

6. ACKNOWLEDGEMENT

This paper was made within *The Knowledge Based Society Project supported by the Sectorial Operational Program Human Resources Development* (SOP HRD), Financed by the European Social Fund and by the Romanian Government under the contract POSDRU /89/1.5/S/56815.

7. REFERENCES

[1] Batterman, R. (2000), *A 'Modern' (Victorian) Attitude towards Scientific Understanding*, The Monist, vol. 83(2), pp. 228 – 257.

[2] Belnap, N.D. & Steele, T.B. (1976), *The Logic of Questions and Answers.* New Haven: Yale University.

[3] Cartwright, N. (1983), *How the Laws of Physics lie*. Oxford: Oxford University Press.

[4] David-Rus, R., (2011), *Explanation through Scientific Models: Reframing the Explanation Topic*, Logos & Episteme vol. II, (2), pp. 177-189.

[5] Duhem, P. (1954), *The Aim and Structure of Physical Theory*. Princeton: Princeton University.

[6] Friedman, M. (1974), *Explanation and Scientific Understanding.* Journal of Philosophy, vol 71, pp. 5-19.

[7] Frigg, R., Hartmann, St., (2005), *Models, Explanation and Understanding.* Draft of the presentation at the 2005 Conference on Philosophical Perspectives on Scientific Understanding, Vrije Universiteit Amsterdam.

[8] Giere, Ronald N. (2004), *How models are used to represent reality*. Philosophy of Science, vol. 71 (5), pp. 742-752

[9] Harre, R. (1983), *An Introduction to the Logic of the Sciences.* London: Macmillan.

[10] Hempel, C. G. (1965), *Aspects of Scientific Explanation and other Essays in the Philosophy of Science.* New York: Free Press.

[11] Kitcher, P. (1989), *Explanatory Unification and the Causal Structure of the World* in P. Kitcher and W. Salmon (Eds.), *Scientific Explanation*. *Minnesota Studies in the Philosophy of Science, volume 13* (pp. 410-505). Minneapolis: University of Minnesota Press.

[12] Kitcher, P. & Salmon, W. (1987), *van Fraassen on explanation*. Journal of Philosophy, *vol. 84*, pp. 315-330.

[13] Morgan, M. and Morrison, M. (Eds.) (1999), *Models as Mediators. Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.

[14] Salmon, W., (1998), *Causality and Explanation,* New York: Oxford University Press.

[15] Salmon, W. (2006), *Four decades of Scientific Explanation,* Pittsburgh: University of Pittsburgh Press.

[16] van Fraassen, B. (1980), *The Scientific Image.* New York: Oxford University Press.

[17] Weinert, F. (1999), *Theories, Models and Constraints.* Studies in History and Philosophy of Science, vol. 30(2), pp. 303-333.

A MULTIDISCIPLINARY APPROACH, IN TERMS OF ECONOMETRICS, STATISTICS AND MECHANICS, TO PREFERENCES FOR DATA SAMPLES

Gheorghe_Săvoiu[1], Victor_Iorga-Simăn[2],
[1]University of Piteşti, e-mail: gsavoiu@yahoo.com, [2]University of Piteşti, e-mail: victoriorgasiman@yahoo.com

Abstract. This article tries to find econometric viable, timely and sufficiently grounded solutions, in the selection of samples or tests and experiments in relation to a standard experiment. The example selected is a real one, and was generated by mechanical experiments concerning pressure and its successive measurements, the identification of a prompt and accurate choice between two samples of experimental data in relation to a standard one, for abnormally spread, unimodal or antimodal, data distributions, requiring, in the end, the simple quality of statistical thinking against a multidisciplinary, both econometric and mechanical, analysis. A brief introduction concerning the delimitation of the problem presented, is followed by a section describing a number of data samples and their analysis, covering different situations, while the final part emphasizes the importance of a multidisciplinary approach within the framework of experimental research.
Keywords: experiment, sample, multidisciplinary thinking and analysis, experiment and sample selection

1.INTRODUCTION

One of the difficult practical issues in experimental research lies in identifying rapid solutions to select samples from several samples and tests, in relation with an experiment considered as a standard. The most difficult questions usually yield unexpected answers, but the analytical efforts fail to consider precisely the simple solutions, as the researcher is always looking for complex and fully justified alternatives, armed with an arsenal of tests and complex validations that have to justify certain choices. Obvious deadlock situations also occur, however, where empirical data sets, and experimental data appear not to be relevant, and selecting a sample or experiment is at least difficult if not impossible. The multidisciplinary approach and seeking solutions as simple as possible seem to be the researcher's best solutions when in a tight spot.

2. STATISTIC ANALYSIS AND SPECIFIC MOTIVATION OF SELECTING A SERIES OF DATA IN CONJUNCTION WITH A STANDARD EXPERIMENT

The results of an experimental investigation, which are described below, have brought about the dilemma of selecting between several samples, which, for methodological purposes, were distinctly named: a) SER01= Experiment data A; b) SER02= Pressure chamber A data variant 1; c)SER03 = Pressure chamber A data variant 2; d) SER04= Experiment data B; e) SER05= Pressure chamber B data variant 1; SER06 = Pressure chamber B data variant 2.

Descriptive statistics of the first set of series of the pressure in the cylinder

Table no. 1

	Experiment data A	Pressure chamber A data variant 1	Pressure chamber A data variant 2
Code	SER01	SER02	SER03
Mean	1.531551	1.392482	1.491696
Median	0.993200	0.700347	0.701360
Maximum	**9.410600**	**8.939087**	**9.813486**
Minimum	0.359600	0.389500	0.390894
Std. Dev.	1.969562	1.898032	2.093992
Skewness	**2.695708**	**2.695446**	**2.694673**
Kurtosis	**9.476500**	**9.379527**	**9.373798**
Jarque-Bera	21300.77	4185.605	4180.221
Probability	0.000000	0.000000	0.000000
Sum	11025.64	2005.174	2148.042
Sum Sq. Dev.	27922.30	5184.033	6309.730
Observations	7199	1440	1440

Software used: EViews

The EViews software package turns to best account the Jarque-Bera test, which denies the normality of the series generated by the experimental data at any test associated probability (often, 0.01 or 0.05). According to the $\chi 2$ distribution, the Jarque-Bera test critical value for a statistical significance threshold of 0.05 is 5.99, and for 0.01 it is 9.21.The Jarque-Bera statistics, calculated for the series of values of variable SER01 is 21,300.77, far greater than 5.99 or 9.21, and the null hypothesis is rejected with a confidence level of 95 or 99 cases out of 100 (or a probability of 0.95 or 0.99). The data series is not normally distributed in the experiment for the 7199 values. Analogously, the SER 02 and 03 series are abnormally distributed according to the values of the JB test. In conclusion, all three series are abnormally distributed, heterogeneous, highly asymmetric and excessively arched. There are no significant differences between the means and variances of the two series, according to the tests further applied (in keeping with the average, median or dispersion).

