Bibliometric analyses of publications from Danish PIs linked to EU Framework Programs 6 and 7.

Jesper W. Schneider & Thomas Kjeldager Ryan
Danish Centre for Studies in Research & Research Policy,
Department of Political Science & Government, Aarhus University
Bartholins Allé 7, DK-8000 Aarhus C, Denmark
jws@ps.au.dk & *tryan@cfa.au.dk*

Introduction

The present bibliometric analyses examine the performance of journal articles affiliated to scholars at Danish research institutions and linked to projects funded by the European Framework Programs FP6 and FP7. We examine the publication outputs, subject profiles and especially the citation impact of these publications and we compare the impact to other funding benchmark units. The main findings presented below are supported by supplementary results and robustness analyses in the appendices; the report ends with a summary and a number of caveats to take into consideration when interpreting the results.

The bibliographic data used in the analyses are validated journal articles (research articles and review articles) indexed in the international citation database Web of Science (WoS). We use the in-house value-added version of WoS at CWTS, Leiden University, the Netherlands (CI-WoS). A thorough validation process has been set up and managed by the Danish Agency for Science, Technology and Innovation (DASTI) at the Ministry of Higher Education and Science, where individual researchers at Danish research institutions funded under the FP6 or FP7 programs were contacted and asked to validate pre-selected publication lists as to whether the publications were linked to their EU funding grant. Since mid-2008 potential funding acknowledgements mentioned in journal articles have been made available for analyses in the WoS. In order to try and enlarge the validated data set of publications, we utilized the WoS funding acknowledgement data and examined the pre-selected publication lists for all *non*-validated projects in order to check whether potential FP6 and FP7 grants were acknowledged. If so, these publications and projects were also included in the analyses thereby extending the data set. Eventually 175 FP6 and 503 FP7 projects and their linked publications were included into the analyses; see Appendix 1 for more details on inclusion and exclusion of projects and publications. It is important to emphasise that not all publications initially validated or identified through the WoS funding acknowledgements were eligible for analysis as only research and review articles are included, and for FP6 projects only publications from 2002 to 2013 are included, and for FP7 projects only publications from 2007 to 2013 are included.

The analyses are based on several different units of analysis. Bibliometric data are characterized by skewed distributions and robust statistics require considerable sample sizes. A common, although arbitrary, threshold is often a minimum of 50 full count publications, but larger samples are preferable. A further consideration with bibliometric data especially from citation databases are the well-known coverage problems. The enhanced citation database we use in this

analysis only index journal articles and mainly English language journals. Hence, research areas where international journals are not the primary medium for reporting research results will have lower coverage in the database and citation analyses in such areas become problematic. Main areas such as the arts and humanities, but also major parts of the social sciences and disciplines such as computer science generally have low coverages in WoS. To obtain a proxy for coverage we examine the reference behaviour in the aggregate units of analysis in the sense that we calculate the proportion of references given to other journal articles indexed in WoS. This number indicates to what extent the unit is depended on international journals in the scientific communication process and eventually the validity of doing citation analysis on such a set of articles. In Appendix 1, we plot publication numbers compared to coverage for FP6 and FP7, first for the projects and secondly for the aggregate "program themes" to which the individual projects are subsumed. It is clear that most projects have relatively few publications although the majority have moderate (> 0.5) to excellent (> 0.8) coverage. At the aggregate level of "program themes" coverage becomes more robust and we also see that several themes have very robust publication volumes. Consequently, for the present analyses we use the following units of analyses, first we examine separately all publications linked to FP6 and FP7 programs. These two publication sets are subsequently compared to overall performance for Danish publications, and publication sets linked to two other Danish funding instruments. At the disaggregate level, we also examine separately for FP6 and FP7 the performance of the individual "program themes".

From previous large-scale bibliometric analyses of two main Danish funding instruments, Centres of Excellence (CoE) funded by the Danish National Research Foundation[1] (DNRF) and various smaller grant types (compared to DNRF) funded by the Danish Council for Independent Research[2] (DFF), we have validated publication sets linked to these funding instruments for roughly the same period as the present analysis of FP6 and FP7. We utilize these publication sets as benchmarks in this analysis because they to some extent can be considered "equal" units of analysis (i.e. publications linked to funding units).

The FP6 program ran from 2002 to 2006 and the FP7 program from 2007 to 2013. In the initially validated publication lists some articles turned out to have publication dates before these programs were initiated. Consequently, we have chosen the following publication windows for the two programs: FP6, all validated articles published from 2002 to 2013, and FP7, all validated

[1] http://ufm.dk/en/publications/2013/files-2013/appendiks-5_bibliometrisk_report_03122013.pdf.
[2] http://ufm.dk/publikationer/2014/filer-2014/analyses-of-the-scholarly-and-scientific-output-from-grants-funded-by-the-danish-council-for-independent-research-from-2005-to-2008.pdf.

articles published from 2007 to 2013. Using such windows probably means that we include validated articles from the early period which is not directly linked to the EU programs, but the numbers are few. We use a citation window of three years including the publication year. This means that articles published after 2011 have shorter windows. We examine the robustness of the overall results when removing publications with shorter citation windows and the findings are robust (e.g., excluding the 2013 publications does not change the overall results). Notice, the same citation windows are applied to the different benchmark units, yet as the DNRF and DFF sets of publications only have a common coverage between 2005 and 2011, the benchmark analyses are carried out with the following publication windows: FP6 form 2005 to 2011, and for FP7 from 2007 to 2011.

Table 1 below presents the standard indicators we use in the analyses. The indicators are defined and constructed by CWTS and tailored to their CI-WoS database. These are the same indicators used in their Leiden Ranking.[3]

Table 1: Overview of standard CWTS bibliometric indicators used in the present analyses.

	Dimension	**Definition**
P	Output	Total number of publications of a unit.
P_{frac}	Output	Fractionalized publications of unit; in the present analysis we fractionalize according to country
Coverage	Validity	Internal coverage. Proxy of oeuvre being covered by Web of Science. Measured by the proportion of cited references in the oeuvre linking to other WoS publications.
MNCS	Impact	Mean normalized number of citations of the publications of a unit (self-citations not included).
$MNCS_{frac}$	Impact	Mean normalized number of citations of the publications of a unit (self-citations not included) based on fractional publication counting on the country level.
MNJS	Journal impact	Mean normalized citation score of the journals in which a research unit has published.
PPtop10%	Impact	Proportion of papers that belong to the top10% highly cited publications in the database.
$PPtop10\%_{frac}$	Impact	Proportion of papers that belong to the top10% highly cited publications in the database based on fractional publication counting on the country

[3] http://www.leidenranking.com/.

		level.
Uncitedness	Impact	Percent of papers uncited

Relative citation indicators based on full publication counts will generally have higher numerical values compared to indicators based on fractional counts. With fractional counting, publication counts sum to unity, whereas full counts necessarily implies double counting. This also means that the average database impact score of 1, often named "the world average" is only strictly viable with fractional counts. With full counts this "average" is somewhat higher, probably 0.2-0.3 points. We should also emphasize that the meaning of the numerical value of an indicator is related to the aggregation level of the unit under study. At higher aggregation levels publication volumes become larger which most often means that it becomes more difficult to have relative impact scores substantially above the database average or the expected proportion of articles among the 10% most cited in the database. This "regression-towards-the-mean" phenomenon is mainly an effect of the underlying skewed citation distributions. At the meso-level (e.g., units with 500-1000 full count publications per year), an MNCS value between 0.8 to 1.2 is generally interpreted as a performance level comparable to the average in the database (i.e., "world average" citation score), whereas values above 1.2 means that the unit's impact as a whole is above the international level, and values of 2 and more, are far above the international level of the "fields" where a unit has published in the examined period. The same yardstick can roughly be used for PPtop10% (full counts), where values above 12% would be considered above the expected performance, and values above 20% far above the expected performance for full counts.

In the next section we present the separate bibliometric analyses of all eligible publications from Danish projects linked to FP6 and FP7 programs. The following section explores the subject profiles of these publications and the subsequent section compares the FP-program performances to the benchmark units. Hereafter we examine the performance at the disaggregate level of "program themes" and we end the report we a small summary and some caveats to consider when interpreting the results.

Bibliometric analyses of total publication sets linked to FP6 and FP7 programs.

Table 2 presents the overall performance statistics for the publication sets linked to FP6 and FP7 programs. Indicators are calculated with two different publication windows, a shorter window where 2013 publications are excluded and a longer window where 2013 publications are included. Even though the publication windows are longer for the FP6 program, the analysed FP6 publication sets are considerable smaller compared to the FP7 set.

Table 2: Overall performance statistics for publication sets linked to FP6 and FP7 programs.

	FP6		FP7	
Publication windows	(2002-2012)	(2002-2013)	(2007-2012)	(2007-2013)
P	1710	2020	2731	3583
P_{frac}	927.3	1083.5	1499.6	1958.9
Coverage	82%	82%	84%	85%
Uncitedness	14.8%	16%	10.1%	13%
MNCS	1.75	1.79	2.07	2.03
$MNCS_{frac}$	1.52	1.52	1.79	1.74
PPtop10%	18.9%	19.6%	22.1%	22.2%
$PPtop10\%_{frac}$	16.8%	17.2%	20.3%	19.8%
MNJS	1.47	1.47	1.56	1.55
International collaboration	71%	71%	69%	69%

The FP6 and FP7 publication sets have coverages slightly above 80% which according to Moed (2005) can be interpreted as “excellent” for the purpose of citation analyses, although the coverage is close to the threshold between “good” and “excellent”. The difference between the two FP6 sets is 310 full count articles, whereas the difference between the FP7 sets is 852, an 18% and 32% rise respectively. The relative difference in fractional counts is roughly similar.

The MNJS indicator reflects the journal publication profile of the unit under investigation. The MNJS indicator measures the average citation impact of the journals in which a set of publications has appeared, where the citation impact has been normalized for the fields to which the journals belong. An MNJS above 1 means that the set of journals on average have been cited more frequently than would be expected based on the average journal citation activity in the respective fields to which the journals belong. The stable MNJS indicators of 1.47 for FP6 and 1.55/6 for FP7 can be considered high. In other words, the FP6 and FP7 publications are on average published in journals with relatively high impact in their respective fields. On an aggregate level, one would expect that publication in higher impact journals will result in higher overall citation impact scores

(although this reasoning does not hold for individual articles). It is interesting to notice that the proportion of non-cited publications after three years of exposure is higher for the FP6 set compared to the FP7 set. It is, however, expected that the sets including 2013 articles will have more uncited articles, this is simply an effect of the short citation window giving these articles a shorter exposure time to receive citations.

When it comes to citation impact, the FP6 and FP7 publication sets differ considerably. With full count MNCS values of 2.07 and 2.03, and PPtop10% values of 22.1% and 22.2%, the FP7 publication sets have an outstanding performance level. The performance level of the FP6 publication sets is also noticeably above the international standard, but also distinctly below the impact level of FP7 set. The variation in indicator values between excluding and including the 2013 publications can be considered diminutive and hence including the 2013 publications seem viable.

Besides the outstanding performance level for the FP7 publications, the most interesting finding from this overall performance analysis is the generally very high proportion of articles with international collaboration both in the FP6 and FP7 sets. Obviously, we would expect that a majority of the articles would be a result of international collaboration given the nature of the EU funding programs combined with the general trend of larger shares of annual publication volumes with international collaboration (e.g., for Denmark this share has been between 55 and 60% in the last decade[4]). Nevertheless, 71% for FP6 and 69% for FP7 is more than expected. Internationally co-authored articles on average have higher citation rates compared to articles with no or national collaboration, this fact no doubt influences the overall impact of the two publication sets. Below we briefly examine the performance of articles from the two publication sets with no, national or international collaboration and in the section where we compare performance to the benchmark units we further examine the influence of international collaboration on impact for the FP6 and FP7 sets.

Table 3 shows the performance statistics for the small set of articles with no, national, or international collaboration in the two publication sets. Notice no collaboration means that all authors are from the same national institution; national collaboration means that at least two national and no international institutions are affiliated with the article; and international collaboration means that at least two institutions from two different countries are affiliated with the article.

[4] http://ufm.dk/publikationer/2015/filer/dfir_scientometric_analysis_final.pdf.

Table 3: Performance statistics for articles with *no* national or international collaboration in the publication sets linked to FP6 and FP7.

	FP6		FP7	
Publication windows	2002-2012	2002-2013	2007-2012	2007-2013
P	308	348	529	689
P_{frac}	308	348	529	689
Coverage	79%	79%	83%	83%
Uncitedness	22.4%	21.2%	12.7%	15.7%
MNCS	1.48	1.49	1.72	1.64
$MNCS_{frac}$	1.48	1.49	1.72	1.64
PPtop10%	16.8%	17.6%	19.3%	19.1%
$PPtop10\%_{frac}$	16.%	17.6%	19.3%	19.1%
MNJS	1.37	1.37	1.34	1.34

If we compare the performance of the articles with no extra-institutional collaboration in Table 3 with the performance of articles with national institutional performance in Table 4, we see that for both the FP6 and FP7 sets the journal publication profile (MNJS) and the MNCS performance are considerably higher for articles with no collaboration.

Table 4: Performance statistics for articles *with* national collaboration *only* in the publication sets linked to FP6 and FP7.

	FP6		FP7	
Publication windows	2002-2012	2002-2013	2007-2012	2007-2013
P	190	213	318	420
P_{frac}	190	213	318	420
Coverage	85.8%	86.1%	87.5%	87.5%
Uncitedness	13.1%	15.6%	9.1%	12.1%
MNCS	1.32	1.32	1.44	1.47
$MNCS_{frac}$	1.32	1.32	1.44	1.47
PPtop10%	14.3%	14.3%	20.0%	18.4%
$PPtop10\%_{frac}$	14.3%	14.3%	20.0%	18.4%
MNJS	1.29	1.28	1.32	1.31

If we then compare the performance of the two previous collaboration types with international collaboration in Table 5, we clearly see that the previous two sets were relatively small in size and that the performance of internationally co-authored articles on average is markedly higher compared to articles with no or national collaboration. Nevertheless, the patterns between the three

collaboration types deviate from the overall characteristics for Danish publications in as much as articles with no collaboration has a higher impact compared to articles with national collaboration.

Table 5: Performance statistics for articles *with* international collaboration in the publication sets linked to FP6 and FP7.

	FP6		FP7	
Publication windows	2002-2012	2002-2013	2007-2012	2007-2013
P	1212	1441	1884	2474
P_{frac}	429.3	503.5	653.1	850.5
Coverage	81.6%	82.45	84.0%	84.4%
Uncitedness	13.1%	14.2%	10.2%	12.3%
MNCS	1.88	1.93	2.28	2.22
$MNCS_{frac}$	1.63	1.63	2.01	1.95
PPtop10%	20.2%	20.9%	23.3%	23.7%
$PPtop10\%_{frac}$	17.9%	18.2%	21.45	21.3%
MNJS	1.52	1.53	1.66	1.66

These results are comparable to the previous findings for the DNRF set of publications; however, in the present case the number of articles for these two types is rather low. What is noticeable is that articles with international collaboration is generally published in journals with higher international impact and have themselves on average much higher impact compared to the two other categories. But it is also remarkable that the performance for the FP7 set is higher in all three categories compared to the FP6 set and considerably higher when it comes to articles with international collaboration.

