UNIVERSIDAD SANTO TOMAS
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA

Comunicaciones en Estadística
Junio 2017, Vol. 10, No. 1, pp. 57–82

Multiple Factor Analysis for Ranking Latinamerican Universities

Análisis factorial múltiple para clasificación de universidades latinoamericanas

Jimmy Antonio Corzo Salamanca[a]
jacorzos@unal.edu.co

Abstract

We use the Multiple Factor Analysis (MFA) to built five classes of Latinamerican Universities from three known university *rankings*. These classes distinguish among universities with high level of specialization and low academic output, universities of excelence with low Scientific Leadership, universities with no good reputation and productive staff, universities with good reputation and few doctoral staff, and productive universities with high impact and low indicators of international collaboration. The factors produced by the MFA reveal some paradoxes corroborated in the classification by the fact that they counterpose the level of specialization vs. Productivity, the scientific leadership vs. impact and quality of the output, and they reveal too the possible inconvenience to include judging criteria, which result independent of leadership and impact.

***Keywords*:** Multiple Factor Analysis, University *rankings*, Universities Classification.

Resumen

Se utiliza la técnica del análisis factorial múltiple (AFM) para proponer cinco clases de universidades latinoamericanas a partir de tres conocidos *rankings*, los cuales distinguen, por ejemplo: universidades con alto grado de especialización y baja productividad, universidades de excelencia con bajo liderazgo científico, universidades que no gozan de buena reputación, pero tienen docentes productivos, universidades con buena reputación y pocos docentes con doctorado y universidades productivas de alto impacto y bajos índices de colaboración internacional. Los resultados del AFM revelan algunas paradojas que se corroboran en la clasificación como el hecho de oponer el grado de especialización a la productividad, el liderazgo científico contra impacto y calidad de la productividad, o la posible

[a]Professor - National University of Colombia, Headquarters Bogotá, Colombia

inconveniencia de incluir criterios de opinión que resultan independientes del liderazgo y del impacto[1].

***Palabras clave*:** Análisis factorial múltiple, *rankings* de universidades, clasificación de universidades.

1 Introduction

rankings are classifications of universities that are based on indicators of academic and research criteria and, in some cases, also on perceptions on the prestige of the universities. The first that can be framed into this definition was published in 1998 by the Centrum für Hoshschulentwicklung (CHE) with that very name. Its objective was to offer information to applicants and students about masters' and Ph.D. programs, in order to show internationalization indicators from a selected group of European universities. Since then, their spreading has been notorious at the point that Sanchez & Moreno (2011) report 21 *rankings*, most of them made from 2007 up to now. An interesting historical review on *rankings* around the globe is found inMartinez Rizo (2011).

The criteria used for the *rankings*, as well as the averages assigned to them, vary according to the objectives of the ranking, the theme fields to which they refer, the geographic regions for which they have been built, and of course the interests of the entities that build them. On the other hand, all of them embrace an implicit university model that has been defined by the producer entity of the ranking, which does not necessarily correspond to academic needs or interests of the institutions. Nonetheless, most of them include at least three common criteria related to productivity (institutional or of the professors-researchers), impact, and quality.

Many have been the questions to such *rankings*, to the criteria or indicators that have been used, to the model of university that is implicit, to their relevance when used in budget or resources assignation, to the fact of being based on English-speaking publications, to the exclusion or low importance assigned to the research in social studies, to the dependence (of some indicators) on the size of the institutions, etc.; see Margison & Ordorika (2010) or Martinez Rizo (2011). Additionally, the institutional interest for obtaining the best places in the *rankings* has raised the risk that universities feel forced to relocate their research paradigms in order to get a better position in such *rankings*, which in turn makes them abandon topics of local, regional or national interest.

[1]Agradezco al Estadístico Jairo Jimenez por su apoyo en el arreglo de la base de datos.

To contribute to the debate, we analyze similarities and differences among three *rankings* (Scimago, Webometrics, and QS), which periodically produce information on Latin American universities. To illustrate the use of MFA, along the first section, a classification based on the simultaneous use of the three *rankings* indicators is proposed.

Through the second section, we make a brief presentation of the criteria used in the three *rankings* used in this analysis. The third section makes a quick revision on the methodology, while the fourth show the main results, to finally conclude in the fifth section with the conclusions and a brief discussion on the results.

2 Description of three classifications that use *rankings*

For this exercise, three *rankings* that gather and periodically publish information on Latin American universities, and partially share some criteria on their construction, were chosen.

University Web *Ranking*(Web)

Also called Webometrics, it is known as a web indicator of institutional presence. It is built upon four indicators: 1. impact on the web (scored with the 50% for the *ranking*) calculation), measured by the number of ingoing links to the university domain; presence (16.66%), which is the number of webpages hosted in the university domain; openness (16.66%), which corresponds to the number of enriched files in .PDF, .DOC, .DOCx, .PPT, that have been published in the institution's website; excellence (16.66), measured by the academic works that have been published in international journals that are in the 10% top of the most cited articles (obtained from the Scimago ranking).

Scimago *Institutions Ranking* (Sci)

Self-defined as an evaluation indicator of the research activity of universities and research institutions, Scimago uses three indicator groups:

Research indicators

Institutional productivity: productivity [2] measured by the number of articles from the institution published in Scopus indexed journals;

International collaboration: production of the institution in collaboration with foreign institutions;

Normalized impact: indicator of "about" or undercitations of the production of an institution with respect of the number of citations of the production world:

$$IN = \frac{\text{Average impact of an institution's scientific production}}{\text{Worldwide average impact}}$$

In which the impact is measured by the number of citations of the scientific production. For example, NI= 1.2 means that the scientific production of the institution is cited 20% above the world citation average, while NI = 0.7 indicates that the productivity of the institution is cited 30% under the citation average of worldwide productivity [3]. Here, it is important to differentiate:

High quality publications: proportion of publications from an institution in journals ranked in Scimago Journal Rank first quartile.

Specialization index: indicates the degree of concentration or topic dispersion of an institution's scientific production, and it is calculated like the Gini index.

Excellence: calculated as the percentage of scientific production from an institution, ranked in the 10% of the most cited papers in its area.

Scientific leadership: percentage of an institution's production in which the institution itself is the main contributor.

