Stephan Stahlschmidt / Dimity Stephen

Comparison of Web of Science, Scopus and Dimensions databases

KB Forschungspoolprojekt 2020

October 2020

German Centre for Higher Education Research and Science Studies

Editor:
German Centre for Higher Education Research and Science Studies (DZHW) GmbH
Lange Laube 12 | 30159 Hannover | Germany | info@dzhw.eu |www.dzhw.eu
POB 2920 | 30029 Hannover | Germany
phone: +49 511 450670-0 | fax: +49 511 450670-960
Chairman of the Supervisory Board:
Ministerialdirigent Peter Greisler
Scientific Director:
Prof. Dr. Monika Jungbauer-Gans
Managing Director:
Karen Schlüter
Registration Court:
Amtsgericht Hannover | HRB 6489
VAT No.: DE291239300
October 2020

Contents

Introduction **1**

Normalised citation analysis **5**

Methodology 5

Results 6

Citation network analysis **22**

Methodology 22

Results 23

Discussion **29**

Introduction

Just as the introduction of Scopus in 2004 challenged the Web of Science's (WoS) monopoly on bibliometric databases, the recent emergence of Digital Science's Dimensions database threatens (or promises) to change the bibliometric landscape once more. Like WoS and Scopus, the Dimensions database has amassed a huge index of scientific publications, however there exists in Dimensions important foundational differences compared to its predecessors, such as the indexing of more varied content across the spectrum of research, including grants, patents and clinical trials, than the typical constraint of bibliometric databases to publications and proceedings. Further, Digital Science's position as part of the Holtzbrinck Publishing Group, which also owns the SpringerNature publishing house, has embedded Dimensions in a content-rich environment which – combined with its relatively open approach to the bibliometric community in offering free research licenses – increases the likelihood that Dimensions will establish itself in the market as an enduring product. In light of the potential impending increase in the uptake of Dimensions for bibliometric studies, it is important then to understand how the diverging coverage of Dimensions from WoS and Scopus might influence the outcomes of bibliometric analyses.

Bibliometric analyses are reflections of their underlying databases' coverage in that the coverage fundamentally defines what is included in the analysis and bibliometric evaluation further contextualises these publications against the database, which is also dependent on the coverage. Since the coverage of the three noted databases differs, we use here a quantitative comparison to determine the extent to which the bibliometric evaluation of publications is influenced by the characteristics of the respective database and how relevant these differences may be. We examine whether the databases offer a slightly varying but essentially homogeneous representation of the general citation network and therefore if the databases are (almost) perfect substitutes, or whether the databases show structurally different perspectives, and whether any potential added value of the Dimensions data can be identified through a comparison of these differences.

In order to understand the varying coverage of the three databases, the underlying philosophies of the databases should be considered. The oldest of the databases is WoS, founded by Eugene Garfield and released by the Institute of Scientific Information in 1964 as the Science Citation Index (SCI). Garfield applied his law of concentration to the SCI, arguing that the majority of significant scientific literature is published in only a small number of journals and only these journals needed to be indexed to achieve sufficient coverage of a scientific discipline [11]. Over time the SCI was expanded and became what is now WoS. The database is now owned and managed by Clarivate Analytics but continues to operate on Garfield's law, indexing only the highest quality journals in each field [7].

In 2004, Elsevier launched Scopus. Elsevier describes Scopus as the "most comprehensive overview of the world's research outputs", curated by a panel of subject matter experts [9]. Scopus is intended to be the largest possible database of research items of sufficient quality. Scopus then differs from WoS in that WoS forsakes quantity for quality, while Scopus attempts to achieve a balance of both. Both Elsevier and Clarivate Analytics offer their databases via subscription.

In 2018 Digital Science and six of its portfolio businesses released Dimensions in response to two issues they saw with current bibliometric databases. First, existing databases focus on publications as the primary scientific output, and secondly, much of the research corpus is unavailable to the research community due to proprietary applications and lack of APIs to extract data [1]. Digital Science and its collaborators created Dimensions with the intent that it would be the largest integrated database of research data available, covering the entire research spectrum from funding to academic, commercial and policy output, and made it mostly freely available to the research community [1][4].

In line with these philosophies, the three databases vary in their coverage and content in four key aspects. The first aspect is the extent and types of documents that are covered in each database. The WoS Core Collection contains documents from journals, books and proceedings Clarivate Analytics judges to be key in the life and biomedical sciences, engineering, social sciences, and arts and humanities. As of June 2020, this included over 76 million records from more than 21,000 journals and 111,000 books, and more than 8 million conference papers [6].

The foundation of Scopus was Elsevier's indexing databases, such as EMBASE and GEOBASE, from which it extracted a large number of records and then enhanced the dataset by additionally indexing the records cited in those extracted records [15]. Scopus continues to expand by indexing serial publications with International Standard Serial Numbers and single-time publications with International Standard Book Numbers from accepted sources [9]. This includes journals, trade journals, book series, books, conference proceedings, and patents in the fields of science, technology, medicine, social science, and arts and humanities [9]. As of January 2020, Scopus contained more than 77.8 million items from more than 23,400 journals, 290 trade journals, and 850 book series, including over 9.8 million conference papers and 44 million patents [9].

Dimensions was initiated by collating content from Crossref, PubMed and other openly available databases and then enhanced by processing the full-texts of items to extract references and identify links to patents, funders, and projects [4][14]. The corpus is regularly updated as new publishers contribute their collections to Dimensions and existing contributors generate new content [4]. As of April 2019, Dimensions held over 153 million items, including around 100 million publications from more than 50,000 journals, around 38 million patents, more than 4.6 million grants, 422,000 policy documents, and 455,000 clinical trials [4].

These differences in the type and number of documents each database covers are important because all three calculate citation counts from indexed items. As such citation counts are influenced by the size of the database, with larger collections likely to return higher citation counts. The document types covered may also be important given there are likely different citation practices between types, such as between a patent and a journal article, which influence the environment against which a publication is normalised.

The second aspect on which the databases differ is the accuracy of internal matching of citing items within the database, which is integral to the accurate calculation of citation counts. Dimensions maintains an extensive network of more than 4 billion citation links between its items, which are generated by matching items based on DOIs from Crossref, PubMed identifiers, and by processing full-texts for references to articles, patents, trials and projects [3][14]. However, the completeness of the network has been found to be somewhat lacking. Visser, Van Eck and Waltmann [22] determined that, when comparing pairs of citation-linked documents between databases, 10.6% of the 205 million links observed in Scopus could not be found in Dimensions, primarily because Dimensions was missing a reference list (69% of cases). This is compared to WoS, where 5.8% of Scopus' links were not found in WoS, of which only 1% was due to missing reference lists in WoS.

Scopus and WoS are not immune to missing references and inaccurate linkages. Van Eck and Waltman [21] found that missing references, incorrect metadata in references, and incorrect matching between cited and citing documents were a problem in WoS, while Scopus had issues with missing citation links between documents, despite both documents being indexed and having sufficient data to be linked. Visser, Van Eck and Waltmann [22] also noted that Scopus was missing 3.4% of the citation-linked pairs identified in Dimensions. As such, all three databases have problems with the completeness of references and linkages, however the scale of the issue appears larger in

Dimensions.

The third important difference between the databases is how content is selected. Both Clarivate Analytics and Elsevier undertake rigorous processes to assess journals on similar criteria to determine their suitability for indexing in WoS and Scopus, in addition to regular assessments to remove journals that no longer meet quality criteria [7][9]. Elsevier is more lenient, accepting approximately 30-70% of journals it assesses each month [8], whereas Clarivate Analytics accepts only around 10% [6]. In contrast, Digital Science does not make editorial decisions about what content enters Dimensions, arguing that Dimensions should be "open to integrate all relevant research objects", excluding predatory or otherwise unfit journals [4].

Clarivate Analytics' and Elsevier's selection processes have resulted in known under-coverage of particular disciplines and geographic regions in WoS and Scopus, with content heavily focused on North America, Europe and the "hard" sciences [17]. Digital Science's removal of judgement about content selection may mean that Dimensions has better coverage of these neglected regions and disciplines. On the other hand, Dimensions – in pulling from open-source databases without assessing quality – is vulnerable to unethical publishing practices, such as the uploading of multiple, low quality, self-citing preprints to repositories to increase the authors' citation rates, which may influence citation counts and the validity of evaluative results [20].

The final difference is how indexed items are classified to disciplines. Complete and accurate classification of items is essential as bibliometric evaluations typically normalise a publication against others in the same discipline to facilitate comparisons between disciplines with different publishing and citation practices. Consequently, differences in classification practices influence bibliometric outcomes.

Clarivate Analytics and Elsevier use different disciplinary classifications and methods to assign items to disciplines. However, both providers classify items according to the discipline(s) of the journal from which they originate, and both have been criticised both for the lack of transparency in how their classifications are applied and for inaccuracies in classification [23]. In contrast, Digital Science classifies each item to one or more disciplines using natural language processing and AI technology that was trained on the Australian and New Zealand Fields of Research classification. This allows classification to occur consistently over time and requires neither that the item is published in a journal, nor that the item reflects the discipline of the journal [4][14], which eliminates the common issues associated with categorising multidisciplinary journals. However, Visser, Van Eck and Waltmann [22] found that more than 16 million items in Dimensions had not been assigned a classification and Bornmann [5] found that a quarter of his own 262 papers were unclassified and many of the remaining 76% were misclassified. Digital Science has acknowledged these issues and is working to improve their classification system [13].

Given these four aspects on which the coverage and content of Dimensions, WoS and Scopus differ, it remains to be seen what impact these differences have on bibliometric analyses. As Dimensions was only released in 2018, relatively few studies have been published examining the differences between it and the two long-standing databases, however we discuss here four that we have identified.

As already noted, Visser, Van Eck and Waltmann [22] matched publications from 2008-2017 in Dimensions and WoS to Scopus to compare coverage and attributes of each database's content. They found that Dimensions contained a substantial amount of exclusive content that was not in Scopus (approximately 9.1 million items or 25% of Dimensions' content). However some of that additional content may be documents of little scientific relevance, such as editorial board lists and journal cover pages as all journal content is classified as articles in Dimensions [22]. Also, as discussed, the study identified significant issues with missing linkages and unclassified documents [22].

Based on smaller studies of subsets of data, Harzing [12], Thelwall [20], and Orduña-Malea and Delgado-López-Cózar [16] each concluded that Dimensions performed similarly to Scopus in terms of coverage and citation counts. Harzing tested Dimensions, Scopus and WoS on the retrieval of her own corpus of business and information science publications and found Dimensions had better coverage of publications than Scopus or WoS, but an equally poor rate of coverage of books and chapters. Using Google Scholar (GS) as a baseline of all available citations of her publications, Harzing [12] found that Dimensions and Scopus equally captured 46% of all GS citations and substantially more than WoS (28%). Accordingly, two-thirds of Harzing's publications had higher citation counts in Dimensions than WoS. The results were more mixed for Scopus, with 66% of publications having similar citation counts in Scopus and Dimensions, while 20% had substantially fewer citations in Scopus and 14% had substantially more in Scopus. Finally, for a selection of 6 business and ethics journals in 2009, all three databases had the same level of coverage, while Dimensions and Scopus recorded counts of citations that were both similar and higher than WoS [12].

Thelwall [20] tested Dimensions' retrieval of a sample of 84,691 food science publications in Scopus published between 2008 and 2018, alongside a random sample of 10,000 records from 2012. More than 90% of the food science articles were captured in Dimensions, and Spearman correlations between the citation counts in both databases were between 0.9 and 1.0 for all but the most recent years [20]. Thewall concluded that Scopus and Dimensions were interchangeable databases in terms of coverage and citation counts for food science publications. Further, strong correlations between citation counts in each database were also observed for the random sample, suggesting the interchangability of Scopus and Dimensions could arguably be extrapolated to the entire database.

Finally, Orduña-Malea and Delgado-López-Cózar [16] used a sample of 20 journals, 262 articles and 28 authors in library and information science to compare coverage and citations between Dimensions and Scopus. The authors found that the coverage in Dimensions exceeded that of Scopus, although Dimensions recorded fewer citations – which may result from the noted issue with missing citation links in Dimensions, strong correlations were observed between citation counts in Dimensions and Scopus – leading the authors to also conclude that Dimensions could be an alternative to Scopus [16].

We then approach our analysis within this context, that three studies of relatively small datasets suggest that Dimensions and Scopus provide similar snapshots of scientific output and its citations. However, the broader study of Visser, Van Eck and Waltmann [22] reveals substantial differences in the content of Dimensions compared to Scopus and underlying issues within Dimensions regarding citation-linkages, which are fundamental to bibliometric evaluations. Further, no study has yet compared Dimensions with WoS on a large scale, as Visser, Van Eck and Waltmann [22] compared each database with Scopus rather than to one another.

As such, we sought in this study to further examine whether Dimensions, Scopus and WoS, due to their different philosophies and content, are structurally different to one another and whether there is any added value of the Dimensions database that could be identified by comparing the differences. We do this by conducting an assessment of the normalised citation impact of duplicate publications across the three databases, and an analysis of the databases' citation networks.

In bibliometrics, priority is given to normalised indicators that evaluate a publication in relation to its environment. Due to the different coverage of the databases, environment-specific differences arise in the evaluation of the same publication. We will therefore analyse the same publications in the different databases and determine how each publication's valuation changes given the environment of the databases against which it is normalised. This varying evaluation of the same content can be used in the sense of a reverse engineering approach to illustrate the structural differences in the databases, which is informative for interpreting bibliometric analyses [17].

