Publications or Citations – Does it Matter? Beneficiaries in Two Different Versions of a National Bibliometric Performance Model, an Existing Publication-based and a Suggested Citation-based Model

Jesper W. Schneider

jws@ps.au.dk
The Danish Centre for Studies in Research and Research Policy, Department of Political Science, Aarhus University (Denmark)

Abstract

The paper discusses the adoption of the Norwegian Publication Model in a Danish context and examines arguments for supplementing or substitution the current mechanism where reward is based on publication activity with one based on citations. Based on national publication data from 2009 from the Danish model, belonging to the science and technology research area, and corresponding citation data, we examine the Danish universities' relative input when it comes to publications and subsequently examine the relative output from these publications, i.e., the "returns on investment" from the model, either the current publication points, or the alternative, citations. Findings support the claims that high-performing units would benefit more from a citation-based approach, but at the same time also show, contrary to what was conjectured, that in the present case the same university also benefits the most from the current publication model. Based on the findings, we discuss the publication versus citation-based models, or hybrids between them, and argue that citation-based models in performance-based funding context are harder to influence and most likely will support already existing cumulative advantages.

Conference Topic

Science policy and research assessment

Introduction

In recent decades several countries have introduced performance-based research funding among their universities (Hicks, 2012). The performance-based research funding systems (PRFS) vary considerably between countries, from panel-based peer review evaluations, to systems based on citation or publication metrics, or various hybrids of these three basic forms (see Hicks, 2012). Generally, peer review systems are considered superior to systems based on bibliometric indicators (see Gläser & Laudel, 2007). Nevertheless, large-scale panel evaluations are very expensive, and several *post hoc* comparisons between panel results and citation metrics, for example from the UK Research Assessment Exercises, suggest that the latter could be an effective, and cost-effective, supplement or even substitute to peer reviews (e.g., Oppenheim, 1996; Moed, 2008). Among PRFS based on bibliometric indicators, citation-based systems are considered by some to be superior due to the assumption that citation indicators to some extent are able to measure aspects of research quality by focusing on impact (Gläser & Laudel, 2007). But citation indicators also have obvious deficiencies especially when implemented in PRFS which in principle are supposed to cover all fields of research (Schneider, 2009). It is well-known that citation indicators are not equally valid across all fields of research and even where relevant, coverage in the citation databases is also restricted (Moed, 2005). Consequently, PRFS based on citation indicators severely restricts the measurable outcome of research basically to journal articles indexed in one of the two major citation databases. But there are other issues with citation indicators which can be considered inadequate when used in PRFS, especially when such systems are supposed to (re)distribute funding on a regular basis, most often annually, and at the same time also give

universities (and their researchers) incentives to improve performance (e.g., Gläser & Laudel, 2007; Schneider, 2009). Citation indicators reflect research done in the past often a considerable number years prior to the actual funding year. It is also very difficult to directly influence citations when conceived of as an incentive system, in fact the well-known cumulative advantages could be detrimental to such an incentive system if it is supposed to be fair for all involved (Merton, 1988). Such features are seen by some as undesirable if PRFS as supposed to cover all research fields with their different publication traditions, and be able to reflect recent research performance in a dynamic model, as well as give transparent behavioural incentives to change performance (Schneider, 2009; Hicks, 2012).

PRFS based on publication activity have been introduced as an alternative to citation-based systems (Butler, 2002; Schneider, 2009). There are some apparent "benefits" with publication-based systems compared to citation-based systems. They can reflect short-term research activity making them more up-to-date when it comes to redistributing funding. In principle they can encompass all desired publication types and they can provide straightforward behavioural incentives. But it is important to emphasise that the two approaches measure different constructs. It would be naïve to suppose that incentives directed at publication behaviour, i.e., quantity and/or supposed status of the publication outlet, encompass the same aspects of perceived "quality" that citation impact is thought to reflect (Schneider, 2009). Experiences from Australia testify to this. In a succession of papers, Linda Butler demonstrated how researchers in Australia responded when funding, at least partially, was linked to publication counts undifferentiated by any measure of supposed "quality" in the early 1990s (e.g., Butler, 2003a; Butler, 2003b). Australian publication output increased considerably with the highest percentage increase in lower impact journals. For a consecutive number of years, this lead to a general drop in overall citation impact for Australia. Since Butler's documentation of the adverse effects, the experience from Australia has stood as a "warning" for what would most likely happen if funding was linked to publication activity. Nonetheless, in the early 2000s a so-called "quality reform" of the higher education sector in Norway introduced a PRFS where publication activity again was linked to funding. The main political intention with the model was in fact to encourage more research activity and thereby also more publication activity, and preferably more international publication activity, in the university sector[1].

