From rankings to ratings: An input-output analysis of productivity of research and teaching by disciplines

Osmo KIVINEN
Research Unit for the Sociology of Education (RUSE), University of Turku
FIN-20014 Turku, Finland

Juha HEDMAN
Research Unit for the Sociology of Education (RUSE), University of Turku
FIN-20014 Turku, Finland

and

Päivi KAIPAINEN
Research Unit for the Sociology of Education (RUSE), University of Turku
FIN-20014 Turku, Finland

ABSTRACT

University rankings such as QS World University Rankings, HEEACT- performance ranking and Academic Ranking of World Universities (ARWU) have reached world-wide reputation among the wide audience. However, there is still much to do if the ranking service providers aim gain any ground on the policy-making arenas, where large amounts of tax-payers money are at stake. The current rankings may serve the global status markets but not so well governments who are in need of evidence for making better strategic decisions as concerns higher education and science policy. From our point of view, successful evidence-based policy requires apt tools for evaluation of productivity of research and teaching; in this respect recent rankings data is one-sided and weak, also the used methods need developing. The paper leans on a digitalized input-output analysis-platform for productivity (and impact) of research and teaching. The paper shows how both national and international data sources (Finnish KOTA-database and International QS-, HEEACT- and ARWU-databases) can serve productivity analysis in higher education, demonstrating two different cases of evidence-based HE-policy. The first case, applying the national Finnish data, presents outlines of a rating-system, which serves as a means to sieve out most productive (A++) university units in research and teaching in nine fields (natural sciences, health sciences, medicine, engineering, social sciences, humanities, business and management, education and psychology) from a pool of 66 Finnish university units. In the second case, the rating-system is applied in a European multinational context in order to sieve out most-productive (A++) university units in research in six fields (natural resources and environment, life sciences, clinical medicine, natural sciences, social sciences and technology) from a pool of 655 European university units.

Keywords: Evidence-based policy, input-output analyses, productivity of research and teaching, productivity ratings, universities

1. INTRODUCTION

This paper asks, how to proceed further from performance indicators to analyzing productivity of scientific action and towards a rating system [cf. 4, 6]. We are interested in making rankings to serve strategic planning and evidence-based policy making. We combine the rankings data with our input-output analyses that we have developed for a nation state framework [5, 7].

European aspirations to create a common European Higher Education and Research Area (EHERA) accompanied for instance with Bologna process call for appropriate tools for evaluation of productivity of scientific action. In this paper we first demonstrate our input-output platform in the national framework utilizing the Finnish data and then show some preliminary outlines applying it to European data extracted from world-wide university ranking on-line-services; ARWU, QS and HEEACT.

2. MEASURES FOR PRODUCTIVITY RATING SYSTEM

In Finland, the Ministry of education maintains a specific HE-data-base (KOTA) providing an access to input and output data of Finnish universities also by disciplines. In the following, we briefly demonstrate our input-output model of productivity analyses of research and teaching in 16 Finnish universities by 9 disciplines utilizing KOTA-data bank. The indicators for output of research are i) refereed articles published in international scientific journals, ii) doctoral degrees and iii) merit-based funding. The indicators for input of research are i) professorial man-years and ii) external funding of research. The indicators for output of teaching, in turn, are i) master's degrees and ii) doctoral degrees. The indicators for input of teaching are i) professorial man-years and ii) other teaching man-years.

The key features of our productivity analysis platform applied in the Finnish case are as follows. Each of the nine disciplines (Medicine, Natural Sciences, Education, Business and Management, Nursing and Health Sciences, Engineering, Humanities, Psychology and Social Sciences) is analyzed. In phase one 66 university units by discipline are assigned with six productivity coefficients for research (a^R, b^R, c^R, d^R, e^R, f^R) per year (5 altogether) by relating unit's share of discipline's total output to unit's share of discipline's total input. In phase two units are ranked 30 times (six coefficients × five years) based on the values of the productivity coefficients. The maximum amount of ranking points are n×30, where n is the number of units ranked. The maximum indicates that a unit receiving n×30 ranking points is the most productive in its discipline on all coefficients in all years. In phase three ranking points are standardized on a scale from 1 to 100.

