Viorel Proteasa

PRESCRIPTION REGARDING HIRSCH'S INDEX: A DISTRIBUTION ANALYSIS IN MEDICINE SCHOOLS IN ROMANIA

*Viorel Proteasa**
West University of Timisoara, Centre for Public Policy
viorel.proteasa@gmail.com

ABSTRACT

In this article I dicuss the distributions of the values of a rational variant of the h index (Anania and Caruso 2013) for the academics within the universities of medicine and pharmacy in Romania. I argue that the averages and the successive h index, two of the alternatives to quantify research impact within the Romanian universities, are determined by completely different properties of the distributions. I adapt the Characteristic Scores and Scale (A. Schubert, Glänzel, and Braun 1987) and I argue that it represents an appropriate mean to compare distributions of the h values across the medical schools considered.

KEYWORDS

- *Rational h index*
- *Higher education policy*
- *Research evaluation*
- *Distribution analysis*

Introduction

Starting with 2007, higher education funding policies in Romania placed increased emphasis on indicators related to research evaluation. This approach can be considered as a measure intended to strengthen the incentives for university research, but, at the same time it has been criticized for relying on too complex and extensive methodologies, which burdened both the universities due to heavy reporting, and the central agency due to validation and computation of the data. The 2015 proposal for a funding methodology contains four indicators for research, one of them quantifying the impact of academics' publications. An alternative consisting of synthetic bibliometric indexes has been proposed from the scholar community (Miroiu, Păunescu, and Vîiu 2015; Păunescu and Hâncean 2013; Vîiu, Vlăsceanu, and Miroiu 2012).

In this article I discuss the significance of some of the main components of the two alternatives, namely averages and second order *h*-type indexes, using an exhaustive data base containing the population of academics in the faculties of medicine in the six health sciences universities in Romania. The distinctiveness of my approach consists

* *This paper is supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU/159/1.5/S/137390.*

of: (1) computing a rational variant of the *h* index, (2) adapting analytical sequences from analyses of individuals' citations or citations per papers to the analyses of rational variants of individuals' *h*-type index, (3) proposing an alternative aggregation of the individual scores which includes a distribution analysis component. In the first section I provide a brief overview of higher education funding policies in Romania. In the second part I introduce some conceptual and methodological information regarding my research. In the second section I analyse some of the particularities of the distributions of a rational version of the *h* index in the six medical schools, focusing on the central tendency of the six distributions and on the significance of the second order, successive *h* index. In the concluding section I link the results of my research with the Romanian policy context.

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 and Miroiu (2013). I 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 programmes used since 2012, and (3) the quality indicators discussed in 2015 and postponed for implementation until 2016.

The "quality indicators": 2003-2011

An importnat milestone in the evolution of higher education funding policies has been the introduction of a formula as basis for allocation in 2001. Before that moment, universities were still funded based on reported costs[1]. Innitialy the formula quantified student in-take, hence it was considered to be driven by „quantity". „Quality indicators" have been added starting with 2003. Both the structure and their weight were subject to change since their introduction. Research outputs, especially publications, were emphasised starting with the 2007 methodology, when an indicator for research was included amongst the „quality indicators". Zulean, Ioni**tă** and Vîiu (2014) analysed the marginal increase due to the research indicators 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 was strongly biased towards the size of the universities, as advanced by Teca (2011) and Miroiu and Vlasceanu (2012). Thus, in the case of health sciences Romanian universities, "Carol Davila" University in Bucharest 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.

Ranking study programmes

The ranking of study programmes 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,

1 A more detailed account is available in Miroiu and Vlăsceanu (2012).

research weighted no less than 60% of the total score. For study programmes in medicine, publications accounted to 75% of the research score, which results in a weight of 45% of the total score (Vîiu and Miroiu 2013).

Quality indicators reloaded: 2015

The funding methodology proposed in 2015 reverts to the concept of quality indicators (MECS 2015). The implementation of the new funding scheme was postponed for 2016. The proposed quality indicators are unique for all fields of studies, but the ranking for the purpose of funding is performed per field of study. The indicator which refers to the *h* index represents 10% of the total score. For health sciences, the individual level of aggregation implies a calculus of the ponderated 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 ponderated averages, at 3/2 power.

