Technological Forecasting & Social Change 123 (2017) 153–160

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Is measuring the knowledge creation of universities possible?: A review of university rankings

Gokcen Arkali Olcay [a,*], Melih Bulu [b]

[a] School of Management and Administrative Sciences, Istanbul Sehir University, Altunizade Mah. Kusbakisi Cad. No: 27, 34662 Uskudar/Istanbul, Turkey
[b] Faculty of Economics and Administrative Sciences, Istinye University, Turkey

ARTICLE INFO

Article history:
Received 30 November 2015
Received in revised form 28 March 2016
Accepted 29 March 2016
Available online 7 April 2016

Keywords:
University rankings
Universities
Ranking indicators
Reputation

ABSTRACT

University ranking indexes are considered very useful benchmarking tools in comparing the performance of universities around the world. Being placed in these prestigious indexes provides a strong advertisement for a university and helps them to attract high-quality students and academicians all over the world. However, there are some important deficiencies of university ranking indexes such as taking into account the whole university as a single unit without differentiating according to different fields of study or research, being limited to some well-known universities, and not considering institutional characteristics such as size or age. This study aims to explore the leading global university rankings to determine the similarities and differences in terms of their ranking criteria, main indicators, modeling choices, and the effects of these on the rankings. Designating the Times Higher Education World Rankings as the base ranking, a comprehensive comparison of the positions of the top universities of the base index with the matched positions of the same universities under other leading indexes including ARWU, QS, Leiden, and URAP is given. Correlations highlight the significant differences among some indexes even in measuring the same criterion such as teaching or research.

1. Introduction

Universities play a central role in the development of societies across the world with their teaching and research missions for centuries. While carrying out these missions, they also create growth strategies and play significant roles in raising the employment of graduates, increasing the education level of society, creating opportunities for individuals, and the development of knowledge and technologies. In this sense, universities develop strategies to fulfill their historic mission of teaching and research and they also undertake a significant role in producing and diffusing new knowledge in today's ever-changing world. Etzkowitz and Leydesdorff (1999) impose a new function of facilitating research and technology transfer on universities in their popularized model of the "triple-helix" (university–industry–government). Benneworth et al. (2009) conceptualize universities as knowledge-explorers, being one of the two sub-systems of regional innovation systems wherein firms form the other sub-system, i.e., the knowledge-exploiters, complementing and interacting with universities, resulting in new regional innovative capabilities.

Given the significant role of universities in the development of societies, measuring and assessing the universities' performances becomes crucial for various stakeholders, including government, industry, and society. University league tables are published each year in the UK in leading newspapers using the statistical data from central Higher Education Statistical Agency, the national funding agencies, and the national Quality Assurance Agency mainly to guide prospective students in their choice of future enrollment (Eccles, 2002).

The world's most prestigious universities have been annually ranked by popular ranking systems such as UK's Times Higher Education (THE) World University Rankings and Quacquarelli Symonds' (QS) World University Rankings starting in 2004. Since 2003, Shanghai Ranking Consultancy and Center for World-Class Universities of Shanghai Jiao Tong University publish annually the Academic Ranking of World Universities (ARWU). CWTS Leiden Ranking is another emerging study published by the Centre for Science and Technology Studies of Leiden University. While many of these international rankings, especially THE World and ARWU, confirm the US universities' leading role among other universities of developed countries, there also exist more than 30 national rankings employed around the world (Saisana et al., 2011).

Having achieved higher rankings in any one of these so called "prestigious" ranking systems is crucial for the university management as they publish this as news or reports in their brochures, catalogs, and annual reports to attract better students and faculty, and increase their public and private funding (Hazelkorn, 2008; Shin and Toutkoushian, 2011). However, many of the good quality universities are left out of the top lists because they are young, focus on a few fields, or are non-English speaking universities (van Raan, 2005; Harvey, 2008). Times Higher Education released the global university rankings for under 50

* Corresponding author.
E-mail addresses: gokcenarkali@sehir.edu.tr (G.A. Olcay), mbulu@istinye.edu.tr (M. Bulu).

http://dx.doi.org/10.1016/j.techfore.2016.03.029
0040-1625/

in 2012 (Soh, 2013) claiming older universities have a wider and deeper alumni network and reputation, biasing the results in favor of these universities (reported in THE, 2015). Similarly, ARWU started releasing global rankings according to the broad subject fields starting in 2007 in order to meet the diversified needs of various stakeholders (Shanghai Ranking Consultancy and Center for World-Class Universities of Shanghai Jiao Tong University, 2015).

Moreover, university rankings also diversified over time such as new rankings focusing on only one criterion which were developed by leading indexes. It is easier to employ research indicators such as counting indexed publications or citations that depend on hard data since measuring teaching is not as straightforward as measuring research. Because teaching indicators are mainly dependent on reputational surveys or data provided by the universities, new rankings such as Leiden Ranking has emerged employing a methodology emphasizing more transparent indicators based on research (Centre for Science and Technology Studies, Leiden University, 2015).

Although university ranking systems have improved and adapted themselves over time, they are generally deficient in responding to different needs of the users in terms of specialized rankings across regions, fields, or subjects with objective measures of research and teaching criteria. Also, these rankings do not adequately reflect academic excellence to the majority (Hurtado, 2012). Moreover, many stakeholders question how comprehensive the global rankings are given that the same universities are repeatedly chosen as the highest performers year after year (Lincoln, 2012).