Table no. 2

Test for Equality of Means Between Series			
Sample: 1 1440			
Method	df	Value	Probability
t-test	2878	1.332149	0.1829
Anova F-statistic	(1, 2878)	1.774621	0.1829
Test for Equality of Medians Between Series			
Sample: 1 1440			
Method	df	Value	Probability
Wilcoxon/Mann-Whitney		2.561753	0.0104

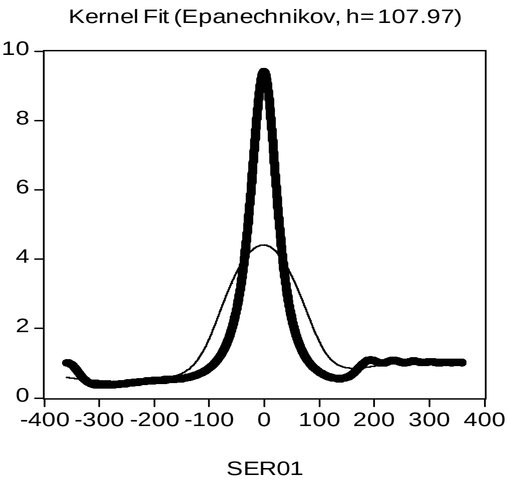
Wilcoxon/Mann-Whitney (tie-adj.		2.561753	0.0104
Med. Chi-square	1	5.512503	0.0189
Adj. Med. Chi-square	1	5.338891	0.0209
Kruskal-Wallis	1	6.562695	0.0104
Kruskal-Wallis (tie-adj.)	1	6.562696	0.0104
van der Waerden	1	8.784998	0.0030
Test for Equality of Variances Between Series			
Sample: 1 1440			
Method	df	Value	Probability
F-test	(1439, 1439)	1.217147	0.0002
Siegel-Tukey		1.575943	0.1150
Bartlett	1	13.86501	0.0002
Levene	(1, 2878)	4.806099	0.0284
Brown-Forsythe	(1, 2878)	1.309178	0.2526

Software used: EViews

The Kernel type graphs of the probability distributions are similar in the three cases, only the arching is different, as can be seen from the maximum values.

SER01 = Data of A experiment

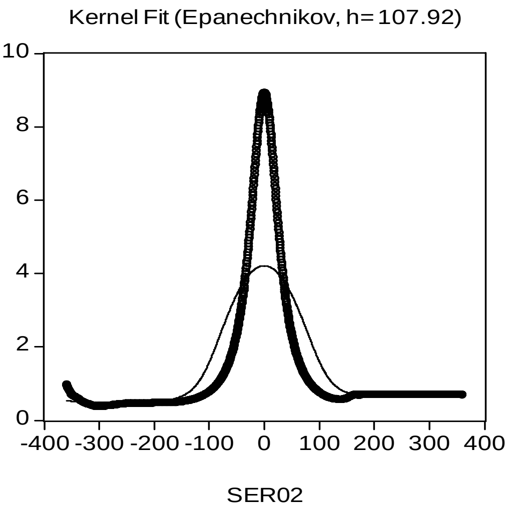
Graph no. 1.



Software used: EViews

SER02=Data of pressure chamber A variant 1

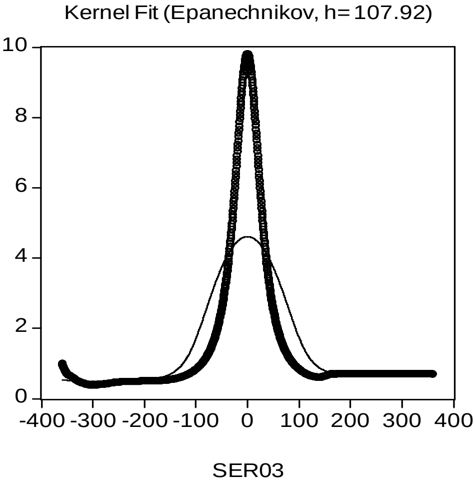
Graph no. 2.



Software used: EViews

SER03=Data of pressure chamber A variant 2

Graph no. 3.



Software used: EViews

In the statistical analysis conducted to identify the criteria for selecting one of the two series were valued the samples in their graphic peaks of distribution curves for the data series, and the range [-16.5, 16.5] was considered representative, where, simultaneously, all the three sets of data show a normal distribution, at the maximum permissible limit of the Jarque-Bera, for a significance threshold of 0.05 (according to the $\chi 2$ distribution, the criticial value of Jarque-Bera for a statistical significance threshold of 0.05 is 5.99). For the experiment only the values corresponding to the series compared were kept.

Taking the three samples of pressure inside the A chamber

Table no. 3.

