Fraunhofer
ISI

Performance and Structures of the German Science System 2014

Carolin Mund, Rainer Frietsch, Peter Neuhäusler

Studien zum deutschen Innovationssystem
Nr. 7-2015

Fraunhofer Institute for Systems and Innovation Research ISI

February 2015

This study was conducted on behalf of the Expertenkommission Forschung und Innovation (EFI). The results and interpretations are the sole responsibility of the institute conducting the study. The EFI exercised no influence on the writing of this report.

Studien zum deutschen Innovationssystem

Nr. 7-2015

ISSN 1613-4338

Publisher:
Expertenkommission Forschung und Innovation (EFI)
Geschäftsstelle
c/o Stifterverband für die Deutsche Wissenschaft
Pariser Platz 6
10117 Berlin
http://www.e-fi.de/

Contact address and further information:

Rainer Frietsch

Fraunhofer Institute for Systems and Innovation Research ISI
Competence Center Policy and Regions
Breslauer Strasse 48
76139 Karlsruhe
Phone: +49-721-6809-197
Fax: +49-721-6809-176
e-mail: rainer.frietsch@isi.fraunhofer.de

Contents

Figures

Tables

0 Executive Summary

The absolute number of scientific publications continued to increase also in 2013 – a trend that holds for the whole observation period since 2003. While the majority of the worldwide growth is still mainly driven by China all countries but Japan increased their annual publication output – however, with varying effects on the other bibliometric indicators. A trade-off becomes apparent especial for countries like the USA, Switzerland and Germany. The conference proceedings are also foremost increasing for China, so that in the end of our observation period nearly every third conference proceedings publication worldwide originates in China. China is also catching up in the citation based indicators like the observed citation rate, Scientific Regard, International Alignment and the Excellence Rate.

Germany's development in the short-term perspective might raise concerns about its performance. The absolute number of publications grew only by 1.8% between 2012 and 2013 and even in a longer-term perspective since 2002 it only managed an average annual growth rate of about 2.4%, while the total worldwide publications grew by 5%. Furthermore, the trend towards higher ranked journals continues (international alignment), but German researchers do not achieve the same citation rates within these journals (scientific regard). It becomes apparent that the additional investments in the science system since the mid 2000s – especially in the context of the High-Tech Strategy – did not yet result in considerably higher publication numbers. However, at least the excellence (top cited publications) and the average quality (in terms of citations) seem to benefit from these investments.

The growth of Chinese publications affects all scientific fields, so that in general, the influence of other countries has sunken. However, the USA still dominate the Social Sciences. China has especially high shares in Optics, chemical fields and Materials research, while Germany's main fields are Medical engineering, Nuclear technology and Physics. Such differences have led to increasing shares in co-publication in these fields, where the country with less expertise in the respective field profits from such knowledge flows. China has in general strengthened its co-publication network, especially with the USA. The latter in turn have withdrawn from many former co-operation partners and show a high share of purely national publications in all scientific fields.

A special focus on acknowledgements in German publications reveals that the DFG – as the largest funding provider – has a similar importance for all scientific fields. This holds even for the humanities, where organizations like the DAAD and BMBF play next to no role. However, in general the importance of other funding sources have increased in the last couple of years, showing in diminishing shares in all publications of the main funding organizations, especially the DFG.

The German regions show a large disparity in absolute publication counts, where Baden-Württemberg, North-Rhine-Westphalia, Bavaria, Hessia and Lower Saxony rank first. However, the share of the regions of Germany's publications is converging, meaning that publications stem more and more often from a variety of German regions. It is interesting to note that especially those regions with high absolute publication counts also show a relatively low citation rate, meaning that they cannot assure quality to such an extend as (publication-wise) smaller regions.

A special focus this year lay on the publication behaviour of German companies to describe structures and trends. In particular, the publications were analyzed in terms of co-authors, scientific disciplines, impact and journal characteristics. The statistics of the companies were compared to that of Germany in total. Because of the small portion of German publications that is emitted by companies, the influence of their behaviour on the average is relatively small. Therefore, most observations in which both showed a similar behaviour are interpreted as the companies complying with the German standards. However, the in-depth analysis showed that similarities are mostly superficial; e.g. the only slightly increasing share of publications shows a different distribution across the scientific disciplines. Also, the citation rates of German companies and Germany in total are nearly equal, but stem from a different set of organizations. It was shown that publications by companies are more often cited by companies than those of other sectors, showing their high utility in the industrial context. In particular, these publications are nearly twice as often cited by other companies than those written by at least one author from a university.

In total, the publications with at least one author from a company have been rising in numbers. However, their share in the German publication output is relatively stable in the observation period.

The co-publication pattern of the companies showed that their publication behaviour is similar to that of the other organizations. Almost 50% of their publications are written with a co-author from a university. In nearly equal shares, international partners are involved. Thus, not all observations made for the respective publications can be fully attributed to the companies. This underlines the hypotheses from literature that the co-publications are an instantiation of co-operations, outsourcing of research or a provision of use-cases. The latter again also would support the signalling effect theory in that it is basically one form of marketing. The findings for the scientific disciplines underline that especially costly or rare resources stimulate collaborations in the sector of industry.

As mentioned above, the publications by companies have citation rates comparable to that of the German average, even though they target journals with a lower visibility, which are also by trend more often German language journals than is the case for other German publications. In these journals they achieve, however, relatively high citation rates in comparison with similar articles (SR value, cf. Figure 39). This counterbalances the lower visibility of these journals and leads to a citation rate which is nearly equal to that of Germany as whole.

Overall, the bibliometric data suggests that the companies indeed use scientific publications as a mean for dissemination and marketing. Of course, it has to be acknowledged that – apart from a prevention of plagiarism – this is generally the main purpose of scientific publications. However, most notable is that companies seem to discover this possibility, reflecting in steadily but slowly increasing publication numbers. There is no indication for qualitative differences to the German average, but the publication behaviour has been adapted to fit to the companies' specific needs and target audience. In that way a different set of journals is used for the publications that relies more (but not exclusively) on national and specialized topics. This especially becomes visible when looking at the specialization (RLA) values. German companies can be shown to publish more often in the fields of engineering as well as medicine and pharmaceuticals and less often in natural sciences, where mostly universities seem to be responsible German positioning with regard to specialization values.

With regard to the characteristics of the publishing firms, it has been found that especially large companies are more prone to publish. Large companies publish significantly more than small firms, which might be a result of the availability of the respective resources in terms of research output that can be generated but could also have to do with different motives to publish for larger firms. Large firms might more often outsource research activities or pursue "defensive" publication strategies. The signalling value of a publication might also be a significant motivation especially for large firms to publish their research results. Also, older firms publish more than younger ones. With regard to the sector specific distribution, we can observe that companies in the manufacturing sector publish more than in other sectors.

1 Introduction to this issue

The scientific system depends on the communication of discoveries and experiences. Scientific publications instantiate such developments and build the foundation for a fluctuant and adaptable knowledge system. Their analysis can shed light on frontier research, co-operations, structures, changes and the role of institutions in science systems.

In this report, the bibliometric performance of a set of 24 countries (see Appendix p. 70) is analyzed. The focus lies on Germany's performance in this global context, but the content of this report has a general broad view. Former reports already focussed on various aspects of Germany and its role in the scientific community. A focus on East Germany (Schmoch and Schulze, 2010) showed the relatively high growth rate in this region that confirmed the success of the respective efforts. In 2011, a closer look on public non-university research organisations was taken (Schmoch et al., 2011), which is complemented by this year's supplementary analysis of companies. Another aspect was the changing publication behaviour of German authors (Schmoch et al. 2012).

Following issues are related to the methodological basis in this study. Firstly, the journal publications are retrieved from the Science Citation Index Expanded (SCIE) and the Social Science Citation Index (SSCI), and conference proceedings are from the CPCI, which are all sub-products of the database Web of Science (WoS). Secondly, the analysis covers "articles", "letters", "notes" and "reviews" for journal papers, and "articles" and "proceedings" for proceeding papers. Thirdly, in contrast to former reports, most analyses use fractional counting of the publications. In that way, the publications are weighted on the relative share of a country. Whole count is used for the co-publication analysis, where a fractional counting is less useful. Fourthly, as it is noted that the external citations are the most relevant for evaluative purpose, this study follows the recommendation of CWTS to exclude self-citations (Nederhof, 1993). As did previous reports in this series, the absolute numbers as well as shares of publications and citations, Scientific Regard (SR), and International Alignment (IA) for selected countries and regions are analyzed in this report. This years' report provides an overview of latest trends and continues most of the indicators collected in earlier years of this reporting system. To keep it concise and handy, we did not include an extensive methodological overview. Interested readers in methods are referred to Michels et al. (2013).

2 Part I: Journal publications in an international comparison

2.1 Number of publications

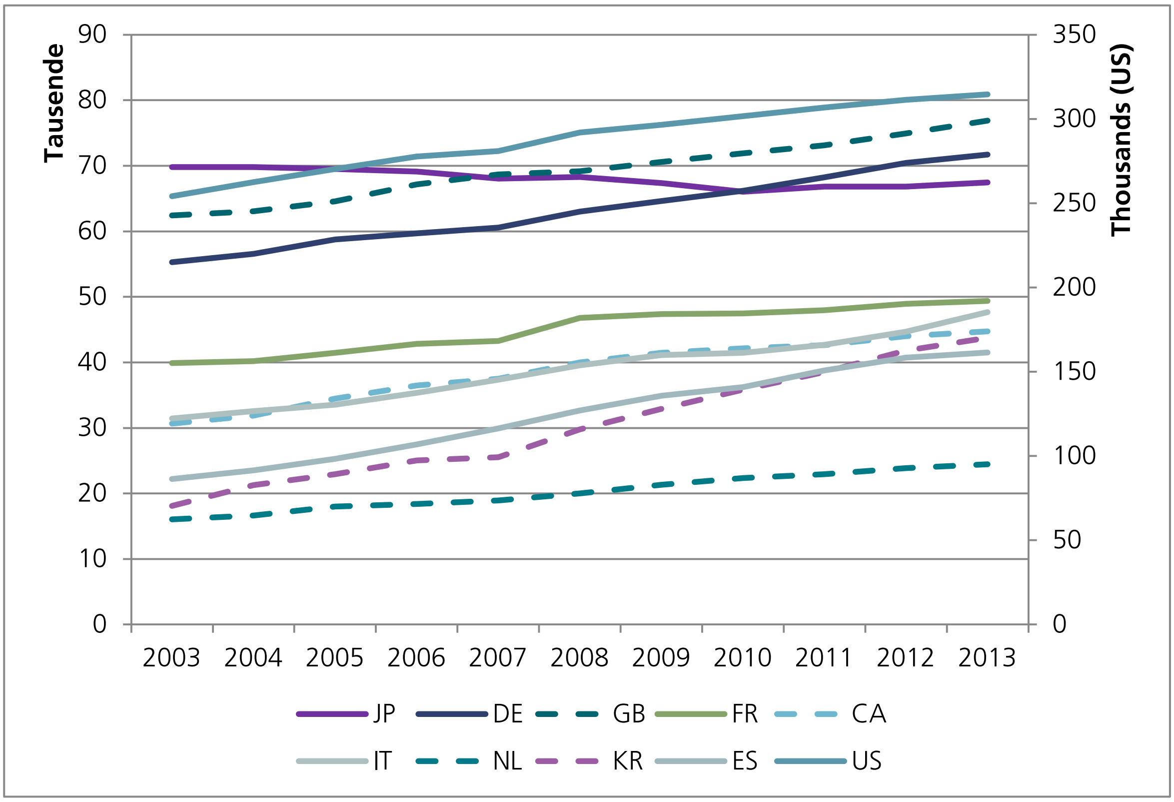
Scientific publications were growing in numbers over the past decades as additional resources were invested in science systems in many countries, additional countries entered the stage of international sciences, and publications even gained in importance as an output scientific activities. In addition, due to growth and database effects (see e.g. Michels and Schmoch, 2012) publication numbers worldwide steadily increase in the observation period. The countries deal differently with the demands for higher gross publication outputs.

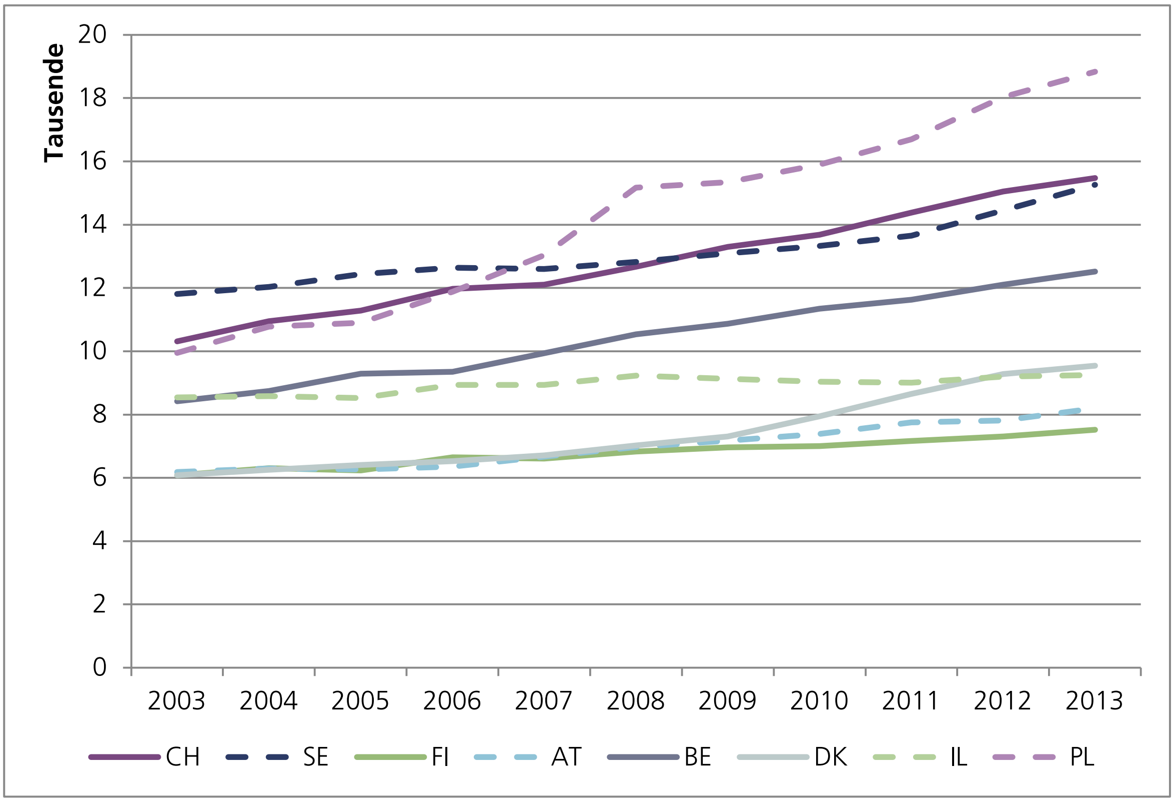
Figure 1 shows the publication output of the selected industrialized countries in the WoS. Since the countries work on very different output levels, the graph had to be adapted to assure a good visibility for all countries: On the one hand, the USA have such a high publication output that they need their own scale (right hand side in the upper figure). On the other hand, we analyse a large number of countries so that the figure was split up. Ten countries that publish the most are shown in the upper figure, while the remaining industrialized countries are shown in the figure in the lower panel.

As the different scales show, the USA still have a yearly publication output far higher than those of the other countries. China is the only country that comes close to their numbers (Figure 2). Regarding only the industrialized countries, Great Britain has the second highest publication number. However, with approximately 315.000 in contrast to 77.000 publications in 2013, the influence of the USA on the publishing market is much higher.

Only few changes can be observed in the ranking of the countries. Germany, Denmark and Poland continue their relatively steep increase in publication numbers. Japan is still the only country with a declining trend in the publication numbers.

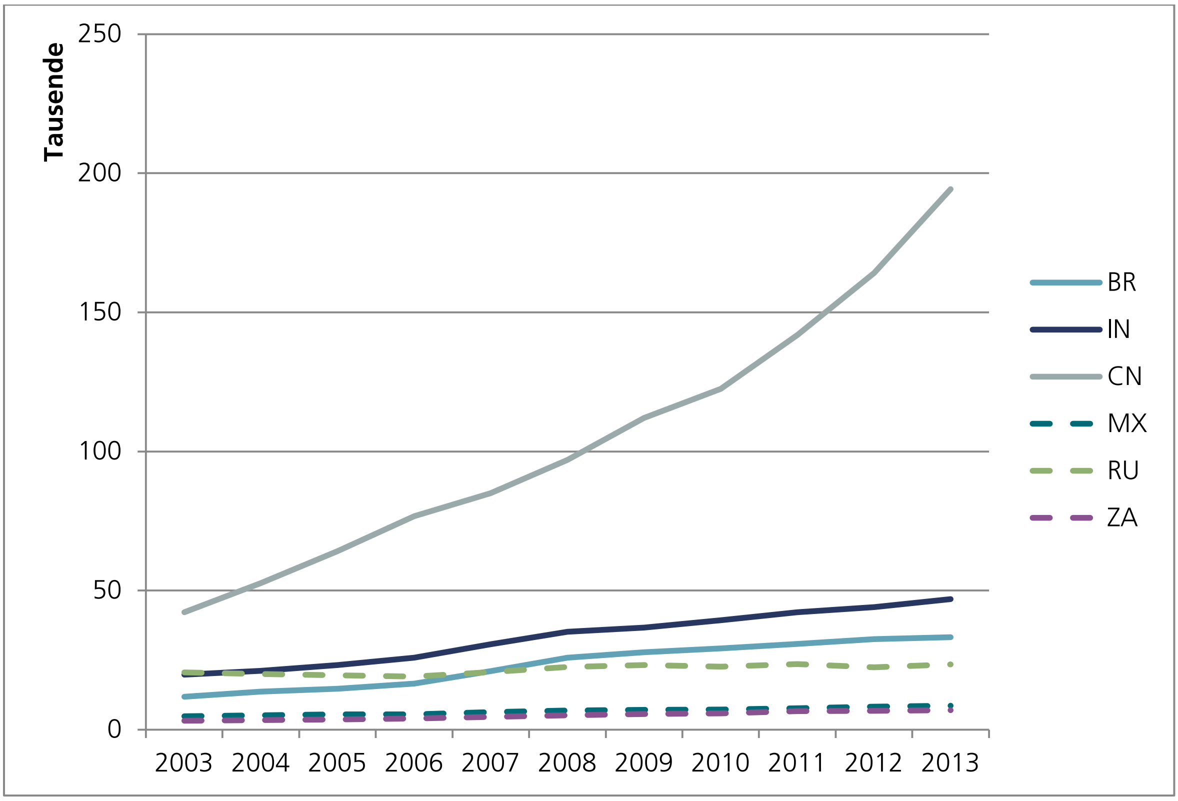
Figure 1: Publication numbers of the selected industrialized countries in the SCIE and the SSCI (fractional counting)





Source: Web of Science, queries and calculations by Fraunhofer ISI

Figure 2: Publication numbers of the BRICS countries and Mexico in the SCIE and the SSCI (fractional counting)



Source: Web of Science, queries and calculations by Fraunhofer ISI

Figure 2 shows the publication numbers for the BRICS countries.[1] China's publication output continued to increase, almost reaching the level of 200,000 publications in sciences and technology in 2013. The compound annual growth rate (CAGR) is 16.3% between 2002 and 2013. The impressive growth even increased after 2010, not only due to a broader coverage of Asian journals in the database, but also due to an increased international orientation of Chinese researchers and a general increase in scientific capabilities in the country. The government continued to heavily invest in the public science system and also kept the incentives high for researchers to publish internationally. India and Brazil were also able to increase their absolute annual publication output, reaching CAGR levels of 9.1% and 10.7%, respectively. India meanwhile reaches a level of about 47,000 publications, which is about half the number of German publications in these areas. Also Mexico and South Africa are able to increase the absolute number of publications by more than 6% on average per year. Only Russia is hardly able to increase its scientific output in the observation period. This is one indication of a far too low investment in the science and innovation system in Russia. It seems to further lose contact with other countries and even within the group of BRICS countries – these were the promising science and innovation nations at the beginning of the 2000s – Russia seems to be decoupled from the overall

1 In consistency with last year's report, the BRICS countries have been supplemented by Mexico.

trends (Schubert et al. 2013). However, all these countries (including China) still reach a level of publication output relative to their number of inhabitants that is still far below the level of most industrialised countries. In consequence, one can expect further increases of the absolute numbers also in the future.

Table 1 shows the absolute publication numbers per year in relation to the number in 2003. The worldwide increase accounts for 64% between 2003 and 2013, heavily affected by the trends in China, Brazil or India. The traditional science-oriented and industrialised countries were hardly able to considerably increase their publication output. The USA, the United Kingdom, France, Finland or Sweden only published 25 to 30% more publications in 2013 than in 2002. Germany was able to increase its publication numbers by 30% as well, with a slightly steeper increase after 2006 – the year of the first implementation of the High-Tech Strategy, which resulted in an enormous growth of public investment also in the science system. This additional investment is so far hardly visible in the absolute numbers of scientific publications, though a number of additional researchers entered the system. Explanations are time lags between input and output, but also a stronger focus on application and innovation, which does not always result in scientific publications, but other output. On the other hand, also other countries have increased their investments in their science and innovation systems in the recent years so that an outstanding effect for Germany cannot be expected.

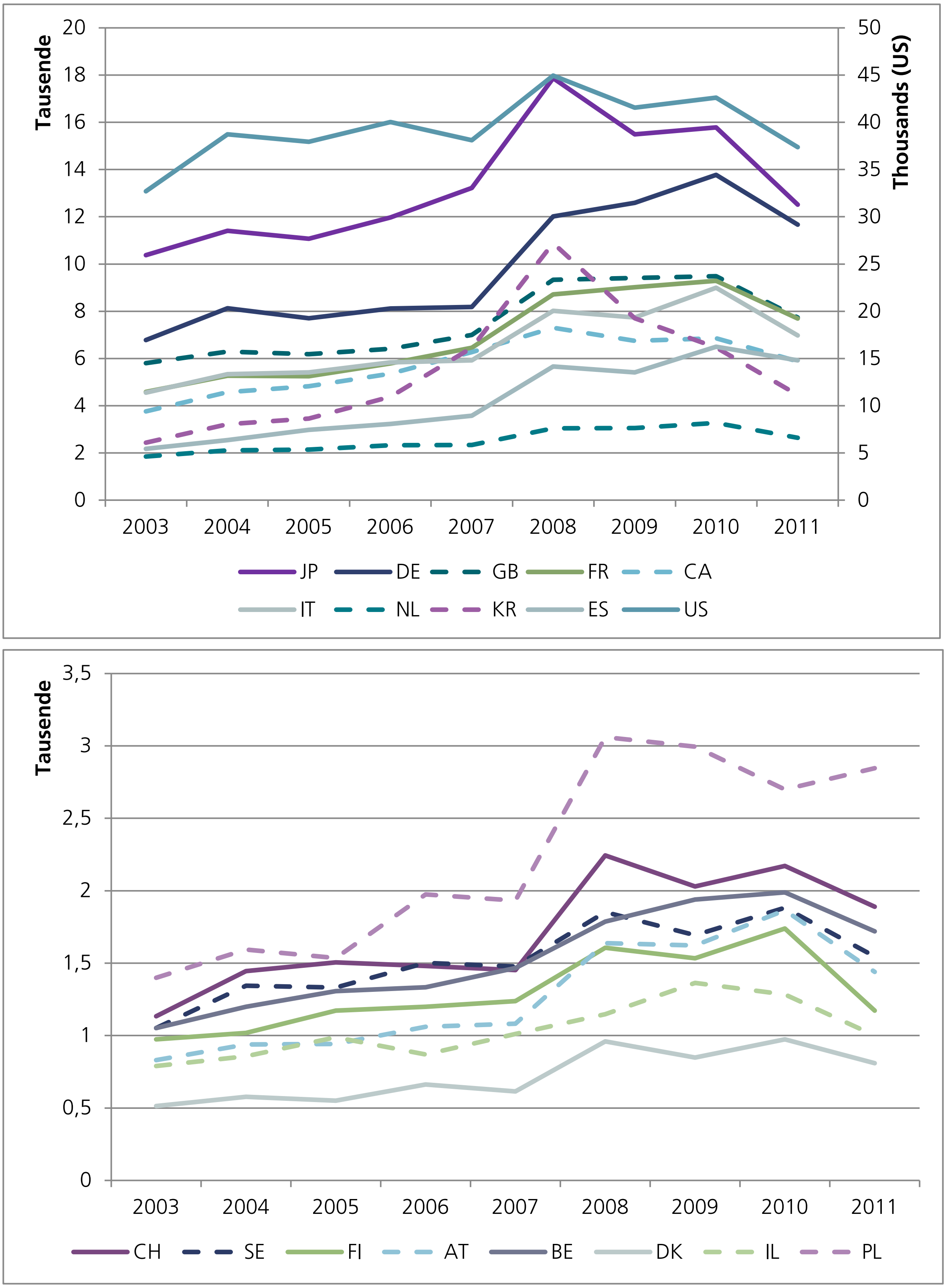
Several also European countries were able to increase their publications even above the world average (Spain, Poland) or close to the world average (Denmark, Netherlands, Italy or Belgium). Japan is the only country under observation here that even publishes fewer articles in 2013 than in 2002. The Japanese government, however, recently took action to overcome the shortcomings of the system and its involvement in international science networks. It needs to be seen, if the action takes effect. Next to the low absolute numbers of scientific publications the low level of international collaboration is a shortcoming of Japan. The government published a white book that addresses exactly this issue.