All these issues point to question the role rankings play in measuring the quality of higher education systems on one hand, and on another hand, how beneficial these ranking systems are to all users since currently this is the only tool excessively used by all stakeholders in measuring the performance of higher education institutions. This issue is clearly related to the indicators and the methodology of the existing leading global ranking systems. Thus, a need emerges to understand the similarities and differences among the ranking systems in terms of both the chosen indicators and data. Their transparency and reflections as to which universities appear in the rankings can then be evaluated.

The remainder of the study is organized as follows. Section 2 provides conceptual arguments to understand the role ranking indexes undertake in measuring higher education quality, while giving a synopsis of the university rankings all over the world. In Section 3, we first determine the main criteria measured by the leading global rankings, grouping the used indicators of the chosen rankings under these criteria. We then compare the positions of the best universities of a chosen base ranking with the matched positions of the same universities under other leading ranking indexes. Lastly, we elaborate on the correlations of the universities' positions across different rankings to put forward the strong and weak points of such rankings and make suggestions for the decision-makers and users of these rankings in Section 4.

2. Ranking indexes in measuring higher education quality

The massification of higher education, increased competition at the national and international levels, and internationalization of higher education created the public concern for measuring the quality of such institutions and as a result the spread of the university rankings has accelerated since the 1990s (Teichler, 2011). While university rankings are one of the essential ways of measuring the quality of higher education, quality measurement in higher education is a multi-dimensional problem that cannot be based solely on rankings.

First of all, defining quality within the context of higher education institutions is challenging, as quality relates to frequently conflicting objectives of meeting or exceeding expectations in two primary functions of higher education institutions: teaching and research. While many institutions in the UK, Germany, South Korea, etc. adapted the American model (so called post-Humboldtian model) of combining research and teaching within the same university, performing well in one function might well result in lower performance in the other, highlighting the difficulty of achieving a balance in both (Shin and Toutkoushian, 2011). Second, measuring quality is another challenge, as there exist various indicators that can be used to measure teaching, research, and service quality in addition to a variety of sizes of institutions, weightings of indicators, and disciplinary, and regional differences of underlying institutions.

2.1. Quality measurement in higher education systems

University rankings emerged as a response to the needs of policymakers, higher education institutes, academicians, and the general public since the beginning of the 1980s when media and research institutions across the world began releasing improved and specified versions of rankings. University rankings are definitely a critical criterion in decision making for various stakeholders, yet there are possible negative side effects of rankings (van der Wende and Westerheijden, 2009; Dill, 2000; Shin and Toutkoushian, 2011). Many university executives focusing on raising their rankings in leading indexes face loosing mission diversity (van der Wende and Westerheijden, 2009).

Given the drawbacks of university rankings, Shin (2011) draws attention to the other mechanisms of quality assurance and accountability along with rankings in measuring organizational effectiveness. Many universities' performance has long been measured by external agencies such as the American Assembly of Collegiate Schools of Business (AACSB) by applying principles of quality management used in the US (Mergen et al., 2000). While many universities adapt the voluntary accreditation mechanisms in North America, many other countries including UK, New Zealand, Sweden, and Hong Kong have been employing new forms of academic accountability, the so called “academic audits” in order to assure the quality of learning and standardization of the degrees offered (Dill, 2000). Östling (1997) draws attention to the significance of academic audits on focusing on the quality of work but not on quality of outcomes since work process is one of the three elements of standardization along with input skills and output (Mintzberg, 1979), which is not really emphasized in many quality assurance mechanisms.

In comparing accountability, quality assurance, and ranking methods, Shin (2011) states the primary goal of rankings is to provide information to their target customers, mainly students, parents, and higher education institutions while on the other hand quality assurance and accountability mechanisms focus on improving quality and financial accountability. In line with this, Shin and Toutkoushian (2011) suggest that future directions of quality measurement in higher education should be combining these mechanisms in order to contribute to enhancing institutional performance in addition to providing information to the target readers of such rankings. A hybrid system embedding quality assurance and accounting mechanisms into ranking would be specific at the country level given the national quality assurance and government styles of the underlying country. However, a global university ranking system summarizes the “quality” of the institution with one metric easy to understand by various stakeholders at any level resulting in the popularization of rankings internationally over the last few decades.

2.2. Rising trend of rankings in measuring higher education quality

Teichler (2011) refers to the prominent role of university rankings in the higher education arena becoming more global and stratified, demanding higher quality in teaching, increased research productivity, and better use of resources. There is no doubt university rankings gained a central place in measuring higher education quality where many media or institutional based rankings attempted to provide better rankings at national and international levels. Among these newly introduced rankings each year, a few of them remained to be the leading ones, while there is little theoretical guidance on the variability of

indicators used and their associated weights (Shin and Toutkoushian, 2011).

