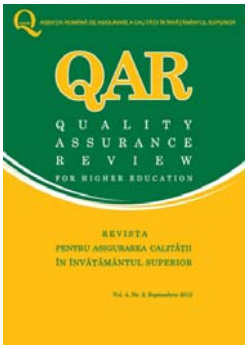
AGENŢIA ROMÂNĂ DE ASIGURARE A CALITĂŢII ÎN ÎNVĂŢĂMÂNTUL SUPERIOR
THE ROMANIAN AGENCY FOR QUALITY ASSURANCE IN HIGHER EDUCATION
*Membră cu drepturi depline în Asociaţia Europeană pentru Asigurarea Calităţii în Învăţământul Superior - **ENQA***
*Înscrisă în Registrul European pentru Asigurarea Calităţii în Învăţământul Superior - **EQAR***



Quality Assurance Review
For Higher Education

Ranking Political Science Departments: The Case of Romania

Gabriel-Alexandru Vîiu, Mirela Vlăsceanu, Adrian Miroiu

Quality Assurance Review for Higher Education, Vol. 4, Nr. 2, 2012, p. 79-97

Publicat de: Consiliul Agenţiei Române de Asigurare a Calităţii în Învăţământul Superior - ARACIS
Locul publicării: Bucureşti, România
Tipul publicaţiei: tipărit, online
Adresa: Bd. Schitu Măgureanu, nr. 1, Sector 1, Bucureşti, cod poştal 050025
Telefon: +40 21 206 76 00; Fax: +40 21 312 71 35
E-mail: qar@aracis.ro
Pagină electronică: http://www.aracis.ro/en/publicatii/qar-magazine/numarul-curent/

Revista Quality Assurance Review for Higher Education este editată din fondurile proprii ale ARACIS şi asigură sustenabilitatea proiectului "Asigurarea calităţii în învăţământul superior din România în context european. Dezvoltarea managementului calităţii academice la nivel de sistem şi instituţional", Contract POSDRU/2/1.2/S/1, cod proiect 3933.

Pentru a reproduce un text din revistă este necesar acordul scris al colegiului de redacţie al revistei Quality Assurance Review for Higher Education.
Răspunderea pentru textele scrise aparţine autorilor.
Conţinutul acestui material nu reprezintă în mod obligatoriu poziţia oficială a ARACIS.

Ranking Political Science Departments: The Case of Romania

Gabriel-Alexandru Vîiu * *gabriel.alexandru.viiu@snspa.ro*
Mirela Vlăsceanu* *mirela.vlasceanu@snspa.ro*
Adrian Miroiu* *admiroiu@snspa.ro*

Abstract

The ranking of the study programs offered by the Romanian universities became a widely debated topic in the past two years in the context of the recent policy changes that followed the enforcement of the new Law of Education. In this paper we propose a simplified ranking methodology and use it to evaluate the political science and international relations departments in the Romanian universities. We also compare our results with the official ranking produced in 2011 by the Ministry of Education.

Keywords: *ranking, scientometrics, political science departments, h-index, g-index*

Rezumat

Ierarhizarea programelor de studii oferite de către universităţile din România a devenit un subiect larg dezbătut în ultimii doi ani în contextul schimbărilor recente de politici ce au urmat implementării noii Legi a Educaţiei. În această lucrare propunem o metodologie simplificată de ierarhizare şi o folosim pentru a evalua departamentele de ştiinţe politice şi relaţii internaţionale din universităţile româneşti. Rezultatele obţinute sunt comparate cu ierarhizarea oficială făcută în 2011 de către Ministerul Educaţiei.

Cuvinte cheie: *ierarhizare, scientometrie, departamente de ştiinţe politice, indice h, indice g*

Introduction: The growing scope of university rankings

Like many other actors from the public sector, in recent years universities have become subject to a growing process of assessment, especially regarding their performance and concrete measurable output (de Bruijn, 2007: 1). Because higher education is also subject to broader attempts at assessing organizational effectiveness (Shin, 2011: 19) more and more instruments are being developed worldwide, both by governments as well as by private entities, in order to evaluate individual universities and compare them. Well-known results of this process are the international rankings of universities which became very popular in the last few decades, partly because they serve as devices to construct legitimacy (Wedlin, 2011: 200), partly because they serve the need for more information. At present, university rankings are also used in many national systems to provide data that inform student choices, and in some countries they also guide the allocation of public funds (Marginson & der Wende, 2007: 319). Consequently, university rankings, be they national or international, have acquired a growing importance in the contemporary higher education landscape.

Although international university rankings certainly have great prominence in an increasingly globalized higher education landscape, they are not the study object of our article. We focus on the rankings created as a result of government policy and intervention. It is our belief that such rankings are a worthwhile research area because, unlike international rankings that may or may not impact the day-to-day operation of universities, the rankings created as a result of government policy have more immediate consequences for the universities under assessment.

Two other characteristics of our research are also essential. First, we narrow our investigation to the Romanian case: we analyze programs offered by both public and private Romanian universities, but do not make comparisons with the higher education programs offered in other

* All authors are affiliated to the National School of Political and Administrative Studies, Bucharest.

countries. Secondly, we do not investigate the ranking of universities as such, but the ranking of the study programs offered by academic departments in the universities. In particular, we take as an example the programs offered by the political science and international relations departments.

The outline of our article is as follows. In section 1 we review some past and present international practices in this field. A special attention will be given to debates concerning the ranking of political science departments in the UK. Section 2 describes our methodology for ranking the political science departments in Romania and compares it with the methodology used in 2011 by the MoE to produce the now official ranking. In section 3 we present and discuss the results. The focus of the analysis is represented by aggregate results at the department level, but we also include some results on the research performance of the individual members of the departments. Section 4 concludes.

Ranking political science departments: a review of practices and critical issues

Several authors have embarked on a quest to evaluate, rank or simply compare the research groups or departments active in various fields of study: Sinha and Macri, for example, ranked Australian economic departments (Sinha & Macri, 2002), Aaltojärvi et al. compared Nordic sociology departments (Aaltojärvi et al, 2008) and van Raan studied chemistry research groups (van Raan, 2006). Numerous other examples of evaluations made at this intermediate level (that is between individuals working in universities and universities taken as a whole) can also be found in scholarly journals. Our own research and ranking of political science and international relations departments can also be circumscribed to the intermediate level of analysis, as will be described at length in the methodological section of the paper.

One of the earliest efforts to assess and compare scholarly performance in the field of political science is made by Morgan and Fitzgerald in 1977. In what is now a 35 years old article, the two authors developed an index of productivity based on (weighted) journal publications and used it to compare political science departments in the United States (Morgan & Fitzgerald, 1977: 345). The methodology employed by the two authors relies on articles published by members of political science departments in what the authors considered "the principal journal" of the discipline (namely the *American Political Science Review*) and in four other regional journals during the years 1964-1973.

The methodology used by the two authors is entirely bibliometric in nature: the total number of publications in the 5 selected journals is the sole indicator used to assess productivity and rank the departments under study. In this sense, the article is perhaps one of the first that tried to use objective measures of assessment (at least as far as political science is concerned), in sharp contrast to the traditional peer-review approach. Indeed, we find three decades later that "research evaluation has evolved from the traditional and relatively simple peer-review process to highly sophisticated benchmarking procedures involving ever-growing numbers of quality criteria and performance standards, as well as immense systems for counting almost everything" (Coryn et al, 2007: 439). Here we clearly meet a fundamental tension that pervades discussions about rankings in general: the use of peer-review (reputational) procedures, versus bibliometric methods. We shall return to this discussion in subsequent paragraphs.

A second noteworthy feature of the Morgan and Fitzgerald approach is the awareness of the bias that the bibliometric methodology employed can produce. They note in particular the potential bias of their index in favor of departments specialized in American politics (in contrast to departments that have greater strength in comparative and international politics). Bias and unintended consequences still constitute a major topic of discussion in the debates surrounding all rankings. As the authors of the Shanghai Academic Ranking of World Universities methodology concede, "any ranking is controversial and no ranking is absolutely objective" (Liu & Cheng, 2005: 135).