Table 1: Development of the publication numbers of the selected countries and regions in the SCIE and the SSCI according to fractional counting (Index 2003=100)

Country/region	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
AT	100	102	101	103	108	113	116	120	125	126	133
BE	100	104	110	111	118	125	129	135	138	144	149
BR	100	116	124	140	178	219	236	247	261	275	281
CA	100	104	112	119	122	131	135	137	139	144	146
CH	100	106	109	116	117	123	129	133	139	146	150
CN	100	125	152	181	201	230	265	290	336	388	460
DE	100	102	106	108	110	114	117	120	123	127	130
DK	100	103	105	107	110	115	120	130	142	153	157
ES	100	106	114	124	135	147	157	163	175	183	187
FI	100	104	103	109	109	112	114	115	118	120	124
FR	100	101	104	107	109	117	119	119	120	123	124
GB	100	101	103	108	110	111	113	115	117	120	123
IL	100	100	100	105	105	108	107	106	105	108	108
IN	100	107	118	131	155	178	186	199	214	223	238
IT	100	104	107	112	119	126	131	132	136	142	152
JP	100	100	100	99	98	98	97	95	96	96	97
KR	100	117	127	138	141	164	182	198	213	231	242
MX	100	107	113	114	132	142	147	150	159	171	178
NL	100	104	112	115	118	125	133	139	143	149	152
PL	100	108	110	119	131	153	154	160	168	181	189
RU	100	97	95	93	101	109	113	110	114	109	114
SE	100	102	105	107	107	109	111	113	116	122	129
USA	100	103	106	109	111	115	117	119	121	122	124
ZA	100	107	113	127	142	160	175	182	207	210	217
EU-12	100	106	111	121	140	162	167	173	178	187	195
EU15	100	103	107	111	115	120	124	127	131	136	140
EU28	100	103	107	112	117	124	128	131	136	140	144
World	100	105	110	117	123	132	138	143	150	157	164

Source: Web of Science, queries and calculations by Fraunhofer ISI

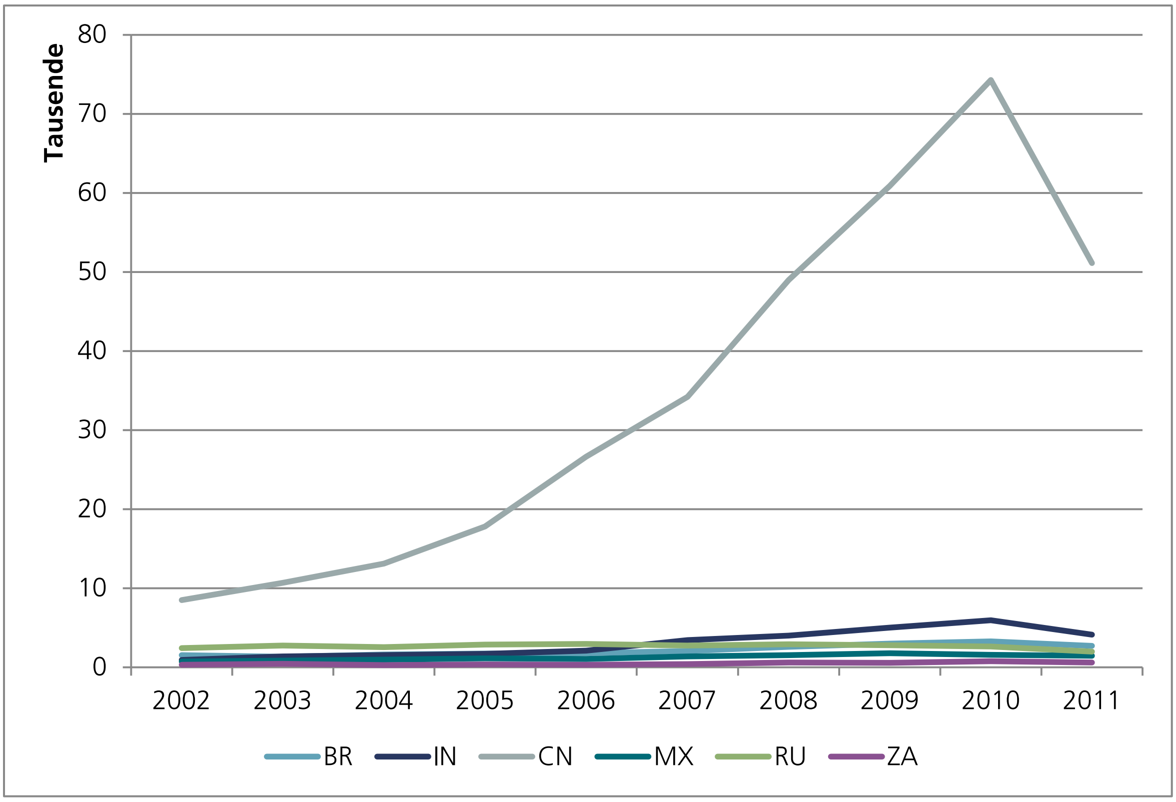
Figure 3: Number of conference proceedings of the selected industrialized countries in the CPCI (fractional counting)



Source: Web of Science, queries and calculations by Fraunhofer ISI

A focus on conference proceedings offers additional insights. They are capable of reflecting short-term and application-oriented research output, but which are especially relevant in certain research fields like computer sciences, electrical engineering or also mechanical engineering. Conference proceedings are more volatile due to the selective data coverage in the Web of Science (cf. Michels and Fu, 2014). Thus, the trends in the conference proceedings are not as clear as in the journal publications. However, here as well China and the USA show exceptionally high numbers (Figure 3 and Figure 4), with China even ahead of the USA. In 2010, China published more than 8 times as many conference proceedings as in 2002. At the current edge, this amount has decreased again to a number of approximately 50,000 conference proceedings in 2011, which is still the highest value in the whole country set.

Figure 4: Number of conference proceedings of the BRICS countries in the CPCI (fractional counting)



Source: Web of Science, queries and calculations by Fraunhofer ISI

2.2 Share of publications

To better relate the numbers of the single countries, Table 2 shows their share of the worldwide publication output. While the USA still hold the highest share in 2013, the influence of China is increasing further. Germany's share has decreased in the observation period from 6.3% to 4.9%, due to the higher growth rates in other countries.

Table 2: Shares of the selected countries and regions in percent in the SCIE and the SSCI within all publications (fractional counting)

Country/region	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
AT	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
BE	1.0	0.9	1.0	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
BR	1.3	1.5	1.5	1.6	2.0	2.2	2.3	2.3	2.3	2.3	2.3
CA	3.5	3.4	3.5	3.6	3.5	3.4	3.4	3.3	3.2	3.2	3.1
CH	1.2	1.2	1.2	1.2	1.1	1.1	1.1	1.1	1.1	1.1	1.1
CN	4.8	5.7	6.6	7.5	7.9	8.3	9.2	9.7	10.7	11.9	13.4
DE	6.3	6.1	6.0	5.8	5.6	5.4	5.3	5.2	5.1	5.1	4.9
DK	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7
ES	2.5	2.5	2.6	2.7	2.8	2.8	2.9	2.9	2.9	2.9	2.9
FI	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.5	0.5	0.5
FR	4.5	4.3	4.3	4.2	4.0	4.0	3.9	3.8	3.6	3.5	3.4
GB	7.1	6.8	6.6	6.5	6.4	5.9	5.8	5.7	5.5	5.4	5.3
IL	1.0	0.9	0.9	0.9	0.8	0.8	0.7	0.7	0.7	0.7	0.6
IN	2.2	2.3	2.4	2.5	2.8	3.0	3.0	3.1	3.2	3.2	3.2
IT	3.6	3.5	3.4	3.4	3.5	3.4	3.4	3.3	3.2	3.2	3.3
JP	7.9	7.5	7.1	6.7	6.3	5.9	5.5	5.2	5.0	4.8	4.7
KR	2.1	2.3	2.4	2.4	2.4	2.6	2.7	2.8	2.9	3.0	3.0
MX	0.6	0.6	0.6	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.6
NL	1.8	1.8	1.8	1.8	1.8	1.7	1.7	1.8	1.7	1.7	1.7
PL	1.1	1.2	1.1	1.2	1.2	1.3	1.3	1.3	1.3	1.3	1.3
RU	2.3	2.2	2.0	1.9	1.9	1.9	1.9	1.8	1.8	1.6	1.6
SE	1.3	1.3	1.3	1.2	1.2	1.1	1.1	1.1	1.0	1.0	1.1
USA	28.8	28.3	27.8	27.0	26.0	25.1	24.3	23.9	23.1	22.5	21.7
ZA	0.4	0.4	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.5	0.5
EU-12	2.7	2.8	2.7	2.8	3.1	3.3	3.3	3.3	3.2	3.2	3.2
EU15	31.4	30.7	30.4	30.0	29.4	28.7	28.3	28.0	27.4	27.2	26.7
EU28	34.3	33.6	33.3	33.0	32.7	32.2	31.8	31.5	30.8	30.6	30.1
World	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Source: Web of Science, queries and calculations by Fraunhofer ISI

Similar observations can be made for the conference proceedings (Table 3). The increase of China's share is even higher, so that in 2013 nearly every third conference publication originates in China. In contrast to that, the USA show steadily decreasing shares, so that they were halved compared to its shares in 2003.

Table 3: Shares of the selected countries and regions in the CPCI within all publications (fractional counting)

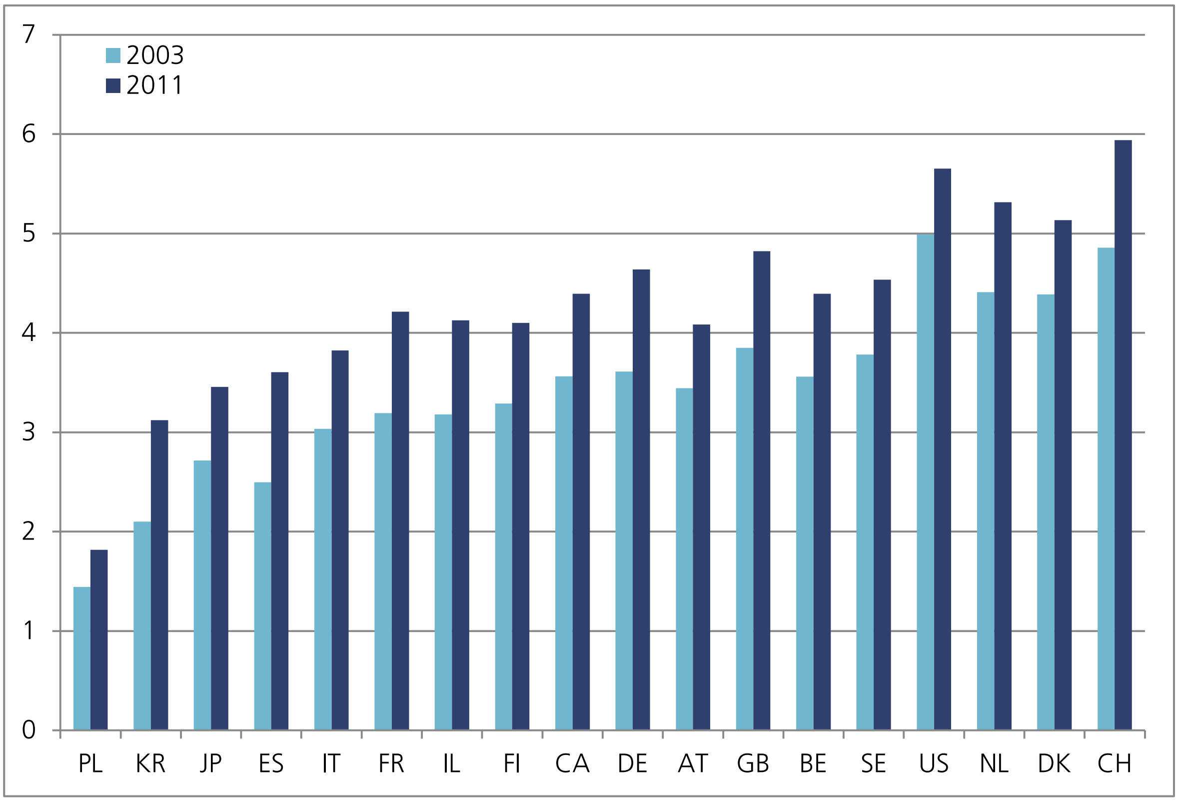
Country/region	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
AT	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.7	0.7	0.6	0.5
BE	0.9	0.9	0.9	0.8	0.8	0.7	0.8	0.7	0.8	0.7	0.6
BR	1.2	1.2	1.2	1.1	1.1	1.1	1.2	1.2	1.2	1.0	1.2
CA	3.2	3.3	3.4	3.3	3.4	3.0	2.7	2.5	2.7	2.3	2.1
CH	1.0	1.1	1.1	0.9	0.8	0.9	0.8	0.8	0.9	0.8	0.7
CN	9.1	9.6	12.4	16.2	18.8	20.0	24.1	26.8	23.5	32.4	31.9
DE	5.8	5.9	5.4	4.9	4.5	4.9	5.0	5.0	5.4	4.6	4.0
DK	0.4	0.4	0.4	0.4	0.3	0.4	0.3	0.4	0.4	0.4	0.3
ES	1.9	1.9	2.1	2.0	2.0	2.3	2.1	2.4	2.7	2.4	2.0
FI	0.8	0.7	0.8	0.7	0.7	0.7	0.6	0.6	0.5	0.5	0.4
FR	3.9	3.9	3.7	3.5	3.5	3.5	3.6	3.4	3.5	3.2	2.7
GB	5.0	4.6	4.3	3.9	3.8	3.8	3.7	3.4	3.6	2.8	2.5
IL	0.7	0.6	0.7	0.5	0.6	0.5	0.5	0.5	0.5	0.4	0.3
IN	1.2	1.1	1.2	1.3	1.9	1.6	2.0	2.2	1.9	2.5	3.7
IT	3.9	3.9	3.8	3.5	3.3	3.3	3.1	3.3	3.2	2.7	2.6
JP	8.9	8.3	7.7	7.3	7.3	7.3	6.1	5.7	5.8	4.7	5.1
KR	2.1	2.4	2.4	2.7	3.5	4.4	3.1	2.3	2.1	2.1	2.1
MX	0.7	0.7	0.8	0.7	0.8	0.6	0.7	0.6	0.7	0.6	0.6
NL	1.6	1.5	1.5	1.4	1.3	1.2	1.2	1.2	1.2	1.0	0.8
PL	1.2	1.2	1.1	1.2	1.1	1.2	1.2	1.0	1.3	1.0	1.1
RU	2.4	1.9	2.0	1.8	1.5	1.2	1.1	0.9	0.9	0.9	1.0
SE	0.9	1.0	0.9	0.9	0.8	0.8	0.7	0.7	0.7	0.7	0.7
USA	28.0	28.3	26.5	24.3	20.9	18.3	16.4	15.4	17.2	14.3	14.0
ZA	0.4	0.2	0.3	0.2	0.2	0.3	0.2	0.3	0.3	0.3	0.3
EU-12	3.8	3.6	3.5	3.9	3.8	4.4	4.7	5.1	6.1	5.3	5.3
EU15	27.5	27.2	26.0	24.8	23.4	24.2	23.7	23.5	24.8	21.3	18.9
EU28	31.6	31.1	29.8	29.1	27.4	28.9	28.7	28.8	31.1	26.7	24.3
World	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Source: Web of Science, queries and calculations by Fraunhofer ISI.

2.3 Number of citations

Citations are used as an indication of the quality of publications, but at least reflect the visibility of publications in the science system. The observed citation rate shows the impact of a publication in the scientific community. It can hint at the role that a country plays in the scientific discourse. However, the citation rates can also be influenced by the fields in which the respective publications are published. Because of that, the following sections deal with bibliometric indicators that take into account field differences. Such indicators use a normalized or relative citation rate or similar derivations to show the relative quality of a country's publication output. In contrast to that, the gross citation rate that is presented in this section reflects the dissemination and utility of a country's work. A so called citation-window is used that allows a comparability of different cohorts of publications. We use a three year citation window that takes into account all the citations in the years of publication and the two subsequent years. This provides equal chances to every publication of being cited. This restricts the observation period to publications of the years up to 2011 as their citations in publications of the years 2012 and 2013 need to be taken into account. Figure 5 shows the countries and their citation rates in the years 2003 and 2011. The countries are ordered according to their citation rate in 2011. Switzerland leads this ranking, followed by the USA and the Netherlands.

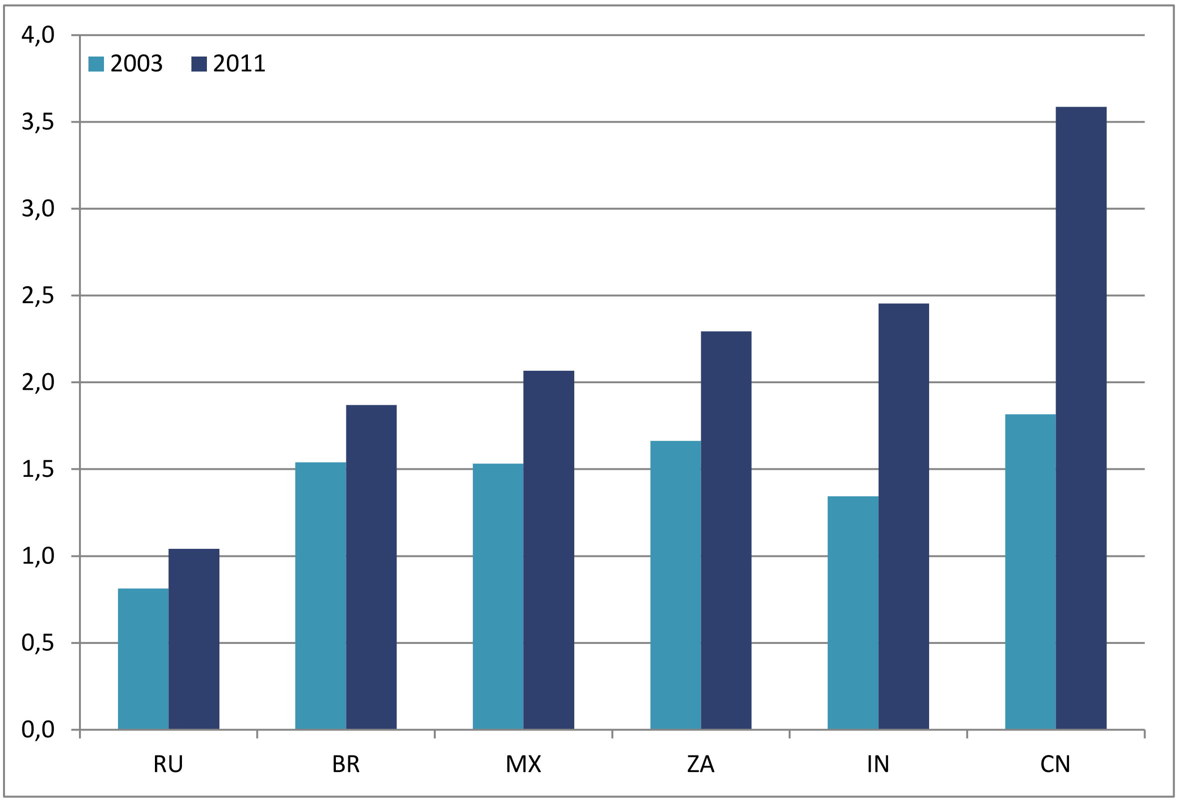
Figure 5: Citation numbers of the selected industrialized countries in the SCIE and the SSCI according to fractional counting (3 year window, no self citations)



Source: Web of Science, queries and calculations by Fraunhofer ISI

Figure 6 analogously shows the citation rates for the BRICS countries. Again, China differs from the other countries with a much higher value. This was yet not the case in 2003, when most of the BRICS countries had an average citation rate between 1.5 and 2.0. Russia is one of the countries with the smallest change between these years and still barely reaches a citation rate of 1.0. Most industrialized countries now have values between 3 and 5, so China caught up also in terms of citations. Poland is by far the country with the lowest citation rate in our set. However, it has nearly doubled its publication output in the observation period, but was not yet able to catch up in terms of quality of the respective publications.

Figure 6: Citation numbers of the BRICS countries in the SCIE and the SSCI according to fractional counting (3 year window, no self citations)



Source: Web of Science, queries and calculations by Fraunhofer ISI

2.4 Journal-specific Scientific Regard (SR) and International Alignment (IA)

The Scientific Regard (SR) and the International Alignment (IA) put the observed citation rate in perspective with the reputation of the publishing journals. While the IA shows whether a country publishes in more or less cited journals (compared with the world average), the SR relates the citation rate of a publication to the average citation rate in that each journal and indexes the average for all publications.

Table 4 shows the SR values for the countries for the years 2003 to 2011. Overall, the value for Germany has been decreasing in the past century. In 2010 it seemed as if this decreasing trend had been stopped, but the value in the most recent year is close to that in 2009, which was Germany's low. Similar observations can be made for the EU28 countries as a whole. In particular, this holds for Austria, France, Finland, Italy and Sweden. Of the EU28 countries, only Spain, Denmark, Poland and the Netherlands show increasing values in 2011 and in all these cases a decline preceded.

Table 4: Index of the journal-specific Scientific Regard (SR) for the selected countries and regions in the SCIE and the SSCIE according to fractional counting

Country/region	2003	2004	2005	2006	2007	2008	2009	2010	2011
AT	7	8	3	5	7	1	4	3	1
BE	6	0	4	5	5	6	5	4	4
BR	-23	-23	-20	-19	-14	-12	-11	-10	-11
CA	3	3	1	1	1	1	1	1	0
CH	14	13	17	14	12	14	12	11	11
CN	-3	1	1	1	3	4	3	3	5
DE	9	8	9	8	7	6	4	6	5
DK	15	12	12	11	9	13	9	9	10
ES	-12	-8	-8	-6	-5	-7	-5	-7	-5
FI	3	1	0	2	-1	-2	0	0	0
FR	1	-1	0	0	0	-1	0	-1	-2
GB	7	8	7	6	5	6	6	7	6
IL	-13	-11	-12	-12	-10	-14	-13	-14	-12
IN	-17	-18	-14	-10	-10	-7	-7	-6	-5
IT	-7	-7	-6	-7	-5	-5	-3	-3	-4
JP	-12	-11	-12	-11	-12	-12	-13	-14	-14
KR	-8	-6	-6	-6	-7	-6	-6	-7	-8
MX	-22	-26	-26	-24	-20	-23	-20	-20	-19
NL	15	10	7	8	6	8	7	7	8
PL	-20	-22	-21	-20	-18	-20	-16	-17	-13
RU	-8	-8	-10	-10	-7	-8	-11	-8	-9
SE	6	6	3	3	2	1	3	2	2
USA	10	10	10	9	8	8	7	7	7
ZA	-9	-12	-9	-5	-7	-3	-5	-2	-7
EU-12	-15	-17	-16	-14	-11	-12	-10	-9	-8
EU15	4	3	3	3	2	2	2	2	2
EU28	2	1	1	1	1	1	1	1	1
World	0	0	0	0	0	0	0	0	0

Source: Web of Science, queries and calculations by Fraunhofer ISI

Switzerland started with a value of 14 in the observation period, which was one of the highest values at that time in the set of observed countries. Even though this value has been decreasing as well, Switzerland has the highest SR value of all countries in 2013, which means that Swiss authors get much higher numbers of citations than the average in the particular journals they are able to place their publications.

A general trend of decreasing SR values can be observed. Apart from the few EU28 countries named above, only China, Israel, India, Japan and Mexico have increased their SR values at the current edge.

Table 5: Index of the International Alignment (IA) for the selected countries and regions in the SCIE and the SSCI according to fractional counting

Country/region	2003	2004	2005	2006	2007	2008	2009	2010	2011
AT	-4	-5	-2	0	-1	2	0	-2	1
BE	-6	-2	-1	1	2	2	3	6	3
BR	-46	-47	-47	-44	-43	-50	-55	-55	-55
CA	3	4	3	4	6	6	6	6	5
CH	24	22	24	23	22	23	24	25	25
CN	-59	-55	-54	-52	-47	-40	-36	-32	-28
DE	-2	0	0	3	4	5	7	8	9
DK	8	12	12	11	13	13	15	15	15
ES	-20	-18	-14	-13	-11	-10	-12	-11	-8
FI	-1	-1	-2	0	-1	0	1	0	1
FR	-6	-5	-5	-3	-2	1	1	2	3
GB	6	7	7	9	8	10	12	11	11
IL	3	4	5	7	6	7	9	7	10
IN	-63	-62	-57	-55	-49	-50	-52	-47	-48
IT	-3	-4	-2	0	0	-1	0	-2	-1
JP	-13	-12	-10	-11	-11	-10	-6	-5	-6
KR	-40	-39	-41	-37	-37	-29	-29	-29	-26
MX	-46	-46	-43	-41	-39	-43	-44	-41	-43
NL	14	17	18	19	20	21	22	21	22
PL	-58	-55	-57	-51	-51	-55	-60	-57	-56
RU	-85	-84	-83	-83	-82	-83	-83	-84	-83
SE	5	6	6	8	8	11	9	10	10
USA	28	28	27	26	26	27	26	26	24
ZA	-54	-50	-47	-44	-44	-45	-44	-45	-40
EU-12	-56	-53	-53	-50	-48	-54	-57	-55	-54
EU15	-2	0	0	2	2	4	4	5	5
EU28	-5	-4	-4	-2	-2	-1	-1	-1	0
World	0	0	0	0	0	0	0	0	0

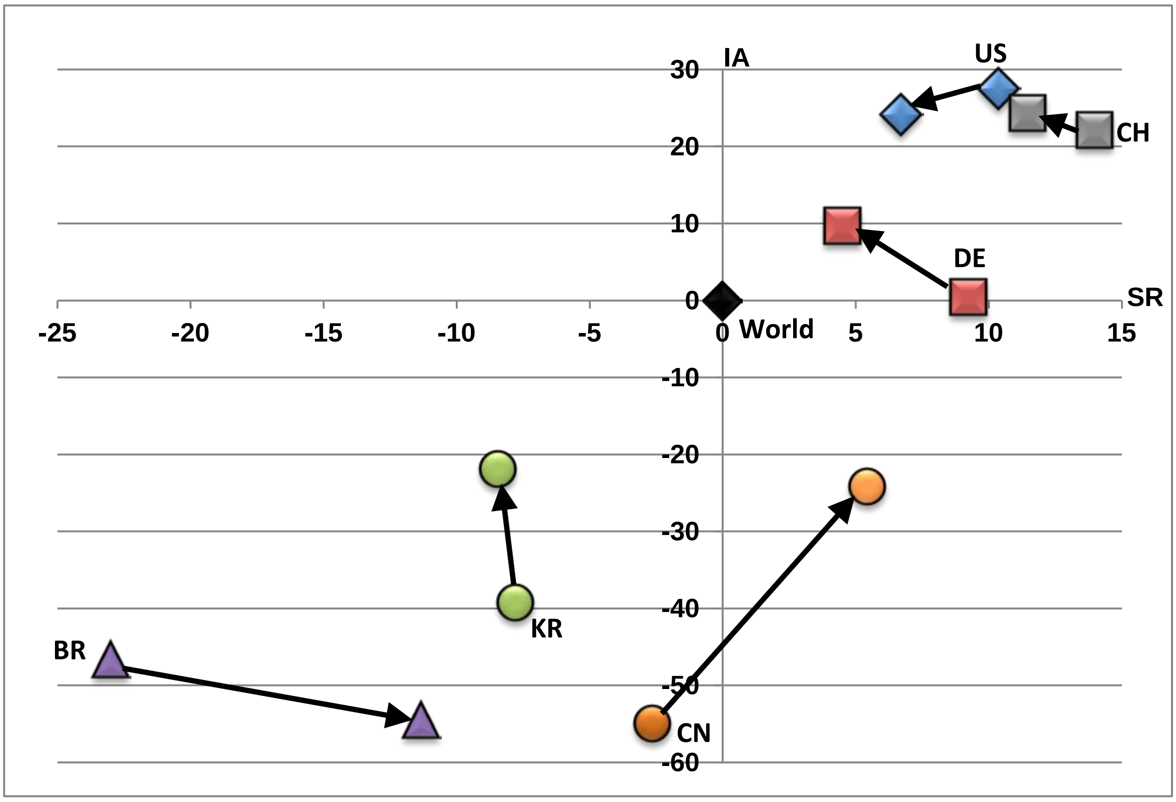
Source: Web of Science, queries and calculations by Fraunhofer ISI

As a supplement to the SR values, Table 5 shows the IA values for the selected countries and regions. In general, the IA values are more dispersed than the SR values, i.e. there are countries with relatively low values (e.g. Russia, Poland and Brazil), but also with relatively high values (e.g. Switzerland, the USA and the Netherlands). Such a high disparity could not be observed for the SR values.

It is interesting to note that several of the countries which increased their SR value in the observation period were also able to increase their IA value – at least in a short-term perspective. China, India, Japan and Israel all increased their IA value from 2010 to 2011, but only the latter does so with a positive initial value. Denmark, Mexico and Poland show no clear trend in the observation period. Germany however demonstrates a clear increase and thus manages to publish in journals with a higher international reputation. The SR values above indicate that the respective reference values in these journals are still (but less) surpassed.

Figure 7 shows the SR and the IA in comparison for six countries. The initial situation in 2003 as well its end situation in 2011 are depicted for each country and connected via an arrow to show the gross development. Both indicators have a value of 0 for the world, which is used as a reference point for the values of the single countries.

Figure 7: Index of the journal-specific Scientific Regard (SR) and the International Alignment (IA) for six selected countries in 2003 and 2011 in the SCIE and the SSCI according to fractional counting



Source: Web of Science, queries and calculations by Fraunhofer ISI

Germany for instance starts with a relatively high SR in 2003, which diminishes during the observation period. At the same time, the IA increases from a value close to 0 to approximately 4.5. Thus, Germany now targets journals with a higher international reputation. It increases its visibility, aiming for journals with a higher international standing. In turn, its relative citation rate in comparison with other articles in its journals decreases. Therefore, the citation rates in these journals are overall higher and Germany cannot maintain its relatively high level in the SR.

Further countries, i.e. Switzerland, the USA, China, South Korea and Brazil were included in the graph to give some other reference points for the German development. Switzerland and the USA, which are the only other countries in the set that also have a positive SR as well as IA value, show similar but not so pronounced developments as Germany; in both cases as well, the SR value decreases. However, while Switzerland also manages to increase its IA value, the USA also lose ground in this indicator. Therefore, they do not only show diminishing shares in the worldwide publication output (cf. Table 2) but also decrease their relative visibility of the publications.

