Cost efficiency and funding of academic research

Marc Luwel

luwel@cwts.leidenuniv.nl

Leiden University, Centre for Science and Technology Studies (CWTS),
Wassenaarseweg 62A, 2333AL Leiden (the Netherlands)

Submission for an oral presentation at either the track '*Evaluation of Research policies and Programs*' or '*Pitfalls and Perils of Internationally Comparative Higher Education Indicators'*

1. Introduction

Over the last decade policy makers and media have become increasingly interested in university rankings and in the efficiency of public investments in these institutes.

Given the universities' multifaceted mission academic rankings and their use have been criticised as overly simplistic. However they are a fact of life and will not disappear very soon. Ranking became a research topic and efforts are made to develop better ones and educate policy makers and the media on their use and potential pitfalls. Rankings can be divides into the more general covering all aspects of the university's mission and more specialised focussing on specific activities. The Shanghai and THES ranking are examples of the former and the Leiden Ranking developed by the Centre for Science and Technology Studies (CWTS) of the latter. This ranking focuses on research performance and is entirely based on bibliometric indicators and covers the 500 largest universities worldwide (in terms of number of publications processed for the Web of Science® (WoS) by Thomson Reuters). Using Leiden Ranking data of European universities van Raan and his collaborators extensively analyzed the relationship between productivity on the one hand and bibliometric indicators on visibility and impact on the other hand, showing size dependent cumulative advantages (see: van Raan, 2008 and references therein).

With the growing societal needs and the recent economic upheaval public expenditures are worldwide under pressure. Governments and the public at large increasingly put emphasis on the accountability of universities and on the efficiency of public investments in higher education and research. In many countries performance driven funding formulae are introduced often aimed at concentrating resources at the best universities and within universities at excellent research groups. The rationale behind these policies are often the tacit believe in economies of scale and economies of scope at the level of universities. However to measure efficiency of public investments in universities, even limited to research activities, turns out to be complex and bedevilled by severe conceptual and technical problems. The knowledge production and application are multi-input and multi-output processes, totally different from standard production activities studied in economic literature. On this subject a body of literature is available. The most relevant papers for the present study were written by Bonaccorsi, Daraio and Simar (see: Bonaccorsi et al, 2005).

In this paper for the largest US research universities the relationship between total research expenditures on the one hand and scientific productivity and impact on the

other hand is studied. In the next section we present the data acquisition. In the section on results the statistical relationship between the above mentioned variables is analyzed and in the preliminary conclusions the tentative implications for science policy and follow up research are discussed.

2. Data description and research question

In the 2008 edition of the Leiden Ranking of the top 250 universities worldwide ranked by the total number of publications 90 are from the United States (http://www.cwts.nl/ranking/). This ranking is based on data from the period 2003-2007 and provides for each university information on the total number of publications, the total number of citations and 'Crown Indicator', the size-independent, field-normalized average impact. This indicator is an expression of a university's citation impact in the precise research areas where the university's researchers publish their results, compared with the global average for the same research areas (for more information see : van Raan et al., 2010).

In the United States the Center for Measuring University Performance publishes the Annual Report of the Top American Research Universities. This report contains an overview of the total research expenditures of top 200 American research universities (*http://mup.asu.edu/index.html*).

Because the publication output lags funding from these reports of 89 US universities[1] in the Leiden Ranking the information on their total research expenditures in 2002, 2003, 2004, 2005 and 2006 was extracted. The one year shift between the funding data and publication data was introduced as a gimmick of the time lag between the initial investment and the publication output[2].

The 89 US research universities have different academic profiles. As the primary aim of the study is to analyze the relationship between funding and research productivity and impact using WoS data, out the 89 universities the subset of broad-based institutes was indentified offering academic programs in humanities, social sciences, natural sciences and engineering, and having a large medical school. Although these 59 universities, further called Fully Fledged US Research Universities, have a fairly similar profile, there are some differences in scope as for example a few do not have a law school and others have a veterinary school.