Excellence and leadership: indicates the number of excellence documents in which the main contributor is the institution.

Scientific talent: number of institution's authors in the institution's whole production.

Innovation indicators

Innovative knowledge: measured by the institution's scientific production that has been cited in patents.

Technologic impact: indicates the percentage of scientific publications cited in patents.

[2]Different from teaching productivity used in the Qs in which it is measured by the number of articles per professor.

[3]Taken from the item ¨oriented field normalized citation score average[a]" produced by the Karolinska Intitutet in Sweden.

Indicators of presence in the web

Size of the web: corresponds to the number of web pages associated to an institution's URL according to Google (https://www.google.com)

Ingoing links to the institution's domain: number of ingoing links to the institution's domain, taken from the known link analysis tool Ahrefs (https://ahrefs.com).

QS Universities *Ranking*

Presented with the intention to show the best universities in Latin America based on seven indicators: 1. reputation among academic peers obtained by a survey; 2. reputation among employers, also obtained by a survey; 3. number of students by professor; 4. number of professors holding a Ph.D.; 5. average number of citations per published article, taken from Scopus; 6. average of articles published by professor; taken from Scopus; 7. Impact on the web, taken from Webometrics.

Weighing of the criteria used for this *ranking* have varied through the years. For example, by 2014, the two indicators of reputation had weighings of 30% and 20% respectively, and the other five had weighings of 10% each, whilst for 2012, weighings were 40 and 10 for the reputations among academic peers and employers respectively, and 20%, 20%, 5%, 5% for the other four criteria, respectively. The former induces certain instability and makes the values of the indicator per year incomparable.

3 METHODOLOGY

The MFA is a multivariate data analysis technique that generalizes the analysis of main components (MCA). It is applied in contexts in which groups of variables of diverse nature in the same set of objects are observed. In such a sense, it is not adequate to perform a MCA of all mixed groups. Another case in which the MFA is used is when groups have big differences in the number of variables, in which case those groups of variables in which there is a higher number of them would dominate in the MCA. The MFA technique is basically to perform MCA separated by the variable groups, using indicators with shared characteristics by the groups of variables (coefficients Lg), in order to select -from the separated analysis- the shared factors. In the end, a global MCA is performed over the shared factors with which the associations among the topics included in the variables are identified. A complete summary of the MFA theory is offered by Escofier (2008)

The database that was used is the fusion (intersection) of three databases obtained from the three *rankings*websites; it contains the 150 universities that occupied

some of such positions according to the Qs *ranking*. Due to the fact that the presence criteria in the Scimago and Qs websites are taken directly or indirectly from Webometrics -or their sources-, such criteria were excluded from these two *rankings* and Webometrics was used directly instead.

The analysis was performed via MFA over the following sets of variables, formed by the indicators of each *ranking*: Webometrics' set of indicators with some indicators; Scimago (Sci) set of indicators with eight indicators; Qs set of indicators with six indicators; set of the country used only as complementary. Later, a hierarchical classification was performed with the first four factors of the MFA. For every calculation Rcode was used, the `FactoMineR`package, as well as the MFA and $HCPC$ procedures.

4 RESULTS AND DISCUSSION

Results are shown in two parts: in the first, a summary of the results of the separated analysis is presented, as well as a basic description of the global analysis factors, and the analysis of the relationships among *rankings* (i.e., among between the *rankings*) indicators set), and the relationships between them and the MFA.

4.1 MFA results

Separated Analysis

Webometrics *Ranking* (Web): figure 1 illustrates how the first factor collects 72% of the variance, and includes its four indicators; that is why it can be adequately summarized by this synthetic indicator. Nonetheless, it is interesting that the second factor opposes presence in the web with the excellence indicator, unveiling a possible paradox which in turn would indicate how counterproductive such presence would be to reaching excellence (see figure 1).

s[scale=0.7]Web

Figure 1: *Factors 1 and 2 Web ranking . Source: own elaboration.*

Scimago *Ranking* (Sci): synthesized in the two factors shown in the figure 2, which include all its indicators and collect the 63% of the variance. The first exposes 44% of the variance, and is explained principally by the indicators of excellence, normalized impact and high-quality publications, which add 69% of the variance to the factor. If, in addition, the indicators of international collaboration and scientific leadership are included, it is obtained the 97% of the variance, although these final two are shared with the second factor. Besides, they add up in opposite directions over the plane, which

highlights certain tension between the international collaboration and the scientific leadership, confirmed by the negative correlation between them (-0.56). This face indicates how counterproductive international collaboration may result if the university is not the main contributor to the institutional production (figure 2).

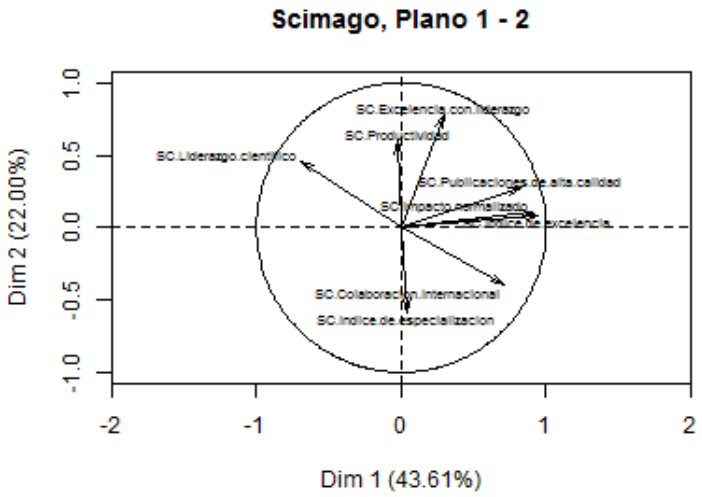

Figure 2: *Factors 1 and 2, Sci ranking Source: own elaboration.*

The second factor that explains the 22% of the variance is typified on its positive part by excellence with leadership and productivity, and on its negative part by the index of specialization in the scientific production, indicating that the latter goes against excellence with leadership and productivity. Nonetheless, this last negative relationship might be explained given the fact that the specialized production may tend to be lower depending on the size of the research group in a determined area.