University rankings measure quality by measuring institutional performance (i.e., teaching, research, and service) given the institutional characteristics such as mission, size, and region although there exists methodological issues that must be addressed in various ranking indices (Shin and Toutkoushian, 2011). Chen and Liao (2012) view the university rankings having carried the competition among universities from an under the table to a paper-based one forming some kind of criteria for measuring educational quality for their various stakeholders. The prior literature (Safon, 2013; Docampo, 2011; Li et al., 2011; Usher and Savino, 2006) points out the existence of a hidden factor in the most influential rankings that has not much to do with quality. In an attempt to determine the hidden factor, Safon (2013) investigated the two most influential global university rankings and concluded that rankings do not have the capacity to assess university quality in all dimensions however they do have different conceptions of university quality.

Comparative studies of global university rankings across employed indicators reveal that research quality is prominently measured through scientific productivity and research impact in many of the rankings (Buela-Casal et al., 2007; Chen and Liao, 2012; Dehon et al., 2009). Despite the shortcomings in covering all dimensions of measuring quality, rankings are likely here to stay as Kivinen et al. (2013) mentioned that universities relate even their existence to whether they enter into the rankings among the ca. 16,000 universities around the world, where only one tenth are recognized by the ranking systems.

3. Major global university rankings

3.1. Main indicators and methodologies of the five leading rankings

QS World University Ranking, THE World Ranking, ARWU, and CWTS Leiden Ranking are the leading university rankings frequently encountered in articles, newspapers, and promotional publications. There exists web-based rankings such as Webometrics (i.e., Ranking Web) and other research based rankings from Turkey, Taiwan, and Australia (Holmes, 2012). Webometrics ranking has been excluded from our analysis since about a half of the indicators it depends on requires visibility to data that were not measured in any other leading indexes.

As for the other prominent research rankings, University Ranking by Academic Performance (URAP) is one of those established in 2009 by the Informatics Institute of Middle East Technical University, which is also included in our analyses. URAP covers more than 2000 universities around the world including many universities not represented in the current leading ranking systems. This allows many institutions to measure their relative locations among the institutions all the over the world based on a multi-criteria ranking system. Thus, with the inclusion of URAP, five leading ranking indexes have been addressed in this study.

Reviewing the performance indicators of the leading university ranking indexes, we grouped the indicators under eight main criteria comprising all criteria used in different indexes through merging some under more general criteria. The criteria of the leading indexes are shown in the first column of Table 1a. Teaching, research, citations, quality of education, quality of faculty, international outlook, and industry outcome are the main criteria in assessing university excellence. Each ranking system measures a different aspect or aspects of these criteria using various indicators as shown in the table. The indexes do not necessarily measure all the criteria we determined here. Besides, in ARWU one indicator is categorized as other to reflect to the size of the university, which is basically the weighted scores of the used indicators to the number of full-time equivalent students.

Weights of the indicators are provided as percentages in parentheses in Tables 1a and 1b. The indicators and their associated weights differ substantially among the various ranking indexes, as it can be seen in Tables 1a and 1b. While citations have been taken into account in all indexes, publications as a total number or proportion have not been assessed in QS World ranking. The quality of education and faculty are the missing criteria in the THE World ranking. ARWU, on the other hand, does not measure teaching, international outlook, or industry outcome in their rankings. Similarly, QS World does not also measure industry income in any form.

Another notable point, which is also perceived as a significant issue by many (Bowman and Bastedo, 2011; Buela-Casal et al., 2007), is the measurement of some criteria through the reputation surveys, questioning the objectivity of rankings. THE World uses surveys for measuring teaching and research with associated weights of 15% and 18%, respectively. For measuring teaching and faculty qualities, THE World employs the so-called Academic Reputation Survey, which explores the perceived prestige of the universities in teaching and research excellence among their peers. Academic Reputation Survey is employed by experienced and published scholars of various disciplines in order to form their opinions on both their institutions and others that they are familiar of. QS World also puts a high weight of 40% on the

Table 1a
Main criteria and the associated weights of the indicators employed in major university rankings.

Criterion	Main indicators		
	THE World	ARWU	QS World
Teaching	Reputation survey (15%) Staff to student, doctorate to bachelor's, doctorates awarded to academic staff ratios, & institutional income (15%)		Student to faculty ratio (20%)
Research	Research productivity (6%) Reputation survey (18%) Research income (6%)	Papers published in Nature and Science (20%)	
Citations	Citations of published work (30%)	Papers indexed in SCI (-expanded) and SSCI (20%)	Citations per faculty (20%)
Quality of education		Alumni winning Nobel Prizes and Fields Medals (10%)	Employer reputation survey (10%)
Quality of faculty		Staff winning Nobel Prizes and Fields Medals (20%) Highly cited researchers in 21 broad subject categories (20%)	Global survey of academic reputation (40%)
International outlook	International to domestic student/staff ratios & international collaboration (7.5%)		International faculty and student ratios (10%)
Industry income	Knowledge-transfer activities (2.5%)		
Other		Weighted scores[a] of the five indicators to the number of full-time equivalent academic staff (10%)	

[a] The weighted scores of the five indicators (research, citations, quality of education and two of them in terms of quality of faculty) divided by the number of full-time equivalent academic staff.

Table 1b
Main criteria and the associated weights of the indicators employed in major university rankings (research-based rankings).