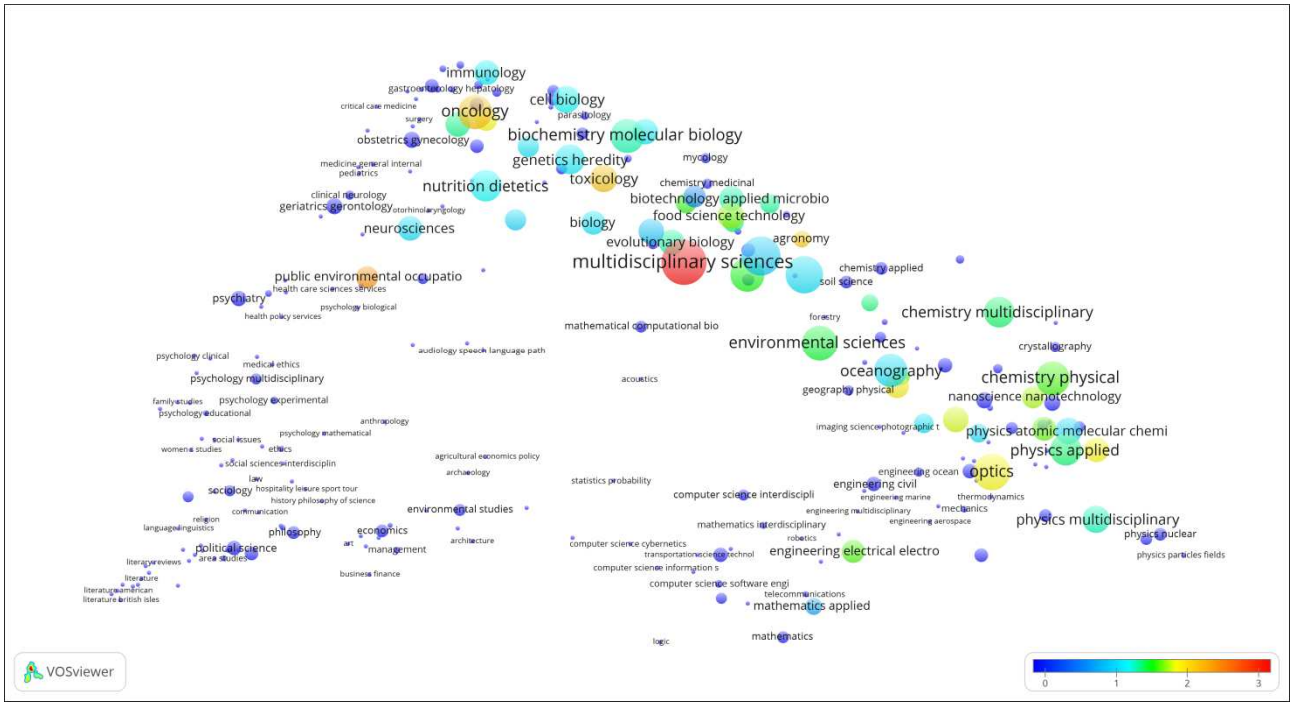
In the next section we examine the overall subject profiles for the two publication sets, first by mapping the volume and impact according to WoS journal subject categories and subsequently we outline relative performance according to OECD main subject fields.

Subject profiles of FP6 and FP7 publications.

We show the subject profiles for the FP6 and FP7 sets based on fractional counts and for the mapping of articles to WoS journal subject categories, we restrict the presentation to MNCS scores.

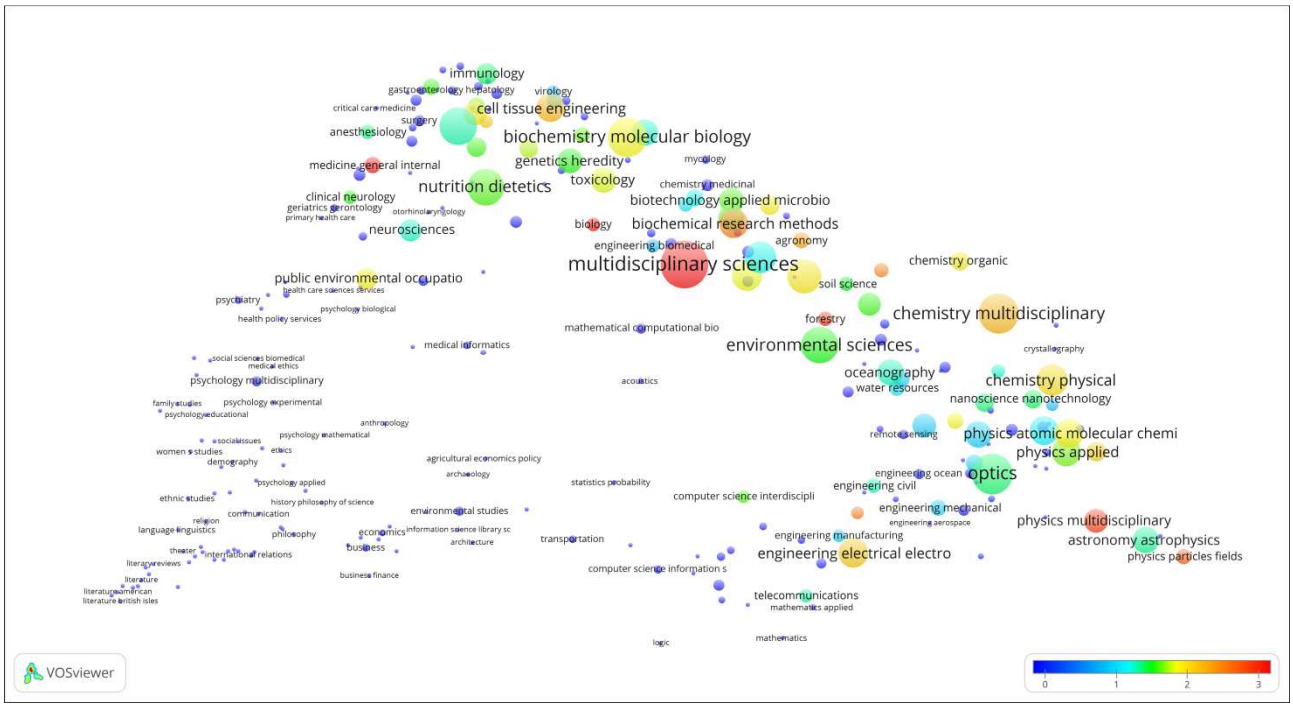
Figures 1 and 2 show the relative distribution of articles in the FP6 and FP7 sets according to the WoS journal subject categories. There are 250 categories in the base map and these categories are positioned according to their mutual cross-citation activities in the whole database (i.e., the base map is shown in Figure A5 in Appendix 2). Notice, these are journal subject categories and not subjects directly related to the topic of an individual article. Hence, we see a category like "multidisciplinary sciences" which includes multidisciplinary journals such as *Science*, *Nature* and *PNAS* (*Proceedings of the National Academy of Sciences*). Overall, the map visualizes the structure of science in as much as categories close to each other can be seen as being more related to each other compared to categories far apart.

Figure 1: Relative distribution of articles linked to FP6 projects among WoS journal subject categories. The size of the circles indicates publication volume (fractional counts) and the colour of the circles indicate mean normalized citation scores (MNCS). No citation score is shown for clusters where fractional publication counts are below 7.



We have projected the FP6 and FP7 publication sets on top of the base map and re-scaled the size of the circles according to the relative FP publication output in the respective categories.

Figure 2: Relative distribution of articles linked to FP7 projects among WoS journal subject categories. The size of the circles indicates publication volume (fractional counts) and the colour of the circles indicate mean normalized citation scores (MNCS). No citation score is shown for clusters where fractional publication counts are below 7.



Further, we have altered the colour code so that it reflects the relative citation impact (MNCS) of the category. The impact intensity is depicted by the "thermometer" in the lower right corner of the maps. Notice, we have excluded citation scores for categories where the fractional publication output is below seven.

What is most interesting in Figures 1 and 2 are the size and impact levels of the "multidisciplinary sciences" category in the two maps. In both maps, this subject category has the largest output and it is also the category with highest relative impact score. The latter is not surprising as the group includes some of the journals with the highest citation traffic in the database. But it is interesting that this category is the single largest subject category in both sets. Given the weight of the output and the impact scores this no doubt also contributes to the overall high performance for the two publication sets.

Tables 6 and 7 depict the performance for the FP6 and FP7 publication sets according to OECD main fields. Notice, the humanities field is excluded due to low publication output and poor coverage.

Table 6: Performance of the FP6 publication set according to six main OECD fields; the 250 WoS subject categories have been exclusively linked to one of the main OECD main fields, the "multidisciplinary sciences" category is subsumed under the "natural science" field. Notice, the "humanities" is excluded due to low publication volume and coverage.

FP6 (2002-2013)	Natural sciences	Engineering and technology	Medical and health sciences	Agricultural sciences	Social sciences
P	1432	310	512	189	75
P_{frac}	757.6	193.9	254.9	113.6	48.3
Coverage	83%	84%	91%	75%	43%
Uncitedness	15.7%	15.2%	8.4%	17.5%	34.7%
MNCS	1.83	1.80	1.70	1.58	1.55
$MNCS_{frac}$	1.52	1.66	1.53	1.43	1.38
PPtop10%	19.3%	21.7%	22.8%	18.8%	16.6%
$PPtop10\%_{frac}$	17.1%	21.3%	20.2%	16.0%	15.0%
MNJS	1.50	1.47	1.38	1.35	1.21
International collaboration	73.4%	63.2%	71.9%	63.5%	58.7%

For FP6, the "natural science" field is by far the largest field, but when it comes to citation impact the picture is more diverse. "Engineering and technology" have a comparable high MNCS score of 1.80 and the fractional count MNCS score is actually higher than the "natural science" field. The latter is probably to a large extent the effect of the "engineering and technology" field having a relatively lower proportion of internationally co-authored articles compared to "natural science". Also noteworthy, is the relatively large proportion of highly cited articles (PPtop10%) for the "medical and health science" field and especially the discrepancy to the MNCS scores for that field.

If we examine the FP7 set in Table 7, we can observe that the relative distribution among fields between the two sets is similar with the "natural science" field constituting around 55% of the publications.

Table 7: Performance of the FP7 publication set according to six main OECD fields; the 250 WoS subject categories have been exclusively linked to one of the main OECD main fields, the "multidisciplinary sciences" category is subsumed under the "natural science" field. Notice, the "humanities" is excluded due to low publication volume and coverage.

FP7 (2007-2013)	Natural sciences	Engineering and technology	Medical and health sciences	Agricultural sciences	Social sciences
P	2454	725	922	302	79
P_{frac}	1286.0	459.0	486.8	184.6	42.9
Coverage	85.1%	77.9%	89.4%	76.6%	61.3%
Uncitedness	11.9%	19.0%	9.4%	15.6%	16.5%
MNCS	2.17	1.47	1.79	1.64	1.69

$MNCS_{frac}$	1.82	1.35	1.63	1.58	1.35
PPtop10%	23.3%	16.3%	20.9%	19.3%	18.7%
$PPtop10\%_{frac}$	20.6%	14.7%	20.4%	17.1%	15.1%
MNJS	1.64	1.38	1.41	1.37	1.26
International collaboration	72.2%	60.1%	69.6%	61.6%	69.6%

Nevertheless, when it comes to citation impact, the pattern in the FP7 set is substantially different. The performance of the “engineering and technology” field is now below the other fields, but most remarkable is the extraordinary high impact of the “natural science” field.

Comparison of restricted publication sets linked to FP6 and FP7 programs to benchmark units.

In scientometric studies, if possible, it is desirable to compare like with like such as a research institution with other research institutions or countries with countries. It is also preferable to compare units of roughly similar size as it is generally so that with larger units indicator values will tend to move closer towards the reference value as mentioned in the introduction. The units of analysis in this report are European funding programs and we include publications in the analyses if they are linked to a project funded by one of these programs. Obviously, publications as discrete units primarily "belong" to authors and institutions, where funders, and there are often several of them, are given an acknowledgement, but otherwise not credited. Nevertheless, we use the funding institution as the unit of analysis and link publications to it. An ideal benchmark unit would obviously be a very similar funding institution. From previous bibliometric analyses of two main Danish funding institutions, Centres of Excellence (CoE) funded by the Danish National Research Foundation (DNRF) and various smaller grant types (compared to DNRF) funded by the Danish Council for Independent Research (DFF), we have validated publication sets linked to these instruments for roughly the same period as the present analysis of FP6 and FP7. We utilize these publication sets as benchmarks in this analysis because they to some extent can be considered "similar" units of analysis (i.e. publications linked to funding units). Such a comparison is however not without problems. The different funding units clearly have different aims and purposes, and are different when it comes to the size of grants. Further, publications may well be linked to several funding institutions and grants making it very difficult to claim any direct link between funding and performance.

The present benchmark analysis is comparable to the ones presented in the previous two analyses of the DNRF and DFF. We have unique validated publication sets linked to DNRF and DFF grants and with the present analysis we also have validated publications linked to FP6 and FP7 projects. In order to make the publication sets comparable we restrict them to the following two time periods: 2005 to 2011 for benchmarking the FP6 set, and 2007 to 2011 for benchmarking the FP7 set. Like the previous analysis we also compare the present units of analysis to the overall Danish performance in the respective time periods. Notice, there is considerable size differences between the units examined in the benchmark comparisons, this should be taken into consideration when interpreting the results.

We present the main performance statistics in Tables 8 (FP6) and 9 (FP9) and the main trends are illustrated in Figures A6 (FP6) and A7 (FP7) in Appendix 2. The performance statistics in the tables come in two parts. First we document the overall performance for all Danish publications (i.e. articles with at least one Danish address) in the period examined and then analyse what happens to the overall impact when we stepwise exclude the various publication sets linked to the funding institutions. Next we compare the overall performance of the three publication sets linked to the funding units (i.e., FP, DFF and DNRF, marked with grey in the tables).

Table 8: Comparison of performance between restricted FP6 publication sets (2005-2011) and benchmark units, the DNRF and DFF funding sets, and the overall Danish set of publications.