Normalised citation analysis

Methodology

The bibliometric data used for our analyses are sourced from the German Competence Centre of Bibliometrics' (KB) in-house versions of the three databases. For WoS, we used the Science Citation Index Expanded, the Social Science Citation Index and the Arts and Humanities Citation Index. Scopus and Dimensions databases are not organised into indices and as such, we used the relevant documents from the entire database. The Dimensions data are a snapshot of the database as of September 2019, while the WoS and Scopus data are from snapshots as of April 2019.

Discipline correspondence

As all three databases use different discipline classifications, we concorded each native classification to the OECD's Fields of Science and Technology (FOS) classification so that results at the discipline level are comparable between databases. The FOS consists of 42 classifications within 6 broad groups of natural sciences, engineering and technology, medical sciences, agricultural sciences, social sciences, and humanities. We concorded Dimensions data from the Fields of Research 4-digit level to the FOS based on a concordance produced by the Australian Bureau of Statistics [18]. WoS and Scopus data were mapped to the FOS from the "traditional" WOS classification and Scopus' All Science Journal Classification (ASJC) based on concordances provided by Clarivate Analytics and Elsevier.

None of the three concordances allocate native disciplines to all of the FOS disciplines. As such, no Dimensions publications are allocated to the FOS disciplines *Other Natural Sciences* or *Other Social Sciences*. No publications in Scopus are allocated to *Environmental Biotechnology*, *Industrial Biotechnology*, *Nano-technology*, or *Agricultural Biotechnology*, and no publications from WoS are allocated to *Health Biotechnology*, *Other Medical Sciences*, or *Agricultural Biotechnology*. Further, the ASJC contains a *Multidisciplinary* category that is not mapped to any FOS discipline, and so these publications are excluded from discipline-based analysis. We also excluded all unclassified publications.

Treatment Effects

To undertake our analyses, we sought to identify all articles published in 2016 affiliated with German institutions that were indexed in all three databases, referred to as "duplicate" publications. As a first step, we retrieved the WoS-Scopus duplicates, which were identified by comparing hash values on a subset of metadata strings between the two databases. We then matched the DOIs of these duplicates from Scopus to the DOIs of German articles published in 2016 indexed in Dimensions to identify the duplicates in all three databases. We used DOIs from Scopus as the basis of the match as Scopus had fewer missing DOIs than did WoS. We validated the DOI-based matches by calculating the Jaro-Winkler distance on the title strings between the three publication versions.

Having identified the duplicates in all three databases, we then analysed the variations in duplicates' normalised citation impact. To assess the effect of the database's environment on the normalised citations, we calculated for every German duplicate from each database the number of citations the article received in the 3-year citation window 2016-2018 (observed citations), and the average number of citations received in the 3-year citation window by all articles published in 2016 that were allocated to the same discipline (expected citations). For exploratory purposes, we then calculated the differences in both observed and expected citation counts between databases aby simply subtracting the citation counts of one database from another. From this, we calculated the average differences in observed and expected citation counts between databases per discipline. Finally, we calculated the normalised citations and the differences in normalised citations between databases as:

$$\Delta \, norm. \, cit. = \frac{obtained \; citations_i^{(s1)}}{expected \; citations_i^{(s1)}} - \frac{obtained \; citations_i^{(s2)}}{expected \; citations_i^{(s2)}}. \qquad (1)$$

where i is each German duplicate, and s is the source of the citation data, i.e. WoS, Scopus or Dimensions. In normalising the observed citations in each database against the expected citations in the same database, we achieved a database-specific valuation of each article. The content of the database is influential here as the inclusion or exclusion of particular articles in the corpus may influence both the citations received by the article and the average citations received by all articles in the discipline, affecting the ratio between the two observations.

Using each source's normalised citations, we then calculated the difference Δ between sources. For each measure of difference between databases, we obtained three sets of results – differences between WoS and Scopus, between WoS and Dimensions, and between Scopus and Dimensions. With these measures, we can examine how the same publication can be valuated differently between databases due to the database's environment. To facilitate identification of the structural differences in the databases and their influence on the article-specific valuations, we aggregated the data to the disciplinary level and present the results for disciplines with at least 200 publications. We also examined the differences in normalised citations by sector of the German science system to identify potential macro-level effects of the database on this bibliometric indicator.

Results

We first present some preliminary exploratory analyses of the databases' overall coverage of disciplines and number of exclusive or duplicate German publications. We then describe the results of the matching procedure between WoS, Scopus and Dimensions publications, followed by the results of the normalised citation analysis and the citation network analysis.

Preliminary exploratory analyses

There were 113,227 articles and reviews published in 2016 that were affiliated with German institutions indexed in WoS, 127,542 in Scopus, and 118,688 articles in Dimensions. To gauge the overall number of exclusive and duplicate German publications in each database, we first removed publications in each database that did not have a DOI (3.8% in WoS, 5.3% in Scopus, and 0.4% in Dimensions), and then matched the remaining publications based on the DOIs in all three databases. Figure 1 shows these results[1]. We identified 83,546 publications present in all three databases, representing 76.7% of the relevant publications in WoS, 70.7% in Dimensions, and 69.1% in Scopus.

[1] In subsequent analysis we found that the 20,874 publications identified in WoS and Scopus but absent from Dimensions occurred as many of these items are missing GRID information and so they were not identified as German publications when querying Dimensions. Publications in this category likely number closer to a few hundred than 21,000.

Dimensions had the largest share of exclusive publications with 20% of its content not contained in WoS or Scopus, compared to Scopus's exclusive share of 6.3% and WoS' 2.1%.

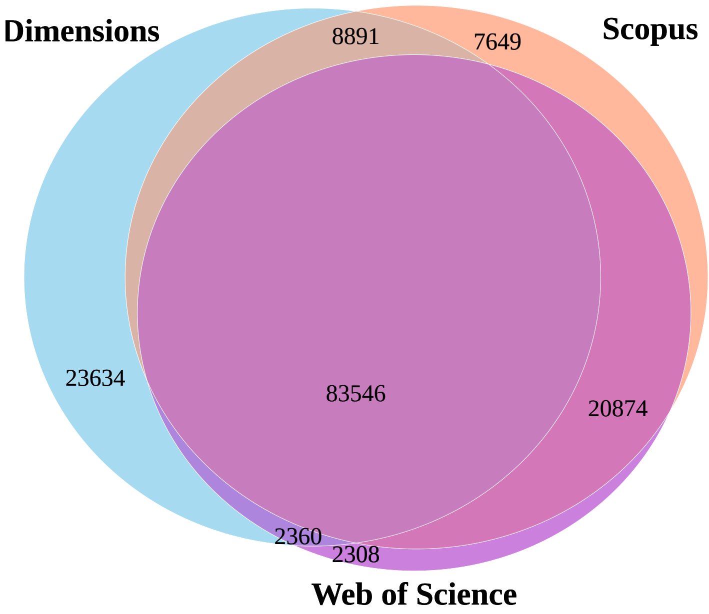

Figure 1: The number of exclusive and duplicate German 2016 publications between WoS, Scopus and Dimensions.

To examine the coverage of the databases by discipline, we show in Figure 2 the number of articles and reviews in Scopus and WoS and articles in Dimensions published in 2016 by FOS discipline. Some double-counting will occur as publications can be assigned to multiple disciplines. This overview of the disciplinary structures of the databases demonstrates the environment against which publications are being normalised.

Overall we see that all three databases had stronger coverage of the natural sciences, medical sciences, and engineering and technology disciplines than they did of the social sciences and humanities. Further, typically Scopus had the highest coverage while WoS had the least. However, Dimensions had the largest coverage of most disciplines in the social sciences and all disciplines in the humanities except the catch-all *Other humanities*. This may reflect that the FOR-FOS correspondence does not map Dimensions data to the *Other social sciences* category, while this category is strongly utilised in the Scopus data, likely due to the mapping of ASJC categories such as *General social sciences* and *Social sciences (miscellaneous)* to the *Other social sciences* category. Such differences reflect the specificity of the native classifications as the FOR classification applies 1,239 categories at its finest level to avoid generalised categories, compared to 334 categories in the ASJC and 254 in WoS.

Some disciplines with large differences emerged: both Scopus and Dimensions held substantially higher numbers of *Clinical medicine* publications than WoS, and Scopus held approximately double the *Materials engineering* publications and nearly treble the publications in *Chemical engineering*, *Health biotechnology*, *Agriculture, forestry, and fishery*, and *Animal and dairy science* than in either WoS or Dimensions.

Some of these differences arose due to genuinely larger coverage in one database. For instance, *Agriculture, forestry, and fishery* contained very similar categories of the three native classifications, however Scopus' 27,072 publications in *Agronomy and crop science* vastly outnumbered the corresponding 11,913 *Crop and pasture production* publications in Dimensions and 9,277 *Agronomy* publications in WoS. However, other differences in coverage occurred due to differences in mapping the native classifications to the FOS structure. For example, Scopus' much larger coverage in *Animal and dairy science* occurred because Elsevier maps the ASJC category *Animal science and zoology* (27,390 publications) to both FOS *Biological sciences* and *Animal and dairy sciences* disciplines, while Clarivate Analytics and Digital Science separate these categories: their *Zoology* categories (12,527 publications in WoS, 11,508 publications in Dimensions) map to *Biological sciences* and WoS' *Agriculture, dairy & animal science* (7,212) publications and Dimensions' *Animal production* (6,231) map to *Animal and dairy sciences*. Further, Elsevier also includes Scopus' 29,265 *Food sciences* publications in *Animal and dairy sciences*, whereas this field is mapped to *Other engineering and technologies* in both Dimensions (12,309) and WoS (23,128). These differences should be considered in interpreting results at the disciplinary level.

Results of matching procedure

For the remaining analyses, we used the KB-matched records between WoS and Scopus. The KB WoS-Scopus matching procedure identified 107,800 of the 113,227 WoS publications (95.2%) and 127,542 Scopus publications (84.5%) as duplicates. Using these records, we then examined the DOI data to ensure matching the WoS-Scopus pairs to Dimensions on DOI would produce reliable results. Of the 107,800 WoS-Scopus duplicates, 1.7% (1,846) Scopus publications and 2.7% (2,905) WoS publications were missing DOIs, as were 495 (0.4%) of the Dimensions records. We excluded publications with missing DOIs from further analysis. Checks revealed that 128 (0.12%) DOIs in Scopus were assigned to 2 records, as were 109 (0.1%) DOIs in WoS and 21 (0.02%) DOIs in Dimensions. In most cases, this occurred as the publication had two entries in the database under different identifiers, but with otherwise identical information. To avoid potential issues with misassigned citation counts, we removed both entries of these records. We also found that 471 (0.4%) WoS-Scopus duplicates had different DOIs recorded between the databases. These mismatches typically occurred as one database recording a letter 'l' or 'o' as numbers 1 or 0, for example, or inclusion of spaces or special characters in one database. As such, we did not remove any records where the DOI was mismatched between WoS and Scopus.

As the DOI data were of sufficient quality, we proceeded to match the WoS-Scopus duplicates to the Dimensions publications using Scopus and Dimensions DOIs. We obtained from this process 84,332 records that were present in all three databases. This represents 74.5% of the total 2016 German publications in WoS, 66.1% in Scopus and 71.1% in Dimensions.

To ensure the validity of the DOI matching, we analysed the similarity of articles' titles across the databases by calculating the Jaro-Winkler distance on the combinations of title strings. In all three comparisons between databases, the distances between titles ranged from 0.0 to 0.73 with a mean difference of 0.02. We examined the 3,879 (4.6%) matches with distances above 0.25 and found that the higher distances resulted from different encoding of special characters, such as in scientific names or formulae, the inclusion of subtitles, or use of a German title in one database. As such, the low mean distance between titles and acceptable reasons for higher distances suggests the matching of publications based on DOI is valid.