The pressure's evolution inside the A chamber			
Interval	Experiment data A	Pressure chamber A data variant 1	Pressure chamber A data variant 2
-16.5	7.497	6.982994	7.659019
-16	7.5947	7.076958	7.763063
-15.5	7.6899	7.170318	7.866379
-15	7.7852	7.26295	7.968813
-14.5	7.8796	7.354702	8.070212
-14	7.9718	7.445426	8.170412
-13.5	8.0624	7.53497	8.269252
-13	8.1518	7.623183	8.366561
-12.5	8.2419	7.709903	8.462169
-12	8.3282	7.794968	8.555875
-11.5	8.4128	7.878213	8.64753
-11	8.493	7.959477	8.736955
-10.5	8.5731	8.038593	8.823974
-10	8.6495	8.115394	8.908405
-9.5	8.7228	8.189713	8.990075
-9	8.7971	8.26139	9.068806
-8.5	8.8622	8.330231	9.144426
-8	8.9254	8.396095	9.216762
-7.5	8.9881	8.458815	9.285651
-7	9.044	8.518244	9.350929
-6.5	9.0981	8.57424	9.412441
-6	9.1464	8.626665	9.470038

-5.5	9.1922	8.675389	9.52358
-5	9.2334	8.720294	9.57293
-4.5	9.2707	8.761266	9.617968
-4	9.3047	8.798201	9.65855
-3.5	9.3361	8.831003	9.694601
-3	9.3596	8.859591	9.726026
-2.5	9.3757	8.88389	9.752747
-2	9.3899	8.903822	9.774667
-1.5	9.4029	8.919339	9.791732
-1	9.4076	8.930413	9.803908
-0.5	9.4095	8.937002	9.81116
0	9.4034	8.939087	9.813486
0.5	9.3966	8.937007	9.811172
1	9.3833	8.930415	9.803898
1.5	9.3637	8.919324	9.791688
2	9.3423	8.90377	9.774595
2.5	9.3125	8.883817	9.752657
3	9.2798	8.859487	9.725933
3.5	9.2418	8.830879	9.694477
4	9.2024	8.798049	9.658404
4.5	9.1548	8.761082	9.617808
5	9.107	8.720089	9.572764
5.5	9.0546	8.675153	9.523412
6	8.9935	8.626407	9.469878
6.5	8.9328	8.573976	9.412297
7	8.8679	8.517985	9.35081
7.5	8.8016	8.45857	9.28557
8	8.7257	8.395873	9.216704
8.5	8.6507	8.330043	9.144399
9	8.5722	8.261229	9.06882
9.5	8.4876	8.189585	8.990137
10	8.4061	8.115276	8.908525
10.5	8.3173	8.038452	8.82418
11	8.2265	7.959277	8.737254
11.5	8.1371	7.877913	8.647927
12	8.0457	7.794521	8.556376
12.5	7.9514	7.709266	8.462778
13	7.8518	7.622305	8.367314
13.5	7.7526	7.533809	8.270148
14	7.6536	7.44392	8.171458
14.5	7.5505	7.352792	8.071408
15	7.4524	7.260572	7.970165
15.5	7.3489	7.167411	7.867889
16	7.2482	7.073446	7.764734
16.5	7.1457	6.97883	7.660853

The test of significance between the experimental sample and the data sample SER02 = Data for pressure in chamber a variant 1 identifies significant differences according to the statistics of the test (*t* - test is equal to 3.284419, and greater than 1.667 *t*-table, the series are significantly different as mean level, or mean - type parameter).

Table no. 4.

Test for Equality of Means Between Series			
Sample: 1 67			
Method	df	Value	Probability
t-test	132	3.284419	0.0013
Anova F-statistic	(1, 132)	10.78741	0.0013
Analysis of Variance			
Source of Variation	df	Sum of Sq.	Mean Sq.
Between	1	4.942938	4.942938
Within	132	60.48421	0.458214
Total	133	65.42715	0.491933

Software used: EViews

Analogously, the tests of significance between the experimental sample and sample SER03 = Data for pressure in chamber A variant 2 identifies significant differences according to statistics of test *t* (*t*-test is equal to 3.740852, and greater than the tabled *t* 1667, and the series are significantly different).

Table no. 5

Test for Equality of Means Between Series			
Sample: 1 67			
Method	df	Value	Probability
t-test	132	3.740852	0.0003
Anova F-statistic	(1, 132)	13.99397	0.0003
Analysis of Variance			
Source of Variation	df	Sum of Sq.	Mean Sq.
Between	1	5.849088	5.849088
Within	132	55.17230	0.417972
Total	133	61.02139	0.458807

Software used: EViews

Tested together, the series of data samples SER02 = Data for pressure in chamber a variant 1, and SER03 = Data for pressure in chamber A varaint 2, are even more clearely defined after the value of test *t* (*t* - test is 7.1101, and greater than 1,667 *t* - tabled).

Table no. 6

Test for Equality of Means Between Series			
Sample: 1 77			
Method	df	Value	Probability
t-test	132	7.110098	0.0000
Anova F-statistic	(1, 132)	50.55349	0.0000
Analysis of Variance			
Source of variation	df	Sum of Sq.	Mean Sq.
Between	1	21.54594	21.54594
Within	132	56.25852	0.426201
Total	133	77.80446	0.584996

Software used: EViews

All this information justifies sampling from the peak of the curves of the data distributions, and increase confidence in the analysis of their descriptive statistic. The criteria for the selection of one of the two series, by comparison with the experiment data series, remain those of homogeneity and normality of the series described by the data samples SER02 = Data for pressure in chamber A variant 1, and SER03 = Data for pressure in chamber A variant 2, and the analysis of the descriptive statistic, of the Jarque-Bera test and the coefficient of homogeneity or uniformity conduce to the following results:

The descriptive statistic of the three samples from the peak of the curves of unimodal distributions

Table no. 7

Sample: 1 67			
	Experiment data A	Pressure chamber A data variant 1	Pressure chamber A data variant 2
Mean	8.626258	8.208407	9.010381
Median	8.797100	8.330231	9.144426
Maximum	9.409500	8.939087	9.813486
Minimum	7.145700	6.978830	7.659019

Std. Dev.	0.670809	0.621256	0.682966
Skewness	-0.535415	-0.489012	-0.490193
Kurtosis	2.039230	1.923168	1.925154
Jarque-Bera	5.778070	5.907442	5.908423
Probability	0.055630	0.052145	0.052120
Sum	577.9593	549.9633	603.6955
Sum Sq. Dev.	29.69900	25.47330	30.78521
Observations	67	67	67

Software used: EViews

The homogeneity of the data SER02 = Data for pressure in chamber A variant 1 is found to be slightly further from the experiment, according to the signals derived from the absolute and relative amplitude, from the value of the standard deviation, and above all, the value of coefficient of homogeneity, and SER03 = Data for pressure in chamber A variant 2 is more similar, as far as the level of all indicators and trend are concerned, to the data series in the experiment.

