The "Black-Box" of Institutional Scores: Analyzing the Distribution of the Values of the H and G Indexes in Medicine Schools in Romania

VIOREL PROTEASA[1]
Carol Davila" University of Medicine and Pharmacy, Romania

MIHAI PĂUNESCU
The National University of Political Studies and Administration, Romania

ADRIAN MIROIU
The National University of Political Studies and Administration, Romania

Abstract

Measuring the university research performance has been an important focus of the higher education policies in past decade in Romania. In the present study we considered alternative methodologies for evaluating quality of research in the faculties of medicine. We set to compare the perspectives of past official evaluations with alternatives based on h and g indexes of the academics within these faculties and subsequent successive indexes and averages. We analyzed the distribution of the values of the individual h and g indexes and we rejected the universality claim hypothesis, according to which all university h- and g-index distributions follow a single functional form, proportional with the size of the universities. However, using the Characteristic Scores and Scales approach, we show that the shape of distributions is quite similar across universities revealing the skewness of scientific productivity. Given the high skewness of all distributions, we conclude that all three collective aggregation rules considered, averages, h- and g-successive indexes fail to provide an accurate measure of the differences between the individual academics within the six medical schools, and fail to provide scientific achievement incentives for the wide majority of the academic staff within the analysed faculties.

Keywords: Medical Schools' Ranking; Higher Education Funding; H Index; G Index.

Introduction

The analysis of research productivity and impact in the wider context of higher education assessment in Romania has captured the attention of a series of scholars (Miroiu, Păunescu & Vîiu, 2015; Păunescu & Hâncean, 2013; Vîiu, Vlăsceanu & Miroiu, 2012). This article represents a contribution to the topic of using scientometric analysis for university research assessment and builds on an extensive data collection exercise performed by the authors in 2014-2015[2].

[1] Postal Assdress: National Institute for Infectious Diseases "Prof. Dr. Matei Balş", European Academy of HIV/SIDA and Infectious Diseases, 3rd Floor, Str. Dr. Calistrat Grozovici no. 1, Sector 2, Bucharest. E-mail Address: viorel.proteasa@gmail.com

[2] We are fully aware of the criticism against placing too much emphasis on research, and especially on using scientometrics to evaluate the quality of research; at the same time, the world-wide influence of scientometrics cannot be ignored. We do not aim with our research to take a clear stand in this debate, but

The stake of this article is to attempt a perspective into the „black–box" of institutional scores, as we consider that too often the attention is captured by the hierarchy the scores yield, without a deeper understanding of the differences that determine the values of the indicators. We focused on the six universities of medicine and pharmacy in Romania. In the first part, we discuss universities' evaluation results within the context of past public (funding) policies. The second part is more technical: we analyze a simplified version of the scientific publication component of the officially proposed methodology and two alternative evaluation methodologies, based on the *h* (Hirsch, 2005) and *g* indexes (Egghe, 2006). For this, we operate with two levels of aggregation: (1) the individual level, where we aggregate data regarding academics' publications and citations, and (2) the collective level, where we will aggregate data on individual academics' performances at the level of faculties. For the first level of aggregation we use citations' sum, the *h* index, and the *g* index, for the second level of aggregation we discuss averages, the *h* successive index (Schubert, 2007), and the *g* successive index (Tol, 2008). We consider this discussion is particularly relevant for the current policy debates: the official methodology for the evaluation of university performance for the purpose of state funding has a similar architecture, and employs the average of the individual *h* index raised at 3/2 power as an aggregation rule (MECS, 2015). In the concluding section we put our results in the current policy context in Romania.

We consider that our research is particularly relevant in the context in which the declared purpose of the initiatives relying on scientometric analysis, such as the 2011 classification of Romanian universities and ranking of study programs or the international rankings, is to make higher education diversity easier to comprehend, while the complex aggregation methodologies often turn these „transparency" initiatives into „black-boxes" for both scholars and the public.

Higher education funding policies: the research component

A research component was included in the different forms of the methodologies for higher education funding since 2007. A detailed account is presented in Vîiu & Miroiu (2013). We will refer in this section to (1) the „quality indicators", especially the research composite indicator IC6, used between 2007 and 2011, (2) the ranking of the study programs used since 2012, and (3) the quality indicators postponed for implementation until 2016.

The "quality indicators": 2003-2011

Both the structure of the quality indicators and their weight in the allocation formula were subject to change since their introduction in 2003. Research outputs, especially publication were emphasized starting with the 2007 methodology, when an indicator for research was included amongst the „quality indicators". Vîiu & Miroiu (2013) ranked universities according to the funding they achieved based on the composite research indicator (IC6) as a proportion of the funding they could have achieved in the absence of the indicator - based solely on enrolments. The maximum score is 200%. We present their results in

Table **1**, noting that the ranking is similar to that produced by Zulean, Ioniță & Vîiu (2014), who used the same methodology while calculating the allocations for the entire period in which the research indicator was used (2007-2011). The maximum score a university can achieve is

rather to inform it through empirical evidence. We consider the main aim of scientometric analysis is finding a proxy for quality, which is easy to measure, thus robust, reliable, and easy to communicate. Attributing significant weight to research productivity is a common feature of widely known university rankings, such as The Academic Ranking of World Universities, widely referred in the academic community as "The Shanghai Ranking" (ARWU, 2015), "Times Higher Education" World University Rankings (THE, 2015), or the less known to the public, yet widely referred by specialists, CWTS Leiden Ranking (CWTS, 2015). The arguments favouring this approach rely on the centrality of the modern universities in the production of knowledge in contemporary societies, but also on feasibility issues: research is more easily quantifiable than the added value of learning. Some of the difficulties related to the measurement of the added value of learning are emphasized both in scholars' and policy makers' community, e.g. Steedle (2010) or (OECD, 2013).

200%, thus the column corresponding to the 2011 funding shows that only the "Victor Babeș" University of Medicine and Pharmacy in Timișoara increased its funding due to the indicator for research. Over the period in which the composite indicator for research was used, none of the six health science universities managed to increase its funding due to the research indicator.