About the research

Following the publication of the official ranking of the study programmes in 2012, a serie of Romanian scholars (Miroiu, Păunescu, and Vîiu 2015; Păunescu and Hâncean 2013; Vîiu, Vlăsceanu, and Miroiu 2012) proposed an alternative evaluation consisting of second order *h*-type (Hirsch 2005) bibliomteric indexes in the logic of the indexes proposed by Schubert (2007) and Ruane and Tol (2008). They argued that second order *h*-type indexes are synthetic indicators that can be easily computed from publicly available data, and they yield results similar to those obtained through extensive and complex methodologies, such as those used in the past higher education funding policies. The 2015 proposal borrows some of the features of the methodologies advanced by the above mentioned academics, in the sense that the indicator regarding the academics' impact is based on first order, individual *h* indexes. At the same time, it differs in terms of way in which individual scores are aggregated at department level. Proteasa, Păunescu, and Miroiu (2015) argued that computing the average of the *h* index for the purpose of comparing the six schools of medicine within the health sciences universities in Romania is problematic due to some of the characteristics of the distributions of the values of the *h* index: (1) the indexes take few, natural values; (2) almost half of the academics have a null value of the indexes; and (3) the distribution of the values of the index is highly skewed and takes the form of a logarithmic function.

The alternative - the second order, successive *h* index, poses a problem of representativeness: the academic staff who contribute to the index represents roughly top 1 to 2% of faculty. Moreover, the majority of these academics have outliers' behaviour.

The research question in this article is two-folded: on one hand, I explore which measures of the central tendency can be used in order to compare meaningfully the distributions of the *h* values within the six medical schools, and, on the other hand, I attempt to shed light on the significance of the second order, successive *h* index, appealing to the same strategy of analyzing the distributions of the *h* values within the academics from the six medical schools.

The data

The analyses in this article are computed on a data base consisting of publication data for the academic staff of the faculties of medicine in the six health sciences universities in Romania, from the Scopus data base. The data collection process took place between November 2014 and May 2015 and the interrogations have been performed per individual, not per institution. This approach, though significantly more laborious, stands chances of being more comprehensive than those based on institutional interrogations, as many authors have multiple affiliations, or simply the name of the university is spelled inconsistently across the Scopus data base. Self-citations of the author whose name was used for the interogation were excluded and the interrogation has been restricted to the field „health sciences". The time window of the interogations has been also limited to the period between 2009 and 2014 in order to exclude the influence of publications' age. Publications with numerous co-authors, in some cases more than a thousand, which represented guidelines, recommendations, definitions, were removed from the data base, following a z-score analysis of the citations of each paper in the data base (more than 22 000): 70% of the publications with a z-score greater than 3 represented such publications. Authors which were considered not relevant for the core business of these universities (such as languages or sports) were also removed from the data base, resulting in a total of 3374 academics in the data base, belonging to the three groups of disciplines outlined in the official regulations: medical, surgical and paraclinical (MS 2009), plus the foundamental sciences.

The research strategy

The inadequacy of the text book central tendency measures for the comparison of the six distributions is obvious when we compute medians – see Table 1 below.

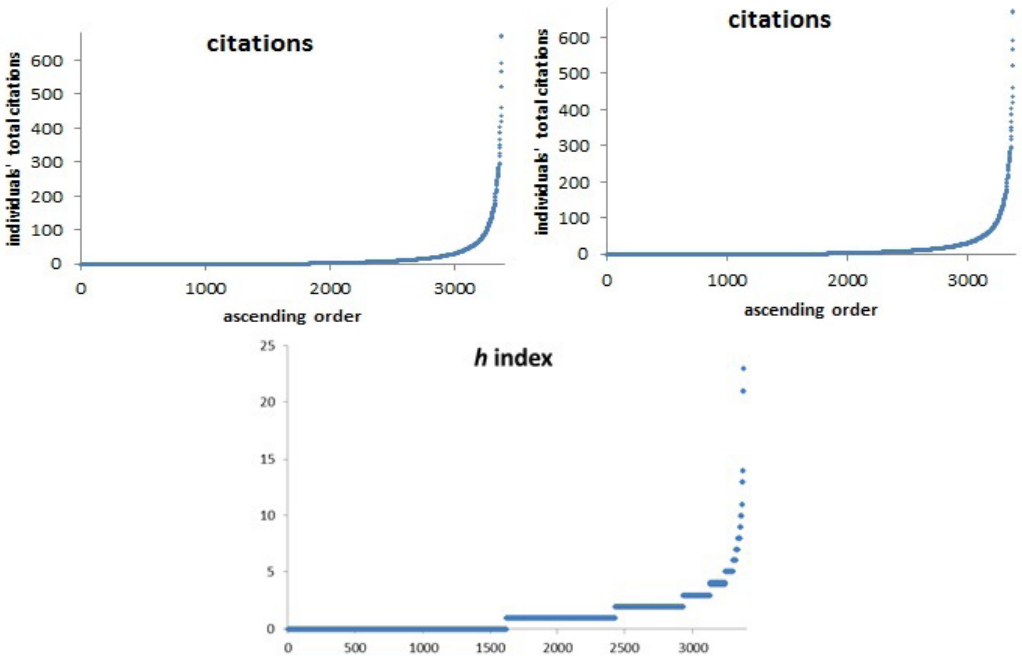
Table 1: Median of the h values of the academics within the six medical schools