The so-called *Norwegian Publication Model* (NPM) is interesting in in relation to PRFS. Obviously, the designers of the NPM were well-aware of the adverse behavioural effects documented in the Australian case. As a consequence, a slightly more sophisticated model was developed (Schneider, 2009; Sivertsen, 2010). A primacy of the model was to reflect the encouragement to publish in international outlets (i.e., international journals and academic book publishers) and at the same time to counter so-called adverse publication effects like the Australian case, where researchers seek to publish more but with less effort. Hence, a differentiated publication model was constructed where publication channels were classified on two levels. Level one comprises in principle all scholarly eligible publication channels, where eligibility criteria are some basic norms such as a standard external peer review process. Level two, is an exclusive number of publication channels, which are deemed to be leading in a field and preferably with an international audience. Level two is exclusive in as much as the number of publication channels designated at any given time to this level should produce roughly one-fifth of the publications produced in a field "world-wide". Correspondingly, three different types of scholarly publications are included in the model: journal publications (articles and reviews), articles in books (contributions to anthologies and

[1] http://www.uhr.no/documents/Rapport_fra_UHR_prosjektet_4_11_engCJS_endelig_versjon_av_hele_oversettelsen.pdf.

conference papers) and books. A two dimensional point system was implemented where the different publication types yield different points within the same level and between the two levels depending on the outlet status. Hence, the basic idea behind this two-tiered classification system is that publications on level two receive more publication points than publications on level one. Finally, publication points are fractioned $1/n$ so that an institution eventually receives $1/n$ points depending on their number of contributing authors.

Eventually the annual sum of publication points for an institution is exchanged for funds, where the exchange rate is determined by the amount of money available for redistribution and the total number of publication points in the system in a given year. A noticeable assumption in the NPM is that publication behaviour, publication activity and publication types across all fields can be treated identically. Consequently, all research fields' eligible research publications are included in the model, which for example means that a level one journal article with one author is worth the same in physics and literature studies. It is assumed that the differentiated point system together with fractionalized counting will level out the major differences in publication behaviour between the fields and also to some extent will discourage researchers to speculate in "easy publications" resulting in a levelling out effect at the aggregate level. Consequently, in the Norwegian PRFS funding is competitive not only between institutions but also across all fields. Hence, the subject composition within and between the research institutions is interesting as performance improvement in one major area, in principle can lead to improved funding at the expense of another major area due to the basic zero-sum situation.

The NPM has recently been "adopted" in several European countries, for example in Denmark, Finland and Flanders (Hicks, 2012; Verleysen, Ghesquière & Engels, 2014). In the present paper we look at the "adoption" of the indicator in Denmark and examine the overall distributional consequences of focusing on publication activity and not impact.

It is important to accentuate that in Norway the publication model was to a large extent developed to support overall political goals, i.e., more international research activity. As it were, Norway's internationalization in research and general citation impact, were considerably lower, than for example Denmark, at the time of the introduction of the model. Since then Norway's international publication output has risen considerably, albeit rise in citation impact has been meagre (e.g., Aagaard, Bloch & Schneider, 2015). Nonetheless, the NPM was developed and implemented with a legitimate goal which to some extent seems to have been achieved seen from the national policy perspective.

During a reform of the Danish research funding system in the mid-2000s it was decided to implement a PRFS officially in order to enlarge competition among universities for funding, although the board of university rectors probably more saw it as management tool that should legitimize their overall research activity to the public (Schneider & Aagaard, 2012). The political process leading to the "adoption" of the NPM in Denmark is complex and documented in Aagaard (2011). It is not totally clear why the choice fell upon the NPM, although its coverage of all areas, transparency and clear incentive system were no doubt deemed viable, yet some actors actually indicated that it would probably be "the one that would cause the least damage" (Aagaard, 2011). Most interesting, contrary to Norway, there were no immediate strategies or goals for research and publication behaviour behind the "adoption" of the NPM in Denmark.

Denmark was the first country to adopt the NPM at a time when the model was still in its infancy in Norway and little empirical evidence of its potential effects was available. The NPM was adopted with very few moderations, as if the model was a one-size fit all package suitable for all contexts. Most notably, the simple two-tiered classification system was kept and considerations about expanding or adapting the classification to a Danish context were not done. Nevertheless, some seemingly minor moderations turned out to be imperative,

including a maximum fractionalization of contributions at 1/10th; but perhaps most important, performance-based publication activity was locked between the major research areas: science and technology, health sciences, social sciences and humanities. Consequently, in the Danish adoption of the NPM, funding is not competitive across areas only within areas. Further, politically it was decided to more or less keep the old annual allocation model between the areas which effectively meant that a publication point, contrary the Norwegian PRRS, have different monetary values across the four main research areas. This is an extremely important deviation from NPM and it gives rise to some questions about the Danish adoption of the NPM, popularly known by the acronym BFI (bibliometric research indicator).

One can argue that the model is transparent, seemingly coherent and all-inclusive when it comes to research areas. All areas are measured with same indicator. But since competition is restricted to within areas and as a consequence publication points have different values across areas, one could also ask why the model still assumes equality of publication practices across areas? And to go further, with the locking of the competition to within areas, there is basically no reason why fields where citation analysis could be a reasonable and indeed preferred indicator could implement such devices either in combination with a publication model or alone. Of course the latter would muddle the overall model, although it would probably satisfy many of the critics of the publication-based model, arguing for more emphasis on impact.