Criterion	Indicators	
	Leiden[a]	URAP
Scientific productivity		Articles published in 2012–2014[b] and indexed by Web of Science (25%)
Research impact	The number/proportion of publications that, compared with other publications in the same field and in the same year, belong to the top 10% most frequently cited	Total number of citations excluding self-citations received in 2012–2014 for the articles published in and indexed by Web of Science (20%)
Research quality		Total number of articles multiplied by the ratio of university to world average of citations per publication in the corresponding field (20%) Total number of citations multiplied by the ratio of university to world average of citations per publication in the corresponding field (25%)
International collaboration	The number/proportion of publications that have been co-authored with one or more other organizations The number/proportion of publications that have been co-authored by two or more countries The number/proportion of publications that have been co-authored with one or more industrial partners The number and the proportion of a university's publications with a geographical collaboration distance of less than 100 km and more than 5000 km	Total number of publications made in collaboration with foreign universities (10%)

[a] Weights of the indicators are not given since only one indicator is used in a chosen ranking.
[b] Since indicators depend on the year measured, descriptions correspond to the most recent year of the rankings (i.e., 2015 for Leiden and 2015/2016 for URAP).

global survey of academic reputation to measure the faculty quality in which academics are asked about the best institutions in their own fields of expertise. QS World also employs a similar global survey to measure the reputation of the universities who are among the best employers for their graduates.

The criteria of the rankings identified in Table 1a do not apply to the two other ratings as they both are measured only on research related factors. Thus, criteria for these rankings were identified as scientific productivity, research impact, research quality, and international collaboration, and are separately given in Table 1b.

In both rankings, publications and associated citations form the basis of their methodology. While there exist different indicators for measuring research impact and international collaboration in Leiden, the actual rankings are given based only on one of these indicators resulting in 100% weight of the chosen indicator. Leiden, is thus not a comprehensive ranking and can be considered a biased one, as the ranking is based on one measure such as the number or proportion of publications belonging to the top most frequently cited. On the other hand, URAP is comparatively more comprehensive in measuring the overall research productivity as it has components from almost all criteria such as scientific productivity, research impact, research quality, and international collaboration with approximately close weights of each.

3.2. Actual rankings of top universities of the base ranking across other leading indexes

THE World is designated as the base ranking of this study. THE World ranking is one of the highly publicized global rankings employed by all types of stakeholders since 2004 as well as studied in prior literature (Marginson and van der Wende, 2007; Saisana et al., 2011; Buela-Casal et al., 2007; Safon, 2013). The rankings of the top 50 universities of the designated (i.e., base) ranking across other leading indexes are given in Table 2.

The top 50 universities of the most recent THE World ranking (2015/2016) is placed in the first column of Table 2. Associated rankings of these 50 universities in other major university ranking indexes are provided in the following four columns. QS World and THE World rankings show similarities in terms of the top ten such that nine of the top ten according to THE find a place in QS World top ten, only with different orders. When the base ranking is compared with ARWU, two universities outside of the top ten of THE World enter into the rankings in ARWU, where the differences substantially increase with Leiden and URAP with respect to the base ranking. Variability among the indexes increases further as more universities are explored (i.e., top 50 or top 100).

Descriptive statistics of the 100 universities of the base ranking have been given in Table 3. The average number of full-time students is 25,807, of which about 23% are international students, and the ratio of female students is about 51%. On average, there are 15.92 students per academic staff. With regards to the locational distribution, 42 universities are located in Europe where the top three countries are United Kingdom, Germany, and Netherlands. 43 universities are located in North America of which 39 of them are in the U.S. Of the rest of the 100 universities, 9 of them are in Asia and 6 of them are in Australia.

Lastly, a comparison of the counts and ranges of the top 100 universities of the base ranking and the matched ones across other indexes are given in Table 4. Out of the 100 best universities, 75, 68, 60, and 71 also appear in the best 100 places of QS, ARWU, Leiden, and URAP, respectively. Only a few of them have the exact same ranking across the other ranking indexes. There exist a few universities that do not even appear in the entire lists of the other leading indexes with the highest number of 6 in URAP. In addition to the frequencies, the corresponding ranges of the top 100 universities of the base ranking have been determined, and these are 1–216, 1–(301 to 400), 1–544, and 1–247 across the indexes QS, ARWU, Leiden, and URAP, respectively. Differences in the counts and ranges of the top 100 universities highlight the role of indicators and their associated weights on the rankings further establishing the need to analyze the relation among the indexes.

3.3. Correlation of the ranks across leading ranking indexes

To understand the interrelation between universities that are ranked best according to the base index and their actual rankings (when they exist) across other leading indexes, Pearson correlation coefficients have been computed. Although Spearman correlation coefficients are used to depict the monotonic ranking relations, Pearson correlations have been employed here since we are interested in the wide range of ranks of the top universities of the base index across others in addition to the orders of these universities across the others.

Correlation coefficients of the positions of top 10, 50, and 100 universities of the base ranking with their positions across the other leading indexes are given separately in Tables 5a, 5b, and 5c, respectively. The significance of the correlations is provided with the *p*-values that are shown in parentheses in Tables 5a, 5b, and 5c. As more universities are included in the analyses, the significance levels considerably

Table 2
Places of the top 50 universities of the base ranking (THE World) in other leading indices of year 2015.