Table no. 8.

	Experiment data A	Pressure chamber A data variant 1	Pressure chamber A data variant 2
Range	2.2638	0.598856	0.669060
Relative range	0.262431	0.072956	0.074254
Coefficient of homogeneity -%	7.776361	7.568533	7.579768

Software used: EViews

Analogously, the data in series SER02 = Data for pressure in chamber A variant 1 can be seen to have both a slightly smaller asymmetry (Skewness) and arching (kurtosis), while the the series SER03 = Data for pressure chamber A variant 2 and SER01 = Data experiment A have more extensive similar trends (tendential similarity in indicators, too, represents a a large enough set of arguments on account of which SER03 = Data pressure in chamber A variant 2 is preferred, as determined by analysing the samples taken from the peak of the curves of distributions).

The case of the analysis of the data series on pressure inside the B chamber, generating antimodal distributions reveals other quantitative aspects and results leading towards the same decisional deadlock in choosing the sample with a greater similarity in relation to the standard experiment.

Descriptive statistics of the first set of series of pressure inside B chamber

Table no. 9.

The pressure's evolution inside the B chamber			
	Experiment data B	Pressure chamber B data variant 1	Pressure chamber B data variant 2
Code	SER04	SER05	SER06
Mean	*0.613536*	*0.563182*	*0.577859*
Median	0.605130	0.530464	0.545044
Maximum	0.704660	0.699983	0.699983
Minimum	0.540860	0.487248	0.511179
Std. Dev.	*0.060959*	*0.076945*	*0.067545*
Skewness	*0.204802*	*0.562395*	*0.603365*
Kurtosis	*1.458467*	*1.722522*	*1.764616*
Jarque-Bera	*189.3235*	*43.21512*	*44.48703*
Probability	0.000000	0.000000	0.000000
Sum	1095.775	201.6191	206.8735
Sum Sq. Dev.	6.633078	2.113629	1.628763
Observations	1786	358	358

Software used: EViews

The Jarque-Bera statistic, calculated for the series of values of variable SER04, is 189.3235, therefore much higher than 5.99 or 9.21, and the null hypothesis is rejected, with a confidence level of 95 or 99 cases out of 100 (or a probability of 0.95 or 0.99). the series of experimental data is not normally distributed for the 1786 values. Analogously, series SER 05 and 06, too, are abnormally distributed in view of the values of the JB test. The three series are abnormally distributed but homogeneous, slightly asymmetric or at the limit of slight positive asymmetry, and of medium arching. No significant differences exist between the means and variances of the two series, in keeping with the tests further applied (in view of dispersion).

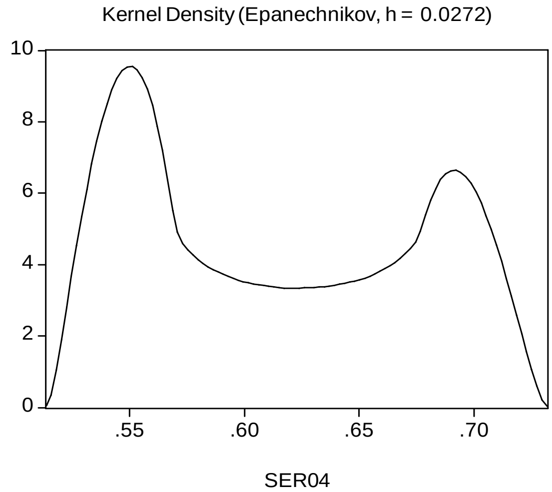
Table no. 10

Test for Equality of Variances Between Series			
Date: 01/12/12 Time: 11:54			
Sample: 1 358			
Method	df	Value	Probability
F-test	(357, 357)	1.297690	0.0140
Siegel-Tukey		9.187565	0.0000
Bartlett	1	6.034981	0.0140
Levene	(1, 714)	12.85911	0.0004
Brown-Forsythe	(1, 714)	5.567339	0.0186

The Kernel type of graphs for probability density distributions are similar in the three cases, only differing in the first portion of the arching, and the subsequent antimodal evolution is done on different minimum levels, as can be seen from the values, which are maximum, at first, and minimum, in the central portion of the graphs.

SER04 = data of experiment B

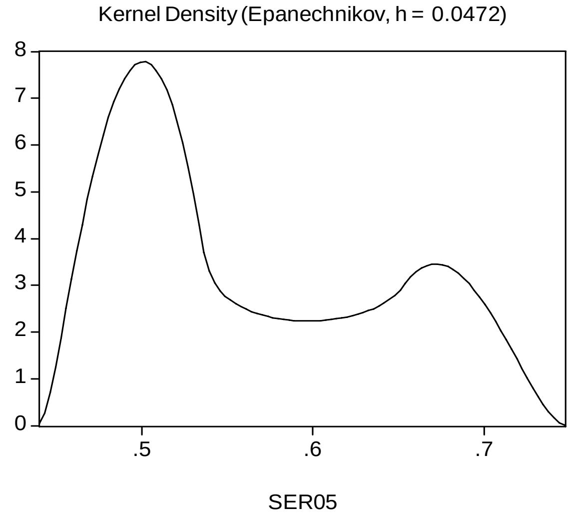
Graph no. 4.



Software used: EViews

SER05 = Data for pressure chamber B variant 1

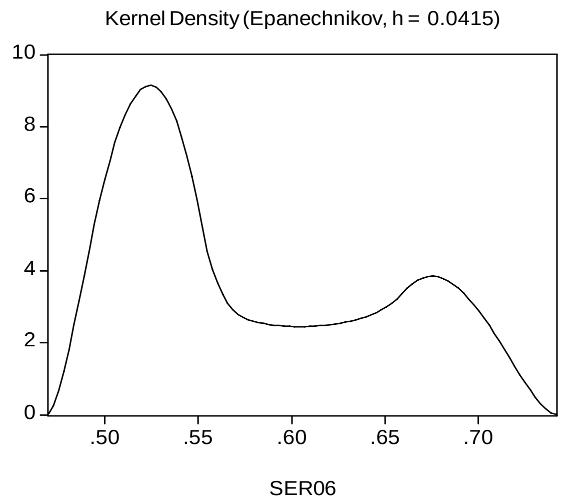
Graph no.5



Software used: EViews

SER06 = Data for pressure chamber B variant 2

Graph no. 6.



