Research-driven Classification and Ranking in Higher Education: An Empirical Appraisal of a Romanian Policy Experience

Gabriel-Alexandru Vîiu[1], Mihai Păunescu[2], and Adrian Miroiu[3]

[1] *gabriel.alexandru.viiu@snspa.ro,* [2]*paunescu.mihai@gmail.com,* [3]*admiroiu@snspa.ro*
National School of Political and Administrative Studies, Povernei Street 6, 010643 Bucharest (Romania)

Abstract
In this paper we investigate the problem of university classification and its relation to ranking practices in the policy context of an official evaluation of Romanian higher education institutions and their study programs. We first discuss the importance of research in the government-endorsed assessment process and analyze the evaluation methodology and the results it produced. Based on official documents and data we show that the Romanian classification of universities was implicitly hierarchical in its conception and therefore also produced hierarchical results due to its close association with the ranking of study programs and its heavy reliance on research outputs. Then, using a distinct data set on research performance we further explore the differences between university categories. We find that our alternative assessment of research productivity – measured with the aid of Egghe's *g*-index – only provides empirical support for a dichotomous classification of institutions.

Conference Topic
University Policy and Institutional Rankings

Introduction
Since the beginning of the 1980s nationally relevant university research coupled with the pressure for accountability have increasingly shaped the policies and priorities of individual universities (Geuna, 2001). Since then, the growing importance of research has been continually underscored by transnational policy documents such as the EU 2020 Strategy, by implementation of performance-based research funding mechanisms which create new competitive pressures within national university systems (Hicks, 2012) and, perhaps most visibly and controversially, by national and international university rankings which fuel debates surrounding 'world-class universities' (Sadlak & Liu, 2007; Salmi, 2009; Shin & Kehm, 2013). It is now well established that "international rankings of universities have become both popular with the public and increasingly important for academic institutions" (Buela-Casal et al., 2007, p. 351). At the same time rankings have also become "successful as an agenda-setting device for both politicians and for the higher education sector" (Stensaker & Gornitzka, 2009, p. 132).

In this paper we present an empirical exploration of the research-driven ranking and classification processes directed toward the Romanian higher education institutions (henceforth "HEIs") in the policy context of a new Law on National Education. In accordance with the new law a comprehensive process of evaluation was conducted in Romania in 2011 with the dual aim of (1) classifying HEIs (at the global, institutional level) and (2) ranking their constituent study programs. The ranking and classification were conducted using a common methodology that heavily emphasized the research productivity of university staff. Our primary objective is to contribute to a better understanding of the relation between the classification and ranking processes by discussing the methodological outline of the official evaluation and by analyzing its results. To achieve this goal we rely on official documents and on data collected with regard to the actual results of the classification and ranking processes. A secondary objective of our paper is to investigate the consistency of the institutional classification categories used in the official evaluation. To do this we employ an alternative data set on research performance, measured using the *g*-index which – for the set of papers of an individual researcher – represents "the largest rank (where papers are arranged in

decreasing order of the number of citations they received) such that the first *g* papers have (together) at least g^2 citations" (Egghe, 2006, p. 144). Our goal is to investigate whether an alternative assessment of research based on this index confirms the official classification of institutions, which was largely determined by research performance.

Background

Theoretical considerations

Higher education in recent years has witnessed the emergence of numerous university rankings, which have been the focus of comprehensive studies that aimed to investigate their methodological underpinnings, theoretical outlook and practical consequences (e.g., Dill & Soo, 2005; Salmi & Saroyan, 2007; Usher & Medow, 2009; Rauhvargers, 2011). In a more recent study Hazelkorn (2013) noted no less than 10 global rankings and at least 60 countries that have introduced national rankings. All these studies highlight (among other aspects) the fundamental importance that ranking systems generally attach to research performance, the deleterious consequences that rankings may have for institutional diversity and quality and, perhaps most importantly, the methodological caution which should be exercised when undertaking and interpreting rankings.

As more and more rankings have been developed over the years and as concerns have mounted regarding their implications and methodological problems (e.g.: van Raan, 2005; Billaut, Bouyssou & Vincke, 2010; Longden, 2011), the adjacent subject of university classification has also received increased attention (see for example Shin, 2009). This has been the case especially at the broader European level where the international ranking impetus has been critically received by scholars and policymakers and carried forward in a new direction with the introduction of the U-Map and U-Multirank initiatives, which, unlike pre-existing commercial rankings, focus on a user-driven approach and emphasize multidimensionality in evaluation.

