Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-024-02915-8 OPEN

Check for updates

# Interdisciplinary research attracts greater attention from policy documents: evidence from COVID-19

Liang Hu[1], Win-bin Huang[1,2] & Yi Bu[1]✉

Interdisciplinary research is increasingly recognized as one of the solutions to today's challenging scientific and societal issues. Many studies have aimed to explore the relationship between the interdisciplinarity of research and the attention they receive from the scientific community as well as society. However, the relationship between interdisciplinarity and attention from policy documents remains unclear. In this study, we utilize publications data on the COVID-19 topic to explore such a relationship. Through the analysis and interpretation of empirical datasets, this research finds that there is a positive correlation between the interdisciplinarity of scientific publications and the attention they receive from policy documents in almost all fields. Among the three dimensions (i.e., variety, balance, and disparity) of interdisciplinarity, variety exhibits the most pronounced positive impact on political attention. This study fills a previous research gap and provides insights for researchers and policymakers, highlighting that interdisciplinary research holds greater potential to impact policy formulation and implementation processes.

[1] Department of Information Management, Peking University, Beijing 100871, China. [2] Peking University Chongqing Research Institute of Big Data, Chongqing 401329, China. ✉email: buyi@pku.edu.cn

## Introduction

As science and technology continue to advance and many complex research issues need to be tackled, research beyond disciplinary boundaries is becoming increasingly important. The modes of research beyond disciplinary boundaries primarily may include three types: multidisciplinarity, transdisciplinarity, and interdisciplinarity. Multidisciplinarity draws on knowledge from different disciplines but stays within their boundaries. Transdisciplinarity integrates the natural, social, and health sciences in a humanities context and transcends their traditional boundaries. Interdisciplinarity analyzes, synthesizes, and harmonizes links between disciplines into a coordinated and coherent whole (Choi & Pak, 2006). The National Academies (2004) defined that "(i)nterdisciplinary research (IDR) is a mode of research by teams or individuals that integrates information, data, techniques, tools, perspectives, concepts, and/or theories from two or more disciplines or bodies of specialized knowledge to advance fundamental understanding or to solve problems whose solutions are beyond the scope of a single discipline or area of research practice" (p. 11153). Considering the capacity of interdisciplinarity to integrate and synthesize diverse disciplinary knowledge, thereby enhancing fundamental understanding or offering innovative solutions to complex issues, we choose interdisciplinarity as the focal point of our research endeavor.

Interdisciplinary research is increasingly recognized as the solution to today's challenging scientific and societal problems (Carayol & Thuc Uyen Nguyen Thi, 2005; Frodeman & Mitcham, 2007). In recent decades, governments and public funding agencies have increasingly called on scientists in universities and public research organizations to demonstrate both the scientific and societal impacts of publicly funded research (Bornmann, 2013; Bozeman & Sarewitz, 2011; Salter et al., 2017). According to the assertion that interdisciplinary research facilitates the recombination of knowledge, encourages atypical combinations of knowledge, and leads to significant scientific breakthroughs (Fontana et al., 2020; Uzzi et al., 2013), many studies have aimed to explore the relationship between the interdisciplinarity of research and the attention they receive from the scientific community as well as society (Larivière et al., 2015; Levitt & Thelwall, 2008; Piwowar, 2013; Wang et al., 2015). However, the relationship between interdisciplinarity and attention from policy documents remains unclear. To address this gap, this study aims to explore whether interdisciplinary research receives more political attention compared to unidisciplinary research.

Given the significant importance of interdisciplinary research, it is crucial to consider the attention it receives from policy documents. Receiving attention from policy documents implies that the research has the potential to influence policy formulation and implementation, in other words, the transformation of scientific research into policy outcomes (Bornmann et al., 2016). The process of transformation may include the assessment and validation of research results, followed by transforming them into policy recommendations, guidelines, regulations, or practical applications, with positive impacts on society (Lewison & Sullivan, 2008). The attention research receives from policy documents facilitates the transformation of scientific knowledge into concrete actions, thereby influencing policy-making and decision-making mechanisms. The transformation of scientific knowledge into concrete actions, which combines scientific research with societal needs, not only achieves the societal impact and value of science but also better utilizes research outcomes to drive societal and economic development, promote technological innovation, and advance society (OECD, 2015).