As it is very difficult to collect standardised funding data on university research, this data set is unique. The scope being fairly well controlled, it allows to explore at institutional level the presence of economies of scale for research productivity and impact.

3. Emperical results

[1] University of California – Los Alamos National Laboratory was not identified in the Annual Report

[2] To check for the robustness of the results a time lag of 2 and 3 years was also used. Due to space limitations these results are not presented.

For the Fully Fledged US Research Universities Figure 1 shows the annual average number of WoS publications and the annual average total research expenditures. For both variables the average was calculated over the above mentioned period. Harvard University with 11.425 publications and only $433 million total research expenditures per annum and Johns Hopkins University with $1.340.630.000 total research expenditures and 5.767 publications per annum are outliers.

This figure shows that both variables are reasonably well linearly correlated; the coefficient of the fitted regression R^2 is 0.87, not weighing the two outliers (on the total sample R^2 is 0.50). There is a difference of six order of magnitude between the input and output values. To test the robustness a linear regression on the log values of both variables was made resulting in fairly similar R^2 values : 0.81 not weighting the two outliers and 0.73 including Harvard University and Johns Hopkins University.

To further test the robustness of the results two alternative approaches are used. Firstly, six universities (±10%) were randomly removed and for the log-log fit the resulting variation in the slope was ±0.012. Secondly least square polynomial fits with increasing fit orders were done on the data. The Pearson and Spearman correlation coefficients were calculated between the data and the fitted values; the correlations became only marginally better.

To further explore scale effects between funding and productivity, local polynomial fits were done. Compared to standard polynomial fits, these fits are done on a subset of the data surrounding the point being estimated (see: www.srs1software.com). Figure 1 shows the results with half the source data as the fit region size and not weighting the two outliers. Except for the universities at the lower end of the total research expenditures having a relative stagnation of their publication output with increasing funding no local trends can be distinguished.

In comparison with the total number of publications the total research expenditures correlate slightly less well with total number of citations: R^2 is 0.74 not weighting the two outliers (R^2 = 0.36 for the full sample).

The total research expenditures and the 'crown indicator' do not correlate very well: R^2 is 0.25 not weighting the two outliers (R^2 is 0.23 for the full sample). Given the large difference between of the values of the dependent and the independent variable this coefficient was also calculated on the log-log values; with 0.21 R^2 is only slightly different. As illustrated in Figure 2 the slope of the linear fit is also very small. On the data local polynomial fits with different fit orders and fit region sizes were made and no local effects were detected.

4. Discussion

The present study extents the work of van Raan (2008) by introduction funding as an extra dimension to Leiden Ranking data and complements the work of Bonaccorsi et al. (2005) on efficiency measurements. Using a sample of US research universities with a similar scope no economies of scale were detected between funding and productivity measured by the number of publications. In the statistics two top universities are outliers: the Harvard University with slightly more than the sample's average research expenditures but nearly four times the average output; Johns Hopkins University has 3.5 times the average research expenditures but only twice the average output. In follow up work additional information on both universities will

be collected. Although the correlation is not as pronounced as for the publications the relation between research expenditures and the citations is fairly similar.

The results for the Crown Indicator are more surprising. The values of this normalised impact indicator are very noisy and the small slope of the fitted curve indicates that, if any, there is only a weak relationship between a university's value of the crown indicator and its research expenditures.

These results should be put in the right context as the universities in the sample obtain most for their research funding in a highly competitive environment and are among the world elite institutes. A hypothesis to be further explored could be that these universities have reached maximum visibility. With additional funding more publications are produced without increasing the overall already high impact of the institute's work.

In ongoing work more sophisticated statistical techniques will be used and data on total research expenditures of a limited number of European top universities studied by van Raan (2008) will be added to the sample to further explore the efficiency of research funding.