Classification of universities has tended to be a much less debated subject than rankings, but these two distinct processes are nonetheless naturally interwoven with each other. On the one hand, due to strictures of comparability "classification is a prerequisite for sensible rankings" (van der Wende, 2008, p. 49). On the other hand, classifications are often interpreted as rankings even though this is clearly against the intentions of the classifying agency. Shulman (2005) and McCormick (2008) provide several examples of how the Carnegie Classification of US HEIs is actually understood as a form of ranking by several types of stakeholders.

A useful analytical distinction made between classifications and rankings involves conceptualizing them in the context of the broader notion of institutional diversity which itself may be divided into vertical diversity and horizontal diversity. According to van Vught (2009), the former refers to differences between higher education institutions owing to prestige and reputation while the latter stems from differences in institutional missions and profiles. In light of this distinction, classifications are "eminently suited to address horizontal diversity" (van Vught & Ziegele, 2011, p. 25) while rankings "are instruments to display vertical diversity in terms of performance by using quantitative indicators" (Kaiser, Faber & Jongbloed, 2012, p. 888).

The Romanian policy of classification and ranking

In 2011, following the provisions of the new law on national education a comprehensive national evaluation was conducted for the first time by the Romanian Ministry of Education with the aim of classifying all accredited HEIs and, additionally, of ranking all accredited study programs offered by the universities. This process was by far the most elaborate evaluation of the Romanian system of higher education and the first one to explicitly

undertake an official classification of HEIs and an official ranking of their study programs on the basis of quantitative indicators.

With regard to the classification process the law stipulated that all universities must be classified as belonging to one of the following three classes: *A – universities focused on education*; *B – universities focused on education and research*; and *C – universities focused on advanced research and education*. This would point toward a functional differentiation with regard to research capacity but the law also stipulated that the allocation of public funding was to be a function of the results of the classification process: universities from class A could only receive public funding for study programs at the bachelor level, those from class B could receive funding for programs at both bachelor and master level, while those from class C were the only ones to receive public funding for all types of programs (including PhD). With regard to the ranking of study programs, the law on education did not contain any detailed provisions. However, a subsequent government decision (789/03.08.2011) established five distinct hierarchical classes A (high quality), B, C, D and E (poor quality). These program ranking classes should not be confused with the university classes.

A detailed methodology for the classification and ranking processes was made public through Ministry of Education Order 5212/ 26.08.2012. This methodology outlined a complex system of criteria, performance indicators, variables and weights. Table 1 provides a simplified account of the evaluation methodology for the particular case of social sciences. At the most general level, four common criteria were used for both classification and ranking purposes: (1) research; (2) teaching; (3) relation to the external environment; and (4) institutional capacity. The most important aspect in the evaluation process was the research performance of the staff working in the universities and/or the study programs under assessment. This is especially significant for our later use of the *g*-index.

Table 1. Criteria, indicators and weights used in the evaluation process for university classification and study program ranking (social sciences).

Criteria and global weights	*Performance indicators and weights within criterion*	*Variables within indicator*
I. Research (weight: 0.50)	Results of scientific research - 0.75	11
	Research funding - 0.10	5
	International recognition - 0.02	2
	PhD programs - 0.13	2
II. Teaching (weight: 0.25)	/	6
III. Relation to external environment (weight: 0.20)	Relation to economic environment - 0.20	2
	Relation to social environment - 0.05	3
	Community development - 0.45	3
	Internationalization - 0.30	9
IV. Institutional capacity (weight: 0.05)	Indicator 1 - 0.34	3
	Indicator 2 - 0.11	3
	Indicator 3 - 0.11	4
	Indicator 4 - 0.11	4
	Indicator 5 - 0.11	4
	Indicator 6 - 0.11	1
	Indicator 7 - 0.11	5

Source: Ministry of Education Order 5212/ 26.08.2012

Within the research criterion four distinct performance indicators were defined but the most important of these four was an indicator dealing with the research output of the staff members employed by the universities. This indicator had a weight of 0.75 while the other three indicators (research funding, international recognition, and PhD programs) had much lower weights (0.10, 0.02, and 0.13). This indicator of research output was itself further broken down into 11 different variables such as the relative influence score of articles, the number of publications in journals indexed in the ISI Web of Knowledge, books, book chapters, etc.

For the ranking of study programs each university reported specific data for all of the distinct programs it operated; then, global indicators were calculated at the level of the study program for the first three criteria listed in Table 1. A separate global indicator was calculated at the university level for the institutional capacity criterion. A further step then entailed the calculation of an overall *aggregated index of ranking* (AIR) based on the four global performance indicators and their attached weights. As a final step in the ranking of a study program, its AIR was compared to the highest one obtained among all the similar study programs and, based on certain predefined intervals, it was finally designated as belonging to one of the five ranking classes.

