What Does it Mean that UCSB is Ranked Number 7 in the World in "Research Impact?"

Ted Bergstrom

March 12, 2012

The Centre for Science and Technology Studies (CWTS) of Leiden University has published the *Leiden Ranking for 2011/2012* that "measures the scientific performance of 500 major universities worldwide." They rank universities by several distinct criteria, based on publications by authors affiliated with the university, the journals in which they publish, and citations to works written in these journals. By the indicator that they call "the most important impact indicator of the Leiden Ranking", the University of California Santa Barbara is ranked as number 7 in the world. It is outranked only by MIT, Princeton, Harvard, Rice, Stanford, and Cal Tech. It outranks all of the other universities in the UC system, including Berkeley and UCLA.

Can this be right? Has UCSB research really surged past that of Irvine, San Diego and Davis, not to mention UCLA, Berkeley, Yale, Chicago, and Michigan? If so, when did it happen and how did they do it? Even the most ebullient among us must be a trifle skeptical.

My sources tell me that the Leiden group has a good reputation in the scientometrics community. The Leiden group's discussion paper [1] and online methodological discussion `http://www.leidenranking.com/methodology.aspx` indicate that they were well aware of the pitfalls in ascribing authorial credit and university affiliation of authors and it appears that they went about this in a competent way. But even if they measured their index with great accuracy, there remains a question of what their index really measures. I will try to address that question here.

One can find the Leiden rankings online at `http://www.leidenranking.com/ranking.aspx` . The principal source of data for these rankings is the Thomson Reuters Web of Science database, which counts journal citations found in all articles from about 10,000 academic journals. The Leiden study considered articles, letters, and reviews that were published between 2005 and 2009 in the sciences and the social sciences. Publications classified as arts and humanities were excluded. The authors of the study had to determine the institutions with which authors of each paper were affiliated. In cases of multiple authorship, they divided the credit to each author for an article by the number of authors listed for the article.

Citation-based measures of total research output

The most direct measure of university research productivity to be found in the Leiden study is a simple count of the number of citations to papers published by authors associated with each university. Table 1 shows that, by this measure, UCSB ranks 44th in the United states and 7th in the University of California system.

Table 1: Citations to UC Faculty Papers

	Number of Papers by faculty	Average Citations per Paper	Total Number of Citations	UC Rank by Number of Citations	US Rank by Number of Citations
UCLA	17158	11.35	194801	1	4
UCSF	11166	14.71	164253	2	8
UCSD	12436	11.78	146528	3	11
UC Berkeley	11713	11.67	136709	4	13
UC Davis	12201	7.85	95778	5	22
UC Irvine	7202	9.79	70474	6	36
UCSB	5265	11.21	58996	7	44
UCR	3701	9.24	34190	8	74
UCSC	2217	10.04	22269	9	102

The Leiden study points out that there are significant differences between disciplines in the average number of citations. In some fields it is the norm for authors to cite many more papers than in others. For example, medicine has very long citation lists, and consequently, papers in medicine tend to be cited more often than those in other disciplines. The Leiden study attempts to control for this effect, by normalizing the count of cite numbers by average disciplinary citation rates. As we see from Table 2, when this effect is corrected for, the ranking of UCSB improves from 44th to 39th in the U.S . As compared with other UC campuses, UCSB remains at the rank of 7, though close in absolute terms to Irvine which is 6th.

Because there may be substantial differences in the average quality of research done at different institutions, a simple count of the total number of papers published by a university's faculty might be a misleading indicator of the importance of its research output. To control for quality differences as well as differences in citation patterns across disciplines, the Leiden group constructed another statistic. This is the fraction of all papers published by a university's faculty that appear in journals with impact factors in the top ten percent for their disciplines. If we multiply this fraction by the number of papers published by the university's faculty, we get an estimate of the number of "high profile" papers published by the university's faculty. At the top of this list, we find

Table 2: Number of Cites Normalized by Discipline

	Number of Papers by faculty	Normalized Citations per Paper	Total Normalized Citations	UC Rank by Normalized Citations	US Rank by Normalized Citations
UCLA	17158	1.52	26065	1	4
UC Berkeley	11713	1.79	20997	2	10
UCSD	12436	1.58	18919	3	12
UCSF	11166	1.58	17694	4	15
UC Davis	12201	1.32	16095	5	21
UC Irvine	7202	1.42	10227	6	38
UCSB	5265	1.91	10044	7	39
UCR	3701	1.52	5611	8	69
UCSC	2217	1.51	3356	9	102

Harvard, Stanford, and the University of Michigan, with UCLA in fourth place.