China manages an increase in both indicators. In that way, it has an SR value even higher than that of Germany. The IA value increases as well, so that not only journals with a higher reputation are used, but also higher relative citation rates are achieved. However, the citation rates are still far below the average in the respective journals.

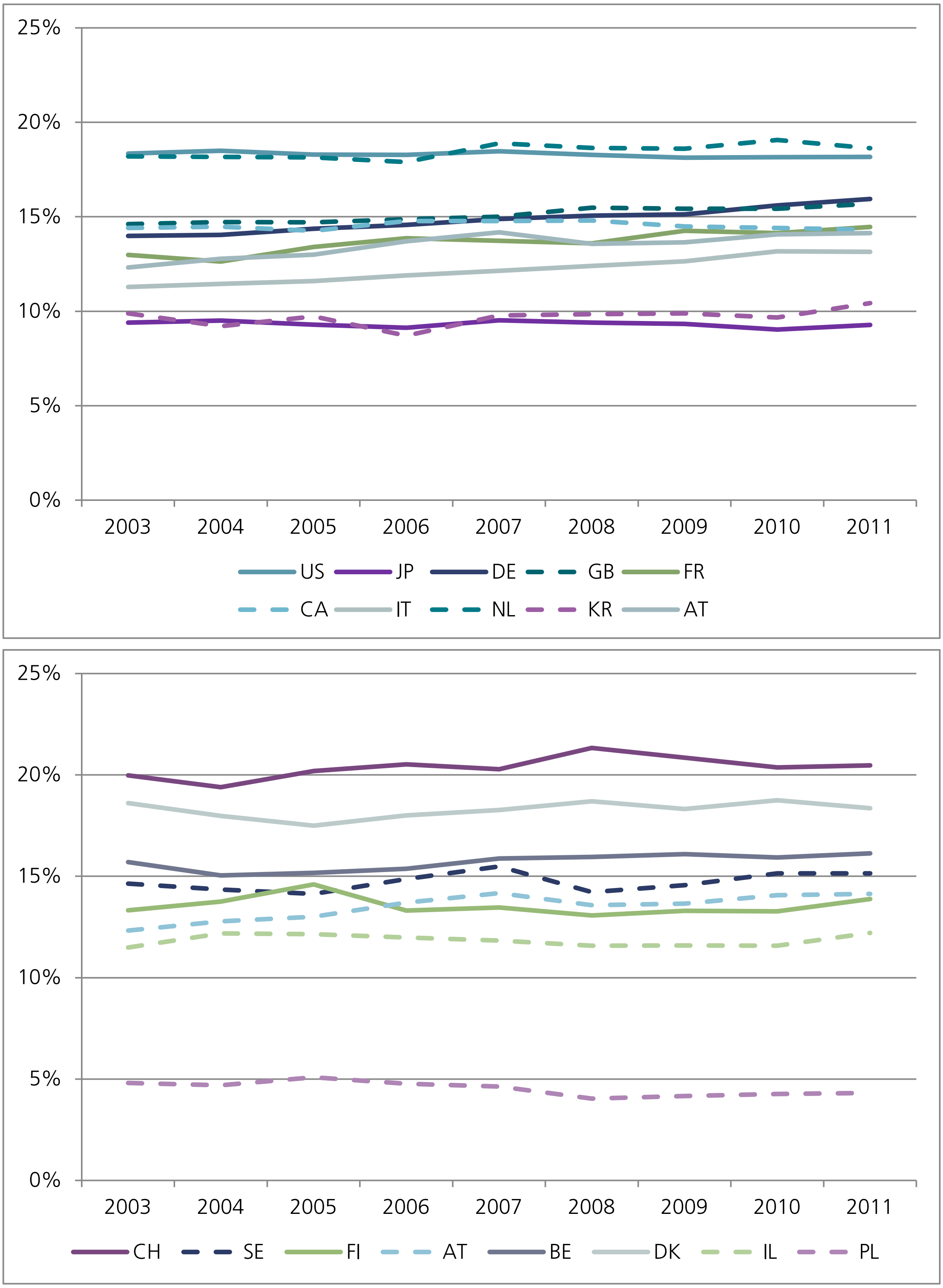
2.5 Share in top cited publications (Excellence Rate)

In this section, the share of publications that belong to the worldwide top cited publications is analyzed. For that purpose, in a first step the 10% top cited publications per field are selected (to account for varying citation rates in the scientific fields). For each country, the number of publications belonging to the top 10% in at least one field is calculated and set in relation to the total number of its publications. In that way, its share of highly cited publications is derived, that is also denoted as Excellence Rate (Bornmann et al., 2012; Waltmann and Schreiber, 2013).

Figure 8 shows the Excellence Rate for the industrialized countries. Only few of them do not reach the reference value of 10%. The Polish publications are only in merely 4 to 5% of the cases in the set of highly cited publications. Japan and South Korea are close to the 10% mark, the latter even tops it in the most recent year.

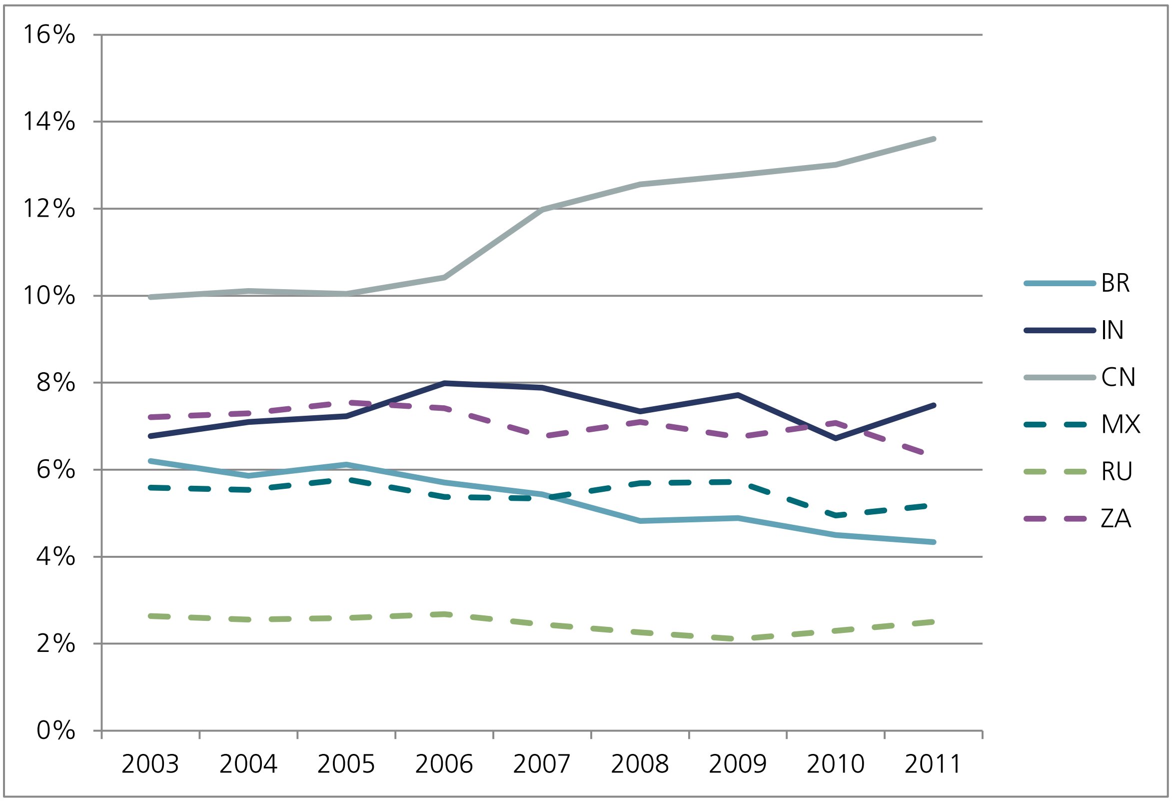
In comparison, the BRICS countries have – again with the exception of China - relatively low values (Figure 9). China achieves values far higher than the other BRICS countries and shows a steep increase after 2006. This again stresses the impressive trend that China managed not only to increase its absolute numbers of publications considerably, but at the same time also increased its quality and visibility of the publications. This, however, needs to be balanced by the fact that national preferences exist – not only in China, but also in the USA, for example – to read and cite publications by researchers from the same country (see e.g. Gondal, 2011). In consequence of the strong increase in publications by Chinese authors, the increase of citations is not that astonishing, given the large absolute number.

Figure 8: Excellence Rate for the industrialized countries according to fractional counting for the years 2003 to 2011



Source: Web of Science, queries and calculations by Fraunhofer ISI

Figure 9: Excellence Rate for the BRICS countries according to fractional counting for the years 2003 to 2011

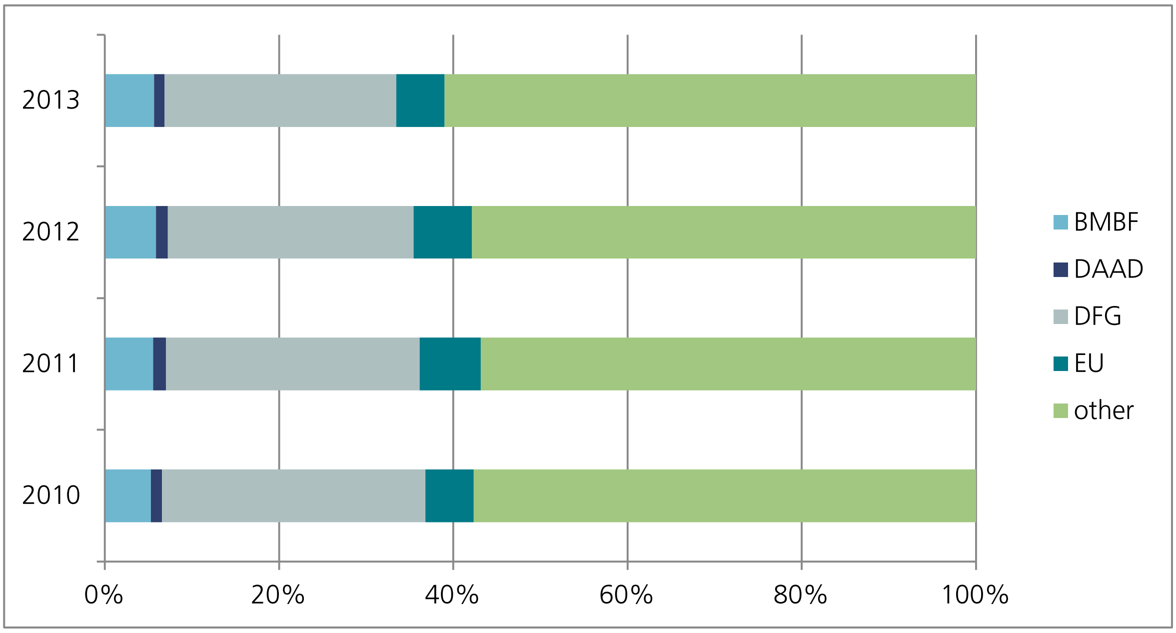


Source: Web of Science, queries and calculations by Fraunhofer ISI

2.6 Acknowledgements

Funding is an important factor also in science. Differences between institutional and project funding can be seen. Changes in the structures of funding in many countries can be detected, based on new or different policy aims. Especially a shift towards a higher application orientation as well as emerging fields is among these policy trends. Analysing the funding acknowledgements, the financial background of the publications can be unveiled. Concerning the methodology, the data requires manual cleaning that currently only allows a restricted analysis. As is the case with the organizations of the authors, the funding organizations are covered as in the publications themselves, i.e. the names are not unified but represented as in the respective publication. Therefore, the data shown in this section is based on thesaurus with manual cleaning of the funding organizations in focus, i.e. spelling variations of BMBF, DAAD, DFG and the EU were queried to aggregate the data.

Figure 10: Shares of German publications mentioning support by the BMBF, DAAD, DFG and the EU according to fractional counting (for Germany)

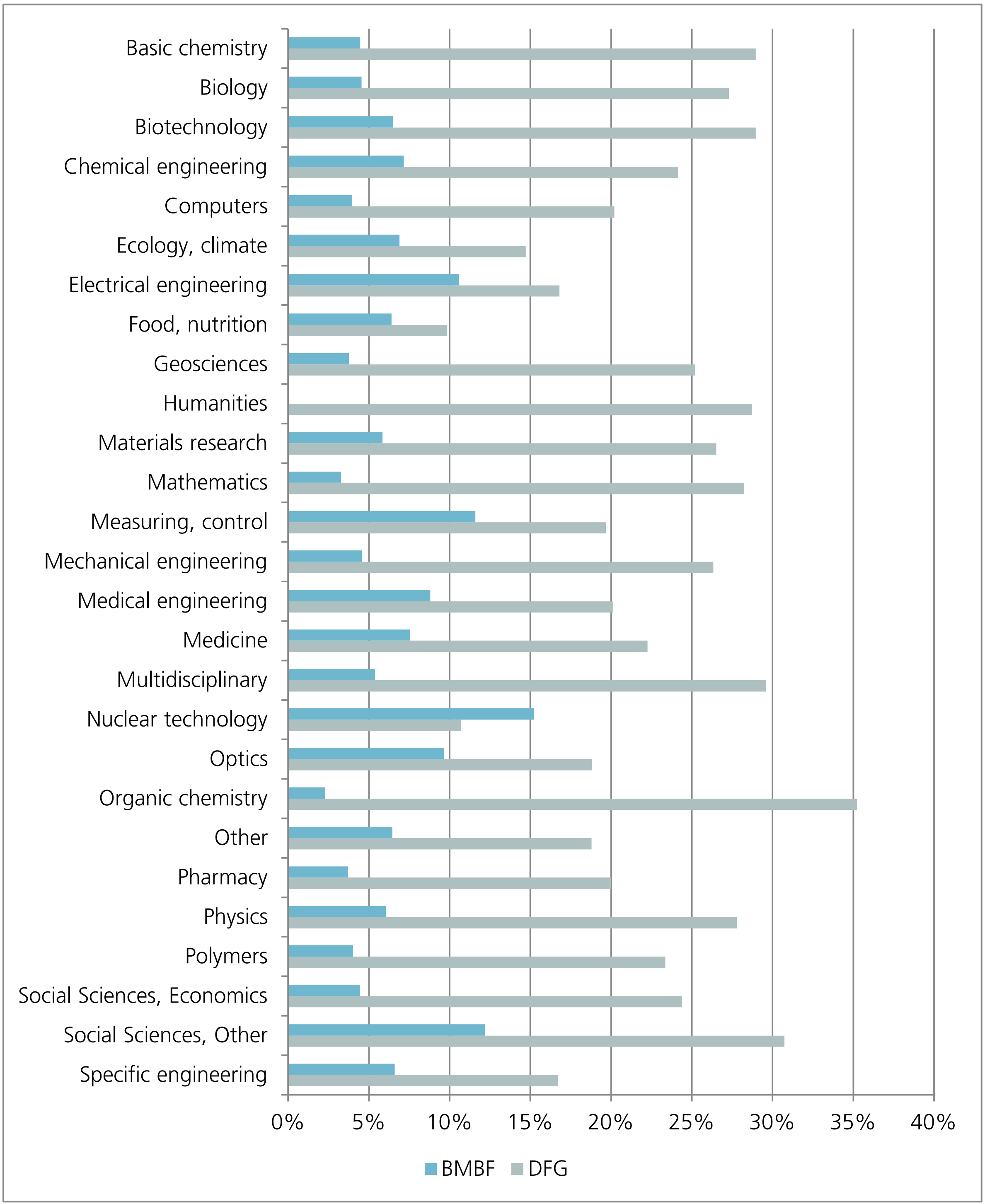


Source: Web of Science, queries and calculations by Fraunhofer ISI

Figure 10 shows the shares of these funding organizations in the German publications. Only few changes can be perceived over time: The importance of other, alternative funding organizations has increased slightly, but at this point in time no specific characteristics of these organizations can be derived. The only other change can be noted for the DFG, which shows a diminishing share in the observation period (starting with 30% in 2010 to 27% in 2013).

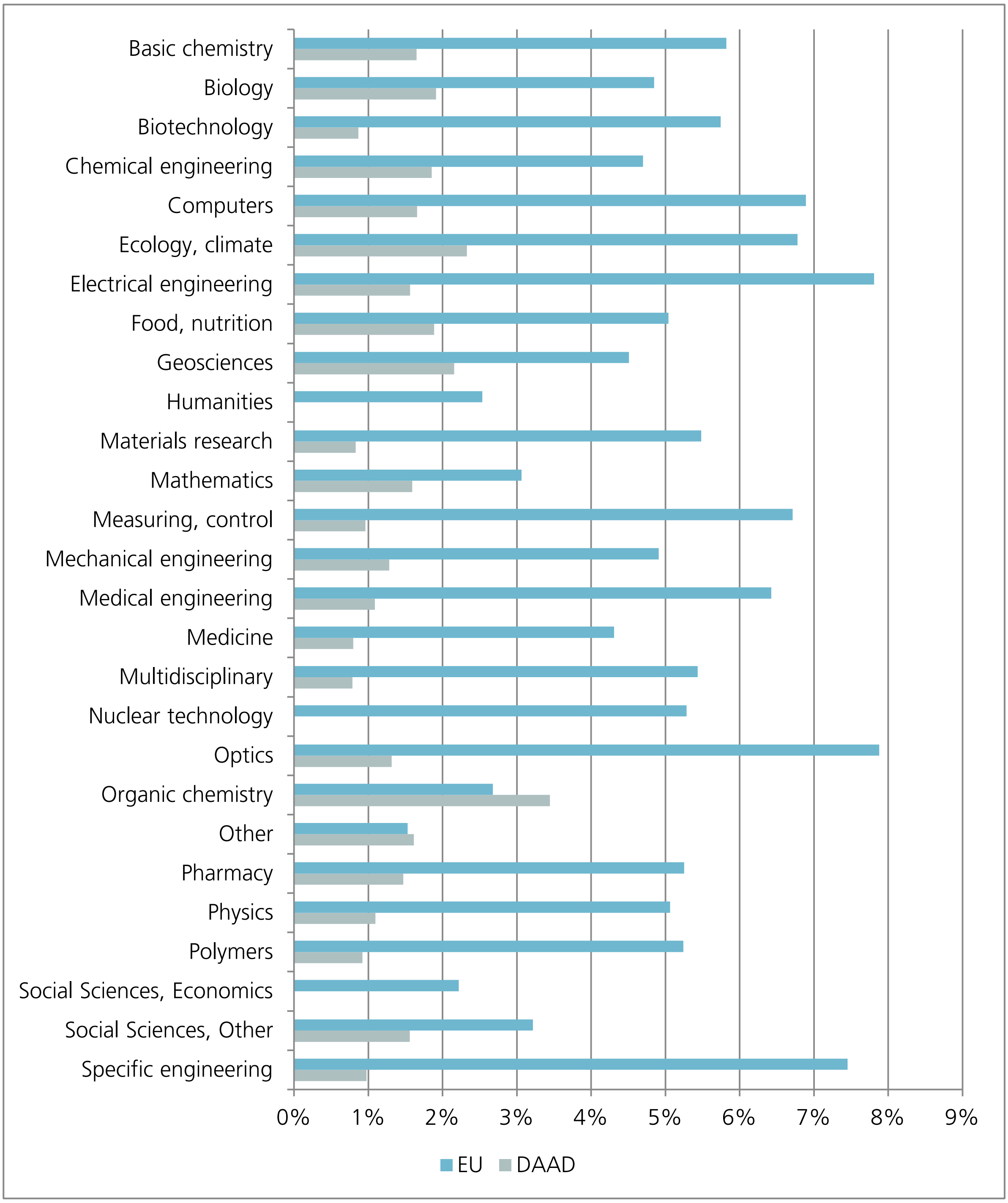
In Figure 11 and Figure 12, the four funding organizations are compared in terms of the respective scientific fields to shed light on the context of the funding. In general, the DFG is higher represented in all fields than the BMBF, except for Nuclear technologies. Humanities are only funded by the DFG and the EU. Exceptionally high presence can also be found in the funding of Organic chemistry: Here funding is supported in 35% of the cases by the DFG. In 3% of the publications, the DAAD is mentioned, which appears in the other fields mostly only in 1% to 2% of the cases.

Figure 11: Shares of German publications in the 26 fields mentioning support by the BMBF and DFG according to fractional counting (for Germany)



Source: Web of Science, queries and calculations by Fraunhofer ISI

Figure 12: Shares of German publications in the 26 fields mentioning support by the EU and DAAD according to fractional counting (for Germany)



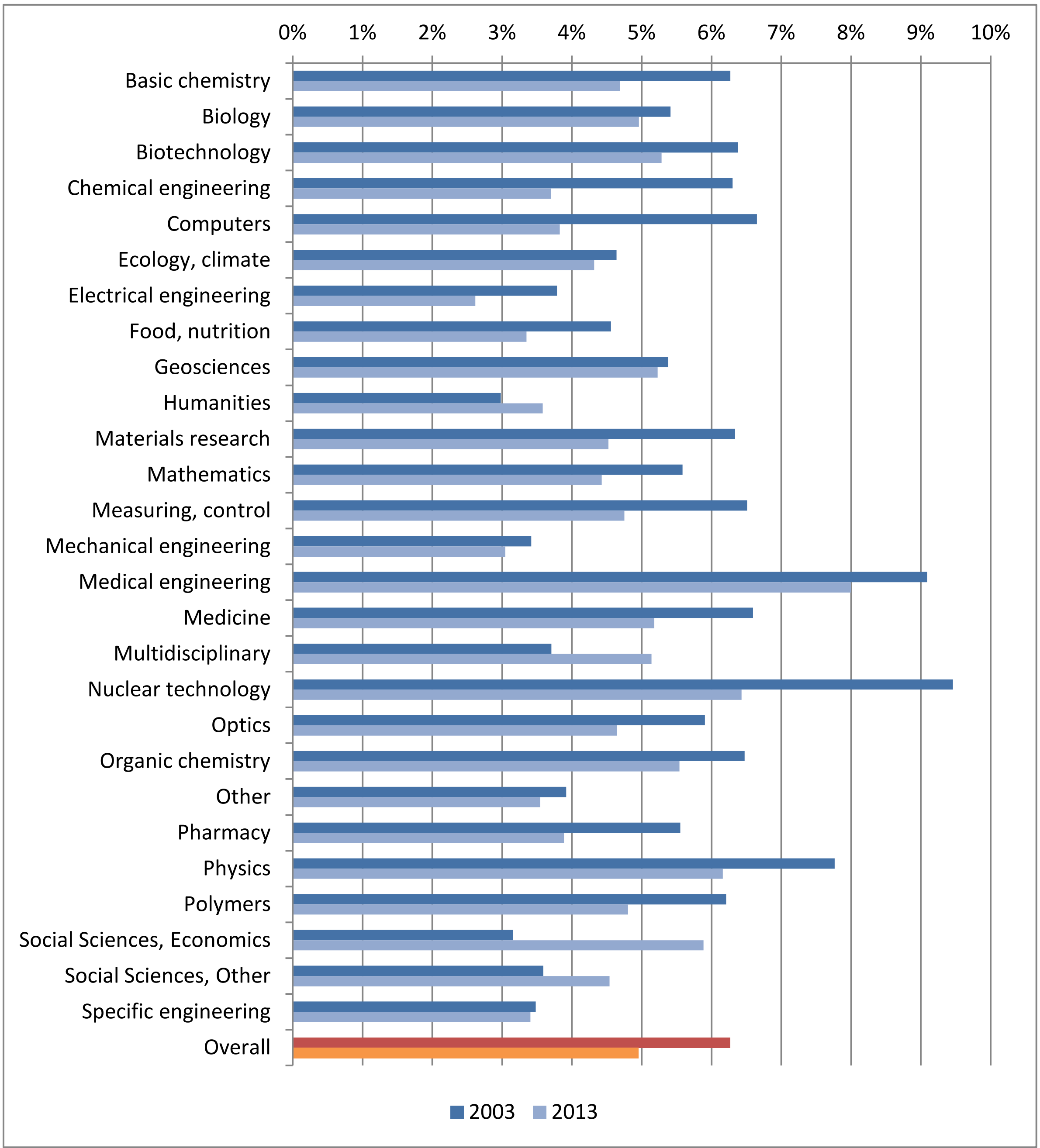
Source: Web of Science, queries and calculations by Fraunhofer ISI

3 Part II: Structure analyses

In this part, the publications and their respective indicators are classified according to the 26 scientific fields. In that way, the scientific profile of the countries can be shown over time and in comparison to each other. Germany is presented with China and the USA as comparative values.

3.1 Shares of publications in the fields

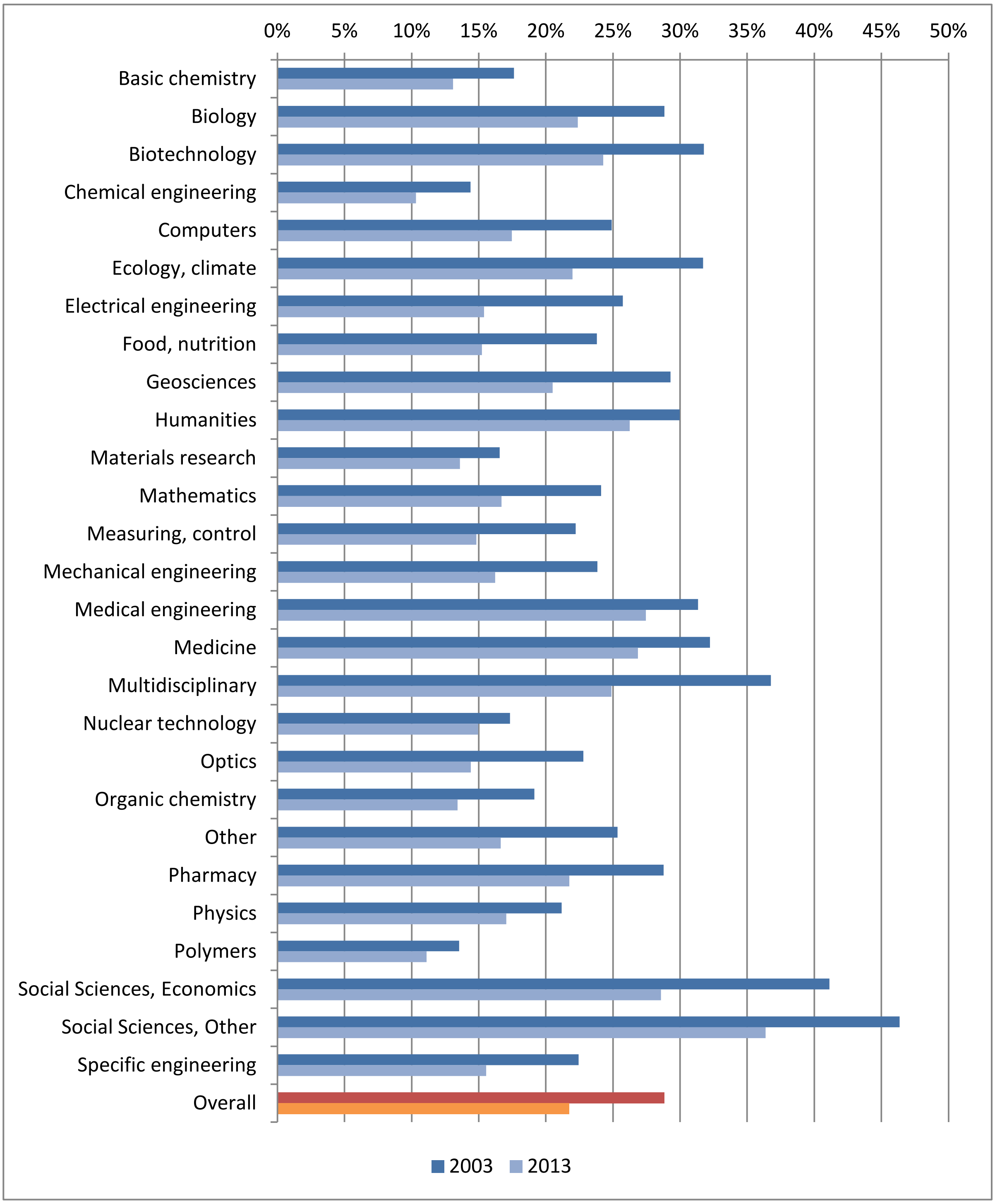
Figure 13: Germany's worldwide shares by fields for publications in the SCIE and the SSCI according to fractional counting for the years 2003 and 2013



Source: Web of Science, queries and calculations by Fraunhofer ISI

In a first analysis, the shares of Germany, the USA and China in the worldwide publication output in the 26 fields are depicted (Figure 13, Figure 14 and Figure 15). Germany has in comparison to the USA a smaller share of total publication output. The main fields, in which it is represented internationally, are Medical engineering, Nuclear technology and Physics.