Restricted publication sets (2005-2011)	P	MNCS	PPtop10%	Uncited	MNJS	P_{frac}	$MNCS_{frac}$	$PPtop10\%_{frac}$	Coverage
Denmark (DK)	78173	1.46	15.5%	19.9%	1.24	51538.9	1.28	13.5%	81%
DK excl FP6	76930	1.45	15.5%	20.0%	1.24	50860.5	1.28	13.4%	81%
DK excl FP6 and FP7	75151	1.44	15.3%	20.2%	1.23	49883.3	1.27	13.2%	81%
DK excl FP 6 and DFF	71113	1.43	15.2%	20.5%	1.22	46990.7	1.25	13.1%	81%
DK excl FP6 and DNRF	70559	1.41	14.9%	21.8%	1.21	46960.5	1.24	12.9%	80%
DK excl FP6, DFF and DNRF	65684	1.40	14.7%	21.3%	1.19	43682.5	1.22	12.7%	80%
Total FP6 set of pubs*	1267	1.82	20.0%	14.6%	1.46	695.0	1.57	18.1%	82.0%
Total DFF set of pubs*	6272	1.81	19.3%	13.4%	1.49	4182.6	1.62	17.9%	87%
Total DNRF set of pubs*	7164	1.88	21.7%	12.0%	1.57	4458.0	1.72	19.6%	88%

*Notice numbers deviate from the Danish set due to non-Danish FP6, DFF and DNRF publications.

It is clear from Tables 8 and 9 that removing the different smaller publication sets linked to the funding units results in a continuous decrease of overall Danish impact. The effect of removing the FP6 set is smaller compared to removing the FP7 set (see Table 9). This is a consequence of the smaller volume of the FP6 publication set but also the lower impact levels compared to the FP7 set. The general drop is more marked when removing the FP7 set. In the previous analyses of DNRF and DFF we discussed how to interpret the seemingly small changes in impact. Significance tests are irrelevant here (cf. Schneider, 2013; 2015), yet resampling techniques where random sets of articles of similar size as the funding units are removed from the overall Danish sets reveal that the changes caused by the funding sets are indeed substantial. Nothing happens to Danish impact when we resample, but removing publications linked to the specific funding units decreases overall

Danish impact. This is so because the funding sets have a substantially higher proportion of highly cited articles[5].

When we compare the overall performance for the FP6 publication set to the benchmark units, we see that the impact levels are obviously substantially higher than the overall Danish impact, comparable to the DFF set, but below the DNRF set. In all instances, the FP6 set is considerably smaller compared to the benchmarks. In the FP6 set of publications, 129 unique articles are also linked to the DFF set, 127 to the DNRF set, and 33 are both linked to the DFF and DNRF sets. Approximately 10% of the FP6 linked publications are also linked to either a DFF grant or a CoE funded by the DNRF.

As stated above, the trends are more marked for FP7 (see Table 9). But perhaps most interestingly are the differences in impact between FP7 and the benchmark units. The difference between the DFF and DNRF was documented in a previous analysis. The publication window is slightly different in the present analysis but the impact scores are similar. Notice both the DFF and DNRF sets have high performance levels and especially the DNRF set is characterized by a very high performance when it comes to the proportion of highly cited articles. Remarkably, the performance of the FP7 set is above that of the DNRF and can be considered outstanding

Table 9: Comparison of performance between restricted FP6 publication sets (2005-2011) and benchmark units, the DNRF and DFF funding sets, and the overall Danish set of publications.

Restricted publication sets (2007-2011)	P	MNCS	PPtop10%	Uncited	MNJS	P_{frac}	$MNCS_{frac}$	$PPtop10\%_{frac}$	Coverage
Denmark (DK)	59130	1.48	15.8%	18.9%	1.26	38490.8	1.29	13.6%	81%
DK excl FP7	57355	1.46	15.6%	19.2%	1.25	37515.7	1.28	13.4%	81%
DK excl FP6 and FP7	56243	1.46	15.5%	19.3%	1.25	36908.8	1.27	13.3%	81%
DK excl FP 7 and DFF	51978	1.44	15.2%	19.8%	1.23	33936.6	1.25	13.0%	80%
DK excl FP7 and DNRF	52591	1.43	15.0%	20.0%	1.23	34647.4	1.24	12.9%	80%
DK excl FP7, DFF and DNRF	48080	1.41	14.8%	20.5%	1.21	31616.4	1.22	12.6%	80%
Total FP7 set of pubs*	1908	2.11	23.0%	10.5%	1.57	1068.1	1.81	21.7%	84%
Total DFF set of pubs*	5841	1.82	19.3%	13.4%	1.50	3895.0	1.63	18.0%	87%
Total DNRF	5638	1.89	22.2%	11.0%	1.58	3421.0	1.72	19.9%	88%

[5] http://ufm.dk/publikationer/2014/filer-2014/analyses-of-the-scholarly-and-scientific-output-from-grants-funded-by-the-danish-council-for-independent-research-from-2005-to-2008.pdf.

set of pubs*									

*Notice numbers deviate from the Danish set due duplets as well as to non-Danish FP7, DFF and DNRF publications.

The PPtop10% is also markedly higher. The marked differences are also visible with fractional counts, but here the degree of internationalization must also be taken into consideration as it influences the scores, not only the fractioning of scores, but also in relation to the fact that internationally co-authored articles on average have higher citation density rates (we explore this below). Noticeably are also the similar journal publication profiles for the FP7 and DNRF sets. The MNJS score confirms the findings presented in the previous section on subject profiles, in as much as the average publication behaviour is directed towards journals with the highest impact in their fields, largest among them is the "multidisciplinary sciences" as documented in Figure 2 above. The volume of the FP7 set is larger than the FP6 set, but at the same time also considerably lower compared to the benchmarks. In the FP7 set, 210 unique articles are also linked to the DFF set, 224 to the DNRF set, and 42 are both linked to the DFF and DNRF sets. Approximately 11% of the FP7 linked publications are also linked to either a DFF grant or a CoE funded by the DNRF.

As already documented in Tables 2 to 5, seven out of ten articles in the FP6 and FP7 publication sets are a result of international collaboration. While we would expect a number somewhere above the 55-60%, which is usually ascribed to Danish publication sets, close to 70% were surprisingly high.

Table 10: Comparison of performance of international co-authored articles between the restricted FP6 and FP7 publication sets and the benchmark units (DNRF, DFF and Denmark).

	(2005-2011)				(2007-2011)			
	FP6	DK	DFF	DNRF	FP7	DK	DFF	DNRF
P_{total}	1267	78173	6272	7164	1908	59130	5841	5638
$P_{int\ collab}$	898	43937	3541	4400	1292	33881	3300	3561
Share of international collaboration	70.9%	56.2%	56.5%	61.4%	67.7%	57.3%	56.5%	63.2%
MNCS	1.98	1.67	2.07	2.05	2.32	1.70	2.08	2.07
PPtop10%	21.0%	18.2%	20.9%	24.2%	23.8%	18.4%	20.9%	24.5%
MNJS	1.49	1.35	1.62	1.69	1.67	1.38	1.63	1.69
Uncitedness	13.0%	16.5%	11.9%	9.8%	10.1%	15.5%	12.0%	9.2%
P_{frac}	326.0	17311.5	1451.6	1694.5	452.1	13249.0	1354.0	1344.5
$MNCS_{frac}$	1.71	1.48	1.88	1.92	2.01	1.49	1.89	1.93
$PPtop10\%_{frac}$	18.7%	16.2%	19.5%	22.5%	21.9%	16.2%	19.5%	22.8%
Coverage	82%	82%	87%	87%	83%	83%	87%	87%

In Table 10 we compare the degree of internationalization between the FP6 and FP7 publication sets and the benchmark units. We also outline the performance for the internationally co-authored articles in these sets. The FP6 set has a markedly higher share of internationally co-authored articles and the DNRF set has the second highest share albeit more than nine percentage points less than the FP6 set. Interestingly, even though the DFF and DNRF sets have considerably lower shares of articles with international collaboration, their impact levels for this group of articles is markedly higher than the FP6 set. On the other hand, the FP7 set has a somewhat lower proportion of internationally co-authored articles compared to the FP6 set, but still a larger proportion compared to the benchmark units; yet the impact for this set is remarkable!

In order to further explore the relationship between the publications' funding links and their citation impact we model this relation controlling for the fact that the proportion of international collaboration varies depending on the funding link. We investigate whether the difference in impact between DFF, DNRF and FP-funded research is mainly driven by a difference in internationalization of research. Two sets of regressions are estimated, the first including publications funded by FP6, DFF, DNRF and Danish publications from 2005-2011 (excluding FP7 funded papers). The second, includes FP7, DFF, DNRF and Danish publications from 2007-2011 (excluding FP6 funded publications). In each of the two sets, three regressions are estimated:

- In the first regression indicators of funding links are regressed on the log of normalized citation scores (NCS). In addition, control variables for well-known correlates of citation scores are included: number of authors, number of references and normalized journal score (NJS).
- The second regression is the same as (1), however, it also includes a dummy variable indicating that a variable is authored by at least two persons in two different countries.
- The third regression is the same as (2) however a variable indicating the number of countries affiliated with the publication is also added.

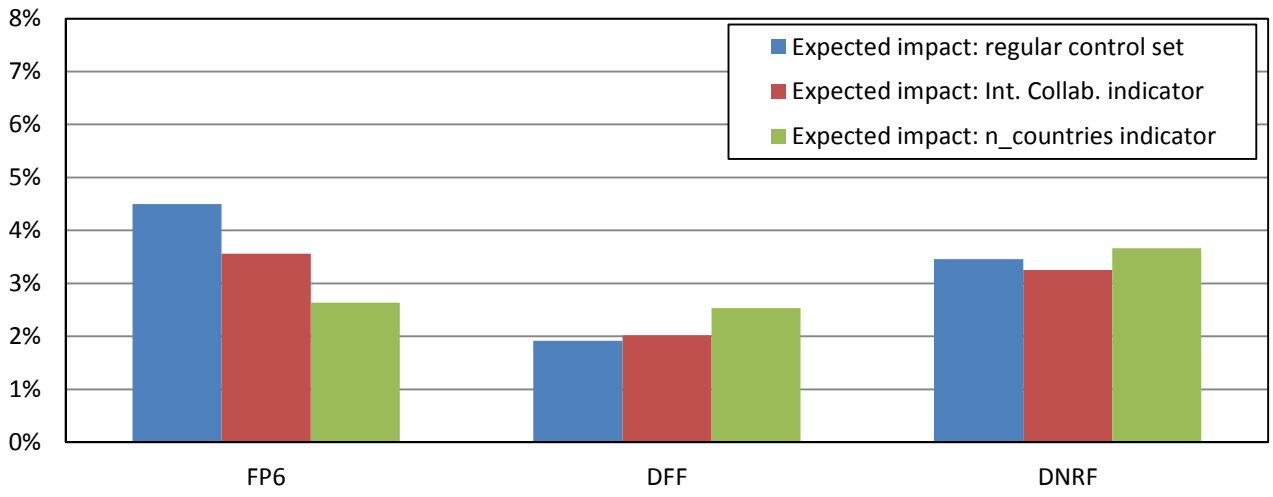
The main interest in each regression is on the coefficient of the dummy-variables which indicate the funding links. The expectation is that the coefficient indicating an FP6 or FP7 grant will decrease when control variables designating international collaboration are included. Since international collaborations, on average, are associated with a higher impact (no matter the funding link) and FP funded research has a relatively high proportion of international collaborations, the coefficient on FP6 and FP7 is expected to be lower when a variable indicating international collaboration is included. The extent to which the coefficient drops is the interesting part of this analysis. The

question is, when we take into account that a large proportion of FP funded publications are international collaborations what is the expected marginal relationship between FP funded publications and their impact?

The technical details including specifications and detailed results are outlined in the Appendix 3; here we present the main findings. Figures 3 and 4 illustrate the estimated marginal effect of funding links on citation impact with and without statistically controlling for international collaboration. The blue bar shows the expected marginal impact (relative to the baseline of Danish publications) of being funded by FP, DFF or DNRF when we control for number of authors, normalized journal impact and number of references. The red bar shows the expected marginal impact of being funded by FP, DFF or DNRF when we include a variable indicating whether the publication is a result of international collaboration or not. The green bar shows the expected marginal impact of being funded by FP, DFF or DNRF when we include the number of countries that are represented in the address field of a publication.

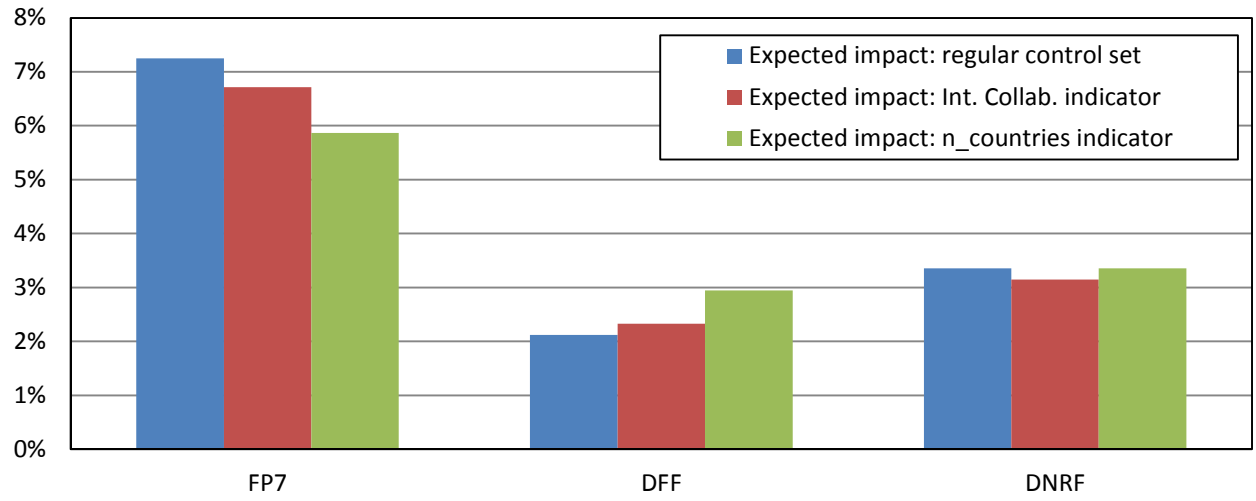
We generally find that the expected impact of publications funded by the FP-programs is higher than publications funded by DFF, DNRF and the set of Danish publications. However, when controlling for whether the publication is an international collaboration, this difference diminishes. When including a variable indicating the number of countries involved in the publication the expected citation impact for FP6 publications is lower than DNRF publications and equal to that funded by DFF.

Figure 3: Results of partial regression effects of FP6, DFF & DNRF. Expected impact of funding links in percentage relative to a baseline of Danish publications from 2005 to 2011.



For FP7 the effect of internationalization is high, however, after controlling for internationalization the expected impact of an FP7 article is still significantly higher than DNRF and DFF funded publications. This underlines the earlier finding, that for FP7 the average impact of international collaborations is extraordinarily high and well above the expected impact of international collaborations.

Figure 4: Results of partial regression effects of FP7, DFF & DNRF. Expected impact of funding links in percentage relative to a baseline of Danish publications from 2007 to 2011.



The dummy variable indicating a DNRF grant does not change in either specification, and is not affected by the inclusion of internationalization indicators. This implies that research funded by DNRF generally has high impact which is more or less independent of internationalization. DFF publications, however, increase in terms of expected impact when controlling for the number of countries. This may indicate that the DFF grants are oriented more towards nationally-oriented research and that this research has an above average impact compared to other national publications.