For purposes of classification a separate *aggregated index of classification* (AIC) was calculated at the global level of each university. The AICs were calculated following a formula which incorporated three factors: (1) the absolute value of the research score obtained at the global level of the HEI; (2) a more complex factor calculated as a sum of the global indicators obtained by each of the study programs organized by the HEIs; and (3) an indicator based on the confidence level given to the HEIs by the Romanian Agency of Quality Assurance in Higher Education (ARACIS) following its periodic evaluations.

Upon calculation of the AICs of all universities the class of a particular HEI could finally be determined. Similar to the process used to establish the ranking classes of study programs, in order to determine a university's class its AIC was compared to the highest one obtained within its category (comprehensive universities were compared to other comprehensive institutions, specialized HEIs were only compared to their counterparts). First, universities were sorted in descending order of their AIC scores. Then, again following predefined intervals, universities were classified in one of the three categories A, B or C.

Without going into further details, it is apparent from even a brief analysis of the methodological outline that the evaluation conducted for purposes of classification actually had the general underpinning of a ranking. This is primarily a consequence of the fact that the classification was based on the composite scores of university performance (the AICs), which were sorted in descending order and clustered in accordance with predefined thresholds. Moreover, the classification relied on the research scores obtained by the constituent study programs of the universities and, therefore, on the partial results of the ranking process of these programs. In effect, research was the object of double counting, once at the individual level of the study programs and once more at the aggregated level of the HEIs. Based only on the analysis of the methodology used in 2011, we may argue that the entire classification process was actually hierarchical in nature and that vertical, not horizontal differentiation was a foreseeable consequence not only at the level of study programs (where ranking was explicit) but also with regard to the more general level of universities (where ranking was disavowed in favour of the more neutral label of 'classification'). However, no empirical analysis has so far been undertaken with regard to the relation between the actual results of the classification and the results of the program rankings. In addition, no independent empirical test of the three classification categories has been conducted, either relying on the performance indicators initially used by the Ministry, or on alternative measures of research performance. In the following paragraphs we will address both issues in an attempt to answer several questions related to the classification and ranking processes.

Research questions

Given the unique nature of the classification and ranking processes undertaken by the Romanian Ministry of Education several important aspects invite questioning and empirical study. We will confine our analyses to the following:

1. Did the overlap in methodology with the program rankings have empirically discernible consequences for the more general process of classification? Is there a significant degree of association between particular classes of universities and particular classes of study programs? If so, which types of programs are more common in which types of university?
2. Since the classification process relied heavily on research outputs, can an alternative assessment of the research productivity of universities confirm the threefold classification? Are there significant differences with regard to the research productivity of faculty members *between* the three university classes? Furthermore, are there significant differences with regard to the research productivity of faculty members *within* the three university classes?

The first set of questions addresses the official classification and ranking processes in tandem and implies an investigation of data on the official results. The second set of questions only addresses the classification process and will be explored using a distinct approach, which will be described in the subsequent section.

Methodology

In order to investigate our first set of research questions we created a comprehensive data set of the results of the ranking process for all the study programs evaluated in 2011. We then added the results of the classification of universities in order to obtain a final data set comprising all the study programs, the ranking class in which they were placed following the evaluation process and the class in which the university managing them was placed following the separate evaluation for classification. This primary data set contains 1056 observations of distinct study programs. To test for the level of association between ranking and classification results we created contingency tables for the occurrence of particular types of study programs (i.e. ranked in class A, B, C, D, E) in the three classes of universities (i.e. class A, class B and class C). Additionally, a chi-squared test was also used to investigate the association between the classification and ranking categories.

To explore the second set of research questions we used a distinct data set composed of information on 1,z385 Romanian faculty members active in the fields of political science, sociology and marketing. Specifically, we used their *g*-index to conduct an alternative assessment of university research output. These staff members represent the full populations of staff employed in political science, marketing and sociology study programs and they are spread out across 64 departments (study programs) and 34 distinct universities. Information on the identity of the staff members was obtained from ARACIS and, for each of the staff members in this second data set, the *g*-index was extracted using Anne Harzing's Publish or Perish software (Harzing, 2007) using a procedure previously employed in Vîiu et al. (2012) in an examination of political science departments. With regard to this secondary data set, the results of the official classification of Romanian HEIs would imply that there are significant differences between the staff employed in the three university classes with respect to their research output. To test this we employ analysis of variance and subsequent Tukey HSD tests to reveal the instances where differences between *g*-indices are significant. We first compare the university classes globally, and then refine our analysis to take into account more granular differences between staff types. We thus compare the four staff types – assistants, lecturers, associate professors and full professors – across the three university classes in order to determine whether or not there is a structural difference between these classes.