As a final step in the matching process, to ensure comparability between the normalised citation counts of the same record between databases, we included only articles (i.e. excluding reviews) that were allocated to the same FOS discipline in all three databases. Notably, this requirement substantially reduced the number of articles available for analysis, as the different discipline classification

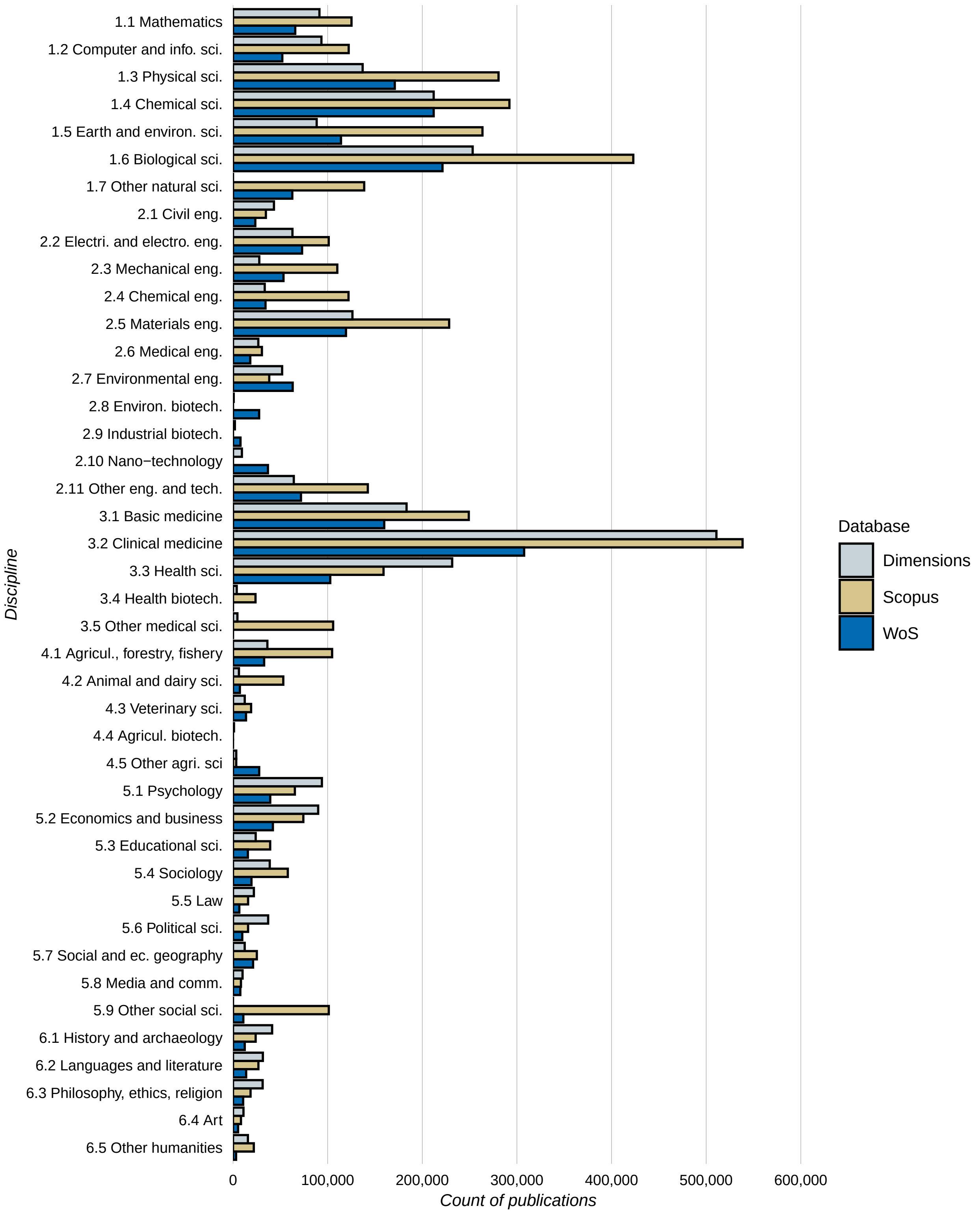

Figure 2: The number of articles and reviews (WoS, Scopus) and articles (Dimensions) published in 2016 by discipline and database

methods between the databases – article-level in Dimensions and journal-level in WoS and Scopus – and different correspondences from native classifications to FOS meant many articles were classified differently. As such, in total, 41,848 German articles published in 2016 and indexed in all 3 databases were included in our normalised citation analysis.

Normalised citation analysis

As a preliminary examination of the differences in citations between the databases, we present on the left-hand side of Figure 3 the observed citation counts of the German duplicate articles, representing the numerators in Equation (1). On the right-hand side are the expected citations of articles in each FOS discipline based on the broader WoS, Scopus and Dimensions databases, representing the denominators. Each panel compares two of the three databases and the diagonal line represents a perfection correlation.

We see in each panel a positive correlation between each database, suggesting all three databases present a relatively similar picture of both observed and expected citations, and the broader bibliometric landscape. However there was less variation between Dimensions and WoS and Dimensions and Scopus than between Scopus and WoS in the observed citation counts: the mean difference between Scopus and WoS $\overline{\text{obtained citations}^{SCP} - \text{obtained citations}^{WoS}}$ was 1.1, compared to $\overline{\text{obtained citations}^{SCP} - \text{obtained citations}^{DIM}} = 0.3$ between Scopus and Dimensions and $\overline{\text{obtained citations}^{DIM} - \text{obtained citations}^{WoS}} = 0.8$ between Dimensions and WoS. The mean differences in expected counts showed a different pattern, with greater variance between Scopus and Dimension $\overline{\text{expected citations}^{SCP} - \text{expected citations}^{DIM}} = 0.6$ than between and Scopus and WoS $\overline{\text{expected citations}^{SCP} - \text{expected citations}^{WoS}} = 0.4$ and between Dimensions and WoS $\overline{\text{expected citations}^{DIM} - \text{expected citations}^{WoS}} = -0.2$. It is these variations from the linear trends in each panel due to the exclusive content of each database that generate changes in the ratio of observed to expected citations, which translates to variations in the normalised citation impact of publications between databases.

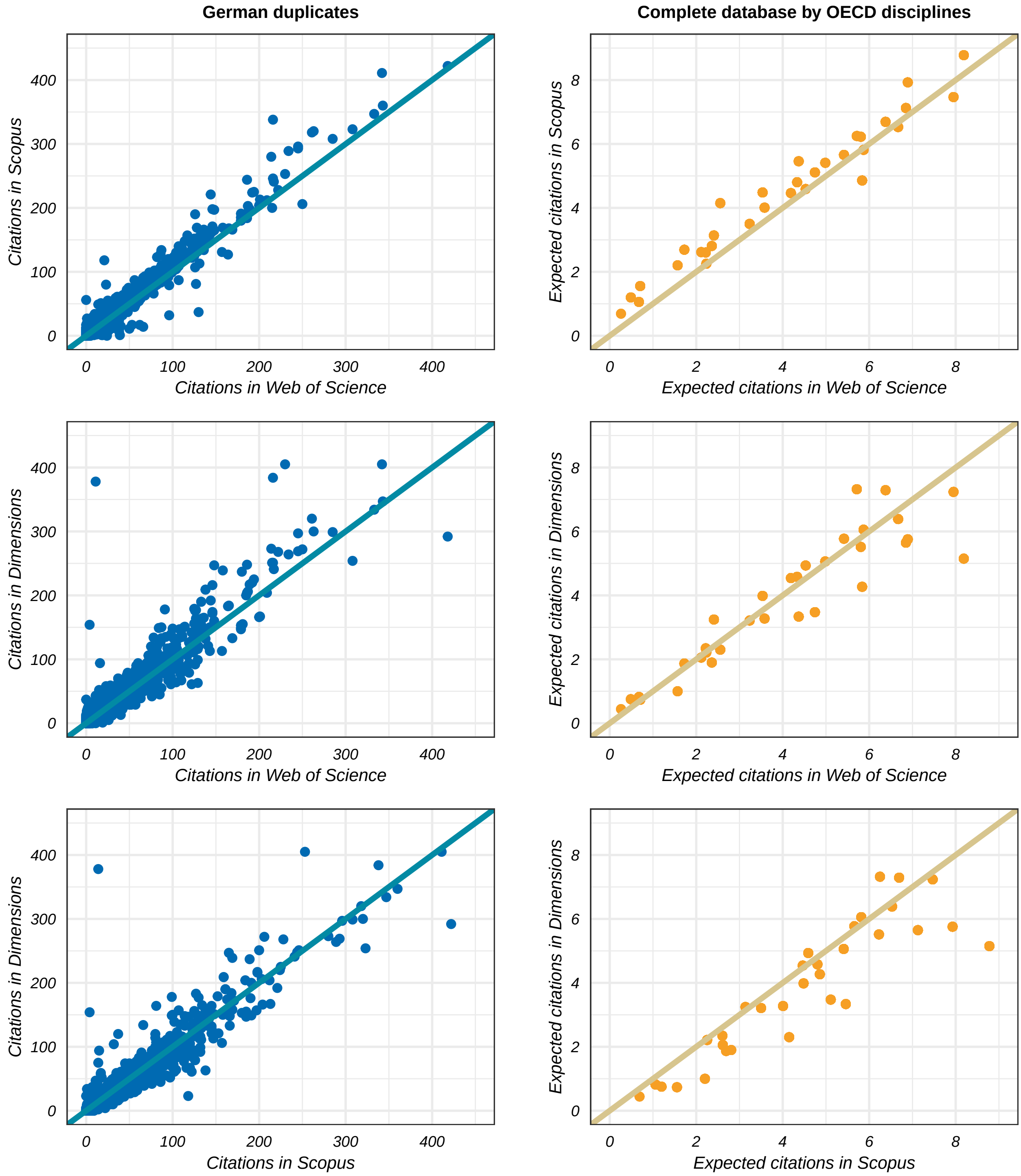

Figure 3: Scatterplots of the absolute counts of citations of duplicate articles (left) and average citations by discipline (right) in each database.

To further examine the differences in observed citations, we present in Figure 4 the differences in citations between the databases for each discipline with more than 200 publications. For each discipline, we show the distribution of the percentage of German duplicates with Δ difference in the number of observed citations between each combination of databases. Greater skewness in the distribution represents greater diversity in the observed citation counts between two databases.

In conjunction with Figure 4, we present in Figure 5 the average differences in each discipline's distribution of observed citations in the left-hand panels, and the average difference in expected citations in the right-hand panels.

In Figure 4, we see across all disciplines that the majority of duplicate publications received the same number of citations in each of the databases. Notably however, when comparing WoS and Scopus, the duplicates in all disciplines tended to receive more citations in Scopus than they did in WoS, as shown by the right-tailed skew in Figure 4 and the positive change in average observed citations in Figure 5. This was particularly pronounced for *Computer and information sciences* and *Economics and business* in which the duplicates received on average 2.6 and 1.9 additional citations respectively in Scopus compared to WoS. The differences in average expected citation counts were mostly small, with differences of less than 0.5 citations for 12 of the 19 disciplines, typically favouring high expected counts in Scopus. Only *Clinical medicine* was particularly higher in WoS than Scopus, with one additional citation expected.

Similarly, duplicates in most disciplines on average received more citations in Dimensions than they did in WoS. As with Scopus, *Computer and information sciences* and *Economics and business* received the largest increase in citations in Dimensions over WoS; on average 2.3 and 2.1 additional citations. Conversely, duplicates in *Chemical sciences*, *Chemical engineering*, and *Materials engineering* on average received higher citations in WoS, although the latter two disciplines received only an additional 0.3 and 0.2 citations. The differences in average expected citations, reflecting the broader content of the databases, were more mixed. For instance, the applied *Clinical medicine* discipline had on average 1.6 expected citations higher in WoS, while conversely *Basic medicine* had an average 1.6 expected citations higher in Dimensions.

Scopus and Dimensions presented a relatively similar picture of observed citations, as most disciplines recorded differences in average citations of less than 0.5 citations. The exceptions were *Chemical sciences*, *Chemical engineering*, *Materials engineering*, which received 2.1, 1.31, and 1.26 additional citations in Scopus, and *Biological sciences*, which received 0.8 additional citations in Dimensions. The average expected citations were also similar across most disciplines, with again only *Chemical engineering* and *Materials engineering* showing substantial increases of 2.2 and 1.5 additional citations in Scopus, and *Basic medicine* had on average 1.1 additional citations in Dimensions.

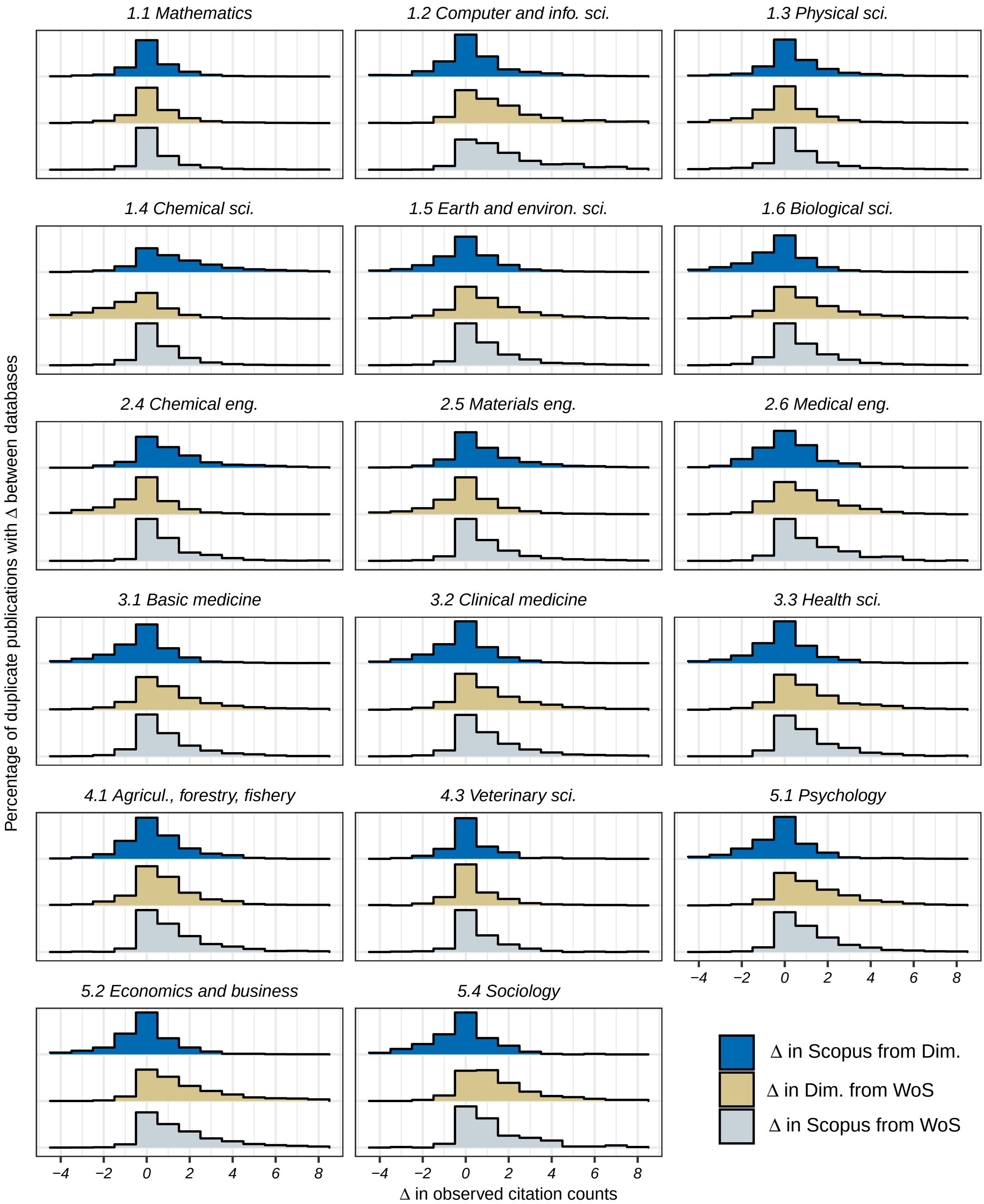

Figure 4: Density plots of the percentage of duplicate publications with each difference in observed citations between the three databases by discipline.