Table 1: *Health sciences universities' scores according to the methodology proposed by Vîiu & Miroiu (2013)*

	University	**Allocated funding 2011** (Vîiu & Miroiu, 2013)	**Allocated funding 2007-2011** (Zulean, Ioniță & Vîiu 2014)
1	UMF "Victor Babeș" Timișoara	108.88%	90.53%
2	UMF "Iuliu Hațieganu" Cluj Napoca	82.88%	74.63%
3	UMF "Gr. T. Popa" Iași	72.98%	68.76%
4	UMF Tg. Mureș	50.93%	65.42%
5	UMF Craiova	46.65%	57.42%
6	UMF "Carol Davila" București	29.91%	37.54%

Zulean, Ioniță & Vîiu (2014) analysed the results and found strong correlation between enrolments and the performance in achieving a large share of the potential funding for research. Their findings provide empirical evidence for the idea that the distribution of funds according to quality indicators is strongly biased towards the size of the universities, as advanced by Țeca (2011) and Miroiu & Vlasceanu (2012), which raises doubts regarding the strength of the incentive, associated with the research indicator, especially for the large universities. Thus, in the case of health sciences Romanian universities, "Carol Davila" University in Bucharest has managed to obtain the highest funding in absolute values, even though they achieved the least of their potential from the distribution according to the research indicator. Zulean, Ioniță & Vîiu (2014) place the above mentioned university among the net benefiters of the distribution of public money in the period they analyzed, which are all universities with large enrolments.

Ranking study programs

The ranking of study programs was instituted by the 2011 education law, and replaced the previous methodology based on quality indicators. The ranking methodology included more than 60 indicators, clustered in four broader dimension, with different weights: (1) research, (2) education, (3) community engagement and (4) institutional capacity (MECTS, 2011). For the universities of medicine and pharmacy, research weighted no less than 60% of the total score. For study programs in medicine, publications accounted to 75% of the research score, which results in a weight of 45% of the total score (Vîiu & Miroiu, 2013).

Similarly to the case of the composite research indicator within the frame of the previous funding policy – the "quality indicators", the quantification of research within the study programs' ranking took into account both quantity and impact, using a blend of absolute and relative measures, called in the official methodology "intensive" and "extensive" indicators. We reproduce in Table 2 the ranking of study programs in medicine, noting that the programs organized by the health sciences universities are all included in the upper groups, while study programs organized in faculties of medicine within comprehensive universities were all included in the lower ranks, and the largest faculty is placed on the leading position.

Table 2: *2012 ranking of study programs in medicine*

Rank	The university organising the medicine study program	Category	Score
	„Carol Davila" University of Medicine and Pharmacy, Bucureşti	A	≥ 50% of the maximum score
	„Gr. T. Popa" University of Medicine and Pharmacy, Iaşi	A	
	„Iuliu Haţieganu" University of Medicine and Pharmacy, Cluj	A	
	University of Medicine and Pharmacy, Craiova	B	< 50% and ≥ 30% of the maximum score
	„Victor Babeş" University of Medicine and Pharmacy, Timişoara	B	
	University of Medicine and Pharmacy, Târgu Mureş	B	
	„Transilvania" University, Braşov	C	< 30% and ≥ 15% of the maximum score
	„Lucian Blaga" University, Sibiu	C	
	„Vasile Goldiş" West University, Arad	C	
	Oradea University	C	
	„Dunărea de Jos" University, Galaţi	D	< 15% and ≥ 5% of the maximum score
	„Titu Maiorescu" University, Bucureşti	D	
	Piteşti University	D	

Quality indicators reloaded: 2015

In 2015 a new funding methodology was proposed, which reverted to the concept of quality indicators for the distribution of the supplementary funding (MECS, 2015). The implementation of the new funding scheme was postponed for 2016, and, as of this writing, this research constitutes the single preview of a component of the 2015 funding methodology, for health sciences. The new funding methodology innovates by proposing the distribution of the supplementary funding according to new quality indicators. The quality indicators are unique for all fields of studies, but the ranking for the purpose of funding is performed per field of study. We will not provide extensive factual information regarding the methodology (MECS, 2015). We will note that the quality indicators are grouped in the following dimensions, each of them with its own weight in the overall score: teaching / learning – 30%, scientific research / artistic creation - 40%, international orientation – 10%, regional orientation and social equity – 20%. The scientific research / artistic creation comprises of: (1) the quality of human resources – 12%, calculated for all teaching staff according to the criteria imposed nationally, for various subject fields, for professors and associated professors; (2) the impact of scientific research / artistic creation – 10%, calculated on the basis of the individual *h* index (Hirsch, 2005) for all academic staff; (3) outstanding performance in scientific research / artistic creation – 12%, calculated on the basis of publications in high impact journals and patents; and (4) funding for scientific research / artistic creation – 6%, calculated as the ratio of funding for research per academic staff. The artistic fields have alternative criteria.

For health sciences, the individual level of aggregation implies a calculus of the weighted average of the h index in three data bases: Web of Knowledge (ISI), Scopus and Google Scholar, with the following proportions: 50%, 30% and, respectively 20%. Individual data is aggregated at university level as the average of the individual weighted averages, raised at 3/2 power to account for the different distribution of citations for higher *h* levels.

We consider noteworthy the fact that this research can be regarded, de facto, an ex-ante analysis of the criterion regarding the impact of scientific research in the official funding methodology. For the sake of simplicity, we analyze the *h* index, but its distribution is similar to that of the *h* index raidsed at 3/2 power. We limited our research to a single data base due to the limited resources. Our preference for Scopus is motivated by issues of accuracy of the interrogations. Tol & Ruane (2007, apud. Tol 2008) found strong correlations between rankings of economics departments in Ireland using the three data bases taken into account in the official

funding methodology in Romania. We did similar computations for a limited and unsystematically chosen number of disciplines and found strong correlations. In our calculations, the values of the successive indexes, *g* and *h*, are highest if calculated using Google Scholar and lowest if calculated using Web of Science.

The scientometric alternative

A group of political scientists published extensively in recent years research papers on the use of scientometric indexes in public policies in Romania (Miroiu, Păunescu & Vîi, 2015; Păunescu & Hâncean, 2013; Vîiu, Vlăsceanu & Miroiu, 2012; Vîiu & Miroiu, 2013). One of their recurrent arguments is that the complex methodologies implying large numbers of indicators and resource exhausting data collection can be replaced with alternatives resting mainly on impact indexes from the *h* type family. Especially the *g* index (Egghe, 2006) is presented as an appropriate alternative, with the argument that it represents a synthetic measure of both quantity and impact, with a significant better discrimination potential than both *h* index and previous measures of research quality, and which yields fairly similar results to the previous official methodologies. The *h* type impact indexes are synthetic measures of quantity and impact, where the *g* index discriminates better different citation patterns and puts more weight on highly cited papers (Rubem, de Moura, & Soares de Mello, 2015) than the widely known *h* index.