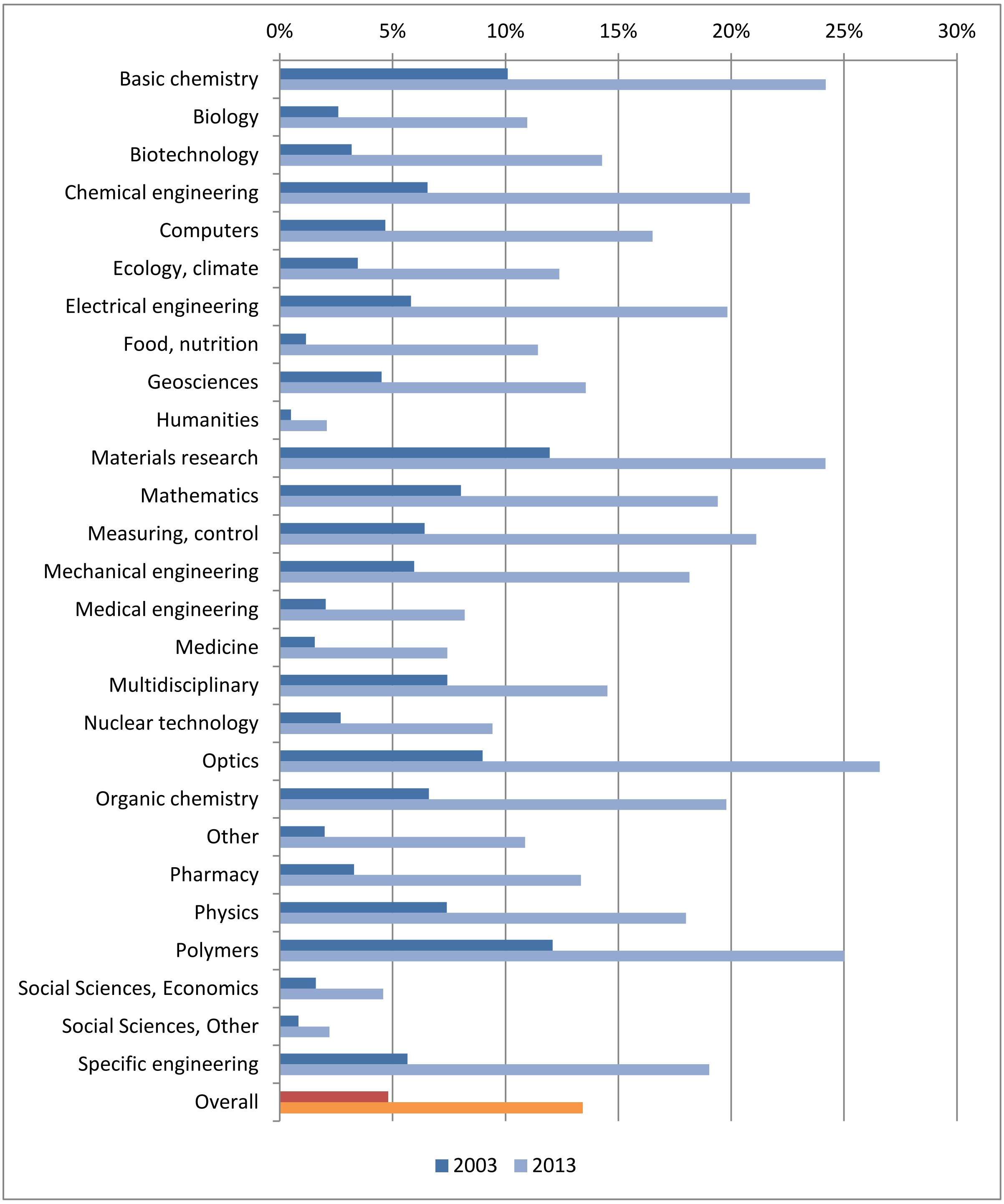
Figure 14: USA's worldwide shares by fields for publications in the SCIE and the SSCI according to fractional counting for the years 2003 and 2013



Source: Web of Science, queries and calculations by Fraunhofer ISI

The USA dominate the Social Sciences with a worldwide share of approximately 30%. In 2003, even more than 40% of the respective publications came from the USA. Both the USA as well as Germany have lost shares in all fields, while China has increased the respective values with no exception. Especially Optics, Chemical fields and Materials research now show relatively high shares of China.

Figure 15: China's worldwide shares by fields for publications in the SCIE and the SSCI according to fractional counting for the years 2003 and 2013

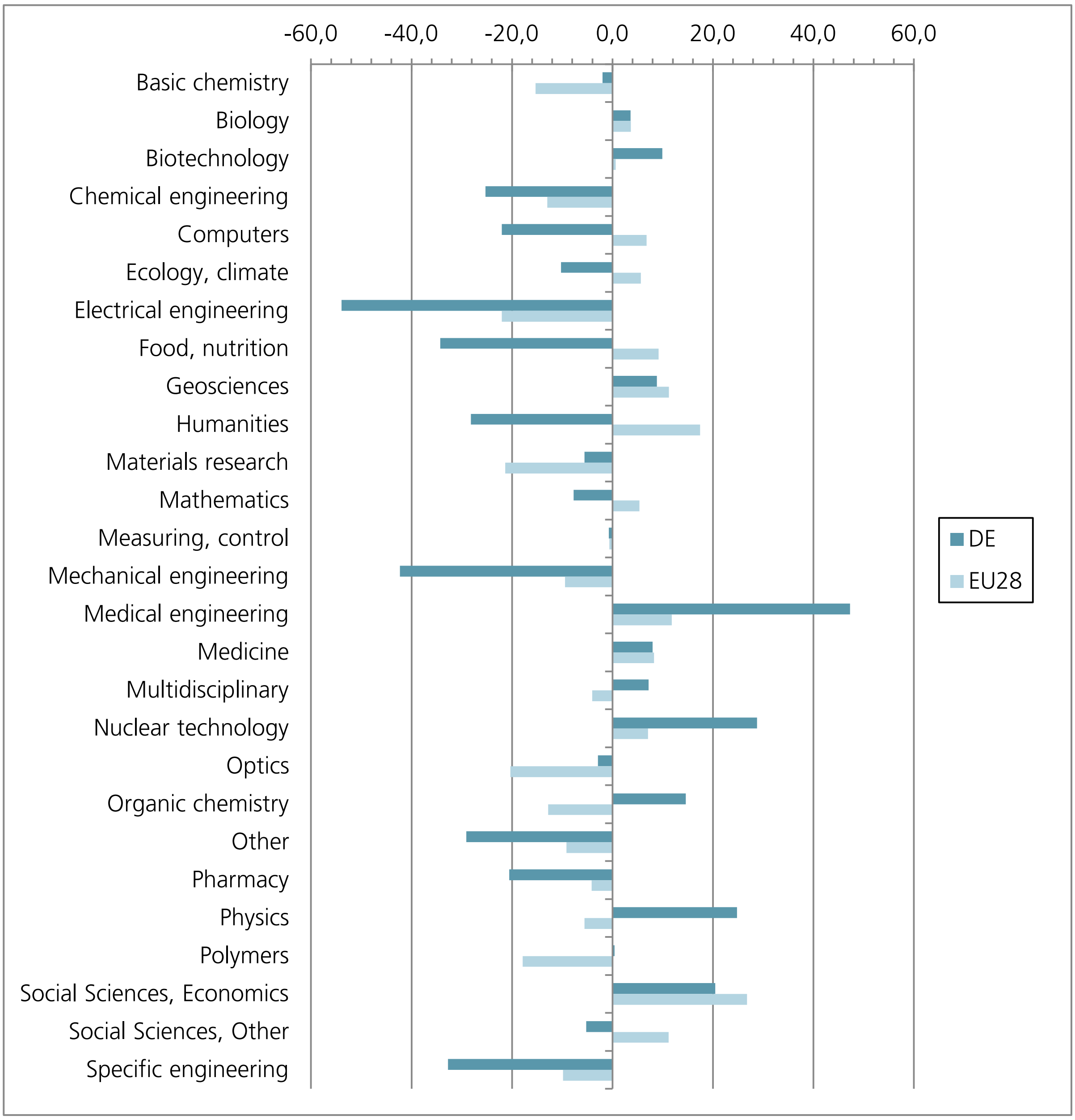


Source: Web of Science, queries and calculations by Fraunhofer ISI

3.2 Revealed Literature Advantage

In contrast to the foregoing analysis, the Revealed Literature Advantage shows the share of a field in a country in comparison to the share of that field in the worldwide publications. The scale is normalized to values between -100 and +100.

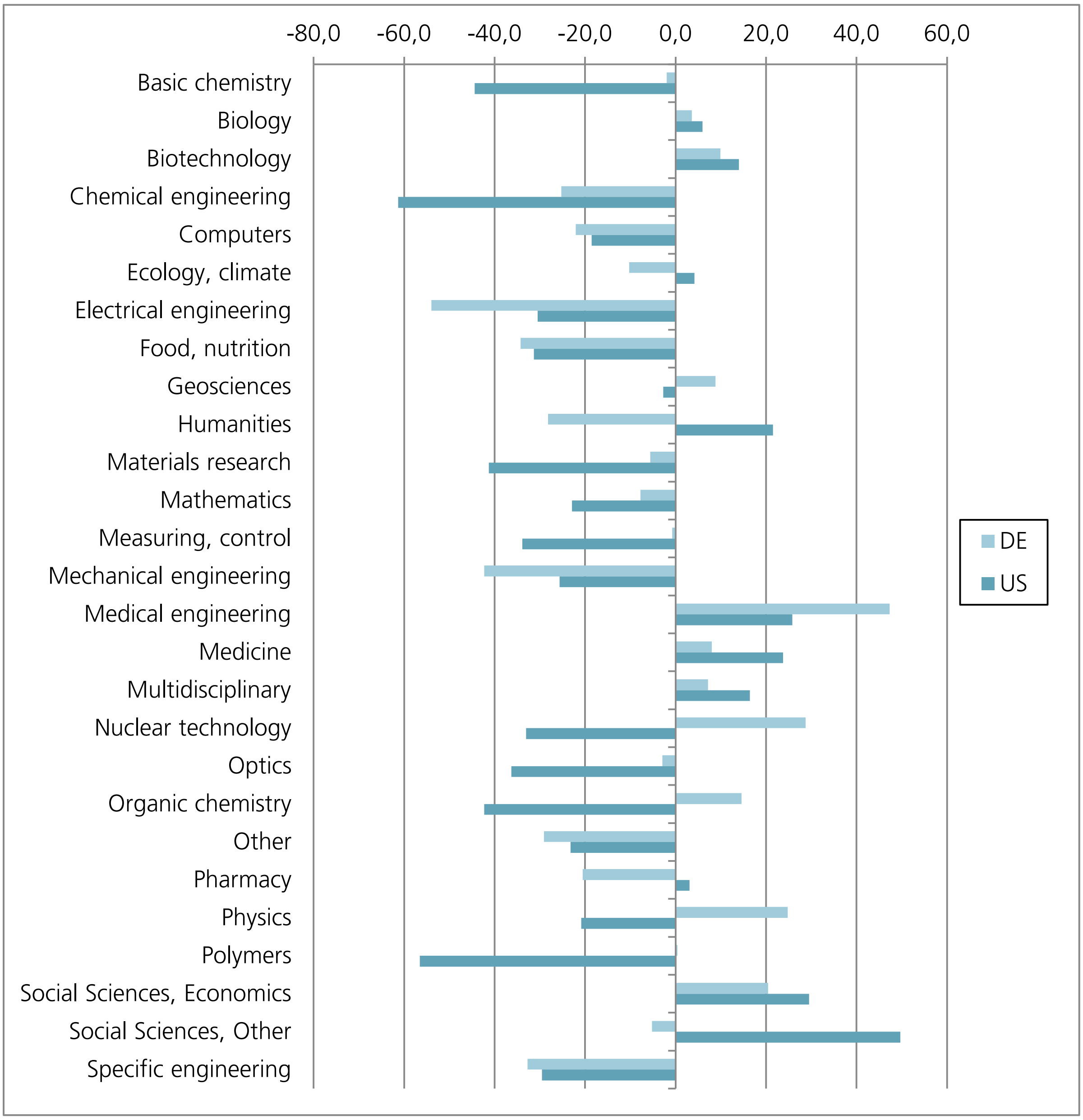
Figure 16: RLA of Germany and the EU28 for publications in the SCIE and the SSCI according to fractional counting in the year 2013



Source: Web of Science, queries and calculations by Fraunhofer ISI

In Figure 16, the RLA values of Germany are compared to that of the EU28 countries. Germany again excels in fields like Medical Engineering and Physics, but shows relative low specialization values in Electrical and Mechanical engineering. Some opposing trends for Germany and the EU28 countries can be found; for instance, fields like Food, nutrition or the Humanities are relatively less represented in German publications than in those of the EU28.

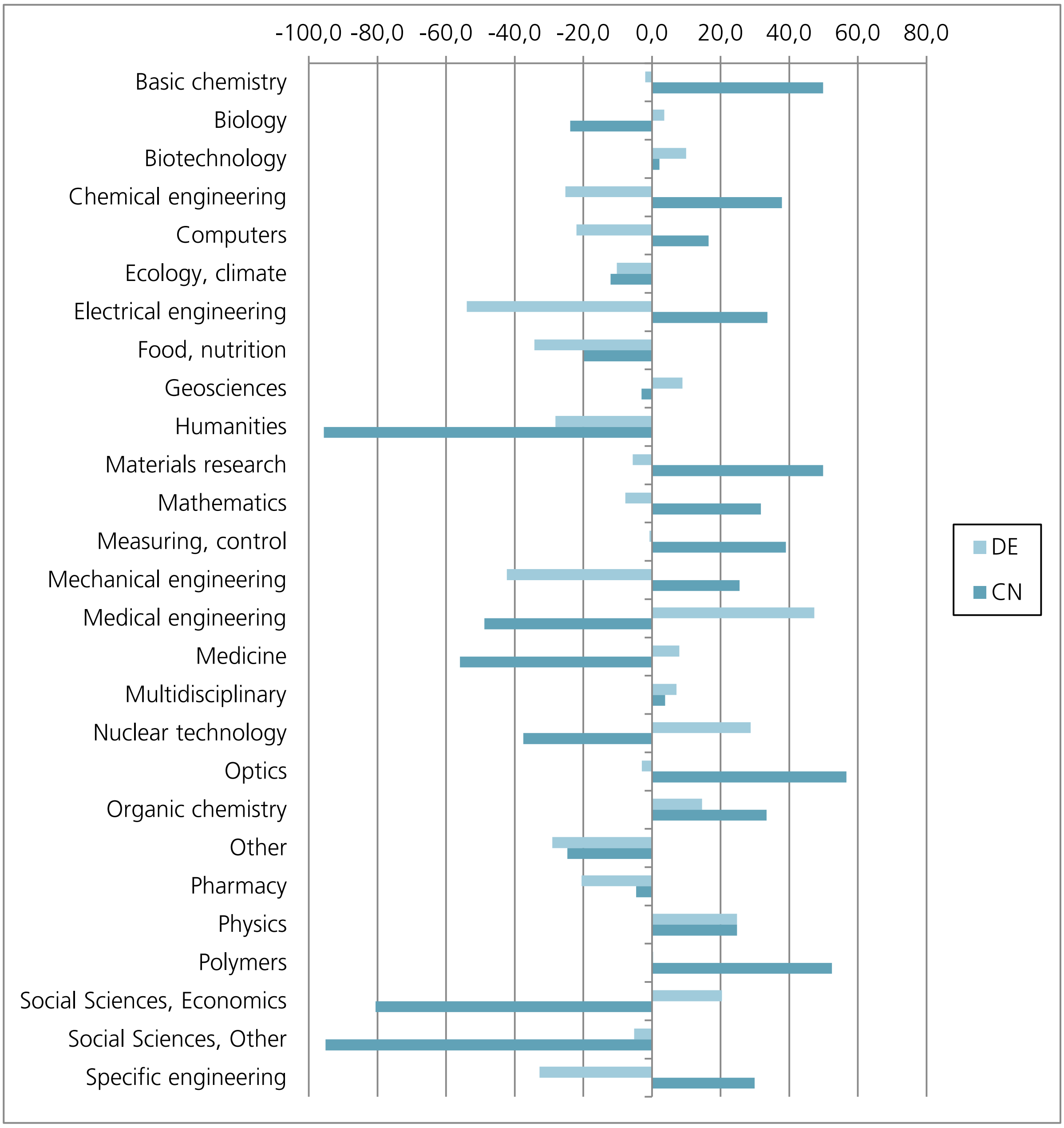
Figure 17: RLA of Germany and the USA for publications in the SCIE and the SSCI according to fractional counting in the year 2013



Source: Web of Science, queries and calculations by Fraunhofer ISI

In comparison to the USA (Figure 17), Germany's focus is set more on technical fields. In contrast to that, China's scientific profile is most prolific in fields like Chemistry and Engineering (Figure 18). In the direct comparison, the profile of the USA is very similar to that of Germany with only few exceptional fields, while the one of China is very different.

Figure 18: RLA of Germany and China for publications in the SCIE and the SSCI according to fractional counting in the year 2013

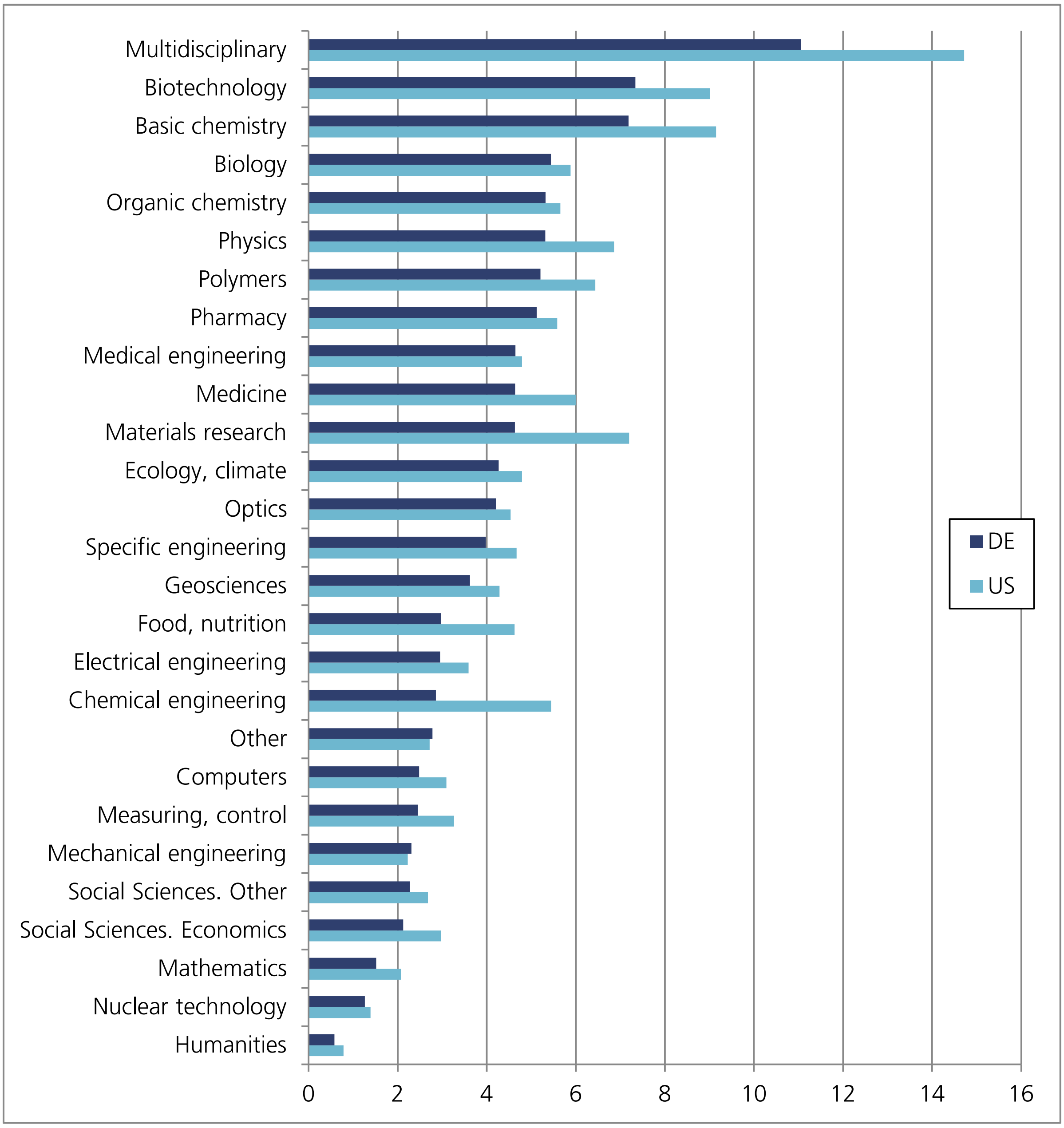


Source: Web of Science, queries and calculations by Fraunhofer ISI

3.3 Citation-based indicators

In general, the citation rates of the USA are higher than that of Germany in all fields. In the multidisciplinary fields, as well as Biotechnology, Basic chemistry, Medicine, Materials research, Food, nutrition and Chemical engineering, the difference is far higher than in the other fields. The only field for which Germany has a higher citation rate is “other”, which collects everything that cannot be clearly assigned to one of the other categories.

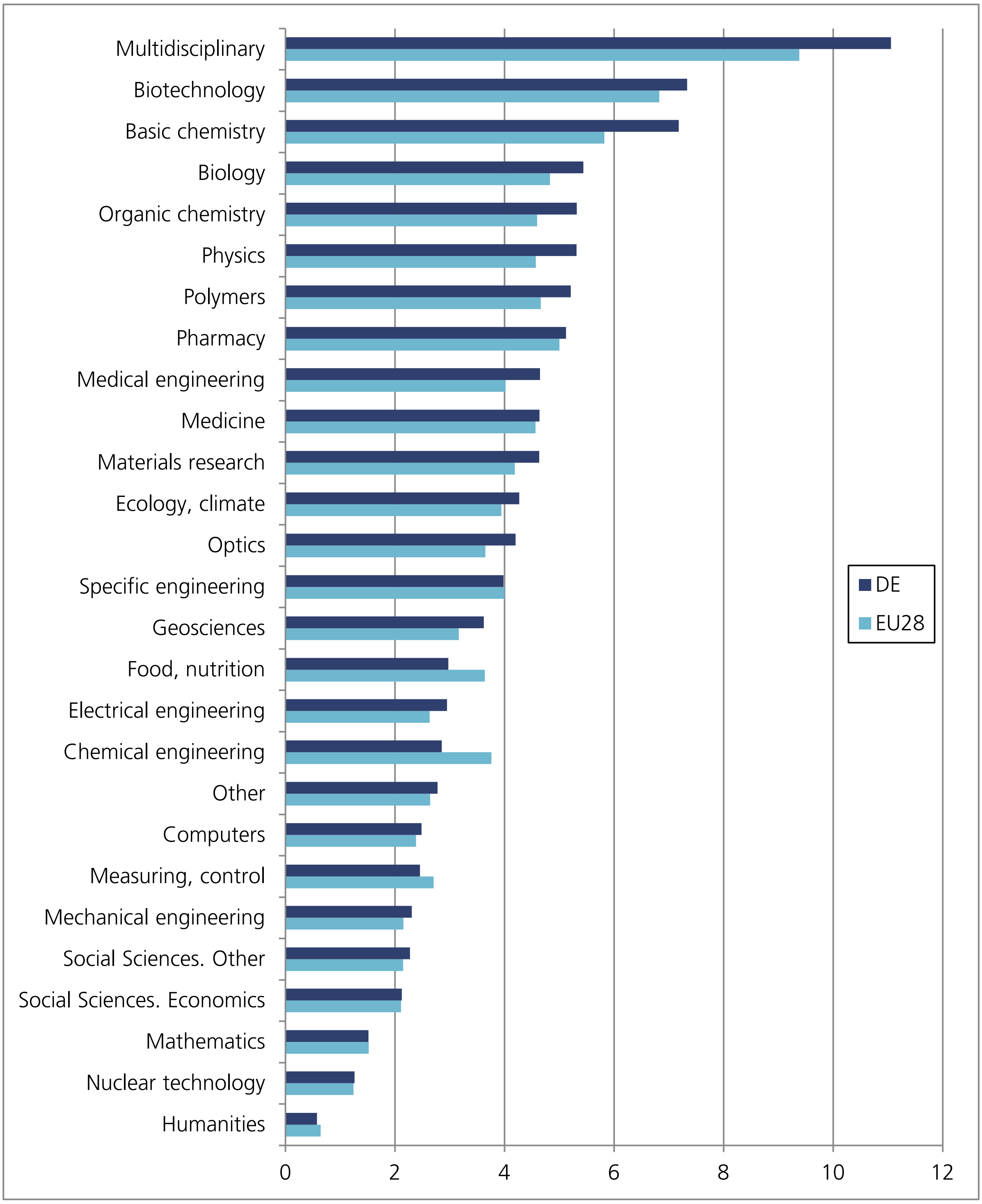
Figure 19: Citation rate per field of Germany and the USA for publications in the SCIE and the SSCI in 2011 according to fractional counting



Source: Web of Science, queries and calculations by Fraunhofer ISI

Yet, Germany has in comparison with the EU28 countries in general a relatively high citation rate in the fields. Here, the only exceptions are Food, nutrition, Chemical engineering and Humanities (Figure 20), which are all fields in which Germany is not specialized (cf. Figure 16). The comparison of Germany's citation rate with that of the worldwide publications is similar to that with the EU28 countries.