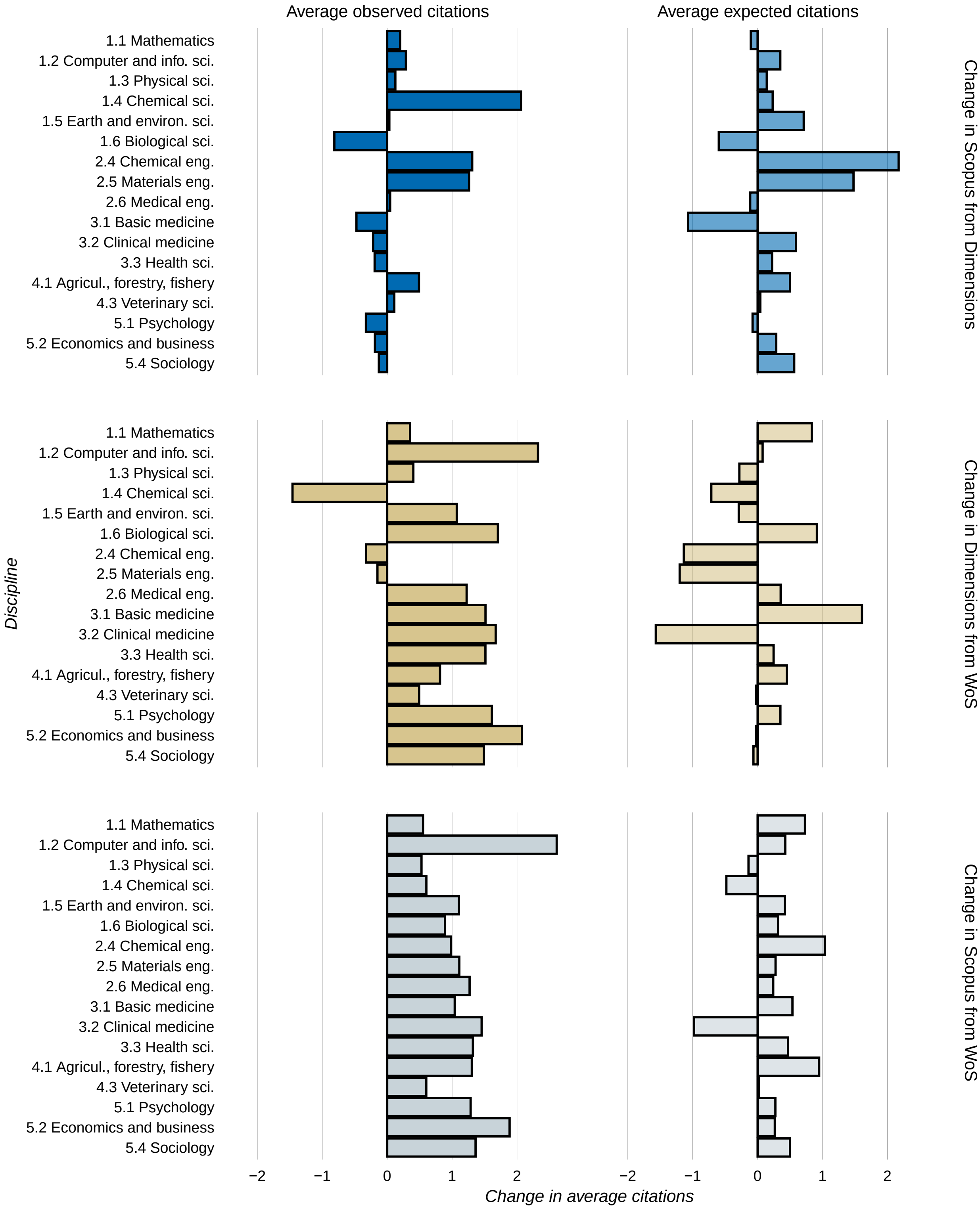

Figure 5: Differences in the average observed citations (left) and average expected citations (right) between the databases per discipline

We show in Figures 6, 7, and 8 the distributions of the differences in normalised citations by discipline between the three databases. The x axis shows the difference in normalised citations between two databases examined in bands of 0.2 and the y axis shows the percentage of duplicate publications with a particular change in normalised citations, while the orange line shows the average of the distribution. Given a normalised citation score of 1 indicates a publication achieved exactly the expected citations for its discipline, a change of 0.1 can be interpreted as a 10% change in normalised impact and as such here a change of 0.1 from 0 indicates a 10% change in the impact of a publication between databases.

Comparing WoS and Scopus in Figure 6, we see that the majority of publications had similar normalised citation scores in both databases. For instance, the impact of approximately 60% of all publications in *Physical sciences* and *Chemical engineering* was different by only 0 and 20% between the databases. The range of differences observed was -64 and 43.8, however only 1.6% of the 41,848 publications examined had differences in normalised citations of more than 2 (200%). On average, the duplicates' impact was slightly higher in Scopus than WoS; publications in most disciplines achieved an average increase in impact of between 2.9% and 25.2% in Scopus. Notably, *Clinical medicine*, *Economics and business* and *Computer and information sciences* fared considerably better in Scopus, increasing their impact by 57.1%, 44.3%, and 40.4% respectively. Only publications in *Agriculture, forestry, and fishery* and *Mathematics* were adversely affected in Scopus, with these disciplines decreasing in impact by 11.0% and 8.3% compared to WoS.

Comparing Scopus and Dimensions in Figure 7, once again most duplicate publications had normalised citation scores within 20% of each other. The range of differences between normalised citations was -155.9 and 16.2, and 1.6% publications had differences greater than 2. On average, 11 of the 17 disciplines we examined had higher citation impact in Dimensions than they did in Scopus. In particular, impact was most greatly improved in Dimensions for publications in *Sociology* (55.1%), *Clinical medicine* (31.6%), *Economics and business* (21.1%), and *Earth and environmental sciences* (19.5%). Conversely, the only disciplines where publications were favourably assessed in Scopus were *Chemical sciences* (24.0%), *Basic medicine* (17.4%), *Mathematics* (9.5%), *Medical engineering* (3.4%), and *Biological sciences* (1.4%), and *Veterinary sciences* (1.1%).

Finally, in Figure 8 we compare Dimensions and WoS. We observed differences in normalised citations of between -36.6 and 156.3, and 3.1% of publications had differences in normalised scores of more than 2. Duplicate publications in the majority of disciplines had improved normalised impact scores in Dimensions over WoS, with *Clinical medicine* (88.8%), *Sociology* (77.5%), *Economics and business* (65.5%), and *Computer and information sciences* (44.4%) particularly strongly affected. Conversely, duplicate publications in only 4 disciplines had higher citation impact on average in WoS: *Mathematics* (17.8%), *Basic medicine* (14.6%), *Chemical sciences* (7.9%), and *Agriculture, forestry, and fishery* (0.9%).

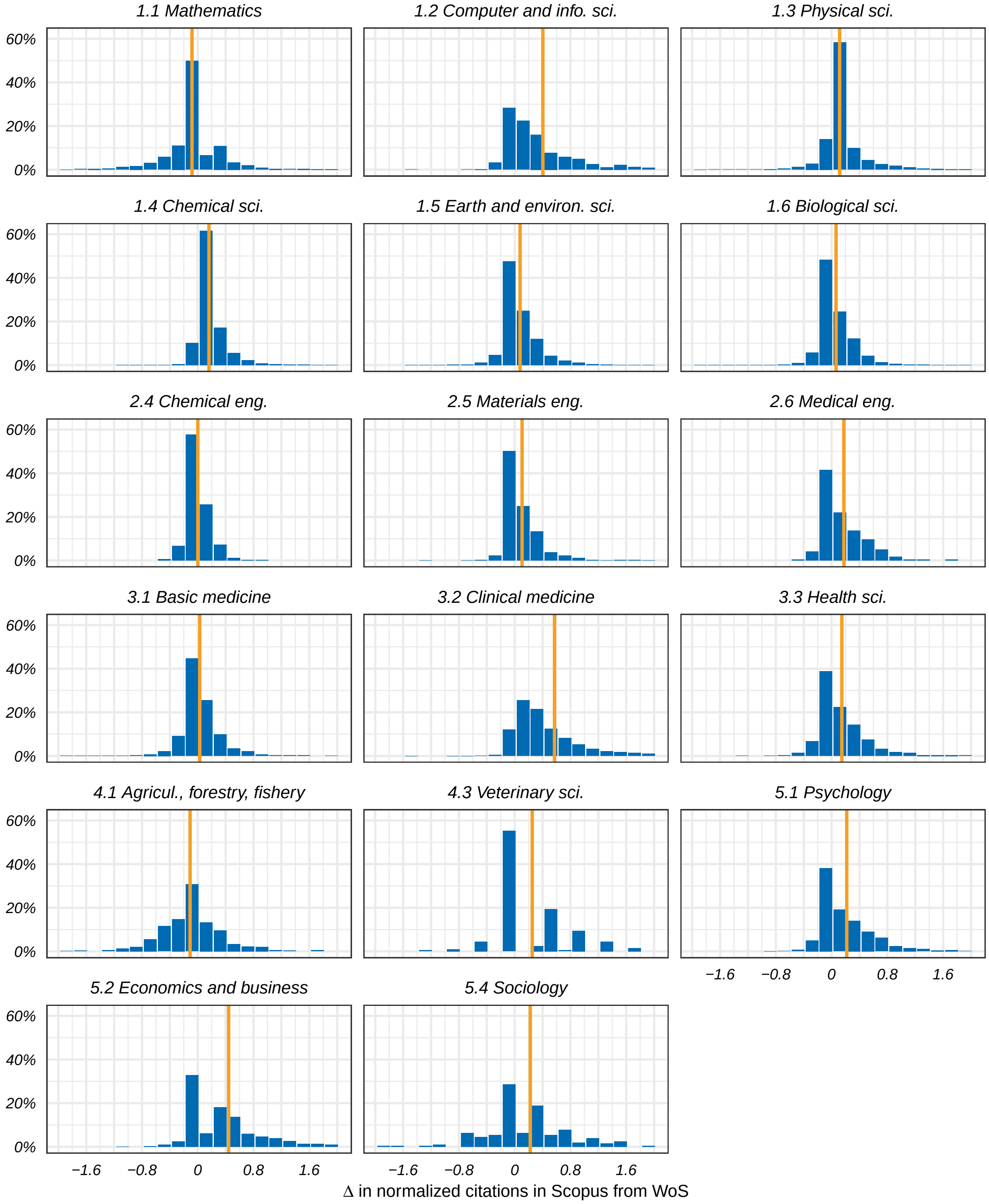

Figure 6: Distribution of differences in duplicates' normalised citations between Scopus and WoS by discipline

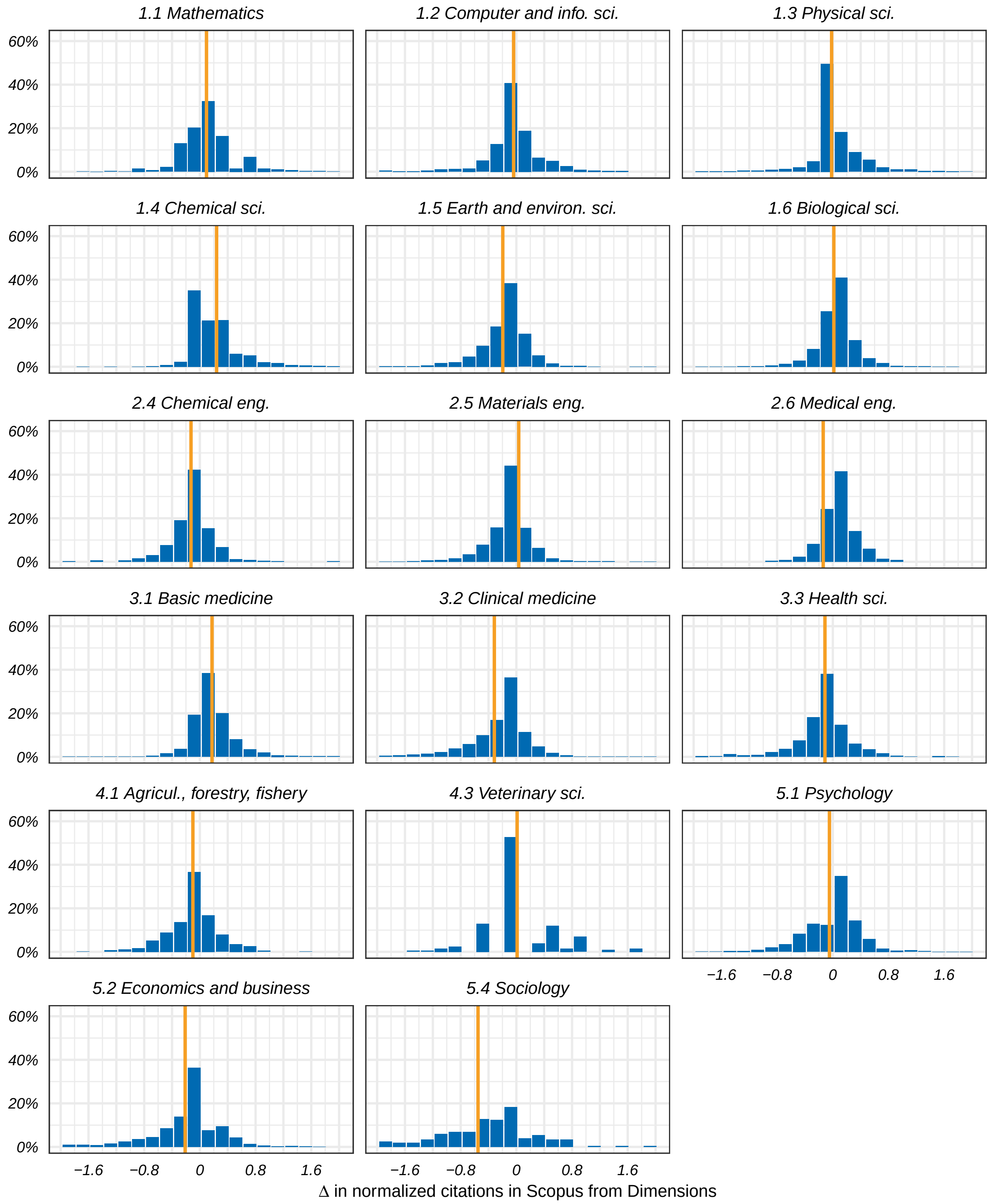

Figure 7: Distribution of differences in duplicates' normalised citations between Dimensions and Scopus by discipline