Qs Ranking: a first analysis would indicate that might be synthesized in two factors that explain the 58% of the variance (see figure 3). The first indicates and associates the institutional reputation (among academic peers and among employers) with the number of professors holding a Ph.D. The second is dominated by the average of articles by professor and, again, by the two types of institutional reputation that the first factor contributes for in similar proportions. The vector bundles, almost perpendicular, show the independence among the corresponding indicators (see figure 3).

Nonetheless, in the former analysis the number of students per professor is left out -which typifies almost in a whole the third factor-, as well as the number of

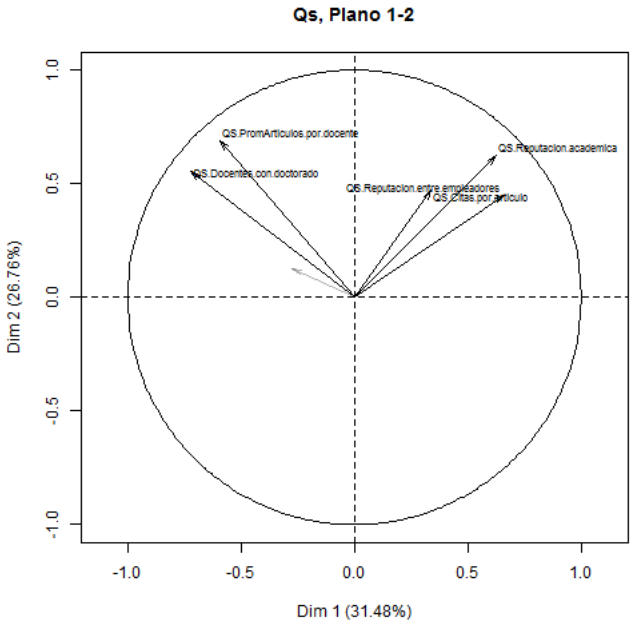

Figure 3: *Factors 1 and 2 Qs ranking. Source: own elaboration.*

citations per article, which typifies completely the fourth factor, gathering among the four factors the 90% of the whole cloud variance. The former analysis indicates the wide spectrum of the criteria that were taken into account in this *ranking*, at the point that it is impossible to reduce it to two factors without losing the valuable information of student density and the impact of the institutional production measured by the citations per article (see figure 4).

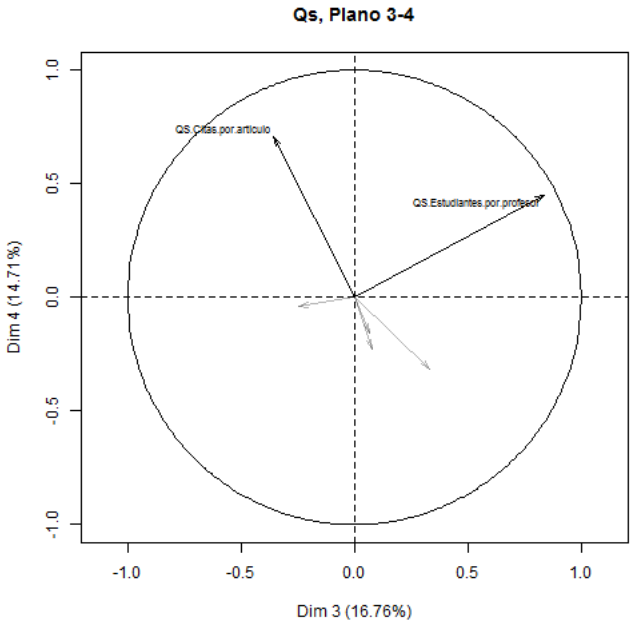

Figure 4: *Factors 3 and 4 Qs ranking. Source: own elaboration.*

Global analysis

For the interpretation phase, the correlations variable-factor were used as the main criterion for inclusion and factor nomination, as they allow to identify in every case at least one variable that correlates negatively with it, indicating its direction and the meaning of its negative values. As in every case, the positive part of the factor was the one that had more contributing variables, then the names of the factors are assigned, giving priority to such part. The first four factors of the global analysis are interpreted, which collect 68% of the variance and are used for the classification.

Factor 1: Excellence, presence in the web and productivity $\longleftrightarrow$ specialization in the scientific production. In this factor, the four indicators from Webometrics are combined, in the zone of positive values, with the indicator of institutional productivity (Scimago) and the indicator of professors' productivity. Scimago's indicator of scientific production specialization contributes to, correlates to, and is well represented in the zone of negative values, which indicates certain tension with the web presence and the productivity. The specialization indicator is shared with the factor 4, but this is not correctly represented (see the horizontal axis in figure 5).

Factor 2: international collaboration, excellence and institutional impact

$\longleftrightarrow$ scientific leadership. In the zone of positive values, this factor is typified by four Scimago indicators: international collaboration, excellence index, normalized impact, high quality publications, and the Qs indicator of citations per article. In the zone of negative values, the indicator of the Scimago *ranking* for scientific leadership opposes to modest contribution close to the half of contributions on the positive side, and not well represented (see vertical axis on figure 5)

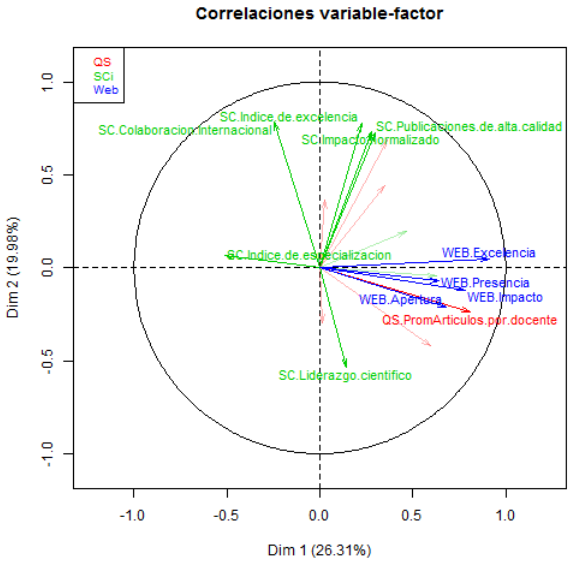

Figure 5: *Factors 1 and 2 of the global analysis. Source: own elaboration.*

Reputation, leadership $\longleftrightarrow$ Impact. This factor is characterized almost completely on its positive part by the two Qs indicators of reputation, while on its negative part is weakly defined by Scimago normalized impact indicator. This fact indicates cerain tension between the institutional reputation and the normalized impact (see plane 1-3 in figure 6). Nonetheless, on the plane 2-3 of the same figure, it is observed that although there is certain tension, in fact reputation and impact are independent, which is confirmed by the correlations between them: $r_{RepParesAca \times impacto} = 0.12$, $r_{RepEmpl \times impacto} = 0.09$.