A third aspect of the Morgan and Fitzgerald endeavor to rank political science departments that is worth mentioning is the general concern with scientific productivity and its impact. Although not

explicitly stated in any part of their article, they work with at least the following two assumptions: a) bibliometric indicators are a powerful alternative to peer-review in assessing scientific productivity; b) the number of published articles (in certain journals) is a good indicator of a political science department's impact. Whether or not these two assumptions are valid and defensible is still the fundamental question up for debate in the recent literature on rankings and research evaluation. The substantial change, however, is that with the advent of modern information technology, new tools have become available to investigate these issues and new theoretical instruments have been created to measure scientific productivity. However, we reserve the discussion concerning these tools for the next sections of the paper.

As may be evident by now, the theoretical effort to understand the possible attempts of ranking political science departments must begin with acknowledging that this is part of the broader discussion about the use of the two main approaches to assess scientific research in general: on one hand the peer-review system, and on the other hand the citation analysis (bibliometric approach) which has developed into the wider field of scientometrics (Thelwall, 2008: 606). All of this, however, must also be viewed against the background of what could be best described as "an increasing emphasis of public policy on institutional auditing and surveillance" (Bratti et al, 2004: 475), coupled with the need for accountability of public entities.

The two main ways in which the peer-review system is used are: 1) the selection of manuscripts for publication in journals; and 2) the selection of fellowship and grant applications (Bornmann, 2011: 145). The underlying logic of the peer review system is that professionals in a certain domain naturally (and uniquely) hold the specialized knowledge necessary to judge and evaluate research within their own field. Since evaluation can only be done by a limited number of highly knowledgeable individuals, the peer-review system is almost by definition a qualitative approach, most useful for purposes of evaluating individual items of scientific output (journal articles, books, etc.). The scientometric approach, on the other hand, and especially recent indicators such as the h-index, g-index and their variants, assesses the merits of researchers in a quantitative, general aggregated form: not just one book or one article, but, ideally, all of these and, more importantly, their overall influence on the scientific community as measured by the number of citations they receive. The main divide between the two approaches is shaped around the idea of subjective versus objective evaluation: whereas the peer-review process is sometimes criticized for its subjective, non-verifiable nature, bibliometric indicators are said to possess a greater degree of objectivity, although several methodological issues demonstrate they too may have inherent biases that must be corrected with specific mathematical and statistical tools.

Two further instances detailing the elements of the peer-review versus scientometric debate in the literature on the ranking of political science departments are also very instructive: the debate surrounding the well-established UK Research Assessment Exercise (RAE)[2], and the system of ranking political science departments proposed in 2004 by Simon Hix. Discussion of these two cases will help illuminate additional issues surrounding the possibility of ranking political science departments, as well as illustrate some of the main limitations associated with using both peer-review and scientometrics as a base for assessing research performance, not only in the field of political science, but also in the larger circle of social sciences. We begin with the UK Research Assessment Exercise, with special emphasis on its consequences for political science, and then proceed to a discussion of Hix's proposal for a global ranking of political science departments.

The UK Research Assessment Exercise is, to date, one of the most comprehensive efforts undertaken at a national level in order to measure, assess and rank higher education institutions according to their scientific output and overall quality. Its declared purpose is "to enable the higher education funding bodies to distribute public funds for research selectively on the basis of quality"[3].

2 The RAE is currently being replaced by the modified initiative called Research Excellence Framework (REF); since for all intents and purposes this refers to the same national research assessment initiative and because our discussion focuses on both exercises, we retain both the RAE and REF short hands.

3 "What is the RAE 2001?", <http://www.rae.ac.uk/2001/AboutUs/>, last accessed on 15.10.2012.

As allocating future funds for research constitutes the very essence of the RAE, the methodology, procedures and criteria used during the assessment are of paramount importance. It is precisely these procedures, criteria and methodologies that constitute an object of scrutiny for several authors which point out limitations and possible unintended consequences.

RAE is primarily based on a system of peer-review evaluation. The most recent document outlining the RAE's procedures states that "as with previous RAEs, the assessment process is based on expert review" (*REF 2014 Panel criteria and working methods*, 2012: 6[4]). At present, there are three distinct criteria taken into account in the evaluation process: outputs[5] (which refer to the quality of submitted research outputs of each unit of assessment), impact (which refers to impact of research on economy, society and culture) and environment (which refers to the vitality and sustainability of the research environment)[6].The final assessment is made by experts working in panels responsible for different research disciplines. Each panel has a limited number of members who are responsible for the evaluation of an entire discipline or sub-discipline. For example, according to the public information available on the 2008 RAE, the Politics and International Studies panel was made up of 15 members; the same panel for the 2014 REF is composed of 19 members.

Despite the fact that this layout is clearly indicative of peer-review as being the preferred method employed for evaluation in the RAE, further analysis conducted by Linda Butler and Ian MacAllister shows that, as far as political science is concerned, "citations are the most important predictor of the RAE outcome" (Butler & MacAllister, 2009: 3). Another more controversial finding of the two authors is that "the second important predictor of outcome for a department is having a member on the RAE panel", something later attributed to the privileged knowledge that panel members possess (Butler & MacAllister, 2009: 9). These results help the two build a case against the use of peer review in ranking political science departments.

The main line or reasoning in their article is that given the fact that citations are the most important predictor of the RAE outcome, and also given the fact that peer-review is an inherently subjective process, there is a dire need to implement a system of evaluation based on more objective parameters which could guarantee greater transparency. In order to support this view they elaborate a statistical modeling of the 2001 RAE in the field of political science. The model excludes the variable measuring RAE panel membership, but takes into account more "objective" measures (citations, department size, research income, national academy membership) and so is closer to a scientometric approach. Because the results of their statistical modeling are fairly close to the actual results obtained in the RAE, they suggest that such a metrics-based system of evaluation should be preferred to the peer-review method, especially since it brings the additional advantage of lower costs.

Andrew Russell, on the other hand, argues that "the critical feature for political science is that there is little consensus between sub-fields" (Russell, 2009: 65). He argues that given the bias inherent in citation index systems like Web of Science and Scopus, some form or peer-review is mandatory in order to correct potentially distorted results of bibliometric indicators. Russell also points to a number of flaws in the arguments presented by Butler and MacAlister, most notably the fact that their results in favor of citation measurements do not eschew the problem of non-journal publications and, furthermore, are based on retrospective data that simply could not have been available at the time the 2001 RAE was conducted (Russell, 2009: 67).

Russell disagrees with Butler and MacAlister's conclusion concerning the impact of panel membership on the final results of the research evaluation exercise. The implication of their argument was that panel members tend to favor the departments from their own university and this therefore has an adverse effect on the final ranking. Russell's criticism is that such a conclusion comes from neglecting the issue of multicollinearity (Russell, 2009: 68): in his view, having a member of the

4 Available at <http://www.ref.ac.uk/media/ref/content/pub/panelcriteriaandworkingmethods/01_12.pdf>, last accessed on 15.10.2012.

5 For each member of staff a maximum of 4 distinct outputs (publications) may be submitted for the period under assessment; the current REF exercise will cover the period from January 2008 to December 2013.

6 The three criteria have a weighting of 65%, 20% and 15%. It is worth stressing the fact that impact as understood in the REF document has no connection with citation impact or journal impact factors.

department in the evaluation panel is more likely a result of department features that make it attractive to good scholars in the first place, prior to their joining the evaluation panel. The implication is that it is quite natural that the panel members in charge of evaluation are selected (even disproportionately so) from departments having excellent performance. These departments then receive good rankings not because they have a member in the evaluating panel, but because they are top performers from the outset. Russell also questions the idea that citation-based metrics would themselves be immune to the possibility of strategic behavior and shows that, despite the purported notion of lower costs, even metrics-based systems have hidden costs associated with data accessibility and processing.

In a subsequent effort to make a case in favor of objective measures (citations in particular) Butler and MacAlister compare the peer-review driven results of the 2001 RAE in political science and chemistry and again find that "a metrics approach to measuring research quality would produce results which are close to those that come from peer evaluation of research, particularly in the natural sciences" (Butler & MacAlister, 2011: 54). In this way they reassert their previous case for objective measurements, but acknowledge the fact that objective indicators are more adequate for STEM (Science-Technology-Engineering-Mathematics) subjects such as chemistry, than for Humanities-Arts-Social Sciences (HASS) such as political science. Another interesting claim they make is that in both disciplines studied in their research, "a measure based on esteem (membership in a national academy) has no predictive power and is unnecessary for any metrics-based system" (Butler & MacAlister, 2011: 55).