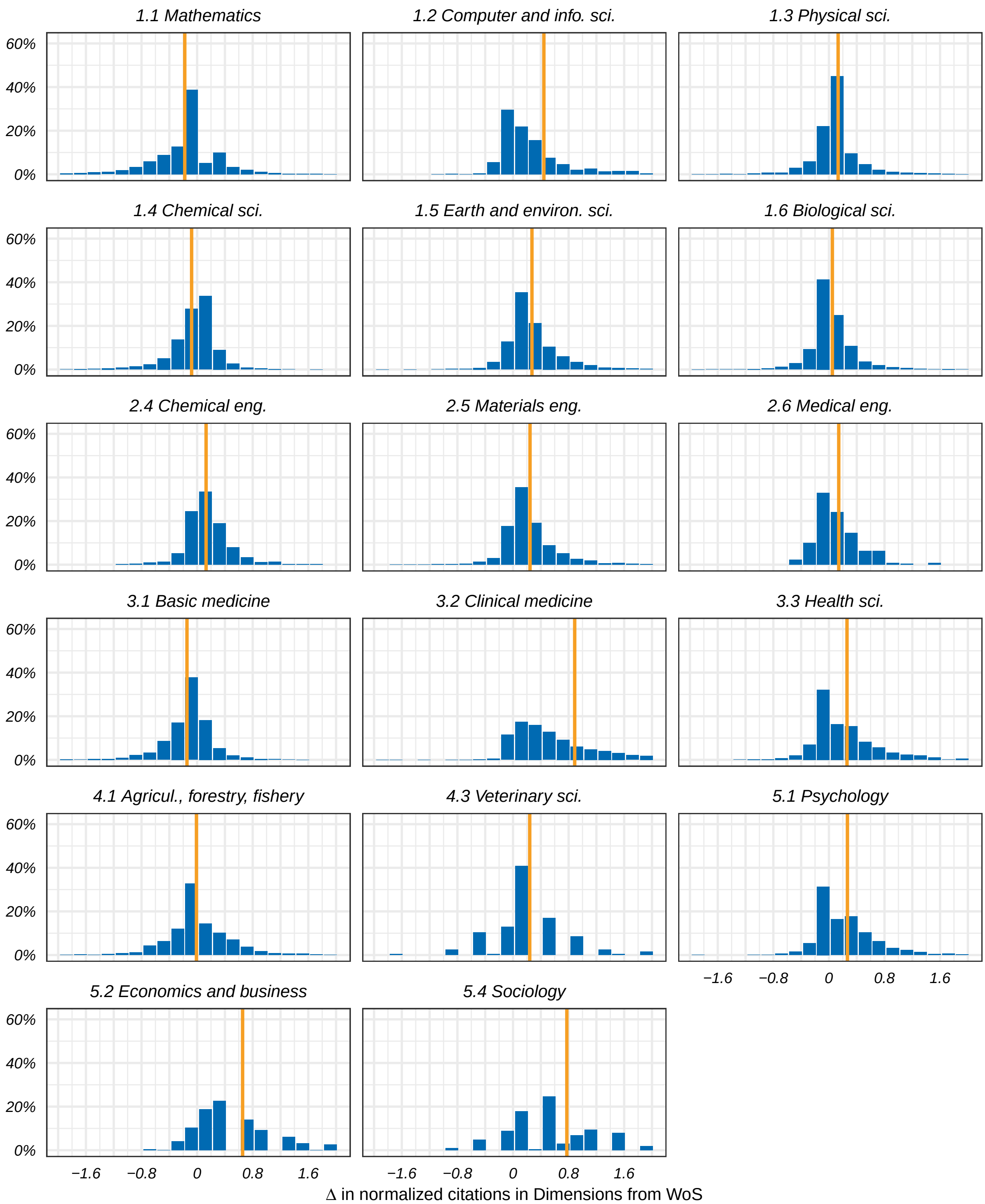

Figure 8: Distribution of differences in duplicates' normalised citations between Dimensions and WoS by discipline

Differences in normalised citations for sectors of the German science system

To identify potential macro-level effects of the structural differences of the databases on normalised citations, we present in Figure 9 the aggregated differences in normalised citations between databases by six sectors of the German science system. Data are presented as the density in the distribution of differences in each sector's 2016 duplicate publications, and also divided into quintiles. The sectors are the Leibniz Association (WGL), the Max Planck Society (MPG), higher education institutions (HEI) including universities and colleges of applied science, the Helmholtz Association (HGF), the Fraunhofer Society (FhG), and institutions affiliated with the German business sector (Economy).

These sectors each have particular research profiles. In the HEI group, the universities undertake both teaching and research across all disciplines, while the colleges focus on technical application in specific areas. The HGF has a health, energy, earth and physical sciences orientation to its research. The WGL conducts both basic and applied research in social sciences, health, natural sciences, mathematics and engineering, while the MPG conducts primarily basic research, and the FhG focuses on applied research.

We see in the bottom panel of Figure 9 that the higher normalised impact in most disciplines in Scopus translates to a skewed distribution at the sector level toward higher impact in Scopus than WoS. The FhG and Economy sectors benefit most strongly from Scopus' content, where the central quintiles received an increase in impact of between 3–12% and up to approximately 25% for the middle-high quintile. The effect on the remaining 4 sectors is similar, although the MPG gains slightly less impact in Scopus than the other sectors and its lower-middle quintile is reduced in impact by up to approximately 6%.

All sectors also appear to benefit in impact from assessment in Dimensions compared to WoS, as shown in the middle panel, although the skew is not as dramatic as between WoS and Scopus. The increased impact is nearly uniform across sectors, with the central quintile improving by between 0% and 6%, and up to around 20% for the middle-high quintile in all sectors. The Economy sector receives the largest boost in impact through the use of Dimensions over WoS. Conversely, the 20% of publications in the lower-middle quintile in all sectors lose up to approximately 10% impact in WoS.

There is less evidence of skew in the distribution of differences between Scopus and Dimensions in the top panel. Notably, the central quintile in all sectors is quite narrow – within a band of approximately 5% – and centred on zero, indicating little variation in valuations of duplicate publications between the two databases. There is in particular a slight improvement in impact for the MPG and FhG, with a greater percentage of publications in these sectors gaining impact than losing, however the improvements are relatively minor.

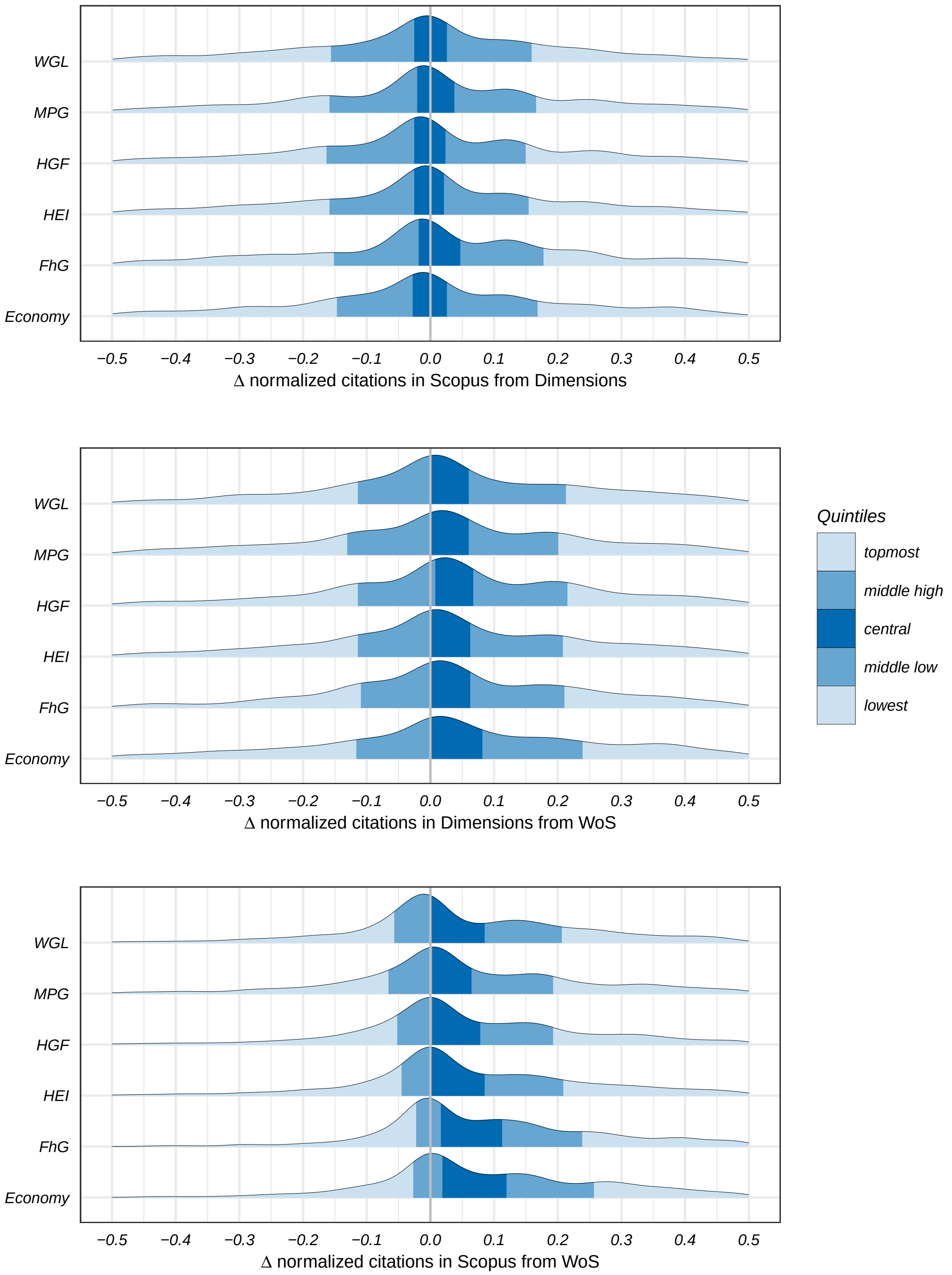

Figure 9: Distribution and quintile of differences in duplicates' normalised citations between data-bases by German sector

Finally, to examine the overall differences in sectoral valuations between the databases, we present in Figure 10 the summed differences and average differences between normalised citations of each sector's duplicate publications between databases. While the summed differences show the total difference in citations for each sector, the average highlights the relative impact of the differences in databases on each sector after accounting for its level of output. For instance, the additional 2,423 citations attributed to the HEI sector in Dimensions over Scopus was far greater than the increases of between 19 and 415 for the remaining sectors, however in terms of impact of the databases' differences, across all sectors on average less than 0.1 separated the normalised citations. As such, overall, sectoral valuations in Scopus and Dimensions were similar.

Comparatively, there was greater variation in Scopus-WoS and Dimensions-WoS valuations than Scopus-Dimensions. Dimensions-WoS were most dissimilar, with higher valuations in Dimensions than WoS of on average between 0.22 and 0.41 citations across the sectors. This translated to between 205 (FhG) and 11,289 (HEI) additional citations over these sectors' corpora. The Economy sector valuation improved most notably via use of Dimensions over WoS (average difference = 0.41), followed by the HEI, MPG and HGF (0.27). The Economy sector also had the largest average difference in citations between WoS and Scopus at 0.35, followed by FhG (0.24) and HEI (0.22).