Results and Discussion

Relation between official ranking and official classification results

With regard to our first set of research questions a review of Table 2 and Figure 1 indicates that universities classified as being focused on education have a limited number of top-performing study programs (90 ranked in A and B, i.e. 17% of all study programs in this class of universities) but cluster the most programs with middle and low performance (those ranked in classes C, D and E add up to 83% of programs managed within the universities focused on education). On the other hand, universities focused on advanced education and research hold a total of 185 study programs and 121 of these (over 65%) are ranked in class A. Another 39 are ranked in class B (thus, over 86% of the programs in this class of universities are ranked in classes A and B) and only less than 5% belong to the lower performing classes D and E. Universities classified as being focused on both education and research have mixed results: out of a total of 344 study programs managed by these universities 189 (55%) are ranked in classes A and B, 28% are in class C and the remaining 17% are ranked in C, D and E.

Table 2. Contingency table of ranking classes of study programs and university classes.

University class	*A - Education*	*B - Education and research*	*C - Advanced research*	*Row total*
Class of study program in official ranking				
A	22 **4.17%**	60 **17.44%**	121 **65.41%**	203
B	68 **12.90%**	129 **37.5%**	39 **21.08%**	236
C	147 **27.90%**	97 **28.20%**	17 **9.19%**	261
D	112 **21.25%**	16 **4.65%**	3 **1.62%**	131
E	178 **33.78%**	42 **12.21%**	5 **2.70%**	225
Column Total	527 **100%**	344 **100%**	185 **100%**	1056

Chi-Square test of ranking classes of study programs and university classes

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	495.433	8	.000
N of Valid Cases	1056		

A more detailed study of the relationship between observed and expected count values of the different classes of study programs within each of the three university classes is also instructive. This study indicates a negative association between programs ranked in classes A and B and universities from class A. A further negative association can also be observed with regard to programs ranked in classes A, D, and E and universities from class B. Finally, universities from class C are negatively associated with study programs ranked in classes B, C, D, and E. On the other hand, a positive association exists between universities from class A and study programs ranked in classes C, D and E. A further positive association exists between universities from class B and programs ranked in classes B and C. Universities from class C are positively associated only with programs ranked in class A.

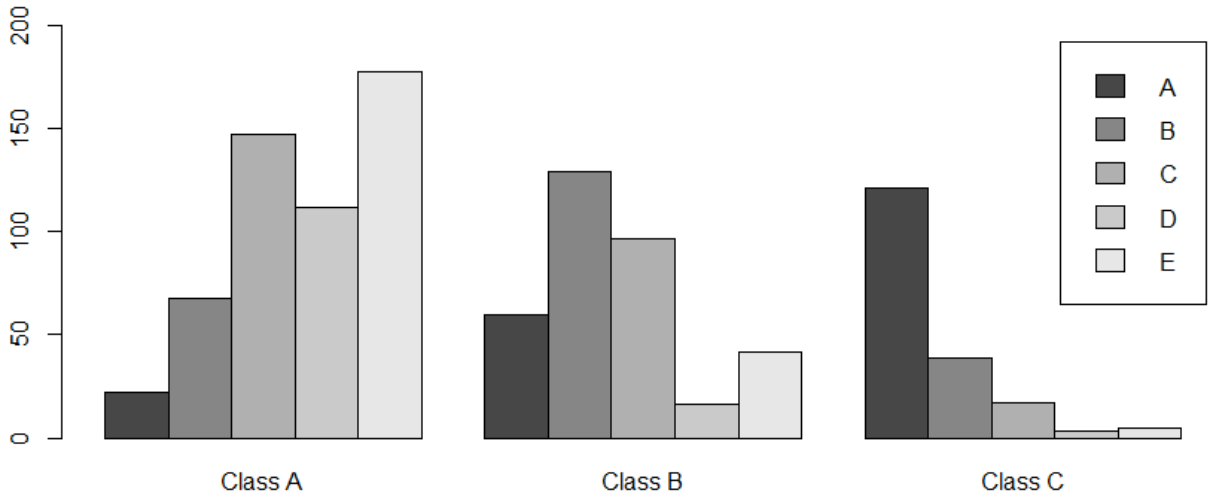

Figure 1. Distribution of study program types across the three university classes.