In this study, we utilize metadata of 159,957 publications related to COVID-19 to explore the relationship between the interdisciplinarity of scientific research and its reception of attention from policy documents. Firstly, we divide publications into groups based on the degree of interdisciplinarity and conduct comparative analyses of how political attention varies with the changes in interdisciplinarity. Secondly, we perform linear probability regression analysis based on fixed effects for disciplines and time on the focal publications. To minimize the potential influence of confounding factors, coarsened exact matching and further regression analysis based on the matching results are conducted. Thirdly, to explore deeper into differences across disciplines, we perform direct regression and regression based on coarsened exact matching (CEM) for various fields and visualize the coefficients for better understanding. Our findings indicate that there is a positive correlation between the interdisciplinarity of scientific publications and the attention they receive from policy documents in almost all fields. The analysis and findings shed new conceptual and empirical light on the factors underlying the relationship between interdisciplinary research and political attention to science. In the following sections, we will begin by reviewing relevant works. Next, we will introduce the data sources and the data processing methods employed in this study, then detail our research methodology and present our research findings. Finally, we will engage in a discussion of the results and highlight the limitations of our current study.

## Related work

**Measurements of interdisciplinarity**. Policymakers and researchers continue to be interested in the quantitative measures of interdisciplinarity (Wagner et al., 2011). Rao (1982) and Stirling (2007) pointed out that diversity is composed of three fundamental elements, namely variety, balance, and disparity. Note that diversity, in this context, is a broader concept compared to interdisciplinarity. Each of these components is essential, yet none alone is adequate to fully define the concept of diversity (Bu et al., 2021; Zhang et al., 2016). This notion and a generic indicator of diversity were then introduced and modified by Rafols and Meyer (2010) to Information Science as a quantitative measurement of knowledge integration to infer interdisciplinarity. Among the three distinct components, variety represents the count of categories to which system elements are allocated. It addresses the question: 'how many different types of elements the system has?' Balance is determined by the distribution pattern of elements among these categories. It answers the query: 'how evenly are the quantities of each type of element distributed?' Disparity relates to the extent and manner in which elements can be distinguished from one another. It responds to the query: 'to what degree do the various types of elements differ from each other?' (Zhou et al., 2021). A substantial portion of the research is dedicated to developing indicators that combine two or three aspects (dimensions) of diversity: Rao-Stirling (RS) diversity (Rao, 1982; Stirling, 2007), DIV (Leydesdorff et al., 2019), and so on. The goal is to create a robust metric that can effectively evaluate and compare the intensity of interdisciplinarity across various entities. In this study, we aim to work as comprehensively as possible so that information loss caused by dimension reduction or integration can be minimized. Therefore, to quantify the interdisciplinarity, we employ single-component (variety, balance, and disparity themselves) indicators as well as comprehensive indicators RS and DIV.

**Attention to interdisciplinary research**. Attention is a profoundly significant concept in many fields, exerting its influence on various phenomena. In quantitative science and technology studies, attention is oftentimes indicated by citation-based metrics, effectively serving as the principal "currency" within the scientific community, and along with other forms of recognition, constitutes the foundation for career advancements and the establishment of scientists' reputations (Petersen et al., 2014). Given the extensive interest and the policy promotion for interdisciplinary research, the question of whether interdisciplinary research garners more attention within the scientific community (referred to as scientific attention) compared to unidisciplinary research has been a topic of ongoing investigation over the long term. For example, Larivière et al. (2015) found that long-distance interdisciplinarity leads to higher scientific impact (note that the measurement methods for impact and attention are often similar, typically measured through citation-based metrics). Levitt & Thelwall (2008) found that interdisciplinary papers received fewer citations in life and physical sciences, where interdisciplinary papers were defined as papers published in journals assigned to multiple subject categories. Different from previous studies compositing various aspects of interdisciplinarity into a single indicator, Wang et al. (2015) used factor analysis to uncover the relationship between distinct dimensions of interdisciplinarity and scientific impact. While there are numerous existing research findings, their conclusions are not consistent. One potential explanation for this divergence is the variations in the measurement methods of interdisciplinarity they employ, as well as the differences in the types of research data and fields they examine (Liu et al., 2022).