Methodological aspects

For the purpose of this research, we collected publication data for the academic staff of the six universities of medicine and pharmacy, from the Scopus database. The data collection process took place between November 2014 and May 2015. We opted for individual author interrogations of the database, as we observed that numerous authors appeared in the database with multiple affiliations. We consider that such an approach, though significantly more laborious than an interrogation based on the affiliation stands chances to be substantially more comprehensible to commensurate the performance of the individual academics in a specific department. We excluded the self-citations of the author whose name was used for the interrogation. We also restricted the interrogation to the field „health sciences". We restricted the time window of the interrogations to the period between 2009 and 2014 in order to limit the influence of publications' age. We also performed a z-score analysis of the database containing individual publication and we observed that no less than 70% of the publications with a z-score greater than 3 represent a specific class: guidelines, recommendations, definitions, with numerous co-authors, in some cases, more than a thousand, and which sometimes can be rather attributed to collective professional bodies than to the authors themselves. We excluded all such publications, although their citations ranged from zero to hundreds. Thus, the primary database came to contain 22188 entries, corresponding to publications and respective citations found in Scopus database for each individual author. We also removed authors who we considered not relevant for the core business of these universities (such as languages or sports), keeping those belonging to the three groups of disciplines outlined in the relevant order of the Minister for Health: medical, surgical and paraclinical (MS, 2009), plus the fundamental sciences. Thus, after the computations of the values of the *h* and *g* index[3], the database came to contain 3467 unique entries for academic individuals, the large majority of the academic staff in the universities of medicine and pharmacy.

The "black-box" perspective: the collective level aggregation

The official funding methodology proposes averages as collective level aggregation rule. Before calculating the average for each of the medical schools within the health sciences sector

[3] We are indebted to Andra-Maria Barbu-Roescu for performing the computations in R Project for Statistical Computing.

in Romania, we discuss shortly the relevance of using this measure. First of all, we note the fact that the distributions of the values of the individual *h* and *g* indexes, as well as the individual sum of citations, present variable levels of skewness. The skewness is one of the indicators of the normal distribution of a population – or, as in our case, the lack of it. It also shows how far are the extremes from the central tendency, calculated as average. The skewness coefficients are calculated in
Table 1, below:

Table 3: *Skewness: h, g and citations' sum*

Faculty of medicine in...	***h* index SKEW**	***g* index SKEW**	**citations SKEW**
București	3.53	6.97	29.40
Cluj	3.52	4.85	18.80
Iași	4.60	3.82	15.20
Timișoara	3.17	3.15	7.32
Tg. Mureș	2.36	3.93	8.88
Craiova	2.60	2.65	5.69

The values of the coefficient indicate that none of the distribution is normal, and that the distributions of citations are the most skewed, while the distributions of the values of the *h* index is the least skewed amongst the three impact indicators. Secondly, we note that almost half of the overall population of academics in the six medical schools has null values for the two scientometric indexes and for the citations' sum. A distribution of the values of the *h* and *g* indexes is presented in Table 4, below:

Table 4: *h and g index values' distribution*

Value of the index	***h* index distributions**	***g* index distributions**
≥8	1.27%	5.36%
5-7	2.77%	7.07%
4	4.61%	4.99%
3	8.77%	7.56%
2	19.12%	10.76%
1	32.68%	16.07%
0	48.20%	48.20%

We consider that both the skewness of the distribution and the numerous null values raise serious doubts regarding the relevance of using the average as collective level aggregation rule.

For illustrative purposes, we represented the distribution of *h* index and *g* index values using a semi logarithmic scale, in Figure 1 and Figure 2, below. Seglen (1992) uses this approach for illustrating the skewness of citations' distribution; we adapted it for the distribution of the values of the two scientometric indexes. We included in the linear regression model approx. 99% of the data, as the large values of the *h* and *g* index are isolated cases which can be regarded as outliers.