	Median	Rank median
Bucharest	1	1
Cluj	1	1
Iaşi	1	1
Timişoara	1	1
Tg. Mureş	0	2
Craiova	1	1

Therefore, this article continues the path opened by Proteasa, Păunescu, and Miroiu (2015) who adapted statistical treatments from analyses of individuals' citations or citations per papers to the distribution of the values of h type indexes, justified by the fact that their distributions are similar. Citations or citations per paper have

been studied by scholars of the "skewness of science" (Albarrán et al. 2011; Perianes-Rodriguez and Ruiz-Castillo 2015; Ruiz-Castillo and Costas 2014; Seglen 1992)or sub-fields. Reference and citation distributions have very different characteristics across sub-fields. However, when analyzed with the Characteristic Scores and Scales (CSS.
I tested empirically the hypothesis of the similarity of the functions that describe the distributions for the values of the *h* index, the citations per author and the citations per paper for all authors. The distributions are presented in the graphics below. For illustration purposes, I ordered ascending the individual academic staff, on the basis of their citations, citations per paper and value of the *h* index, respectively, and I removed some of the extreme outliers:

Figure 1: Distribution of individuals' total citations, citations per paper, and h index value



I computed the histograms, using the unit as interval of frequencies for the distribution of citations per paper and for the values of the *h* index, while for the distribution of total citations I used a ten unit interval scale. I extracted the natural logarithm from the calculated frequencies and computed a linear regression. I excluded from the regression the approximation of the top 1% of the values. The R-squared coefficients, as well as other data relevant for the regression models are presented in Table 2, below:

Table 2: Regression models for citations, citations per paper, and the values of the h index

	R-squared	Population included in the regression models	Values included in the regression
Citations	0,82	98,46%	0 - 199
Citations per paper	0,81	98,96%	0 – 18,99
h index	0,99	98,73%	0 - 7

The values of the R-squared coefficient indicate that an impressive share of the variation within the distributions presented in Figure 1 can be explained by the proposed model, which constitutes a strong argument in favor of adapting analytical strategies from the study of individuals' citations or citations per paper to the study of the distribution of the values of the *h* index, for the six medical schools in the data base.

One of the limits of the analysis of Proteasa, Păunescu, and Miroiu (2015) is the fact that the *h* index takes few, natural values. In order to overcome this limit, in this article I compute the rational version of the *h* index proposed by Anania and Caruso (2013):

$$h_{rational} = h + [1 - h^2 / totcit], \text{ for } h > 0 \text{ and}$$
$$h_{rational} = 0, \text{ for } h = 0.$$

Amongst the variants to this rational version of the *h* index I mention those advanced by Anderson, Hankin, and Killworth (2008), and Ruane and Tol (2008). All these variants preserve the monotonicity properties of the original *h* index, and have a wider, yet incontinuous specter of values. They operate differences in the interval between *h* and *h + 1*, preserving, thus, the general order generated by the *h* index, and adding discriminative power at subunit level.

I opted not to use the variant proposed by Anderson, Hankin, and Killworth (2008) because its construction appeals to not so frequently used matemathical concepts (Dufree squares) and I considered this incongruent with the conceptual simplicity of the original *h* index. The variant proposed by Ruane and Tol (2008), which quantifies the minimum distance (in citations) to the next value of the *h* index (*h + 1*) was also discarded because it is less inclusive than the version proposed by Anania and Caruso (2013).