Table 3 reports the estimated number of high profile papers published by the faculty of each university in the UC system and shows their rank among all universities in the United States.

Table 3: Number of "High Profile" papers by Faculty

	Number of High Profile Papers	UC Rank	US Rank
UCLA	2994	1	4
UC Berkeley	2461	2	8
UCSD	2185	3	10
UCSF	2159	4	12
UC Davis	1702	5	18
UC Irvine	1191	6	27
UCSB	1136	7	29
UCR	627	8	39
UCSC	387	9	44

Of all the statistics produced by the Leiden group this measure seems to be the one that can most reasonably be called a measure of the "research impact" of a university. By this measure, UCSB ranks 29th nationally and 7th among UC campuses, slightly behind Irvine. It is impressive to see that four UC campuses rank in the top 12 in the nation and 9 of them rank in the top 45.

Adjusting for Size of Universities

The tables that we have considered so far measure total output of institutions. This seems appropriate if we really want to measure "research impact of an institution". But for many purposes, it is of interest to control for the size of the universities. A very large institution, might, for example, have mediocre per capita performance and still produce a larger number of high-quality research articles than a smaller elite university like Princeton or Cal Tech. Controlling for size is important if we want to know whether taxpayers or donors are getting good value for their money. In this case, we might want to divide the absolute measures of output by the size of "the research budget." or by the number of full time equivalent academic faculty employed. If we want a measure of the average research output of the university's faculty, we would again want to divide a measure of total research productivity by the number of faculty. If a prospective graduate student wants an estimate of the degree if exposure he or she will have to high quality work, we might want to divide by the number of graduate students, or by some weighted average of the numbers of graduate and undergraduate students.

Adjustment by Average Quality of Citations

The Leiden group was aware of the importance of controlling for size, but did not supply any direct measures of university size as measured by enrollment, budget, or faculty size. In a technical report [1] that discusses their methodology, they remark that

> "The Leiden ranking is based exclusively on output variables...Input variables such as the number of research staff of a a university or the amount of money that a university has for research, are not taken into account. Ideally, scientific performance should be measured on both input and output variables...However, internationally standardized data on input variables is not available..."

The Leiden group evaded this difficulty by constructing scale-independent quality measures that use citation data only, without the need to obtain data on inputs. They do this, using two measures of the average quality of the articles that have been published by a university's faculty. One of these measures is the average number of citations per paper (rather than the total number of citations) to papers from authors from this university, where cites are normalized to account for differences in citation practices between disciplines. A second measure of the quality of work by a university's faculty is the fraction of all papers written by university faculty members that are published in journals that rank in the top 10% of their discipline by impact factor. When universities are ranked by the percentage of their faculty papers that are published in journals in the top 10% for their disciplines, the top 6 universities are MIT, Princeton, Harvard, Rice Stanford, and Cal Tech. Remarkably, this distinguished group is joined in 7th place by UCSB. When universities are ranked by normalized

4

average citations per faculty member, UCSB fares even better, being ranking in fourth place in the United States.

Tables 4 and 5 show the scores for the Leiden group's two quality measures for universities in the UC system, along with rankings relative to the UC and relative to universities in the entire US.