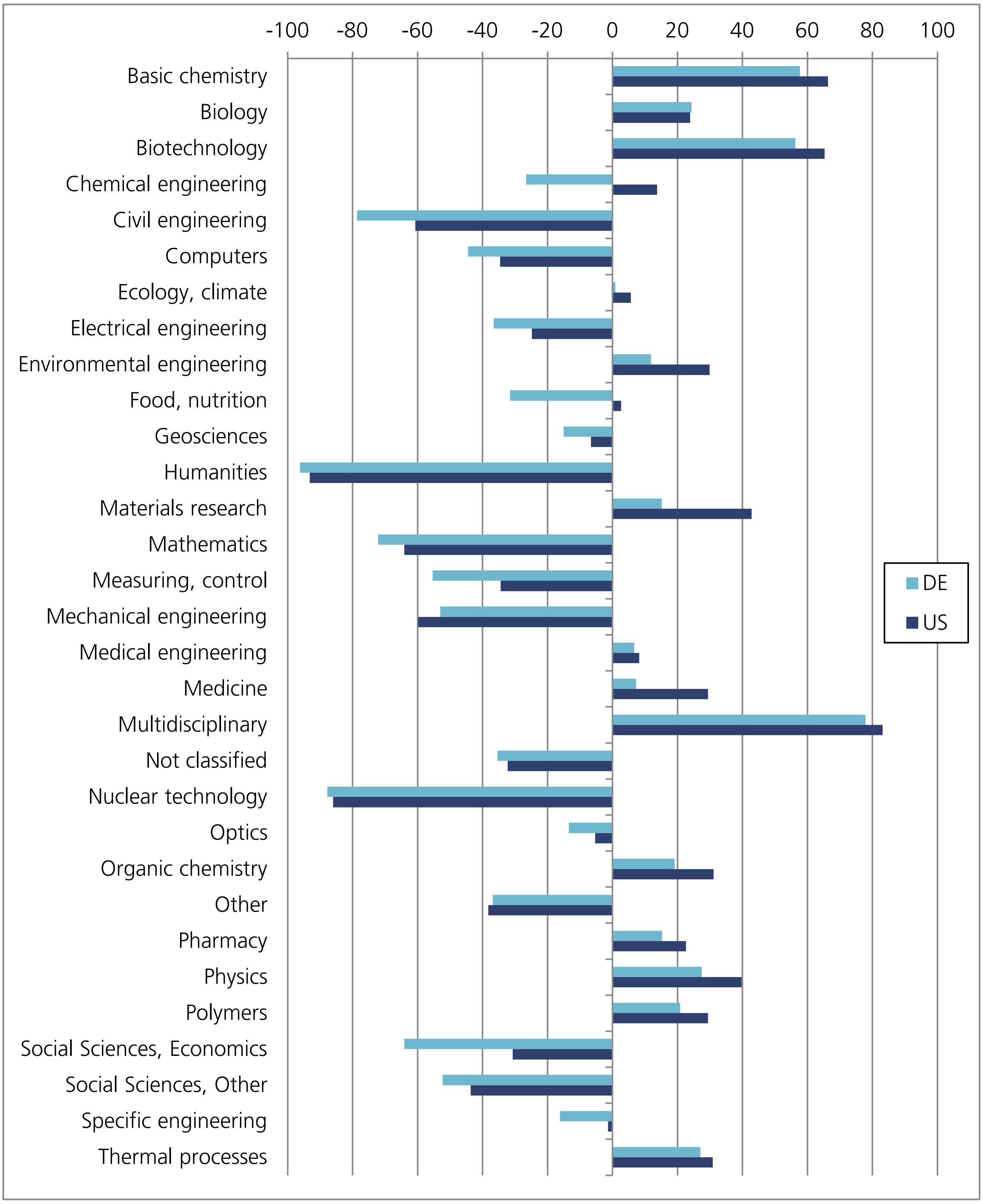
Figure 20: Citation rate per field of Germany and the EU28 for publications in the SCIE and the SSCI in 2011 according to fractional counting



Source: Web of Science, queries and calculations by Fraunhofer ISI

Figure 21 to Figure 24 show the IA and SR index of Germany, China and the USA in comparison. Again, the USA and Germany share a very similar profile. Since the IA is calculated based on a comparison to the worldwide average citation rate, it can be expected to show negative values in fields with a general low citation rate, e.g. the Social Sciences. Therefore, the similar values can be expected. Also in this indicator, the field Food, nutrition, in which Germany publishes relatively seldom, deviates negatively.

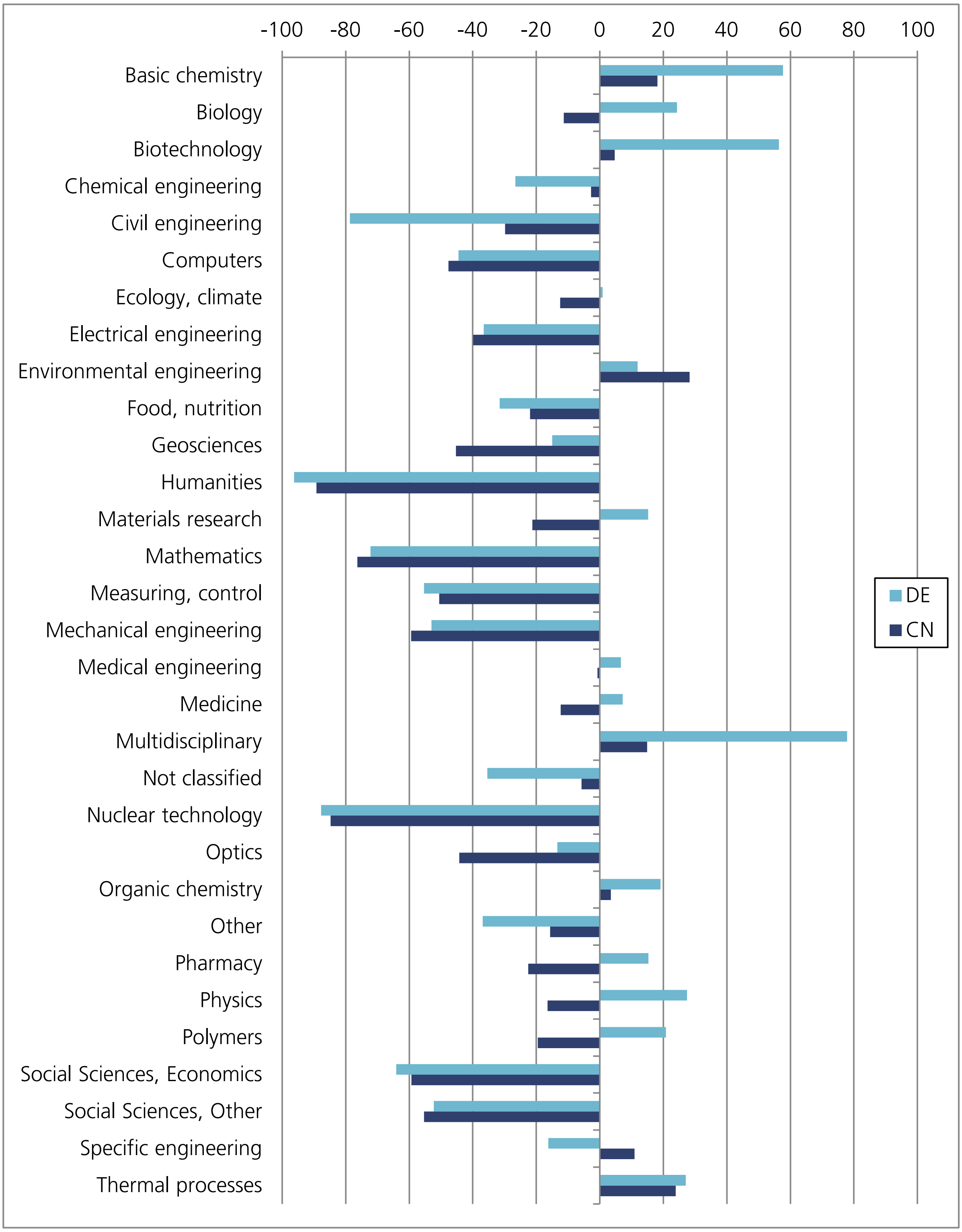
Figure 21: IA per field of Germany and the USA for publications in the SCIE and the SSCI in 2011 according to fractional counting



Source: Web of Science, queries and calculations by Fraunhofer ISI

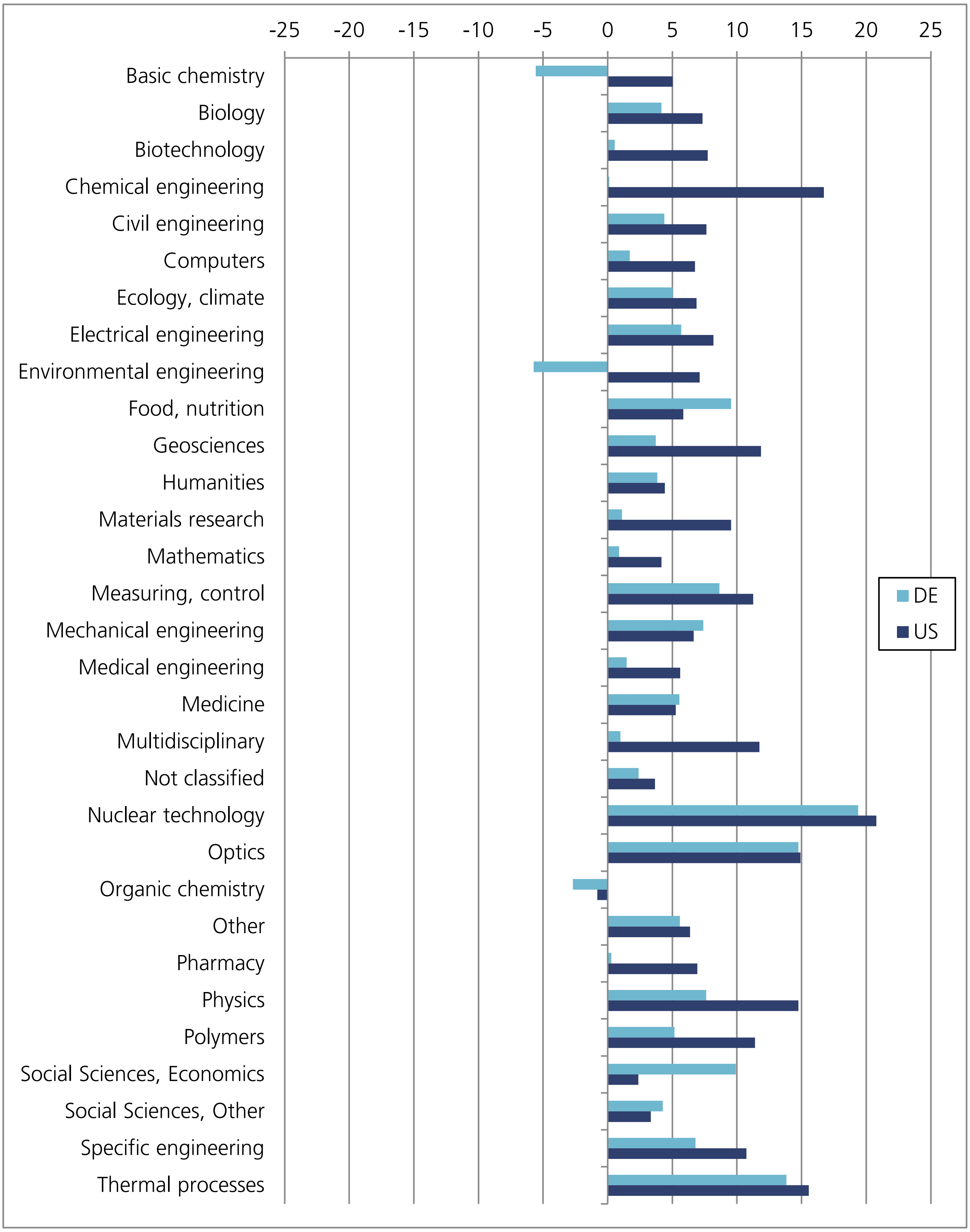
As was already to be expected based on the aggregated IA value shown above, China has a negative IA value for most of the fields (Figure 22). In most of them, it manages to be cited more often than other publications in the respective journals (Figure 24).

Figure 22: IA per field of Germany and China for publications in the SCIE and the SSCI in 2011 according to fractional counting



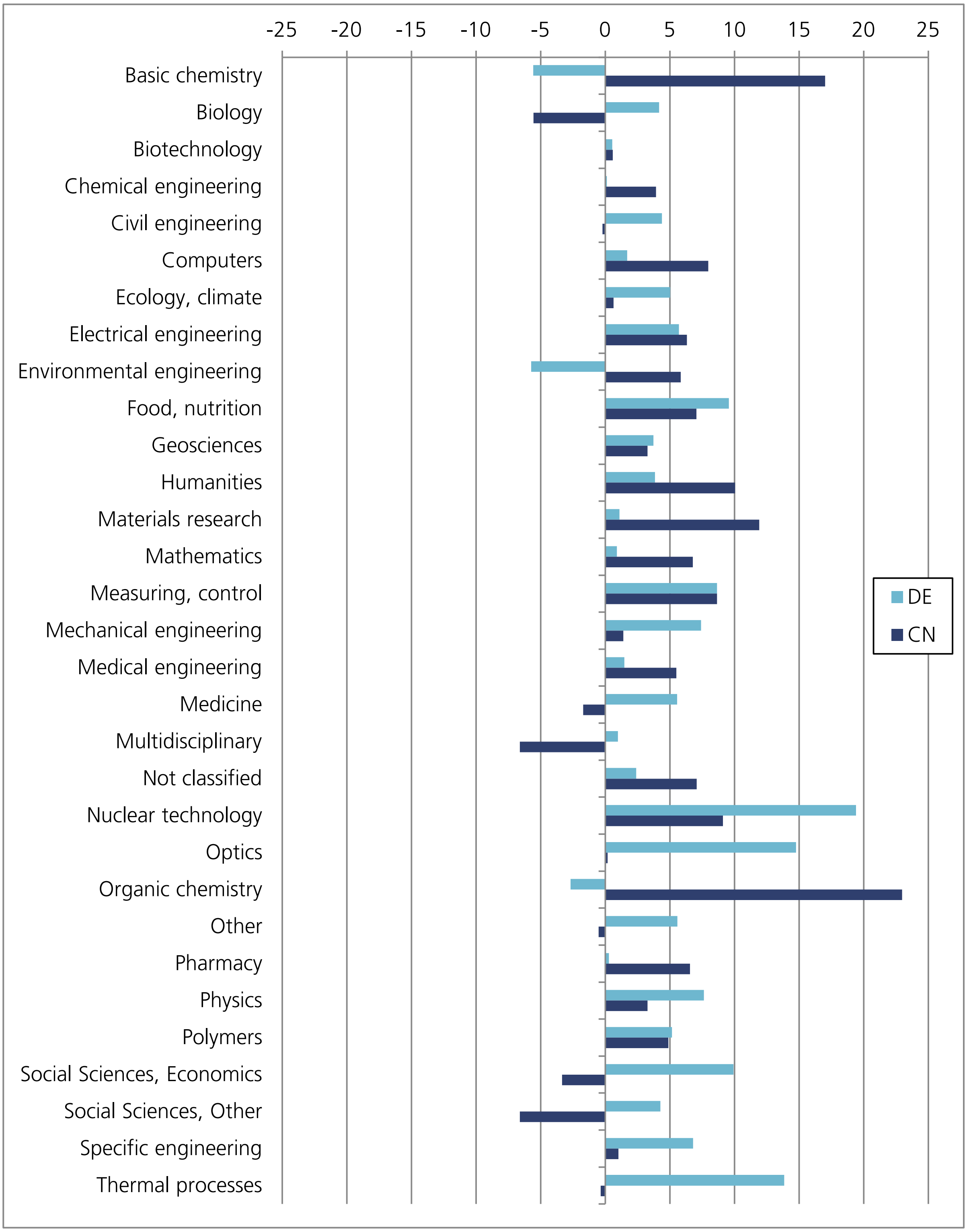
Source: Web of Science, queries and calculations by Fraunhofer ISI

Figure 23: SR per field of Germany and the USA for publications in the SCIE and the SSCI in 2011 according to fractional counting



Source: Web of Science, queries and calculations by Fraunhofer ISI

Figure 24: SR per field of Germany and China for publications in the SCIE and the SSCI in 2011 according to fractional counting



Source: Web of Science, queries and calculations by Fraunhofer ISI

4 Part III: Co-publication analysis of Germany, China and the USA

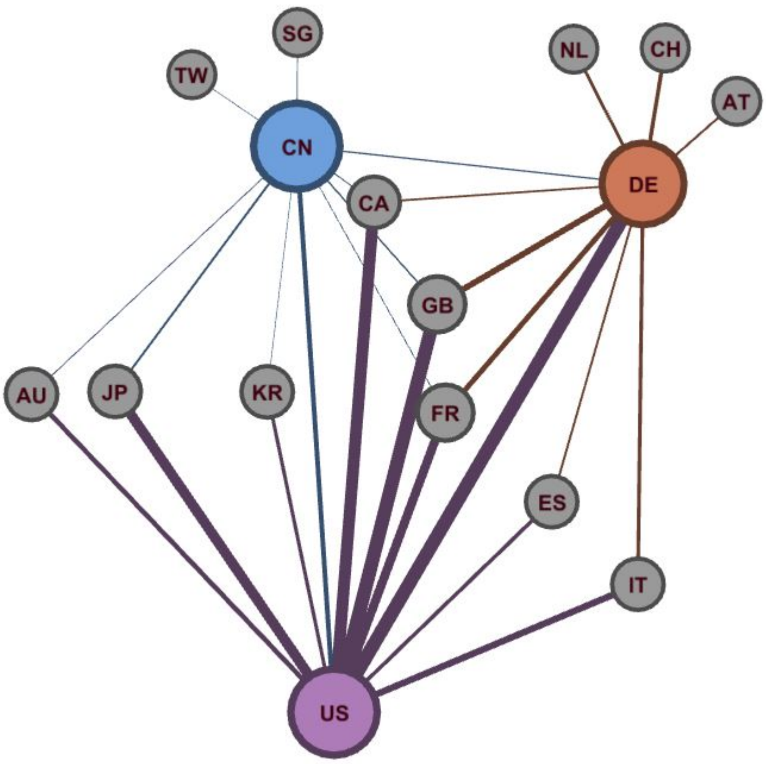
International co-publications are able to reflect international collaborations between institutions from different countries, thereby giving an indication of international knowledge exchange and internationally acknowledged scientific knowledge. Based on co-publication data networks and collaboration patterns between countries can be analysed, using the author addresses listed on the publications. The main aim of the analysis provide in this section is to compare the international co-publication patterns between Germany, China and the USA, respectively.

4.1 Co-publication networks

Figure 25 and Figure 26 compare the co-publication patterns in 2000 and 2013. The country set consists of Germany, the USA and China and their respective top 10 co-publication partners. Since the top 10 co-publication partners for the three countries overlap, 16 countries are shown in total.

In comparison between the two observation years 2000 and 2013, the USA show diminishing amounts of co-publications with most countries. The only exception is China, with which they have intensified their collaborations. In contrast, Germany and China have relatively stable connections to the other partner countries.

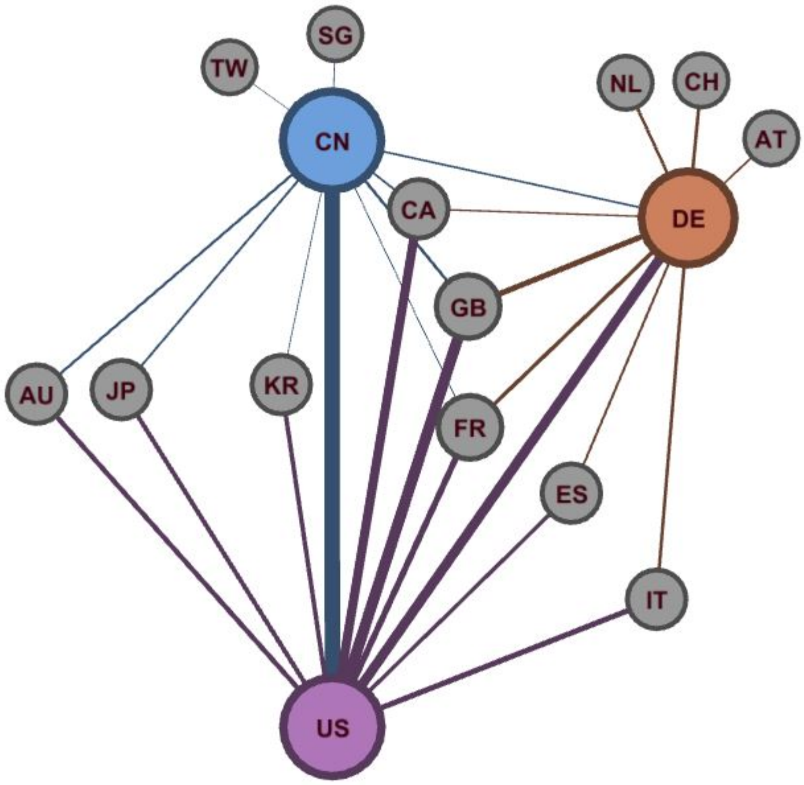
Figure 25: Co-Publication pattern of Germany, the USA and China for publications in the SCIE and the SSCI according to whole counting for the year 2000



Source: Web of Science, queries and calculations by Fraunhofer ISI

There are many countries that belong to the most frequently used co-publication partners for all three of the main countries. China and especially Germany use their geographical neighbours in addition. The closest and biggest neighbour of the USA, Canada, is also among the top 10 collaboration partners of Germany.

Figure 26: Co-Publication pattern of Germany, the USA and China for publications in the SCIE and the SSCI according to whole counting for the year 2013



Source: Web of Science, queries and calculations by Fraunhofer ISI

4.2 Impact of co-publications

For a further test, the citation rates for the three countries, the USA, Germany and China, were compared with regard to international collaboration partners. For that purpose, the citation rates were calculated for the co-publications with the other countries as well as for the publications in which no international partner was involved. Table 6 shows for each of the three countries the citation rate for co-publications with other countries, that of purely national publications and that of all publications. In general, publications without an international partner are cited less. The visibility is higher for international publications and thus positively influences the citation rate. No clear trend can be seen in that certain countries elevate the citation rate in general. For instance, China achieves the highest citation rate in co-publications with Brazil, Germany does so with Canada and the USA reach their maximum value in Table 6 with Switzerland. However, generally speaking high values are achieved in the co-publications with Belgium, Denmark, Spain and Finland.

Table 6: Comparison of citation rates of the USA, Germany and China for international co-publications and publications in the SCIE and the SSCI according to whole counting for the year 2011

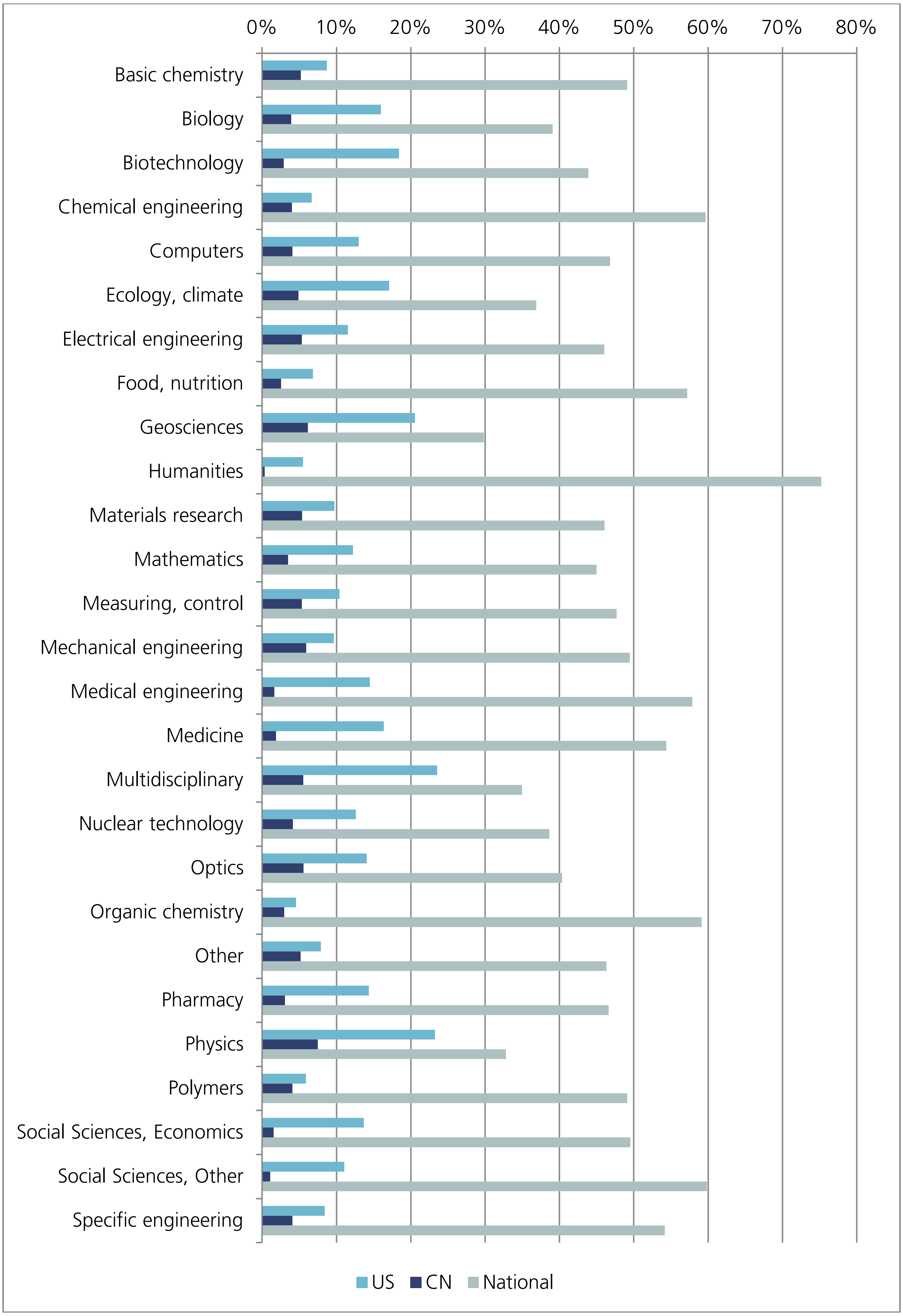
	US	DE	CN
US		10.4	6.6
DE	10.4		8.2
CN	6.6	8.2	
AT	11.0	8.6	10.5
BE	11.9	11.6	10.6
BR	7.1	9.4	16.9
CA	9.6	13.6	6.3
CH	12.6	9.5	11.1
DK	11.7	12.0	13.2
ES	10.3	10.8	12.5
FI	11.0	12.2	12.1
FR	10.6	10.6	8.1
GB	11.0	10.4	7.0
IL	9.1	12.2	13.5
IN	7.2	7.5	11.9
IT	10.1	11.1	10.3
JP	9.9	10.5	5.6
KR	6.8	12.1	6.3
NL	12.0	10.9	9.7
PL	10.2	9.4	14.8
RU	7.6	6.7	10.7
SE	11.3	11.6	9.3
ZA	9.4	10.3	12.6
Other Countries	6.8	7.1	5.8
National Publications	5.3	4.0	3.4
Total	5.7	4.6	3.6

Source: Web of Science, queries and calculations by Fraunhofer ISI

4.3 Co-publications in scientific fields

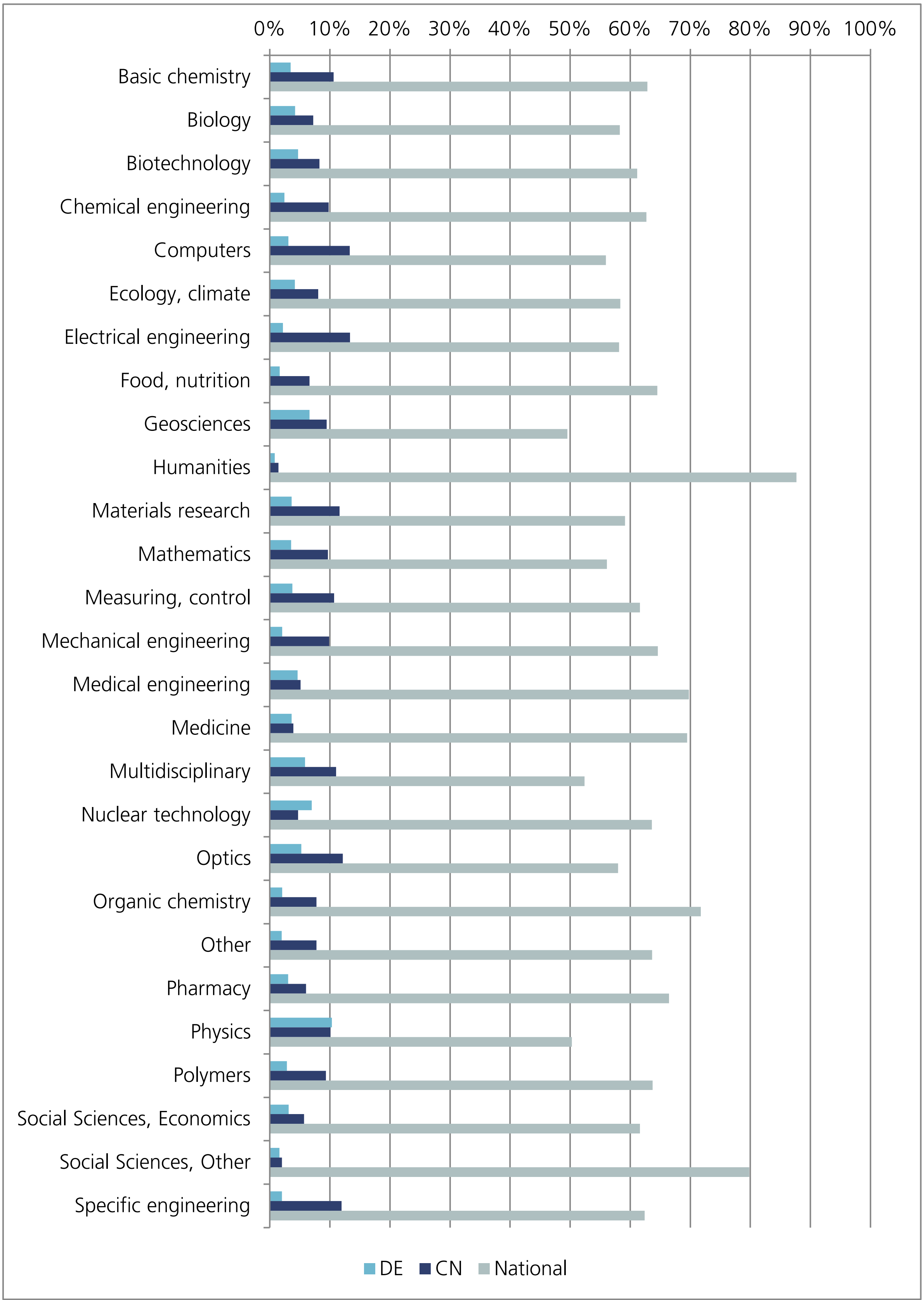
Figure 27 shares the co-publication shares of Germany with the US and China in the different scientific fields. In addition, the shares of publications with only German institutions in each field are shown. The latter is especially high for fields that have to take into account national characteristics, e.g. law, humanities or food. Shares of co-publications are especially high in those fields in which Germany is specialized, particularly Physics, Nuclear technology and Medical engineering.