Policymakers expect science to demonstrate its value to society but not limited to academia (Bornmann, 2013). In addition to attention from the scientific community, scientific research, including interdisciplinary research, also attracts attention from society and policy documents (referred to as societal attention and political attention, respectively). The assessments of the societal impacts and societal attention of research outputs are prompting a search for alternative quantifiable measures and potential complementary metrics. Altmetrics (Gunn, 2013) is also considered an interesting option for assessing the societal impact or societal attention of research as they offer new ways to indicate (public) engagement with research output (Piwowar, 2013). There are several quantitative studies on the relationship between interdisciplinarity and societal attention—for example, Chavarro et al. (2014) showed that papers with higher scores for certain dimensions of interdisciplinarity are associated with a stronger focus on research that addresses local issues.

Despite numerous studies aiming to explore the relationship between interdisciplinarity and scientific as well as societal attention, there is currently a lack of research concerning whether interdisciplinary research receives more political attention compared to unidisciplinary research. This is precisely the focus of our study. Referring to citation-based bibliometric indicators, if the citing entities transition from scientific publications to policy documents, implying that policy documents cite scientific publications, it reflects that the attention research receives from policy documents. Consequently, to assess the political attention of scientific research, analogous to measuring scientific attention, one can similarly construct citation-based metrics. For instance, one can use the policy citation count of a scientific publication to gauge the extent of the political attention of that scientific publication.

## Methodology

**Data**. The empirical data employed in the analysis mainly comes from OpenAlex, a new and fully open scientific knowledge graph, launched to replace the discontinued Microsoft Academic Graph (MAG) (Priem et al., 2022). MAG is a comprehensive and constantly evolving research knowledge base developed by Microsoft Research. It encompasses a vast number of records of academic publications dating back to the 19$^{th}$ century, including citation relationships between these publications, as well as related metadata about authors, institutions, journals/conferences, and research fields, making it a valuable resource for researchers, students, and institutions worldwide (Wang et al., 2020). OpenAlex (MAG) employs a bottom-up approach, operating at the individual paper level, for its field categorization process. This involves quantifying the semantic "distance" between textual paragraphs from two publications. The resultant semantic representations are then clustered, forming the foundation of concepts that represent fields, domains, or disciplines. This process leads to the automatic clustering of concepts into six levels of granularity. Notably, the top two levels of concepts (Levels 0 and 1) are manually defined to create a coherent hierarchical structure that aligns with various categorization systems (Wang et al., 2020). In this structure, Level 0 encompasses 19 distinct fields such as chemistry and economics, while Level 1 comprises 292 subfields, e.g., biochemistry and macroeconomics. To better capture the political attention of scientific research, we limit our research scope to the topic of "COVID-19". Due to the rapid emergence and development of a significant amount of new research as a result of the outbreak and prevalence of the COVID-19 pandemic, the COVID-19 field has become a highly valuable subject of study. Furthermore, to address the COVID-19 pandemic, governments worldwide have implemented numerous policy documents that heavily reference scientific publications, which has provided a substantial increase in the sample size for this research. The time frame is set from January 1, 2020 to December 31, 2021, because this period corresponds to the outbreak and spread of the COVID-19 virus, during which a significant number of related policy documents were issued. By searching for the keywords "COVID-19" or "Coronavirus" or "Corona virus" or "2019-nCoV" in the titles of articles within the OpenAlex dataset, we obtain a total of 486,471 scientific publications within the specified time frame.

To obtain the instances where these scientific publications are referenced by policy documents, we utilize the Overton policy document database. Overton is the world's largest collection of policy documents, parliamentary transcripts, government guidance, and think tank research which contains a core set of policy documents with sufficient citation linkage to academic publications to support various citation analyses that may be informative in research evaluation, impact assessment, and policy review (Szomszor & Adie, 2022). Overton uses machine learning techniques to extract topics from the full text of each policy document they index. Therefore, with "covid-19" as the topic, we obtain the DOIs of 32,379 scientific publications along with the number of times these publications are cited in policy documents from the Overton database. Subsequently, by using their DOIs, we match the retrieved Overton data with the entire OpenAlex dataset, which finds that out of the 32,379 publications from Overton, a total of 31,105 publications appear in the entire OpenAlex dataset. Distributions of some variables of the policy documents (number of policy documents in each year, discipline, etc.) can be found in Szomszor and Adie (2022).