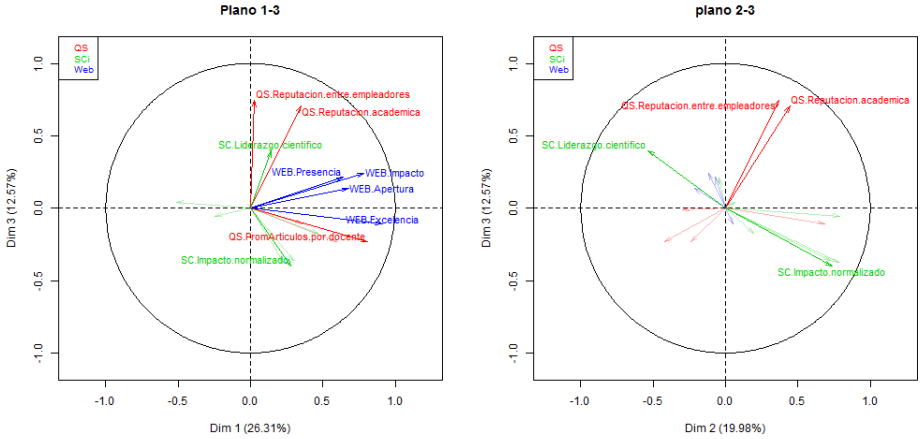

Figure 6: *Factor 3 (vertical) of the global analysis. Source: own elaboration.*

human capital, specialization $\longleftrightarrow$ low web presence. This factor correlates, in a high degree and with its positive part, the indicator of students per professor (0.8), the specialization index (0.44) and the indicator of professors holding a Ph.D. (0.33), and on its negative part with all the indicators of web presence impact (see figure 7).

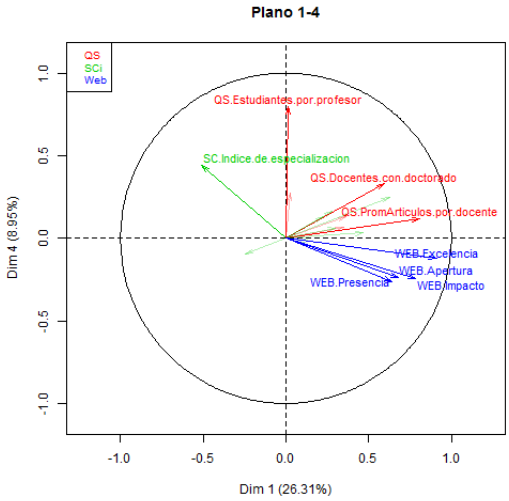

Figure 7: *Factor 4 (vertical) of the global analysis. Source: own elaboration.*

The fact that both Webometrics and Scimago define or typify a factor of global analysis means that they measure concepts relatively independent.

On the other hand, as the Qs indicators correlate both Webometrics and Scimago indicators -and also contribute to the characterization of the four factors-, it can be interpreted as an indicator of the diversity or dispersion of its criteria.

It is convenient to say that the correlations among the interpreted factors in the separated analysis from the three sets of variables $R_{Sci2\times Web1} = 0.58$, $R_{Qs2\times Web1} = 0.56$ $R_{Sci2\times Qs2} = 0.53$, y $R_{Sci1\times Qs3} = -0.32$ indicate that the criteria used in the construction of *rankings* are not very or barely related and, because of that, the institutions end up being differently classified in many cases. The former explains why (Spearman) correlations are so low among the positions occupied by universities in two *rankings*: $R_{(Web,Qs)} = 0.55$, $R_{(Sci,Qs)} = 0.46$; the high correlation $R_{(Web,Sci)} = 0.85$ is easily explainable as both *rankings* use criteria from the other.

Dimensionality and homothety (Lg coefficients)

These two characteristics are measured by the Lg coefficient, which is the degree of similarity or deformation with respect to a center (homothety) among the sets of indicators and when it is calculated for only one of the sets. It is known as the dimensionality indicator (see appendix). The value of the coefficient $Lg_{(Web)} = 1.05$ for Webometrics clearly indicates that its dimensionality is one, that is to say, it can be synthesized in only one factor; $Lg_{(Sci)} = 1.4$ indicates that Scimago might have dimensionality one or two, and the value $Lg_{(Qs)} = 2.3$ indicates the Qs has two -or even three- dimensions or factors that characterize it..

Lg coefficients among the sets of *rankings* indicators show, for example, that the Qs might partially share a factor with the Sci $Lg_{(Qs,Sci)} = 0.43$, and with the Webometrics $Lg_{(Qs,Web)} = 0.37$, whilst Sci and web are not very similar sets as $Lg_{(Sci,Web)} = 0.22$.

In addition, Lg coefficients of the three sets with the MFA: $Lg_{(AFM,Qs)} = 1.71$ indicates that the Qs shares up to two factors with the multiple factor; $Lg_{(AFM,Sci)} = 1.14$ and $Lg_{(AFM,Web)} = 0.92$ indicate that Sci and web share a factor with MFA. Finally, the $Lg_{(AFM)} = 2.09$ indicates that it can be synthesized in at least two factors.

lg coefficients with the supplementary variable $Lg_{(Pais,Qs)} = 0.98$, $Lg_{(Pais,Sci)} = 0.67$ y $Lg_{(Pais,Web)} = 0.33$ show that this could share at least one factor with the Qs group. On the other hand, with MFA, the coefficient $Lg_{AFM,Pais} = 1$ is also one, which indicates that the variable "pais" shares a factor with the first.

Association among the sets of indicators (Rv coefficients)

The association among indicator sets is measured with the Rv coefficient, which is an extension to matrixes of the Spearman correlation coefficient (defined in the appendix). Coefficients $Rv_{(Qs,Sci)} = 0.24$, $Rv_{(Qs,Web)} = 0.24$ and $Rv_{(Sci,Web)} = 0.19$ indicate that among the indicators of the matrixes formed by the universities with the three *rankings*, there are no meaningful associations, which means that the criteria used by the three *rankings* include topics that are little associated among themselves.