Figure 27: Co-publication shares in percent of Germany in the 26 fields for publications in the SCIE and the SSCI according to whole counting for the year 2013



Source: Web of Science, queries and calculations by Fraunhofer ISI

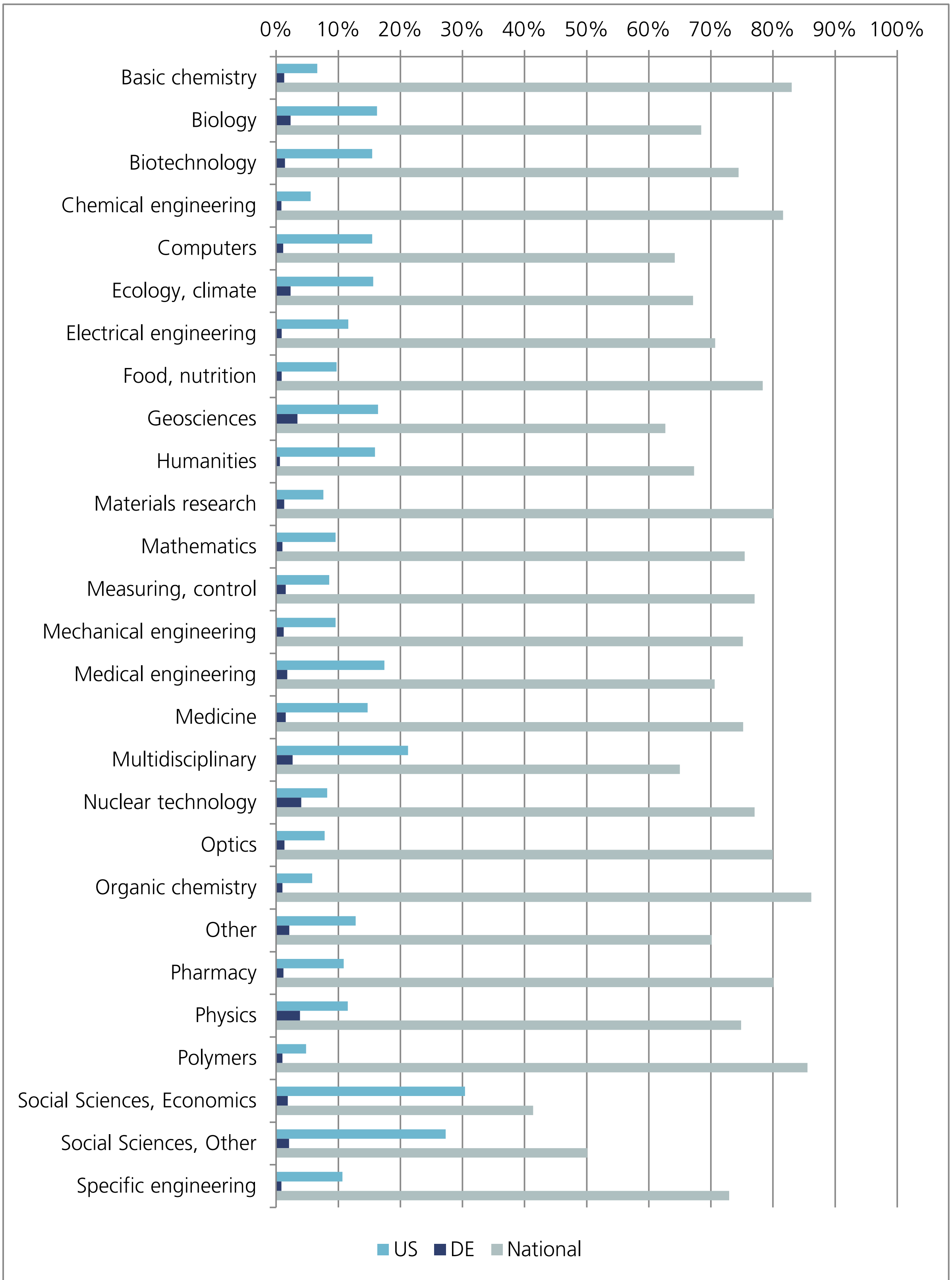
Figure 28: Co-publications of the USA in the 26 fields for publications in the SCIE and the SSCI according to whole counting for the year 2013



Source: Web of Science, queries and calculations by Fraunhofer ISI

In contrast to Germany, the USA and China have much higher shares of purely national co-publications (Figure 28 and Figure 29). However, their grown connection is also showing in the different scientific disciplines. China oftentimes co-publishes with the USA in their main discipline, Social Sciences. The USA also co-operate often with China in Optics, several fields in Engineering as well as Materials research. Similarly, the highest shares of Germany for the USA and China can be found in its main fields Nuclear technology and Physics. Geosciences are another field in which Germany is in demand.

Figure 29: Co-publications of China in the 26 fields for publications in the SCIE and the SSCI according to whole counting for the year 2013

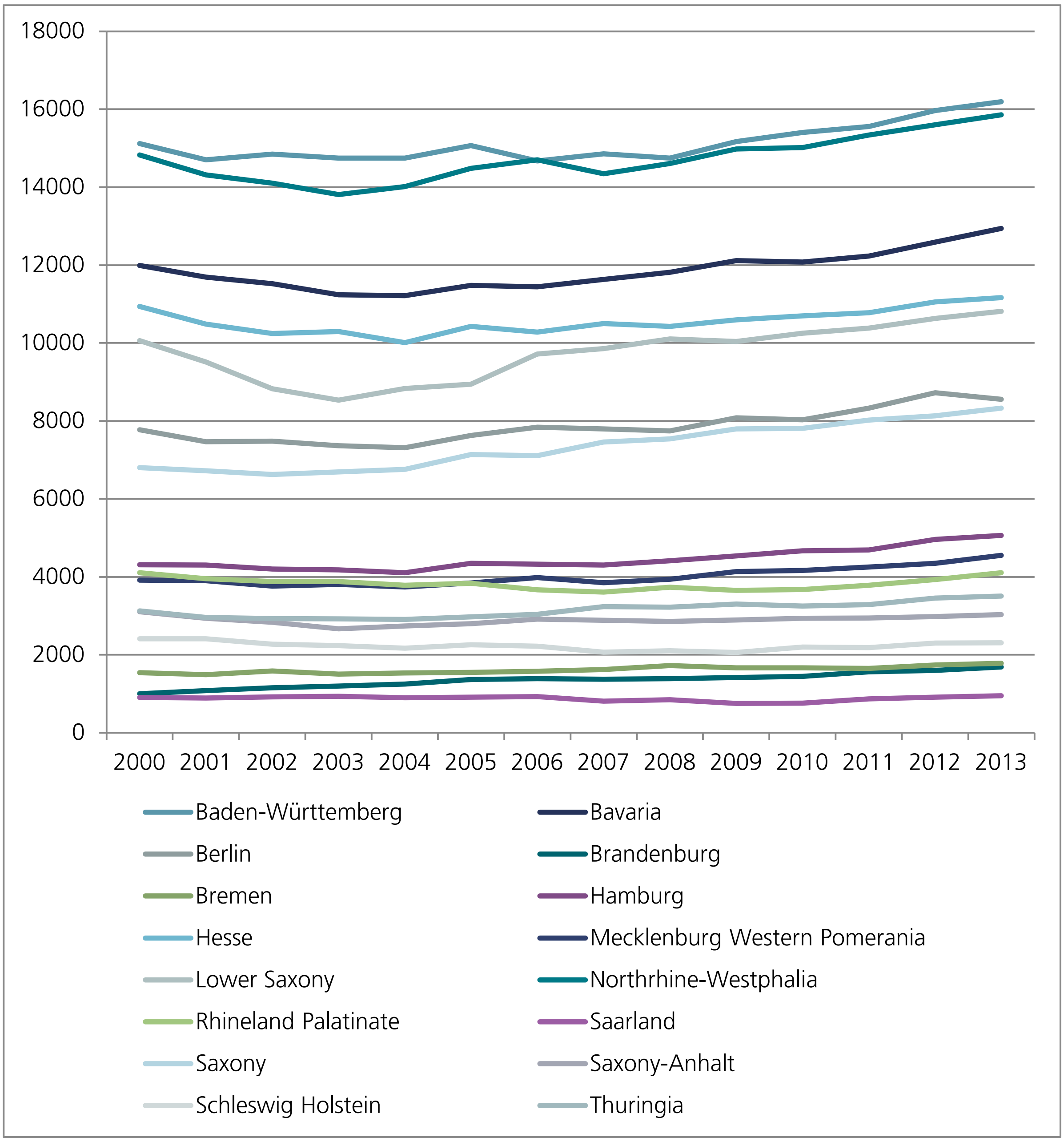


Source: Web of Science, queries and calculations by Fraunhofer ISI

5 Part IV: Regional analysis of Germany

In this part, the publication numbers for the 16 German regions are shown. The absolute numbers are depicted in Figure 30. Baden-Wuerttemberg and North-Rhine-Westphalia are the regions with the most publications, followed by Bavaria, Hessia and Lower Saxony.

Figure 30: Publication numbers of the German regions in the SCIE and the SSCI (fractional counting) for the years 2000-2013



Source: Web of Science, queries and calculations by Fraunhofer ISI

To preclude size effects, these absolute numbers were normalized with the number of inhabitants of the regions (Table 7). In that case, Berlin, Bremen, Hamburg, Mecklenburg-Western Pomerania, Saxony and again Hessia have the highest number of publications per 1 million inhabitants. These values are rather stable over time.

Table 7: Number of publications per 1 million inhabitants of the German regions in the SCIE and the SSCI according to fractional counting

Region	2005	2006	2007	2008	2009	2010	2011	2012	2013
Baden-Württemberg	1,376	1,404	1,366	1,382	1,372	1,412	1,432	1,442	1,473
Bavaria	901	921	916	929	944	968	963	971	994
Berlin	2,160	2,248	2,303	2,282	2,256	2,346	2,321	2,379	2,460
Brandenburg	486	533	544	541	550	564	579	628	641
Bremen	2,314	2,337	2,376	2,451	2,604	2,520	2,526	2,500	2,618
Hamburg	2,368	2,493	2,467	2,431	2,489	2,556	2,614	2,606	2,735
Hesse	1,642	1,711	1,693	1,729	1,719	1,748	1,764	1,769	1,808
Mecklenburg Western Pomerania	2,175	2,253	2,351	2,293	2,365	2,505	2,536	2,600	2,671
Lower Saxony	1,105	1,119	1,217	1,236	1,271	1,266	1,294	1,312	1,343
Northrhine-Westphalia	775	802	815	797	814	838	842	859	874
Rhineland Palatinate	932	945	905	893	927	911	919	947	983
Saarland	850	867	891	785	821	735	750	861	908
Saxony	1,573	1,670	1,672	1,768	1,799	1,870	1,883	1,939	1,967
Saxony-Anhalt	1,098	1,132	1,193	1,195	1,199	1,227	1,258	1,272	1,298
Schleswig Holstein	766	797	783	730	744	728	777	770	810
Thuringia	1,235	1,274	1,315	1,414	1,423	1,468	1,455	1,481	1,565

Source: Web of Science, Eurostat, queries and calculations by Fraunhofer ISI

Given the publication output of the individual regions, their share in German and worldwide publications can be calculated (Table 8 and Table 9). The shares of the regions in respect to Germany vary between 3.1% and 29.4%. The shares for most regions have increased in the observation period. Some (especially larger) institutions cannot be attributed to single regions and are thus distributed among them. Therefore, instead of a shift between the regions, a spread among them can be observed that leads to a higher participation of all the regions in the German publication activity.

When measured by the worldwide publication output, the maximum share amounts to 1.9% for both Baden-Wuerttemberg and North-Rhine-Westphalia (Table 9). Here, the shares are mostly stable for the observation period, showing that even though the distribution in Germany has disseminated, the relative weight in all publications in the scientific landscape has remained the same.

Table 8: Share of the German regions on national publications in the SCIE and the SSCI according to fractional counting[2]

Region	2005	2006	2007	2008	2009	2010	2011	2012	2013
Baden-Württemberg	27.3%	26.6%	26.9%	26.7%	27.5%	27.9%	28.2%	28.9%	29.4%
Bavaria	20.8%	20.8%	21.1%	21.4%	22.0%	21.9%	22.2%	22.8%	23.5%
Berlin	13.8%	14.2%	14.1%	14.0%	14.6%	14.6%	15.1%	15.8%	15.5%
Brandenburg	2.5%	2.5%	2.5%	2.5%	2.6%	2.6%	2.8%	2.9%	3.1%
Bremen	2.8%	2.9%	2.9%	3.1%	3.0%	3.0%	3.0%	3.2%	3.2%
Hamburg	7.9%	7.8%	7.8%	8.0%	8.2%	8.5%	8.5%	9.0%	9.2%
Hesse	18.9%	18.6%	19.0%	18.9%	19.2%	19.4%	19.5%	20.0%	20.2%
Mecklenburg Western Pomerania	7.0%	7.2%	7.0%	7.1%	7.5%	7.6%	7.7%	7.9%	8.3%
Lower Saxony	16.2%	17.6%	17.9%	18.3%	18.2%	18.6%	18.8%	19.3%	19.6%
Northrhine-Westphalia	26.3%	26.7%	26.0%	26.5%	27.2%	27.2%	27.8%	28.3%	28.8%
Rhineland Palatinate	7.0%	6.7%	6.6%	6.8%	6.6%	6.7%	6.9%	7.1%	7.4%
Saarland	1.7%	1.7%	1.5%	1.5%	1.4%	1.4%	1.6%	1.7%	1.7%
Saxony	12.9%	12.9%	13.5%	13.7%	14.1%	14.2%	14.5%	14.7%	15.1%
Saxony-Anhalt	5.1%	5.3%	5.2%	5.2%	5.2%	5.3%	5.3%	5.4%	5.5%
Schleswig Holstein	4.1%	4.0%	3.8%	3.8%	3.7%	4.0%	4.0%	4.2%	4.2%
Thuringia	5.4%	5.5%	5.9%	5.9%	6.0%	5.9%	6.0%	6.3%	6.4%

Source: Web of Science, queries and calculations by Fraunhofer ISI

Table 9: Share of the German regions on worldwide publications in the SCIE and the SSCI according to fractional counting

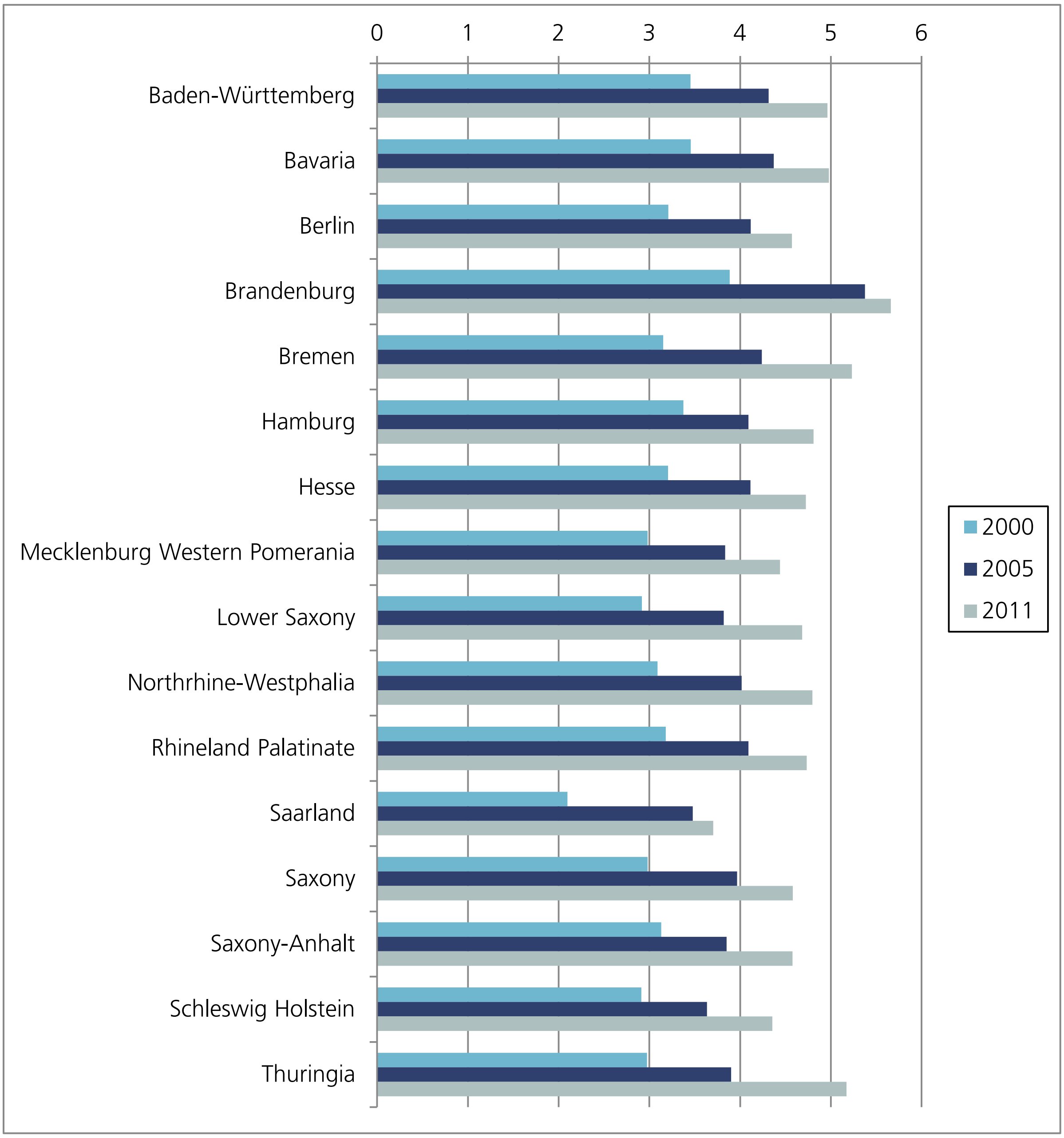
Region	2005	2006	2007	2008	2009	2010	2011	2012	2013
Baden-Württemberg	1.8%	1.7%	1.8%	1.8%	1.8%	1.8%	1.8%	1.9%	1.9%
Bavaria	1.4%	1.4%	1.4%	1.4%	1.4%	1.4%	1.5%	1.5%	1.5%
Berlin	0.9%	0.9%	0.9%	0.9%	1.0%	1.0%	1.0%	1.0%	1.0%
Brandenburg	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Bremen	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Hamburg	0.5%	0.5%	0.5%	0.5%	0.5%	0.6%	0.6%	0.6%	0.6%
Hesse	1.2%	1.2%	1.2%	1.2%	1.3%	1.3%	1.3%	1.3%	1.3%
Mecklenburg Western Pomerania	0.5%	0.5%	0.5%	0.5%	0.5%	0.5%	0.5%	0.5%	0.5%
Lower Saxony	1.1%	1.2%	1.2%	1.2%	1.2%	1.2%	1.2%	1.3%	1.3%
Northrhine-Westphalia	1.7%	1.7%	1.7%	1.7%	1.8%	1.8%	1.8%	1.9%	1.9%
Rhineland Palatinate	0.5%	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%	0.5%	0.5%
Saarland	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
Saxony	0.8%	0.8%	0.9%	0.9%	0.9%	0.9%	1.0%	1.0%	1.0%
Saxony-Anhalt	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.4%	0.4%
Schleswig Holstein	0.3%	0.3%	0.2%	0.3%	0.2%	0.3%	0.3%	0.3%	0.3%
Thuringia	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%

Source: Web of Science, queries and calculations by Fraunhofer ISI

2 Shares sum up to values higher than 100%, since some organizations cannot be uniquely assigned to single regions.

In contrast to the distribution of the publication output, other regions excel in the citation rate (Figure 31). Especially Brandenburg; Bremen and Thuringia, regions that have a relatively low publication count, show very high citation rates of more than 5 citations per publication. Saarland has shown a huge increase in the citation rate from 2000 to2005, collecting then approximately 3.5 citations per publication. Since the average citation rate in Germany in 2005 was 4.0, the deviation was not that far from the reference value.

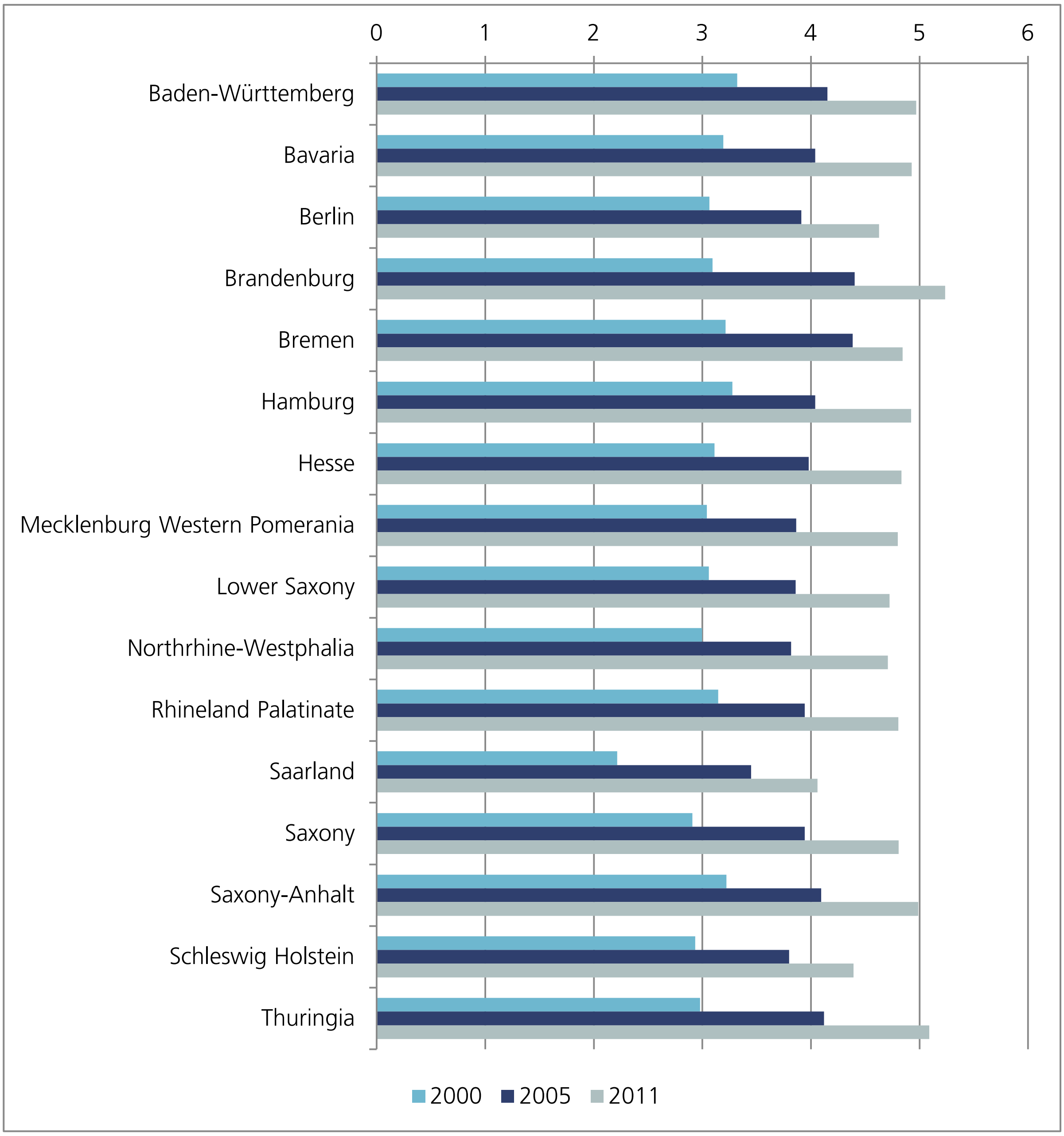
Figure 31: Observed citation rate of the German regions in the SCIE and the SSCI (fractional counting) for the years 2000, 2005 and 2011



Source: Web of Science, queries and calculations by Fraunhofer ISI

The expected citations show the average citation rate in the fields used for publication (Figure 32). The interdependency between the expected and the observed citation rate can be seen in the comparison of Figure 31 and Figure 32; publications in fields with a high/low expected citation rate can influence the citation rate positively/negatively. There are some exceptions to that rule, but mostly in that way that regions gain fewer citations than the average in the field. In particular, Mecklenburg Western Pomerania, Saarland and Saxony-Anhalt gain fewer citations than other publications in the respective fields. However, Brandenburg and Bremen acquire more citations than similar publications. For the former, this observation was even more pronounced in the years 2000 and 2005.

Figure 32: Expected citation rate of the German regions in the SCIE and the SSCI (fractional counting) for the years 2000, 2005 and 2011



Source: Web of Science, queries and calculations by Fraunhofer ISI

6 Supplementary Part: Corporate Publishing

6.1 Introduction

Knowledge production in companies has evolved to one of the most important competitive factors in industry in the past century (Nonaka and Takeuchi, 1999). In the context of financial profit, knowledge in companies can play a role in many aspects. On the one hand, a knowledge advantage can help to better assess (risky) situations and decisions in an uncertain environment (cf. Rosenberg, 1990; Fleming and Sorenson, 2004). In a similar vein, first-mover advantages are enabled by self-conducted fundamental research (Rosenberg, 1990).

Form an economic point of view, collaborations with universities enable companies to outsource and/or enhance research activities and bring many advantages similar to self-conducted research. Particularly, the externally created knowledge can be incorporated in the internal R&D processes. With this background, companies strive to increase the collaboration effort (e.g. Liebeskind et al., 1996; Powell et al., 1996; Vallas and Kleinman, 2008) while downsizing their own R&D departments (Chesbrough, 2003). Even though such projects might start with a clear focus in mind, the results might be applicable to other problems later on as well (Rosenberg, 1990; Stokes, 1997).