In spite of this and other pleas for so-called objective measures, however, it seems that the use of "objective" measures in the UK research exercise, citations in particular, will not become a dominant practice in the near future (certainly not in the current, ongoing exercise). In fact, the *REF 2014 Panel criteria and working methods* document clearly states that "given the limited role of citation data in the assessment, the funding bodies do not sanction or recommend that higher education institutions rely on citation information to inform the selection of staff or outputs for inclusion in their submissions" (p. 8). This does not mean, however, that citations have no place in the current REF exercise; for certain sub-panels (Economics and Econometrics for example) a limited use of citation data is sanctioned, but only as supplementary information, not as a primary tool. For most other sub-panels, including the Politics and International Studies sub-panel, citation data or any other bibliometric indicators cannot be used. In this sense, it seems that for the time being, the peer-review process is still the firmly established mechanism for quality and research performance evaluation in the UK, despite past criticism and despite proposals for other evaluation and ranking methodologies.

Simon Hix initiated what is probably the most ambitious project in ranking political science and international relations departments. He proposes nothing less than a global ranking of top political science departments. Hix's ranking method is based on quantity and impact of publications, starting from the premise that peer evaluation is "subjective, biased toward established institutions and costly in terms of time and money" (Hix, 2004: 293). Unlike Morgan and Fitzgerald (but with the same logic in mind) Hix identifies not one or five, but 63 "main journals in political science"[7] and studies articles published in a ten-year window (1993-2002) in order to assess the impact of political science departments around the world. Only main articles and research notes were included in the analysis; multiple-authored articles were counted by taking into account the institutional affiliation of each author (for example a paper with two authors belonging to different institutions was counted as 0.5 for each institution) and articles were also weighted in accordance with a journal impact score that Hix himself devises using statistical tools and which he claims is highly correlated with SSCI impact scores (Hix, 2004: 300), at least for the 54 SSCI indexed journals from his selected 63.

Using the data about publications and department membership, Hix constructs four intermediate political science department rankings and an overall rank calculated as the average

7 Most (54) were drawn from Thomson-Reuter's *Social Science Citation Index (SSCI)*, the other 9 were added by Hix on grounds that they constitute national political science associations journals or because they constitute major sub-field journals not indexed in SSCI

position of a department on the other four rankings. These four intermediate rankings refer to quantity of published articles (ranking 1), impact (ranking 2), a quantity/faculty size ratio (ranking 3) and an impact/faculty size ratio (ranking 4). The first ranking refers to the total number of articles in the 63 journals authored by scholars from a particular institution in a five-year period. The second ranking is basically ranking 1 multiplied with the journal impact score created by Hix. The third ranking is a ratio between the first rankings and the faculty size of the political science department of an institution. Ranking 4 is a ratio between the second ranking and the faculty size of the political science department. As far as the size of the department is concerned, only "full-time staff with a rank of full, associate or assistant professor (or equivalent)" was taken into account; information for each university was obtained from the public web-site of the department under study but, in the case of British universities, data from the 2001 RAE were used. Based on the data collected, Hix presents a global ranking of the top 200 political science departments from around the world.

Although fairly rigorously constructed, Hix's proposed method was subject to much criticism shortly after becoming public. One of the first aspects that elicited critique was the lack of coherence in the staff counting process; in particular, the fact that data from the RAE were used in the case of UK universities produced two major distortions (Bull & Espindola, 2005: 27-29): it put non-British universities at a disadvantage (because, during the RAE, British departments tend to report fewer staff, omitting their weaker researchers) and, in the case of some UK universities, it ignored the contribution of political scientists who were not submitted to the Politics and International Relations unit of assessment (from which Hix drew his data), but to other panels that Hix did not take into account. As Bull and Espindola point out, Hix's ranking "is not a ranking of political science 'departments' but a ranking of political science output produced by *institutions*" (Bull & Espindola, 2005: 29, emphasis in original text). Moreover, a single method of counting staff should be used and, in their view, website counts may be a fairer alternative despite their flaws. Hix subsequently accepted all these lines of criticism against his method and also made a few suggestions on how the initial methodology might be improved (Hix, 2005).

Dale and Goldfinch also present a critical perspective on Hix's methodology, pointing in particular to at least three major faults (Dale & Goldfinch, 2005: 426-27): a misleading picture of the political science discipline created by defining only a limited number of journals as political science journals; a North American bias that pervades Hix's methodology through overrepresentation of North American journals in Hix's selected list; insufficient comparability between the departments ranked due to their different publishing culture which may not be necessarily focused on articles, but on books, which are not taken into account in Hix's methodology due to the technical limitations of the SSCI database he used in his analysis.

These lines of criticism are also supported by further comments from Ronald Erne who also highlights several other flaws in Hix's methodology. First, Erne points out the inflated importance of a single indicator (the number of articles published in the 63 selected journals) on all intermediate rankings produced by Hix and, implicitly, on the final result (Erne, 2007: 308). Secondly, Erne suggests that the rankings produced by Hix are biased in favor of big institutions, something particularly evident in the case of rankings 1 (quantity) and 2 (impact) which Erne sees as biased proxies for university size and material resources. Erne raises several other issues, but his main criticism is the arbitrary nature of the ranking procedure employed by Hix, coupled with the general concern that "quantitative performance metrics will never be able to measure all aspects of multifaceted academic work" (Erne, 2007: 312).

A methodology for ranking political science departments in Romania

Regardless of its multifaceted nature, academic work is nonetheless increasingly assessed by appealing to indicators intended to measure the output of academics at the individual level and the overall achievements of university departments or even institutions taken as a whole. This is particularly true in the case of Romania where the first comprehensive national evaluation of the performance of higher

education institutions was completed in 2011. The evaluation process was carried out by the Ministry of Education, with the technical support of the European University Association[8]. The purpose of the evaluation process was two-fold: first, to establish a classification of all universities with respect to their research and teaching capacity into one of three broad categories and, secondly, to offer a ranking of all study programs existing in Romanian universities into five ranking classes (with A representing the top performing study programs and E the opposite). In this manner, the first ranking of (inter alia) Romanian political science programs was produced. The results of the ranking were included in the new Methodology for the financing of the public universities proposed by the National Council for The Financing of Higher Education (CNFIS). According to it, the amount of funds allocated to a university for each of the students enrolled in a given program was correlated with the ranking class in which the university's programs were included[9].

It should be mentioned that during the evaluation process carried out by the MoE, data pertaining to all international relations study programs were submitted within the domain of political science, thus yielding a single common ranking for both disciplines. In order to have sufficient comparability to the national ranking, we too focused on both disciplines and aggregated data in a corresponding manner.

One more fact should be explicitly mentioned. According to the official MoE ranking, all the study programs at the bachelor, master and doctorate level offered by a university in a domain (like, e.g., political science), were placed in the same class. However, universities develop a large variety of internal mechanisms to manage these study programs. In some universities all the programs in the field of political science are offered by one department (included in a single faculty), while in other universities the political science programs are associated with more departments and more faculties[10]. Similarly, our methodology does not primarily concern academic departments, conceived simply as formal administrative structures in universities. Nonetheless, since both the official ranking and the alternative one we propose are heavily supported by data concerning the teaching and research staff reported by the universities, we decided to speak about rankings of the academic departments, understood as generic structures in the universities which are mainly responsible for the study programs offered in a given field like political science.

Our study seeks to compare the results of the national ranking of study programs done in 2011 for the field of Political Science and International Relations with results obtained by using a simplified methodology[11]. Our main effort was driven by the idea that a few, but synthetic and pertinent indicators suffice for purposes of comparison and may yield results fairly similar to the official ranking. The official methodology evaluated four dimensions (research, teaching, relation to external environment and institutional capacity[12]) and used nearly 70 individual performance indicators for a retrospective period of 5 years. This is a substantial leap even in contrast with the methodology used by the Romanian Agency for Quality Assurance in Higher Education for provisional authorization and accreditation of study programs, a methodology which only uses 43 indicators[13].