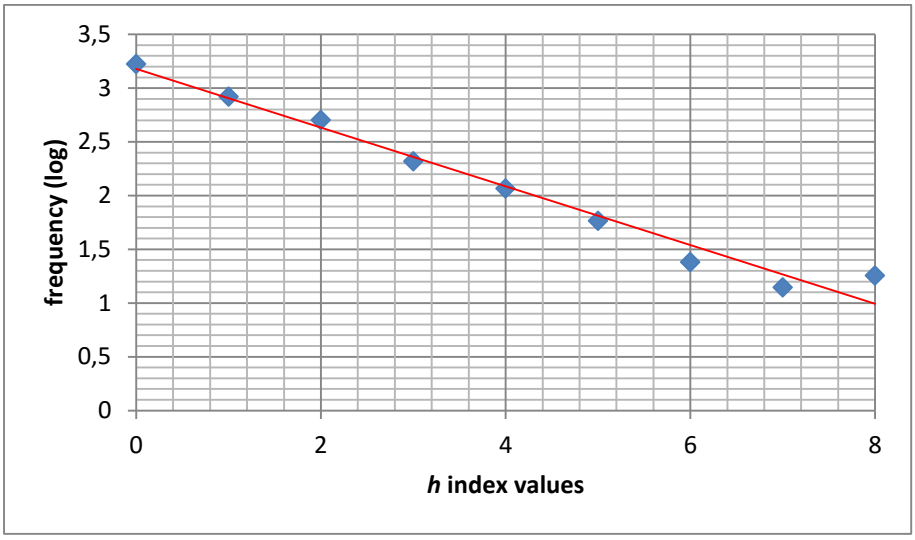

Figure 1: *h index values' distribution (semilog)*

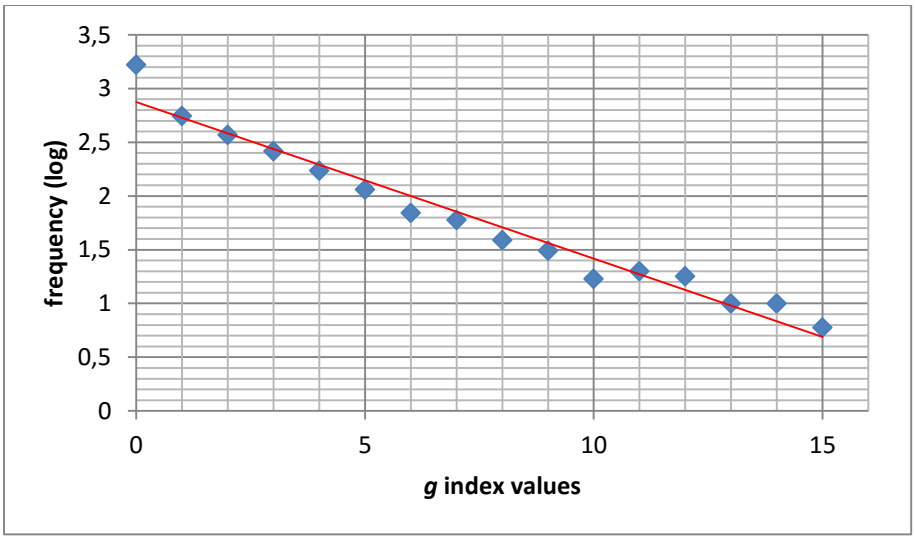

Figure 2: *g index values' distribution (semilog)*

The linear model explains a very large proportion of both semi-logharitmated distributions, as the values of the squared R-squared coefficient indicate: 0.97 for the *h* index and 0.96 for the *g* index, thus the distribution of the values of both indexes of the individuals from the six medical schools can be described by exponential functions. Seglen (1992) and other authors argue that the distribution of citations per articles is also described by an exponential curve – „the skewness of science". We will use in this section of the article analytic instruments developed for the study of the distribution of citations, as the literature on the topic is richer than that on the distribution of the measures of scientific productivity and impact in the communities of scholars.

We calculated averages of the *h* and *g* index, per university, and they correlate moderately with size, whereas both the successive *h* and *g* indexes correlate significantly with size: for the successive *h* index, the value of the coefficient is 0.82, while for the *g* index the value of the coefficient is 0.94. See

Table **5** for the calculations of the averages and for the values of the successive *g* and *h* indexes; the calculated values of the Pearson product-momentum correlation coefficient are presented in the bottom row of the table. The size of the faculty is represented by the total number

of academics in the data base, which represents the wide majority of the staff within the analyzed faculties, as explained in the section dedicated to the methodological aspects of the research.

Table 5: *h and g averages, successive h and g indexes, per university and correlation with size*

Faculty of medicine in...	The size of the faculty	Average *g* index	Average *h* index	Successive *h* index	Succesive *g* index
București	1134	2.04	1.12	9	20
Cluj	615	2.30	1.36	8	16
Iași	569	1.83	1.21	7	14
Timișoara	515	1.76	1.13	7	14
Tg. Mureș	345	0.93	0.61	4	9
Craiova	289	1.72	1.15	6	8
Correlation with the size of the faculty		0.55	0.30	0.82	0.94

The rankings yielded by the two successive indexes are similar, with the notable exception of the "smallest" two universities. The hierarchies are inverted for them, and, as a consequence, the *g* index, which is more sensitive to the impact of highly cited papers, appears to correlate stronger with size than the *h* index, which is more sensitive to productivity. Comparing the hierarchies in Table **5** with the hierarchy of the official ranking of the study programs in Table 2, we observe that the successive *h* and *g* index position the large universities on the top positions, whereas the average of the *h* and *g* individual indexes yields a slightly different hierarchy (the first two positions and inverted).

Besides the seeming correlation of the successive indexes with size, another property of these aggregation rules can be noticed at a close superposing of Table 4 and Table **5**: there are roughly 1 to 2% of the academic staff who actually contribute towards the value of the successive indexes, while the rest of the population could, from a mathematical perspective, have null values of the indexes and would not affect the results. This is the consequence of both the low values of the individual scientometric indices and of the aggregation rule. Especially the *h* successive index is affected by this apparent limited representativity of the individual academics composing the department.

We consider that the characteristics outlined above are sufficient reasons to proceed to an in-depth analysis of the distribution of the values of both *g* and *h* index in the population of academics in the six faculties of medicine within the specialized health sciences universities in Romania.

Into the "black-box": a distribution analysis

In order to complement the perspective offered by the collective level rules of aggregation, be those either the size-independent average indices or the successive indices that are size-dependent, we extended the focus of the analysis to the distribution of the entire population of individual academics in the faculties considered for this research. Firstly, we examined the data by partitioning the total population in groups delimited by the values of the individual scientometric indicators. The grouping corresponds for the top 1%, 10%, and 50% of the overall population, determined by the value of the individual scientometric index. The grouping is inspired by the impact indicators used in the CWTS Leiden Ranking (CWTS, 2015) and represents an adaptation from an analysis of the distribution of top cited publications. The number of academics from each faculty in the partitions determined by the above mentioned percentages is presented in the tables below, along with their averages.

Table **6** represents the calculations for the h index, while Table 7 presents the results from the same analysis, performed on the values of the individual *g* index.

Table 6: *h index: academics from each of the six faculties in the top 1%, 10% and 50% partitions of the total population and their average*

	No. of academics in top 1% / their average	**No. of academics in top 10% / their average**	**No. of academics in top 50% / their average**
București	13 /11.62	156 / 4.49	570 / 2.25
Cluj	12 / 9.18	109 / 4.53	339 / 2.42
Iași	6 / 11.67	73 / 4.30	345 / 2.00
Timișoara	7 / 11.00	76 / 4.55	259 / 2.25
Tg. Mureș	0 / 0.00	18 / 3.83	124 / 1.70
Craiova	6 / 8.67	32 / 4.91	159 / 2.09
Percentage of the total population	1.27%	13.38%	51.80%
Cut-off value of the h index	8	3	1
Average h index (population)	10.43	4.48	2.18

Comparing the distribution of the values of the *h* indexes outlined above, in Table **6**, with the values of the successive *h* indexes in Table **5**, we can observe that most of the academics who contribute to the latter are de facto amongst the top 1% of the academics in the field, with the notable exception of the Faculty in Tg. Mureș, whose scores are significantly lower. The averages for the top 1% academics exhibit a relatively low level of differentiation across the six faculties, a situation which adds to the concerns that the successive *h* index may be biased with size.