In a similar vein, each unit gets a score also on productivity of teaching. Units are assigned with two productivity coefficients (a^T, b^T) per year and then ranked 10 times. The maximum amount of ranking points is now n×10. Let us emphasize that the standardization procedure enables a simultaneous study of productivity of research and teaching, even though both have their own scales; productivity of research is measured from 1 to n×30, while productivity of teaching is measured from 1 to n×10. In both cases, however, the middle-point of standardizing is the divide between productive (+) units and other units (-). The standardization allows us to use productivity coefficients (see b^R and a^T) in both orderings identically. The input-output relationship between doctoral degrees and professorial man-years bears significance in both analyses, because the nexus of research, teaching and researcher training is in the core of university action.

In searching for objective indicator basis for an evidence-based policy, the sum-and-weight approach, preferred in global university rankings, is problematic because it incorporates hidden preconceptions in the calculations. Sum-and-weight procedure is also inevitably irreversible, because the published set of scores represents a specific case only and other score sets are untraceable *post facto* [c.f. 3]. Consequently, the scores and weights tell (at least) two contradictory stories about how to make sense of various performance indicators [see 4]. It is evident that evidence-based policy calls for methodological refinements of rankings [see also 1, 2, 8 and 9].

Next we present Finnish results of productivity in research and teaching in nine disciplines. The rating system based on the results of productivity analyses sieves out the most productive units in their own disciplines out of the pool of 66 units studied. Figure 1 illustrates, an example of the Finnish university units in the field of Technology: that the units scoring 50 or more (1 representing lowest score and 100 highest score) in productivity of both research and teaching (Research + and Teaching +) position themselves in quadrant I and hence receive rating of A^{++}. In our example Helsinki University of Technology is the only Finnish technology unit to reach this quadrant getting rating A^{++} with scores 66 for research and 68 for teaching.

Units scoring 50 or more on productivity of research (Research +), but scoring less than 50 for teaching (Teaching -) position themselves in quadrant II with a rating of A+. Only the technology unit of the Swedish language university Åbo Akademi is positioned in this quadrant with scores 70 on research and 41 on teaching.

Units which score less than 50 in productivity of research (Research -), but score 50 or more for teaching (Teaching +) place themselves in quadrant III with a rating of A. In the example Lappeenranta University of Technology (score 32 on research/score 62 on teaching) and Tampere University of Technology (39/52) fall into this A-rate quadrant. These units meet the 'standards' of the discipline in teaching, but have all the reason to strengthen their research basis.

Those units, scoring less than 50 in productivity of both research and teaching (Research -, Teaching -) are placed in quadrant IV with a rating of A_0. In the example such units are found in Oulu (38/39), Turku (46/14) and Vaasa (26/4). These units show lesser evidence of productive research and teaching than other units. It is important to note, however, that even those units which have received an A_0 rating have still managed to gain results noted by the science community, albeit on a smaller scale than the units within the other three quadrants. Note, that the size of the marker (circle) is in relation to the size of inputs. The tiny circles indicate universities with very small inputs in the field in question and hence those universities naturally have very little importance for the education policy. While the universities with big inputs, indicated by big circle markers, like Helsinki University of Technology, in our example, have a lot of importance when government decisions are made.