The FP7 set includes ERC and Marie Curie grants. These grants are different from the more strategic or topic specific FP7 program themes (see next section). Together, the publications linked to ERC or Marie Curie grants constitute 27% of all FP7 linked publications, and 33% of the restricted FP7 set used for the benchmark analyses. Table 11 shows what happens to the impact scores when we remove these grants from the restricted FP7 set.

Table 11: Consequences of removing ERC and Marie Curie grants from the restricted FP7 set of publications used for the benchmark analyses.

	(2007-2011)			
	FP7	FP7 without ERC & Marie Curie	DFF	DNRF
Share of international collaboration	67.7%	68.0%	56.5%	63.2%
MNCS	2.11	1.91	1.82	1.89
PPtop10%	23.0%	21.1%	19.3%	22.2%
MNJS	1.57	1.49	1.50	1.58
$MNCS_{frac}$	1.81	1.61	1.63	1.72
$PPtop10\%_{frac}$	21.7%	19.7%	18.0%	19.9%

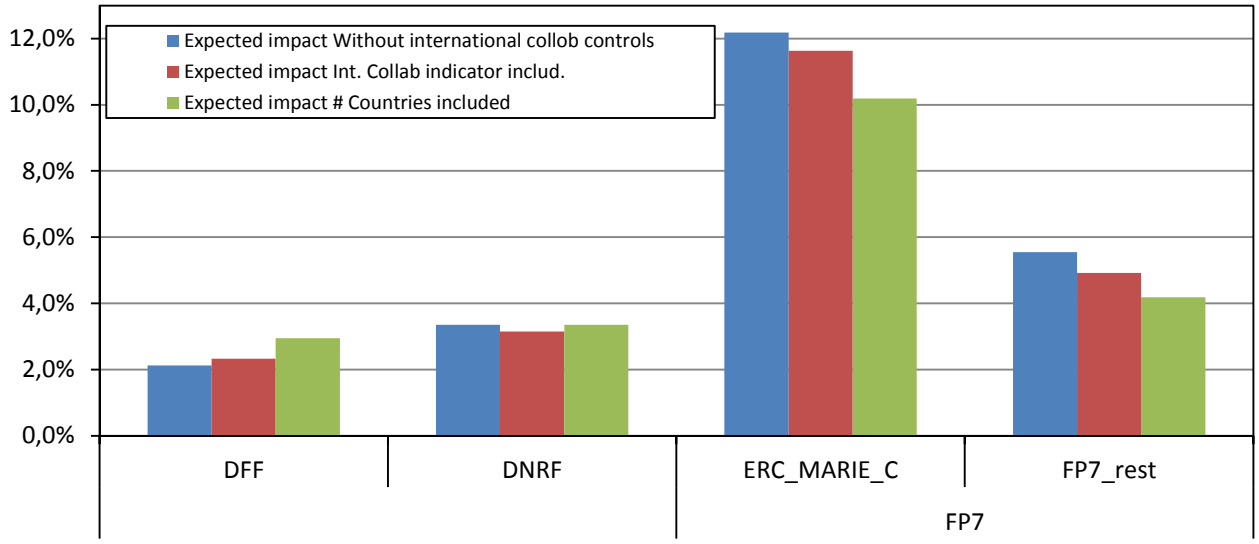
It is clear that the degree of internationalisation is not affected, but both full and fractional count MNCS and PPtop10% scores drop markedly, and so does the MNJS indicator. The latter suggest that the publication profile for these specific grants is in journals with very high international visibility, perhaps the most important factor influencing citation impact on the aggregate level of publication sets. Depending on whether we focus on indicators based on full or fractional counts, the performance level for the remaining FP7 publications is on level with the DNRF with full counts for the MNCS, but slightly below with PPtop10%. With fractional counts the MNCS score is considerably below the DNRF level but similar to DFF. Noticeably here is that the more robust indicator for the proportion of highly cited articles (PPtop10%frac) suggests that the drop is most marked in the average-based indicators (MNCS) as they are less robust in relation to outliers' influence on indicator values. Consequently, the average-based indicators in the FP7 set is more "vulnerable" because the subset of ERC and Marie Curie linked publications include some very highly cited outliers. Notice, there is no overrepresentation of ERC or Marie Curie linked publication that also have links to either the DFF or DNRF sets. For ERC there is a 9% overlap with DFF and 11% with DNRF; for Marie Curie, there is again a 9% overlap with DFF but only 8% with DNRF.

A further regression seems to confirm this general finding, that ERC and Marie Curie linked publications to a large extent can explain the remaining gap between FP7 and the DNRF when we have controlled for international collaboration.

We use the same specification as above however, we split the FP7 indicator into two; one indicating ERC or Marie Curie funding and the other indicating other FP7 funding. The technical details including specifications and detailed results are outlined in the Appendix 3. Figure 5

illustrates the estimated marginal effect of funding links on citation impact with and without statistically controlling for international collaboration. The figure is directly comparable to Figure 4, however; the marginal effect for the two funding groups ERC and Marie Curie and the other themes under FP7 are illustrated on the right instead of FP7.

Figure 5: Results of partial regression effects of FP7, ERC & MARIE_C, FP7_rest, DFF & DNRF. Expected impact of funding links in percentage relative to a baseline of Danish publications from 2007 to 2011.

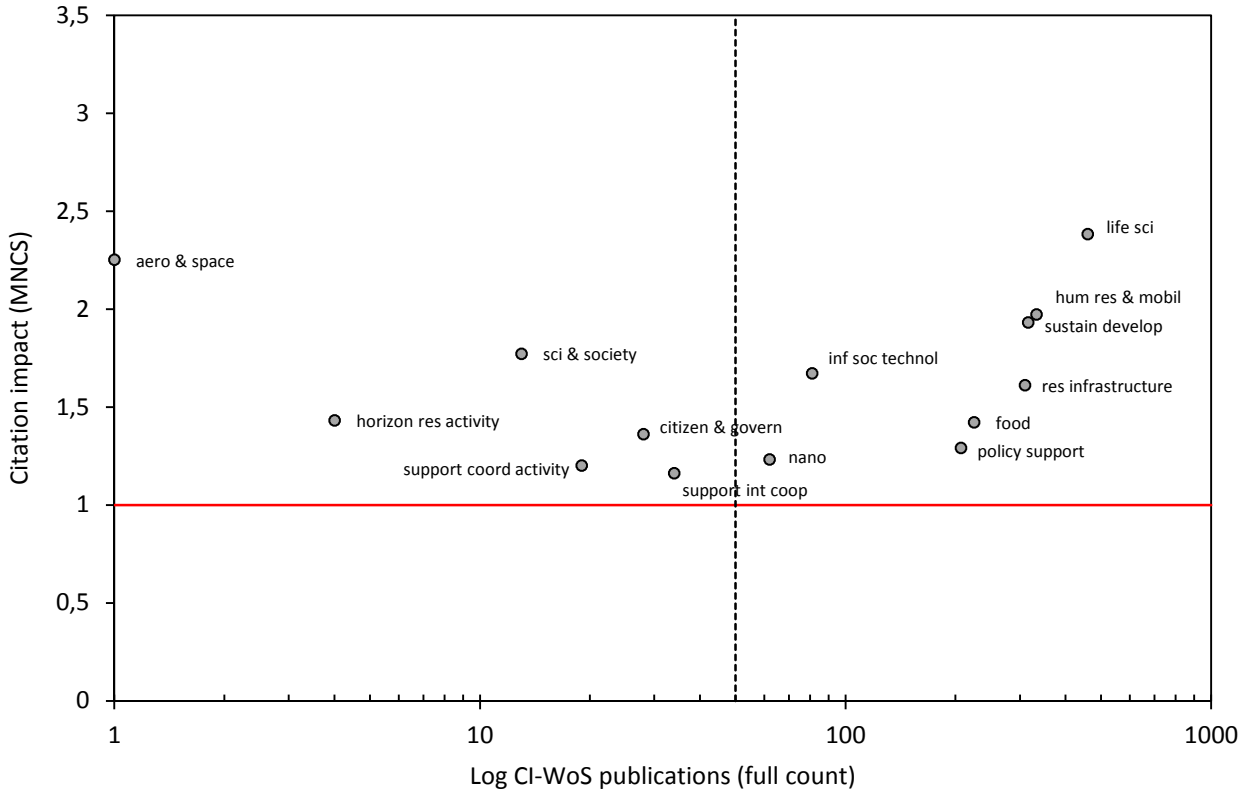


We find that the expected impact of ERC and Marie Curie is larger than other FP7 themes combined that in turn have an expected impact that is slightly higher than DNRF and DFF and that all funding sets have a higher expected impact compared to the baseline of Danish publications. The results indicate that is it is meaningful not only to distinguish between the funding sets but also within the funding sets as their aim and functions vary considerably. Even though ERC and Marie Curie drives impact to a considerable degree the expected impact of the other FP 7 themes are still above that of DFF and DNRF. The marginal impact of FP7 themes can therefore be divided into a group with a very high marginal effect (ERC and Marie Curie) and a group with a moderate marginal effect (FP7_rest). The results point to the possibility that there may be some specific characteristics of ERC and Marie Curie Actions that influence impact of publications which are not included in the present model.

Bibliometric analyses of publication sets linked to program themes under FP6 and FP7 programs.

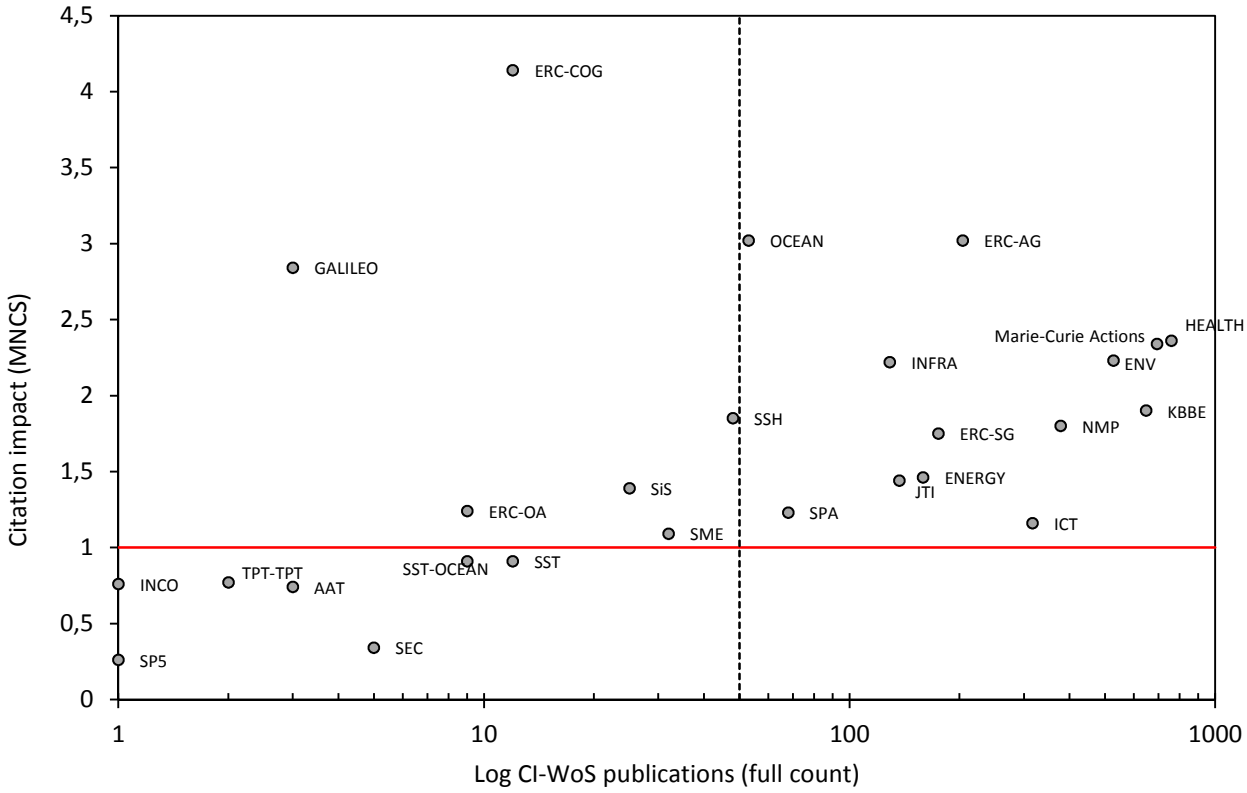
As a final performance analysis we disaggregate the FP6 and FP7 publication sets to the level of "program themes". In the Appendix we demonstrate the validity when it comes to coverage and publication volume of using this disaggregate unit of analysis compared to usage of individual projects. Below in Figures 6 and 7 we present results for full count MNCS scores plotted as a function of output for the individual "program themes". By plotting impact to output it becomes easier to interpret the importance and robustness of the individual indicators. We have plotted a grid line corresponding to 50 full count publications on the log-scaled x-axis (output); this rather arbitrary threshold can be used as a guideline when interpreting the results. Results on or just below the threshold should be treated carefully and results far below should be discarded.

Figure 6: Mean normalized citation scores (MNCS) as a function of publication output (full counts) for FP6 program themes (FP6 project publications are distributed according the themes their parent projects belong to).



In the Appendix we present two tables that explain the abbreviations used in the plots. Also in the Appendix are two figures illustrating the PPtop10% indicator as a function of output as well as two overall tables presenting the main performance statistics for the FP6 and FP7 "program themes".

Figure 7: Mean normalized citation scores (MNCS) as a function of publication output for FP7 program themes (FP7 project publications are distributed according the themes their parent projects belong to).



From Figure 6 we can see that eight "program themes" have publication outputs above 50 and all eight also have impact scores on or above 1.20. One theme, "life sci" has an impressive impact score of 2.30 and at the same time this "theme" is the largest among the 14 examined in this analysis when it comes to publication output.

In Figure 7, 14 "program themes" have outputs from approximately 50 up to 760 full count publications. Thirteen of these themes have indicator values above 1.20 and 6 with MNCS scores above 2. The two highest performing themes among those with robust publication outputs are OCEAN and ERC-AG with impressive MNCS scores of 3.02. It is noticeable that there seemingly is a broad variation among the 14 themes when it comes to "types" among the 14 most robust themes. There is a mixture of ERC grants, Marie Curie-grants, infrastructure and topical themes, all with high performance – consequently no single type seems to stand out.