The results of this analysis paint a rather clear and polarized picture in which universities focused on education generally cluster study programs with poor performance while universities focused on advanced research cluster the programs with high performance. In addition, universities focused on advanced research are fewer and more selective (accounting for a total of only 185 study programs) as compared to universities focused on education (which manage a total of 527 programs). A certain hierarchy is implicit: universities focused on advanced research seem to be 'better' than those focused on both education and research which, in turn, are 'better' than those focused solely on education. However, as we mentioned earlier, these results were to be expected since both the classification and the ranking evaluation relied on a common methodology, which was mostly concerned with research performance. This leads us to our second set of research questions.

Differences in research productivity across and within university classes

We now move to explore whether our secondary data set enables us to distinguish between three university classes. In particular, what we want to see is whether the average *g*-index of all academic staff in class A universities is significantly lower than the average *g*-index of staff in class C universities and also in class B universities. The ANOVA procedure yields the results presented in Table 3. The subsequent Tukey HSD test indicates significant differences between all three means (although the confidence level for the class A – class B distinction is lower, but still above 95%) and therefore seems to provide empirical ground for the threefold classification, which was legally mandated in 2011.

Table 3. ANOVA of *g*-index with regard to university class (N=1,385).

Model summary for ANOVA of g-index with regard to university class

	Sum of Squares	df	Mean Square	F
Between Groups	953	2	476.3	81.62
Within Groups	8065	1382	5.8	Sig.
Total	9018	1384		0.000

Tukey HSD values for ANOVA of g-index with regard to university class

Comparison	Difference	Lower bound	Upper bound	*p*-value
Class A – Class C	-2.119	-2.513	-1.725	0.000
Class B – Class C	-1.714	-2.136	-1.293	0.000
Class B – Class A	0.405	0.054	0.756	0.019

However, the results presented in Table 3 only provide information on the global differences between university classes with regard to the *g*-indices of their entire staff, without further consideration of academic titles. Therefore, in order to test the consistency of the threefold model of classification imposed by the 2011 law, we must explore in greater depth the differences between universities, taking into account not only their classes, but also more granular differences between their academic staff. We thus set out to test not only the global aggregate differences, but also the *structural* patterns of the three classes of universities, taking into account the academic titles of the teaching staff.

In other words, bearing in mind the results of the official evaluation from 2011, we wish to know whether, for example, associate professors from class A universities are significantly different from associate professors in class B universities and from those belonging to class C and, still further, if the associate professors from class B institutions are different from those from class C. Similarly, we also wish to know whether assistants, lecturers and full professors from one class of universities are different from those belonging to the other two classes of universities. Based on such analyses we may draw more general conclusions regarding the degree of structural differentiation that exists between the three classes of universities.

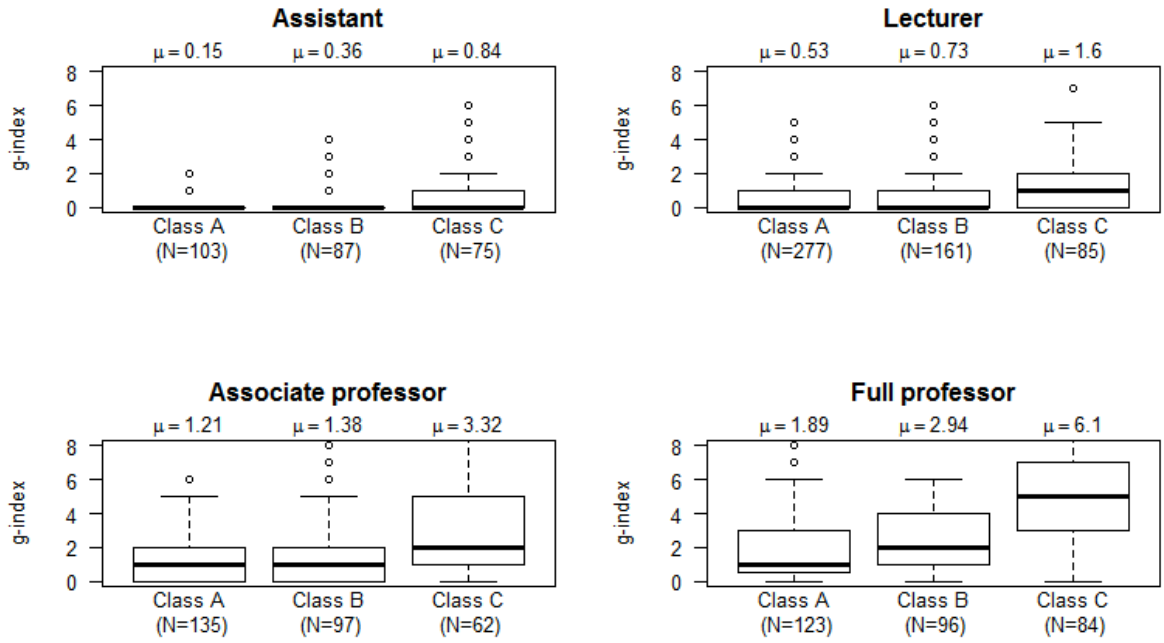

Figure 2. Distribution of g indices by academic title and university class.