On the other hand, the coefficients among the sets and the MFA $Rv_{(AFM,Qs)} = 0.78$, $Rv_{(AFM,Qs)} = 0.66$ AND $Rv_{(AFM,Web)} = 0.61$ indicate that, even being not similar among themselves, the three *rankings* do have a considerable degree of association with the MFA, that is, the representation on the planes generated by the MFA is adequate.

Relationship of the separated analysis factors with the MFA.

Webometrics factor 1 (impact/institutional presence on the web), Qs factor 2 (articles per professor, reputation among academic peers and professors holding Ph.D.), and Scimago factor 2 (excellence-leadership-productivity) mostly correlate with the factor 1 of the MFA. The extreme positive values of this factor can be associated with presence-impact on the web, excellence, leadership, productivity, and reputation (see figure 8).

Scimago factor 2 (excellence, impact, high quality publications and international collaborations) correlates in first place with MFA factor 2, followed by Qs factor 1 (reputation among employers and academic peers). In such a way, positive values of this factor are indicators are indicators of desirable characteristics for a university, accompanied generally by good reputation.

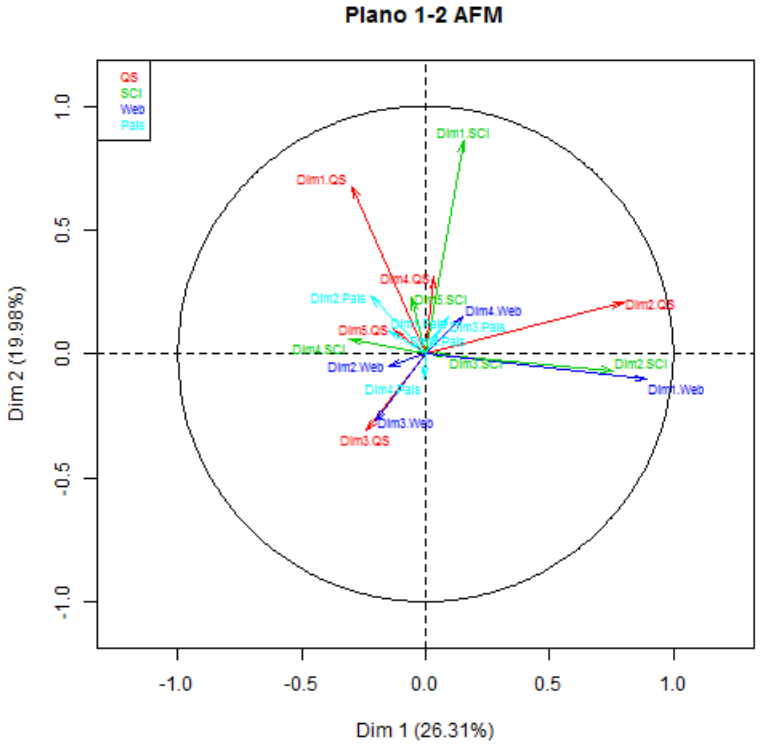

Figure 8: *Projection of the axes of partial analysis on factors 1 and 2 of MFA. Source: own elaboration.*

For MFA factor 3, the biggest correlation comes from Qs factor 1; that is why the positive values of such indicator are indicators of institutional reputation and education for professors. For MFA factor 4, the main contributor is Qs factor 3, which is the indicator of student density in the university. From the former, it can be deduced that the influence of Qs *ranking* is enough to dominate two of the MFA, and corroborates the spectrum width covered by its indicators on its particular construction on the university *ranking*.

Relationships of the variable sets (rankings) with the MFA

With any of the criteria to be used (coordination, contribution, cos2, correlation), Webometrics associates almost exclusively the first factor. Scimago, on the other hand, has the higher values of all criteria regarding factor 2, and can be moderately associated with factor 1 also for its high correlation with it. Finally, Qs has very low square cosines in all the factors (it is not well represented in any of them), and its other three indicators are slightly higher in factor 3. In a nutshell, *web* group dominates exclusively on factor 1, *Sci* group takes control over factor 2, factor 1 is shared between them, and *Qs* cannot be clearly associated with any of the MFA factor as dominant.

4.2 Results of the classification

For the classification of the built classes, it was included the following two indicators of class average comparison ($\bar{x}_{clase}$) with respect to the average of the complete group $\bar{X}$, and those of the standard deviation (σ_{clase}) of the class with the standard deviation of the complete group σ from the 150 universities included in the exercise:

Indicator of the class average with respect to the general average:

$$I_{\bar{x}} = \frac{\bar{x}_{clase}}{\bar{X}}$$

It indicates if a class surpasses ($I_{\bar{x}} > 1$) or not ($I_{\bar{x}} < 1$) the general average of all the universities that have been analyzed.

Indicator of standard deviation of the class with respect to the general standard deviation:

$$I_{Sd} = \frac{\sigma_{clase}}{\sigma}$$

It indicates if a class is more heterogeneous $I_{Sd} > 1$ or more homogeneous $I_{Sd} < 1$ ten the complete group of the universities that have been analyzed.

For each class, the values for $I_{\bar{x}}$ and I_{Sd} are shown for those ranking indicators which the test value is

$$\frac{\bar{x}_{clase} - \bar{X}}{\sigma_{clase}} > 2$$

Class 1. Very specialized universities, with good reputation among employers, high student density, and low productivity. The universities in this category have averages that are above of the complete group of the 150 universities in the indicators of specialization, reputation among employers, and number of students per professor; nonetheless, their averages are below the group of the 150 regarding

Scimago productivity index, articles per professor (Qs), and publications in high impact journals (Webometrics). Regarding their dispersion, this class in general is less disperse than the complete group, mainly in terms of its institutional productivity (Sci) and productivity per professor (Qs).