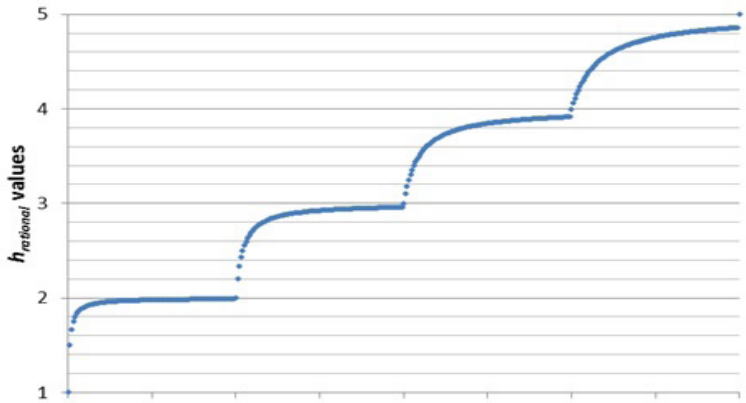
The version proposed by Ruane and Tol (2008) has the following formula:

$h^A = h + 1 - m / (2h + 1)$, where m represents the sum of additional citations needed for an increase of the value of *h* with one unit and takes values between 1 and $2h + 1$.

Thus, the version of Ruane and Tol (2008) is computed on the basis of the top *h + 1* papers of an author, while the version of Anania and Caruso (2013) is computed on the basis of all cited papers of the same author. I consider the version of Anania and Caruso (2013) to be more adequated for authors with low values of the *h* index, both conceptually, because it takes into account all citations, but also from the perspective of the discriminative power: theoretically, its spectre of values is infinite.

I plotted in Figure 2 the values of a $h_{rational}$ index function; only the first five units are included, alongside with the first 100 subunit values in each unit.

Figure 2: Theorethical values of the $h_{rational}$ index



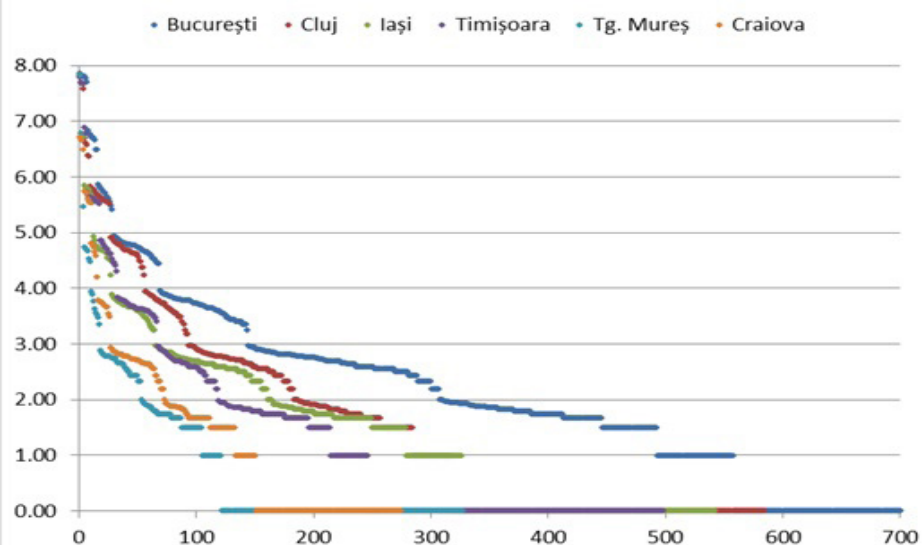
The sub-unit spectre of the $h_{rational}$ index presents wider distances between its values towards its lower end and tends to be continuous towards the upper end. The distances at the lower end of the spectre are wider for the lower values of the unit - the value of the respective *h* index.

The analysis

I illustrated graphically the distributions of the values of the $h_{rational}$ index in Figure 3, below. For illustrative purposes, the graphic does not include values of the index which are greater than 8 – those values that were excluded from the regression model for the entire population (Table 2), and does not include all the null values for the medical school in Bucharest.

A visual inspection of the data reveals that the curves have a logarithmic shape. Further on, the function appears to be influenced by the size of the faculty, as the curves decrease smoother in the case of the larger medical schools. It can be also observed that the specter of the *h* rational index in the six medical schools tends to be more continuous between values 1,5 and 3 – where most of the non-null values of the distributions tend to concentrate.