Table 4: Ranking by Fraction of papers published in High-profile Journals

	Fraction of Papers in Top Journals	Rank in UC System	Rank in US
UCSB	21.6%	1	7
UC Berkeley	21.0%	2	8
UCSF	19.3%	3	10
UCSD	17.6%	4	17
UCLA	17.5%	5	18
UCSC	17.4%	6	19
UCR	17.0%	7	26
UC Irvine	16.5%	8	31
UC Davis	14.0%	9	56

Table 5: Ranking by Average cites per paper, Normalized by Discipline

	Mean Normalized Citations per Paper	Rank in UC System	Rank in US
UCSB	1.91	1	4
UC Berkeley	1.79	2	8
UCSF	1.58	3	14
UCLA	1.56	4	16
UCSD	1.52	5	19
UCR	1.52	6	20
UC Irvine	1.42	8	37
UC Davis	1.32	9	50
UCSC	1.51	7	21

There is room for concern about whether these measures reflect average quality of a university's faculty or some other difference in publishing norms, perhaps reflecting differences in academic disciplines. For example, if we consider two universities with the same number of faculty where each university has two types of faculty–an A list, who publish in top journals and a B list who publish in more pedestrian journals. Suppose that the A list faculty in

the two universities are equally productive, but in the first university, the B list faculty are active, publishing frequently in less prestigious journals and in the second university the B list faculty rarely publish. By the Leiden measure, the second university would be ranked as more productive than the first. This is true despite the fact that the first university produces more research and would seem to harbor a more positive research environment.

Further concern arises from the fact that another scientometric group, known as Scimago `http://www.scimagoir.com/`, found very different results when constructing a conceptually similar ranking using a different data source. Scimago, which is sponsored by the world's largest academic publisher, Elsevier, counts citations from a larger set of journals than the Thomson-Reuters Web of Sciences used by the Leiden group, and classifies the journals by discipline in a slightly different way. Beyond this, Scimago copied the Leiden quality index by calculating the fraction of all articles published by a university faculty that appear in the most cited 10% of journals in the author's field. The Scimago score ranks UCSB 508th among institutions in the world. The main reason for this very low ranking is that the Scimago study includes a large number of small research institutions and research hospitals that do little or on teaching, and which, though they may be very small, have very high average performance by their researches. By the Scimago score, UCSB ranks a quite respectable 34th among US universities. The fact that a conceptually identical measure using a slightly different method of citation counts and of classification of disciplines changes the ranking of UCSB from 7th to 34th suggests that the quality index used by Leiden may not be a reliable and robust indicator.

Adjustment by Direct Size Measures

In an effort to determine whether UCSB's ranking by the Leiden scale of "size-adjusted quality" reflects an important general truth or is just an artifact of one particular normalization, it seems appropriate to seek alternative ways of adjusting citations data to account for the sizes of universities. The Leiden group reports that they were unable to find "standardized measures of inputs" for universities worldwide. I have, however, found a few standardized measures of scale that are available for all universities in the United States and some that are available for all universities in the UC system and a handful of other universities.

Enrollment statistics are available for all US universities from the NCES survey of US colleges and universities, found at `http://nces.ed.gov/`. Table 6 below reports enrollments along with rankings by the ratio of high-profile papers to the total number of students. The ratio of high-profile papers by faculty to the total number of students at a university might be taken as a rough measure of the degree of exposure to highly productive faculty that a random university student could expect. Alternatively, we might take the view that faculty who produce high profile articles devote most of their teaching efforts to graduate students. In this case, the ratio of high-profile papers by faculty to the number of graduate students could be a useful indicator of the

Table 6: High-profile articles and Student Enrollment

	Total Enrollment	Graduate Enrollment	US Rank by high-profile articles per student	US Rank by high-profile articles per grad student
UCSF	3024	3024	1	3
UCLA	38172	11995	19	17
UCSD	29176	5513	21	7
UCB	35833	10293	23	20
UCD	31392	24670	27	16
UCSB	22218	3032	31	8
UCI	26994	5018	32	21
UCR	20692	2450	38	14
UCSC	17187	1519	41	15

expected quality of instruction that can be expected by a potential graduate student. Table 6 reports the total number of students and number of graduate students at each UC campus. This table also reports rankings in the UC system and in the United States of UC institutions by high-profile papers per student and high-profile papers per graduate student.