Software used: EViews

In the statistical analysis conducted to identify the criteria for selecting one of the two series, the samples in the central or antimodal area of the three curves were turned to used, centered on value of - 272, and the range [-285; -261] was considered representative, where, simultaneously, all the three data series show a normal distribution at the maximum permissible limit of the Jarque-Bera test, for a significance level of 0.05 (according to $\chi 2$ distribution, the Jarque-Bera critical test for statistical significance level of 0.05 is 5.99). (Note: for the experiment only the values corresponding to the series compared with a 0.5 to 0.5 leap were kept). Sampling normally distributed population samples observed the criterion of the intersection of the three graphs in the antimodal area, which is virtually the larger portion of the antimodal curve of distributions.

Taking the three samples of pressure in chamber B

Table no. 11

The pressure's evolution inside the B chamber			
Interval	Experiment data B	Pressure chamber B data variant 1	Pressure chamber B data variant 2
-285	0.67245	0.630362736	0.63620366
-284.5	0.67174	0.629157593	0.63511052
-284	0.67058	0.627946016	0.63401078
-283.5	0.67035	0.626728583	0.63290493
-283	0.66935	0.625507279	0.63179507
-282.5	0.66861	0.624283636	0.63068297
-282	0.66827	0.623057798	0.62956907
-281.5	0.66731	0.621829503	0.62845313
-281	0.66669	0.62059879	0.62733506
-280.5	0.66618	0.619364459	0.62621355
-280	0.6651	0.61812863	0.62509098
-279.5	0.66473	0.616892494	0.62396911
-279	0.66398	0.615652542	0.62284448
-278.5	0.66347	0.614404324	0.62171205
-278	0.66283	0.613146095	0.62056942
-277.5	0.66203	0.611880869	0.61941934
-277	0.66131	0.610614177	0.61826762
-276.5	0.66014	0.60935008	0.61711874
-276	0.65967	0.608089438	0.61597392
-275.5	0.65896	0.606824568	0.6148255
-275	0.65786	0.605542714	0.61366051
-274.5	0.65728	0.604239211	0.61247422
-274	0.65635	0.60292043	0.61127316
-273.5	0.65508	0.601596487	0.61006749
-273	0.65491	0.60027431	0.60886393
-272.5	0.65375	0.598957351	0.60766559
-272	0.6532	0.597646593	0.60647321
-271.5	0.65261	0.596340998	0.60528559
-271	0.65133	0.595038299	0.60410029
-270.5	0.6507	0.59373616	0.60291478
-270	0.64964	0.59243304	0.60172741
-269.5	0.64896	0.591126718	0.60053602
-269	0.64839	0.589814653	0.59933823
-268.5	0.64774	0.588495751	0.59813311
-268	0.64677	0.587171012	0.5969217
-267.5	0.64581	0.585844389	0.59570793
-267	0.64505	0.58452103	0.59449694
-266.5	0.64467	0.583204349	0.59329223
-266	0.64369	0.581894843	0.59209467
-265.5	0.64319	0.580591139	0.5909034
-265	0.64206	0.579291509	0.58971747
-264.5	0.64123	0.577994875	0.58853652
-264	0.64073	0.576701127	0.58735958
-263.5	0.63916	0.575410833	0.58618533
-263	0.63866	0.574124624	0.5850122
-262.5	0.6378	0.572852675	0.58385061
-262	0.63657	0.571606463	0.58271365
-261.5	0.63605	0.570385593	0.58160249
-261	0.63481	0.569179412	0.58050633

The tests of significance between the experimental sample and the sample of pressure data inside B, variant 1, identify significant differences according to t test statistics (t-test is equal to 17.73026, and greater than 1.676, and t-table series are significantly different as medium level, or medium type parameter).

Table no.12.

Test for Equality of Means Between Series			
Sample: 1 49			
Method	df	Value	Probability

t-test	96	17.73026	0.0000
Anova F-statistic	(1, 96)	314.3622	0.0000
Analysis of Variance			
Source of Variation	df	Sum of Sq.	Mean Sq.
Between	1	0.072474	0.072474
Within	96	0.022132	0.000231
Total	97	0.094606	0.000975

Software used: EViews

Analogously, the test of significance between the experimental sample and the sample of pressure data inside B, variant 2, identifies significant differences according to t test statistics (t-test is equal to 15.96466, and greater than t tabled 1676, the series are significantly different).

Table no. 13.

Test for Equality of Means Between Series			
Sample: 1 49			
Method	df	Value	Probability
t-test	96	15.96466	0.0000
Anova F-statistic	(1, 96)	254.8704	0.0000
Analysis of Variance			
Source of Variation	df	Sum of Sq.	Mean Sq.
Between	1	0.051397	0.051397
Within	96	0.019359	0.000202
Total	97	0.070757	0.000729

Software used: EViews

Tested together, pressure data series of samples for data inside B, variant 1 and variant 2, are also different according to *t* test value (*t* - test is 7.1101, and greater than *t* tabled 1676), but, in point of limit, they can be compared with the differences between each single data sample, and the data in the experimental sample.

Table no. 14.

Test for Equality of Means Between Series			
Sample: 1 49			
Method	df	Value	Probability
t-test	96	2.418029	0.0175
Anova F-statistic	(1, 96)	5.846865	0.0175
Analysis of Variance			
Source of variation	df	Sum of Sq.	Mean Sq.
Between	1	0.001806	0.001806
Within	96	0.029657	0.000309
Total	97	0.031463	0.000324

Software used: EViews

All this information warrants sampling in the antimodal area of the data distributions curves, and increase confidence in their descriptive statistical analysis. The criteria for selecting one of the two series by comparison with the experiment data series, are the same, i.e. homogeneity and normality of the series described by the samples of data for pressure inside chamber B, variant 1, and variant 2, and the analysis of the descriptive statistic, of the test Jarque-Bera and the coefficient of homogeneity or uniformity leads to the following results:

Descriptive statistics of the three samples from the peak of the distributions curve

Table no. 15.