Caveats for the interpretation of the data

In the interpretation of the data presented in this report it should be kept in mind that measuring the properties of science is a difficult exercise. Bibliometric data can contribute to this exercise with important insights, but cannot stand alone. Indicators measuring citation impact capture the short term reception of journal articles in the scholarly communication system. This is termed academic impact, but it is important to realize that there is not a one-to-one relation between impact and research quality. Under reasonable circumstances, impact may be seen as a partial or indirect measure of quality. As a consequence of this partial and one-dimensional nature of the indicators, a single indicator is often not reliable. However, when various complementary indicators suggest similar insights more convincing evidence about the property observed is offered (Martin, Nigthingale & Rafols, 2014). Furthermore, the indicators have to be appropriate to the property under investigation and be applicable to the whole system under study. The limitations with regard to this are well-known within the humanities and major parts of the social sciences, but also apply to certain areas of the hard sciences. Finally, bibliometric indicators are unreliable below certain levels of aggregation and need careful mathematical normalization to be used across diverse research areas. However, these normalization procedures are by no means perfect. As a consequence comparisons across fields should be treated with caution. The interpretation of the data in this report should in other words be done with care, but despite these limitations, bibliometric data do have a lot to offer when examining academic performance.

One specific issue needs to be emphasised in relation to the presented analysis: The potential systematic selection bias in the examined data. To begin with some 345 researchers with projects being funded by the FP6 program and 1026 researchers with FP7 projects, all affiliated to Danish research institutions, were contacted in order to validate publication lists as to whether the publications were a result of the EU funding. In the end 175 (FP6) and 503 (FP7) projects ended up in the analyses and 171 (FP6) of 461 (FP7) of these had eligible journal publications. In that process, and in order to enlarge the study, we utilized the WoS funding acknowledgement information indexed since mid-2008 to include projects and publications not manually validated. These data collection processes do neither provide apparent populations nor constitute random samples. Although we seem to have included approximately half of the originally appointed FP6 and FP7 projects we cannot rule out systematic bias. We cannot expect the missing projects and their affiliated publications to be an exact mirror of those included. Further among the included

projects we must reckon with some false positive as well as missing publications. In Appendix 1 we have tried to guess to what extent the included data may be biased. Table A1 shows the number projects included and excluded and Table A2 disaggregates this to "program themes". It is interesting that there seems to be a fairly even distribution among included and excluded projects both at the program level but also the theme level. When the researchers were contacted a pre-selected publication list was presented to them for validation. In Tables A3 and A4 we have calculated overall indicators for the remaining non-validated publications, knowing that many of them are not linked to FP6 and FP7 funding. The results show generally lower impact scores compared to the validated publication sets included in the analysis. This is by no means surprising given the fact that the FP-funded publications on average have considerably higher impact scores than non-FP-funded publications and we can expect numerous of the latter to be included in the non-validated sets. The unanswered question is therefore: would the citation distributions significantly change if the FP6 and FP7 publication sets were substantially enlarged? Since the sets are already fairly robust given their numbers and the experience we have with the larger DFF and DNRF sets, were are inclined to say that an enlargement will probably not change the distributions and thus impact levels in any substantial way.

Finally we have also estimated the difference between the impact of the validated publications and those extra publications included based on the WoS funding acknowledgements. These findings are reported in Table A5 in Appendix 1. Here we see that the FP7-validated set has slightly larger impact compared to the extra included publications but it is the other way around for the FP6 set. Consequently, the results presented in this report should be interpreted carefully as systematic bias cannot be excluded, however, we have good indications that the results are indeed robust and to a large extent reliable.

Summary of main findings

In this section we briefly summaries the main findings of the bibliometric analyses. Overall the publication sets linked to FP6 and FP7 examined in this study perform above and far above the international performance levels respectively when it comes to citation impact. The impact level is generally high, although the impact level for the FP7 set can be considered outstanding. Noticeably for both sets are the degree of internationalization, although anticipated, the actual proportions of articles with international collaboration were higher than expected. Also characteristically for both sets is the publication profile when it comes to output in the "multidisciplinary sciences" subject category. For both groups, this is the single largest subject category when it comes to output and it is also the category with highest average citation impact. On the other hand, the two sets vary to some degree when their subject profiles are characterized according to OECDs main research fields. Not surprisingly, the natural science field is the largest in both sets, but in the case of FP6 the field of engineering and technology perform at the same level as the natural science field, where as in the case of FP7 the natural science field markedly outperforms all other fields.

A main finding of the present analyses is the generally outstanding performance level of the FP7 linked publications. As a set it has higher performance levels compared to all the benchmark units including the DNRF. As in previous analyses of funding units, we also see that in the present case removing either the FP6 or the FP7 set causes a decrease in overall Danish impact. The results are robust yet the decrease is most marked for the FP7 set and confirms that the sets contain a relatively larger share of articles with higher citation rates compared to the overall distribution of Danish articles. Interestingly, the two sets differ when it comes to the actual impact of the articles with international collaboration. As expected, impact is generally high and considerably higher than the impact of publications with no or national collaboration. Nevertheless, the impact level for the FP6 set is below the levels for the two funding units used as benchmarks, whereas the FP7 set outperforms them all. Statistical modelling suggests that substantial parts of the impact received by FP6 and FP7 linked publications are associated with the high level of international collaboration. But again there are differences between the two sets. The expected marginal impact for FP6 linked publications are on the same level as the DFF set but below the DNRF set when controlling for international collaboration. Without the statistical control, the expected impact level for FP6 linked publications are on level with or slightly below the benchmark sets. Consequently, international collaboration with its derived impact effects, to a large extent seems to statistically explain the performance level of the FP6 set.

International collaboration is, however, not the only explanation for the outstanding performance levels for the FP7 set. Controlling for international collaboration reduces the expected marginal impact, nevertheless even after the statistical control the expected impact of FP7 linked publications is markedly higher than the two benchmark sets. Consequently, the high impact of FP7 publications cannot be explained as primarily an effect of international collaboration, other factors are at play. Publications linked to ERC and Marie Curie grants are included in the FP7 set and constitute 27% of the total FP7 set and 33% of the restricted set used for the benchmark analyses. Removing the ERC and Marie Curie linked publications causes a considerable drop in impact for the remaining FP7 linked publications. Depending on whether we focus on indicators based on full or fractional counts, the performance level for the remaining FP7 publications is on level with the DNRF with full counts, or below the DNRF but on par with the DFF set using fractional counts. Noticeably here is that the more robust indicator for the proportion of highly cited articles (PPtop10%frac) suggests that the drop is most marked in the average-based indicators (MNCS) as they are less robust in relation to outliers' influence on indicator values. Subsequent modelling confirms these findings in as much as controlling for ERC and Marie Curie grants seems to explain most of the gap to the DNRF set, although the expected marginal impact is still slightly higher for the FP7 set even after controlling for these specific grants.

Finally, the disaggregate analyses at the level of "program themes" reveals that no single type of theme seems to dominate performance, high impact levels are spread among various different funding themes and types. It is interesting to observe that other "themes" than ERC and Marie Curie grants both have large volume and high impact in the FP7 set.

Appendix 1: Coverage and robustness

Figure A1. Coverage and validity: Publication output of included FP6 projects compared to their overall internal coverage.

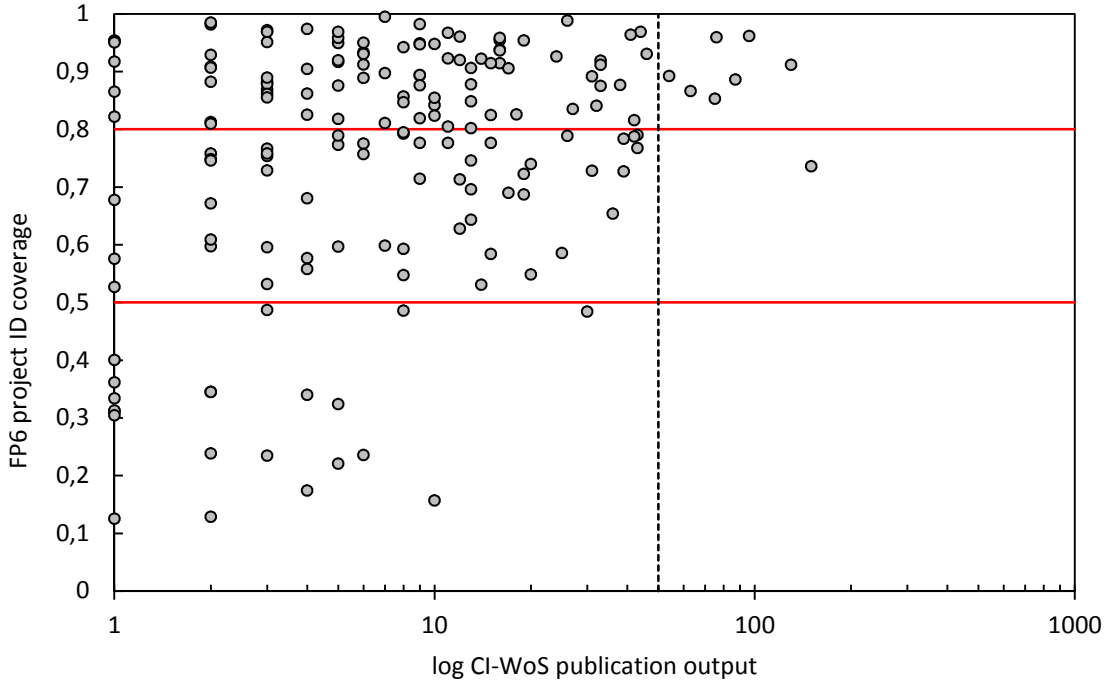


Figure A2. Coverage and validity: Publication output of included FP7 projects compared to their overall internal coverage.

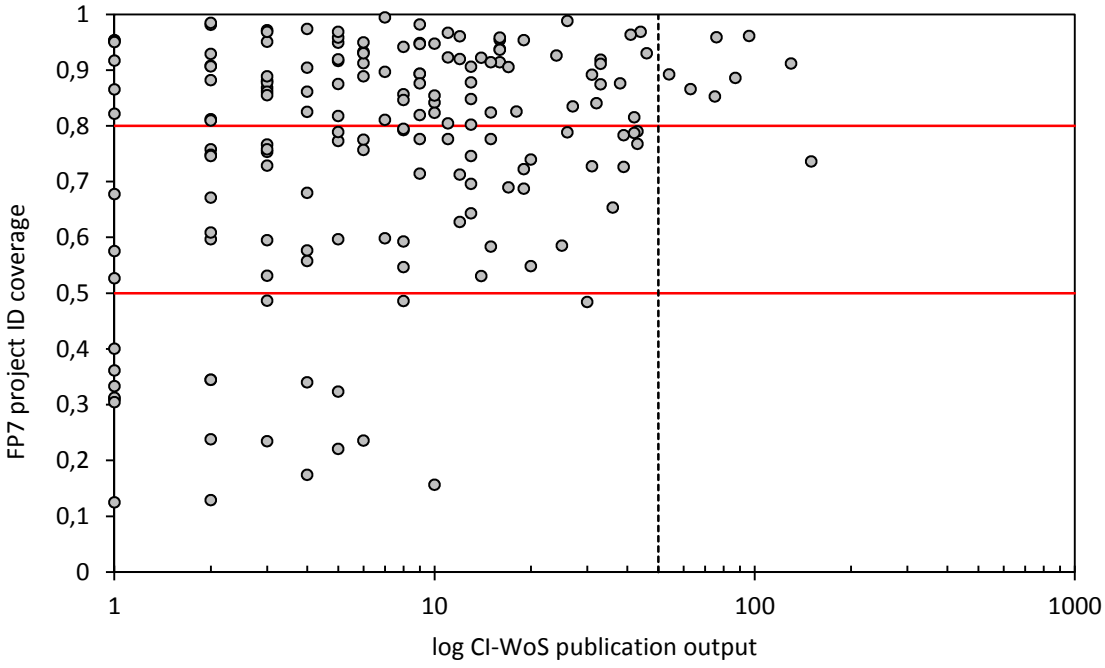


Figure A3. Coverage and validity: Publication output of included FP6 themes compared to their overall internal coverage.

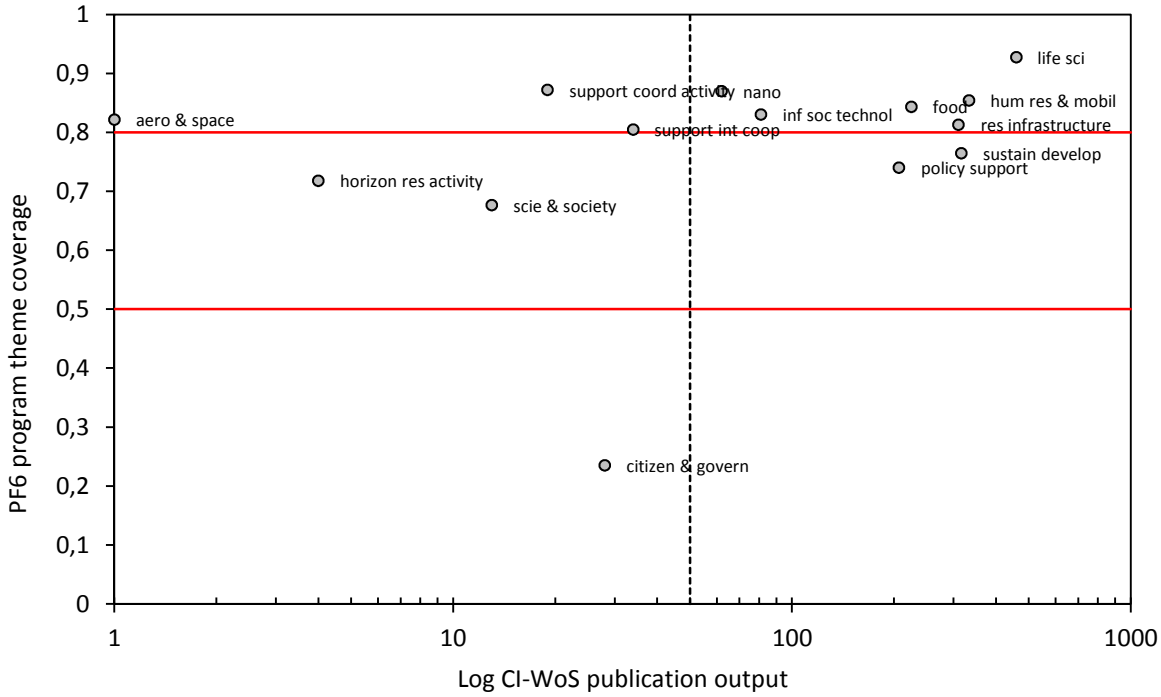


Figure A4. Coverage and validity: Publication output of included FP7 themes compared to their overall internal coverage.