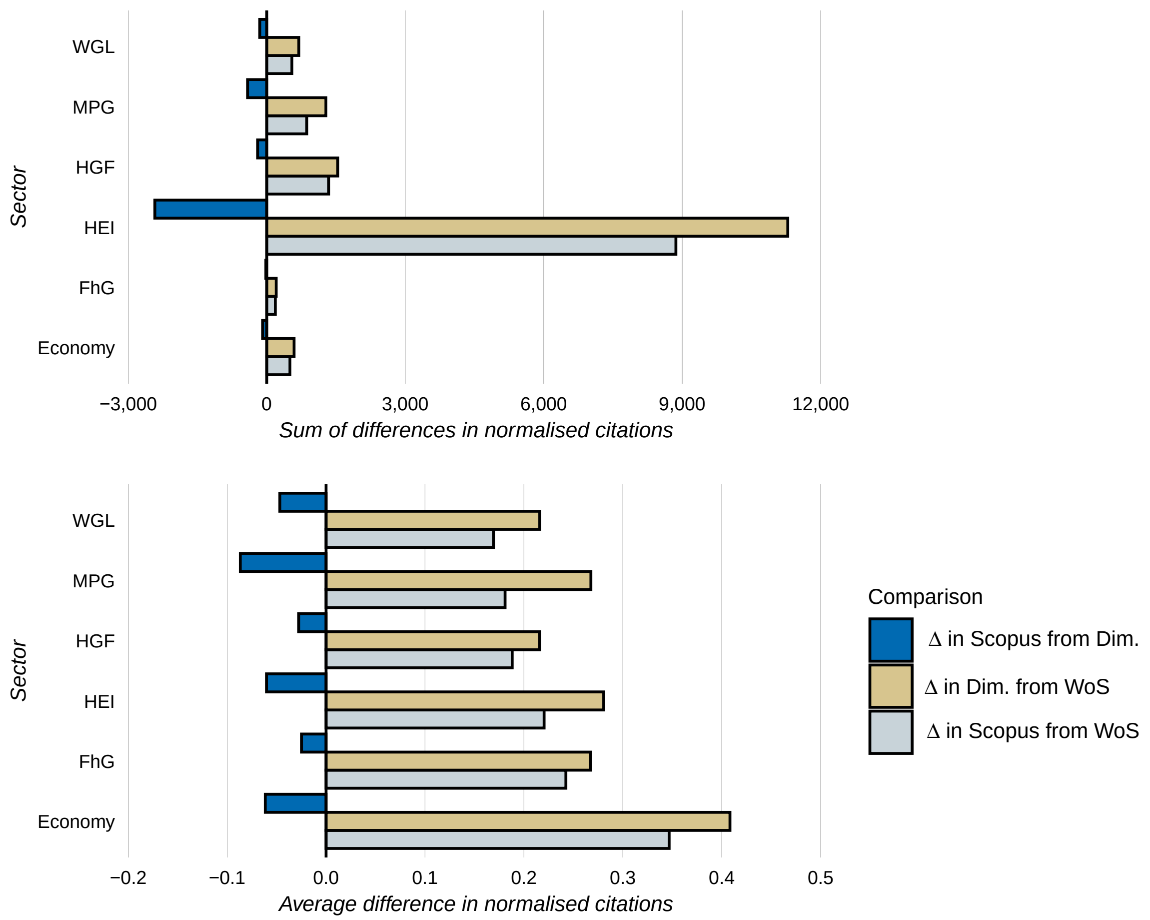

Figure 10: Average and sum of differences between databases in normalised citations by sector

Citation network analysis

Previous studies [22] show that the divergent coverage between the databases results in different subsets of publications indexed in only one, two or all three databases. The power set of these subsets consists of the average set (set of publications and citation links between them, which are uniformly contained in all databases), partial intersections of publications and related citations in two of the three databases, and residual sets of exclusive publications and related citations indexed in only one database. In this section, we will investigate the role of these subsets for themselves and other subsets. With respect to citation analyses, role is understood as relevance and impact, which can be measured by reducing (or assuming stochastic determination of further functions) citations to their Mertonian ideal ("give credit, where credit is due"), by citations between and within these subsets. Citations between subsets represent how information flows and how the publications indexed exclusively in Dimensions are embedded in this information flow. The role of these publications in the overarching citation network, which is unknown in its population, can be identified and the added value of the Dimensions data as a reflection of previously unobserved scientific communication can be approximated.

Methodology

The approach is implemented by the indicators (1) out/in-degree, as a network analytical perspective on the micro level of publications, which is translated to the level of the subsets described above by representing the respective distributions, and (2), as far as possible the indicator internal coverage, as an aggregated value on the database level. With this approach, in addition to the expected increased coverage, i.e. the pure "more" of communication, also communicative characteristics of the Dimensions data and thus differences observed in the normalised citation analysis can be partially quantified.

As before we restrict our analysis to publications published in 2016 and account for a three year citation window 2016-2018. Due to the general interest to observe differences between the databases we do not restrict the set of publications to German ones, but make use of all 2016-2018 articles and reviews indexed in the three databases.

While also WOS and SCP assign the same document to different document types, the Dimensions database differs in this respect notably. In detail, Dimensions holds more documents of type article in the respective years than the intersection of core publications jointly indexed by WOS, SCP and DIM does. However, most of these documents do not include any references and Dimensions has been observed to assign the article type to all journal publications [22]. Contrary only 1% (4%) of WOS (SCP) articles and reviews from 2016-2018 hold no source reference and only 0 (9) articles hold no reference at all.

Hence, we restrict our comparison to articles and reviews indexed in SCP and/or WOS and to articles in DIM with at least one reference. Thus for DIM we improve the separation between substantial scientific contributions building upon and highlighting former contributions via references and other journal content to improve the validity of the database comparison. Still the stated requirement for DIM articles constitutes a lower bound rather than a solution to the insufficient document type classification in DIM, but as noted before also WOS and SCP disagree sometimes upon the document type.

We join the three databases WOS, SCP and DIM via an exact string matching procedure based on DOI identifiers. DOIs uniquely identify publications and are therefore suited for matching purposes. But DOIs recorded in bibliometric databases have partially been observed to include errors [2],

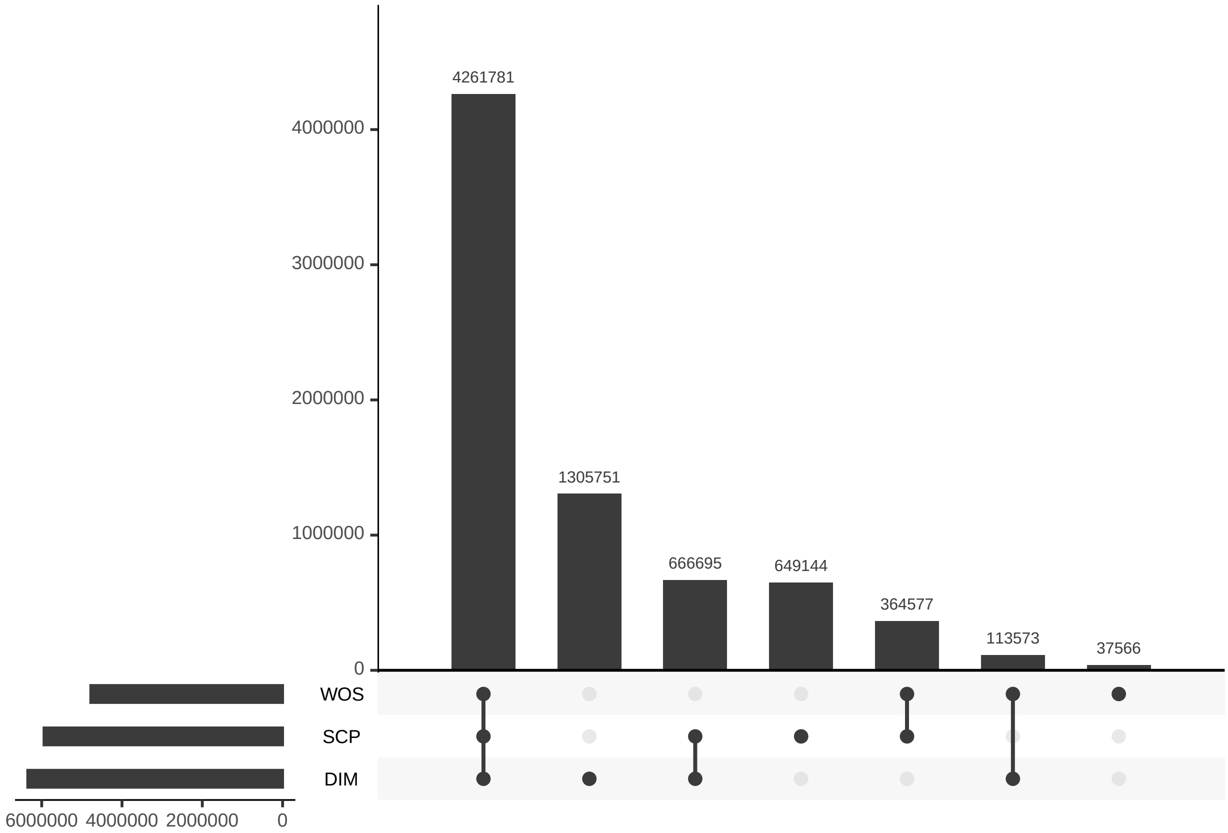

Figure 11: Size of databases in terms of articles and reviews with a DOI in 2016-2018 (lower left panel), magnitude of intersections (upper panel) and relation between total databases and intersections (lower right panel)

e.g. non-unique DOIs or misread string characters in optical character recognition processes. But these issues might arise randomly and might not adversely affect any structural differences between databases. Instead DOI matching has been observed to produce highly valid matching results [10] and less than 7%, 10%, respectively 1% of articles and reviews in 2016-2018 indexed in WOS, SCP or DIM miss a DOI. Given these relatively low counts of missing DOIs and the time, respectively validity constrains of a more complex, tailor-made algorithmic matching procedure, we stick with the DOI approach for this analysis.

Results

Intersections

Given the selection of publication years, document types and the matching approach Figure 11 details the size of each database as well as the intersections between the databases. For the period 2016-2018 DIM indexes the largest number of articles with more than 6.3 million. SCP comes close with 5.9 million, while WOS includes substantially fewer articles and reviews with 4.8 million respective documents. As explained before the applied and knowingly imperfect lower bound of at least one reference for DIM articles might allow other DIM publications of varying document types to be included as articles and the difference between DIM and SCP might actually be smaller than reported here.

Most publications are indexed in several databases and the degrees of freedom to differentiate a database by its exclusively indexed publications seem limited. Indeed 4.2 million publications can be identified in all three databases and this set of core publications is by far the largest intersection (67%, 71%, 88% of the entire DIM, SCP, respectively WOS corpus). Furthermore another 1.1 million publications are indexed by combinations of any two of the three databases. On top of the jointly indexed publications DIM adds another 1.3 million, or about 20% of its entire corpus, exclusive publications differentiating itself from the other databases. SCP exclusively indexes 0.6 million (approx. 10% of its corpus) publications, while the 37k exclusive WOS publications constitute less than 1% of its corpus and hence WOS differentiates from other databases not by exclusively indexed publications, but apparently by foregoing the indexation of more publications. It might also be observed that contrary to its inclusive indexing policy DIM apparently does not index some 1.1 million article and reviews, which can be identified only in WOS and/or SCP.

Internal Coverage

Given the sizable differences between the databases, we now analyse how the internal coverage of core publications varies due to the different indexation practices. By observing if a reference in an indexed publication is, or is not, itself indexed in the same database we compare the relevance attribution of authors with the relevance attribution of database providers. An especially pronounced difference indicates that a database only partially captures the communication flow perceived relevant by authors and hence a bibliometric analysis might be biased on any such out-of-sync dataset. By adding the missing publications or restricting the coverage to a well interlinked citation network component with relatively few external links, the balance between the authors' and the databases providers' relevance attributions can be improved.

As articles and reviews in the core set of jointly indexed publications have unanimously been found relevant by the three database providers and hence allow for a comparison, we apply their reference lists to observe potential differences in the database specific coverage described above. Especially references to non-core publications will differentiate the databases in this respect.

Figure 12 reports the database specific percentage of indexed, or so-called source references via a density plot across all 4.2 million core publications. The upper panel reports on the overall internal coverage of 2016-2018 publications, while the middle panel restricts the references to publications published up to 2004 and earlier and the lower panel depicts the share of indexed references published in 2016.

The overall internal coverage in the upper panel includes a wide variety with values from zero to one hundred percent. All three distributions are skewed to the left indicating that a large share of references in core publications are indexed, while some indexed core publications find few of their references indexed in the respective database. Especially the social sciences and humanities with their minor focus on journal articles as the primary communication device have been observed to lack internal coverage [19]. Apart from this discipline specific effect affecting all databases alike, also notably differences can be observed in the overall internal coverage between databases. Whereas DIM and WOS produce relatively high internal coverage values with many core publications having over 90% of their references indexed, SCP shows a lower agreement with the authors' relevance attribution on the references.