However, the methodology we propose assesses the two major components, education and research, with a total of only 9 indicators[14]. (Note that indicators grouped under the criterion

8 The main documents detailing this process are Government Decision no 789/ 03.08.2011 concerning the evaluation methodology for the purpose of university classification and study program ranking (available at <http://chestionar.uefiscdi.ro/docs/hg789.pdf>, last accessed on 15.10.2012) and, in particular, MoEOrder no 5212/ 26.08.2012 concerning the methodology used for processing the data collected from the universities (available at <http://www.edu.ro/index.php/articles/16066> last accessed on 15.10.2012). This latter document provides the detailed indicators, weights and formulas used by the Ministry in order to complete the classification and the ranking process. The official document presenting the final rankings for all domains can also be found at <http://chestionar.uefiscdi.ro/docs/programe_de_studii.pdf>, last accessed at 15.10.2012.

9 See MoE Order no 3998/05.05.2012.

10 In the case of some universities a very difficult task is to identify the relevant formal administrative structures and the proper staff involved in political science programs.

11 The relevant data used to construct the ranking presented in the paper are available at www.csap.ro.

12 The associated weights that were used in the official ranking process for the political science domain are as follows: research - 0.50; teaching - 0.25; relation to external environment - 0.20; institutional capacity - 0.05.

13 See Romanian Agency for Quality Assurance in Higher Education "Methodology for External Evaluation", <http://www.aracis.ro/fileadmin/ARACIS/Proceduri/Methodology_for_External_Evaluation.pdf>, last accessed 15.10.2012

14 We also considered a peer-review component aiming to evaluate the reputation of different departments among the

"education" include aspects of the relation of the universities with the environment, as well as a measure of their institutional capacity.) **Table 1** summarizes the criteria and indicators we used in our ranking process, together with the individual and global weights attached to each.

Table 1 Criteria, indicators and associated weights used in our ranking process

Criteria	Indicators	Individual indicator weights	Criteria global weights
Research	Successive g-index of department	0.77	0.50
	Research projects / department staff ratio	0.10	
	PhD students / total students ratio	0.13	
Education	Department size index	0.40	0.50
	Staff / student ratio	0.20	
	Graduate / undergraduate student ratio	0.10	
	ERASMUS students / total students ratio	0.10	
	ARACIS confidence score for study programs	0.10	
	ARACIS confidence score for the university	0.10	

Research, the first component of our methodology, is evaluated by appealing to three indicators: the successive g-index of the department, the ratio of research projects to the department staff and the ratio of PhD students to the total number of students. Although the latter two are fairly intuitive and straightforward, the first requires some explanation. As we have seen in our general discussion regarding rankings, the bibliometric component they use refers to citations of published academic work, usually articles from journals indexed in major international databases like Thompson-Reuters' Web of Science or Elsevier's Scopus. Indeed, until recently, the numbers of publications, citations and various impact factors have been the standard bibliometric instruments used for the assessment of research output[15]. In recent years, however, classical bibliometric indicators such as these are no longer the sole criteria for assessing scientific productivity and performance. In particular, after the introduction of scientific impact indices like the Hirsch h-index in 2005, it too has become a widely used bibliometric indicator and has become so popular that it is currently included in the online platforms of both Thompson-Reuters and Elsevier.

The h-index was introduced by Jorge Hirsch with a fairly simple underlying logic: "A scientist has index h if h of his or her Np papers have at least h citations each and the other ($Np– h$) papers have $\leq h$ citations each" (Hirsch, 2005: 16569). A more intuitive way of understanding this would be to have in mind the list of papers by an author arranged in decreasing order of their number of received citations and see the h-index as the cut-off value where the n-th paper has at least n citations[16]. Because of its intuitive nature and simplicity, the h-index immediately triggered a wave

academics in the field. A questionnaire was developed and data began to be collected, but the process has not yet been concluded. In order not to have results biased due to a statistically inadequate sample size and through disproportionate representation of some departments that did answer, we decided not to include this reputational component in our analysis.

15 While the methodology used in the official ranking in Romania did not directly use citations in the measurement of the scientific research component, it did however incorporate the relative influence score of articles, the number of publications in journals indexed by ISI Web of Knowledge (with and without impact factor), the number of publications in journals indexed in international databases as well as six other numeric indicators dealing with other types of publications.

16 For example, if an author has 4 papers, each with 20, 10, 5 and 2 citations, her h-index would be 3 because her third

of discussion regarding its possible application beyond the level of individuals, most notably for journals (Braun et al, 2006), but also for institutions (Prathap, 2006).

Although several of its features attracted praise, some of its properties were considered less desirable. In particular, its insensitivity to very highly cited work (something Hirsch himself had acknowledged) prompted some authors to introduce other scientific impact indices. One example is the g-index, proposed by Leo Egghe. The g-index of a researcher is defined as "the highest number g of papers that together received g^2 or more citations" (Egghe, 2006: 8). Egghe's main argument in favor of this new index was its higher sensitivity and, therefore, greater discriminative power[17].

We can now define the successive g-index for a department, as used in our ranking process: it is simply a g-index of the second order, calculated on the bases of the g-index (of the first-order) of each of the staff members of a department[18]. In other words, in order to calculate it we first determine each individual author's g-index and then calculate the successive (second order) g-index based on the initial results, treating first-order g-indices of the department members in the same manner as the number of citations is treated in the calculation of the first- order g-index[19]. Just like the original index introduced by Egghe for individual researcher, the successive g-index (for departments) also has greater discriminative power than other comparable alternatives.

Moreover, in order to have a maximum degree of differentiation, we appealed to an even more complex indicator, the *rational* successive g-index, as defined by Richard Tol (Tol, 2008: 151)[20]. We decided to adopt it for evaluating the research output of departments because it has the advantage of allowing differentiation between departments that would otherwise have an identical score based on the simple successive g-index[21] (which is a natural number)[22].

In calculating the first order g-indices of the political science department members we used Anne Harzing's *Publish or Perish* program (Harzing, 2007) which has the Google Scholar database as a source for citations measurement. This database was also comparatively used by Tol & Ruane who reported good correlation between its results and the ones obtained when appealing to other major platforms like Web of Science, Scopus and EconLit (Ruane & Tol, 2007: 310). Google Scholar was also used by Aaltojärvi et al. (2008) in their analysis of sociology departments and the *Publish or Perish* interface to Google Scholar has itself already been used in a number of studies dealing with evaluation of journals (Hodge & Lacasse, 2011) or the evaluation of academic departments (Altanopouloua et al, 2012) that found it to be a particularly useful tool due to screening options

most cited paper has 5 (>3) citations, but it would not be 4 because her forth paper only has 2 (<3) citations; in other words, the third paper has at least 3 citations (in this case more), but the fourth does not have at least 4 citations.

17 This can readily be seen with reference to our previous example: the same hypothetical author with an h-index of 3 would have a g-index of 6 because 37 (the sum of the citations for all 4 papers) is greater than 36.

18 The idea of a successive g-index was first described and used by Richard Tol, "A Rational, Successive G-Index Applied to Economics Departments in Ireland"*, Journal of Informetrics,* 2/2 (2008), 149-155.

19 Suppose for example that a hypothetical department consists in only 5 members; if inspection of each reveals g-indices of 9, 4, 2, 2 and 2, then the successive g-index of this department would be 4 because the cumulative score of the first 4 members (17) is greater than 4^2; it could not be 5 because the cumulative score of all the 5 members (19) is not greater than or equal to 5^2.

20 In formal terms, Tol defines the successive g-index as a solution to t$\boldsymbol{max}_{g_1} \sum_{i=1}^{g_1} g^i \geq g_1^2$, where g^i refers to the researchers in a department (denoted by their first order g-index); Tol then defines the rational successive g-index (g_1^{Δ}) as $\boldsymbol{g_1^{\Delta} = g_1 + \frac{\sum_{i=1}^{g_1+1} g^i - g_1^2}{(g_1+1)^2 - g_1^2}}$, which, by definition, is $\geq \boldsymbol{g_1}$, but $< \boldsymbol{g_1 + 1}$.

21 Suppose we want to compare the hypothetical 5 member department described above (with members having g-indices of 9, 4, 2, 2 and 2) with another 5 member department whose members have g-indices of 9, 9, 2, 2 and 2. Comparison based on the simple successive g-index would yield a tie, as both departments have a g-index value of 4, although there is clearly a difference between them. Comparison with the aid of the rational successive g-index, however, yields a more differentiating outcome: the first department has a g value of 4.33, but the second has 4.88, therefore reflecting the greater contribution made by one of its researchers.