Figure 3: The distributions of the values of $h_{rational}$



I run regression models for each of the six medical schools, similar to that performed for the distribution of the values of the *h* index for the entire population. For reasons of comparison across medical schools, I did not include in the regression models the academics in the approx. top 1% in the overall populations. These individuals are all situated at a large distance from the average of the overall population (zscore ≥ 4,51), and generally at similar distances from the average of the individual distributions, in the case of four of the medical schools. The faculty in Tg. Mureş constitutes and exception, due to the inclusion of the entire population in the regression model. The other exception is the medical school in Cluj, which has a more compact distribution of the values of the $h_{rational}$ index.

Table 3: Regression models for the $h_{rational}$ index in the six medical schools

	Population included in the regression models	RSQ	zscore ≥ ...
Bucharest	98,84%	0,9882	4,68
Cluj	97,99%	0,9784	3,69
Iaşi	99,09%	0,9777	4,33
Timişoara	98,61%	0,9579	4,58
Tg. Mureş	100,00%	0,9429	:
Craiova	97,86%	0,9666	4,52
integrated	98,73%	0,9961	4,51

The R-squared coefficients indicate that the individual distributions of the values of the $h_{rational}$ index can also be described using a logarithmic function. I consider these characteristics of the individual distributions justify an analytical strategy of

partitioning the data for identifying the central tendency.

The central tendency

For the purpose of identifying the central tendency, I adjusted 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).

The adjustment consists, first of all, in defining the "outstanding" contributors as the academics with a value of the $h_{rational}$ index which is greater or equal to 8. As mentioned before, these academics represent roughly 1% of the total population and will not be included in the computations in this subsection, due to the fact that they may induce disturbances in the data.

The CSS consists in computing the average of all the values of the $h_{rational}$ index and, afterwards, the average of the population with values of the $h_{rational}$ index which are greater than the average, which will be further referred to as the second average. The first and the second average, plus the cut-off point of $h_{rational} = 8$, delimit the following categories, labeled in quasi-accordance with the original version of the CSS:

- The "low" contributors – the academics with a value of the $h_{rational}$ less or equal to the average;
- The "fair" contributors – the academics with a value of the index which is greater than the average and lower than the second average;
- The "high" contributors - the academics with a value of the $h_{rational}$ higher than the second average, but lower than 8. In the original version of the CSS, this category is labelled "very high". I considered the values of the index of the academics to fall in this category does not justify the original label;
- The "outstanding" contributors - the academics with a value of the $h_{rational}$ index which is greater or equal to 8.

The percentages of the academics belonging to each of the four categories outlined are presented in Table 4 below. The two averages divide the distributions in fairly simillar proportions: the coefficients of variation in the columns in Table 4 are 0,08 for the proportions of the category labeled as "low", 0,11 for "fair" and 0,14 for "high". Most of the variation is due to the distinctiveness of the proportions in the medical school in Tg. Mureş, which is rather particular compared to the rest of the medical schools.

Table 4: The proportions of the academics belonging to the four categories in the six schools of medicine

	average	2nd average	"low"	"fair"	"high"	"outrstanding"
Bucharest	1,32	2,84	54,99%	28,97%	14,88%	1,16%
Cluj	1,50	3,13	54,87%	27,68%	15,44%	2,01%

	average	2nd average	"low"	"fair"	"high"	"outrstanding"
Iaşi	1,43	2,62	48,26%	29,80%	21,02%	0,91%
Timişoara	1,27	2,80	56,24%	27,52%	14,85%	1,39%
Tg. Mureş	0,84	2,26	63,00%	21,10%	15,90%	0,00%
Craiova	1,33	2,62	50,71%	25,71%	21,43%	2,14%

As the exponential character of the distributions is determined especially by their „heavy tails" (see Figure 3), I tested the „fair" and „high" partitions for linearity. The results are presented in Table 5, below. The R-squared coefficients are very high, especially for the fair category. Overall, the coefficients are higher for the „fair" category, compared to the „high" category of the corresponding school of medicine.

Table 5: Regression models, „fair", „high", and „fair" and „high"

	RSQ „high"	RSQ „fair"	RSQ „fair" and „high"
Bucharest	0,84	**0,95**	0,81
Cluj	0,90	**0,96**	0,86
Iaşi	0,80	**0,89**	0,80
Timişoara	0,84	**0,89**	0,81
Tg. Mureş	0,65	**0,88**	0,75
Craiova	0,77	**0,82**	0,77

The linearity of the distributions within the „fair" categories and the rather continuous specter of the $h_{rational}$ index in the interval between the average and the second average constitute arguments in favour of computing the average of these segments of the distributions. Moreover, the coefficients of variation within the „fair" categories are situated between 0,17 and 0,21, indicating similar levels of intermediate variation. Given the proportion of the academics in the „fair" category - close to 25% and the position of the „fair" category within the entire population, I compared the averages of the category with the third quartile of the entire population, „outstanding" category included. The computations are presented in Table 6. The median is inadequated for a comparison of the six distributions on the basis of their central tendency due to the discontinuities of the specter of the $h_{rational}$ index between 1 and 1,5, were most of the medians are situated, and due to the high frequencies of those few values of the $h_{rational}$ in that part of the specter.