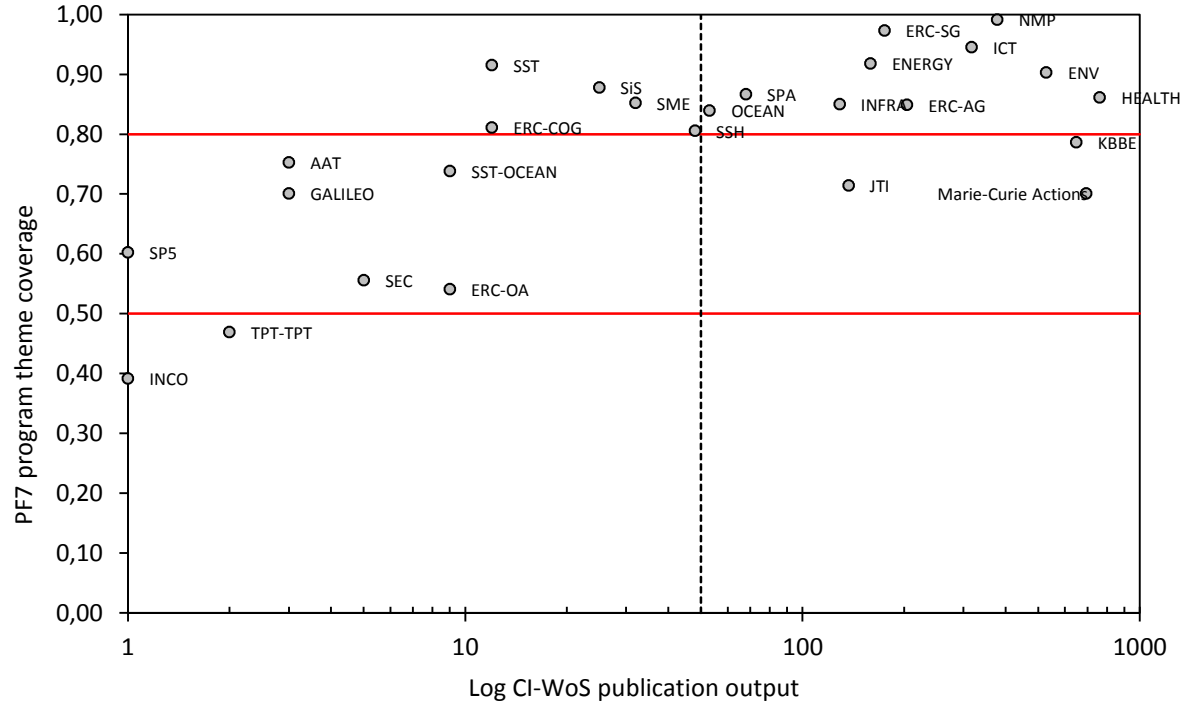


Table A1. Robustness and bias: Number of project-id's included and excluded for the bibliometric analysis [inclusion criteria: a) validated by PIs, or b) EU funding acknowledgement identified in WoS publications from 2008 onwards].

	Included (eligible publications)	Excluded
FP6	175 (171)*	170
FP7	503 (461)*	523

*Notice attrition due to illegible publications

Table A2. Robustness and bias: Distribution of included and excluded FP7 projects according to their main framework theme affiliation.

FP7 themes	Included	Excluded
AAT	2	1
ENERGY	24	27
ENV	48	51
ERC-AG	16	17
ERC-COG	1	0
ERC-OA	1	0
ERC-SG	9	8
GALILEO	1	1
HEALTH	61	64
ICT	44	53
INCO	1	1
INFRA	7	6
JTI	18	23
KBBE	74	76
Marie-Curie Actions	96	81
NMP	48	41
OCEAN	2	4
SEC	2	3
SiS	6	10
SME	17	20
SP5	1	0
SPA	10	10
SSH	6	13
SST	4	6
SST-OCEAN	2	0
TPT-TPT	2	2
n/a	0	4
Total	503	522

Table A3. Robustness and bias: Publication numbers and citation impact for excluded non-validated potential FP6 publications from 2002 to 2013.

	Publications	MNCS	PPtop10%
Full count	6327	1.51	16.6%
Fractional count	4482.4	1.38	11.1%

Table A4. Robustness and bias: Publication numbers and citation impact for excluded non-validated potential FP7 publications from 2007 to 2013.

	Publications	MNCS	PPtop10%
Full count	11704	1.59	17.2%
Fractional count	8458.9	1.39	11.1%

Table A5. Robustness and bias: Differences in overall impact for FP6 and FP7 publications with and without identified WoS publications with funding acknowledgements to EU from 2008 onwards, and various publication windows ending in 2012 or 2013 (full counts).

	$FP6_{pub\ win\ 2012}$	$FP6_{pub\ win\ 2013}$	$FP7_{pub\ win\ 2012}$	$FP7_{pub\ win\ 2013}$
$Publications_{incl.\ WoS\ pubs}$	1710	2020	2131	3583
$MNCS_{incl.\ WoS\ pubs}$	1.74	1.79	2.07	2.03
$PPtop10\%_{incl.\ WoS\ pubs}$	18.9%	19.6%	22.1%	22.2%
$Publications_{excl.\ WoS\ pubs}$	897	985	946	1201
$MNCS_{excl..\ WoS\ pubs}$	1.67	1.66	2.24	2.12
$PPtop10\%_{excl..\ WoS\ pubs}$	19.4%	19.3%	23.2%	22.2%
$Publications_{WoS\ pubs\ alone}$	813	1035	1785	2382
$MNCS_{WoS\ pubs\ alone}$	1.83	1.91	1.99	1.98
$PPtop10\%_{WoS\ pubs\ alone}$	18.4%	20.0%	21.6%	22.2%

Appendix 2: Supplementary tables and graphs

Figure A5. Subject profiles: Base map of WoS journal subject categories based on their cross-citation activities.

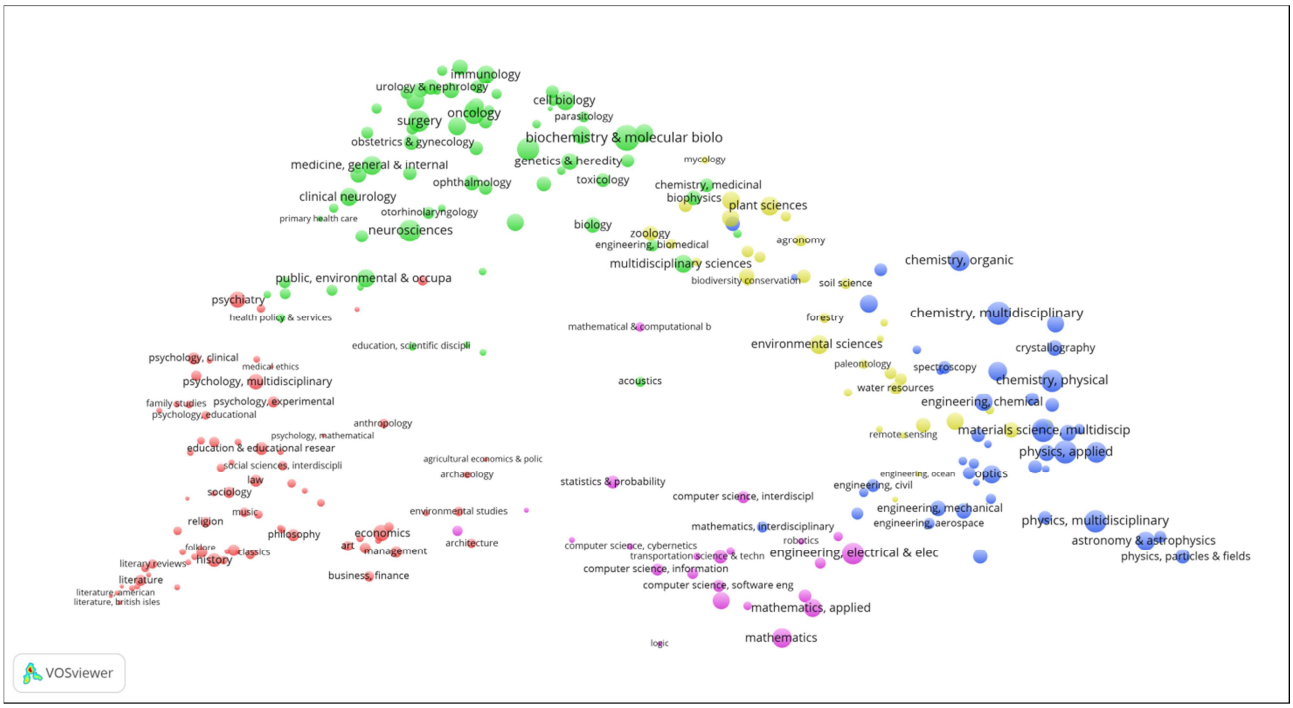


Figure A6: Benchmark analyses: Trend in overall Danish impact (MNCS) when removing various publication sets linked to FP6, DFF and DNRF. The size of the circles illustrates the relative difference in publication size between the different units of analysis.

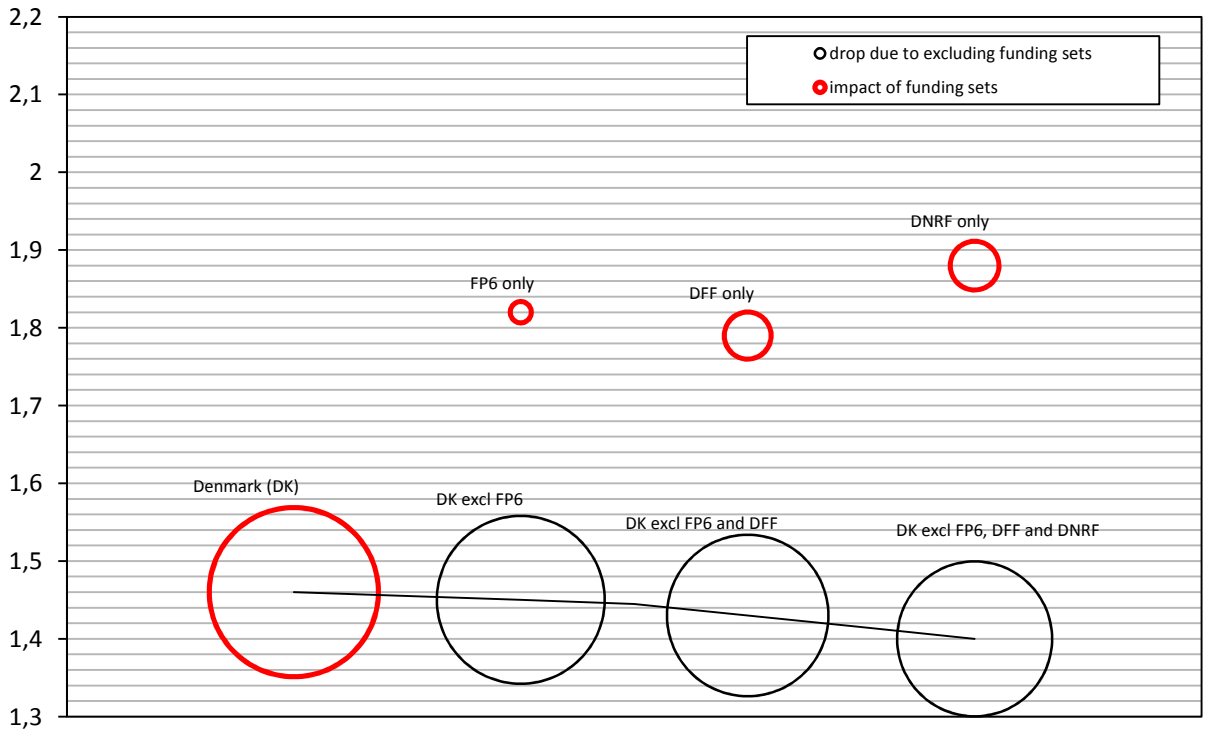


Figure A7: Benchmark analyses: Trend in overall Danish impact (MNCS) when removing various publication sets linked to FP7, DFF and DNRF. The size of the circles illustrates the relative difference in publication size between the different units of analysis.

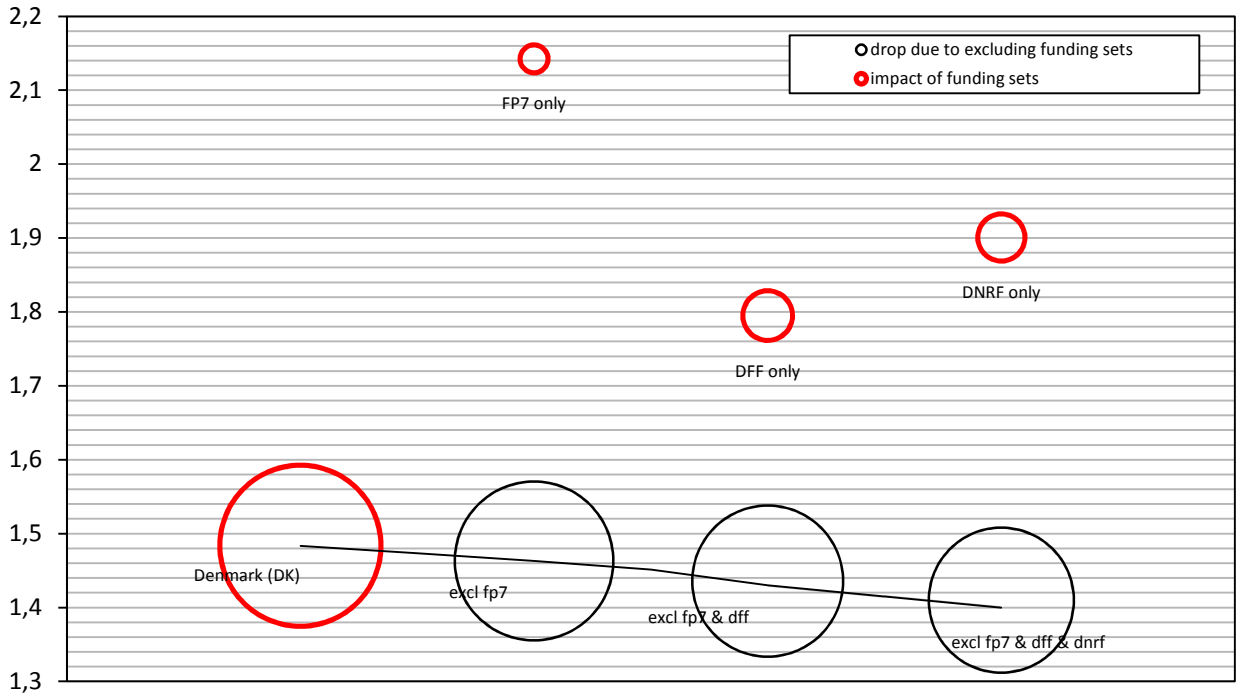


Table A6. "Program themes": Abbreviations of FP6 "program themes".

Abbreviation	Name of FP6 overall "program themes"
life sci	Life sciences, genomics and biotechnology for health
inf soc technol	Information society technologies
nano	Nanotechnologies and nanosciences, knowledge-based multifunctional materials and new production processes and devices
aero & space	Aeronautics and space
food	Food quality and safety
sustain develop	Sustainable development, global change and ecosystems
citizen & govern	Citizens and governance in a knowledge-based society
horizon res activity	Horizontal research activities involving SMEs
hum res & mobil	Human resources and mobility
policy support	Policy support and anticipating scientific and technological needs
res infrastructure	Research infrastructures
sci & society	Science and society
support int coop	Specific measures in support of international cooperation
support coord activity	Support for the coordination of activities

Table A7. "Program themes": Abbreviations of FP7 "program themes".