Another factor of scientific publications is the signalling effect in regard to other organizations. Hicks (1995) assumed that the main targets of scientific publications by companies are research partners, which would stand in line with the outsourcing and corporation of research as a main focus point. However, with the increasing facilitated access to publishing sources and publications, publications can also be used as a marketing tool, sending signals about current assets and potentials to venture capitalists, prospective partners, consumers and also competitors. While the collaborations of companies are a topic more often analyzed in various aspects, the signalling effect of the publications is a less well explored topic. However, it has been shown by Simeth and Cincera (2013), that the resulting signalling effect positively affects the market value.

Therefore, a closer look will be taken on the publication pattern of companies in Germany. In particular, the co-authors, the publishing journals and their impact will be analyzed. Furthermore, the characteristics of the publishing companies will be unveiled: Are the size of a company and its publication output correlated or do small companies use the chance of a less costly marketing effect more often? Do the characteristics of the company affect its bibliometric performance?

The following analysis tries to answer these and other questions by detecting publishing companies in the Web of Science and the Hoppenstedt database. Apart from comparing the publication behaviour of the different sectors, an analysis of the co-publication patterns of companies might unveil their collaboration affinity and also show the quality of the resulting research.

In order to do that, a methodology for identifying company names in the organization names of scientific authors was developed. Furthermore, the organizations, which were identified in that way, were matched to those stored in the Hoppenstedt database. The following section explains how these two steps were implemented for the purpose of this report.

6.2 Methodology

6.2.1 Identification of companies in the Web of Science

For this report, the organizations registered in Thomson Reuter's Web of Science were classified as universities, companies and others according to their name. This automatic procedure relied heavily on the explicit mentioning of legal forms (e.g. GmbH, AG) or other indicators (e.g. "university").

Firstly, the organization names given in the Web of Science were queried for any mentioning of a legal form of a company. A list with national and international notations of these forms was used. The list covered abbreviated as well as full notations. For instance, "gmbh" (in any form of upper and lower cases) and "gesellschaft mit beschränkter haftung" were search terms contained in this list.[3] All organization names for which a text match could be found were then labelled accordingly, i.e. in the previous example, all matches were labelled as a company. Universities could be identified in the majority by spelling variations of "university" itself (or rather the German equivalent "Universität"). Other types of organizations were more difficult to detect, because they seldom share specific keywords denoting their sector. Therefore, most analyses in this report focus on publications of companies in comparison with universities and/or the whole of Germany. Nonetheless, lists with spelling variations of hospitals and non-university research institutions were used as well to sort out these organization types as well. In particular, some of these organizations used legal forms (especially "GmbH"), so that a distinction was necessary.

Thus, separate lists were used for universities, hospitals, schools and other non-academic institutions. This separation helped in developing and improving the keyword lists. In particular, the results of the matches were checked and corrected after each execution of the search. In this iterative manner, the keyword lists were enhanced by and by, whereas the introduced changes were "tailor-made" for the data in the Web of Science: After each iteration of the textual match, organization names which remained unassigned or were assigned to multiple sectors were checked manually for new patterns. In that way, the list of textual indicators enhanced with each iteration until no further patterns could be detected. In the end, 70% of the German organizations in the Web of Science could be classified unambiguously.

3 To account for umlauts and their spelling alternatives (or their erroneous conversion by OCR software to other letters), wildcards were used in place of these letters in the database queries.

6.2.2 Matching with the Hoppenstedt database

In order to gather more information about the companies themselves, a probability matching of the publishing companies identified in the Web of Science with company names from the German company database Hoppenstedt was performed.

The aim of the matching procedure is locating information on organizations in the Web of Science that corresponds to an entry within the Hoppenstedt company database or has a high similarity with it. For this purpose, the similarity between organization names in the Web of Science and each entry of company names in the Hoppenstedt database is calculated. In case a certain degree of similarity is exceeded, the respective pair of Hoppenstedt and Web of Science organizations is stored as a “match". All data associated with the entry in the Web of Science can now be adopted for the entry in Hoppenstedt, and vice versa. In the case of name variations, name changes and the like, it may happen that several Web of Science entries are assigned to one entry in the Hoppenstedt database. The related information is then aggregated and the Web of Science entries will be treated in a combined form.

In the following, the three steps of the method are described in more detail. In the first step, the organization names in both databases are cleaned via the same procedure. This is to prevent that a different notation leads to a lower similarity. In a first step, the entire text is converted to lower case letters. Then, all the special characters are removed or replaced; umlauts are replaced by the corresponding vowel, i.e. Ü, Ä and Ö to U, A and O. Special characters are replaced by a space. Special characters include all characters that are not a letter or any number, including punctuation. In the next step, all occurrences of multiple spaces are replaced by a single space.

In a final step, legal forms were removed from the names. This applies to both, abbreviated as well as formulated text. Analogous to the term “with” also the term “and” is removed to ensure a greater similarity to texts where the “&” sign was removed as a special character in the previous step.

In the second part of the process, the similarity between the cleaned names was calculated. For this purpose, a variant of the Levenshtein distance was implemented. The Levenshtein distance originally is a measure to determine the dissimilarity between two text strings. It accounts for the number of changes in letters (i.e. edits), which are needed to make two text strings equal. In order to calculate a similarity measure with a value between 0 and 1, the Levenshtein distance is normalized with the length of the longer of the two text strings and subtracted from 1. In that way, a similarity measure for the organization names in the two different databases is calculated.

In the final step, for each Web of Science entry, the entry that has the highest similarity is selected. To assure that the text strings are similar enough, only matches with a similarity of 0.75 and higher were chosen. These matches are then stored in a separate table in the database for further use. The result is a table which connects the publication data in Web of Science with the company information in Hoppenstedt.

With the help of this dataset, differentiated multivariate analyses are possible, which enable us to interpret the effects of certain characteristics, e.g. size, from other characteristics we control for, e.g. sectors.

6.3 Bibliometric Analysis

For this report, publication counts are measured as whole count. In that way, co-publications are not penalized by the counting method. As the assumption is also that many of the publications by companies are written specifically for the purpose of exhibiting a co-operation, co-publications should be quite common. Fractional counting would thus be counterproductive.[4] The observation period ends in 2012, since the affiliation data are not yet complete for 2013. This leads to slight drops for all types of organizations (universities and companies) in 2013 and the data were thus excluded from the analyses.[5]

Figure 33: Share of articles with company involvement in Germany in the SCIE and the SSCI (whole counting) for the years 2000-2012

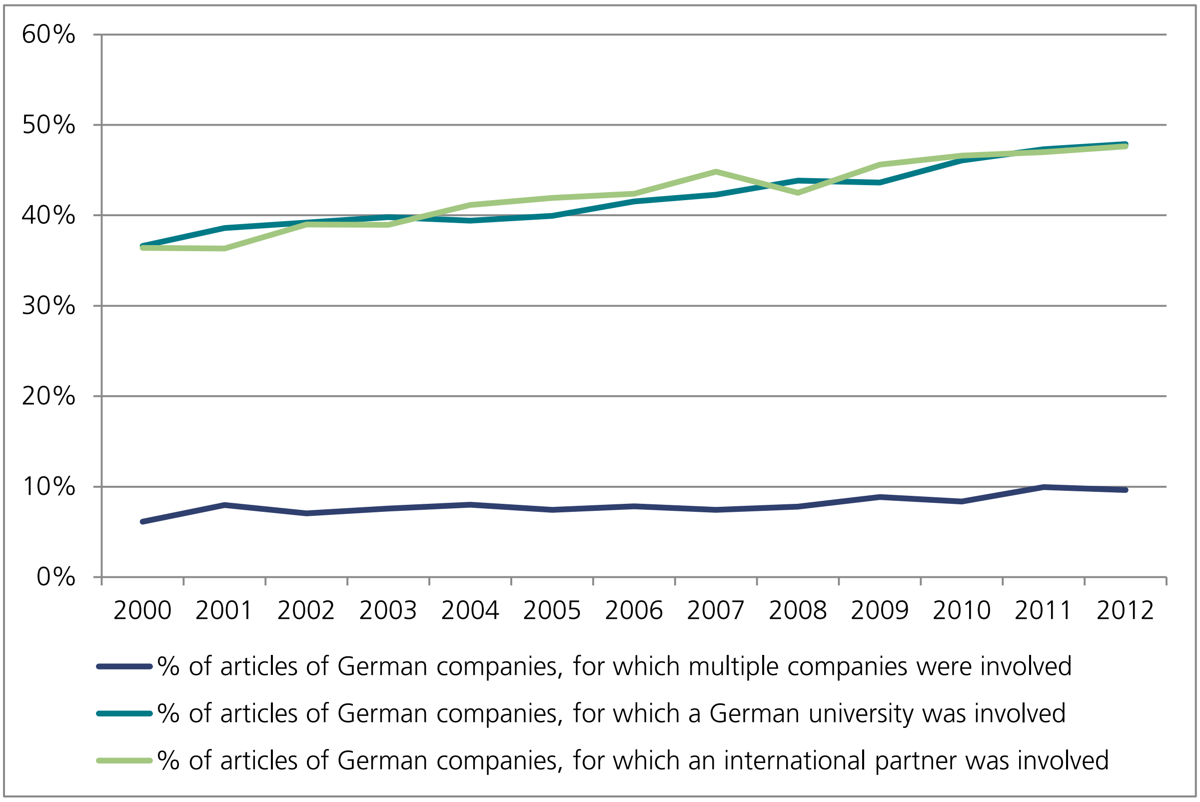
Source: Web of Science, queries and calculations by Fraunhofer ISI.

4 I.e. the more actors are involved in a co-publication, the less it is worth in the publication count.

5 The loss in publication counts accounts for approximately 1%, which might seem low, but in most cases breaks a trend of rising numbers.

Figure 33 shows the absolute and relative number of articles published in Germany, which were written with at least one author from a German company. While the absolute value has been steadily increasing throughout the observation period, the share of publications by at least one German company has slightly decreased. Thus, the increase in absolute numbers is proportionally lower than that for Germany as a whole. However, as a result the relative value decreases only slightly and still amounts for approximately 6.8% in 2012.

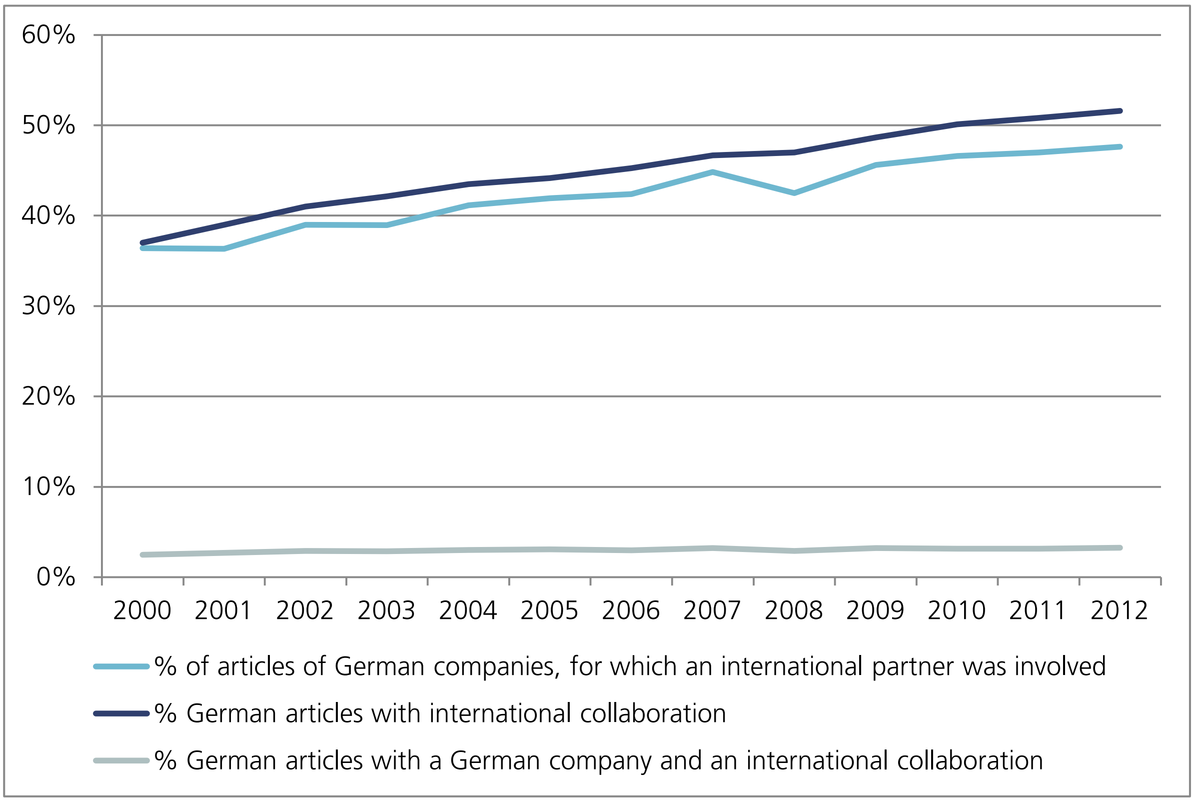
Figure 34: Share of articles of German companies which are the result of collaborations with universities, other companies and international partners in the SCIE and the SSCI (whole counting) for the years 2000-2012



Source: Web of Science, queries and calculations by Fraunhofer ISI

Considering the role of scientific publications as an instantiation of co-operations, a closer look is taken at the collaboration pattern of German companies (Figure 34). The first thing to note in this respect is that most collaborations involve at least one university. At the end of the observation period, nearly 50% were written in collaboration with a university. Thus, the ties between these two actors have tightened. The same can be said for collaborations with international partners, for which the numbers are almost the same. We cannot assess whether these foreign partners are universities or companies, but the companies' research seems to benefit from the import of knowledge. Yet, when setting these shares in relation to Germany's international collaboration pattern (Figure 35), we can see that in general international collaborations are even more common.

Figure 35: Share of German articles which are the result of collaborations with universities, other companies and international partners in the SCIE and the SSCI (whole counting) for the years 2000-2012

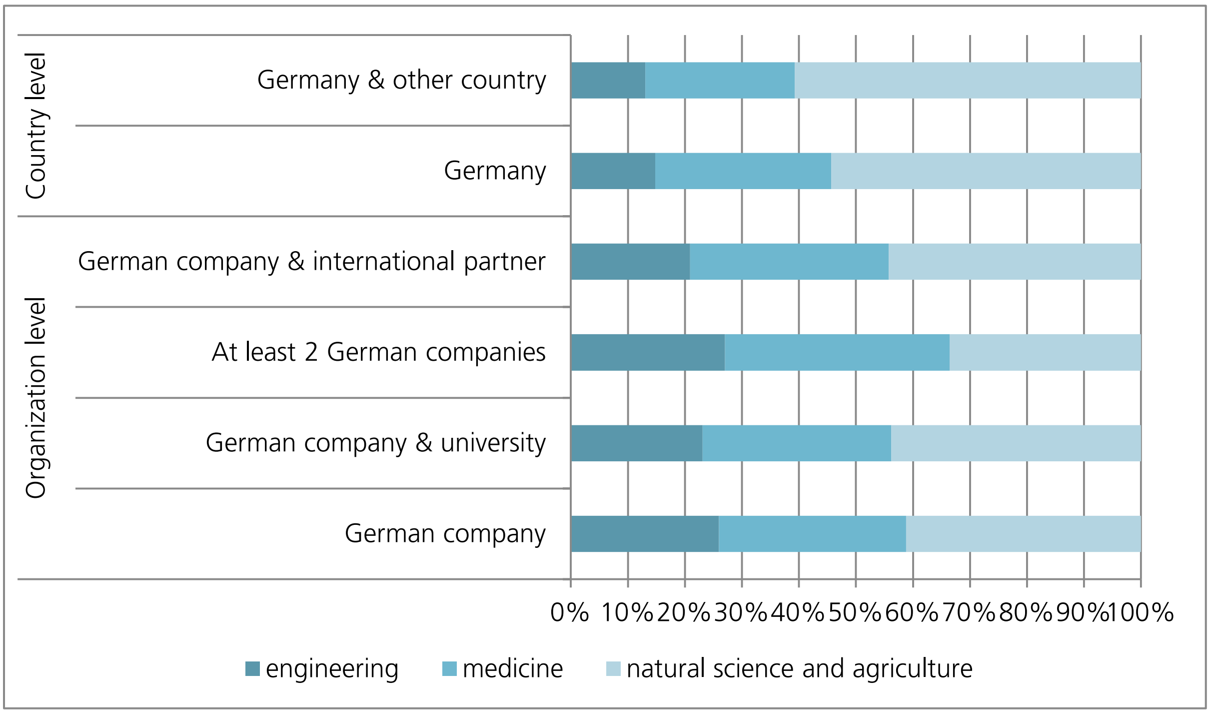


Source: Web of Science, queries and calculations by Fraunhofer ISI

Counting all German articles, in 2012 approximately 3.5% are created with the involvement of both a German company and an international partner (line at the bottom of the figure). When comparing the statistics for German publications in total with those of German companies in regard to international collaboration, the observations are very similar. All in all, Figure 35 shows that companies choose a foreign co-publication partner less often than the German average.

Collaborations between two different (German) companies play only a minor role (Figure 34). In 2012, approximately 8.5% of the papers emitted by companies are the result of a partnership between multiple companies. One explanation could be that the competitiveness is too high in most cases to consider a partnership. A reasonable scenario is that (project) partners from different fields seek a collaboration to complement each others methods and resources, but it is in most cases unlikely that the results of this co-operation are also publishable in a scientific context (especially if no academic partner is involved) or that the partners would want to advertise their co-operation. However, again this would be a form of outsourcing research activities as described in the literature behind industrial publications.

Figure 36: Share of publications of different publishing organizations according to scientific disciplines in the SCIE and the SSCI according to whole counting in 2012

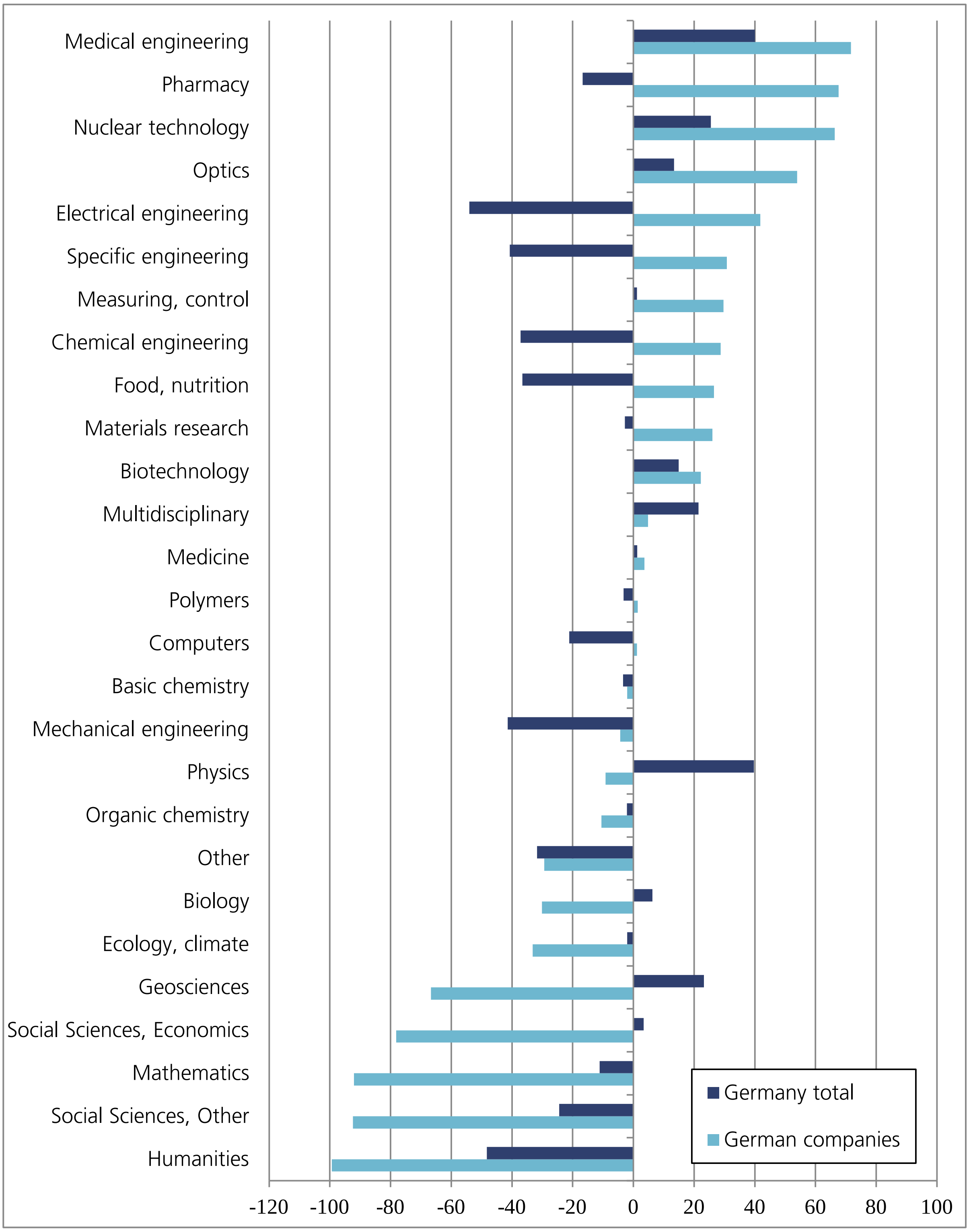


Source: Web of Science, queries and calculations by Fraunhofer ISI

The scientific fields were analyzed to test for further differences in publication behaviour. Figure 36 shows the percentage of German publications in the three scientific main fields “engineering”, “medicine” and “natural science and agriculture”. The majority of Germany’s publications fall into the latter category. This is even more pronounced in collaborations with authors from other countries. However, publications by German companies deal relatively seldom with this field: 41% of the publications by German companies can be associated with the field “natural science and agriculture”, for Germany in general the share amounts to 54%. In contrast, German companies focus more on medicine and especially engineering.

Another indicator for differences in the fields is the Revealed Literature Advantage (RLA, see Hinze and Grupp, 1996). In Figure 37, the share of German publications (by companies and as a whole) is compared to the world average in the respective scientific fields and transferred to a range between +/-100. In comparison to the total German publications, the German companies publish more often in the fields of engineering as well as medicine and pharmaceuticals and less often in natural sciences. Characteristic for the industrial sector is the neglect of the Humanities and Social Sciences.

Figure 37: Field specialization index (RLA-index) of German companies and total German publications with regard to the world average, 2012



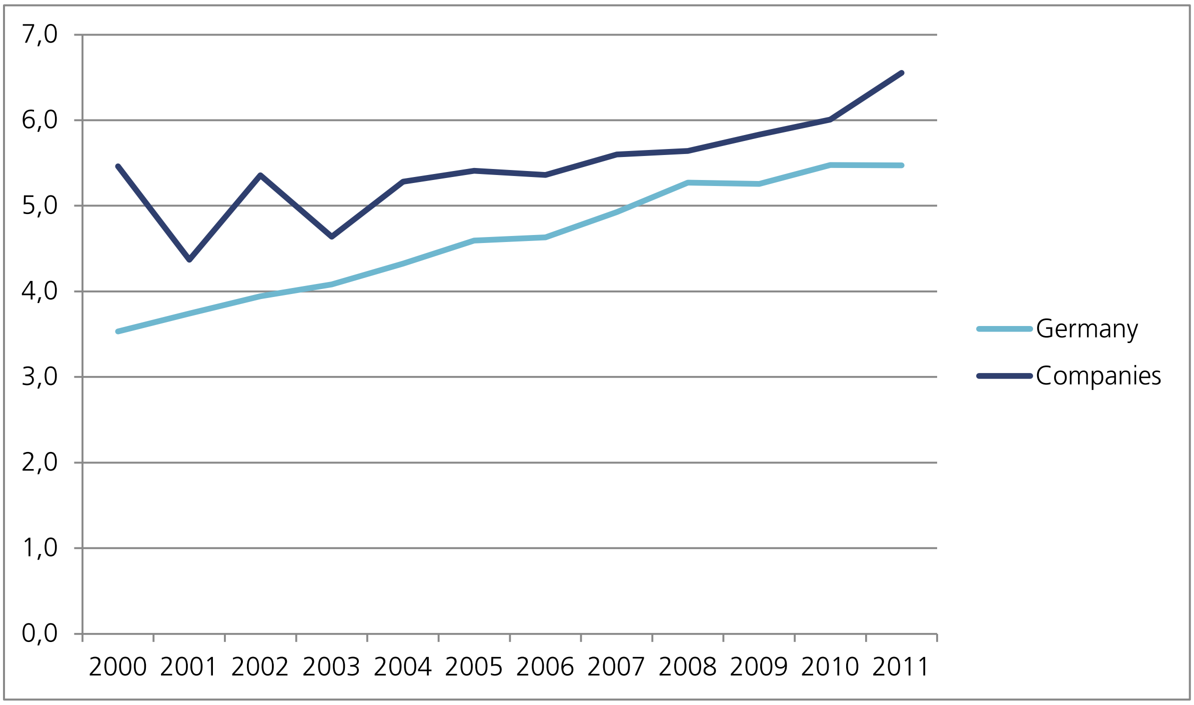
Source: Web of Science, queries and calculations by Fraunhofer ISI

Citation rates for articles by German companies are slightly higher than those of Germany in total in the observation period. For instance, in 2011, the citation rate of companies accounts for 8.7 citations per document, while that of Germany in total is 7.6. Overall, the trend of increasing citation rates is observable for both publication sets in an equal manner.

If we exclude self-citations, the gap between German companies and Germany becomes even wider (Figure 38). Thus, we can conclude that a smaller proportion of the citations for the companies' publications are based on self-citations. This observation complies the usage of company publications as a means for marketing instead of research itself: The follow-up on former work and thus the self-reference is not as important as the publication itself.

To analyze this aspect further, we classified the citing publications as well into groups according to the involvement of companies and universities. In that way we can also see if there is a transfer of research from companies to universities and vice versa or if both groups rather keep to themselves.

Figure 38: Citation rates (self-citations excluded, 3-year citation window) for publications of companies and Germany in the SCIE and the SSCI according to whole counting in 2012



Source: Web of Science, queries and calculations by Fraunhofer ISI

Table 10 shows both the publishing as well as the citing source of German articles. The last line denotes all German articles, while the lines above are split up in regard to the involvement of universities and companies. German articles in general attract especially the attention of German universities and of foreign organizations, acknowledging 48% and respectively 53% of the published articles. Citations by German companies are in the minority, as only 6% of the German articles are cited (at least once) by a German company. This also holds for articles by universities, as 6% of the articles written by universities are cited by companies. More frequently, articles in which at least one company is involved are cited by other companies (28%). Thus, even though the citation rates of publications by

German companies are similar to that of Germany in total, the set of emitting institutions behind these citations differs. Thus, findings by other companies have a higher significance for industrial research than fundamental research from universities. Similarly, the highest share of citations by universities is observable for articles from universities (with and without partners from other sectors).

Table 10: Sources of citations for publications of companies and Germany in the SCIE and the SSCI according to whole counting in 2012

	% of articles cited by universi-ties	% of articles cited by compa-nies	% of articles cited by foreign organizations	% of articles cited only by German organi-zations
German articles by companies and universities	41%	18%	50%	45%
German articles by at least 2 companies	39%	25%	48%	39%
German articles by universities	54%	6%	54%	40%
German articles by companies	37%	28%	48%	35%
German articles	48%	6%	53%	38%

Source: Web of Science, queries and calculations by Fraunhofer ISI

Another aspect which can influence and explain the background of the observed citations concerns the publishing journal; its reputation and visibility can affect the citation of the therein published articles in both a negative as well a positive way. Therefore, two metrics are calculated to analyze the journals used:

- The International Alignment (IA) compares the citation rate of the publishing journal with that worldwide, showing whether the journal has a high international reputation or not.
- The Scientific Regard (SR) compares the citation rate of an article with the citation rate of other articles published in the same journal in the same year.