Then we obtain 489,361 publications by taking the union of the 486,471 publications obtained earlier from the OpenAlex dataset and the 31,105 publications obtained from Overton. Furthermore, we retain only journal publications and exclude those without references recorded, resulting in 175,950 scientific publications and their related metadata, e.g., authors, numbers of policy citations, and research fields, as the focal dataset. In the subsequent steps, we further process the data, including excluding publications with reference counts less than 3 and those with missing values regarding variables to be considered in regression

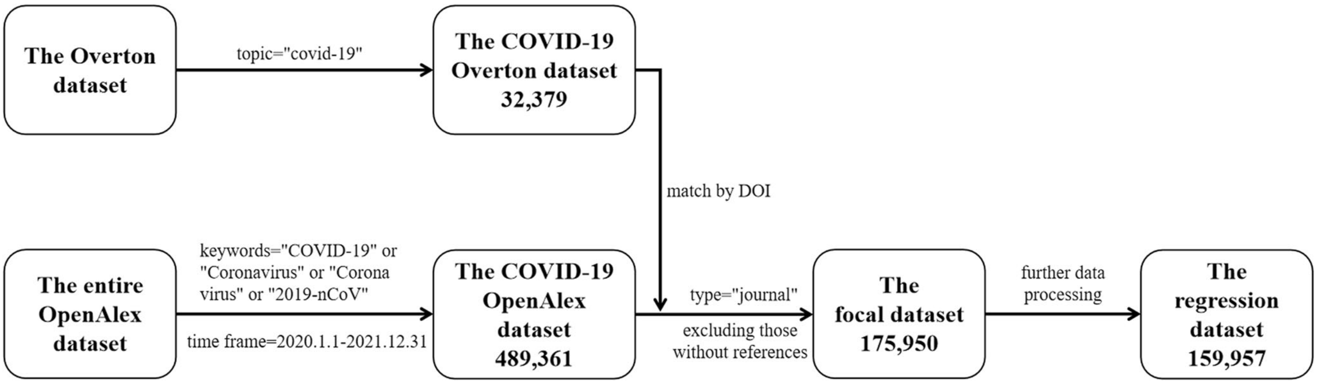


**Fig. 1 Flowchart of data acquisition.** This figure illustrates the process of data acquisition in detail. Focusing on the topic of COVID-19, publication data and policy document citation data were obtained from OpenAlex and Overton, respectively, and the focal dataset was obtained by matching the two through DOI.

**Table 1 Measurements for interdisciplinarity.**

| Measure | Description |
|---|---|
| *M* | A 292*292 citation matrix in which $M_{ij}$ equals the number of times all publications from field *i* cite publications from field *j*. |
| $d_{ij}$ | $1-\cos(row_i, row_j)$ (1), where $row_i$ represents the row vector corresponding to field *i* in matrix *M*. |
| *Gini* | $\frac{\sum_i^n (2i-n-1)x_i}{n\sum_i^n x_i}$ (2) |
| *variety* | $\frac{n}{N}$ (3) |
| *balance* | $1-Gini$ (4) |
| *disparity* | $\sum_{ij(i\neq j)}^{n} \frac{d_{ij}}{[n*(n-1)]}$ (5) |
| *RS* | $\sum_{ij(i\neq j)}^{n} p_i p_j d_{ij}$ (6) |
| *DIV* | $variety * balance * disparity$ (7) |

models (see in the later sections). The entire process of data acquisition is illustrated in Fig. 1.

## Methods

*Interdisciplinarity*. Our research employs two distinct indicators to measure the interdisciplinarity of scientific publications: the RS index proposed by Rao (1982) and further discussed by Stirling (2007), and the DIV index introduced by Leydesdorff et al. (2019). Before commencing the calculations, publications with fewer than 3 references are excluded (a total of 11,628 articles) to ensure the accuracy of the interdisciplinarity indicators. Furthermore, we categorize the references of all articles in the focal dataset into 292 fields based on the Level 1 discipline classification system of OpenAlex. A 292*292 citation matrix *M* is constructed with the entire OpenAlex dataset, in which $M_{ij}$ equals the number of times all publications from field *i* cite publications from field *j*. Then we use 1 minus the cosine similarity between any two rows in the citation matrix to capture the disparity between the two corresponding fields, as shown in formula (1), Table 1.