Figure 2 illustrates the distribution of academic staff in our secondary dataset with respect to academic titles and also with regard to the class of university they belong to. Mean values are presented in the upper sections as μ. An initial visual inspection of the data would seem to indicate that in the case of assistants, lecturers and even associate professors there are no substantial differences between class A universities and those from class B. On the other hand, all three staff types working in class C universities seem to have substantially different *g*-indices compared to the ones from both class A and class B universities. A somewhat more nuanced picture emerges when looking at full professors. In this case the *g*-indices are more easily distinguishable between university classes and there indeed seem to be differences not only between class C and the other two university classes, but also between these two.

Based on the information contained in Figure 2 and on the ANOVA procedures presented in Appendix 1 we may now answer our secondary research questions. In the case of all staff members (be they assistants, lecturers, associate or even full professors) the parametric

statistical procedures show that universities classified within the official evaluation of 2011 as focused on advanced research (class C) are indeed significantly different from the other two types. In other words, assistants, lecturers, associate and full professors working in these universities focused on advanced research have significantly higher *g*-indices than their counterparts from education-centred universities, as well as from those in universities focused on both research and education. Beyond the clear distinction of staff members working in class C universities, statistical procedures also confirm something that Figure 2 reveals in a more intuitive manner: virtually no statistically significant distinction can be made between class A universities and the universities classified in 2011 as belonging to class B: assistant staff from class A universities are in no way significantly different form assistant staff working in class B universities, lecturers from one are in no way different from lecturers in the other and neither are associate professors. Even the apparent differences described by Figure 2 between full professors from class A universities and those from class B universities do not seem to be statistically meaningful either, as can be observed in Appendix 1. This suggests that a dichotomous classification would fit the data better than the threefold model imposed by the 2011 law.

So far we have argued that the data we have available clearly indicate significant *inter*-university differences (at least insofar as class C universities are made up of staff with higher indices than both class A and B universities). We now turn to *intra*-university differences. We have a reasonable expectation that within research universities there is a greater gap between the four staff types with regard to their scientific productivity. In other words, within class C universities we expect that the *g*-indices of assistants, lecturers, associate and full professors show greater dispersion than the corresponding indices of the equivalent staff that are employed in class A and class B universities. If we review the mean *g*-index values in Figure 2 we can observe that they appear to confirm our expectation. Whereas in the case of class A universities the gap between an average assistant and an average full professor is 1.74 and in the case of class B universities this gap is 2.58, in class C universities the difference is no less than 5.26. This indicates that full professors in research-centred universities have a substantially larger scientific contribution in their fields of study, not only when compared to staff employed in class A and class B universities, but also in comparison to their colleagues from the same university class. This suggests more competitive selection mechanisms of highly qualified academic staff in the research-centred universities compared to the other two university classes. These more competitive selection mechanisms may actually explain the institutional differences.

Concluding Remarks

The boundaries between classification and ranking of higher education institutions are often hard to establish and it is even harder to properly communicate the differences to intended stakeholders. When classification and ranking processes are carried out simultaneously and using common criteria the task of disambiguation becomes virtually impossible and the risk that a classification is perceived as a ranking increases exponentially. In the case of the evaluation conducted in Romania in 2011 the boundaries between classification and ranking were weak from the very inception of these evaluation processes in the law on education. The official methodology for classification and ranking further obscured the differences between the two due to its reliance on common criteria and indicators, most notably the research performance of academic staff employed by the HEIs.

By analysing the official methodology we have shown that the classification of Romanian HEIs carried out in 2011 had the underpinning of a ranking. By further analysing the results of both the classification and ranking processes we have shown that there is a clear association between the outcomes of the global process of classification and those of the more

specific process of program ranking: a polarized landscape thus emerges in which HEIs classified as focused on education cluster the overwhelming part of poor performing programs, while universities classified as focused on advanced research cluster the better part of the top performing programs.