Indeed, the Technical University of Denmark (DTU) has been an ardent critic of the adoption of the NPM in Denmark. A common argument goes: Why implement an incentive model that reward publication activity in international outlets when "we" already do that and do it well? More generally the critics stated that the behavioural goals with the model in Norway were irrelevant in a Danish context, because Denmark, contrary to Norway, has 1) for decades consistently been among the top five highest performing countries when it comes to impact; 2) has consistently four of its eight universities in the top 200 of the Leiden Ranking[2]; and 3) the Danish research system has had a long trajectory of internationalization (e.g., Karlsson & Persson, 2012). According to DTU, what should be procured and rewarded is impact and not publication activity. While the argument is relevant, it is also self-serving. DTU happens to be the highest performing Danish university when it comes to impact and is ranked in the top 50 of the Leiden Ranking. DTU has a very strong focus upon science and technology and close to no medical, social or humanistic research activities. Also, DTU has the lowest student to researcher ratio in Denmark. Obviously, DTU would fit very-well to a model based on citations. DTU has continued the criticism over the years claiming that they are the actually "losers" in the current Danish PRFS. According to DTU, universities are reward for quantity and not "quality" which should always be the focus in research. Why risk the current impact status by increasing output for some marginal gains? This cannot be a national interest.

So goes the argument - what we examine in this paper is to what extent the argument holds. Who benefits from the current Danish publication-based model and is DTU the current "losers"? What would be the differences if a citation-based approach was applied instead?

The aim of the analysis is to examine the universities' "return on investment". We take a simple approach where we examine the relative input of the universities when it comes to publications and subsequently examine the relative output from these publications, i.e., the rewards in the model, either the current publication points, or the alternative, citations. We keep the analysis simple using basically a zero-sum approach, like the current model, where gains somewhere mean losses elsewhere.

[2] www.leidenranking.com

The next section briefly presents the data and main methods and indicators used for the analyses. The subsequent section presents main results, and the final section contains a brief discussion of the findings.

Data and methods

The paper examines the first full publication year (2009) used for redistributing funds in the Danish model. We are able to measure the citation impact of the Danish journal publications from 2009 and make comparisons between the Danish universities and examine their potential gains and/or losses by using either differentiated publication counts or citations. We compare publication counts and points derived from the BFI model between Danish universities, and we likewise compare the impact between these universities for the 2009 journal publications indexed in Web of Science (WoS). As argued in the introduction section, locking the main research areas in principle means that the current publication-based model could be adapted to specific behaviours and wishes, or even supplemented or exchanged with a citation approach, in the individual areas, although citations would only be relevant in the areas: science and technology and medical and health sciences. In this paper we focus the analysis on the main research area of science and technology. We do this because the issue concerning citation impact versus publication activity raised by DTU is directly linked to this area due to DTUs research profile. We have done a corresponding analysis for the medical and health sciences but due to limited space we will not address them in this paper.

The publication activity in 2009 in the main research area of science and technology is around 8700 publications of all types eligible in the BFI model, books constituted 2%, articles in books 19% and journal articles 79%. It is reasonable to argue that (international) journal publication is the primary publication activity in this area, which means that citation analysis of eligible articles is a sensible endeavour. However, as the area includes some fields known to have their main publication activity in conference proceedings (i.e., articles in books), we do scrutinize the influence of proceedings papers on the total number of BFI points acquired for the individual universities and discuss that in relation to the citation analysis where proceedings papers are excluded. Notice, we do not include conference papers in the citation analysis due to the meagre quality of the current proceedings citation indices.

All journal publications published in 2009 reported by the universities to the BFI-indicator were extracted from the BFI database. Subsequently, paper titles were extracted, and so were first author names and journal names. These parameters were used to match the publications with Danish WoS journal publications from 2009 using CWTS's in house version of WoS. Eligible publication types are research articles and reviews. The match rate is 77% of the initial journal articles. Among the non-matched publications were non-English language articles, as well as false positive articles, articles not eligible for the BFI model, but still succeeded in accruing points.

As indicated in the introduction section, the BFI model applies a fractional counting method at the institutional level where articles are fractioned up to $1/10^{th}$ among the participating institutions. We do not apply the exact same counting formula for the WoS publications going into the citation analysis. Here we simply do a straightforward fractional counting on the institutional level. As will be clear from the results section, this small deviance had no practical relevance on relative publication shares.

We use standard CWTS citation indicators from the Leiden Ranking (www.leidenranking: P_{frac} (fractionalized publications), TNCS (total number of normalized citations), MNCS (mean normalized citation score) and PPtop10% (proportion of papers for a unit among the 10 percent most cited in the database) (Waltman et al., 2012).