The formulas for calculating the three dimensions of interdisciplinarity are presented in formulas (3) to (5), Table 1, where *n* represents the number of field categories included in the references of the focal publication while *i* and *j* represent specific field categories within the *n* field categories, $p_i$ denotes the proportion of references in the focal publication belonging to field *i* (similarly for $p_j$), $d_{ij}$ signifies the disparity between fields *i* and *j* which has been calculated before, and *N* stands for the total number of fields in the Level 1 discipline classification system (i.e., $N = 292$). *Gini* denotes the Gini coefficient of the references' field categories and is calculated in formula (2), Table 1, where *n* still represents the number of field categories, *i* is the index and $x_i$ is the number of references to the *i*-th field category when the field categories are sorted by $x_i$ in a non-decreasing order. The Gini coefficient was originally proposed to measure income inequality, and has been used to capture the unevenness and unbalance of the distribution of references across involved disciplines (Leydesdorff & Rafols, 2011). Finally, the calculation formulas of RS and DIV are presented in formulas (6) and (7), Table 1. Note that, for the focal publications with only one reference publication category (i.e., $n = 1$), the calculation according to equation (3) may result in a denominator of 0. However, based on the actual meaning of interdisciplinarity, we set the DIV value to 0 in such cases, which indicates having only one reference publication category signifies no interdisciplinarity.

**Variables and regression models**. This research explores the impact of interdisciplinarity on the attention scientific research receives from policy documents through regression analysis with Python and StataSE 17. Firstly, we set the dependent variable, $policy\_cited_{i,t,j}$, as whether a certain publication *i* published in month *t* and under field *j* is cited by a certain policy documents (referred to as $policy\_cited_{i,t,j}$). The independent variables, $interdisciplinarity_i$, include comprehensive indicators of the interdisciplinarity of the focal publication, namely RS and DIV, as well as three dimensions of interdisciplinarity: variety, balance, and disparity (see operationalization in the previous section). Through a review of relevant literature (Xie et al., 2019), we build four control variables that are potentially related with $policy\_cited_{i,t,j}$, including $team\_size_i$ the number of co-authors in publication *i*, $scientific\_citations_i$ the number of citations of *i* from other scientific publications, $references\_count_i$ the number of references of in publication *i*, and $journal\_impact\_factor_{i,t}$ the impact factor of the journal in which the publication *i* is located for its year of publication, as shown in Eq. (8).

Additionally, fixed effects for both field and time are also taken into consideration. In more detail, we divide the focal publications into 24 months based on the time frame, from January, 2020 to December, 2021. Simultaneously, we categorize the focal publications into 19 fields using the Level 0 discipline classification system of OpenAlex. The model employs multiple linear regression, specifically a linear probability model, due to the dependent variable $policy\_cited_{i,t,j}$ being a binary variable. As shown in equation (8), $\delta_t$ represents the time-fixed effects and $\theta_j$ represents the field-fixed effects. After defining all the variables, we remove data with any missing variables, resulting in a final dataset of 159,957 observations to be used for regression analysis (as shown in Fig. 1).

$$\begin{aligned} policy\_cited_{i,t,j} = \beta_0 + \beta_1 * interdisciplinarity_i + \beta_2 * team\_size_i \\ + \beta_3 * scientific\_citations_i + \beta_4 * references\_count_i \\ + \beta_5 * journal\_impact\_factor_{i,t} + \delta_t + \theta_j + \varepsilon \end{aligned} \tag{8}$$

*Coarsened exact matching (CEM)*. In addition to performing direct regression analysis, we also conduct regression analysis utilizing matched samples as robustness tests. Firstly, we define a variable named *Inter* to indicate whether a focal publication is interdisciplinary. If both the DIV and RS values of a publication are greater than their respective medians, the publication is considered interdisciplinary ($Inter = 1$); otherwise, it is considered unidisciplinary ($Inter = 0$), as shown in formula (9). Using the method, 95,430 unidisciplinary and 64,608 interdisciplinary publications are obtained. Moreover, we also use the quartile instead of the median to judge whether a publication is interdisciplinary. Specifically, if both the DIV and RS values of a publication are within the greatest 25%, it will be classified as an interdisciplinary publication ($Inter' = 1$), if they are within the lowest 25%, it will be classified as a unidisciplinary publication ($Inter' = 0$), as shown in formula (10). Then we obtain 29,569 unidisciplinary and 27,670 interdisciplinary publications in this part. After this, we match interdisciplinary focal publications with similar unidisciplinary focal publications. The purpose of this matching process is to create two groups of focal publications that are identical in terms of co-variates, except for their treatment. This step is crucial as it enables us to estimate the impact of interdisciplinarity on policy citation more accurately by minimizing the potential influence of confounding factors.