Abbrevation	Name of FP7 overall "program themes"
AAT	Aeronautics and air transport

ENERGY	Energy
ENV	Environment (including Climate Change)
ERC-AG	Advanced Grants
ERC-COG	Consolidated grants
ERC-OA	Other activities
ERC-SG	Starting Grants
GALILEO	Support to the European global satellite navigation system (Galileo) and EGNOS
HEALTH	Health
ICT	Information and Communication Technologies
INCO	Activities of International Cooperation
INFRA	Research Infrastructures
JTI	Joint Technology Initiative
KBBE	Food, Agriculture and Fisheries, and Biotechnology
Marie-Curie Actions	Marie-Curie Actions
NMP	Nanosciences, Nanotechnologies, Materials and new Production Technologies - NMP
OCEAN	OCEAN.2010/2011
SEC	Security
SiS	Science in Society
SME	Research for the benefit of SMEs
SP5	EURATOM
SPA	Space
SSH	Socio-economic sciences and Humanities
SST	Sustainable surface transport (INCLUDING THE ‘EUROPEAN GREEN CARS INITIATIVE’)
SST-OCEAN	OCEAN.2010/2011
TPT-TPT	HORIZONTAL ACTIVITIES for implementation of the TRANSPORT PROGRAMME

Figure A8. "Program themes": Proportion of the 10 percent most highly cited articles (PPtop10%) as a function of publication output (full counts) for FP6 program themes (FP6 project publications are distributed according the themes their parent projects belong to).

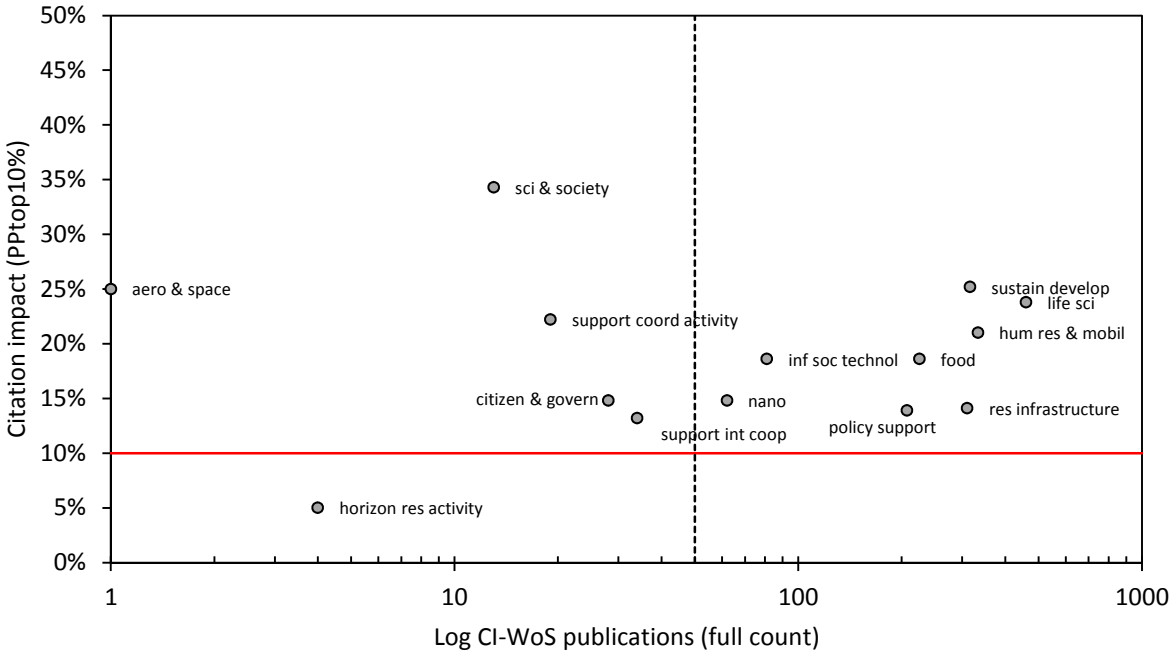


Figure A9. "Program themes": Proportion of the 10 percent most highly cited articles (PPtop10%) as a function of publication output (full counts) for FP7 program themes (FP7 project publications are distributed according the themes their parent projects belong to).

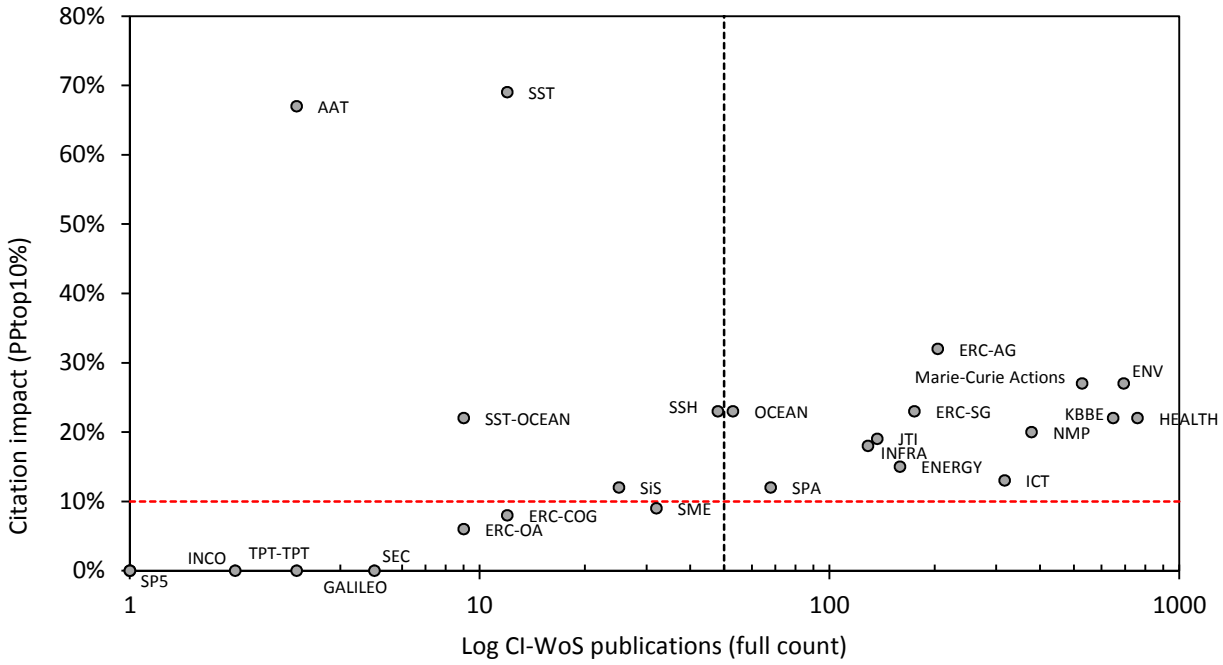


Table A8. "Program themes": Performance statistics for FP6 program themes.

	P	MNCS	PPtop10%	Uncitedness	MNJS	P_{frac}	$MNCS_{frac}$	$PPtop10\%_{frac}$	Coverage
Life sciences, genomics and biotechnology for health	460	2.38	23.8%	7.6%	1.67	250.3	1.58	18.1%	93%
Information society technologies	81	1.67	18.6%	21.0%	1.63	56.6	1.84	20.1%	83%
Nanotechnologies and nanosciences, knowledge-based multifunctional materials and new production processes and devices	62	1.23	14.8%	19.4%	1.08	40.2	1.24	15.4%	87%
Aeronautics and space	1	2.25	25.0%	0.0%	1.34	1.0	2.25	25.0%	82%
Food quality and safety	225	1.42	18.6%	13.3%	1.36	98.8	1.21	13.2%	84%
Sustainable development, global change and ecosystems	316	1.93	25.2%	10.4%	1.32	161.5	1.71	22.2%	76%
Citizens and governance in a knowledge-based society	28	1.36	14.8%	39.3%	1.17	19.9	1.37	15.9%	23%
Horizontal research activities involving SMEs	4	1.43	5.0%	0.0%	0.98	3.5	1.52	5.7%	72%
Human resources and mobility	333	1.97	21.0%	14.7%	1.76	204.0	1.90	20.9%	85%
Policy support and anticipating scientific and technological needs	207	1.29	13.9%	18.8%	1.35	114.0	1.14	11.3%	74%
Research infrastructures	310	1.61	14.1%	28.1%	1.32	132.0	1.36	13.0%	81%
Science and society	13	1.77	34.3%	30.8%	0.93	8.3	1.47	26.8%	68%
Specific measures in support of international cooperation	34	1.16	13.2%	11.8%	1.50	16.6	1.15	10.2%	80%
Support for the coordination of activities	19	1.20	22.2%	10.5%	1.01	14.6	1.08	19.0%	87%

Table A9. "Program themes": Performance statistics for FP7 program themes.

	P	MNCS	PPtop10%	Uncitedness	MNJS	P_{frac}	$MNCS_{frac}$	$PPtop10\%f_{rac}$	Coverage
AAT	3	0.74	0.0%	33.3%	0.91	1.3	0.76	0.0%	86%
ENERGY	159	1.46	15.5%	20.1%	1.23	105.6	1.27	14.4%	76%
ENV	528	2.23	26.9%	11.0%	1.59	257	1.74	21.6%	80%
ERC-AG	204	3.02	32.2%	8.8%	2.23	110.6	2.71	28.6%	91%
ERC-COG	12	4.14	68.6%	0.0%	3.78	7	3.86	68.2%	99%
ERC-OA	9	1.24	22.2%	22.2%	1.35	4.7	0.95	16.1%	97%
ERC-SG	175	1.75	23.0%	6.3%	1.38	52.1	1.84	28.2%	83%
GALILEO	3	2.84	66.7%	0.0%	0.96	3	2.84	66.7%	96%
HEALTH	760	2.36	22.2%	7.9%	1.68	381.8	1.82	19.1%	92%
ICT	317	1.16	13.2%	28.1%	1.18	183.3	1.12	12.7%	72%
INCO	1	0.76	0.0%	0.0%	1.73	0.5	0.76	0.0%	85%
INFRA	129	2.22	17.8%	20.2%	1.45	57.5	2.08	20.2%	88%
JTI	137	1.44	18.8%	13.1%	1.33	92.6	1.35	17.9%	84%
KBBE	648	1.9	22.0%	11.0%	1.49	362.1	1.68	20.8%	82%
Marie-Curie Actions	694	2.34	26.6%	9.4%	1.71	409.5	2.03	23.8%	86%
NMP	378	1.8	19.6%	11.6%	1.62	237.2	1.72	18.5%	89%
OCEAN	53	3.02	23.1%	9.4%	2.43	35	1.88	19.2%	80%
SEC	5	0.34	0.0%	60.0%	0.73	3.2	0.35	0.0%	94%
SiS	25	1.39	11.8%	24.0%	1.34	14.3	1.2	6.5%	57%
SME	32	1.09	9.5%	31.3%	1.31	22	1.1	11.4%	82%
SP5	1	0.26	0.0%	0.0%	1.88	0.3	0.26	0.0%	56%
SPA	68	1.23	12.3%	16.2%	1.09	39.5	1.05	8.9%	79%
SSH	48	1.85	22.8%	14.6%	1.32	19.3	1.36	15.4%	69%
SST	12	0.91	8.3%	58.3%	0.98	8	0.5	2.1%	53%
SST-OCEAN	9	0.91	5.6%	33.3%	1	7.6	0.73	1.6%	41%
TPT-TPT	2	0.77	0.0%	50.0%	2.7	1.3	0.38	0.0%	56%

Appendix 3: Regression analysis

The aim of the regression analysis is to estimate the partial relationship between funding links and impact and to identify the importance of international collaboration in this relationship. We observe that the publications linked to FP funding have very high degree of internationalization and on average generally very high impact levels. By controlling for international collaboration we seek to identify the extent to which the internationalization of research has an importance for the overall impact of FP-funded publications.

The multivariate regression analysis has the main advantage that it is possible to "isolate" the relationship between two variables to a larger extent than with tabulations. It allows us to look at a relationship holding specific variables fixed. In this analysis we want to observe the relationship between citation impact and funding origins holding international collaboration fixed. The data used for the regression analysis are presented in Table A10 below, the restricted benchmark analyses in the period from 2005 to 2011 for FP6 and 2007 to 2011 for FP7. Notice, the set of Danish publications is reduced by isolating the subsets of FP, DNRF and DFF.

The dependent variable is the normalized citation score (NCS). The independent variables include three dummies indicating funding origins, DFF, DNRF and FP6 or FP7. A set of control variables are included to control for the fact that citations are affected by a number of factors such as number of authors, number of references and the visibility and reputation of a journal measured by normalized journal score (NJS). In addition to these variables two indicators of international collaboration are included: a dummy variable that is equal to 1 if the publication is authored by at least two authors in at least two different countries and a count variable indicating the number of different countries in which the authors are employed. The interpretations of the two variables are similar but different, the first indicates whether or not there is international collaboration and the latter indicates the extent of internationalization of the publication.

Table A10: Definition of variables.

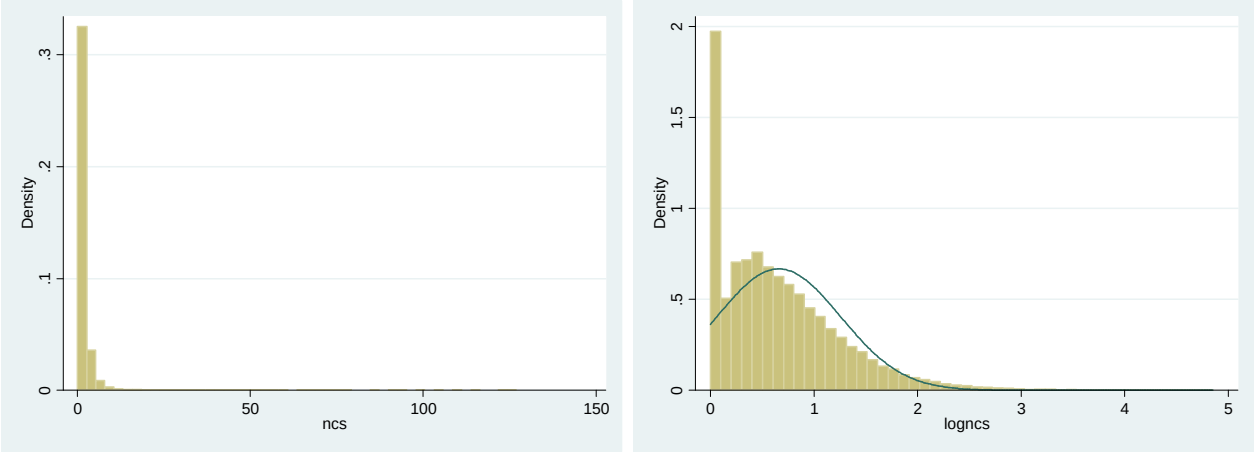
Variable	Definition
NCS	Field Normalized Citation Score
Log(NCS+1)	Natural logarithm of NCS+1
FP6	Dummy indicating funding from FP6
FP7	Dummy indicating funding from FP7
DFF	Dummy indicating funding from
DNRF	Dummy indicating funding from DNRF
NJS	Normalized Journal Score
REFS (	Number of references in publication

AUs	Number of authors
Collab	Dummy indicating international Collaboration
Countries	Count variable indicating number of countries involved in publication
ERC_MARIE_C	Dummy indicating funding from ERC or Marie Curie Actions
FP7_rest	Dummy indicating funding from FP7 and not ERC or Marie Curie Actions

Model specification

The dependent variable NCS has the unfortunate property of being highly skewed with a power tail and a large number of zero-observations. Thus, NCS is an exponential function of explanatory variables rather than a linear function of explanatory variables. To adapt to the distribution of NCS we log-normalize NCS by adding 1 and taking the log. Log (NCS+1) then becomes a linear function of explanatory variables **X**. The reason for adding 1 before taking the log is to keep the zero observations in the dataset. This is an imperfect solution, however we prefer to keep the observations rather than exclude them creating a truncation bias. In the figures below the distribution of NCS and log(NCS+1) is illustrated.