Ranked by the ratio of high-profile papers published to the total number of students, UCSB continues to rank at about 30th in the US and 6th in the UC system. However, if universities are ranked by the ratio of high-profile papers to the number of *graduate students*, UCSB ranks 8th in the US and 3d in the UC system.

The National Science Foundation `http://www.nsf.gov/statistics/nsf11313/` has collected data on total reported expenditure (of money collected from all sources) on Research and Development by each US university. If we divide the number of high profile papers, by total Research and Development expenditures, we have a measure of "bang-per-buck" for each university. Table 7 shows the results of ranking universities by the ratio of the number of high profile papers to R&D expenditures. By this measure, UCSB 1st in the UC system and 5th in the United States, behind Harvard, University of Texas at Dallas, Rice, and Princeton.

I have not been able to find standardized estimates of the faculty size for all universities in the U.S. However, statistics on the number of full time equivalent academic employees for universities in the UC system can be found at `http://www.ucop.edu/ucophome/uwnews/stat/`. I have used these FTE figures to compare the number of high profile articles and the normalized number of citations per academic FTE at UC institutions.

Tables 8 and 9 show that UCSB substantially outperforms all other UC institutions by either measure of research output per full-time-equivalent faculty.

Table 7: High Profile Papers and R&D Expenditure

	Expenditure on R& D (in $1000)	High Profile Papers per dollar R&D	Rank in UC System	Rank in US
UCSB	215728	5.27	1	5
UCR	130187	4.82	2	7
UC Berkeley	652479	3.77	3	17
UC Irvine	325493	3.66	4	19
UCLA	889955	3.37	5	24
UCSC	144052	2.68	6	31
UCSD	873357	2.50	7	36
UC Davis	681618	2.50	8	37
UCSF	947697	2.28	9	41

Table 8: High profile articles and Academic FTEs

	Academic FTEs	High Profile Papers per FTE	Rank in UC System
UCSB	2364	0.48	1
UC Berkeley	5613	0.44	2
UCSF	5401	0.40	3
UCLA	7815	0.38	4
UCSD	5857	0.37	5
UCR	1848	0.34	6
UC Irvine	4149	0.29	7
UC Davis	6194	0.28	8
UCSC	1537	0.25	9

Table 9: Normalized Citations and Academic FTEs

	Academic FTEs	High Profile Papers r per FTE	Rank in UC System
UCSB	2364	9.4	1
UC Berkeley	5613	3.74	2
UCLA	7815	3.41	3
UCSF	5401	3.28	4
UCSD	5857	3.23	5
UCR	1848	3.04	6
UC Davis	6194	2.60	7
UC Irvine	4149	2.46	8
UCSC	1537	2.18	9

We have also seen that by the measurements found in Table 7 that UCSB produces a larger number of high-profile research papers per dollar of the university R&D budget than any other university in the UC system.

Table 10: Students per FTE faculty

	Academic FTEs	Student Enrollment per FTE	Rank in UC System
UCR	1848	11.2	1
UCSC	1537	11.2	2
UCSB	2364	9.4	3
UC Irvine	4149	6.5	4
UC Berkeley	5613	6.4	5
UC Davis	6194	5.1	6
UCSD	5857	5.0	7
UCLA	7815	4.9	8
UCSF	5401	0.6	9

It is important to remember that universities produce undergraduate and graduate education as well as research papers. Another indicator of the cost-effectiveness of universities in the production of their intended outputs is the number of students per faculty member. By this measure, UCSB is also highly productive. In the UC system, only Riverside and Santa Cruz have higher ratios of students to faculty. The UCSB ratio of students to faculty is almost 50 per cent higher than that of Irvine, more than 50 percent higher than those of Berkeley, Davis, UCSD, and UCLA . Thus the data supplied in Tables 7-

10 seems to support a claim that that, in terms of cost-effective provision of research and teaching, UCSB ranks at the top of the UC system.