$$Inter_i = \begin{cases} 1, DIV_i > median \cap RS_i > median \\ 0, DIV_i \leq median \cup RS_i \leq median \end{cases} \tag{9}$$

$$Inter'_i = \begin{cases} 1, DIV_i \in top25\% \cap RS_i \in top25\% \\ 0, DIV_i \in bottom25\% \cap RS_i \in bottom25\% \end{cases} \tag{10}$$

The matching method we use is the CEM in StataSE 17. CEM is a statistical method used in the field of causal inference and observational studies and aims to reduce bias in treatment effect estimation by creating balanced comparison groups (Ho et al., 2007; Iacus et al., 2011, 2012). The matching conditions include team size, time (publishing month), and field (Level 0 field of the publication). The matching results can be found in Table S2 of the supplementary information.

## Results

**Interdisciplinarity and political attention**. Due to the focus on the topic of "COVID-19," there is inevitably an imbalance in the distribution of the focal publications across different fields. As shown in Fig. 2A, the distribution of publications across different fields based on the Level 0 discipline classification system of OpenAlex demonstrates a significant imbalance, where geology has the least number of publications, with only 20 papers, while medicine has 108,262 papers. To better capture the disciplinary differences, we define a simpler discipline classification method with only five fields, medicine, natural science (excluding medicine), social science, engineering and technology, and humanities and arts. The correspondence between the discipline classification method we define and the Level 0 discipline classification system of OpenAlex can be found in the supplementary information (Table S1). For simplicity, let's refer to natural sciences (excluding medicine) as "natural sciences". The inner plot of Fig. 2A depicts the distribution of publications after adopting the new taxonomy. Despite the continued imbalance, the field with the fewest publications, humanities, and arts, now encompasses 662 pieces of publications.

After redefining the discipline classification, we explore the relationship between interdisciplinarity and political attention coarsely and attempt to capture the disciplinary differences. Specifically, we conduct equal-depth binning and split the focal publications into 5 bins for each field according to RS and DIV, respectively, then calculate the corresponding average polity citations. Figure 2B, C, respectively, reflect how the average policy citations change with the variation in RS and DIV values in different fields. Firstly, regardless of the indicators, the average policy citations for medicine, natural science, and social science consistently surpass those for engineering and technology, as well as humanities and arts. This is likely because policy documents citing COVID-19-related publications are more focused on the societal issues arising from the coronavirus pandemic and the scientific theories contributing to addressing these issues. Furthermore, the average policy citations in natural science exhibit a trend of initially decreasing and then increasing with enhanced interdisciplinarity. On the contrary, the other four fields show an upward trend in average policy citations, which suggests that, in the vast majority of fields, higher levels of interdisciplinarity are associated with a greater likelihood of receiving political attention. This result provides an initial glimpse into the association between interdisciplinarity and political attention.

**Regression analysis**. After the preliminary descriptive analysis concerning interdisciplinarity and political attention, we proceed with further regression analysis. The descriptive statistics of the main variables are shown in Table 2. We also calculate the correlation coefficient matrix of the main variables in Table 3. From the correlation coefficient matrix, we observe that the correlations between variables are generally not strong, except for the correlations between certain interdisciplinarity indicators, which aligns with our expectations.

The regression results are presented in Table 4. Models 1, 2, and 3 involve multiple linear regression with independent variables varying across the three dimensions of interdisciplinarity, leading to comprehensive indicators RS and DIV. From Model 1, it can be observed that variety and disparity exhibit a positive correlation with policy citation, while balance exhibits a negative correlation with policy citation. As for RS and DIV, it can be observed from Models 2 and 3 that both of the two indicators exhibit a positive correlation with policy citation. For instance, in model 2, the probability of scientific publications being cited by policy documents increases on average by 8.2% with every unit increase in RS. Furthermore, the regression results grounded on CEM as robustness tests are presented in Table 5 and the threshold of judging interdisciplinary publications is median for Models 4, 5, and 6 and quartile for Models 7, 8, and 9. From Models 4 and 7, it can be observed that both variety and disparity consistently exhibit a positive correlation with policy citation, whereas the initially negative correlation between balance and policy citation transforms into a weak positive correlation after CEM. As for RS and DIV, it can be observed from Models 5, 6, 8, and 9 that both the two indicators consistently exhibit a positive correlation with policy citation, regardless of CEM. The regression results demonstrate good robustness. This is specifically manifested by the fact that the signs of most coefficients for the independent variables remain unchanged. Only their magnitudes have been altered, with the exception of disparity. The results of regression analysis indicate a