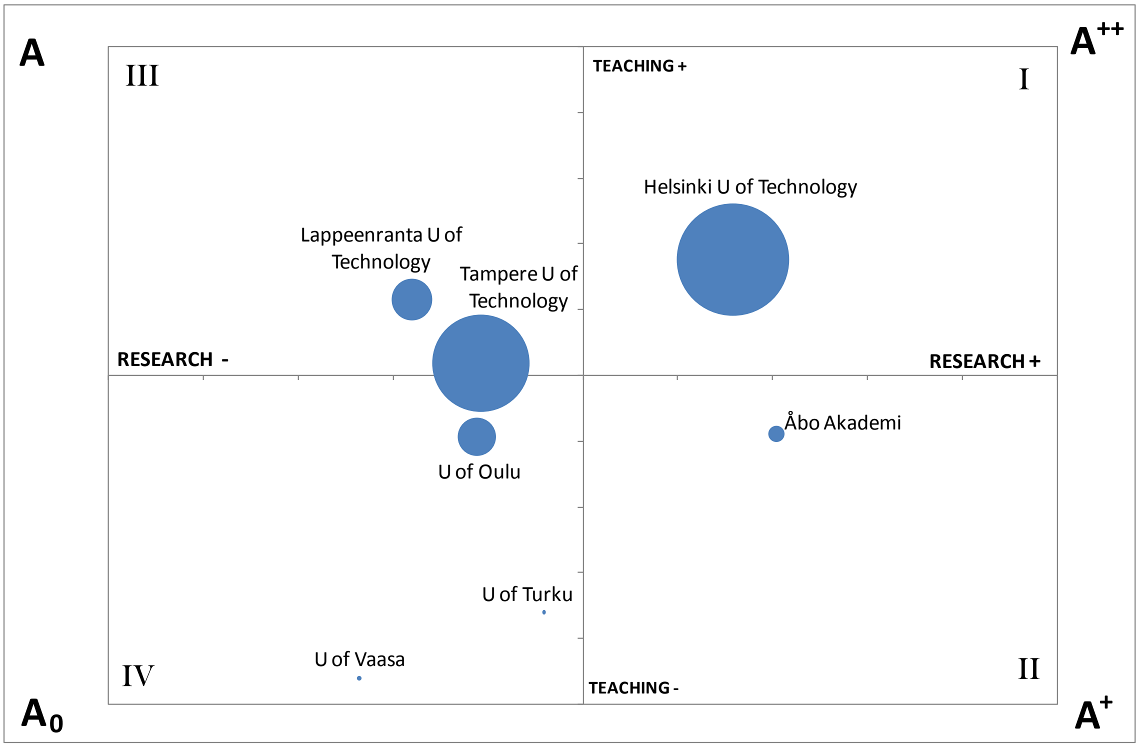

Figure 1. A rating system for productivity of research and teaching: example of the university units in technology in Finland.

3. AN APPLICATION ON EUROPEAN DATA

Next we present some tentative results of our input-output analysis in a multinational context. The analysis is based on the QS-, HEEACT- and ARWU-data by disciplines on European universities in 2009. In this multinational context, we have to confine ourselves with productivity of research only since data on teaching is not available. However, we take into account also impact of research, which enriches the analysis. Hence, in the next rating we have had to replaced teaching with research impact.

The input-data does not play any role in university rankings. If the rankings should serve the demands of competition-relevant and productivity-oriented policy mere outputs without corresponding inputs are not enough. In the following we apply our input-output analysis in productivity (and impact) analysis of research in European universities, of 45 EHERA countries.

ARWU on-line service 2009 is methodologically based on measuring academic performance consisting of quality of education, quality of faculty, research output and per capita performance. The academic performance of 3,000 universities is operationalized by a score composed from six indicators and their corresponding weights. QS-THES 2009-edition (http://www.topuniversities.com) on-line service is methodologically based on measuring relative strength of leading universities consisting of academic peer review, employer review, faculty-student ratio, citations per faculty, international faculty and international students. The relative strength of 2,000 universities is operationalized by a score composed from six indicators and their corresponding weights HEEACT 2009-edition (http://ranking.heeact.edu.tw) is methodologically based on measuring research performance consisting of research productivity, research impact and research excellence, research output and per capita performance. The research performance of 3,500 universities is operationalized by a score composed from eight indicators and their corresponding weights.

A closer inspection of the performance indicators of these three different ranking data reveals that many indicators are almost totally substitutable with each other, while other indicators in their elusiveness are too specific (e.g. Nobel Prizes and Fields Medals) to yield appropriate base for large-scale productivity analysis. By constructing six fields (Natural Resources and Environment, Clinical

Medicine, Technology, Life Sciences, Natural Sciences and Social Sciences), however, we have the criterion needed to unequivocally identify our complete set of 655 European ranked units from a total of 184 European universities in 15 European countries, amounting to the human resources total of 390 000 in terms of faculty. Moreover, in our data on ranked units the human resources is distributed so that 22 % of is in the service of Clinical Medicine, 19 % in the service of Natural Sciences, 17 % Life Sciences, 15 % Technology, 14 % Natural Resources and Environment and 12 % Social Sciences.