References

Bonaccorsi, A., Daraio, C., & Simar L. (2005). Advanced Indicators of Productivity of Universities. An Application of Robust Nonparametric Methods to Italian Data. *Scientometrics*, 66(2), 389-410.

Van Raan A.F.J. (2008). Bibliometric Statistical Properties of the 100 Largest European Research Universities: Prevalent Scaling Rules in the Science System. *JASIST*, 59(4), 565-576.

Van Raan, A.F.J., Van Leeuwen, T.N., Visser, M.S., Van Eck, N.J., & Waltman, L. (2010). Rivals for the crown: Reply to Opthof and Leydesdorff. *Journal of Informetrics*, *4*(3), 431–435.

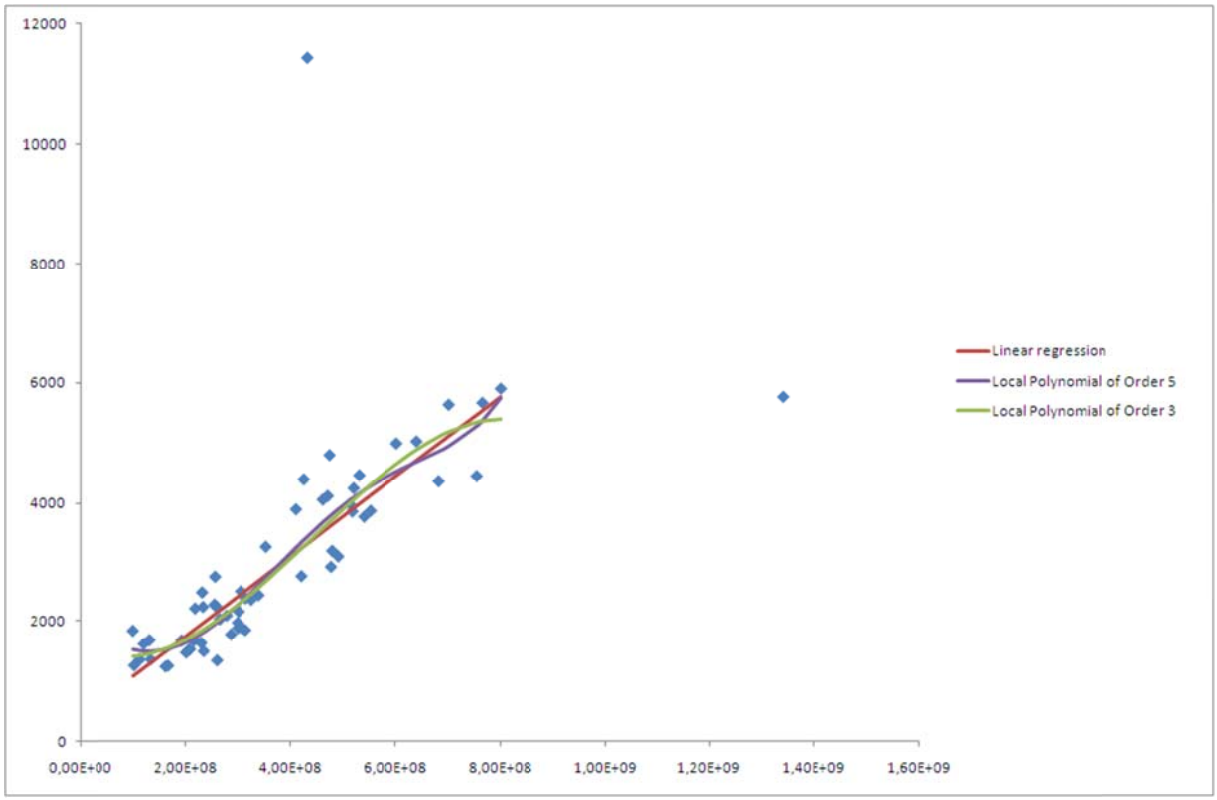

Figure 1. Correlation of the number of publications (P) with the total research expenditures (F) in US$ for 49 Fully Fledged US Research Universities (P= 425,98+ 6,66E-06 * F; R^2=0.87). For the regression curves the two outliers were not weighted and for the Local Polynomial Fit the FitRegionSize is 0.5.