Eight universities are included in the Danish PRFS. The universities differ considerable in both subject/faculty composition and size. We have two "old" universities basically covering

all four main research areas included in the BFI model: Copenhagen University (KU) and Aarhus University (AU). These universities are also the largest universities in Denmark with long research traditions and strong science faculties. University of Southern Denmark (SDU) is a younger university, but its subject/faculty composition is basically a reflection of KU and AU, although the size is considerably lower. Roskilde University (RU) and Aalborg University (AAU) are even younger, from the mid-1970s. These universities have regional obligations with a substantial emphasis on teaching. Nevertheless, both universities have developed unique research profiles, both universities have focused on interdisciplinary research, where RU has a strong focus on the social sciences and AAU has focused strongly on engineering. Both universities have science and technology faculties, albeit at RU the size is only comparable to a large department. The Information-Technology University is the youngest and smallest university in Denmark. Their focus is mainly outside the science and technology areas but we include them here for numbers to add up. Likewise, Copenhagen Business School (CBS) is also included for matters of completeness in the analyses, their publication activity in the science and technology area are scanty. Finally, as discussed in the introduction, the Technical University of Denmark (DTU) is basically a "mono-faculty" university, albeit its activities are spread between science and technology. It is important to emphasise that while the university is known for primarily educating engineers, it has a considerable research activity in what would be considered basic natural science fields as well. In fact DTU can be dated back to the early nineteenth century where it was part of Copenhagen University, making it the second oldest university in Denmark. We recapitulate, DTU has been particularly dissatisfied with the Danish PRFS arguing that - for them at least - citations would be a more appropriate and valid performance-based indicator. In the next section we examine the consequences of this claim.

We calculate basic statistics based on individual articles both for the publication-based model and the simple citation approach we apply. As stated in the introduction, we take a simple approach where we examine the relative input of the universities when it comes to publication shares and subsequently examine the relative "rewards" the universities archives from these publications, i.e., the output in the model, either shares of the total publication points, or the alternative, shares of the total number of citations. Also, we keep the analysis simple using basically a zero-sum approach, like the current PRFS, where gains somewhere mean losses elsewhere.

Results

Table 1 below shows the eight universities' total number of matched fractionalized WoS publications belonging to the science and technology area, as well as their accumulated number of normalized citations after four years. Notice, these are fractionalized WoS publications, the absolute number of publications is 6,117.

Table 1 also shows relative citation performance for the eight universities using the MNCS and PPtop10% field normalized indicators.

The three main actors measured by volume is not surprisingly KU (32.9%), DTU (28.7%) and AU (21.4%), the volumes for AAU and SDU are considerably lower, both universities have a share of 7.2% of the total volume. DTU has the largest number of normalized citations among the eight universities. It is noticeable that DTU's share of citations (34.8%) is markedly higher than their share of publications (28.7%). Obviously, this is also reflected in the relative citation indicators. The MNCS at 1.66 is considerably higher than the average of the database and a score that would rank DTU among the top 30 in the Leiden Ranking if we only focused on science and technology, and among the top 50 for all fields combined.

Table 1. Science and technology: Number of fractionalized publications in WoS, total number of citations and relative citation indicators.

	WoS pubs (P_{frac})	TNCS	MNCS	Share of total P_{frac}	Share of total no. of NCS	PPtop10%
AAU	225.3	284.4	1.26	7.2%	6.6%	12.3%
AU	673.0	874.5	1.30	21.4%	20.3%	14.6%
CBS	13.2	12.2	0.93	0.4%	0.3%	
DTU	904.9	1498.8	1.66	28.7%	34.8%	17.0%
ITU	11.5	9.1	0.79	0.4%	0.2%	
KU	1035.9	1281.0	1.24	32.9%	29.7%	13.4%
RU	56.9	61.3	1.08	1.8%	1.4%	10.7%
SDU	227.7	284.8	1.25	7.2%	6.6%	15.8%
Total	3148.2	4306.3		100%	100%	100%

Interestingly, we also see that the minor universities, SDU and AAU, have relative citation indicator scores comparable to the larger universities KU and AU. In fact, SDU has more of their 2009 publications among the 10% most cited in the database compared to KU and AU. Overall, these results confirm what we suspect and are essentially the basis for the argument about including citations in the BFI model advanced by DTU.

In order to examine "return on investment", i.e., the institutions' reward for their publication input, we have calculated their share of BFI publications and BFI points for 2009 for the science and technology area, as well as the shares of fractionalized WoS publications and the total number of field normalized (fractionalized) citations. We thereby assume that shares of BFI points and shares of normalized citations can be treated equally. In the final discussion section we reflect upon this. We do, however, think that the straightforward approach taken can give a rudimentary indication of potential differences in "returns" for the individual institutions if one was to apply a citation based approach instead of or as a supplement to the current differentiated publication-based indicator in the science and technology area.

Table 2 below shows the shares of BFI publications and BFI points, where all publication types used in the science and technology fields are included (e.g., also conference proceedings), as well as shares of fractionalized WoS journal articles and normalized citations.

Table 2. Science and technology: Distribution and shares of BFI-points, BFI-publications, plus fractionalized publications from WoS and total number of normalized citations; notice <u>all</u> BFI-publication types are included.