Correspondingly in our data 32 % of the articles are published in the field of Clinical Medicine, 24 % in the field of Life Sciences, 22 % in Natural Sciences, 12 % in Technology, 6 % in Natural Resources and Environment and 3 % in Social Sciences. As concerns articles in high-impact journals (see http://ranking.heeact.edu.tw/en-us/2010/Page/Indicators) 31 % of the articles are published in the field of Clinical Medicine, 22 % in the field of Natural Sciences, 21 % in the field of Life Sciences, 12 % in the field of Technology, 11 % in the field of Natural Resources and Environment and 2 % in the field of Social Sciences.

Those units that make it to the international rankings (655 altogether) are assigned with a productivity coefficient, which relates, by fields, unit's share of outputs to share of inputs. Impact coefficient, in turn, relates, by fields, unit's share of scientific impacts to share of inputs. Units are then ranked twice within their fields (NR-ENV, MED, LIFE, SCI, TEC and SOC) based on the values of the productivity coefficients and impact coefficients respectively. The maximum amount of ranking points vary by countries (UK, GER, ITA, ESP, FRA, NL, SWI, IRL, BEL, AUT, SWE, DEN, NOR, FIN, RUS) depending on how many ranked units in each field a country has. Hence, the standardized scores on productivity and impact are determined by relative placements of country X's ranked units in field-specific rankings. The maximum (100) number of standardized points would indicate that country's ranked units make the top placements in all field rankings that they are involved in. The minimum (1) would indicate that country's ranked units make the bottom placements in all field rankings that they are involved in. As a result, each country scores on productivity of research and its scientific impact on a scale from 1 to 100. These two scores are then combined into one as indicated by the following formula,

$$SCORE = \frac{100\sqrt{P^2 + I^2}}{\sqrt{(\max P)^2 + (\max I)^2}}.$$

For example out of the 45 countries belonging to the European Higher Education Area (EHERA) as many as 30 countries do not have any ranking notified units in any of the six fields. Hence, based on ranking data we are able to evaluate the top 15 European countries in six fields leaning on obtained scores on productivity and scientific impact of research.

As table 1 shows, when all six fields are considered the top seven countries scoring 50 or more are NL (64), SWE (60), UK (58), BEL (55), DEN (53), FRA (52) and SWI (51). It is important to note however that the differences in input, or in 'critical mass', even among these seven leading countries are significant. For example UK's input is more than 7 times of that of Denmark's or Belgium's. Also France's and Netherland's input is more than twice the size of Denmark's or Belgium's. Sweden's and Switzerland's input is a bit larger than that of Denmark's or Belgium's.

Next when we look at table 1 by disciplines, the countries scoring 50 or more in productivity and scientific impact of research in the field of natural resources & environment are Denmark (91), Austria (75), Belgium (71), Finland (67), the Netherlands (59), Sweden (55), Italy (51) and the United Kingdom (50). In the field of clinical medicine the top countries are the Netherlands (89), Denmark (62), United Kingdom (61), Sweden (58), Switzerland (51) and Finland (53). In the field of technology the top countries are Norway (76), Belgium (72), United Kingdom (70), Switzerland (68), Sweden (67), France (63) and the Netherlands (53). In the field of life sciences the top countries are Finland (79), Austria (75), Netherlands (57), Sweden (53) and Italy (52). In the field of natural sciences the top countries are France (68), Switzerland (64), United Kingdom (62), Sweden (61), Belgium (56), the Netherlands (54), Denmark (53) and Spain (50). In the field of social sciences the top countries are Netherlands (71), Sweden (63), United Kingdom (61) and Austria (52).

Table 1 Top 15 European countries in terms of productivity of research and its impact in six fields: Clinical Medicine (MED), Technology (TEC), Natural Sciences (SCI), Social Sciences (SOC), Life Sciences (LFS) and Natural Resources & Environment (NR-ENV), (scores 1–100).