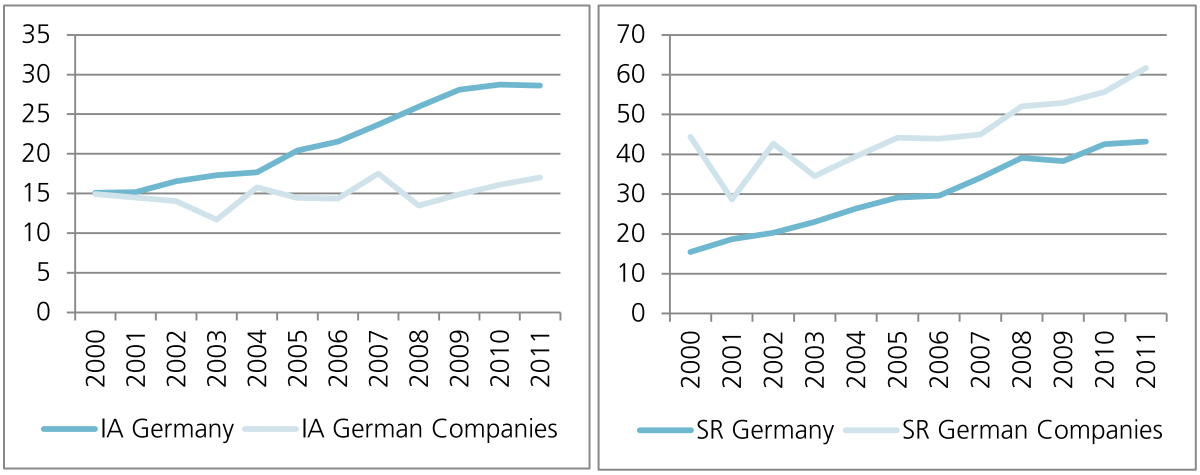
Therefore, the IA denotes the visibility of the journals in which the articles were published and the SR shows how the articles fare in comparison with other articles in the same journals (see Grupp et al., 2001).

Figure 39 shows the IA and the SR for Germany and the German companies in comparison. While the development of the IA for German companies differs from that of Germany in total, the SR value shows a similar trend. The IA suggests, that German companies by trend publish in journals with a lower reputation. These journals might on the one hand have lower barriers than in other journals, i.e. might not be as selective, but on the other hand be also more specialized or application-oriented – application-oriented journals are on average less often cited than basic research oriented journals (Boyack et al., 2013).

While the reputation of the journals for German publications has increased during the observation period, that of German companies varies only slightly over the years.

The SR of Germany is lower than that of the subset of German company publications. Thus, in relation to other articles published in the same journal, German companies are cited more frequently. In regard to the citation rates of Germany and the German companies being nearly equal, this shows how the companies have found their niche in the publication outlets: They aim for journals with a smaller readership. Because of the probably also more narrowed focus of these journals, the articles have both a higher chance to be published and also to be seen by the target audience. Regarded only in the context of their selected journals, the companies are "overperformers" in terms of citations. Thus, while the selection of journals with a low (international) visibility might seem counterproductive in regard to a signalling effect, the articles get a high attention and thus might still be used for this purpose.

Figure 39: IA and SR of publications with and without company involvement in Germany in the SCIE and the SSCI (whole counting) for the years 2000-2011 (ratio of averages, self-citations included)



Source: Web of Science, queries and calculations by Fraunhofer ISI

Another factor that can influence the visibility, the reputation and in turn also the citation rate is the language of the article. Naturally, publications in English achieve higher impact as they can be read and used worldwide. Publications in other languages have a smaller target audience and are thus also less disseminated.

Figure 40 shows the share of German publications in total and German companies. Regarding these figures, one has to bear in mind that the overall fluctuation can be a database effect reflecting the coverage of German language publications in the Web of Science. Thus, the absolute value of these shares should not be regarded solely. However, the shares of Germany and the German companies can be compared to each other and over time.

Figure 40: Share of German language publications for German companies and Germany in total in the SCIE and the SSCI (whole count) for the years 2000 to 2012

Source: Web of Science, queries and calculations by Fraunhofer ISI

Overall, German companies are – in comparison with other organization forms – more prone to publish in German. Thus, the journals used for publication are more often only covering German articles. This also partially explains the relatively low IA values (cf. Figure 39): For journals restricted to a German-language readership, this visibility-dependent indicator shows a value far below that of the in general more internationally publishing German authorship. Still, this does not only reflect on the visibility but also on the focus of the journals concerned. Journals in a native language also target more often issues that are of national interest. Furthermore, it is an indication of a large national community and a result of high specialisation (Schmoch et al., 2012; Schubert and Michels, 2013). On the one hand, this accommodates the use of the German language but on the other hand, narrows the target audience further down. Please note, however, that this still concerns only a minority of the publications by German companies. The tendency is just higher for them than for other publishing organizations, with a difference to the German average by approximately 1 to 4%.

Reasons for this tendency could be that the journals used for publication are a better marketing instrument for the German companies (as their focus is more aligned to national issues) or have lower entrance barriers. The latter assumption is corroborated by the observations made for the IA and SR again; by using journals with a lower reputation, in which the publications more often outshine the other articles in terms of citations, the placement of the articles should be easier in the first place.

6.4 Characteristics of publishing companies

Within this section, we will dig deeper into the characteristics of publishing companies within the German economy. Are there structural differences with regard to the size and sectoral distribution of publishing companies? Are high-technology firms more prone to publishing than their counterparts from less R&D-intensive sectors? Are there differences with regard to the age of companies, i.e. are younger firms publishing more than older firms or vice versa? Are there structural differences with regard to citations, too? These questions will be analyzed in more detail with the help of a multivariate model, which will be discussed in the following.

6.4.1 Variables and Summary Statistics

Table 11 shows the summary statistics for the publishing firms in the sample. In total, 5,209 companies could be assigned to a corresponding record in Hoppenstedt via the text string comparison described above. As with the former analyses, only journal articles were analyzed to focus on the participation of companies in the scientific discourse.

Table 11: Summary statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
# publications	15,627	1.23	4.58	0.00	135
# citations	15,627	8.87	53.87	0.00	1563
Citations per publication	7,809	5.04	8.57	0.00	270.5
Firm age (in years)	15,342	29.79	40.14	1.00	551
Firm size (categorical)	9,327	0.87	0.88	0	2
Sector (1-digit) (dummy)	15,573	9	5	1	19
Sector (2-digit) (dummy)	15,573	53	22	1	96
Publication year	15,627	2009	0.82	2008	2010

Source: Web of Science, Hoppenstedt, queries and calculations by Fraunhofer ISI

For each of the 5,209 companies, publication information was collected from the Web of Science for the years 2008 to 2010. Thus, our dataset consists of 15,627 firm-year observations. The indicators calculated from the Web of Science were the number of publications per firm in the given years as well as the citations received by those publications. We additionally calculated the average number of citations per publication, which gives us an impression of the quality of firm publications absent size effects. These three variables serve as dependent variables in our models. From the Hoppenstedt database, several firm-structure related variables were compiled, which enter our models as explanatory variables. This is first of all a dummy variable capturing the size of a given firm. Companies with less than 50 employees and sales values below 10 Mio. Euros per year were classified as small firms (coded “0”). Companies with 50 to 249 employees and sales values between 10 and 50 Mio. Euros were classified as medium-sized (coded “1”). The remaining companies were categorized as large firms (coded “2”). Since not all firms have information on employees and/or sales, this reduces our sample size for the final models. Furthermore, the age of the company in years was derived from the Hoppenstedt database, based on the year of foundation. Finally, we added the sector information (NACE Rev. 2, 1-digit[6] and 2-digit[7]) to our dataset in order to control for different publishing behaviour in the industrial sectors. The number of publications, citations and the average number of citations per publication across industrial sectors are displayed in Table 12. As we can see from the table, the largest shares stem from “Manufacturing", “Professional, scientific and technical activities", and “Wholesale and retail trade, repair of motor vehicles and motorcycles". Within the manufacturing sector, especially the ICT, pharmaceuticals and chemical industries have the highest publication shares.

6 The 1-digit level sectors are: A. Agriculture, forestry and fishing, B. Mining and quarrying, C. Manufacturing, D. Electricity, gas, steam and air conditioning supply, E. Water supply; sewerage, waste management and remediation activities, F. Construction, G. Wholesale and retail trade; repair of motor vehicles and motorcycles, H. Transportation and storage, I. Accommodation and food service, activities, J. Information and communication, K. Financial and insurance activities, L. Real estate activities, M. Professional, scientific and technical activities, N. Administrative and support service activities, O. Public administration and defence; compulsory social security, P. Education, Q. Human health and social work activities, R. Arts, entertainment and recreation, S. Other service activities

7 The 2-digit level subsectors are: 10 Manufacture of food products, 11 Manufacture of beverages, 12 Manufacture of tobacco products, 13 Manufacture of textiles, 14 Manufacture of wearing apparel, 15 Manufacture of leather and related products, 16 Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials, 17 Manufacture of paper and paper products, 18 Printing and reproduction of recorded media, 19 Manufacture of coke and refined petroleum products, 20 Manufacture of chemicals and chemical products, 21 Manufacture of basic pharmaceutical products and pharmaceutical preparations, 22 Manufacture of rubber and plastic products, 23 Manufacture of other non-metallic mineral products, 24 Manufacture of basic metals, 25 Manufacture of fabricated metal products, except machinery and equipment, 26 Manufacture of computer, electronic and optical products, 27 Manufacture of electrical equipment, 28 Manufacture of machinery and equipment n.e.c., 29 Manufacture of motor vehicles, trailers and semi-trailers, 30 Manufacture of other transport equipment, 31 Manufacture of furniture, 32 Other manufacturing, 33 Repair and installation of machinery and equipment

On the basis of the variables described above, we ran three models with the number of publications, the number of citations and the average number of citations as the respective dependent variables. As the number of publications and citations are count-variables, negative-binomial regressions were applied, which account for the peculiarities in this data structure. In the case of the average number of citations an OLS model was estimated. As explanatory variables, we use the firm size, firm age and sector (1-digit) information. Additionally, we include time-dummies in our models to control for potential period specific effects. We repeated this modelling for the manufacturing sector only, replacing the 1-digit by the 2-digit sector information to get a more differentiated picture of the publication patterns within the manufacturing sector.

Table 12: Publications, citations and citations per publication by industrial sectors

NACE Rev. 2	# publications	# citations	Citations per publication
Agriculture, forestry and fishing [A]	1%	0%	4.3
Mining and quarrying [B]	0%	0%	4.6
Manufacturing [C]	37%	38%	4.9
Manufacture of food products	2%	1%	4.3
Manufacture of beverages	0%	0%	7.0
Manufacture of tobacco products	0%	0%	n.a.
Manufacture of textiles	1%	0%	6.1
Manufacture of wearing apparel	0%	0%	5.7
Manufacture of leather and related products	0%	0%	2.0
Manufacture of wood and of products of wood and cork, except furniture	0%	0%	3.8
Manufacture of paper and paper products	0%	0%	1.1
Printing and reproduction of recorded media	1%	0%	2.1
Manufacture of coke and refined petroleum products	0%	0%	2.4
Manufacture of chemicals and chemical products	23%	28%	5.1
Manufacture of basic pharmaceutical products and pharmaceutical preparations	21%	29%	7.7
Manufacture of rubber and plastic products	1%	1%	4.3
Manufacture of other non-metallic mineral products	1%	0%	2.1
Manufacture of basic metals	3%	1%	1.9
Manufacture of fabricated metal products, except machinery and equipment	3%	2%	4.2
Manufacture of computer, electronic and optical products	25%	25%	5.9
Manufacture of electrical equipment	4%	2%	3.8
Manufacture of machinery and equipment n.e.c.	8%	4%	3.2
Manufacture of motor vehicles, trailers and semi-trailers	2%	1%	4.6
Manufacture of other transport equipment	1%	0%	2.8
Manufacture of furniture	0%	0%	1.5
Other manufacturing	4%	3%	4.9
Repair and installation of machinery and equipment	0%	0%	3.7
Electricity, gas, steam and air conditioning supply [D]	1%	1%	3.6
Water supply; sewerage, waste management and remediation activities [E]	0%	0%	2.8
Construction [F]	2%	1%	3.6
Wholesale and retail trade, repair of motor vehicles and motorcycles [G]	12%	12%	5.3
Transportation and storage [H]	1%	0%	3.2
Accommodation and food service activities [I]	0%	0%	3.3
Information and communication [J]	5%	4%	4.2
Financial and insurance activities [K]	5%	5%	4.6
Real estate activities [L]	1%	1%	5.6
Professional, scientific and technical activities [M]	27%	30%	5.6
Administrative and support service activities [N]	4%	4%	5.1
Public administration and defence; compulsory social security [O]	0%	0%	3.6
Education [P]	0%	0%	4.5
Human health and social work activities [Q]	3%	2%	5.2
Arts, entertainment and recreation [R]	0%	0%	3.8
Other service activities [S]	1%	0%	4.6
Act. of households as emp.; undiff. goods- & services-producing act. of households [T]	0%	0%	n.a.
Activities of extraterritorial organizations and bodies [U]	0%	0%	n.a.

Source: Web of Science, Hoppenstedt, queries and calculations by Fraunhofer ISI

6.4.2 Results of the Correlation Analysis and the Multivariate Models

A look at the correlation between the variables shows that citations are positively correlated not only with publication output but also with the size of the company as well as firm age (Table 13). However, both correlations are significantly negative in the case of the average citation rate. The regression models can reveal if this is a result of the also increasing publication output for larger companies as shown in the correlation analysis.

In order to look for differences in the publication behaviour, a first negative binomial regression model with the absolute number of publications as the dependent variable was estimated (Table 14). Significant coefficients with regard to the age of the companies can only be found in the model with the number of publications as the dependent variable. Therefore, it can be stated that the publication output of firms is rising with the age of the firm. The citation behaviour, on the other side, does not seem to be affected by the age of the enterprise. Significant effects can also be found for firm size. The number of publications is increasing on average with the size of a company. Large companies publish significantly more than small firms, which serve as the reference category in this model, and also receive more citations in absolute numbers. This might be a result of the availability of the respective resources in terms of research output that can be generated. However, it also could be connected to the fact that larger firms are more often outsourcing research activities, e.g. in collaboration with universities, leading to the publication of a journal article. Also signalling effects might be at play here, for example for the marketing of certain technologies. Finally, a publication can be seen as "defensive", stating the novelty of a given technology with incurring costs for obtaining a patent. This might also serve as an explanation on the significantly negative coefficient for large firms on the number of citations they receive per publication, i.e. large firms receive fewer citations per publication than small firms do.

Table 13: Correlation Analysis

	# publications	# citations	Citations per publication	Firm age (in years)
# publications	1			
# citations	0.916***	1		
Citations per publication	0.099***	0.294***	1	
Firm age (in years)	0.094***	0.071***	-0.027**	1
Firm size (categorical)	0.120***	0.088***	-0.0311**	0.450***

Significance Level: ***p<0.01, **p<0.05, *p<0.1, † marginally significant.

Source: Web of Science, Hoppenstedt, queries and calculations by Fraunhofer ISI

Considering the industrial sectors, we find various significant coefficients. In most sectors, the publication output is smaller than in manufacturing as resembled by a significantly negative coefficient. In mining and quarrying (B) and wholesale and retail trade (G), however, we find positive coefficients, indicating that the publication output here is larger than in manufacturing. With regard to the number of citations, similar effects can be observed. Most sectors receive significantly fewer citations than the firms in manufacturing. The only exceptions are once again mining and quarrying (B) and wholesale and retail trade (G) – which is not surprising since the absolute number of citations is highly correlated to the number of publications – as well as professional, scientific and technical activities (M). Looking at the citation rate reveals fewer significant industry effects, i.e. the industries do not differ so heavily in the citations received per publication than in absolute terms (number of publications and number of citations). Firms in the sector professional, scientific and technical activities (M) receive most citations per applications, which is due to the fact that these companies are rather close to science per se. Negative coefficients, i.e. lower citation rates than in manufacturing can be found in construction (F), transportation and storage (H) and information and communication (J). Especially in information and communication this can probably be attributed to the fact that firms within this sector mostly publish in more applied journals, which are – on average - less highly cited (Boyack et al., 2013). Another effect might be that conference proceedings are more frequently used than articles to advertise ongoing work and new developments in this sector.

A closer look at the manufacturing sector reveals similar effects with regard to firm age and firm size on all three outcome variables. The only difference is the negative coefficient for medium sized enterprises on the number of publications as well as citations. In manufacturing, there thus seems to be a u-shaped size effect. Mostly small as well as large firms publish – and get cited – whereas medium sized firms publish less than their respective counterparts. Once again, a large number of significant (sub-)sector effects can be observed. In terms of publications and citations, all of these effects are negative, implying that firms in the pharmaceuticals sector (the base category) have the largest output in terms of publications and citations. This is also true for the citation rate. Pharmaceutical firms receive the highest number of citations per publication, followed by manufacturers of computer, electronic and optical products (26), manufacturers of chemicals and chemical products (20), manufacturers of motor vehicles, trailers and semi-trailers (29), other manufacturing (32) and manufacturers of rubber and plastic products. In sum, we can state that there are differences in the publication behaviour of sectors. A clear distinction between high-technology and less R&D intensive sectors, however, cannot be made at the basis of these results.

Table 14: Regression results

All firms

	dV: # publications			dV: # citations			dV: Citations per publication		
	Coef.		S.E.	Coef.		S.E.	Coef.		S.E.
Firm age (in years)	0.623	*	0.325	0.416		0.510	-0.369		2.356
Firm size									
Medium sized (50-249)	0.161	***	0.047	0.105		0.086	-0.304		0.364
Large (250+)	0.859	***	0.042	1.033	***	0.080	-0.543	*	0.320
Sector									
A.	0.074		0.206	-0.216		0.388	-0.934		1.668
B.	0.597	**	0.243	0.767	†	0.479	-0.322		1.902
D.	-0.780	***	0.154	-1.629	***	0.266	-1.393		1.162
E.	-0.997	***	0.260	-1.953	***	0.396	-1.790		1.864
F.	-0.046		0.104	-0.491	***	0.188	-1.731	**	0.781
G.	0.168	***	0.055	0.276	***	0.103	0.442		0.438
H.	-0.554	***	0.203	-1.136	***	0.350	-2.695	*	1.570
I.	-0.077		0.332	-0.508		0.581	0.265		2.629
J.	-0.254	***	0.076	-0.519	***	0.135	-1.299	**	0.581
K.	-0.497	***	0.146	-0.881	***	0.255	-1.704		1.074
L.	-0.733	***	0.227	-1.056	***	0.345	-0.969		1.651
M.	0.031		0.047	0.247	***	0.087	1.386	***	0.355
N.	-0.606	***	0.134	-0.942	***	0.221	-1.403		0.968
O.	0.116		0.672	0.052		1.217	-0.507		4.539
P.	-0.651	**	0.320	-0.431		0.486	-1.330		2.358
Q.	-0.529	***	0.091	-0.517	***	0.161	0.386		0.685
R.	-1.076	**	0.536	-1.578	**	0.80	-1.269		4.065
S.	-0.446	†	0.274	-1.115	**	0.456	-1.713		1.865
Constant	-0.891	***	0.327	1.080	**	0.518	5.425	**	2.376
Year Dummies	YES			YES			YES		
Observations	9225			9225			4711		
Pseudo R²/R²	0.026			0.011			0.012		
Chi²/F	718.91			404.45			2.41		
Prob Chi²/Prob > F	0.000			0.000			0.000		

Manufacturing only

	dV: # publications			dV: # citations			dV: Citations per publication		
	Coef.		S.E.	Coef.		S.E.	Coef.		S.E.
Firm age (in years)	0.002	***	0.000	0.001		0.001	-0.005		0.004
Firm size									
Medium sized (50-249)	-0.167	**	0.084	-0.438	***	0.139	-0.781		0.559
Large (250+)	0.803	***	0.073	0.775	***	0.132	-0.712		0.504
Sub-Sector									
10	-1.580	***	0.174	-2.116	***	0.289	-3.326	***	1.135
11	-2.796	***	0.888	-2.889	**	1.188	1.038		5.885
12	--		--	--		--	--		--
13	-1.401	***	0.300	-1.422	***	0.487	-2.485		1.803
14	-1.275	*	0.772	-1.275		1.181	-0.577		4.807
15	-1.039		1.035	-2.689		1.712	-6.191		8.296
16	-1.535	***	0.525	-3.181	***	0.804	-5.675	*	3.172
17	-1.759	***	0.299	-3.093	***	0.482	-6.341	***	1.967
18	-0.448		0.340	-1.783	***	0.638	-5.963	**	2.669
19	-1.043	***	0.377	-1.602	**	0.687	-5.021	*	2.805
20	-0.461	***	0.103	-0.729	***	0.201	-3.051	***	0.764
22	-1.058	***	0.190	-1.856	***	0.334	-3.336	**	1.291
23	-1.511	***	0.195	-2.767	***	0.328	-5.622	***	1.272
24	-1.182	***	0.150	-2.662	***	0.279	-6.095	***	1.077
25	-1.074	***	0.140	-2.128	***	0.247	-4.135	***	0.939
26	-0.538	***	0.092	-0.923	***	0.179	-2.268	***	0.651
27	-0.934	***	0.150	-1.794	***	0.279	-4.400	***	1.065
28	-1.213	***	0.103	-2.075	***	0.193	-4.655	***	0.726
29	-1.133	***	0.193	-1.775	***	0.354	-3.049	**	1.410
30	-0.953	***	0.241	-2.221	***	0.441	-5.111	***	1.737
31	-1.938	***	0.739	-5.343	***	1.397	-8.044	*	4.807
32	-1.075	***	0.141	-1.778	***	0.248	-3.100	***	0.934
33	-1.253	***	0.385	-2.267	***	0.617	-4.371	†	2.673
Constant	0.436	***	0.102	2.889	***	0.194	8.588	***	0.706
Year Dummies	YES			YES			YES		
Observations	3702			3702			1958		
Pseudo R²/R²	0.058			0.028			0.044		
Chi²/F	677.43			444.28			3.31		
Prob Chi²/Prob > F	0.000			0.000			0.000		

Significance Level: ***p<0.01, **p<0.05, *p<0.1, † marginally significant.

Note: Small firms (1-49 employees and sales < 10m Euros) is the base category for the size dummies. Sector “C. Manufacturing” is the base category for the sectors. Subsector “21 Pharmaceuticals” is the base category for the subsectors.

Source: Web of Science, Hoppenstedt, queries and calculations by Fraunhofer ISI

All in all we can thus summarize the characteristics of publishing companies in regard to the introductory questions:

- There are structural differences with regard to the size, age and sectoral distribution of publishing companies: Typically, older firms publish more than younger firms. This, however, does not affect the citations received. Larger firms publish more than small and medium sized firms and publications are more common in manufacturing than in most other sectors.
- There is no clear cut between high-technology firms and their counterparts from less R&D-intensive sectors with regard to publication activity.
- The only structural differences with regard to citations could be found for the field of professional, scientific and technical activities, which is clearly in an advantage in regard to scientific credibility and dissemination and thus is also cited on average more often than other sectors.

References

Bornmann, L./de Moya Anegón, F./Leydesdorff, L. (2012). The new Excellence Indicator in the World Report of the SCImago Institutions Rankings 2011. *Journal of Informetrics*, 6, 333-335.

Boyack, K.W./Klavans, R./Patek, M./Yoon, P./Lyle, H.U. (2013): An Indicator of Translational Capacity of Biomedical Researchers. Berlin: 18th International Conference on Science and Technology Indicators, Sept. 4-6, 2013.

Chesbrough, H. (2003): *Open Innovation – The New Imperative for Creating and Profiting from Technology*. Boston: Harvard Business School Press.

Fleming, L./Sorenson, O. (2004): Science as a map in technological search. *Strategic Management Journal*, 25, 909-928.

Gondal, Neha (2011): The local and global structure of knowledge production in an emergent research field: An exponential random graph analysis. *Social Networks*, 33, 20–30.

Grupp, Hariolf/Schmoch, Ulrich/Hinze, Sybille (2001): International Alignment and Scientific Regards as Macro-Indicators for International Comparisons of Publications. *Scientometrics*. 51, 359-380.

Hicks, D. (1995): Published Papers, Tacit Competencies and Corporate Management of the Public/Private Character of Knowledge. *Industrial and Corporate Change*, 4, 401-424.

Hinze, S./Grupp, H. (1996): Mapping of R&D structures in transdisciplinary areas: New biotechnology in food sciences. *Scientometrics*, 37, 313-335.

Liebeskind, J.P./Zucker, O./Brewer, M. (1996): Social networks, learning and flexibility: Sourcing scientific knowledge in new biotech firms. *Organization Science*, 7, 428-443.

Michels, C./Fu, J. (2014): Systematic analysis of coverage and usage of conference proceedings in web of science. *Scientometrics*, 100, 307-327.

Michels, C./Fu, J./Neuhäusler, P./Frietsch, R. (2013): *Performance and Structures of the German Science System 2012* (Studien zum deutschen Innovationssystem No. 6-2013). Berlin: Expertenkommission Forschung und Innovation (EFI) (ed.).

Michels, C./Schmoch, U. (2012): The growth of science and database coverage. *Scientometrics*, 93, 831-846.

Nederhof, A. J./Meijer, R. F./Moed, H.F./ van Raan, A.F.J. (1993). Research Performance Indicators for University Departments - a Study of an Agricultural University. *Scientometrics*, 27, 157-178.

Nonaka, Ikujirō/Hirotaka Takeuchi (1995): *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford: Oxford university press.

Powell, W.W./Koput, K.W./Smith-Doerr, L. (1996): Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, 41, 116-145.

Rosenberg, N. (1990): Why firms do basic research (with their own money). *Research Policy*, 19, 165-174.

Schmoch, U. /Mallig, N./Michels, C./Neuhäusler, P./Schulze, N. (2011): *Performance and structures of the german science system in an international comparison 2010 with a special analysis of public non-university research institutions* (Studien zum deutschen Innovationssystem Nr. 8-2011). Berlin: Expertenkommission Forschung und Innovation (EFI).

Schmoch, U./Michels, C./Schulze, N./Neuhäusler, P. (2012): *Performance and Structures of the German Science System 2011, Germany in an international comparison, China's profile, behaviour of German authors, comparison of the Web of Science and SCOPUS* (Studien zum deutschen Innovationssystem Nr. 9-2012). Berlin: Expertenkommission Forschung und Innovation (EFI).

Schmoch, U./Schulze, N. (2010): *Performance and Structures of the German Science System in an International Comparison 2009 with a Special Feature on East Germany* (Studien zum deutschen Innovationssystem Nr. 9-2010). Berlin: Expertenkommission Forschung und Innovation (EFI) (Hrsg.).

Schubert, T./Michels, C. (2013): Placing articles in the large publisher nations: Is there a "free lunch" in terms of higher impact?, *Journal of the American Society for Information Science and Technology*, 64, 596-611.

Schubert, T./Rammer, C./Frietsch, R./Neuhäusler, P. (2013): *Innovationsindikator 2013*, Deutsche Telekom Stiftung; BDI (Hrsg.), Bonn: Deutsche Telekom Stiftung.

Simeth, Markus/Cincera, Michele (2013): Corporate Science, Innovation and Firm Value. Innovation and Firm Value (August 28, 2013).

Stokes, D.E. (1997): *Pasteur's quadrant: Basic science and technological innovation.* Washington DC: Brookings Institution.

Vallas, S.P./Kleinman, D.L. (2008): Contradiction, convergence and the knowledge economy: the confluence of academic and commercial biotechnology. *Socio-Economic Review*, 6, 283-311.

Waltman, L./Schreiber, M. (2013): On the calculation of percentile-based bibliometric indicators. *Journal of the American Society for Information Science and Technology*, 64, 372-379.

Appendix: Country Code list

Country	Country code
Austria	AT
Belgium	BE
Brazil	BR
Canada	CA
China	CN
Denmark	DK
Finland	FI
France	FR
Germany	DE
Great Britain/United Kingdom	GB
India	IN
Israel	IL
Italy	IT
Japan	JP
Mexico	MX
Netherlands	NL
Poland	PL
Russian Federation	RU
South Africa	ZA
South Korea	KR
Spain	ES
Sweden	SE
Switzerland	CH
United States	US