This peculiarity of SCP vanishes if we restrict the set of references to the publication years of 2004 and earlier. In the respective middle panel all three databases show a similar albeit lower coverage on this reduced set of references. The above described additional content in SCP and DIM only slightly increases the internal coverage compared to WOS. However, the quality of SCP

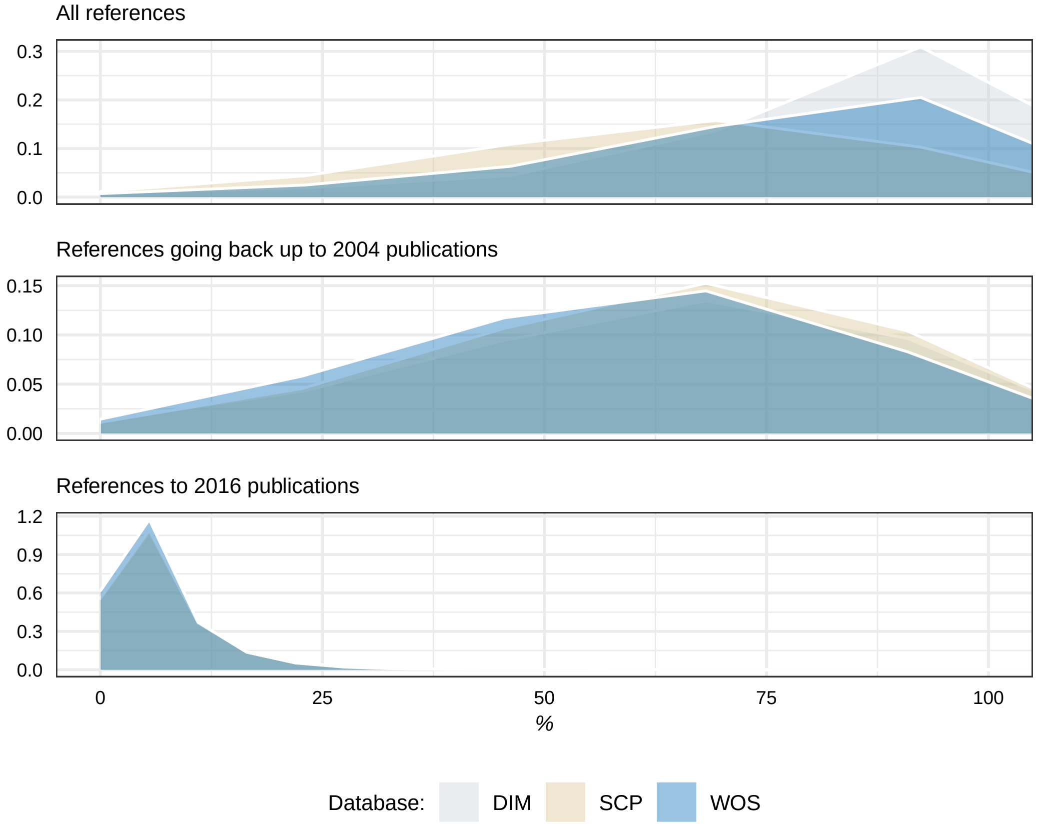

Figure 12: Internal coverage of references in jointly indexed journal articles and reviews published between 2016 and 2018 with at least one reference

data, a service launched in 2004 and covering publications back to at least 1996, has been observed to improve from publication year 2004 onwards [19] and hence the lower coverage of SCP in the upper panel might not result from cross-sectional differences in coverage, but rather differences in coverage over time[2].

Finally, the lower panel adopts a three year citation window perspective and references from 2016-2018 publications are restricted to 2016 publications, i.e. the 2016 publications might accumulate citations from 2016-2018 publications along the lines of the above analysis on normalised citations. As in the middle panel, no pronounced visual difference can be observed between the databases. However it might be noted that on average only 5% of all references are taken into account and most other signals of relevance attribution via references are discarded in any such analysis. As the depicted scale of the x axis might conceal actual differences between the databases, we will focus now especially on this lower end of 2016 publications. In doing so we adopt a maximum contrast approach comparing core publications indexed in all three databases with exclusive publications indexed solely in one of the three databases.

Internal communication of core and exclusive publications

We commence the analysis of references to 2016 publications by comparing internal communication patterns separated by core and database specific exclusive publications. As Digital Science currently

[2] Currently only SCP publications back to 1996 are provided in the KB infrastructure and available backfilling provided by Elsevier might decrease the observed difference.

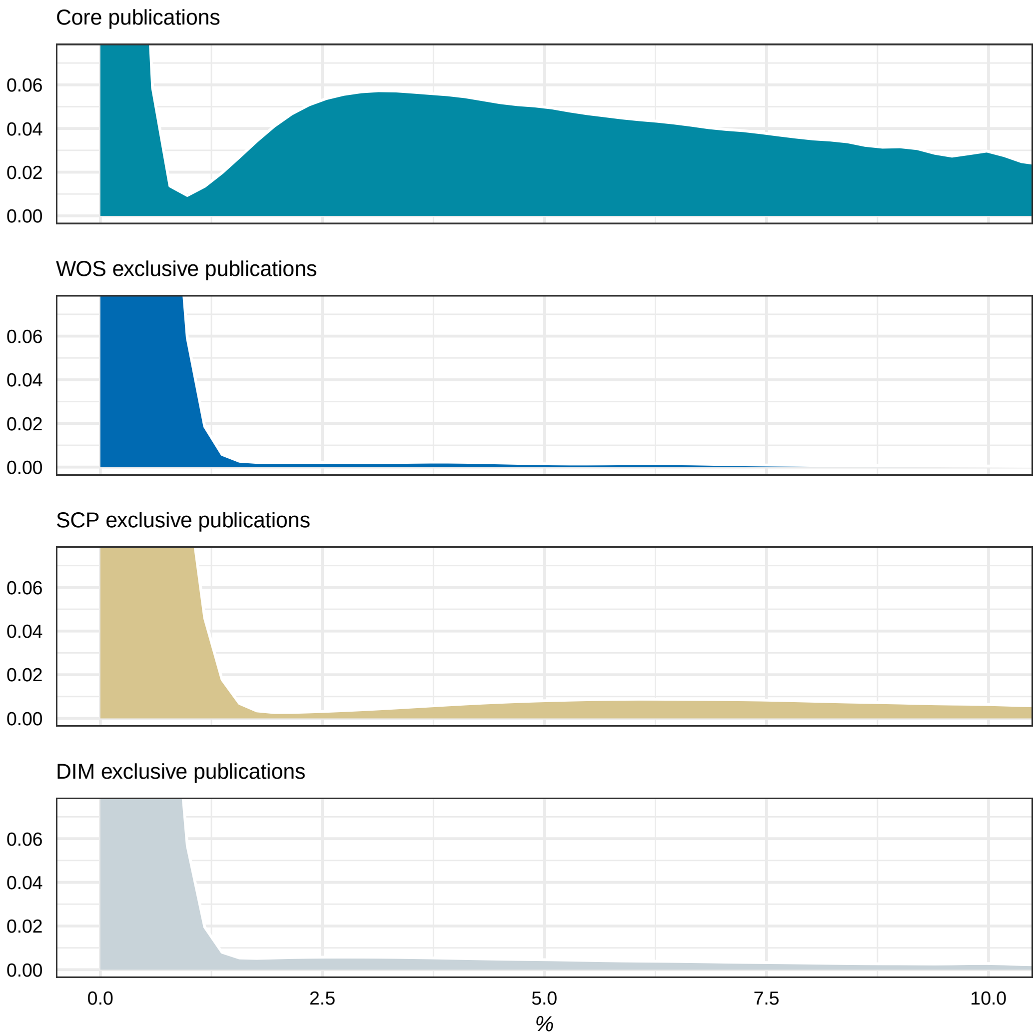

Figure 13: Share of internal references from core and exclusive publications from 2016-2018 to core respectively exclusive publications of 2016

only provides source references for publications indexed in DIM, the total number of references in exclusive DIM publications is unknown and for comparability we report instead the share of 2016 publications among all source references of 2016-2018 publications.

Figure 13 differentiates between the reference-based internal communication among core publications (top panel) and among the database specific exclusive publications (lower panels). As to be expected by the overall low share of references to 2016 publications depicted in the lower panel of Figure 12, the share of internal communication is rather small and the vast majority of 2016-2018 publications miss to reference a single 2016 publication. This observation holds for all databases alike and hence WOS, SCP and DIM can only be compared on the residual share of publications which do include references to 2016 publications.

Comparing the diverse panels it might be observed that core publications possess a much higher degree of internal communication with a sizeable number of core publications referencing other core publications. Hence these publications define an interlinked network component, which might be defined as highly self-referential. This component has been identified and indexed by all three databases and, as described above, constitutes between 67% (DIM) to 88% (WOS) of the respective databases.

To the contrary, exclusive publications solely indexed in one particular database differentiate the databases. But these publications refer to each other to a much lower degree and the internal communication among the set of database specific exclusive publications is hardly visible if compared with core publications. They seldom refer to themselves and compared to core publications are only loosely connected by citation links if at all.

But among the exclusive publications noticeable differences emerge. WOS exclusive pubs might be defined as the least interconnected publications and the much larger set of DIM exclusive publications communicates considerably more internally if compared to this low WOS baseline. SCP exclusive publications might be differentiated by their notably higher degree of internal communication. Especially the observed increase and local maximum around 6% indicates that SCP indexes an additional component of the underlying citation network which is considerably more densely connected than the WOS or DIM equivalents. While WOS exclusive publications hardly communicate internally, especially SCP has uniquely identified additional publications, which show a relatively high degree of internal communication in parallel to the core publications and might constitute a separate component of the citation network.

Outer circle referencing core publications

After analysing the internal communication in each strata, we now describe how core and exclusive publications are interlinked. In Figure 14 we start with relevance of the core publications to exclusive publications by computing the share of references in exclusive publications citing core publications.

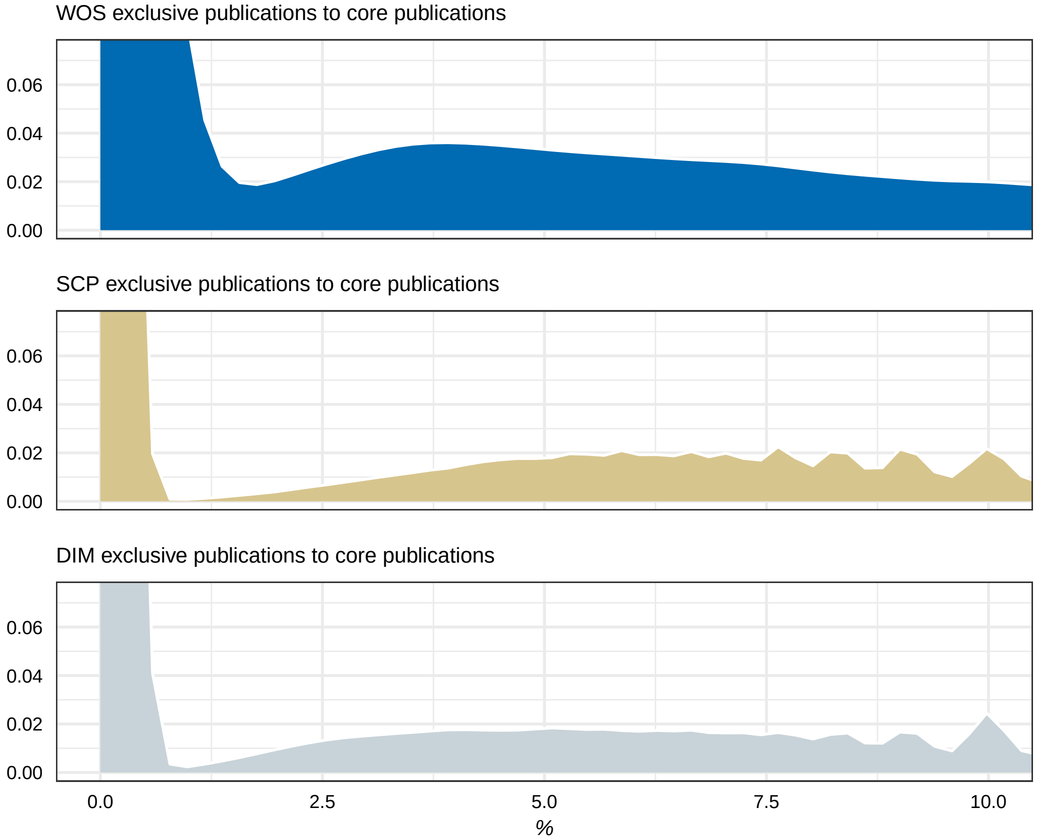

Figure 14: Share of references in exclusive publications from 2016-2018 to 2016 core publications

As before the majority of exclusive publications in 2016-2018 do not cite any 2016 (core) publication. Hence most signals of relevance attribution of 2016-2018 exclusive publications are discarded. However the share of references from exclusive publications to core publications is substantially higher than the share of references from exclusive publications to other exclusive publications depicted in Figure 13. This observation holds unanimously for all three databases. Hence, the content of core publications seems to carry more relevance than the content of exclusive publications and exclusive publications might be understood as an outer, consecutive circle building upon the content provided in core publications.

Comparing the three databases especially WOS exclusive publications exhibit a strong dependence on core publications resulting in a much denser overall database than SCP and DIM. But also exclusive publications in these two larger databases rely to a large and similar extent on core publications. Taking into account the different sizes of exclusive publications (Figure 11), WOS apparently foregoes indexing more publications, but offsets this with a more dense citation graph. DIM identifies more exclusive publications than SCP, which show similar dependence on core publications as SCP exclusive publications (Figure 14), but lag the SCP exclusive publications in internal communication (Figure 13). Hence, the larger set of exclusive publications in DIM is offset by less internal consistency if measured by internal communication among exclusive publications.

Impact of exclusive publications on core publications

To complete the interlinkage between core and exclusive publications, we finally depict in Figure 15 the share of references in core publications citing exclusive publications. As we observe the total number of references of these core publications via WOS and SCP, we normalize by the total reference count and not the source reference count used before.

A noticeable share of core references links out to exclusive publications across all databases. However, taking into account the actual percentage, often less than 2.5%, a low relevance of exclusive publications for core publications might be observed. The respective range for internal references among core publications (upper panel in Figure 13) and the share of core publications in references of exclusive publications (Figure 14) is much higher, often above 5%.