Table 7: *g index: academics from each of the six faculties in the top 1%, 10% and 50% partitions of the total population and their average*

	Academics in top 1% / their average	**Academics in top 10% / their average**	**Academics in top 50% / their average**
București	16 / 18.73	156 / 8.75	570 / 3.98
Cluj	6 / 27.50	105 / 8.90	339 / 4.24
Iași	5 / 25.00	60 / 8.87	345 / 3.08
Timișoara	6 / 19.00	61 / 8.25	259 / 3.49
Tg. Mureș	1 / 17.00	16 / 7.31	124 / 2.60
Craiova	1 / 17.00	33 / 7.88	159 / 3.13
Percentage of the total population	1.01%	12.43%	51.80%
Cut-off value of the g index	16	5	1
Average g index (population)	17.14	8.60	3.61

The *g*-index averages of the academics in the top 1% present a higher differentiation, which fades significantly when the segment of the population for which the average is calculated extends to approx. 10%. The distributions of the academics in the three intervals are almost identical for the two indexes, *h* and *g*.

In order to better understand the association between size and the values yielded after the aggregation of the data at the collective level, we further explored the similarity between the distributions of the values of the individual *h* and *g* index within the six faculties in the database. For this purpose we adapted the test of universality of university citations distribution, proposed

by Radicchi, Fortunato & Castellano (2008). Briefly, the original test consisted in considering a population of papers attributed to different universities and checking for proportionality in different intervals of the hierarchy determined by the citations of each paper; the null hypothesis consists in positing that the distribution of citations is similar for all the universities, thus the differences are only a matter of size. Instead of papers ranked by their citations, we used academics ranked by the value of their individual *h*, respectively *g* index. Assuming the individuals in any of the six faculties within the health sciences universities in Romania (academics having a certain *h* and *g* index) are scattered uniformly along the rank axis, for any value z% we would expect the average relative frequency of the number of individuals in any university to also be z% with a standard deviation $\sigma_z=\{[z(100-z)\Sigma_i(1-N_i)]/N_c\}^{1/2}$, which is equation (2) in Radicchi, Fortunato & Castellano (2008), and where N_c and respectively N_i are the number of universities and the number of academics in the *i*-th university. The results of the tests are presented in Table 8 and Table 9, below.

Table 8: *Test of universality for the values of the individual h index*

h≥ …	z%	STDEV Th.	STDEV Th./z%	STDEV Calculated	STDEV Calculated/z%
8	1.27%	0.0521	0.0411	0.6827	0.5379
5	4.04%	0.0930	0.0230	1.7606	0.4360
4	7.38%	0.1257	0.0170	2.5991	0.3520
3	13.38%	0.1692	0.0126	3.8540	0.2880
2	27.89%	0.2441	0.0088	5.6549	0.2027
1	51.80%	0.3323	0.0064	7.6750	0.1482

Table 9: *Test of universality of the values of the individual g index*

g≥ …	z%	STDEV Th.	STDEV Th./z%	STDEV Calculated	STDEV Calculated/z%
16	1.01%	0.0465	0.0461	0.4079	0.4040
9	4.24%	0.0953	0.0225	1.4048	0.3313
6	9.11%	0.1397	0.0153	3.3971	0.3727
5	12.43%	0.1631	0.0131	3.7476	0.3015
4	17.42%	0.1930	0.0111	4.6247	0.2655
3	24.98%	0.2310	0.0092	5.4906	0.2198
2	35.74%	0.2762	0.0077	6.3582	0.1779
1	51.80%	0.3323	0.0064	7.6750	0.1482

The null hypothesis is rejected for both *h* and *g* indexes, and the calculated standard deviations and coefficients of variation (Stdev/z%) depart significantly from the theoretical predictions. These results reject the hypothesis that the medical schools within the six health sciences universities in Romania represent similar distributions of academic staff, in terms of values of the individual *h* and *g* indexes, which differ mostly in terms of magnitude. Additionally, we can observe that the differences between the theoretically estimated values of the coefficients of variation and the calculated ones are significant, especially for the top intervals (those with lower values of z%), which indicate that the differentiation between the six medical faculties occurs mainly at the top of the hierarchy of individual values of the *g* and *h* indexes.

Further on, we focused the analysis on the upper part of the distributions. We thus used an alternative aggregation rule, following an adaptation of the Characteristic Scores and Scale (CSS hereafter) first introduced by Schubert, Glänzel & Braun (1987). This method has also been used

by Albarrán, Perianes-Rodríguez & Ruiz-Castillo (2015) and Albarrán et. al. (2011), for distributions across 219 sub-fields, respectively 36 countries, Ruiz-Castillo & Costas (2014) for distributions over 30 broad scientific fields, Perianes-Rodriguez & Ruiz-Castillo (2015) for distributions across 500 universities in the Leiden ranking (CWTS, 2015).

We translated the original version of the CSS, which was applied by the previously cited authors to citation distributions either across fields, universities or countries, to the distribution of academics' individual *h* and *g* indexes. We obtained an alternative aggregation rule which consists in partitioning the distributions into three classes of academics with respectively low, fair and very high values of the index, thus contributing differently to the faculty level of aggregation. The „low" contributors are the academics with a value of the index which is lower than the average, the „fair" contributors are those with a value of the index which is greater than the average and lower than the average of those above the average, while the „very high" contributors are the academics with a value of the index which is greater than the average of those above the average. The results are presented in Table 10, organized in two symmetrical parts, both of them consisting of three parts, each corresponding to the three partitions delimited by the two averages calculated within the CSS. In each of the three parts of the tables we included the average of the value of the *g* index, respectively *h*, of the academics within that partition, the proportion of the academics within each faculty within the total number of staff, and the proportion of citations cumulated by the academics within that partition, within each faculty, within the cumulated citations of the academics within that faculty. The bottom rows contain the Pearson correlation coefficient between the averages data set and the staff proportions data set, respectively between the averages and the citation proportions.