University	THE World	QS	ARWU	Leiden[a]	URAP
California Institute of Technology	1	5	7	6	56
University of Oxford	2	6	10	17	3
Stanford University	3	3	2	3	8
University of Cambridge	4	3	5	23	5
Massachusetts Institute of Technology	5	1	3	1	7
Harvard University	6	2	1	2	1
Princeton University	7	11	6	5	89
Imperial College London	8	8	23	33	15
ETH Zurich	9	9	20	25	39
University of Chicago	10	10	9	18	21
Johns Hopkins University	11	16	16	36	4
Yale University	12	15	11	13	20
University of California, Berkeley	13	26	4	4	9
University College London	14	7	18	32	6
Columbia University	15	22	8	19	14
University of California, Los Angeles	16	27	12	20	12
University of Pennsylvania	17	18	17	24	13
Cornell University	18	17	13	28	25
University of Toronto	19	34	25	86	2
Duke University	20	29	31	31	24
University of Michigan	21	30	22	48	10
Carnegie Mellon University	22	62	61	65	229
LSE	23	35	101–150	112	NA
University of Edinburgh	24	21	47	63	45
Northwestern University	25	32	27	21	38
National University of Singapore	26	12	101–150	144	32
King's College London	27	19	55	35	54
Karolinska Institute	28	NA	48	114	60
LMU Munich	29	75	52	110	52
New York University	30	53	27	30	64
EPFL	31	14	301–400	15	103
University of Washington	32	65	15	27	11
University of Melbourne	33	42	44	117	30
University of British Columbia	34	50	40	107	22
KU Leuven	35	82	90	71	23
University of Illinois at U-C	36	59	29	67	72
Heidelberg University	37	66	46	146	47
McGill University	38	24	64	149	34
University of California, San Diego	39	44	14	16	17
University of California, Santa Barbara	39	129	38	7	131
Georgia Institute of Technology	41	84	101–150	49	136
Peking University	42	41	101–150	379	44
University of Tokyo	43	39	21	415	18
University of California, Davis	44	85	57	74	41
University of Hong Kong	44	30	151–200	272	149
University of Texas at Austin	46	77	37	47	67
Tsinghua University	47	25	101–150	250	48
Wageningen UR	47	135	101–150	92	167
Humboldt University of Berlin	49	126	NA	184	69
University of Wisconsin-Madison	50	54	24	56	28

[a] Size independent ranking has been given. Selected indicator is scientific impact measured as the proportion of top 10% publications.

improve between the correlation pairs and even more pairs become significantly correlated that were not before with fewer universities.

Actual positions of the top 100 universities of the base ranking THE World are strongly and significantly correlated with the positions of the same universities in ARWU and QS World. Strong correlations pinpoint the closeness of indicators of the underlying indexes, although different measures are used, that result in approximately similar rankings of the top universities of THE World and the corresponding positions across ARWU and QS World. Moderate correlations exist among the index pairs ARWU and URAP, ARWU and QS, and THE World and URAP. On the other hand, weak correlations also occur between the pairs THE World and Leiden, ARWU and Leiden, and QS and URAP. Lastly, two of the pairs have very weak correlations or insignificant ones. The correlation between Leiden and URAP is not significant, and the correlation between QS and Leiden is very weak at a very low significance level.

Table 3
Descriptive statistics of the top 100 universities of the base ranking (i.e., THE World).

Statistics	Number of students[a]	Student to academic staff ratio	International students (%)	Female students (%)
Avg.	25,807	15.92	22.57	50.60
Std. Dev.	12,443	10.50	11.85	8.17
Min	2243	3.60	6.00	26.00
Max	66,198	70.40	70.00	70.00
Median	24,680	13.95	20.00	52.00

Country of the institution	Continent	Count (country)	Count (continent)
Australia	Australia	6	6
Belgium	Europe	1	42
Denmark		1	
Finland		1	
France		1	
Germany		9	
Netherlands		8	
Sweden		3	
Switzerland		2	
United Kingdom		16	
Hong Kong	Asia	2	9
Japan		2	
Singapore		2	
South Korea		1	
China		2	
Canada	America	4	43
United States of America		39	

[a] Full-time students.

3.4. Analysis of the rankings

This study analyzed the ranking indexes from two perspectives. First, the main indicators of the leading global ranking indexes were categorized under more general criteria. Secondly, this study explored the relationship between the places of the top 100 universities of the base ranking and matched places of the same universities across other leading indexes.

3.4.1. The main criteria of the leading indexes

Based on the indicators used in determining rankings of the five leading indexes, we have come up with seven criteria for assessing the university qualities. Indicators of the indexes are collected in the proper criteria for each index. Analysis of Tables 1a and 1b, where the criteria and indicators of the underlying indexes were shown, reveals that the existing indexes can be misleading at various points.

Most of the leading indexes, except for ARWU and Leiden, do not take into account the size of the institution. In ARWU, weighted scores of the used indicators to the number of full-time equivalent academic staff have been used as a proxy for size. Whereas in Leiden, size independent ranking are calculated using the proportion of the underlying metric with one university property, so both smaller and larger universities perform well in such rankings. However, the other rankings do not consider the size where many indicators' measurements would naturally increase with the size of the institution. Besides, the field of the subject may be limited to only a few for relatively smaller universities resulting in fewer outcomes measured through the used indicators of the indexes. In general, smaller universities in terms of institution size are usually young universities that are recently established, and they naturally lag in terms of the metrics measured by many indicators such as top cited publications, international collaboration, and reputation surveys.