But apart from covering approximately the same range, the three databases differ in the magnitude of the density in this range. The approximately 37k WOS exclusive publications might not show up in the 4.2 million core publications as often as the 1.3 million exclusive DIM publications do. Or, put differently, varying core publications refer to different exclusive publications, but in any case the share of these references is relatively low.

Hence, we observe the same low relevance for all exclusive publications and the particular choice of exclusive publications by a database results in no visual difference to the core publications. No database apparently manages to identify particularly relevant publications for the core publications, but instead all three databases find different exclusive publications of the same low relevance to the core. As a seeming paradox, exclusive publications in this sense do not distinguish databases from one another, but constitute different samples with the same characteristics.

As a consequence, core publications might not be enlarged in any of the three directions offered by WOS, SCP and DIM. None of the sets seems particularly relevant to the core, while all exclusive publications incorporate lower internal communication than the core publications (Figure 13) and hence would compromise the observed internal consistency in communication flows among core publications.

Figure 15: Share of references in core publications from 2016-2018 to 2016 exclusive publications

Discussion

Based on descriptions of WoS, Scopus and Dimensions from Clarivate Analytics, Elsevier and Digital Science, we determined there are four particular differences between the databases that could influence the results of bibliometric analyses that use them. These differences lie in the number and type of documents indexed, the completeness of citation linkages between indexed documents, the inclusion criteria applied to content, and how items are classified to disciplines. Each of these aspects is important as the environment of a database both defines how any particular publication is valued in terms of citations within the database, and also the context against which publications are normalised.

As Dimensions was only released in 2018, only a small number of previous studies have examined the differences between it and the long-standing Scopus and WoS. One larger study examined macro-level differences in the content of Dimensions and WoS compared to Scopus and found important differences in the number of items indexed, the classification of documents to disciplines, and the completeness of citation linkages [22], however this study did not examine the potential effects of these differences on bibliometric indicators. Another three smaller scale studies determined that Scopus and Dimensions were essentially interchangeable in their coverage and citation-based indicators of publications in food science, library science, and business, and a random sample of 10,000 publications [12][20][16], while no study had yet directly compared WoS and Dimensions.

In this study, we conducted first a comparative analysis of the normalised citations of 41,848 duplicate German articles published in 2016 in the three databases, examining the outcomes at both the disciplinary and sectoral levels, to assess the differences incurred in the impact of the same publication when contextualised against the unique content of each database due to the noted differences.

A notable finding, although one that has previously been described [17], is that the impact of publications in nearly all disciplines and sectors is valuated more favourably in Scopus than in WoS. This occurs as Scopus' larger index of publications, particularly in the social sciences (Figure 2), means that all items have a greater opportunity to be cited, which increases the average observed citation count, as seen in the left-hand side of Figure 5. Meanwhile, the similar criteria against which Elsevier and Clarivate Analytics test content for indexing and the similar classification of document types means that the indexed articles are of similar quality and generate comparable average expected citation counts, as we observed in Figure 5. Consequently, the higher observed citations in Scopus but similar expected citations produced higher normalised citation impact in Scopus for nearly all disciplines. In particular, we see that *Clinical medicine*, *Economics and business*, and *Computer and information sciences*, which had nearly double the amount of content in Scopus, were particularly favoured due to these circumstances. Importantly, *Clinical medicine* was one of only three disciplines that had higher expected citations in WoS, suggesting a structural difference may be present in the content indexed in WoS versus Scopus. Only *Agriculture, forestry, and fishery* and *Mathematics* improved in impact in WoS over Scopus. While this is not readily reconcilable, as Scopus had substantially higher coverage of both disciplines and also higher both observed and expected citations, the increase in impact was marginal at 11.0% and 8.3% respectively.

Although there was a similar situation of increased impact in Dimensions over both Scopus and WoS for the majority of disciplines, there is less explanatory evidence for this pattern. For some disciplines, the same situation of greater coverage in Dimensions which facilitated higher observed citations alongside similar expected citations can explain the improved impact in Dimensions. Further, for eight disciplines with more than 10% improvement in Dimensions over Scopus, the driving force of differences in valuations appeared to be the expected citations. These disciplines tended to have minimal differences in the observed citations but higher expected citations in Scopus, generating higher impact in Dimensions. The difference in expected citations may stem from the lack of disambiguation of document types in Dimensions. As noted, expected citations in Dimensions are calculated based on the original document type classification, which does not split out articles from other content of a journal as occurs in Scopus and WoS. This inclusion in the 'article' document type of other material which is unlikely to receive the same citations as articles, such as editorials, comments or lists of journal boards, is likely to influence the expected citations counts. However, there is little consistency in patterns of expected citations between Dimensions and WoS and Scopus, hindering a generalisable explanation for the generally higher impact in Dimensions for most disciplines.

Although there are reasonable – if inconsistent – differences in valuations between databases at the level of disciplines, the influence of the database was comparatively less at the level of German sectors. In particular for Scopus and Dimensions, the distribution of differences in impact between the databases was close to normal and across all sectors the average difference in normalised citations was less than 0.1. Overall then our analysis here aligns with the three aforementioned studies that suggest Dimensions and Scopus produce similar results when conducting assessments at the macro-level. However, in contrast to those studies, we see definitive differences in assessment of impact of German publications between the databases at the discipline level, which suggests Dimensions and Scopus might produce different results if an institution had a particularly strong research profile in a single or small number of disciplines.

Further, there were larger differences in valuations between WoS and Scopus and WoS and

Dimensions, indicating that WoS produces a different picture than the other databases. WoS appears to provide a more conservative valuation of German sectors and skews toward basic research, whereas Scopus more favourable assesses all sectors, but in particular the Fraunhofer Society and Economy sectors with their focus on applied research. The same effect was seen between WoS and Dimensions also, with all sectors having an average increase in normalised citations of between 0.2 and 0.4 in Dimensions, with a particular increase for the Economy sector. As such, while Scopus and Dimensions report similar impact on the macro level, this is not the case for WoS and Scopus or Dimensions.

Next, in the second part of the study we identified core publications jointly indexed in all three databases and exclusive publications solely indexed in one of the three databases. In a maximum contrast approach the separately recorded communication flows via references were compared across these sets to observe the analytical value each database reveals by its particular indexation practices. As noted above the resulting differences in coverage are assumed to partially explain the observed differences in the normalised citation analysis and consequently the network analysis supports the normalised citation analysis in that it offers a potential explanation for the observed differences.

Not completely unexpectedly, the indexation of more publications is accompanied by an offset in the density of communication flows in the resulting citation graph. Irrespective of the database, exclusive publications have been observed to mark an outer ring resting substantially upon core publications. They barely bring additional relevance to core publications, but the core might be characterised by its relatively high degree of self-reference, where former core publications constitute the base or frame for new core publications.

Hence knowledge is transferred from the core to the outer circle and this transfer is especially visible in SCP and DIM. As the amount of parallel communication among exclusive publications seems negligible and the particular choice of exclusive publications does not affect its relation with the core, a star model seems to emerge of an interconnected core producing knowledge in a self-referential mode, which is then disseminated to the only loosely interconnected outer circle.

The exact frontier between core publications and the outer ring might be difficult to define as a continuum between two poles and several jump discontinuities arising from highly connected discipline or geographic components in between impede an exact binary definition of each part. Especially Indian or Chinese local journals with Crossref-independent DOIs might explain the slightly more extreme position SCP adopts on the continuum if compared to the slightly fuzzier position of DIM. Given the relatively small difference between SCP and DIM in absolute article counts, the imperfect document type resolution in DIM and the likely supposition that the article document type in DIM has not been singled out perfectly by our reference requirement but is actually over-reported, a sizeable amount of incertitude accompanies the current study. Especially improvements to the data quality in the comparably young DIM database, alike to what has been observed for SCP in the past, might alleviate these limitations in the future.

Hence WOS with its restrictive indexation policy and SCP with its selective indexation policy constitute two separate self-imposed stances with a distinct message: WOS largely represents the well interconnected core citation network component, while SCP allows to observe some transfer from the core to the periphery. DIM with its laissez-faire indexation policy conveys, apart from the currently substantial data quality issues, more coverage but a less decisive but similar message to SCP. However, Digital Science has communicated that the indexed publication set in DIM constitutes a basis and not the final offering for citation analysis [14]. Accordingly DIM provides for example journal lists to refine the citation network by a quality criterion. Furthermore other indexed data types like clinical trials or policy documents potentially enhance the perspective on transfer beyond what is currently available in other databases.

References

[1] J. Adams et al. *Dimensions – A collaborative approach to enhancing research discovery*. Tech. rep. Digital Science, 2018. DOI: 10.6084/m9.figshare.5783160.v1.

[2] A. Akbaritabar and S. Stahlschmidt. "Merits and Limits: Applying open data to monitor Open Access publications in bibliometric databases". In: *Proceedings of the 17th Conference of the International Society for Scientometrics and Informetrics*. Ed. by G. Catalano et al. Vol. 2. Rome: Edizioni Efesto, 2019, pp. 1455–1461.

[3] C. Bode et al. *A guide to the Dimensions data approach: A collaborative approach to creating a modern infrastructure for data describing research: where we are and where we want to take it*. Tech. rep. Digital Science, 2018. DOI: 10.6084/m9.figshare.5783094.

[4] C. Bode et al. *A guide to the Dimensions data approach: A collaborative approach to creating a modern infrastructure for data describing research: where we are and where we want to take it*. Tech. rep. Digital Science, 2019. DOI: 10.6084/m9.figshare.5783094.

[5] L. Bornmann. "Field classification of publications in Dimensions: a first case study testing its reliability and validity". In: *Scientometrics* 117.1 (2018), pp. 637–640. DOI: 10.1007/s11192-018-2855-y.

[6] Clarivate Analytics. *Web of Science: Summary of Coverage*. Retrieved on March 3, 2020. n.d. URL: https://clarivate.libguides.com/webofscienceplatform/coverage.

[7] *Editorial selection process*. Retrived on March 5, 2020. n.d. URL: https://clarivate.com/webofsciencegroup/solutions/editorial/.

[8] Elsevier. *Scopus: Content Coverage Guide*. Retrieved on March 5, 2020. 2017. URL: https://www.academia.edu/41511063/Scopus_Content_Coverage_Guide_Updated_August_2017_.

[9] Elsevier. *Scopus: Content Coverage Guide*. Retrived on July 13, 2020. 2020. URL: https://www.elsevier.com/__data/assets/pdf_file/0007/69451/Scopus_ContentCoverage_Guide_WEB.pdf.

[10] N. Fraser and A. Hobert. *Report on Matching of Unpaywall and Web of Science*. Tech. rep. 2019.

[11] E. Garfield. "The mystery of transposed journal lists – wherein Bradford's law of scattering is generalised according to Garfield's law of concentration". In: *Current Content* 17 (1971). Reprinted in Essays of an Information Scientist, pp. 222-223, by E. Garfield, 1977, Philadelphia, USA: ISI Press.

[12] A-W. Harzing. "Two new kids on the block: How do Crossref and Dimensions compare with Google Scholar, Microsoft Academic, Scopus and the Web of Science?" In: *Scientometrics* 120.1 (2019), pp. 341–349. DOI: 10.1007/s11192-019-03114-y.

[13] C. Herzog and B. K. Lunn. "Response to the letter 'Field classification of publications in Dimensions: a first case study testing its reliability and validity'". In: *Scientometrics* 117.1 (2018), pp. 641–645. DOI: 0.1007/s11192-018-2854-z.

[14] D. Hook, S. Porter and C. Herzog. "Dimensions: Building Context for Search and Evaluation". In: *Frontiers in Research Metrics and Analytics* 3.2 (2018). DOI: 10.3389/frma.2018.00023.

[15] P. Jacso. "As we may search – Comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases". In: *Current Science* 89.9 (2005), pp. 1537–1547.

[16] E. Orduña-Malea and E. Delgado-López-Cózar. “Dimensions: re-discovering the ecosystem of scientific information”. In: *El Profesional de la Información* 27.2 (2018), pp. 420–431.

[17] S. Stahlschmidt, D. Stephen and S. Hinze. “Performance and Structures of the German Science System. Studien zum deutschen Innovationssystem”. In: EFI, 2019. Chap. Studie 5-2019.

[18] Australian Bureau of Statistics. *Australian and New Zealand Standard Research Classification (ANZSRC), 2008*. cat. no. 1297.0. Canberra, Australia, 2008.

[19] D. Stephen, S. Stahlschmidt and S. Hinze. “Performance and Structures of the German Science System 2020. Studien zum deutschen Innovationssystem”. In: EFI, 2020. Chap. Studie 5-2020.

[20] M. Thelwall. “Dimensions: A competitor to Scopus and the Web of Science?” In: *Journal of Informetrics* 12.2 (2018), pp. 430–435. DOI: 10.1016/j.joi.2018.03.006.

[21] N. J. Van Eck and L. Waltman. “Accuracy of citation data in Web of Science and Scopus”. In: *Proceedings of the 16th International Conference of the International Society for Scientometrics and Informetrics*. Ed. by R. Rousseau, W. Glänzel and Z. Rongying. Wuhan: HSE, 2017, pp. 1087–1092.

[22] M. Visser, N. J. Van Eck and L. Waltmann. *Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic*. 2020. arXiv: 2005.10732 [cs.DL].

[23] Q. Wang and L. Waltman. “Large-scale analysis of the accuracy of the journal classification systems of Web of Science and Scopus”. In: *Journal of Informetrics* 10.2 (2015), pp. 347–364. DOI: 10.1016/j.joi.2016.02.003.