Sample: 1 49			
	Experiment data B	Pressure chamber B data variant 1	Pressure chamber B data variant 2
Mean	0.654445	0.600056	0.608643
Median	0.654910	0.600274	0.608864
Maximum	0.672450	0.630363	0.636204
Minimum	0.634810	0.569179	0.580506
Std. Dev.	0.01030	0.018380	0.016735
Skewness	-0.084226	-0.027377	-0.030311
Kurtosis	1.793653	1.780113	1.780161
Jarque-Bera	3.029119	3.044376	3.045520
Probability	0.219905	0.218234	0.218109
Sum	2.06780	29.40276	29.82348
Sum Sq. Dev.	0.005917	0.016215	0.013442
bservations	49	49	49

Software used: EViews

It was found that the coefficient of homogeneity for the data series of pressure data inside chamber B variant 1 is slightly larger, analogously the signals derived from the absolute and relative amplitude, from the value of standard deviation, but above all, of the value of the homogeneity or uniformity coefficient, describing a relatively small distance of that series from the experiment, while the data series for pressure inside B variant 2 is more like, in point of the level of indicators and trend, the data series of experiment.

Table no. 16.

	Experiment data B	Pressure chamber B data variant 1	Pressure chamber B data variant 2
Range	0,0376401	0,061185	0,055698
Relative range	0,0575145	0,101965	0,0915118
Coefficient of homogeneity -%	1,696552	3,0630474	2,7495593

Software used: EViews

Analogously, it can be noticed that the data series SER05 on the pressure inside a B, variant 1, also have a slightly lower vaulting (kurtosis), while the data series on the pressure inside chamber B, variant 2, and the data in experiment B have similar but more extended trends (the trend and indicator similarity is a set of arguments consistent enough, for which the data series SER06 = data for pressure inside B variant 2 is preferred, as determined by the analyses of the samples taken from the common or value intersected area, i.e. the antimodal area of the distributions curves) .

3. A FINAL REMARK

Following the approaches of a multidisciplinary type, one can select appropriate samples from the data series of experimental character that simplify and motivate the reasons o scientific research itself. An approch that is simultaneously statistical through testing, econometric through modelling, and mechanical through selective and experimental impact may result in simple solutions with quick and efficient effect.

4.REFERENCES

[1] Greene, W. H., (2000), *Econometric Analysis*, 4th ed., Prentice Hall International,.

[2] Gujarati, D. N., Porter, D.C., (2009), *Basic econometrics*, Boston: McGraw Hill, 5th Edition,.

[3] Maddala, G.S., (2001), *Introduction to Econometrics*, Wiley, 3rd ed.

[4] Săvoiu G., (coord), Iorga-Simăn I., (coord), Constantin Andronache, David-Pearson, A-M., Ecker-Lala, W., Schjaer-Jacobson, H., Gligor, M., Costea, C., Chişleag, R., (2009), *Exploratory Domains of Econophysics News, EDEN I & II,*Ed. Universitară, Bucureşti.

[5] Săvoiu, G., (coord), Ghereş M., (coord), (2010), *Economia mediului. Tratat.* Ed. Universitară, Bucureşti.

[6]Săvoiu, G, (2010),*Gândirea statistică aplicată.Sisteme de indicatori rezultaţi din documente şi situaţii statistice financiar contabile* Ed.Universitară, Bucureşti.

[7] Săvoiu, G., Iorga–Simăn, I., Manea, C., Ştefănescu I., (2010), *Econometrics models versus physics models and their final connections with social economic reality, the educational system and scientific research* Economy. Transdisciplinrity. Cognition, Publisher: George Bacovia University, Bacău, vol XIII, issue no 2 /2010, pp. 32-44. http:// www. ugb.ro/

[8] Săvoiu G., Manea C., Iorga–Simăn, V.,Enescu F.M., Čudanov M.,Jaško O.,Jaško A, (2011), *A Corporate Model of Similitude for SMEs Reunion into a Corporation, Viewed from the Angle of Physical Thought, and Its Complex Economic and Social Impact*. Amfiteatru Economic Journal, Vol.XIII,No.29, pag. 10-28, http://www.amfiteatrueconomic.ro /.aspx?

[9] Săvoiu G., (2011), *Econometrie*, Ed. Universitară, Bucureşti.

IMPROVED EXPONENTIAL RATIO CUM DUAL TO RATIO TYPE ESTIMATOR OF POPULATION MEAN

Subhash_Kumar_Yadav

Department of Mathematics and Statistics, RML Avadh University, Faizabad, U.P., India

e-mail: drskystats@gmail.com

Abstract. *In the present paper, an efficient exponential ratio cum dual to ratio type estimator has been proposed to estimate the population mean of the variable under consideration by using simple random sampling scheme. The bias and mean squared error of the proposed estimator have been discussed up to the first order of approximation. A comparison has been made with existing similar estimators obtained by prominent researchers engaged in this area of interest. An improvement has been reflected in terms of mean squared error (MSE). The numerical demonstration has been presented to gain the better insight into efficiency criterion of the estimator under study.*

Key words: *Exponential estimator, dual to ratio estimator, bias, MSE, efficiency.*

1. INTRODUCTION

The use of auxiliary information increases precision of the estimate of the parameter of variable under study. Many authors have proposed improved estimator in terms of greater precision. Cochran (1940) used auxiliary information and proposed the usual ratio estimator of population mean. Robson (1957) and Murthy (1964) worked out independently on usual product estimator of population mean. Searls (1964) and Sisodia and Dwivedi (1981) used coefficient of variation of study and auxiliary variables respectively. Srivenkataramana (1980) first time proposed the dual to ratio estimator for estimating population mean. Singh and Tailor (2005), Tailor and Sharma (2009) worked on ratio-cum-product estimators. Sharma and Tailor (2010) proposed a ratio-cum-dual to ratio estimator for the estimation of finite population mean of the study variable y. Bahl and Tuteja (1991) were the first to suggest an exponential ratio type estimator for the estimation of population mean of the variable under study using auxiliary information.