The intermediate class of universities focused on both education and research presents mixed results. However, by conducting an alternative assessment of the research performance of the individual staff employed by Romanian universities in three fields of study we have shown that the threefold classification may not have a sufficiently robust empirical grounding, at least insofar as social sciences are concerned. By using the *g*-index as a concise measure of research performance we have illustrated the fact that the intermediate universities focused on both education and research may not be sufficiently distinct from the universities focused on education and therefore this intermediate class might have a certain degree of redundancy. When looking in our data set of 1385 staff members only at the aggregate results across university classes we do find some empirical grounding for the three classes defined in 2011. However, when analysing in greater detail the structure based on the academic titles and positions, we find less empirical grounds for the threefold classification as most of the staff employed in class A and class B universities are virtually indistinguishable from one another (i.e. assistants, lecturers and associate professors). It is only full professors that seem to make a more substantial difference between class A and class B universities, thus narrowly substantiating a threefold classification, which might otherwise well be a simpler dichotomous one. Previous extensive studies on the quality of Romanian higher education (Păunescu et al., 2012; Vlăsceanu et al., 2011; Miroiu & Andreescu, 2010) revealed the structural isomorphism of the Romanian higher education organizations. The undifferentiated set of standards that all institutions must comply with for purposes of accreditation and public funding led the institutions to adopt similar strategies for achieving these objectives. This is reflected in the poor differentiation and homogeneity of HEIs as shown by their similar scores in the external evaluation of the accreditation agency, similar missions, similar achievements on various performance indicators, etc. While there is empirical support for the vertical differentiation between advanced research universities (usually traditional, older universities) and the rest (more recent, including all private initiatives), the actual structures of the bulk of HEIs, including class A and class B universities, reveal more similarities than differences. These findings should of course be considered under the due caveat that our results are based only on data collected for social sciences.

Acknowledgments

Financial support from the National Research Council (grant number PN-II-ID-PCE-2011-3-0746) is gratefully acknowledged by Gabriel Vîiu and Adrian Miroiu.

References

Billaut, J.-C., Bouyssou, D. & Vincke, P. (2010). Should you believe in the Shanghai ranking? An MCDM view. *Scientometrics*, 84, 237–263

Buela-Casal, G., Gutiérrez-Martínez, O., Bermúdez-Sánchez , M.P. & Vadillo-Muñoz, O. (2007). Comparative study of international academic rankings of universities, *Scientometrics*. 71, 349–365

Dill, D. & Soo, M. (2005). Academic quality, league tables, and public policy: A crossnational analysis of university ranking systems. *Higher Education*, 49, 495–533

Egghe, L. (2006). Theory and practise of the G-index. *Scientometrics*, 69, 131–152

Geuna, A. (2001). The changing rationale for European university research funding: Are there negative unintended consequences? *Journal of Economic Issues*, 35, 607–632

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

Hazelkorn, E. (2013). How rankings are reshaping higher education. In Climent, V., Michavila, F. & Ripolles, M. (Eds.), *Los rankings univeritarios: Mitos y realidades*. Ed. Tecnos

Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, 41, 251–261

Kaiser, F., Faber, M. & Jongbloed, B. (2012). U-Map, university activity profiles in practice. In Curaj, A., Scott, P., Vlăsceanu, L., Wilson, L. (Eds.), *European Higher education at the Crossroads: Between the Bologna Process and National Reforms* (pp. 887–903). Dordrecht: Springer

Longden, B. (2011). Ranking indicators and weights. In Shin, J.C., Toutkoushian, R.K. & Teichler, U. (Eds.), *University Rankings. Theoretical Basis, Methodology and Impacts on Global Higher Education* (pp. 73–104). New York: Springer

McCormick, A. (2008). The complex interplay between classification and ranking of colleges and universities: Should the Berlin principles apply equally to classification? *Higher Education in Europe*, 33, 209–218

Miroiu, A. & Andreescu, L. (2010). Goals and instruments of diversification in higher education. *Quality Assurance Review for Higher Education*, 2, 89–101

Păunescu, M., Florian, B. & Hâncean, M.-G. (2012). Internalizing quality assurance in higher education: Challenges of transition in enhancing the institutional responsibility for quality. In Curaj, A., Scott, P., Vlăsceanu, L., Wilson, L. (Eds.), *European Higher education at the Crossroads: Between the Bologna Process and National Reforms* (pp. 317–338). Dordrecht: Springer.