22 An alternative we considered in the earlier stages of our research was a composite *hg*-index which would be calculated as the maximum number such that there are at least h individuals with a g-index of at least g in a given department. Such a measure would, however, prove fairly insensitive to high g-indices and would therefore not discriminate much between departments with medium g-indices and those where some staff members have higher than average g scores.For example, in our previous illustration, the composite *hg*-index of the hypothetical 5 member departments is 2, completely ignoring the high g-index of the first members. Note also that the *hg*-index we originally considered is different from the *hg*-index proposed by Alonso et al. ("HG-Index: A New Index to Characterize the Scientific Output of Researchers Based on the H- and G-Indices", *Scientometrics,* 82/2 (2010), 391—400) which, for an individual, is simply a geometric mean of her h-index and g-index.

not directly available in Google Scholar. This program also has the advantage of a direct option to calculate g-indices for individual authors.

We need to acknowledge some important limitations of both the database we constructed and of the program used. Most notable are the problem of self-citation, which is not properly dealt with, and the issue of multiple references to the same item. When computing individual g-indices for the political science department staff we did not attempt to correct for either of these based on the assumption that they are randomly occurring events that would not substantially alter the final results. Due to the nature of the g-index, we also did not restrict results to any specific time window: all items retrieved by the *Publish or Perish* software were factored into the g-index for all individuals, thus taking into account the entire output produced for each. In additions, due to data management issues, we did not distinguish between single-authored papers and multiple-authored papers, allowing for multiple-authored papers to be factored into the calculation of each individual g-index.

We computed these indices only for the full-time staff members whose main teaching position is within a particular university. In order to ensure that we correctly selected the members of the staff, we used the personnel lists submitted by the universities themselves to ARACIS, the Romanian Agency for Quality Assurance in Higher Education, for the authorization and accreditation of the study programs within the fields of political science and international relations. However, since the ARACIS evaluations are done every five years and the dates when these staff lists were submitted varied, we cross-referenced these lists with the lists of full-time staff members posted on the websites of universities, where these were available, and added the additional staff in the two fields to the initial lists. In cases where it was not possible to find lists of staff members on the websites, or where the lists were combined for several different fields in the Social Sciences and we could not discern who are associated with a Political Science or International Relations department, we chose to keep the ARACIS lists because we considered them more reliable. Nonetheless, despite our best efforts, given the complexity of the task we acknowledge that our final lists may not be fully correct in all cases. There is invariably at least a small degree of mobility between institutions that cannot be properly documented.

The following clarification should also be made: as pointed out by Sinha and Macri, there are two approaches that can be used to measure the research output of departments: a stock approach which assigns articles to an academic's present affiliation and a flow approach which gives credit to the institution with which the author was affiliated at the time of publication (Sinha & Macri, 2002: 137). We used the stock approach, partly acknowledging the validity of the two authors' argument that publications are a type of human capital which an academic carries with him or her to any institution, partly because the use of the alternative approach would have complicated the empirical efforts of data collection to an unmanageable degree.

For the number of research projects, we extracted data on the submitted projects from the lists of the Executive Agency for Higher Education, Research, Development and Innovation Funding (UEFISCDI)[23], the Romanian national body responsible for managing the national competitions for research funds. We analyzed three years (2007, 2008 and 2009) and three types of projects: the "Young Teams" and "Post-Doctoral Research" projects funded under the Human Resources Development line, and the "Ideas" projects funded through the Exploratory Research Projects competition. Note that the indicator takes into account only the submitted projects, irrespective of whether they have received funding. Although we admit that the number of successful research projects would have been a more relevant indicator, data on which projects were funded were not available for the competitions previous to 2011; the marks each project received do not provide sufficient information, since funding is not given over a cut-off value, but depending on the available budget. In addition, implementation of the projects was also difficult to assess, given that some grants were discontinued after projects began due to lack of funding. Finally, the data indicates that only some of the members of departments under consideration (and only in some universities) participated in these competitions; therefore, the practice of elaborating research projects is not yet so widespread that it would render irrelevant an indicator measuring submissions.

23 See <http://uefiscdi.gov.ro/Public/cat/568/FINANTARE-CDI.html>, last accessed on 15.10.2012.

The three types of projects selected are most relevant for our study due to the fact that they are rather complex projects involving research teams, and due to the conditions these grants impose on project directors wishing to apply (they must have completed a PhD, must have the institutional support of a university, and must also meet a series of requirements concerning scientific activity and output). Although there are a few other types of projects funded through national grants, these were not included because they addressed Romanian researchers from abroad, individual PhD level research (as for example small travel or research grants for one PhD student), or were specifically addressed to technical or medical fields. We also excluded the international research projects, because the participation of Romanian universities is generally modest, especially in the fields discussed here, and it would have therefore been an indicator which distorted the results, giving a premium only to the extremely few universities involved in such projects.

Education is the second component of our proposed methodology. It comprises six indicators: a score for department size, the staff/student ratio, the ratio of MA students to those at BA level, the number of ERASMUS students (both incoming and outgoing, BA and MA) within the total number of BA and MA students, and scores for the ARACIS evaluations of the BA study program, on the one side, and for the overall institutional quality assessment of a university, on the other. Data regarding the number of students are derived from the official information submitted in 2011 by the universities during the national evaluation process[24] and all refer specifically to the field of political science and international relations[25].

The department size index was measured on a 5-point scale: the score for departments with more than 4 professors supervising PhD programs was set to 5; the score for the departments with at least four professors was set to 4; for departments with at least three professors or associate professors it was set to 3; if a department included at least two lecturers or associate professor or professors it received score 2, and departments where none of these conditions were met received score 1. It is important to observe that the size of a department is not a purely quantitative indicator: it includes information about the "quality" of the teaching staff as given by their academic titles. This entails that two departments with a similar number of staff may largely differ in size, as defined here.

The ARACIS score for BA programs was calculated as a weighted average of the number of study programs, with a weight of 1 for authorized and accredited programs with "(full) confidence", and 0.5 for those with "limited confidence". Similarly, the overall institutional evaluations were measured on a 4-point scale, where 4 stand for "high degree of confidence", 3 for "confidence", 2 for "limited confidence" and 1 for "lack of confidence". Data for the two indicators were collected from ARACIS.

As outlined in Table 1, our methodology yields two intermediate scores – one for research and one for education – which are then aggregated into a final score for each university. The research score represented 50% of the total score and, within it, the successive g-index has a weighting of 77%, while the projects/total staff ratio has a weight of 10% and the ratio of PhD students within the total student population has a weighting of 13%. Within the education component, also weighted globally at 50%, the department size index constitutes 40%, the staff/student ratio 20%, and each of the other three indicators 10%. Based on the total scores, universities were ranked from A to E in the same manner that was specified by the official methodology used in the national evaluation exercise. First, each department's score was divided by the highest score obtained out of all departments in order to compute the ratio of its score to the highest possible. Secondly, departments were included in a ranking class based on the ratio of their score to the highest one obtained, following 5 distinct intervals: departments with a ratio to the highest score higher than 73% were ranked in class A; those with a ratio less than 73% but above 50% were ranked in class B; those with a ratio less than 50% but above 30% were ranked in class C; those with a ratio less than 30% but above 10% were ranked

24 Available at http://chestionar.uefiscdi.ro/public5/index.php?page=punivlist, last accessed on 15.10.2012.

25 As already mentioned, data for 5 years were submitted by the universities. We did not, however, aggregate the data for the entire 5 year window. Instead, we took into consideration data in accordance with the duration of the study programs corresponding to the Bologna system (3-years undergraduate; 2-MA; 3-PhD). Therefore, in the case of undergraduate and PhD students, only data for the academic years 2007-2008, 2008-2009 and 2009-2010 were taken into account, and in the case of MA students only those from the academic years 2008-2009 and 2009-2010 were counted.

in class D; finally, class E comprises the departments with a ratio to the highest score less than 10%. We retained this procedure and the corresponding ranking intervals, as outlined by the official methodology used in the national ranking process, largely for comparative purposes.

At the time we initiated our research we found political science and/or international relations study programs in 26 Romanian universities. **Table 2** presents these universities in alphabetical order. The four universities that are highlighted were not included in our study for the following reasons: three of them („Bogdan-Vodă" University, „Sapientia" Hungarian University of Transylvania and the „Titu Maiorescu" University) could not be included because they did not submit data for the political science domain during the national evaluation process and therefore no comparable data could be obtained for several of the indicators we used; this is a consequence of the fact that at the time of the evaluation the political science and/or international relations study programs within these universities had not yet been accredited and were therefore not included in the evaluation process. The fourth („Mihail Kogălniceanu" University from Iaşi) was disregarded due to the fact that its political science study program has been discontinued. This left us with a total of 22 departments to analyze.