Table 6: Comparison: „fair" averages and third quartiles

	QRT3 – entire population		average „fair" – „outstanding" excluded		Median – entire population	
	values	rank	values	rank	values	rank
Bucharest	2,54	3	2,10	3	1,00	2
Cluj	2,71	1	2,28	1	1,00	2
Iaşi	2,56	2	2,15	2	1,50	1
Timişoara	1,96	5	1,98	5	1,00	1
Tg. Mureş	1,67	6	1,52	6	0,00	3
Craiova	2,46	4	1,80	4	1,00	1

The third quartile and the average of the "fair" category are highly correlated: the Pearson coefficient of is 0,82, while the Spearman correlation is 1.

Further on, given that in most of the medicine schools the „high" categories exhibit a reasonable linearity, I further compared the distributions in the „fair" category with the upper category. The computations are presented in Table 7. Though the rankings are different, the values of the averages are highly correlated: 0,97.

Table 7: Comparision: „fair" and „high" categories

	„fair"		„high"	
	averages	rank	averages	rank
Bucharest	2,10	3	4,28	4
Cluj	2,28	1	4,66	2
Iaşi	2,15	2	4,67	1
Timişoara	1,98	5	4,32	3
Tg. Mureş	1,52	6	3,25	6
Craiova	1,80	4	3,60	5

I consider that the linearity of the distributions within the „fair" categories, the similar degrees of variation, and the high level of correlation with the third quartile and with the averages of upper category consitute the main arguments for using the average of the „fair" category, or the third quartile – however more convenient, as measures of the central tendency of the distributions for purposes of comparison across the six medical schools. The main limit of this measure remains the discontinuity of the specter of the index.

However, the six distributions of the „fair" categories present notable differences, which are indicated by the different values of the skewness coefficient and by the differences between the average and the medians. Generally the distributions are tail heavy, with the exception of the distribution corresponding to the medical

school in Cluj, which appears to be balanced, and the distribution corresponding to the medical school in Tg. Mureş, whose negative coefficient indicated a top heavy distribution. The computations are presented in Table 8, below.

Table 8: The distributions within the „fair" category

	average	median	skewness	Coefficient of variation
Bucharest	2,10	1,93	0,24	0,22
Cluj	**2,28**	**2,33**	0,01	0,20
Iaşi	2,15	1,80	0,74	0,17
Timişoara	1,98	1,83	0,72	0,21
Tg. Mureş	**1,52**	**1,67**	-0,64	0,21
Craiova	1,80	1,67	1,20	0,17

Last but not least, for policy purposes it is useful to compare the average of the "fair" category with the averages of the entire population, as proposed in the official methodology (MECS 2015). I computed the average of the $h_{rational}$ indicator instead of the original version of the index, for issues of comparability. The rankings present notable similitudes and the correlation is strong: 0,88.

Table 9: Averages of the total population and of the ‚fair' category

	Average total population	rank	Average „fair"	rank
Bucharest	1,44	4	2,10	3
Cluj	1,69	1	2,28	1
Iaşi	1,54	2	2,15	2
Timişoara	1,41	5	1,98	5
Tg. Mureş	0,84	6	1,52	6
Craiova	1,50	3	1,80	4

Reverting to the decision of excluding the academics in the „outstanding" category from the total population, I mention that its impact on the overall results are limited: the correlation coefficient between the averages of the „fair" category „outstanding" excluded and, respectively, „oustanding" included is very high: 0,94. The results of computations for the six distributions partitioned in only three categories determined by the average and the average of those above the average, thus „outstanding" included are presented in Annex 1. Comparing the results in Annex 1 with those from the tables in this section, we can observ that the exclusion of the „outstanding" category has resulted in: (1) less variation of the proportions of the „fair" category in the distributions of the six medical schools is higher (0,17 compared to 0,11); (2) less variation inside the „fair" categories; (3) significantly higher

level of linearity of the „high" categories. The major advantage of the exclusion of the „outstanding" category consists in a higher level of comparability between the segments of the distributions. Moreover, the category „outstanding" can prove helpful in explaining the significance of the second order, succesive *h* index.