Comparison of the criteria across different indexes also reveals the variability in measuring the performance of the two basic functions of the universities, i.e., teaching and research. While teaching has been measured using some input measures such as ratios of student to academic staff or doctorates awarded to academic staff, and the outcomes of reputation surveys in only two of the indexes, teaching has not

Table 4
Comparison of the counts and ranges of the universities across the base (THE World) and the other leading indexes.

Out of 100 best (acc. to THE) that				
	Number of universities			
	QS	ARWU	Leiden	URAP
Also appear in:	75	68	60	72
Have the exact same rankings in:	4	2	2	0
Are not available in:	1	2	1	6

Corresponding range of the best 100 (acc. to THE) in the other leading index			
QS	ARWU	Leiden	URAP
1–216	1–(301 to 400)	1–544	1–247

been assessed in ARWU, URAP, and Leiden. From the viewpoint of research, while all indexes measure research productivity, the indicators differ among them even with different weights for the same indicators. Furthermore, two of the indexes focus only on research outcomes completely ignoring the other aspects of the quality of a higher education, as shown separately in Table 1b.

The drawback of the variability in measuring teaching and research related criteria also shows itself in other indicators and their associated weights. It appears that the industry income has almost never been measured in the existing major ranking indexes. Only in one index, i.e., THE World ranking, has knowledge-transfer activities of the institution been measured with a small percentage of 2.5% in the overall score.

Another emerging measure is using the count of Nobel Prizes or Fields Medals (for alumni or academic staff) for measuring the quality of education or quality of faculty in ARWU. Both measures have pretty good weights in the overall methodology and do not appear in any other indexes further questioning the use of such measures. Lastly, two of the rankings use reputation surveys for measuring teaching, research, quality of education and quality of faculty with at least a weight of 10% each. Since reputation surveys are based on opinions of the respondents, although respondents are well-known scholars or employers in their respective fields, employing these measures will result in biases towards the well-known institutions in developed countries.

3.4.2. Pairwise interrelations of the leading indexes

We explore the rankings of the top 100 universities of the base index and the matched positions of these universities across other leading indexes, as shown in Table 2. The variability among the base ranks and the matched positions, as highlighted in Table 4, further require pairwise correlations to see whether the variability can be explained through the differences in selected indicators of such indexes.

The strength as well as the significance of the pairwise correlations improved as more universities are included in the analysis. Only two pairs of indexes, which are between THE World and ARWU and THE World and QS indicate a strong and significant correlation (with values above 0.70). The base ranking appears in both pairs, and all three indexes have indicators of measuring various aspects of higher-education, since Leiden and URAP measure only research related criteria. Despite the variability in indicators used across these three indexes for measuring the same criteria (such as the reputation surveys of THE World and QS and Nobel Prizes and Field Medals of ARWU), the strong correlations indicate that different indicators may indeed measure the similar criteria to the same level of effectiveness.

It is also worth noting that there is no correlation between the indexes Leiden and URAP and a very weak correlation exists between QS and Leiden. Although both Leiden and URAP are research-based rankings, their indicators do not correlate at all. Moreover, Leiden also does not correlate with QS, where research is evaluated based only on citations. However, we found that the other sole research based index (i.e., URAP) correlates significantly with other indexes at moderate or low levels, which can be explained through the diversification of its indicators relative to the Leiden, even though all are purely research related.

4. Conclusions and recommendations for decision makers

This study has implications for the decision makers in higher education across the world. The synopsis of the major university ranking indexes and how these indexes can be used to derive the most benefit for the university are the two main implications.

The synopsis of the leading university rankings reveals the variability in the actual places of the best universities across different indexes. Some top universities in one leading index do not even take a position in the list of another leading index. The variability in the actual lists partially lead to the variety and weighting of the indicators used. While some rankings use only hard data, some others rely on both hard and

Table 5a
Pearson correlation coefficients of the top ranked universities of the base ranking (THE World) and other leading ranking indices.

Rankings of the top 10 universities in THE World and corresponding rankings in other leading indexes					
	THE	QS	ARWU	Leiden	URAP
THE	1				
QS	0.5855 (0.0753)	1			
ARWU	0.4803 (0.1600)	0.5844 (0.0760)	1		
Leiden	0.4215 (0.2251)	0.4422 (0.2008)	0.8512 (0.0018)	1	
URAP	0.1723 (0.6341)	0.6458 (0.0437)	0.1680 (0.6427)	−0.1107 (0.7607)	1

Table 5b
Pearson correlation coefficients of the top ranked universities of the base ranking (THE World) and other leading ranking indices.

Rankings of the top 50 universities in THE World and corresponding rankings in other leading indexes					
	THE	QS	ARWU	Leiden	URAP
THE	1				
QS	0.7324 (<0.0001)	1			
ARWU	0.6303 (<0.0001)	0.6193 (<0.0001)	1		
Leiden	0.5783 (<0.0001)	0.1984 (0.1717)	0.3734 (0.0162)	1	
URAP	0.4104 (0.0034)	0.5269 (0.0001)	0.4324 (0.0048)	0.1428 (0.3276)	1

Table 5c
Pearson correlation coefficients of the top ranked universities of the base ranking (THE World) and other leading ranking indices.