A final note regards the way we counted the staff members. In calculating the ratio indicators relating to staff members (namely the staff/student ratio within the education component and the projects/staff ratio within the research component) we only counted those individuals with a title of lecturer and higher. Furthermore, we only counted them if they held the academic title of PhD. University assistants were not included in the calculation of the two ratio indicators. All staff members were, however, irrespective of their position or title, taken into account when computing the successive g-index of the departments. Overall, data for more than 500 individuals were retrieved with the *Publish or Perish* software[26].

Table 2. Romanian universities with political science and/or international relations programs

1	„Alexandru Ioan Cuza" University – Iaşi
2	„Andrei Şaguna" University – Constanţa
3	„Babes-Bolyai" University – Cluj-Napoca
4	„Bogdan-Vodă" University – Cluj-Napoca
5	„Constantin Brâncoveanu" University – Piteşti
6	„Constantin Brâncuşi" University – TârguJiu
7	„Danubius" University – Galaţi
8	„Dimitrie Cantemir" Christian University – Bucharest
9	„Hyperion" University –Bucuresti
10	„Lucian Blaga" University – Sibiu
11	„Mihail Kogălniceanu" University – Iasi
12	National School of Political and Administrative Studies – Bucharest
13	„Nicolae Titulescu" University – Bucureşti
14	„Ovidius" University – Constanţa
15	„Petre Andrei" University – Iaşi
16	„Petru Maior" University – TârguMureş
17	„Sapientia" Hungarian University of Transylvania
18	„Spiru Haret" University – Bucharest
19	„Stefan cel Mare" University – Suceava
20	„Titu Maiorescu" University – Bucharest
21	University of Bucharest
22	University of Craiova

26 The process of data collection was carried out during the first two weeks of October 2012.

23	University of Oradea
24	University of Piteşti
25	„Vasile Goldis" West University – Arad
26	West University of Timişoara

Source: Government Decision no. 707 from 18.07.2012

Results and discussion

As we mentioned at the beginning of our paper, our main goal was to compare the results of the national evaluation and ranking process conducted in the field of political science and international relations with the results obtained by appealing to our substantially simplified methodology. As already described in the previous section, our study included departments from 22 universities, all of which participated in the official ranking process. **Table 3** shows the class comparison between the official ranking and the one produced as a result of our research. In our study the highest aggregate score was obtained by the political science and international relations department of „Babes-Bolyai" University of Cluj-Napoca; therefore, consistent with the ranking procedure used in the official ranking exercise, we computed the ratio of the score obtained by all other departments against the score of this university in order to obtain values that would allow grouping these political science departments into ranking classes.

Looking at the overall picture painted by Table 3, let us compare the different rankings produced by the methodologies under comparison and their associated indicators. First, only 4 universities received the same ranking class in both cases: „Babes-Bolyai" University, the National School of Political and Administrative Studies, the University of Bucharest and the „Alexandru Ioan Cuza" University[27]. As for the other universities, 10 were promoted to a higher class by our ranking process, but 8 others were demoted to a lower class. In general, the magnitude of the changes arising from our proposed methodology is incremental in nature: out of the 18 differing outcomes, 11 universities changed their class with strict proximity, either moving up to the immediately higher class or moving down to the immediately lower class. In the 7 other cases, differences are more visible: five universities ascended by two classes („Nicolae Titulescu" University[28], „Petru Maior" University, „Andrei Şaguna", „Vasile Goldis" West University and „Constantin Brâncoveanu") and two descended by 2 classes (the University of Craiova and the West University of Timişoara). We believe that the more noticeable differences for these seven departments rest heavily on our use of the successive g-department indicator. Nonetheless, we note that no extreme shifts occur (such as an A department falling to D or E, or an E one moving up to the B class or even to A).

Table 3. Comparison between the official ranking of the political science departments and the alternative ranking produced in our study

Rank	Political science departments (host institutions)	Official ranking class	Alternative ranking class	Changes and magnitude
1	„Babes-Bolyai" University	A	A	↔
2	National School of Political and Administrative Studies	A	A	↔
3	University of Bucharest	A	A	↔

27 The „Constantin Brâncuşi" University came very close to maintaining a C ranking because its ratio to the highest score was 0.296. Unfortunately, only rounding this number up to 0.3would actually have achieved this result.

28 The results for the "Nicolae Titulescu" University could be expected: in the official ranking it was heavily downgraded because in 2011 it did not pass through the ARACIS institutional evaluation.

4	„Nicolae Titulescu" University	D	B	↑2
5	University of Oradea	B	C	↓1
6	„Lucian Blaga" University	B	C	↓1
7	„Stefan cel Mare" University	B	C	↓1
8	„Petru Maior" University	E	C	↑2
9	„Alexandru Ioan Cuza" University	C	C	↔
10	University of Craiova	A	C	↓2
11	„Petre Andrei" University	D	C	↑1
12	„Dimitrie Cantemir" University	D	C	↑1
13	„Spiru Haret" University	D	C	↑1
14	West University of Timişoara	A	C	↓2
15	„Andrei Şaguna" University	E	C	↑2
16	„Hyperion" University	D	C	↑1
17	„Vasile Goldis" West University	E	C	↑2
18	„Constantin Brâncoveanu" University	E	C	↑2
19	„Ovidius" University	B	C	↓1
20	„Constantin Brâncuşi" University	C	D	↓1
21	University of Piteşti	E	D	↑1
22	„Danubius" University	C	D	↓1

Another important result is that whereas in the official ranking 6 departments were relegated to the E ranking class, our ranking produced no such result. In other words, none of the scores obtained by the universities had a value less than 10% of the top score obtained by the highest ranking department. We suspect this is the cumulative consequence of two specific provisions in the official methodology: the first stated that all indicators used in the ranking process would be multiplied with a coefficient of either 1, 0.70, 0.40, 0.15 or 0.10 according to the confidence level awarded by ARACIS in light of the most recent institutional evaluation process. The second provision was, however, more severe, stating that all indicators for the first 3 criteria in the official methodology (research, teaching and relation to external environment) would also be multiplied with a coefficient of either 0, 0.90 or 1, depending on the number of full-time staff members allocated to the study program under evaluation. The 0 weighting was valid in cases where this full time staff number was below 5. Taken together, the two global coefficients these two provisions stipulate undoubtedly had a substantially negative impact on small departments that were part of institutions with weak ARACIS confidence levels[29]. The fact that our methodology does not provide such severely penalizing measures might explain why no aggregate scores in our study dropped below the 10% threshold, although a few did come close to it.

Moving to the overall distributions of departments in the two ranking processes, the official ranking seems to have produced more balanced overall results than our own. During the national evaluation, out of the 23 political science programs that were evaluated, 5 were ranked within class A, 4 within B, 3 within C, 5 within D and the remaining 6 in E. Our results, however, point toward a much more polarized situation: out of the 22 departments, only 3 received A class ranking, only one received B, 15 were ranked in class C and 3 in D. This polarization is also readily apparent through comparison of the individual scores received by the departments we studied. **Figure 1** presents these

29 Between the two, the second coefficient might have particularly dire consequences. Consider for example the hypothetical case of a department with 4 members: because all the indicators of the first three criteria are multiplied with 0, its aggregate ranking index is reduced from the very beginning to the value of only the fourth component evaluated by the official methodology. Somewhat ironically for this situation, this fourth component (institutional capacity) has a negligible weighting of only 0.05 in the final score.

scores for all 22 departments. From left to right, each data point in the figure corresponds to the department rank presented in Table 3, in precisely the same order as given by Table 3 (for example, the fourth point, with a score of 2.96 refers to the „Nicolae Titulescu" University; the sixteenth to "Hyperion", etc.). Numbers below the data points refer to the ratio of the particular university's score against the highest score obtained in our study.