Table 1: *indicators of the class 1 average. Source: own elaboration.*

	$I_{\bar{x}}$	I_{Sd}
SC.Specialization index	1.71	1.01
QS.Reputation among employees	1.22	0.74
QS.Students per professor	1.19	1.06
SC.High Quality Publications	0.80	1.05
SC.Normalized impact	0.79	0.79
WEB.Openness	0.76	1.04
SC.Excellence index	0.74	0.81
WEB. Impact	0.74	0.91
WEB.Presence	0.72	1.05
QS.Quotations per article	0.72	0.89
SC.Excellence with leadership	0.68	0.98
QS.Professors holding Ph.D.'s	0.65	0.82
WEB.Excellence	0.49	0.75
QS.Prom. Articles per professor	0.27	0.32
SC.Productivity	0.12	0.10

From the 38 universities in this category, 29% are from Argentina, as shown in table 2: 2:

Table 2: *number of class 1 universities per country. Source: own elaboration.*

	1	2	3	4	5	6	7	8	9	10
Pais	ARG	COL	CHL	MEX	VEN	BRA	PAN	PER	URY	CRI
Freq	11	6	5	4	3	2	2	2	2	1

Aiming at highlighting extreme characteristics of the categories, the first are calculated for the indicators of *ranking* that are obtained -for example, for the specialization indicator "Scimago SC.Specialization.Index"-, and the same name is kept in order to facilitate the notation in any university of the class, just as follows:

$$\text{SC.Specialization.Index} = \frac{\text{SC.Specialization.Index}}{\bar{x}_{\text{category in specialization.index}}}$$

Similarly, the average indicator for "SC.Productivity" is calculated as follows:

$$\text{SC.Productivity} = \frac{\text{SC.Productivity}}{\bar{x}_{\text{SC.Productivity}}}$$

This indicators are calculated, from now on, for all the categories and for the five universities that have the higher values of the indicator $I_{\bar{x}}$ in the category.

The universities with the highest specialization indexes, comparedto the class 1 average and its corresponding indicators of productivity average are shown in table 3:

Table 3: *Las más especializadas del grupo 1. Fuente: elaboración propia.*

	SC.Specialization index	SC.Productivity
UNIVERSIDADDECIENCIASEMPRESARIALESYSOCIALES(UCES) - ARG	1.65	0.03
EXTERNADO DE COLOMBIA - COL	1.63	0.22
TORCUATO DI TELLA - ARG	1.58	0.42
COLEGIO DE MEXICO - MEX	1.58	0.54
ALBERTO HURTADO - CHL	1.58	0.47

Colombian universities in class 1: the six Colombian universities that are included in this category are private and shown in table 4 with five of the $I_{\bar{x}}$ indicators that distinguish this class.

Table 4: *Colombian universities in class 1. Source: own elaboration.*

	Sc.Espcial	Sc.Productv	Qs.Reput.Empl	Qs.EstxProf	Qs.ArtxDoce
PONTIFICIA BOLIVARIANA - COL	0.07	1.63	1.04	0.64	1.43
EAFIT - COL	0.29	1.03	1.07	0.45	2.02
LA SABANA - COL	0.79	0.80	1.28	0.76	1.37
LA SALLE - COL	1.15	0.36	0.99	0.37	0.29
JORGETADEOLOZANO - COL	1.33	0.41	1.09	0.54	0.77
EXTERNADO DE COLOMBIA - COL	1.63	0.22	1.25	0.82	0.32

In the first two columns, it is noticeable how the indicator of productivity (SC. Productv) decreases as the specialization index increases. Besides, it is observed that EAFIT, UPB and La Sabana universities surpass the average of the 150 that were analyzed in the indicator of number of articles per professor (Qs.ArtxDoce), and at the same time, those universities have the lowest specialization scores, while the universities with highest specialization indexes are Externado, Jorge Tadeo Lozano, and La Salle, which also happen to have the lowest indexes of articles per professor (Qs.ArtxDoce) and institutional productivity (Sc.Productivity). The former confirms, at least for the Colombian universities, the type of tension that had already been notices in the first factor of the global analysis.

Class 2. Excellence universities, impact, international collaboration, citations per article, high quality publications, and low scientific leadership.
The universities in this category double in the excellence index the average of the 150, and their indicators of impact, international collaboration, citations per article, high-quality publications, and specialization index are between 41% and 76% above the group of the 150 institutions analyzed. Among their weak points

it is possible to highlight the scientific leadership indicator and three out of four Webometrics indicators. The category is more heterogeneous than the group of the 150 regarding the indicators of excellence, normalized impact, and scientific leadership, which indicates that some of the indicators of universities included may be high.

Table 5: *Indicadores de promedio de la clase 2. Fuente: elaboración propia.*

	$I_{\bar{x}}$	I_{Sd}
SC.Excellence.index	2.03	1.48
SC.Normalized.Impact	1.76	1.48
SC.International.Collaboration	1.73	0.85
QS.Citation.per.article	1.58	0.78
SC.High.Quality.Publications	1.49	0.87
SC.Specialization.Index	1.41	1.00
WEB.Impact	0.87	0.69
WEB.Openness	0.84	1.10
WEB.Presence	0.84	0.95
SC.Scientific.Leadership	0.69	1.46

In this category, there is no dominant country, as seen on the following chart.

Table 6: *Number of class 2 universities per country. Source: own elaboration.*

	1	2	3	4	5	6	7	8	9	10
Pais	CHL	ECU	MEX	URY	ARG	BOL	COL	PER	PRI	PY
Freq	3	2	2	2	1	1	1	1	1	1

The five universities with the highest excellence index in this category, together with the values of average indicators within the category, are:

Table 7: *class 2 excellence. Source: own elaboration.*

	SC.excelencia	SC.Lid.cientifico
IBEROAMERICANA (UIA) - MEX	1.86	0.76
SAN FRANCISCO DE QUITO - ECU	1.84	0.34
DEL NORTE - COL	1.35	0.00
TECNICA FEDERICO SANTA MARIA - CHL	1.23	1.26
MONTEVIDEO - URY	1.12	1.10

Colombian universities in class 2: in this group, the only university is Universidad del Norte, with scores above the average of the complete group in excellence, impact, and international collaboration, and with a high specialization index.