Conclusion

A press release by the UCSB Office of Public Affairs states that "The Centre for Science and Technologies Studies at Leiden University in the Netherlands has ranked UC Santa Barbara number 7 on its annual list of the top 500 major universities in the world...UC Berkeley and UC San Francisco were ranked 8 and 10, respectively." This is the kind of thing the Office of Public Affairs is paid to say. They are not paid to wonder how it is plausible that UCSB is somehow a greater university than Berkeley, UCLA , Yale, Chicago, and Michigan, or, on a less grand scale, that it outranks UCSD, Davis, and Irvine. On a closer look, the Leiden study does not claim that UCSB is "number 7 on a list of the top 500 major universities in the world." The Leiden study publishes several indicators of university performance and ranks universities by many of these indicators. One of the measures that the Leiden study used was the fraction of those papers published by university authors that were published in journals that ranked among the most prestigious ten percent of articles in their discipline. Doing well by this criterion is certainly a favorable sign, but there is no reason to believe that being ranked number 7 by this rather odd criterion gives a university any claim to being "number 7 on a list of the top 500 universities in the world."

In defense of the university publicists, the authors of the Leiden study do assert that the ranking on which UCSB scored number 7 is "the most important impact indicator of the Leiden Ranking". But a close look at the Leiden rankings data and a bit of thought makes one wonder what the Leiden group can possibly mean by this claim.

Certainly if we rank the impact of universities by total productivity as measured by either the number of citations to papers written by university faculty or the number of high profile papers written by UC faculty , UCSB does not come close to being number 7 among the top universities in the world. According to the Leiden statistics, UCSB ranks 39th in the United States and 68th in the world by number of normalized citations to works of its faculty and it ranks 29th in the United States and 59th in the world in terms of the number of high profile papers written by its faculty.

Even if there is reason to have reservations about the two quality measures by which the Leiden group places UCSB in the world's top 7, it is possible that these measures point toward significant virtues of this university. Our comparison of university inputs to outputs suggests that although UCSB's total scholarly output is overwhelmed by that of larger institutions like UC Berkeley, UCLA, Harvard, or Michigan, UCSB outperforms these institutions in terms of "bang-per-buck. " We found that the ratio of the number of high profile articles by UCSB authors to the size of the university research and development expenditures is higher than for these larger institutions. We also found that the

ratio of scholarly output, as measured by the Leiden group, to the number of full time equivalent faculty is significantly higher for UCSB than for any other university in the UC system. Finally we found that not only does UCSB perform comparatively well in producing scholarly output per unit of input, but it also educates a much higher number of students per full time equivalent faculty than its sister institutions, Berkeley, UCLA, Davis, UCSD, Irvine, and UCSF.

Legislators, potential donors, and system-wide administrators should find it interesting that UCSB seems to be unusually efficient in turning resources into scholarly output and instruction, whether we measure output per dollar's worth of research expenditure or output per full time faculty member.

There remains reason to be cautious about these assertions. Universities are made up of a large number of relatively independent departments which differ greatly in their missions and by their success in filling these missions. Rankings that aggregate the successes of all its constituent parts into a single university ranking are likely to be influenced in unforeseen ways by a weighting system that of necessity involves many arbitrary choices. Perhaps UCSB's favorable rankings are an artifact of the particular mix of disciplinary specializations at this institution, such as the fact that UCSB does not have professional schools such as a law school, a medical school, or a business school. While potential students and potential new faculty will place some value on the reputation and intellectual tone of the university as a whole, they be far more interested in the performance of UCSB in their own disciplines rather than in a poorly understood aggregate index. A better understanding of the comparative strengths and weaknesses of universities would arise from disaggregated discipline-by-discipline measures.

A department-by-department breakdown of the kinds of performance data calculated by the Leiden group would not only be valuable to students and scholars, but would also be likely to provide some explanations for the sometimes surprising differences found in the aggregate measures. Constructing such measures on a large scale like that of the Leiden study, would be extremely costly. However, such a study confined to UC universities and a few other reference institutions might be manageable with relatively modest resources.

References

[1] Ludo Waltman et al. The Leiden ranking 2011/2012: Data collection, indicators and interpretation. Technical report, Centre for Science and Technology Studies, Leiden University, The Netherlands, 2012. available online at `http://arxiv.org/abs/1202.3941`.