Rankings of the top 100 universities in THE World and corresponding rankings in other leading indexes

	THE	QS	ARWU	Leiden	URAP
THE	1				
QS	0.7364 (<.0001)	1			
ARWU	0.7582 (<.0001)	0.6239 (<0.0001)	1		
Leiden	0.4747 (<.0001)	0.1718 (0.0907)	0.3734 (0.0019)	1	
URAP	0.5115 (<.0001)	0.4323 (<.0001)	0.6393 (<0.0001)	0.1215 (0.2434)	1

soft data. Soft data are highly qualitative in nature and depend on ideas, knowledge, experience and opinions. Using hard data is good in order to obtain more objective results, however, hard data are limited to certain criteria, universities, and regions. While many of the criteria of university excellence can be explained through soft data, both the high cost of obtaining such data and the reliability of measures used make the ranking systems to turn towards hard data. In terms of the use of hard and soft data, both can be used in forming the overall scores, however, the transparency and accuracy of the soft data need to be particularly assured.

Another significant issue with respect to the variety of indicators, is the measurement of the two most important functions of a university, research and teaching, together in assessing quality. Teaching and research are historically perceived as two functions that go hand in hand and support each other, yet, tensions occur in excelling at both functions as rewarding research may take time from teaching or vice versa from the standpoint of academicians (Serow, 2000). Thus, accumulating the performance of such criteria in one basket as a single unit of performance also raises the question of how healthy it is to analyze universities of different areas of emphasis in one list of rankings. This also gives rise to indexes that focus only on research such as Leiden and URAP that are also explored in this study.

While the variability in the actual lists can be explained partly by the variety of indicators used in measuring the fundamental criteria of higher education quality such as teaching, research, international outcome, etc., some research-based indexes even do not correlate with each other. All these point to the need for more diversified rankings with respect to the size of the institutions, weighting of the used indicators, and disciplinary differences, as Shin and Toutkoushian (2011) also elaborate on the future of university rankings. Thus, in terms of viewing and reading the indexes correctly, only universities similar in size, age, or specialized field of subject could be compared with each other. Comparing the major fields of study or research in place of the colleges or faculties would make sense in most cases as universities also differ in terms of their impact or excellence in certain fields of study.

A major weakness in almost all indexes explored, is the lack of instruments measuring the industry income such as the knowledge-transfer activities. To fill in this gap, university-ranking systems could create separate indexes specific to certain subject fields such as entrepreneurship and innovation.

The second main implication of this study is related to obtaining the most benefit out of reading and viewing the university rankings. The institution needs to clearly question its role of appearing and climbing in the leading international university rankings. Monitoring a university's trend in the rankings over time can be beneficial in developing strategies to increase the recognition of the university. This strategy is closely related to the university's mission of attracting international students and faculty. Thus, the institution is suggested to work towards raising its position in such rankings as it would be a critical element of its marketing efforts. Incorporating a faculty with a Nobel Prize or providing the correct data at the correct time to such institutions would help to make the rankings climb significantly.

Future studies could consider intercorrelations among ranking indexes in a specialized field of subjects as universities' strength in terms of teaching and research might differ significantly across different fields of study. Another possible area of future research might be using factor analysis or other advanced statistical techniques to further determine the common factors that position the same universities across different ranking indexes.

References

Benneworth, P., Coenen, L., Moodysson, J., Asheim, B., 2009. Exploring the multiple roles of Lund University in strengthening Scania's regional innovation system: towards institutional learning? Eur. Learn. Stud. 17 (11), 1645–1664.

Bowman, N.A., Bastedo, M.N., 2011. Anchoring effects in world university rankings: exploring biases in reputation scores. High. Educ. 61 (4), 431–444.

Buela-Casal, G., Gutierrez-Martinez, O., Bermudez-Sanchez, M.P., Vadillo-Munoz, O., 2007. Comparative study of international academic rankings of universities. Scientometrics 71 (3), 349–365.

Centre for Science and Technology Studies, Leiden University, 2015y. CWTS Leiden ranking 2015. Retrieved Jan 19, 2016, from http://www.leidenranking.com.

Chen, K.-h., Liao, P.-y., 2012. A comparative study on world university rankings: a bibliometric survey. Scientometrics 92 (1), 89–103.

Dehon, C., McCathie, A., Verardi, V., 2009. Uncovering excellence in academic rankings: a closer look at the Shanghai ranking. Scientometrics 83 (2), 515–524.

Dill, D.D., 2000. Capacity building as an instrument of institutional reform: improving the quality of higher education through academic audits in the UK, New Zealand, Sweden, and Hong Kong. J. Comp. Policy Anal. Res. Pract. 2 (2), 211–234.

Docampo, D., 2011. On using the Shanghai ranking to assess the research performance of university systems. Scientometrics 86 (1), 77–92.

Eccles, C., 2002. The use of university rankings in the United Kingdom. High. Educ. Eur. 27 (4), 423–432.