Rauhvargers, A. (2011). *Global University Rankings and Their Impact*. Brussels: European University Association

Sadlak, J. & Liu, N.C. (2007). *The World-class University and Ranking: Aiming beyond Status*. Bucharest: UNESCO-CEPES

Salmi, J. (2009). *The Challenge of Establishing World-class Universities*. Washington DC: World Bank

Salmi, J. & Saroyan, A. (2007). League tables as policy instruments: uses and misuses. *Higher Education Management and Policy*, 19, 31–68

Shin, J.C. (2009). Classifying higher education institutions in Korea: A performance-based approach. *Higher Education*, 57, 247–266

Shin, J.C. & Kehm, B. (Eds.). (2013). *Institutionalization of World-class University in Global Competition*. Dordrecht: Springer

Shulman, L.S. (2005). Classification's complexities. *The Chronicle of Higher Education* (November 11, 2005), 52, p. B20

Stensaker, B. & Gornitzka, A. (2009). The ingredients of trust in European higher education. In Kehm, B.M., Huisman, J. and Stensaker, B. (Eds.), *The European Higher Education Area: Perspectives on a Moving Target* (pp. 125–139). Rotterdam: Sense Publishers

Usher, A. & Medow, J. (2009). A global survey of university rankings and league tables. In Kehm, B.M. and Stensaker, B. (Eds.), *University Rankings, Diversity, and the New Landscape of Higher Education* (pp. 3–18). Rotterdam: Sense Publishers

van der Wende, M. (2008). Rankings and classifications in higher education: A European perspective. In Smart, J. C. (Ed.) *Higher Education: Handbook of Theory and Research* (pp. 49–72), Vol. XXIII, Springer.

van Raan, A. F. J. (2005). Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62, 133–143

van Vught, F. (2009). Diversity and differentiation in higher education. In van Vught, F. (Ed.) *Mapping the Higher Education Landscape. Towards a European Classification of Higher Education* (pp. 1–16). Dordrecht: Springer

van Vught, F. & Ziegele, F. (Eds.).(2011). *Design and Testing the Feasibility of a Multidimensional Global University Ranking. Final Report*. Consortium for Higher Education and Research Performance Assessment, CHERPA-Network

Vîiu, G.-A., Vlăsceanu, M., & Miroiu, A. (2012). Ranking political science departments: the case of Romania. *Quality Assurance Review for Higher Education*, 4, 79-97

Vlăsceanu, L., Miroiu, A., Păunescu, M. & Hâncean, M.-G. (Eds.). (2011). *Barometrul calității 2010. Starea calității în învățământul superior din România.* Brașov: Editura Universității Transilvania din Brașov.

Appendix 1. Tests of difference for g-index across academic titles and university classes.

*1.Model summary for ANOVA of g-index of **assistant staff** with regard to university class*				
	Sum of Squares	df	Mean Square	F
Between Groups	20.68	2	10.341	13.29
Within Groups	203.82	262	0.778	Sig.
Total	224.50	264		0.000
Tukey HSD values for ANOVA of g-index of assistant staff with regard to university class				
Comparison	Difference	Lower bound	Upper bound	*p*-value
Class A – Class B	0.212	-0.090	0.515	0.224
Class C – Class A	0.684	0.369	1.000	0.000
Class C – Class B	0.472	0.144	0.799	0.002
*2.Model summary for ANOVA of g-index of **lecturers** with regard to university class*				
	Sum of Squares	df	Mean Square	F
Between Groups	73.7	2	36.85	25.39
Within Groups	754.8	520	1.45	Sig.
Total	828.5			0.000
Tukey HSD values for ANOVA of g-index of lecturers with regard to university class				
Comparison	Difference	Lower bound	Upper bound	*p*-value
Class A – Class B	0.195	-0.085	0.475	0.232
Class C – Class A	1.062	0.710	1.413	0.000
Class C – Class B	0.867	0.487	1.246	0.000
*3.Model summary for ANOVA of g-index of **associate professors** with regard to university class*				
	Sum of Squares	df	Mean Square	F
Between Groups	204.8	2	102.40	24.44
Within Groups	1219.2	291	4.19	Sig.
Total	1424			0.000
Tukey HSD values for ANOVA of g-index of associate professors with regard to university class				
Comparison	Difference	Lower bound	Upper bound	*p*-value
Class A – Class B	0.166	-0.475	0.808	0.813
Class C – Class A	2.107	1.367	2.847	0.000
Class C – Class B	1.941	1.157	2.725	0.000
*4.Model summary for ANOVA of g-index of **full professors** with regard to university class*				
	Sum of Squares	df	Mean Square	F
Between Groups	914	2	457.0	34.83
Within Groups	3936	300	13.1	Sig.
Total	4850			0.000
Tukey HSD values for ANOVA of g-index of full professors with regard to university class				
Comparison	Difference	Lower bound	Upper bound	*p*-value
Class A – Class B	1.053	-0.108	2.215	0.084
Class C – Class A	4.212	3.005	5.420	0.000
Class C – Class B	3.159	1.884	4.433	0.000