	BFI-points	BFI-publications (P)	Share of BFI-points	Share of total BFI P	Share of P_{frac} (WoS)	Share of total no. of TNCS
AU	1814.9	1766	19.1%	20.4%	21.4%	20.3%
CBS	6.9	6	0.1%	0.1%	0.4%	0.3%
DTU	2854.1	2378	30.1%	27.5%	28.7%	34.8%
ITU	117.4	107	1.2%	1.2%	0.4%	0.2%
KU	2730.9	2457	28.8%	28.4%	32.9%	29.7%
RUC	185.9	157	2.0%	1.8%	1.8%	1.4%
SDU	571.0	572	6.0%	6.6%	7.2%	6.6%
AAU	1203.6	1219	12.7%	14.1%	7.2%	6.6%
	9484.8	8662	100%	100%	100%	100%

Table 3 below shows the same variables as Table 2, but in this case we *only* use the BFI publication type journal articles and the points derived from these articles. Table 3 is included for comparison because the citation analysis in reality only deals with journal articles. Notice, the BFI journal articles include non-WoS indexed articles, which give points in the indicator, however, the numbers are very low, the coverage of the science area in WoS is very high.

Table 3. Science and technology: Science and Technology: Distribution and shares of BFI-points, BFI-publications, plus fractionalized publications from WoS and total number of normalized citations; notice only the BFI-publication type journal article is included.

	BFI-points (journals only)	BFI-publications (P) (journals only)	Share of BFI-points (journals only)	Share of total BFI P (journals only)	Share of P_{frac} (WoS)	Share of total no. of NCS
AU	1526.2	1515	21.9%	23.5%	21.4%	20.3%
CBS	6.9	6	0.1%	0.1%	0.4%	0.3%
DTU	2007.4	1663	28.8%	25.8%	28.7%	34.8%
ITU	53.3	39	0.8%	0.6%	0.4%	0.2%
KU	2166.8	2047	31.1%	31.8%	32.9%	29.7%
RUC	139.5	126	2.0%	2.0%	1.8%	1.4%
SDU	420.1	442	6.0%	6.9%	7.2%	6.6%
AAU	657.5	596	9.4%	9.3%	7.2%	6.6%
Total	6977.7	6434	100%	100%	100%	100%

For analytical and illustrative reasons we plot the results from Table 2 and 3 in Figures 1 and 2 below. Figure 1 shows the results based on all BFI publication types, whereas Figure 2 shows the results where only BFI journal articles are included.

The figures are simple plots were the shares of the total number of publications (i.e., both BFI publications and fractionalized publications from WoS) for the eight universities constitute the x-axis, this is the "input", i.e. what the individual institutions "invested" in the Danish performance-based model for science and technology in 2009. The y-axis shows the shares of BFI points and citations, this is the "output", i.e. the institutions' "return on their investment" in the Danish performance-based model for science and technology in 2009. The axes are symmetrical and the diagonal shows the point where the institution has the same relative share of input (publications) and output (BFI points or citations). The distance from the university to the diagonal suggests whether input is larger than the return (output), which means that the institution will be below the diagonal, or the return (output) is larger, in which case the university is placed above the diagonal. Further, each university is plotted two times, one for the BFI data and one for the WoS citation data. Significant changes between these two representations for a university up and down the diagonal, suggest that the university receives a substantial number of BFI points from publication types other than journal articles. Notice in order to avoid confusion when examining the figures, shares of BFI publications on the x-axis should be compared to shares of BFI points on the y-axis, and likewise shares of WoS publications on the x-axis should be compared with shares of citations on the y-axis.

It is clear from Figure 1 that RU, CBS and ITU are not interesting for the current analysis as their numbers and shares are too low. We are interested in the other five universities, which all have a faculty of some size within science and technology. Interestingly, from Figure 1, where *all* BFI publication types are included, we can see that DTU actually has a larger output than input with a ratio of 1.09. This is somewhat unexpected and contrary to the conjecture that DTU is not gaining much from the current model. If we then turn to the

citation analysis, then we can see an even larger distance from the diagonal to DTU, compared to the BFI data, but also all other universities. The ratio is 1.25, so in line with the previous findings, DTUs WoS publications receive considerably more citations than the other Danish universities in 2009 but also the average paper in the WoS database. If a citation-based indicator of some sort were constructed where points were given based on citations, as implied in the arguments from DTU, then it seems that DTU would benefit from such a model, obviously conditioned on how it was designed. However, the most interesting finding here is perhaps that DTU within the science and technology area also seems to be the largest beneficiary when it comes to BFI points earned per input publication. Notice, like the current PRFS, we also treat it as a zero-sum game. If all universities improve then we have status quo. As it is in Figure 1, only DTU seems to really benefit from the citation approach. While KU seems to be in balance with the BFI data, they experience a smaller drop in returns on their input in the citation approach. Perhaps the most remarkable result from Figure 1 is the dramatic drop on the diagonal between BFI data and WoS citation data for AAU. We return to this below.