Let $U = (U_1, U_2, ..., U_N)$ be the finite population of size N out of which a sample of size n is drawn with simple random sampling without replacement technique. Let y and x be the variable under study and the auxiliary variables respectively. Let $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ & $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$ be the population means of study and the auxiliary variables and $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ & $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ be the respective sample means.

Cochran (1940) proposed the classical ratio estimator for estimating population mean as

$$\bar{y}_R = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right) \tag{1.1}$$

With mean squared error

$$MSE(\bar{y}_R) = f\,\bar{Y}^2\left[C_y^2 + C_x^2(1-2K)\right] \tag{1.2}$$

where

$$C_y^2 = \frac{S_y^2}{\bar{Y}^2}, \quad C_x^2 = \frac{S_x^2}{\bar{X}^2}, \quad K = \frac{C_{xy}}{C_x^2} = \rho\frac{C_y}{C_x},$$

$$S_y^2 = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \bar{Y})^2, \; S_x^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X})^2,$$

$$S_{xy} = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \bar{Y})(X_i - \bar{X}), \quad f = \frac{1}{n} - \frac{1}{N}$$

and ρ is the correlation coefficient between x and y.

Srivenkataramana (1980), using the transformation $x_i^* = \frac{(N\bar{X} - nx_i)}{(N-n)}$ or

$x_i^* = (1+g)\bar{X} - gx_i$, $i = 1,2,......,N$,

which usually gives $\bar{x}^* = (1+g)\bar{X} - g\bar{x}$ where $g = \frac{n}{N-n}$, obtained dual to ratio estimator as

$$\bar{y}_R^{(d)} = \bar{y}\left(\frac{\bar{x}^*}{\bar{X}}\right) \tag{1.3}$$

with mean squared error

$$MSE(\bar{y}_R^{(d)}) = f\,\bar{Y}^2\left[C_y^2 + gC_x^2(g-2K)\right] \tag{1.4}$$

Sharma and Tailor (2010) suggested a ratio-cum-dual to ratio estimator as

$$\hat{\bar{Y}}_{bkl} = \bar{y}\left[\alpha\left(\frac{\bar{X}}{\bar{x}}\right) + (1-\alpha)\left(\frac{\bar{x}^*}{\bar{X}}\right)\right] \tag{1.5}$$

where α is a suitably chosen scalar. If $\alpha = 1$ and $\alpha = 0$, $\hat{\hat{Y}}_{bkl}$ reduces to estimators $\bar{y}_R$ and $\bar{y}_R^{(d)}$ respectively.

The mean squared error of the estimator $\hat{\hat{Y}}_{bkl}$ is

$$MSE(\hat{\hat{Y}}_{bkl}) = f\,\bar{Y}^2\left[C_y^2 + \alpha_1 C_x^2(\alpha_1 - 2K)\right] \tag{1.6}$$

Its minimum MSE is equal to usual linear regression estimator. Bahl and Tuteja (1991) proposed an exponential ratio type estimator as

$$t_1 = \bar{y}\exp(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}) \tag{1.7}$$

with mean squared error as

$$MSE(t_1) = f\,\bar{Y}^2[C_y^2 + C_x^2(\frac{1}{4} - K)] \tag{1.8}$$

The exponential dual to ratio type estimator is as follows

$$t_2 = \bar{y}\exp\left(\frac{\bar{x}^* - \bar{X}}{\bar{x}^* + \bar{X}}\right) \tag{1.9}$$

where:

$X_i^* = \dfrac{(N\bar{X} - nX_i)}{(N-n)}$ or

$X_i^* = (1+g)\bar{X} - gX_i$, $i = 1,2,......,N$, which usually gives $\bar{x}^* = (1+g)\bar{X} - g\bar{x}$ where $g = \dfrac{n}{N-n}$.

The mean squared error of the estimator t_2 is

$$MSE(t_2) = f\,\bar{Y}^2[C_y^2 + gC_x^2(\frac{g}{4} + K)] \tag{1.10}$$

2. PROPOSED ESTIMATOR

Motivated by Sharma and Tailor (2010), the following exponential ratio-cum-dual to ratio estimator has been proposed to define as

$$t = \bar{y}\left[\alpha\exp(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}) + (1-\alpha)\exp\left(\frac{\bar{x}^* - \bar{X}}{\bar{x}^* + \bar{X}}\right)\right] \tag{2.1}$$

where α is a real constant to be determined such that the MSE of t is minimum. For α = 1, t reduces to the estimator t_1 and for α = 0, it reduces to the estimator t_2 .To obtain the bias and mean squared error (MSE) of the estimators, let

$\bar{y} = \bar{Y}(1+e_0)$ and $\bar{x} = \bar{X}(1+e_1)$ such that $E(e_i) = 0, i = 0,1$ and $E(e_1^2) = f\,C_x^2$ and $E(e_0 e_1) = f\,C_{xy} = f\,\rho C_y C_x$

Expressing (2.1) in terms of e's, we have

$$t = \bar{Y}(1+e_0)\left[\alpha\exp\left(\frac{-e_1}{2}\right) + (1-\alpha)\exp\left(\frac{ge_1}{2}\right)\right] \tag{2.2}$$

Expanding the right hand side of (2.2) and retaining terms up to second powers of e's, and then subtracting $\bar{Y}$ from both sides, we have

$$t - \bar{Y} = \bar{Y}\left[1 + e_0 + \alpha_1\frac{e_1}{2} + \alpha_2\frac{e_1^2}{8} + \alpha_1\frac{e_0 e_1}{2}\right] - \bar{Y} \tag{2.3}$$

where $\alpha_1 = [g - (1+g)\alpha]$ and $\alpha_2 = [g^2 + (1-g^2)\alpha]$.