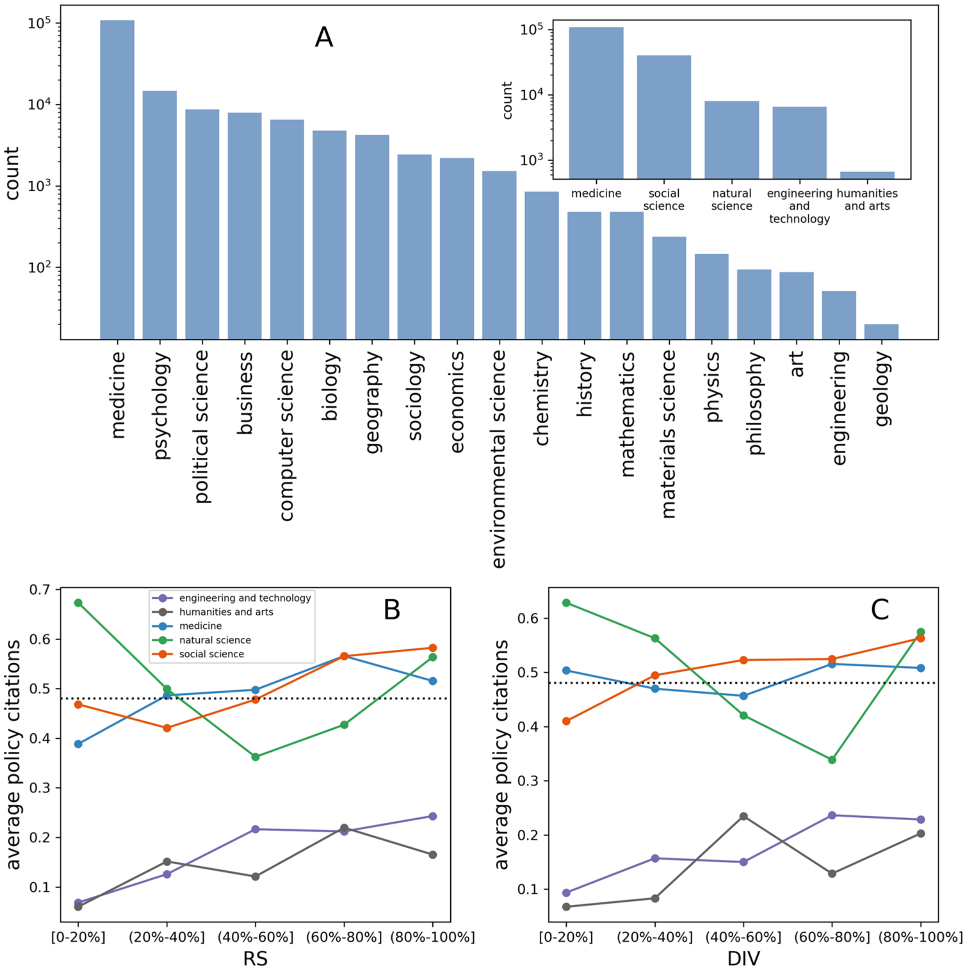


**Fig. 2 The distribution of publications and policy citations changes with variations in interdisciplinarity across different fields.** Panel **A** displays the distribution of publications across different fields based on the Level 0 discipline classification system of OpenAlex and the inner plot of panel **A** displays the distribution of publications based on the discipline classification method we define. Note that the vertical axes for both have been logarithmically scaled. Panels **B** and **C** reflect how the average policy citations change with the variation in RS and DIV values, respectively, in different fields. The black horizontal dashed lines both represent the average policy citations for all focal publications.