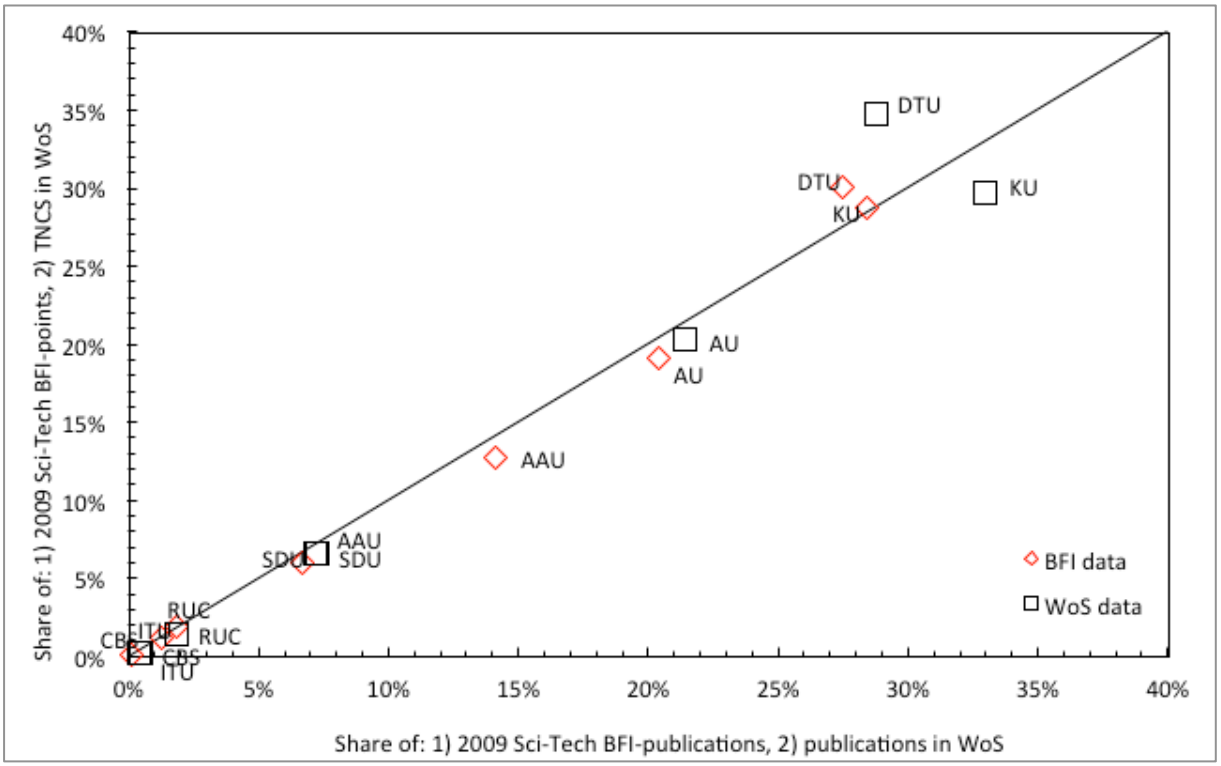

Figure 1. Science and technology: 1) Shares of 2009 Sci-Tech BFI publications as a function of shares of 2009 BFI-points; and 2) Shares of 2009 Sci-Tech WoS publications as a function of shares of total number of citations to these publications; notice BFI data includes all publication types.

Figure 2 depicts the same analysis but this time we have reduced the BFI data to include only journal publications in order to compare like with like, i.e., BFI journal data with WoS journal data. Obviously, the WoS data are identical to Figure 1, what is changing is the relative shares of BFI data (i.e., shares of publications and shares of points). There are some minor repositions, but the two major differences are the large drop on the diagonal for AAU and the corresponding smaller drop above the diagonal for DTU. Notice, the input-output is in balance for AAU, whereas DTU still has a substantial "return on investments" when it comes BFI journal data.