Table 10: *CSS, individual h and g index*

***h* index**	**"Low" contributors**			**"Fair" contributors**			**"Very high" contributors**		
	average	*count*	*citations*	*average*	*count*	*citations*	*average*	*count*	*citations*
București	0.31	71.52%	5.48%	2.31	21.25%	24.29%	5.68	7.23%	70.57%
Cluj	0.33	66.50%	5.10%	2.29	21.95%	24.73%	5.60	11.54%	70.18%
Iași	0.43	68.72%	7.45%	2.00	18.28%	19.25%	4.30	13.01%	73.30%
Timișoara	0.34	74.76%	7.79%	2.40	17.09%	23.12%	5.85	8.16%	69.09%
Tg. Mureș	0.00	63.77%	0.00%	1.00	20.29%	14.60%	2.61	15.94%	85.40%
Craiova	0.38	71.97%	6.99%	2.19	20.42%	28.18%	5.90	7.61%	64.84%
CORREL		*0.64*	*0.96*		*-0.08*	*0.84*		*-0.91*	*-0.96*
***g* index**	**"Low" contributors**			**"Fair" contributors**			**"Very high" contributors**		
	average	*count*	*citations*	*average*	*count*	*citations*	*average*	*count*	*citations*
București	0.45	72.40%	3.83%	4.02	19.58%	19.52%	11.63	8.02%	76.66%
Cluj	0.51	68.94%	3.92%	3.98	19.51%	18.12%	10.34	11.54%	77.96%
Iași	0.38	63.09%	3.48%	2.64	26.19%	19.76%	8.48	10.72%	76.76%
Timișoara	0.24	65.24%	6.62%	2.74	22.72%	17.08%	8.25	12.04%	80.89%
Tg. Mureș	0.00	63.77%	0.00%	1.35	22.90%	14.70%	4.78	13.33%	85.30%
Craiova	0.32	65.74%	3.37%	2.71	22.49%	20.69%	7.88	11.76%	75.94%
CORREL		*0.63*	*0.52*		*-0.65*	*0.56*		*-0.85*	*-0.74*

We also illustrated the CSS partitions in Figure 3 and Figure 4. The red lines represent the cut-off points between the three CSS partitions in the overall population of academic staff within the six schools of medicine.

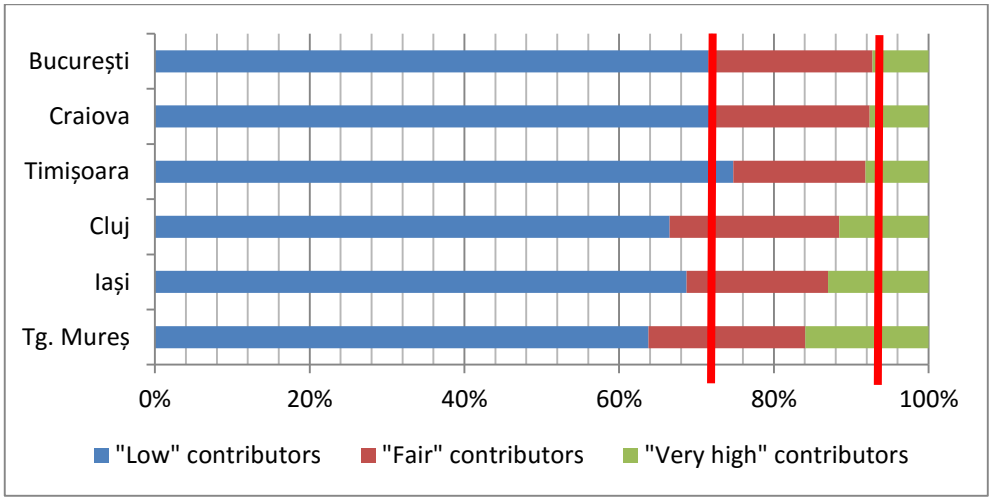

Figure 3: *CSS distributions, h index*

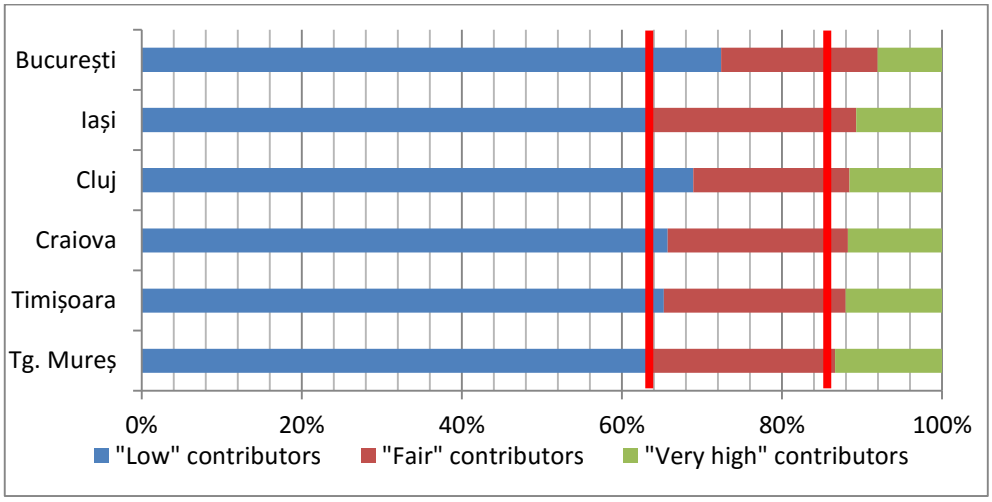

Figure 4: *CSS distributions, g index*

The CSS aggregation reveals interesting aspects of the successive *g* index. First and foremost, we note the similarity of universities *h* or *g* indexes distributions. They are all highly skewed, with the numerous "low" contributors accounting for a very small fraction of the total citations, even null in the case of one faculty, the fair contributors accounting for a proportion of the total citation comparable to their proportion in the total staff, whereas the wide proportion of citations is cumulated by a small percentage of "very high" contributors.

Particularly for the *g* index CSS, we note the significant degree of correlation between the hierarchy generated by the successive *g* index and the hierarchy generated by the average of the „very high" contributors. We also note the robustness of the hierarchy in the top half of the *g* index distributions: the hierarchy within the „fair" contributors' category is fairly similar to that in the „very high" contributors' hierarchy. Last but not least, we note the significant values of the correlation coefficient between the size of the category and its average. The negative sign of the coefficient indicates that more skewed distributions are associated with higher averages. We consider these findings as arguments in favor of considering the successive g index, or the

average of the „very high contributors" as a robust and synthetic aggregation rule and also as a confirmation of the skewness of science in this field in Romania.

For the *h* index we can observe the same significant, negative correlation between the average of the „very high" contributors' category and its size, but the hierarchy within the two categories in the upper half of the distribution presents major differences. The differences can be a consequence of the fact that the *h* index presents lower and less varied values than the *g* index, hence the categorization induced by the two averages within the CSS is less sharp. Last but not least, we note that the averages of the "very high" contributors within the *h* index distributions yield different hierarchies compared to both the successive *h* index and the averages of the "very high" contributors within the *g* index distributions.