Outstanding contributors

In the search for the significance of the differences between the values of the second order, succesive *h* index, I compared it with the number of academics each of the six medical school has in the „outstanding" category. I opted to calculate the succesive *h* index using the rational alternative proposed by Ruane and Tol (2008), as I considered it to be a more exact measure, with higher discriminatory power, compared to the original version. I present the values of the index and the number of academics from each of the six medical schools in the „outstanding" category in Table 9, below.

Table 10: Comparison: second order $h_{rational}$ and „outstanding" category

	Second order $h_{rational}$	rank	„outstanding"	rank
Bucharest	9,8421	1	13	1
Cluj	8,8235	2	12	2
Iaşi	6,9231	4	5	5
Timişoara	7,9333	3	7	3
Tg. Mureş	4,7778	6	0	6
Craiova	6,9231	4	6	4

The two rankings are quasi-identical, and, moreover, the correlation between the values is very strong: 0,98. However, in scientometry the top 1% is not just an arbitrary percentage, the CWTS Leiden Ranking uses an indicator which counts the number of articles in the top 1% of a population of widely cited journal articles (CWTS 2015). Thus, the „outstanding" category can be regarded as an adaptation of CWTS Leiden Ranking's indicator to the population of academics in the six medical schools, where the attribute is the value of the individual h index.
In this context it is no surprise that the ranking produced by one of the measures of the central tendency and that produced by the second order h index are different, as the results presented in Table 10.

Table 11: Comparison: second order $h_{rational}$ and averages of the „fair" category

	Second order $h_{rational}$	rank	average "fair"	rank
Bucharest	9,8421	1	2,10	3
Cluj	8,8235	2	2,28	1
Iaşi	6,9231	4	2,15	2

	Second order $h_{rational}$	rank	average "fair"	rank
Timişoara	7,9333	3	1,98	5
Tg. Mureş	4,7778	6	1,52	6
Craiova	6,9231	4	1,80	4

Conclusions

The analysis of the distributions showed that the alternatives discussed in this article, the average proposed by the official methodology for higher education funding and the second order, successive *h* index, represent two different properties of the population of academics in the six faculties of medicine within the universities of medicine and pharmacy in Romania. The averages represent problematic measures of the central tendency, while the succesive *h* index is given by the academics with outstanding scores in the overall population and it is influenced by the size of the department, as pointed out by Proteasa, Păunescu, and Miroiu (2015).

The partitioning of the populations in the four categories: „low", „fair", „high" and „outstanding", has proven useful both for the identification of the central tendency, and for the explanation of the second order, succesiv *h* index. I consider this approach can be useful for the study of highly skewed distributions, similar to those analysed in this article. I will briefly enumerate some of its strengths: (1) the partitions are determined by attributes of the entire population (a unique cut-off point for the „outstanding" category for all the distribution), complemented with attributes of the individual distributions (the average and the second average); (2) it creates two quasi-continuous distributions, „fair" and „high", which are linear and where the central tendency measures have proven rather robust. I suggest further research on the matter, as I assume that the distributions in other fields of science in Romania share most of the characteristics of those studied in this article.

Last but not least, I conclude that the rational variant of the *h* index proposed by Anania and Caruso (2013) has proven useful for the study of a population with low values of the original *h* index and with numerous academics with the same (low) value of the index, such as the academics from the six medical schools under scrutiny in this article.

Acknowledgments: I am indebted to Adrian Miroiu, Mihai Păunescu and Andra Roescu for the useful discussions we had in various stages of this research project.

Annex 1

	Average	Second average	Proportion in the total population „fair"	Coefficient of variation „fair"	average „fair"	RSQ high	RSQ fair
Bucharest	1,44	3,06	31,11%	0,23	2,16	0,55	0,96
Cluj	1,69	3,59	26,85%	0,20	2,43	0,53	0,95
Iaşi	1,54	2,95	33,27%	0,20	2,20	0,42	0,95
Timişoara	1,41	3,08	29,31%	0,22	2,04	0,60	0,90
Tg. Mureş	0,84	2,26	21,10%	0,21	1,52	0,65	0,88
Craiova	1,50	2,91	37,14%	0,24	2,09	0,84	0,93

References

Albarrán, P., Crespo, J., Ortuño, I., and Ruiz-Castillo, J. (2011), "The Skewness of Science in 219 Sub-Fields and a Number of Aggregates", *Scientometrics,* 88(2), 385–97.