Figure 1. Scores of Political Science Departments and ratios to the highest score obtained in the ranking

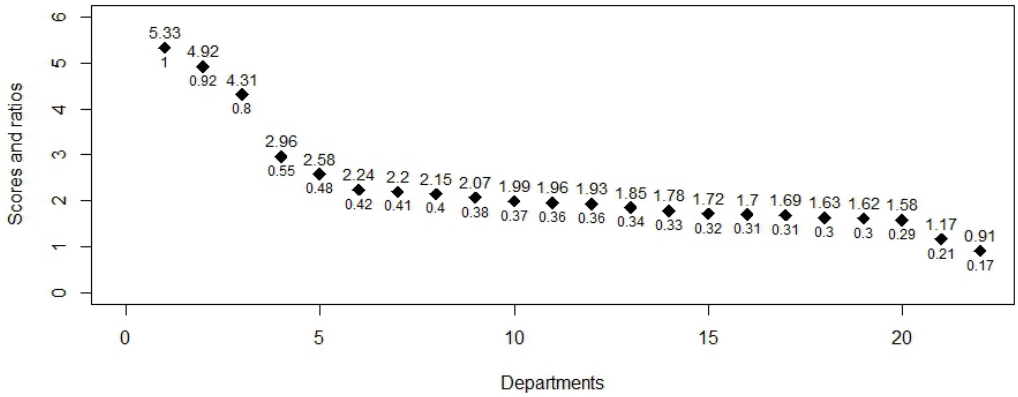


Figure 1 illustrates another interesting aspect: there seems to be a high degree of uniformity in the case of most departments. As the above figure shows, most of the departments received a score ranging between approximately 1.6 and 3. What this means is that in terms of the indicators used in our methodology, with the exception of the first three top departments (University "Babeş-Bolyai" Cluj-Napoca, National School of Political and Administrative Studies and the University of Bucharest), most Romanian political science (and international relations) departments do not seem to differ substantially from one another. A particular subject worth further investigation in this context would be the degree of differentiation of these departments in terms of research productivity and visibility. This is all the more relevant given the fact that emphasis on research was one of the driving forces in the national evaluation process. As we have already mentioned in a previous footnote, within the official methodology, the aggregated research component was weighted at 50% of the total score for all political science programs. In the cases of natural sciences, the research component was weighted at 60% and even in the case of humanities and arts it did not fall below 40%.

Figure 2. Individual Research Scores of Political Science Departments

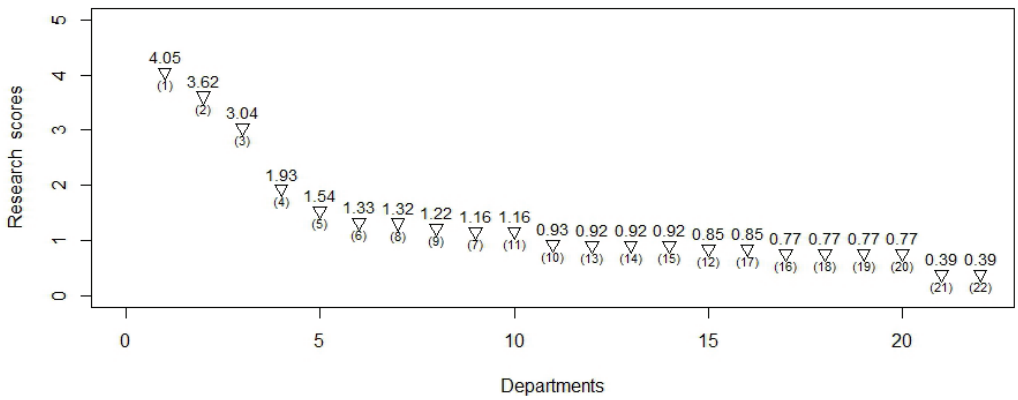


Figure 2 illustrates the scores received in our study by each department only on the research

component given by the three indicators that were used to measure it and their associated weights. These research scores are presented in decreasing order from left to right. In the case of Figure 2, however, the order of the data points no longer (fully) corresponds to the department ranks previously presented in Table 3. Each data point (department) is, however, identified in Figure 2 by using the rank position previously given by Table 3. These corresponding rank positions are given in parentheses under each data point (for example, the eighth data point in Figure 2 from left to right refers to the "Alexandru Ioan Cuza" University, ranked 9 in Table 3) . The reader may therefore note that a ranking based solely on research scores differs from the overall ranking given in Table 3 because the latter takes into account both criteria proposed by our methodology[30].

Figure 2 shows, again with the exception of the three top departments, an even greater uniformity than the aggregate scores previously presented. As can be seen above, most departments (16 out of the total 22) received a score between 0.77 and 1.54. What is also remarkable about the figure above is the convergence of several separate clusters of political science departments towards near perfectly identical scores. In particular, 10 departments can be observed converging at an overall score ranging between 0.77 and 0.93, while a separate cluster of 5 departments converge at a score ranging between 1.16 and 1.33. Given the surprising nature of this finding we took a closer look at the three indicators used to assess the research component and found that this significant homogeneity is almost fully attributable to the values the departments received on the successive g-index indicator: the 5 departments clustered around the score of 1.16–1.33 all have a rational successive g-index of 3 (or slightly higher, but below 4), while the 10 departments clustered around the score of 0.77 to 0.93 all have a rational successive g-index of 2 (or slightly higher, but below 3). This means that in the case of the first cluster of departments, none had at least 4 staff members whose combined individual g-indices outweighed (or at least equaled) a value of 16. For the second cluster, none of the 10 departments had at least 3 staff members whose combined g-indices outweighed (or equaled) a value of 9.

In order to test the stability of the ranking procedure we employed, we also ran a series of 8 distinct simulations based on incremental, 5% modifications to the global weights attached to the education and research criteria. The two most extreme cases we considered had the following global weights: case 1: education- 70%, research- 30%; case 8: education- 30%, research- 70%. All other 6 cases considered intermediate values (for example 65%-45%, 55%-45%, etc.)[31]. These simulations prove our methodology yields rather stable results, especially concerning the top four and the bottom two scoring departments. Also, none of these simulations yield results that imply departments ranked in the E class.

Unlike the top and bottom departments, some of the overall scores and final ranking classes of the other departments we presented in Table 3 are, however, less stable. In particular, after running the simulations we noticed that higher weights attached to the research component tended to produce a greater number of departments ranked in the D class (up to 11 in the extreme case). Conversely, higher weights of the education component tended to cluster many departments in class C (even up to 15). As we have already discussed, the final weighting of 50-50 that we actually used (this is very similar to the official weighting) in our ranking process yields 15 departments in class C. The reader may note that the information contained in the two figures presented so far, one depicting the overall aggregate scores and the other the intermediate scores for research, already cumulatively imply that there is also a high degree of uniformity with regard to the scores obtained by the 22 departments on the intermediate score for education. Indeed, in the case of the scores obtained

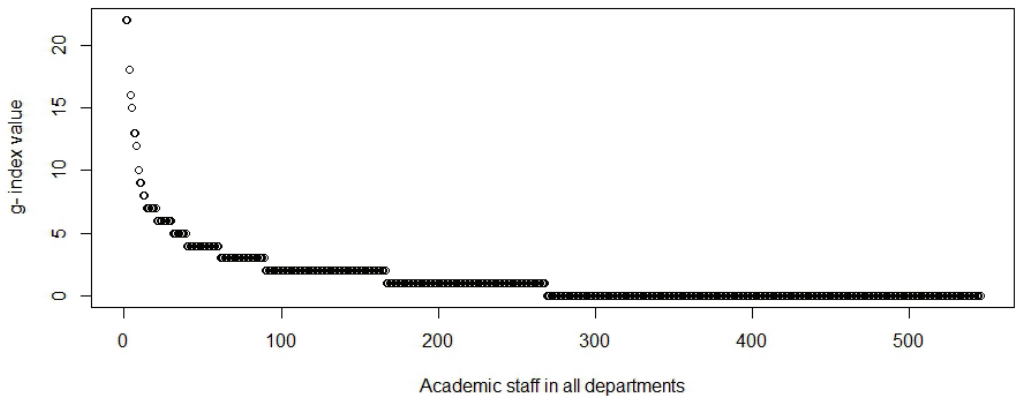
30 We note in passing that there seems to be a fair (and somewhat predictable) correlation between the intermediate score received by the universities for the education component and the one received strictly for the research component. The data in our study point to a correlation of 0.74, significant beyond the 0.01 statistical threshold; however, given the fact that the number of cases under study is necessarily small (only 22 departments) the corresponding (95%) confidence interval for the correlation value given above ranges between 0.44 and 0.88.