Taking expectations on both sides of (2.3), we get the bias of the estimator t up to the first order of approximation, as

$$B(t) = f\bar{Y}\left[\alpha_2\frac{C_x^2}{8} + \frac{\alpha_1}{2}\rho C_y C_x\right] \tag{2.4}$$

From equation (2.3), we have

$$(t - \bar{Y}) \cong \bar{Y}(e_0 + \alpha_1\frac{e_1}{2}) \tag{2.5}$$

Squaring both sides of equation (2.5) gives

$$(t - \bar{Y})^2 = \bar{Y}^2\left[e_0^2 + \alpha_1^2\frac{e_1^2}{4} + \alpha_1 e_0 e_1\right] \tag{2.6}$$

and now taking expectation, we get the MSE of the estimator t, to the first order of approximation as

$$MSE(t) = f\,\bar{Y}^2[C_y^2 + \alpha_1^2\frac{C_x^2}{4} + \alpha_1\rho C_y C_x] \tag{2.7}$$

which is minimum for optimum value of α as

$$\alpha = \frac{2K+g}{1+g} = \alpha_{opt}\ (say)$$

and the minimum MSE of t is

$$MSE_{\min}(t) = f\,\bar{Y}^2 C_y^2(1-\rho^2) = MSE(t)_{opt} \tag{2.8}$$

which is same as that of traditional linear regression estimator.

3. EFFICIENCY COMPARISON

We know that the variance of the sample mean $\bar{y}$ is

$$V(\bar{y}) = f\,\bar{Y}^2 C_y^2 \tag{3.1}$$

Now we have from (3.1) and (2.8), that

$$V(\bar{y}) - MSE(t)_{opt} = \rho^2 \geq 0 \tag{3.2}$$

Showing that proposed estimator t is better than the per unit estimator of population mean.

From (1.2) and (2.8), we have

$$MSE(\bar{y}_R) - MSE(t)_{opt} = (C_x - \rho C_y)^2 \geq 0 \qquad (3.3)$$

Which shows that proposed estimator is better than the traditional ratio estimator of Cochran (1940).

From (1.4) and (2.8), we have

$$MSE(\bar{y}_R^{(d)}) - MSE(t)_{opt} = (gC_x - \rho C_y)^2 \geq 0 \qquad (3.4)$$

Thus t is better than the estimator $\bar{y}_R^{(d)}$ due to Srivenkataramana (1980).

From (1.6) and (2.8), we have that both the estimators are equally efficient.

From (1.8) and (2.8), we have

$$MSE(t_1) - MSE(t)_{opt} = (\frac{C_x}{2} - \rho C_y)^2 \geq 0 \qquad (3.5)$$

Showing that proposed estimator t is better than the Bahl and Tuteja (1991) estimator t_1.

From (1.10) and (2.8), we have

$$MSE(t_2) - MSE(t)_{opt} = (\frac{C_x}{2} + \rho C_y)^2 \geq 0 \qquad (3.6)$$

Which shows that proposed estimator t is better than the estimator t_2.

1. THE EMPIRICAL STUDY

We have used the data in Koyuncu and Kadilar (2009) to compare the efficiencies between the previous and the proposed estimator for the population mean under simple random sampling.

Data Statistics

Table no. 1

N = 923	*n = 180*	$\bar{Y}$ =436.4345	$\bar{X}$ =*11440.4984*
Cy=171833	*Cx=1.864528*	ρ = 0.9543	*g = 0.24226*

The percentage relative efficiencies of previously developed and proposed estimators with respect to usual unbiased estimator $\bar{y}$ of population mean $\bar{Y}$ have been computed and presented in table 2, below.

Percentage relative efficiencies of different estimator's w.r.t. $\bar{y}$

Table no.2

Estimator	Values of
$\bar{y}$	100.000
$\bar{y}_R$	939.649
$\bar{y}_R^{(d)}$	176.247
t_1	386.307
t_2	78.856
$(t)_{opt} = (\hat{\bar{Y}}_{bkl})_{opt}$	1123.596

5. CONCLUSION

In the light of above numerical demonstration, we conclude that the proposed estimator t performs better than the usual estimator $\bar{y}$, Cochran (1940) usual ratio estimator $\bar{y}_R$, Srivenkataramana (1980) dual to ratio estimator $\bar{y}_R^{(d)}$, Bahl and Tuteja (1991) exponential ratio type estimator and exponential dual to ratio type estimator. So the proposed estimator should be preferred over above estimators for the estimation of population mean of the study variable using auxiliary information.

6. REFERENCES

[1] Bahl, S. and Tuteja, R.K. (1991), *Ratio and product type exponential estimator.* Information and optimization sciences, Vol. XII, I, pp. 159-163.

[2] Cochran, W.G. (1940), *The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce.* Jour. Agri. Sci., vol. 59, pp. 1225-1226.

[3] Kadilar, C. and Cingi, H. (2006), *Ratio estimators for the population variance in simple and stratified random sampling.* Applied mathematics and computation, vol 173, pp. 1047-1059.

[4] Murthy, M.N. (1964), *Product Method of estimation*. Sankhya, A, vol. 26, pp. 294-307.

[5] Robson, D.S. (1957), *Application of multivariate polykays to the theory of unbiased ratio-type estimation*. Jour. Amer. Stat. Assoc., vol. 59, pp. 1225-1226.

[6] Sharma, B. and Tailor, R. (2010), *A new ratio-cum-dual to ratio estimator of finite population mean in simple random sampling.* Global Journal of Science Frontier Research, 10, 1, 27-31.

[7] Tailor, R. and Sharma, B. (2009), *Modified ratio-cum-product estimator of finite population mean using known coefficient of variation and known coefficient of kurtosis.* Statistics in Transition-new series, vol. 10(1), pp. 15-24.

[8] Searls, D.T. (1964), *The utilization of known coefficient of variation in the estimation procedure*. Jour. Amer. Stat. Assoc., vol. 59, pp. 1125-1126.

[9] Singh, H.P. and Tailor, R. (2003). Use of known correlation coefficient in estimating the finite population mean Statistics in Transition, 6, 4, 555-560.

[10] Singh, H.P. and Tailor, R. (2005), *Estimation of finite population mean using known correlation coefficient between auxiliary characters*. Statistica, Anno vol LXV, (4), pp. 407-418.

[11] Sisodiya, B.V.S. and Dwivedi, V.K. (1981), *A modified ratio estimator using coefficient of variation of auxiliary variable.* Jour. Ind. Soc. Agri. Stat., vol. 33, pp. 13-18.

[12] Srivenkataramana, T. (1980), *A dual to ratio estimator in sample surveys*. Biometrika, vol. 67 (1), pp. 199 - 204.