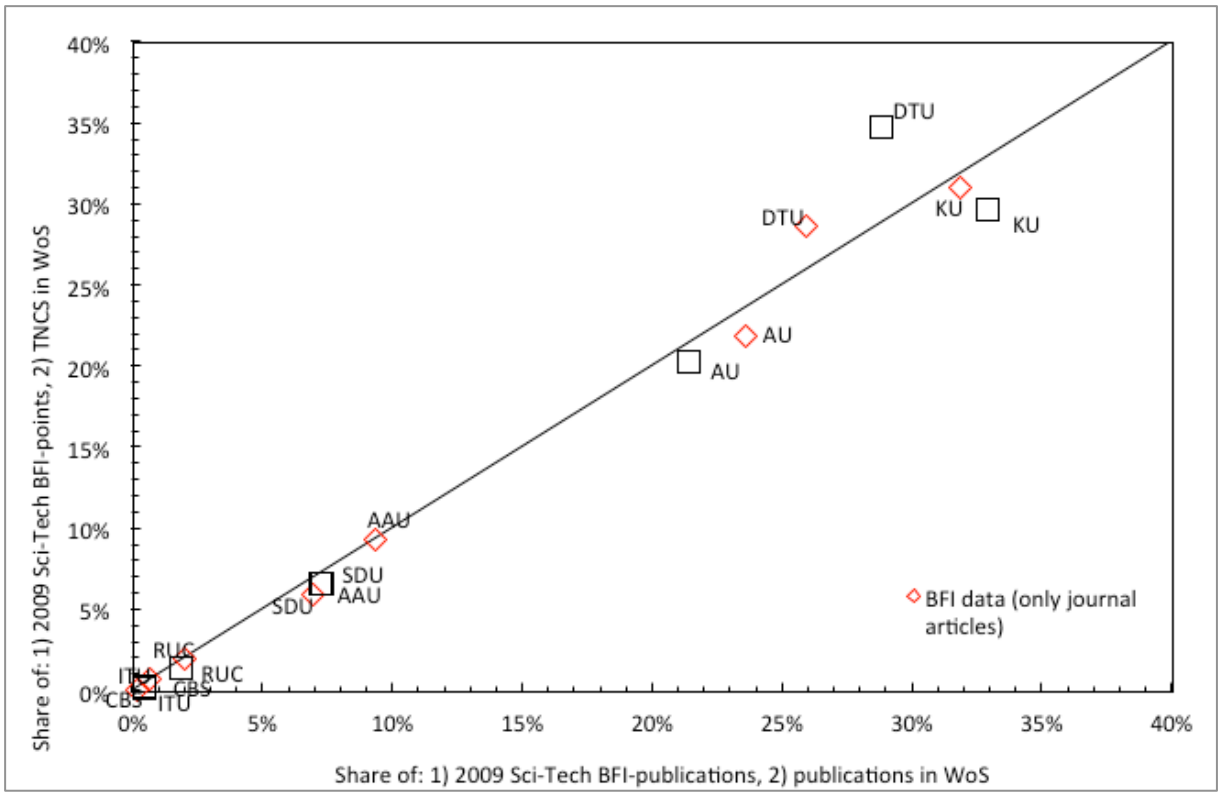

Figure 2. Science and technology: 1) Shares of 2009 Sci-Tech BFI publications as a function of shares of 2009 BFI-points; and 2) Shares of 2009 Sci-Tech WoS publications as a function of shares of total number of citations to these publications; notice BFI data only includes the publication type journal articles.

The drop of AAU along the diagonal was foretold in the WoS data in Figure 1. Here we saw a considerable distance between the BFI data when they included all publication types and the restricted WoS journal data needed for the citation analysis. For obvious reasons, this gap has been shortened considerably in Figure 2 since both data sets are restricted to journal articles.

The discrepancy in Figure 1 and the drop in Figure 2 are caused by the deviant publication profile for AAU compared to the other four universities with substantial publication activity in the science and technology area. Interestingly, 41% of the BFI publication activity in 2009 for AAU is in the category "articles in books", which in this case essentially means conference papers, and 49% is journal articles. For a comparison, 21% of DTUs activity is in "articles in books" and 70% in journal articles. These are both universities with strong focus on the technical sciences where publication in conference proceedings is very important. To contrast these profiles, the three other universities, KU, AU and SDU, all have more traditional science faculties and their relative publication activity in "articles in books" is 9%, 9% and 14% respectively. For these universities, due to their strong focus on science and less focus on technology, journal publication is the main activity 83% for KU, 86% for AU and 77% for SDU. However, we can also see that DTU does indeed have a strong science focus judged from their strong journal publication profile.

Considering the impetus for DTU to argue for a citation model, it is interesting to notice that while DTU clearly has the highest citation performance among the eight universities based on the 2009 journal publications, as we expected, they also have the highest performance when it comes to BFI publication points. Indeed, it seems that DTU would benefit even more in the science and technology area if they were to be rewarded for their relative share of the total number of citations, but contrary to the expected and suggested, DTU also benefit the most when it comes shares of BFI publication points compared to their relative input in the science

and technology area. DTU seems not only to be the most efficient when it comes to citations, this is also the case when it comes to BFI publication points. For example, the size of KUs activity in the science and technology area is larger than DTUs, but DTUs average point per publication is 1.20 for both of the above-mentioned analyses, considerably higher than KUs at 1.11.

Discussion

The main immediate findings in the present case study is that DTU will most probably benefit from a citation model, but perhaps more important, that they also seem to be the relatively most efficient university when it comes to BFI publication points. What are the more general implications of these findings seen in relation the current spread of the NPM to a number of European countries? The Danish case is special because competition is locked within the main areas this opens up for adapted models across areas including citation models where relevant. In Sweden a citation model is currently in use encompassing all fields. This is undesirable for several reasons; one of them is clearly demonstrated in this analysis, the desire to embrace all major publication behaviours, one of the rationales for the original NPM. A citation model alone restricts data to journal articles indexed in one of the two major citation databases. It was clear from Figure 1, that a university with an emphasis on technical sciences, like AAU, will be reduced in relative size when it comes to sharing the output.