Conclusions

The distribution analyses did not particularly explain the bias of the successive *g* and *h* indexes towards the size of the faculties, which is indicated by the correlation of the values of the successive indexes with the size of the faculties. A possible explanation is that the relation between size and impact is more subtle, as advanced by Dienes (2015), who argues that the scientific impact is a function of the scientific community the researcher/academic is part of, including size. Applied to our analyses, a possible explanation that needs empirical testing is that a larger university represents larger networking potential, thus a larger citation base.

The alternative aggregation rules, such as the CSS of the distributions are considerably more relevant than the average, especially when the distributions are highly skewed as is the case of the scientific productivity. The hypothesis of the universality of *h* or *g* indexes distributions is rejected; especially for lower values of z% (approx. 1%, 4%, 9%, 12%) the individuals are not scattered uniformly across the rank axis; this feature is more distinctive in the case of the *g* index, where the dispersion of the values is more pronounced. For instance within the first 1% of individuals ranked according to their *g*-index, those from Bucharest represent 1.41% of their university faculty members, while those from Tg. Mureș and Craiova represent 0.29% and respectively 0.35% of their faculty members. The dispersion is considerably higher than the theoretical one calculated according to equation (2) in Radicchi, Fortunato & Castellano (2008).

Secondly, all university distributions of either *h* or *g* indexes are highly skewed. For the *g*-index distribution, the average is 14.2 points above the median, while, overall, approx. 12.43% of the individuals account for approx. 82.93 % of citations. The universities that have higher averages and lower percentages of individuals in the last partition, and thus present a more pronounced skewness, are also better performing across the successive *g* index. This suggests that the ranking is rather given by the upper segment of the distributions within each university, which confirms the skewness of productivity research which has been argued in similar studies (Perianes-Rodriguez & Ruiz-Castillo, 2015; Ruiz-Castillo & Costas, 2014; Seglen, 1992). Given the high skewness, we consider that taking the Scopus database as reference is a better choice than Web of Science, as for the latter probably the values of the individual indexes would be lower and thus the skewness even more pronounced. Ruiz-Castillo & Costas (2014) note that Scopus is used in scientometric analyses as an alternative, which is more comprehensible than Web of Science. See examples of analyses performed on Scopus in Ioannidis, Boyack & Klavans (2014) or SCImago (2007). Furthermore, we note that Ruiz-Castillo & Costas (2014) found that 68% of authors in all fields published only one Web of Science indexed article over a period of nine years; their research spans over the period 2003-2011 and includes an impressive sample of 17.2 million disambiguated authors. The previously quoted authors found that the proportion of authors with a single article indexed in Web of Science within the nine years span is 71.5% in health sciences, 66.8% in clinical medicine, 64.8% in biomedical sciences and 64.8% in basic medical sciences. All these authors could not exceed an *h* or *g* index of 1, in the most advantageous circumstances.

Thirdly, we note that our study is limited by the fact that the values of individual *h* and *g* values are natural numbers, hence presenting a discrete spectrum, which increases the risk of statistical artefacts as the percentage intervals (percentiles) cannot be accurately calculated across universities. For further research, we even suggest the use of rational variants of the *h*

index (Anderson, Hankin & Killworth, 2008; Ruane & Tol, 2008) and *g* index (Tol, 2008) that better discriminate between lower values of *h* and *g,* and whose spectrum of values is more dense than that of the original version of the indexes. An alternative avenue of research is to explore indexes which have a higher discrimination potential among authors with low numbers of citations, as already proposed by Anania & Caruso (2013). The rules of aggregation we discussed in this article place emphasis on the "very high" contributors, whereas the contribution of the wide majority of the academics may seem irrelevant. The g index is eloquent in this sense: currently, more than 65% of academics have *g* index values below average, while the scientific productivity of an institution is largely given by 10% of its academics. This elitism exhibited by the discussed aggregation rules reinforces the "skewness trap" as there are no incentives for those below average to become more productive, given the poor discriminative power of *h* and *g* indexes. Alternative metrics may help better discriminate between average and below average academics, and thus to allow meaningful competition based on scientific productivity and impact, in wider intervals of the population of academics.

Our results indicate that the metrics used for the evaluation of research are not neutral: they tend to provide incentives only at the top of the distribution and they assume that the research performance of a social entity such as a department/faculty is a function of the most performing core and not of the all individuals composing it. Thus, the younger, less productive researchers, with a lower impact, contribute rather marginally to the aggregate scores of the universities, unless co-opted in the core production teams.

References

1. Albarrán, P., Crespo, J. A., Ortuño, I., & Ruiz-Castillo, J. (2011). The skewness of science in 219 sub-fields and a number of aggregates. Scientometrics, 88(2), pp. 385–397. http://doi.org/10.1007/s11192-011-0407-9

2. Albarrán, P., Perianes-Rodríguez, A., & Ruiz-Castillo, J. (2015). Differences in citation impact across countries. Journal of the Association for Information Science and Technology, 66(3), pp. 512–525. http://doi.org/10.1002/asi.23219

3. Anania, G., & Caruso, A. (2013). Two simple new bibliometric indexes to better evaluate research in disciplines where publications typically receive less citations. Scientometrics, 96(2), pp. 617–631. http://doi.org/10.1007/s11192-013-0951-6

4. Anderson, T. R., Hankin, R. K. S., & Killworth, P. D. (2008). Beyond the Durfee square: Enhancing the h-index to score total publication output. Scientometrics, 76(3), pp. 577–588. http://doi.org/10.1007/s11192-007-2071-2

5. ARWU. (2015). Academic Ranking of World Universities 2015. Retrieved October 10, 2015, from http://www.shanghairanking.com/ARWU2015.html

6. CWTS. (2015). CWTS Leiden Ranking 2015. Retrieved October 5, 2015, from http://www.leidenranking.com/methodology/indicators

7. Dienes, K. R. (2015). Completing h. Journal of Informetrics, 9(2), pp. 385–397. http://doi.org/10.1016/j.joi.2015.01.003

8. Egghe, L. (2006). Theory and practise of the g-index. Scientometrics, 69(1), pp. 131–152. http://doi.org/10.1007/s11192-006-0144-7

9. Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America, 102(46), pp. 16569–16572. http://doi.org/10.1073/pnas.0507655102

10. Ioannidis, J. P., Boyack, K. W., & Klavans, R. (2014). Estimates of the continuously publishing core in the scientific workforce. PLoS ONE, 9(7). http://doi.org/10.1371/journal.pone.0101698