Albarrán, P., Perianes-Rodríguez, A., and Ruiz-Castillo, J. (2015), "Differences in Citation Impact across Countries", *Journal of the Association for Information Science and Technology,* 66(3), 512–25.

Anania, G., and Caruso, A. (2013), "Two Simple New Bibliometric Indexes to Better Evaluate Research in Disciplines Where Publications Typically Receive Less Citations", *Scientometrics* 96(2), 617–31.

Anderson, T., Hankin, R., and Killworth, P. (2008), "Beyond the Durfee Square: Enhancing the H-Index to Score Total Publication Output", *Scientometrics* 76(3), 577–88.

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

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), 16569–72.

MECS (2015), 454 Monitorul Oficial *ORDIN Privind Aprobarea Metodologiei de Alocare a Fondurilor Bugetare Pentru Finanțarea de Bază Şi Finanțarea Suplimentară a*

Instituţiilor de Învăţământ Superior de Stat Din România, Pentru Anul 2015. Ministerul Educaţiei şi Cercetării Ştiinţifice.

MECTS (2011), *ORDIN Privind Aprobarea Metodologiei de Prelucrare a Datelor Şi Informaţiilor Colectate În Vederea Realizării Evaluării Primare a Universităţilor Şi a Evaluării Programelor de Studii Universitare Cu Scopul Clasificării Universităţilor Şi Ierarhizării Prog.* Ministerul Educaţiei, Cercetării, Tineretului şi Sportului.

Miroiu, A., Păunescu, M. and Vîiu, G. (2015), "Ranking Romanian Academic Departments in Three Fields of Study Using the G -Index." *Quality in Higher Education* 8322 (December), 1–24.

Miroiu, A. and Vlăsceanu, L. (2012), "Relating Quality and Funding: The Romanian Case." In *European Higher Education at the Crossroads*, eds. Adrian Curaj, Peter Scott, Lazăr Vlăsceanu, and Lesley Wilson. Springer Netherlands, 791–807.

MS. (2009), *Ordin Nr. 1509/2008 Privind Aprobarea Nomenclatorului de Specialităţi Medicale, Medico-Dentare Şi Farmaceutice Pentru Reţeaua de Asistenţă Medicală*. Ministerul Sănătăţii.

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

Perianes-Rodriguez, A., and Ruiz-Castillo, J. (2015), "University Citation Distributions.", 129–38.

Proteasa, V., Păunescu, M. and Miroiu, A. (2015), "The 'black-Box' of Institutional Scores: Analyzing the Distribution of the Values of the H and G Indexes in Medicine Schools in Romania", *Journal of Social Research & Policy,* 6(1).

Ruane, F., and Tol, R. (2008), "Rational (Successive) H-Indices: An Application to Economics in the Republic of Ireland", *Scientometrics,* 75(2), 395–405.

Ruiz-Castillo, J., and Costas, R. (2014), "The Skewness of Scientific Productivity", *Journal of Informetrics,* 8(4), 917–34.

Schubert, A., Glänzel W. and Braun, T. (1987), "A New Methodology for Ranking Scientific Institutions", *Scientometrics,* 12, 267–92.

Schubert, A. (2007) "Successive H-Indices", *Scientometrics,* 70(1), 201–5.

Seglen, P. O. (1992), "The Skewness of Science." *Journal of the American Society for Information Science,* 43(9), 628–38.

Ţeca, Mircea. (2011). "Viziune de Ansamblu Asupra Modelului Matematic de Construcţie Şi Utilizare a Indicatorilor Relativi de Calitate În Finanţarea Învăţământului Superior Utilizat În Perioada 2003 – 2011", *Quality Assurance Review for Higher Education,* 3(1), 81–92.

Vîiu, G., and Miroiu, A. (2013). "Evaluarea Cercetării Universitare din România . Abordări Metodologice Alternative", *Revista de politica ştiinţei şi scientometrie,* 2(2), 1–20.

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

Zulean, M., Ioniţă, I. and Vîiu, G. (2014). *Raport de Evaluare a Guvernanţei Sistemului Public de Cercetare, Dezvoltare Şi Inovare Din România 2007 – 2013*.