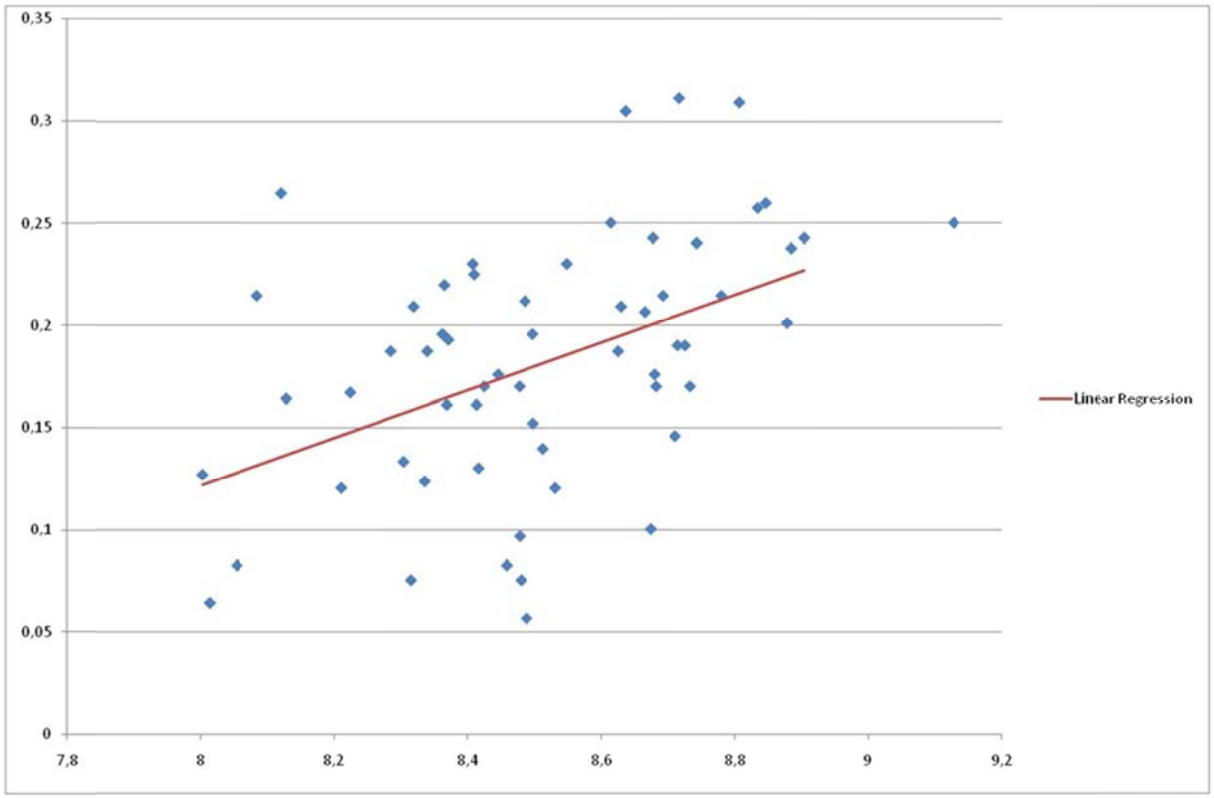

Figure 2. Correlation of the Crown Indicator (CPP/FCSm) with the total research expenditures (F) in US$ for 49 Fully Fledged US Research Universities using for both variables a logarithmic scale (log10(Cpp/FSCm)= -0.813 + 0.117 * log10(F); R^2=0.21). For the regression curve the two outliers were not weighted.