Figure A10: Distribution of normalized citation scores (NCS) before and after log+1 transformation.



Initially two sets of regressions are estimated, the first including publications funded by FP6, DFF, DNRF and Danish publications from 2005-2011 (excluding FP7 funded papers). The second, includes FP7, DFF, DNRF and Danish publications from 2007-2011 (excluding FP6 funded publications). In each of the two sets, three regressions are estimated. By modelling the impact (NCS) as a function of funding and international collaboration we show that adding a dummy for international collaboration diminishes the size of the coefficient between FP6 and FP7 funding on citation impact. This is operationalized by running three regressions and comparing the coefficient

of FP6 before and after including a variable indicating international collaboration. The expectation is that $\beta_{FP6} > \beta'_{FP6} > \beta''_{FP6}$, because internationalization is correlated with higher NCS and the number of countries means larger internationalization.

Table A11: Model specification for the regression analyses on the FP6 set.

Regression set 1 – Coefficient of interest $\boldsymbol{\beta_{FP6}}$
Regression 1 – Years 2005-2011
$\boldsymbol{Log(NCS+1)_i = \beta_0 + \beta_{FP6}FP6_i + \beta_{DFF}DFF_i + \beta_{DNRF}DNRF_i + \beta_{NJS}NJS_i + \beta_{REFS}REFS_i + \beta_{AUs}AUs_i}$
Regression 2 – years 2005-2011
$\boldsymbol{Log(NCS+1)_i = \beta`_0 + \beta'_{FP6}FP6_i + \beta`_{DFF}DFF_i + \beta`_{DNRF}DNRF_i + \beta`_{NJS}NJS_i + \beta`_{REFS}REFS_i + \beta`_{AUs}AUs_i + \beta`_{Collab}Collab_i}$
Regression 3 – years 2005 – 2011
$\boldsymbol{Log(NCS+1)_i = \beta''_0 + \beta''_{FP6}FP6_i + \beta''_{DFF}DFF_i + \beta''_{DNRF}DNRF_i + \beta''_{NJS}NJS_i + \beta''_{REFS}REFS_i + \beta''_{AUs}AUs_i + \beta''_{Collab}Collab_i + \beta''_{countries}Countries_i}$

Table A12: Model specification for the regression analysis on the FP7 set.

Regression set 2 – Coefficient of interest $\boldsymbol{\beta_{FP7}}$
Regression 1 – Years 2007-2011
$\boldsymbol{Log(NCS+1)_i = \beta_0 + \beta_{FP7}FP7_i + \beta_{DFF}DFF_i + \beta_{DNRF}DNRF_i + \beta_{NJS}NJS_i + \beta_{REFS}REFS_i + \beta_{AUs}AUs_i}$
Regression 2 – years 2007-2011
$\boldsymbol{Log(NCS+1)_i = \beta`_0 + \beta'_{FP7}FP7_i + \beta`_{DFF}DFF_i + \beta`_{DNRF}DNRF_i + \beta`_{NJS}NJS_i + \beta`_{REFS}REFS_i + \beta`_{AUs}AUs_i + \beta`_{Collab}Collab_i}$
Regression 3 – years 2007 – 2011
$\boldsymbol{Log(NCS+1)_i = \beta''_0 + \beta''_{FP7}FP7_i + \beta''_{DFF}DFF_i + \beta''_{DNRF}DNRF_i + \beta''_{NJS}NJS_i + \beta''_{REFS}REFS_i + \beta''_{AUs}AUs_i + \beta''_{Collab}Collab_i + \beta''_{countries}Countries_i}$

A third set of regressions is estimated to explore to what extent ERC and Marie Curie publications are associated with a higher expected impact and therefor are main drivers of impact of FP7. We define two dummy variables where ERC_MARIE_C indicating funding by either an ERC or a Marie Curie grant and FP7_rest indicating funding by any other FP7 grant. We use the same control variables as in the two first sets and repeat the three regressions. Here we are interested in to what extent the coefficients $\beta_{ERC\_MARIE\_C}$ & β_{FP7_rest} differ from each other.

Table A13: Model specification for the regression analysis on the FP7 set.

Regression set 3 – Coefficients of interest $\beta_{ERC\_MARIE\_C}$ & β_{FP7_rest}

Regression 1 – Years 2007-2011

$$Log(NCS+1)_i = \beta_0 + \beta_{ERC\_MARIE\_C} ERC\_MARIE\_C_i + \beta_{FP7\_rest} FP7\_rest_i + \beta_{DFF} DFF_i + \beta_{DNRF} DNRF_i + \beta_{NJS} NJS_i + \beta_{REFS} REFS_i + \beta_{AUs} AUs_i$$

Regression 2 – years 2007-2011

$$Log(NCS+1)_i = \beta_0 + \beta_{ERC\_MARIE\_C} ERC\_MARIE\_C_i + \beta_{FP7\_rest} FP7\_rest_i + \beta_{DFF} DFF_i + \beta_{DNRF} DNRF_i + \beta_{NJS} NJS_i + \beta_{REFS} REFS_i + \beta_{AUs} AUs_i + \beta"_{Collab} Collab_i + \beta"_{countries} Countries_i$$

Regression 3 – years 2007 – 2011

$$Log(NCS+1)_i = \beta_0 + \beta_{ERC\_MARIE\_C} ERC\_MARIE\_C_i + \beta_{FP7\_rest} FP7\_rest_i + \beta_{DFF} DFF_i + \beta_{DNRF} DNRF_i + \beta_{NJS} NJS_i + \beta_{REFS} REFS_i + \beta_{AUs} AUs_i + \beta"_{Collab} Collab_i + \beta"_{countries} Countries_i$$

Results

The results for the FP6 set are reported in Table A14. Regression 1 includes dummies for funding from FP6, DNRF, and DFF, NJS indicating the average impact of the journal in which the publication was published, number of references in the publication and number of authors of the publication. The coefficient of FP6 is 0.044 the interpretation is that a publication funded by a FP6 program has an expected NCS+1 that is 4.5 % (exp(0.044)) higher than a publication from the random sample of Danish publications. A publication funded by DNRF has an expected NCS+1 that is 3.4 % higher than a publication from the Danish sample. Thus a publication funded by FP6 has an expected NCS+1 that is 1.1 percentage points higher than a publication funded by DNRF.

Regression 2 includes a dummy indicating whether the publication was an international co-publication (written by authors employed in two or more different countries). Including this dummy decreases the coefficient of FP6 to 0.035. This is expected since there is a higher volume of international collaboration within this group of publications. Thus, when controlling for the fact that publications with FP6 grants tend to be international collaborations and therefore have a higher expected NCS, the expected value of NCS+1 is 3.6 % higher for a publication financed by FP6 compared to a Danish publication not funded by DFF, DNRF or FP6. The interesting part is that the coefficient on DNRF only changes slightly; therefore the difference in expected NCS+1 is only 0.03 percentage points higher for FP6 financed publications relative to DNRF publications.

Regression 3 includes in addition to a dummy variable indicating international collaboration a count variable of the number of countries the authors of the publication work in. This indicates not only whether the publication is international but also how many different countries are involved. The effect of including this variable is that the dummy variable indicating international

collaboration becomes insignificantly small from 0.06 to 0.006 this is because of the high correlation between the two variables.

The interpretation of the coefficient on countries is that for every extra country that is involved in the publication the expected value of NCS+1 will increase by 2.8 %. The coefficient on FP6 decreases by the inclusion of the Countries variable which implies that a large part of the correlation between FP6 funding and citation impact is related to the fact that there are many publication with many countries and that publications with many countries generally have a higher citation impact.

Table A14: Results of OLS regression analyses on the FP6 set.

	Model (1)	Model (2)	Model (3)
Dependent variable:	Log(ncs+1)	Log(ncs+1)	Log(ncs+1)
FP6	0.044	0.035	0.026
	(0.015)	(0.015)	(0.015)
DNRF	0.034	0.032	0.036
	(0.007)	(0.007)	(0.007)
DFF	0.019	0.020	0.025
	(0.007)	(0.007)	(0.007)
NJS	0.217	0.215	0.210
	(0.008)	(0.008)	(0.008)
REFS	0.004	0.004	0.003
	(0.0002)	(0.0002)	(0.0002)
AUs	0.0002	0.0002	-0.0002
	(0.0000)	(0.0000)	(0.0000)
Collab		0.062	0.006
		(0.004)	(0.005)
Countries			0.028
			(0.002)
Constant	0.250	0.221	0.204
	(0.010)	(0.010)	(0.010)
Observations	77,308	77,308	77,308
R-squared	0.253	0.256	0.262

Note: Robust standard errors are reported in parentheses. Robust Standard errors are calculated because of the presence of heteroskedasticity. The regression includes data from 2005-2011. Publications funded by FP7 are excluded except in the case where the publication is also funded by FP6.

The three regressions indicate that the citation impact of FP6 funded publications largely stems from the fact that they are to a higher degree international collaborations compared to other funded publications and that international collaborations in general have a higher citation impact. At the

same time the coefficients of DNRF and DFF are affected (weakly) positively when including the variables indicating internationalization. The reason for this may be that especially DFF is more directed at national interests and that the publications have a limited number of countries, and that the impact of national (mono country) research is higher than average. Therefore when controlling for number of countries DFF granted publications are relatively high performers.

Table A15 presents the same regressions as above; however, this time includes a dummy for FP7 participation instead of FP6. The same pattern is observed, however, the impact of FP7 is much higher than FP6 and even after controlling for international collaboration FP7 has a considerably higher impact than DNRF, DFF and other Danish publications. This indicates that FP7 has a general higher impact than FP6. The coefficient of DNRF and DFF indicate that DNRF has a general stable and high impact regardless of level of collaboration while DFF has a higher impact in terms of national and small scale international collaboration.

Table A15: Results of OLS regression analyses on the FP7 set.

	Model (1)	Model (2)	Model (3)
Dependent variable:	Log(ncs+1)	Log(ncs+1)	Log(ncs+1)
FP7	0.070	0.065	0.057
	(0.013)	(0.013)	(0.013)
DNRF	0.033	0.031	0.033
	(0.008)	(0.008)	(0.008)
DFF	0.021	0.023	0.029
	(0.008)	(0.008)	(0.008)
NJS	0.215	0.212	0.208
	(0.009)	(0.009)	(0.009)
REFS	0.004	0.004	0.003
	(0.0002)	(0.0002)	(0.0002)
AU's	0.0002	0.0002	-0.0002
	(0.0000)	(0.0000)	(0.0000)
Collab		0.061	0.006
		(0.005)	(0.006)
Countries			0.027
			(0.002)
Constant	0.253	0.224	0.208
	(0.012)	(0.011)	(0.011)
Observations	58,933	58,933	58,933
R-squared	0.265	0.268	0.274

Note: Robust standard errors are reported in parentheses. The regression includes data from 2007-2011. Publications funded by FP6 are excluded except in the case where the publication is also funded by FP7.

Table A16 shows the third set of regressions that include the same variables and observations as the prior regression; however, the FP7 dummy is now replaced by two disaggregated dummies

indicating either 1) an ERC or Marie Curie grant or 2) any other FP7 grant. The expected NCS+1 of a publication funded by ERC or Marie Curie grant is 10 % higher than the baseline of Danish publications. While publications funded by other FP7 grants have an expected NCS+1 that is 4 % higher than the baseline. Thus dividing FP7 into these two groups gives a more diverse picture of the expected impact of publications given their funding links. In the prior regressions we found that the expected marginal effect of FP7 funding is 6 % relative to the baseline set. In this regression we can observe that the effect of ERC and Marie Curie grants are higher at 10 % while other FP7 grants have a marginal effect on NCS that is 5 %. This suggest that impact of publications funded by ERC and Marie Curie is larger than other FP7 funded publications after controlling for internationalization and they in turn have an expected impact that is higher than DNRF and DFF and that all these funding regimes result in a higher expected impact compared to the baseline of Danish publications. The results also indicate that it is meaningful not only to distinguish between overall types of funding sets but also between differences within funding sets as their aims and purposes differ considerably. These aims and methods may be correlated with the impact of the publications.

Table A16: Results of OLS regression analyses on the FP7 set with two dummies.

	Model (1)	Model (2)	Model (3)
Dependent variable:	Log(ncs+1)	Log(ncs+1)	Log(ncs+1)
MARIE_CURIE_C	0.115	0.111	0.098
	(0.029)	(0.029)	(0.029)
FP7_rest	0.054	0.048	0.042
	(0.014)	(0.014)	(0.014)
DNRF	0.031	0.029	0.031
	(0.008)	(0.008)	(0.008)
DFF	0.021	0.023	0.029
	(0.008)	(0.008)	(0.008)
NJS	0.215	0.212	0.208
	(0.010)	(0.009)	(0.009)
REFS	0.004	0.004	0.003
	(0.000)	(0.000)	(0.000)
AUs	0.0002	0.0002	-0.0002
	(0.000)	(0.000)	(0.000)
Collab		0.061	0.006
		(0.005)	(0.006)
Countries			0.0265
			(0.002)
Constant	0.253	0.224	0.208
	(0.012)	(0.011)	(0.011)
Observations	58,933	58,933	58,933
R-squared	0.266	0.268	0.274

Note: Robust standard errors are reported in parentheses. Robust Standard errors are calculated because of the presence of heteroskedasticity. The regression includes data from 2005-2011. Publications funded by FP7 are excluded except in the case where the publication is also funded by FP6.

Literature

Martin, B. R., Nigthingale, P., & Rafols, I. (2014). Response to the call for evidence to the independent review of the role of metrics in research assessment.

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

Schneider, J. (2015). Null hypothesis significance tests. A mix-up of two different theories: The basis for widespread confusion and numerous misinterpretations. *Scientometrics, 102*(1), 411-432.

Schneider, J. W. (2013). Caveats for using statistical significance tests in research assessments. *Journal of Informetrics, 7*(1), 50-62.