31 It is worth mentioning that the two most extreme weightings considered (70% research–30% education and the reverse case) are incidentally the ones that would have produced the most numerous identical ranking classes with respect to the official list (7 departments would have had the same ranking class as the one given by the MoE if either weighting would have been used).

on this dimension, the range of observed values only extends from 0.53 to 1.30. This also better explains the somewhat contrasting results obtained in the simulations. It seems that most of the departments analyzed are even more uniform in terms of the education component than they are as far as research is concerned. This leads us to the conclusion that research has a significantly more important role than the educational component in differentiating between the departments analyzed, regardless of its attached weight.

Given the degree of uniformity depicted by Figures 1 and 2 above we considered that a final digression from our primary goal of comparing ranking results was warranted: what is the overall performance and visibility, as measured by the g-index, for the entire population of staff members working in political science and/or international relations departments in Romania? **Figure 3** illustrates the distribution of all 545 staff members investigated, according to their individual g-index. Our findings are as follows: out of the 545 staff members for which we collected data by means of the *Publish or Perish* software, only 268 have a g-index of at least 1 (basically meaning they have at least one paper with one citation), 166 have a g-index of at least 2, 89 have a g-index of at least 3, 60 have a g-index of at least 4 and only 39 have a g-index of at least 5 or higher (which means that their publications were cited at least 25 times).

Figure 3. Overall distribution of g-indices for staff members of all political science departments



Concluding remarks

Our paper focused on the manner in which different ranking methodologies, criteria and indicators may affect the result of a ranking process. By devising a parsimonious methodology based on a limited number of indicators we compared and ranked the political science and international relations departments in the Romanian universities, in accordance with the guidelines set out in the methodology used during the official ranking process carried out in 2011. Despite having used only 9 indicators (as opposed to the official ranking exercise that used 67), we obtained fairly comparable overall results: several departments maintained their ranking class and most of the others shifted their rank by a single position. The ranking scores we obtained also pointed both to a high polarization (three top universities received a very high score) and also to a high degree of uniformity among the other departments. We make no claim that outside the indicators specifically used in our study this is necessarily true. The official ranking has certainly shown that more variability is possible and the addition of further indicators to our own methodology might indeed point towards greater differences existing between the various departments analyzed.

Nonetheless, for the time being it was not our intention to develop an exhaustive set of indicators meant to distinguish between all the numerous aspects that make up the activity of a department. To the contrary, our methodology appealed to indicators we consider both pertinent and synthetic

in nature. It also weights the research component, particularly evaluated by means of the rational successive g-index. Given the high sensitivity of this indicator, we would have expected to see greater differentiation between the departments analyzed. The fact that, with the exception of the top three universities, this was not the case indicates a degree of homogeneity between the departments that invites further study, possibly through comparison with departments from other study fields[32].

References

Aaltojärvi, Inari, et al., "Scientific Productivity, Web Visibility and Citation Patterns in Sixteen Nordic Sociology Departments", *Acta Sociologica,* 51/1 (2008), 5-22

Altanopouloua, Panagiota; Dontsidoua, Maria and Tseliosa, Nikolaos: "Evaluation of ninety-three major Greek university departments using Google Scholar", *Quality in Higher Education,* 18/1 (2012), 111-137

Alonso, S., et al., "HG-Index: A New Index to Characterize the Scientific Output of Researchers Based on the H- and G-Indices", *Scientometrics,* 82/2 (2010), 391–400

Bornmann, Lutz, "Peer Review and Bibliometric: Potentials and Problems" in Jung Cheol Shin, et al. (eds.): *University Rankings Theoretical Basis, Methodology and Impacts on Global Higher Education* (New York , Springer, 2011), 145-164

Bratti, Massimiliano, et al., "Higher Education Outcomes, Graduate Employment and University Performance Indicators", *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 167/3 (2004), 475-496

Braun, Tibor; Glänzel, Wolfgang and Schubert, András, "A Hirsch-type index for journals", *Scientometrics,* 69/1 (2006), 169–173

Bull, Martin and Espindola, Roberto, "European Universities in a Global Ranking of Political Science Departments: A Comment on Hix", *European Political Science*, 4/1 (2005), 27-29

Butler, Linda and McAllister, Ian, "Metrics or Peer Review? Evaluating the 2001 UK Research Assessment Exercise in Political Science", Political Studies Review, 7/1 (2009), 3–17

Butler, Linda and McAllister, Ian, "Evaluating University Research Performance Using Metrics", *European Political Science*, 10/1 (2011), 44-58,

De Bruijn, Hans, *Managing performance in the public sector* (London and New York, Routledge, 2007)

Coryn, Chris L. S., et al., "Models and Mechanisms for Evaluating Government-Funded Research: An International Comparison", *American Journal of Evaluation*, 28/4 (2007), 437-457

Dale, Tony and Goldfinch, Shaun, "Article Citation Rates and Productivity of Australasian Political Science Units 1995–2002", *Australian Journal of Political Science,* 40/3 (2005), 425–434

Egghe, Leo, "An Improvement of the h-Index: the g-Index", *ISSI Newsletter*, 2/1 (2006), 8-9

Erne, Ronald, "On the Use and Abuse of Bibliometric Performance Indicators: A Critique of Hix's 'Global Ranking of Political Science Departments' ", *European Political Science*, 6/3 (2007), 306-314

Harzing, A.W. (2007) Publish or Perish, available from http://www.harzing.com/pop.htm

Hirsch, Jorge, "An index to quantify an individual's scientific research output", *Proceedings of the National Academy of Sciences,* 102/46 (2005), 16569-16572

Hix, Simon, "A Global Ranking of Political Science Departments", *Political Studies Review*, 2/3 (2004), 293–313

Hix, Simon, "European Universities in a Global Ranking of Political Science Departments: A Reply to Bull and Espindola", *European Political Science,* 4/1 (2005), 30-32

Hodge, David and Lacasse, Jeffrey, "Evaluating Journal Quality: Is the H-Index a Better Measure than Impact Factors?", *Research on Social Work Practice*, 21/2 (2011), 222-230

32 Financial support from the European Union through the European Social Fund is gratefully acknowledged by the authors: Mirela Vlăsceanu's work was supported through the project POSDRU/107/S/76844; Adrian Miroiu's work was supported through the project POSDRU/21/1.5/G/16838

Liu, Nian Cai and Cheng, Ying, "Academic Ranking of World Universities by Broad Subject Fields", *Higher Education in Europe*, 30/2 (2005), 127-136

Marginson, Simon and Van Der Wende, Marijk, "To Rank or To Be Ranked: The Impact of Global Rankings in Higher Education", *Journal of Studies in International Education*,11/3-4 (2007), 306-329

Morgan, David R. and Fitzgerald, Michael R., "Recognition and Productivity among American Political Science Departments", *The Western Political Quarterly,* 30/3 (1977), 342-350

Prathap, Gangan, "Hirsch-type indices for ranking institutions' scientific research output", *Current Science,* 91/11 (2006), 1439

Van Raan, Anthony, "Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups", *Scientometrics,* 67/3 (2006), 491-502

Ruane, Frances and Tol, Richard, "Centres of Research Excellence in Economics in the Republic of Ireland", *The Economic and Social Review,* 38/3 (2007), 289–322

Russell, Andrew, "Retaining the Peers: How Peer Review Triumphs over League Tables and Faulty Accounting in the Assessment of Political Science Research", *Political Studies Review*, 7/1 (2009), 63–72

Shin, Jung Cheol, "Organizational Effectiveness and University Rankings" in Jung Cheol Shin, et al. (eds.): *University Rankings Theoretical Basis, Methodology and Impacts on Global Higher Education* (New York, Springer, 2011) 19-34

Sinha, Dipendra and Macri, Joseph, "Rankings of Australian Economics Departments, 1988–2000", *The Economic Record,* 78/241 (2002), 136–146

Thelwall, Mike, "Bibliometrics to Webometrics", J*ournal of Information Science*, 34/4 (2008), 605-621

Tol, Richard, "A Rational, Successive G-Index Applied to Economics Departments in Ireland", *Journal of Informetrics,* 2/2 (2008), 149-155

Wedlin, Linda, "Going global: Rankings as rhetorical devices to construct an international field of management education", *Management Learning,* 42/2 (2011), 199-218