Etzkowitz, H., Leydesdorff, L., 1999. The future location of research and technology transfer. J. Technol. Transfer. 24 (2/3), 111–123.

Harvey, L., 2008. Rankings of higher education institutions: a critical review. Qual. High. Educ. 14 (3), 187–208.

Hazelkorn, E., 2008. Learning to Live with League Tables and Ranking: The Experience of Institutional Leaders. High. Educ. Policy 21 (2), 193–215.

Holmes, R., 2012. Eight years of ranking: what have we learned? University World News, 207, 5 February. [Online] Available: http://www.universityworldnews.com/article.php?story=20120131120436 74&query=Holmes.

Hurtado, M.E., 2012. Latin Americans challenge international rankings, University World News, 224, 30 May. [Online] Available: http://www.universityworldnews.com/article.php?story=20120530142927343.

Kivinen, O., Hedman, J., Kaipainen, P., 2013. Productivity analysis of research in natural sciences, technology and clinical medicine: an input–output model applied in comparison of Top 300 ranked universities of 4 North European and 4 East Asian countries. Scientometrics 94 (2), 683–699.

Li, M., Shankar, S., Tang, K.K., 2011. Why does the USA dominate university league tables? Stud. High. Educ. 36 (8), 923–937.

Lincoln, D., 2012. Rankings: an idea whose time has come, and gone. Inside-Higher Ed, 28 February. [Online] Available: http://www.insidehighered.com/blogs/world-view/rankings-idea-whose-time-has-come-and-gone.

Marginson, S., van der Wende, M., 2007. To rank or to be ranked: the impact of global rankings in higher education. J. Stud. Int. Educ. 11 (3–4), 306–329.

Mergen, E., Grant, D., Widrick, S.M., 2000. Quality management applied to higher education. Total Qual. Manag. 11 (3), 345–352.

Mintzberg, H., 1979. The Structuring of Organizations. Prentice Hall, Englewood Cliffs, NJ.

Östling, M., 1997. Self-evaluation for audit — a tool for institutional improvement? INQAAHE Conference. Berg-en-Dl, South Africa, May

Safon, V., 2013. What do global university rankings really measure? The search for the X factor and the X entity. Scientometrics 97 (2), 223–244.

Saisana, M., d'Hombres, B., Saltelli, A., 2011. Rickety numbers: volatility of university rankings and policy implications. Res. Policy 40 (1), 165–177.

Serow, R.C., 2000. Research and teaching at a research university. High. Educ. 40, 449–463.

Shanghai Ranking Consultancy & Center for World-Class Universities of Shanghai Jiao Tong University, 2015e. The academic ranking of world universities: 2015. Retrieved Jan 19, 2016, from http://www.shanghairanking.com.

Shin, J.C., 2011. Organizational effectiveness and university rankings. In: Shin, J.C., Toutkoushian, R.K., Teichler, U. (Eds.), University Rankings, The Changing Academy: The Changing Academic Profession in International Comparative Perspective (Vol. 3). Springer Science, Dordrecht.

Shin, J.C., Toutkoushian, R.K., 2011. The past, present, and future of University Rankings. In: Shin, J.C., Toutkoushian, R.K., Teichler, U. (Eds.), University Rankings, The Changing Academy: The Changing Academic Profession in International Comparative Perspective (Vol. 3). Springer Science, Dordrecht.

Soh, K., 2013. Times Higher Education 100 under 50 ranking: old wine in a new bottle? Qual. High. Educ. 19 (1), 111–121.

Teichler, U., 2011. The future of university rankings. In: Shin, J.C., Toutkoushian, R.K., Teichler, U. (Eds.), University Rankings, The Changing Academy: The Changing Academic Profession in International Comparative Perspective (Vol. 3). Springer Science, Dordrecht.

The Times Higher Education Supplement, 2015. The Times Higher Education World university rankings 2015–2016. Retrieved Jan 19, 2016, from https://www.timeshighereducation.com/world-university-rankings/.

Usher, A., Savino, M., 2006. A World of Difference: A Global Survey of University League Tables. Educational Policy Institute, Toronto, ON.

van Raan, A.F.J., 2005. Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. Scientometrics 62 (1), 133–143.

van der Wende, M.C., Westerheijden, D.F., 2009. Rankings and Classifications: the Need for a Multidimensional Approach. In: van Vught, F.A. (Ed.), Mapping the Higher Education Landscape. Towards a European Classification of Higher Education. Springer, pp. 71–87.

Gökçen Arkalı Olcay is currently an Assistant Professor at the School of Management at İstanbul Sehir University. She received her BS in Electrical Engineering from Bilkent University, and her MS and PhD degrees in Operations Management from the University of Texas at Dallas. Her research interests include empirical research in operations management, innovation management, and R&D management. She teaches courses on Operations Management, Statistics, Statistical Analysis, and Quantitative Management.

Melih Bulu worked at various sections of the private sector both as a professional and as an entrepreneur. Since 2004, he has been the General Secretary of the International Competitiveness Research Institute (URAK), an NGO working on economic competitiveness of cities and countries. He taught strategy-related courses at Istanbul Sehir University. He is currently the Founding Rector at Istinye University. His main interest areas are city competitiveness, sectoral competitiveness, entrepreneurship and innovation. He has various publications in academic and popular media.