11. MECS. ORDIN privind aprobarea Metodologiei de alocare a fondurilor bugetare pentru finanţarea de bază şi finanţarea suplimentară a instituţiilor de învăţământ superior de stat din România, pentru anul 2015 [Order regarding the approval of the Methodology for the distribution of public budgets for basic and suplemmentary funding of state higher education institutions in Romania (2015)]. Retrieved August 25, 2015, from http://www.cnfis.ro/wp-content/uploads/2012/08/ordin-3889-din-2015.pdf,

12. MECTS. ORDIN privind aprobarea metodologiei de prelucrare a datelor şi informaţiilor colectate în vederea realizării evaluării primare a universităților și a evaluării programelor de studii universitare cu scopul clasificării universităților și ierarhizării prog (2011) [Order regarding the approval of the methodology for the processing of data and information collected for the primary evaluation of the universities and of the study programe, for the purpose of the classification of the universities and the ranking of the study programmes]. Retrieved August 25, 2015, from http://www.edu.ro/index.php/articles/16066.

13. Miroiu, A., Păunescu, M., & Vîiu, G.-A. (2015). Ranking Romanian academic departments in three fields of study using the g-index. *Quality in Higher Education*, 1–24 (in press).

14. Miroiu, A., & Vlasceanu, L. (2012). Relating Quality and Funding: The Romanian Case. In A. Curaj, P. Scott, L. Vlasceanu, & L. Wilson (Eds.), *European Higher Education at the Crossroads* (pp. 791–807). Springer Netherlands. http://doi.org/10.1007/978-94-007-3937-6_41

15. MS. Ordin nr. 1509/2008 privind aprobarea Nomenclatorului de specialități medicale, medico-dentare şi farmaceutice pentru rețeaua de asistență medical [Order regarding the approval of the register of medical, medical-denthal and pharmacy specialities, for the health system] (2009). Retrieved August 25, 2015, from https://www.eshg.org/fileadmin/www.eshg.org/documents/countries/RomaniaFin1.pdf

16. OECD. (2013). Assessment of Higher Education Learning Outcomes Feasibility. Study Report, Volume 3: Further Insights. Retrieved August 25, 2015, from http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume3.pdf

17. Păunescu, M., & Hâncean, G.-M. (2013). Ranking the Romanian Departments of Sociology. Comparative Results of Different Evaluation Methodologies. *Quality Assurance Review for Higher Education*, 5(1-2), pp. 5–17.

18. Perianes-Rodriguez, A., & Ruiz-Castillo, J. (2015). University Citation Distributions. Paper presented at Istanbul: 15th International Society of Scientometrics and Informetrics Conference, Istanbul.

19. Radicchi, F., Fortunato, S., & Castellano, C. (2008). Universality of citation distributions: toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences of the United States of America*, 105(45), pp. 17268–17272. http://doi.org/10.1073/pnas.0806977105

20. Ruane, F., & Tol, R. S. J. (2008). Rational (successive) h-indices: An application to economics in the Republic of Ireland. *Scientometrics*, 75(2), pp. 395–405. http://doi.org/10.1007/s11192-007-1869-7

21. Ruiz-Castillo, J., & Costas, R. (2014). The skewness of scientific productivity. *Journal of Informetrics*, 8(4), pp. 917–934. http://doi.org/10.1016/j.joi.2014.09.006

22. Rubem, A. P. dos S., de Moura, A. L., & Soares de Mello, J. C. C. B. (2015). Comparative analysis of some individual bibliometric indices when applied to groups of researchers. *Scientometrics*, 102(1), pp. 1019–1035. http://doi.org/10.1007/s11192-014-1428-y

23. Schubert, A. (2007). Successive h-indices. *Scientometrics*, 70(1), pp. 201–205. http://doi.org/10.1007/s11192-007-0112-x

24. Schubert, A., Glänzel, W., & Braun, T. (1987). A new methodology for ranking scientific institutions. *Scientometrics*, 12, pp. 267–292.

25. SCImago. (2007). SJR - SCImago Journal & Country Rank. Retrieved August 24, 2015, from http://www.scimagojr.com

26. Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science*, 43(9), pp. 628–638.

27. Steedle, J. (2010). On the foundations of standardized assessment of college outcomes and estimating value added. In K. Carey & M. Schneider (Eds.), *Accountability in American higher education* (pp. 7–32). New York: Palgrave Macmillan.

28. THE. (2015). World University Rankings 2015-2016, Times Higher Education. Retrieved October 10, 2015, from https://www.timeshighereducation.com/world-university-rankings/2016/world-ranking?gclid=CjwKEAjw7uKwBRDUlJvRo-z6rgMSJACbmSBh6FCtXcDEfC88lP8PGPEujbhwYfS9hS7FiPe3m_JC1xoChMjw_wcB#!/page/0/length/25

29. Tol, R. S. J. (2008). A rational, successive g-index applied to economics departments in Ireland. *Journal of Informetrics*, 2(2), pp. 149–155. http://doi.org/10.1016/j.joi.2008.01.001

30. Țeca, M. (2011). Viziune de ansamblu asupra modelului matematic de construcție și utilizare a indicatorilor relativi de calitate în finanţarea învăţământului superior utilizat în perioada 2003 – 2011 [Bird-eye view over the mathematic model of the construction and usage of the relative quality indicators in the funding of higher education between 2003 and 2011]. *Quality Assurance Review for Higher Education*, 3(1), pp. 81–92.

31. Vîiu, G.-A., & Miroiu, A. (2013). Evaluarea cercetării universitare din România . Abordări metodologice alternative [Evaluation of University Research in Romania. Alternative Methodological Approaches]. *Revista de Politica științei și Scientometrie*, 2(2), pp. 1-20.

32. Vîiu, G.-A., Vlăsceanu, M., & Miroiu, A. (2012). Ranking the Romanian Political Science Departments. *Quality Assurance Review for Higher Education*, 8(5212), pp. 1–15.

33. Zulean, M., Ioniță, I., & Vîiu, G.-A. (2014). Raport de evaluare a guvernanţei sistemului public de cercetare, dezvoltare și inovare din România 2007 – 2013 [Report regarding the governance of public research, development and innovation in Romania 2007 – 2013]. Retrieved August 25, 2015, from http://www.incd2020.ro/sites/default/files//Raport_de evaluare a guvernantei sistemului public de CDI din Romania 2007-2013.pdf