	ALL 6 FIELDS		MED		TEC		SCI		SOC		LFS		NR-ENV	
1	**NL**	**64**	NL	89	NOR	76	FRA	68	NL	71	FIN	79	DEN	91
2	**SWE**	**60**	DEN	62	BEL	72	SWI	64	SWE	63	AUT	75	AUT	75
3	**UK**	**58**	UK	61	UK	70	UK	62	UK	61	NL	57	BEL	71
4	**BEL**	**55**	SWE	58	SWI	68	SWE	61	AUT	52	SWE	53	FIN	67
5	**DEN**	**53**	SWI	54	SWE	67	BEL	56	BEL	48	ITA	52	NL	59
6	**FRA**	**52**	FIN	53	FRA	63	NL	54	SWI	46	BEL	49	SWE	55
7	**SWI**	**51**	NOR	48	NL	53	DEN	53	NOR	44	UK	48	ITA	51
8	**FIN**	**48**	GER	47	ITA	46	ESP	50	DEN	38	FRA	47	UK	50
9	**NOR**	**46**	BEL	44	DEN	45	NOR	44	ESP	33	GER	43	SWI	44
10	**ITA**	**42**	ESP	41	FIN	42	FIN	41	IRL	27	ESP	43	FRA	43
11	**ESP**	**40**	ITA	40	GER	34	GER	39	GER	23	SWI	41	NOR	41
12	**GER**	**39**	FRA	36	IRL	34	ITA	37	FIN	20	NOR	40	ESP	41
13	**AUT**	**36**	AUT	9	ESP	34	AUT	37	ITA	11	DEN	37	GER	35
14	**IRL**	**23**	IRL	5	AUT	32	RUS	27			IRL	26	IRL	29
15	**RUS**	**17**			RUS	7	IRL	17						

4. CONCLUSIONS

In this paper we have demonstrated our input-output analyses of research and teaching on Finnish data. Preliminarily we also have tested our model in multinational context by analyzing productivity of research and its impact in European universities leaning on international ranking-data. Since current rankings based on indicators can only give some indications on universities' position in the international status markets we have suggested that in order to serve evidence-based policy they could be replaced by productivity analyses and ratings system. There is clearly a need of appropriate tools for evaluation of productivity of research and teaching in various research-policy arenas around the globe for being able to make evidence-based science policy. Finally, let us point out, that the currently weak data base can not be allowed to be any real obstacle for future evidence-based policy on the international scale.

5. REFERENCES

[1] J.C. Billaut, D. Bouyssou & P. Vincke, "Should You Believe in the Shanghai Ranking?" **Scientometrics:** Vol. 84, No. 1, 2010, pp. 237–263.

[2] C. Dehon, A. McCathie & V. Verardi, "Uncovering Excellence in Academic Rankings: A Closer Look at the Shanghai Ranking." **Scientometrics**: Vol. 83, No. 2, 2010, pp. 515–524.

[3] R.V. Florian, "Irreproducibility of the Results of the Shanghai Academic Ranking of World Universities." **Scientometrics:** Vol. 72, No. 1, 2007, p. 25–32.

[4] O. Kivinen & J. Hedman, "World-wide University Rankings – A Scandinavian Approach." **Scientometrics***:* Vol. 74, No. 3, 2008, pp. 391–408.

[5] O. Kivinen & J. Hedman, "From university rankings towards productivity analysis of scientific action and evidence-based policy." **QS Journal of International Education** 2011. [accepted].

[6] O. Kivinen, J. Hedman & P. Kaipainen, **Academic Performance and Scientific Productivity: A Methodological Assessment of the Academic Ranking of World Universities.** Refereed roundtable paper presented at the 102nd Annual Meeting of the American Sociological Association, August 10–14, 2007 New York.

[7] O. Kivinen, J. Hedman & K. Peltoniemi, **Productivity of research by disciplines in Finland in 2002-2006.** Research Unit for the Sociology of Education (RUSE), Research report 73. Turku: University of Turku.

[8] A. Kroth, & H.D. Daniel, "International University Rankings – A Critical Review of the Methodology." **Zeitschrift fur Erziehungswissenschaft:** Vol. 11, No. 4, 2008, pp. 542–558.

[9] A.F.J. Van Raan, "Fatal Attraction: Conceptual and Methodological Problems in the Ranking of Universities by Bibliometric Methods." **Scientometrics**: Vol. 62, No. 1, 2005, pp. 133–143.

How to cite this paper:

Kivinen, O.; Hedman, J. & Kaipainen, P. 2011. **From rankings to ratings: An input-output analysis of productivity of research and teaching by disciplines.** In: N. Callaos, H-W Chu, W. Lesso, M.J. Savoie, F. Welsh & C.D. Zinn (Eds.) The 2nd International Multi-Conference on Complexity, Informatics and Cybernetics, Proceedings Volume II, March 27th-30th, 2011, Florida: International Institute of Informatics and Systemics, 268–273.