The NPM is a differentiated publication indicator where points are graded for where you publish. Incentives to improve performance are clear and straightforward. Citation indicators reflect short term impact upon the scientific communication system. Citation indicators are retrospective and quite stable. It is very difficult to directly try to improve performance when it comes to impact. While one can argue that a publication-based model support the publish and perish culture with the ever increasing publication pressure, one could also argue that a citation model at the university level, due to its stability or conservative nature, and the fact that preferential attachment is at play for some universities, most likely would give cumulative advantages to those "who already have plenty", and potential changes brought about by incentives, are certainly not a short term phenomena.

There have been suggestions in Denmark to meet some of the requirements from DTU to focus more on citation impact. In order to keep the existing differentiated publication model intact, suggestions have been presented to bring in a third level especially in relation to journal outlets. This should be a category for the few hyped journals and publishing in these should be rewarded more lavishly. There may be good reasons for extending the levels in the model, but it is a flawed argument to claim to compensate wishes for more focus on impact by rewarding publication activity in "high impact" outlets. As it is well-known, article citation rates and journal citation impact have meagre correlations and the latter is a rather poor predictor of the former (Seglen, 1997).

A citation-based indicator or a hybrid indicator based on both publications and citations can be conceived in many ways, the question is whether the former or the latter is desirable. As discussed in the introduction, publication activity and citation impact are two different phenomena with substantially different prospects when it comes to incentives and behavioural adjustments. In the present analysis we could of course have experimented with more sophisticated citation-based approaches, for instance by constructing a mirror of the current publication-based model, where an arbitrary system allocates points according to which percentile group in the citation distribution they belonged to. We actually did that with a three-tiered point system, both the results were in line with the ones presented here.

As it is, based on the 2009 data, the BFI model in Denmark seems to work. Claims of more focus on citation impact seem only to speed up the cumulative advantage for "those who

already have" and at the same downgrade the influence of certain publication behaviours and muddling the transparent incentive structure.

References

Aagaard, K. (2011). Kampen om basismidlerne. Historisk institutionel analyse af basisbevillingsmodellens udvikling på universitetsområdet i danmark. PhD, Aarhus University, Aarhus.

Aagaard, K., Bloch, C., & Schneider, J. W. (2015). Impacts of performance-based research funding systems: The case of the Norwegian publication indicator. Research Evaluation, 24(2), 106-117.

Butler, L. (2002). A list of published papers is no measure of value - the present system rewards quantity, not quality - but hasty changes could be as bad. Nature, 419(6910), 877-877.

Butler, L. (2003a). Explaining Australia's increased share of ISI publications - the effects of a funding formula based on publication counts. Research Policy, 32(1), 143-155.

Butler, L. (2003b). Modifying publication practices in response to funding formulas. Research Evaluation, 12(1), 39-46.

Gläser, J., & Laudel, G. (2007). The social construction of bibliometric evaluations. In R. Whitley & J. Gläser (Eds.), The changing governance of the sciences (Vol. 26, pp. 101-123): Springer Netherlands.

Hicks, D. (2012). Performance-based university research funding systems. Research Policy, 41(2), 251-261.

Karlsson, S. & Persson, O. (2012). The swedish production of highly cited papers Vetenskabsrådets lilla rapportserie. Stockholm, SWE.

Merton, R. K. (1988). The Matthew effect in science, ii: Cumulative advantage and the symbolism of intellectual property. Isis, 79(4), 606-623.

Moed, H. F. (2005). Citation analysis in research evaluation. Dordrecht, NL: Springer.

Moed, H. F. (2008). UK research assessment exercises: Informed judgments on research quality or quantity? Scientometrics, 74(1), 153-161.

Oppenheim, C. (1996). Do citations count? Citation indexing and the research assessment exercise (rae). Serials: The Journal for the Serials Community, 9(2), 155-161.

Schneider, J. W. (2009). An outline of the bibliometric indicator used for performance-based funding of research institutions in norway. European Political Science, 8(3), 364-378.

Schneider, J. W., & Aagaard, K. (2012). "Stor ståhej for ingenting" - den danske bibliometriske indikator. In K. Aagaard & N. Mejlgaard (Eds.), Dansk forskningspolitik efter årtusindskiftet (pp. 229-260). Aarhus: Aarhus Universitetsforlag.

Seglen, P. O. (1997). Citations and journal impact factors: Questionable indicators of research quality. Allergy, 52(11), 1050-1056.

Sivertsen, G. (2010). A performance indicator based on complete data for the scientific publication output at research institutions. ISSI Newsletter (International Society for Scientometrics and Informetrics), 6(1), 22-28.

Verleysen, F. T., Ghesquière, P., & Engels, T. (2014). The objectives, design and selection process of the Flemish academic bibliographic database for the social sciences and humanities (vabb-shw). In W. Blockmans, L. Engwall & D. Weaire (Eds.), Bibliometrics: Use and abuse in the review of research performance (pp. 117-127). London: Portland Press Ltd.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E. C. M., Tijssen, R. J. W., van Eck, N. J., . . . Wouters, P. (2012). The leiden ranking 2011/2012: Data collection, indicators, and interpretation. Journal of the American Society for Information Science and Technology, 63(12), 2419-2432.