Table 8: *Class 2 Colombian Universities $I_{\bar{x}}$Indexes . Source: own elaboration..*

	Ind Excel	Impac Nor	Colab Int	Citas x Art	Publ calid	Specializa
DEL NORTE - COL	1.35	1.21	1.31	0.26	0.65	2.16

Class 3. Universities with productive professors and Ph.D., high-impact produc-

tion, low values in the reputation indexes. It can be typified as the average of articles per professor and the number of professors holding a Ph.D. is between 38% and 35% above the complete group of universities, but they do not have good reputation neither among academic peers nor among employers. Tis category is the most homogeneous in the whole group, which indicates that the universities that belong to it are very similar among them in the three criteria used by the three *rankings*.

Table 9: *Average indicators of Class 3. Source: own elaboration.*

	$I_{\bar{x}}$	I_{Sd}
QS.PromArticles.Per.Proffesor	1.38	0.75
QS.Professors.with.PhD	1.35	0.71
WEB.Excellence	1.10	0.41
WEB.Openness	1.09	0.54
SC.High.quality.publications	0.90	0.60
SC.normalized.impact	0.88	0.46
SC.Excellence.index	0.81	0.43
QS.citations.per.article	0.75	0.63
SC.specialization.index	0.69	0.61
SC.International.Collaboration	0.67	0.65
QS.Reputation.among.employers	0.59	0.78
QS.Academic.reputation	0.54	0.59

In this category, dominated by Brazilian universities, there are universities from only five countries:

Table 10: *number of universities per country in class 3. Source: own elaboration.*

	1	2	3	4	5
Pais	BRA	MEX	CHL	ARG	VEN
Freq	27	7	5	2	1

The top five universities in this class show a meaningful increase in reputation with a moderate decrease in the average of articles per professor, as observed in table 11.

Table 11: *Universities with most productive professors in class 3. Source: own elaboration.*

	QS.PromArt.x.docente	QS.Rep.academica
FEDERAL DE LAVRAS - BRA	1.64	0.59
ESTADUAL DO NORTE FLUMINENSE - BRA	1.63	0.22
FEDERAL DE VIÇOSA - BRA	1.63	1.10
LA FRONTERA (UFRO) - CHL	1.59	1.46
FEDERAL DO PERNAMBUCO - BRA	1.54	2.14

Colombian universities in class 3: there is no Colombian university in this class.

Clase 4. Universities with good reputation among academic peers and employers, with impact on the number of citations, web presence, international collaboration, lowly specialized, low index of professors holding a Ph.D., and low student density. This group is pretty much more homogeneous than the complete group of universities that were analyzed, and the academic peer reputation indicator moderately stands out.

Table 12: *indicators of class 4 average. Source: own elaboration.*

	$I_{\bar{x}}$	I_{Sd}
QS.Academic.reputation	1.34	0.61
QS.Citations.per.article	1.19	0.81
WEB.Presence	1.18	0.39
QS.Reputation.among.employers	1.17	0.73
SC.International.Collaboration	1.14	0.66
WEB.Excellence	1.14	0.45
WEB.Impact	1.13	0.53
WEB.Openness	1.12	0.47
QS.Students.per.professor	0.70	0.72
QS.Professors.with.PhD	0.66	0.75
SC.Specialization.Index	0.59	0.58

In this category, as well as in category one, there is a majority of Argentinean universities:

Table 13: *number of universities per country in class 4. Source: own elaboration.*

	1	2	3	4	5	6	7	8	9
Pais	ARG	MEX	COL	CHL	VEN	CRI	PER	CUB	ECU
Freq	10	6	5	4	3	2	2	1	1

The top five universities in this group are:

Table 14: *Class 4 best academic reputation universities. Source: own elaboration.*

	QS.Reputacion.academica	SC.Indice.de.especializacion
BUENOS AIRES - ARG	1.26	0.94
NACIONAL DE LA PLATA - ARG	1.25	0.10
PONTIFICIA JAVERIANA - COL	1.24	0.37
NACIONAL DE CORDOBA - ARG	1.24	0.63
SANTIAGO DE CHILE (USACH) - CHL	1.24	1.62

Colombian universities in group 4: in this group, there are three public (state) universities and one private:

Table 15: *Colombian universities in class 4. Source: own elaboration.*

	ANTIOQUIA - COL	PONTIFICIA JAVERIANA - COL	DEL ROSARIO - COL	DEL VALLE - COL	INDUSTRIAL DE SANTANDER - COL
Rep Aca	1.19	1.24	0.82	1.10	0.77
Qs Cit x Art	1.19	0.60	1.31	0.59	0.89
Web Presen	1.01	0.99	1.08	1.00	0.84
Qs REp Empl	1.19	1.40	1.34	0.86	1.03
Sc Col Int	1.11	0.98	0.86	1.08	1.01
Web Excel	1.13	0.92	0.88	0.99	0.96
Web impct	1.05	1.02	0.69	0.92	0.75
Web Apert	1.06	1.01	0.90	1.03	1.04

It is observed that Universidad de Antioquia has most of its indicators above the average, with the exception of reputation among academic peers and employers, which are higher for Universidad Javeriana.

Class 5. Elite universities: productive, productive professors and with Ph.D. education, reputation among academic peers, excellence and leadership, and with low indexes of international collaboration and specialization. In this group, the Scimago productivity indicator highlights, and is almost four times the average of the 150, and the Qs indicator of professors' (Qs.PromArticulos.por.docente) productivity almost doubles the average of the complete group of the analyzed universities. Besides, with the exception of the indicators of specialization and international collaboration, all the other indicators surpass the average of all universities.

Table 16: *Average indicators of class 5. Source: own elaboration.*

	$I_{\bar{x}}$	I_{Sd}
SC.Productivity	3.73	1.97
QS.PromArticles.per.professor	2.00	0.53
QS.Professors.with.PhD	1.63	0.38
QS.Academic.Reputation	1.49	0.54
WEB.Excellence	1.43	0.30
SC.Excellence.with.leadership	1.28	0.58
QS.citations.per.article	1.28	0.75
WEB.impact	1.27	0.29
QS.reputation.among.employers	1.26	0.86
WEB.Presence	1.25	0.39
QS.Students.per.professor	1.25	0.74
WEB.openness	1.20	0.28
SC.International.collaboration	0.81	0.62
SC.Specialization.index	0.70	0.66

In this category, as in class 3, Brazilian universities are top:

Table 17: *number of universities per country in Class 5. Source: own elaboration.*

	1	2	3	4
Pais	BRA	CHL	COL	MEX
Freq	14	3	2	2

The top 5 universities of group 5 are:

Table 18: *most productive from class 5. Source: own elaboration.*

	SC.Productividad	SC.Indice.de.especializacion
DE SAO PAULO (USP) - BRA	4.74	1.67
NACIONAL AUTONOMA DE MEXICO (UNAM) - MEX	1.92	1.32
UNIVERSIDADE ESTADUALPAULISTA "JULIODEMESQUITAFILHO" - BRA	1.71	0.00
ESTADUAL DE CAMPINAS (UNICAMP) - BRA	1.67	2.24
FEDERAL DO RIO DE JANEIRO - BRA	1.49	2.06

Colombian universities in group 5: two universities are included in the elite group:

Table 19: *Colombian universities in group 5. Indicators $I_{\bar{x}}$. Source: own elaboration.*

	LOS ANDES COLOMBIA - COL	NACIONAL DE COLOMBIA - COL
Sc prod	0.23	0.57
Qs Art x Doc	0.98	0.83
Qs Doc Drado	0.91	0.75
Qs Rep Aca	1.13	1.13
Web Excel	0.96	0.94
Sc Excel Lid	0.73	0.71
Qs Cit x Art	1.33	0.39
Web Impct	0.93	1.05
Qs Rep Empl	1.31	1.31
Web Presen	0.91	1.08
Qs Est x Prof	0.80	0.65
Web Aprtur	0.96	1.04

5 Conclusions

The nuances that distinguish the classification let see the diversity of institutional profiles and discover some paradoxes about indicators or criteria included in the *rankings* and that were used in this exercise. To begin, the fact that in class 1 it is possible to find universities with a high index of specialization and low productivity, whilst in class 5 there are universities which have low specialization indexes, shows how counterproductive that the grade of specialization may be in order to have good scores in the productivity indicators. From the former, several interpretations are possible, for example, that the range of impact of the specialized scientific production is much more reduced than that of universities with less concentrated production in some topic fields, or that the institutions that publish about specialized topics are less productive.

In class 3, it is possible to find productive universities with teachers that hold Ph.D. degrees and have high-impact production, but that also have bad reputation. That is when comes up the question about the pertinence of opinion indicators -such as the Qs reputation among academic peers and employers- when what it is expected is to classify universities by their academic and scientific activity.

Two more paradoxes come up from classes 2 and 4: from class 2, it can be concluded that scientific leadership, measured by the percentage of production of an institution in which the last id the main contributor, is not necessarily a criterion of excellence or impact and, additionally, that even having high values in the index of international collaboration, an institution may have high-quality publications without performing scientific leadership. Class 4, on the other hand, it can be deduced that Ph.D. education or low student density are not necessarily factors of reputation, impact, or institutional collaboration.

The first factor of global analysis shows certain tension between the degree of institutional specialization and excellence indicators, presence on the web, and professors productivity, explainable in part by the fact that high degrees of specialization in a topic field reduce the circle of impact to researchers in the same field.

The second factor shows certain paradox when opposing scientific leadership - percentage of institutional production in which the institution is the main contributor- to impact indicators, production quality, institutional production and professor production, as if leadership were counterproductive for a good score in such indicators.

The projection over the factors two and three of the indexes of leadership, impact, and institutional reputation shows a negative association between the first two and, at the same time, shows them as independent of institutional reputation.

Regarding the fourth factor, it is important to note the fact that student density is a criterion that does not alter none of the indicators, with which it can be stated that, for university *rankings* on academic and research activity, such factor would not be of interest, at least not for the analyzed universities.

The dimensionality of the three *rankings* (one for Webometrics, two for Scimago, and four for Qs) is an indicator of diversity of the criteria used in the *rankings*, whilst the low correlations among the factors produced by the isolated analyses indicate that none of them classifies integrally the universities.

Received: April 11, 2016
August 16, 2016

References

Escofier, B. P. J. (2008), *Analyses Factorielles Simples et Multiples*, 4 edn, DUNOD, Paris, ISBN 978-2-10-053809-6.

Escoufier, Y. (1973), 'Le traitement des variables vectorielles', *Biometrics* **29**, 751–760.

Herve Abdi, L. J. W. & Valentin, D. (2013), 'Multiple factor analysis: principal component analysis for multitable and multiblock data sets', *WIREs Comput Stat 2013* **doi: 10.1002/wics.1246**.

Margison, S. & Ordorika, I. (2010), *Hegemonía en la era del conocimiento: competencia global en la educación superior y la investigación científica*, 2 edn, Academic Press, Universidad Nacional Autónoma de México.

Martinez Rizo, F. (2011), 'Los rankings de universidades: una visión crÃtica', *Revista de la educación superior* **XL(I)**(157), 77–97.

Sanchez, J. & Moreno, G. (2011), *Caracterización de rankings universitarios*, 1 edn, Universidad Nacional de Colombia, sede Bogotá, Universidad Nacional de Colombia, sede Bogotá.

APPENDIX 1. SOME IMPORTANT CONCEPTS.

Rv coefficient: defined by Escoufier (1973). It is a measure of similarity between two semi-defined positive matrixes (sdp), and corresponds to an extension for matrixes of the Spearman correlation coefficient. Be $X_{N,J}$ and $Y_{N,K}$ real matrixes. Then, matrixes $S = XX'$ and $T = YY'$ are sdp. Coefficient R_v between S and T is defined by:

$$R_v = \frac{Traza(S'T)}{\sqrt{Traza(S'S)Traza(T'T)}}.$$

In MFA, matrixes have the same number of lines; in consequence, homothecy can also be calculated by taking the matrixes of the co-variances $S = X'X$ and $T = Y'Y$, That is why it can be interpreted as the sum of the products of the variances and co-variances of each group of variables.

Lg coefficient: a measurement of the relation between pairs of tables that indicates how many characteristics are shared between them in a homothetic sense[4]. The more related the variables of one are to the variables of the other, the bigger the Lgcoefficient. It is defined by:

$$L_g = \frac{Traza(S'T)}{\alpha_1^2 \times \lambda_1^2}.$$

in which α_1 and λ_1 are the first own value for S and T respectively. See Herve Abdi & Valentin (2013).

[4]Projection with respect to a certain point of a (homothetic) figure in bigger or smaller proportions, in a direct or inverse way.