**Table 2 Descriptive statistics of the variables.**

| VarName | Obs | Mean | SD | Min | Q1 | Median | Q2 | Max |
|---|---|---|---|---|---|---|---|---|
| policy_cited | 159,957 | 0.142 | 0.349 | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 |
| variety | 159,957 | 0.653 | 0.168 | 0.132 | 0.532 | 0.643 | 0.766 | 1.000 |
| disparity | 159,957 | 0.520 | 0.140 | 0.000 | 0.443 | 0.537 | 0.617 | 0.880 |
| balance | 159,957 | 0.015 | 0.009 | 0.000 | 0.008 | 0.012 | 0.019 | 0.121 |
| RS | 159,957 | 0.706 | 0.126 | 0.000 | 0.652 | 0.724 | 0.784 | 0.999 |
| DIV | 159,957 | 0.033 | 0.021 | 0.003 | 0.017 | 0.027 | 0.045 | 0.445 |
| science_citation | 159,957 | 18.196 | 142.242 | 0.000 | 0.000 | 2.000 | 10.000 | 25322.000 |
| reference_count | 159,957 | 28.042 | 32.124 | 0.000 | 9.000 | 20.000 | 36.000 | 4666.000 |
| team_size | 159,957 | 6.350 | 11.426 | 1.000 | 3.000 | 5.000 | 8.000 | 1557.000 |
| journal_impact_factor | 159,957 | 6.222 | 13.380 | 0.001 | 2.155 | 4.000 | 7.070 | 505.000 |

statistically significant positive correlation between interdisciplinarity and policy citation—the stronger the interdisciplinarity of scientific research, the more likely it is to attract attention from the political sphere.

**Regression analysis for different fields**. To further capture the disciplinary difference, multiple linear regressions with and without CEM for various fields based on the previously defined discipline classification method are conducted. To avoid too few samples in each field, the median is used as the criteria for judging interdisciplinarity before CEM. Meanwhile, it is worth noting that, due to the differentiation of fields in regressions, field-fixed effects are not considered in the regression analysis. Similarly, in CEM, the matching criteria no longer include field and only include team size and time when a scientific paper was published. The matching results can be found in Table S3 of the supplementary information, and the specific regression results can be found in Tables S4–S9 of the supplementary information.

**Table 3 Pearson's correlation matrix.**

| | Policy_cited | DIV | RS | variety | disparity | balance | science_citation | reference_count | team_size | journal_impact_factor |
|---|---|---|---|---|---|---|---|---|---|---|
| policy_cited | 1 | | | | | | | | | |
| DIV | 0.022*** | 1 | | | | | | | | |
| RS | 0.028*** | 0.315*** | 1 | | | | | | | |
| variety | −0.010*** | −0.504*** | −0.175*** | 1 | | | | | | |
| disparity | 0.029*** | 0.555*** | 0.688*** | −0.001 | 1 | | | | | |
| balance | 0.025*** | 0.885*** | 0.436*** | −0.159*** | 0.740*** | 1 | | | | |
| science_citation | 0.154*** | 0.002 | 0.014*** | −0.043*** | −0.002 | −0.013*** | 1 | | | |
| reference_count | 0.009*** | 0.718*** | 0.163*** | −0.630*** | 0.195*** | 0.411*** | 0.021*** | 1 | | |
| team_size | 0.073*** | 0.021*** | −0.001 | −0.102*** | −0.012*** | −0.014*** | 0.061*** | 0.045*** | 1 | |
| journal_impact_factor | −0.025*** | 0.120*** | 0.043*** | −0.122*** | 0.051*** | 0.081*** | −0.011*** | 0.127*** | 0.015*** | 1 |

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

**Table 4 Regression results.**

| | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| | Dependent variable: policy_cited | | |
| variety | 1.065*** | | |
| | (17.024) | | |
| disparity | 0.041*** | | |
| | (5.665) | | |
| balance | −0.054*** | | |
| | (−8.166) | | |
| RS | | 0.082*** | |
| | | (13.059) | |
| DIV | | | 2.000*** |
| | | | (18.656) |
| science_citation | 0.000*** | 0.000*** | 0.000*** |
| | (46.884) | (47.405) | (47.414) |
| reference_count | −0.000*** | 0.000*** | 0.000*** |
| | (−6.693) | (12.207) | (5.911) |
| team_size | 0.002*** | 0.002*** | 0.002*** |
| | (30.391) | (31.449) | (31.391) |
| journal_impact_factor | −0.000*** | −0.000*** | −0.000*** |
| | (−5.909) | (−5.081) | (−5.318) |
| field-fixed effect | YES | YES | YES |
| time-fixed effect | YES | YES | YES |
| constant | 0.097*** | 0.058*** | 0.081*** |
| | (9.729) | (7.817) | (11.739) |
| observations | 159,957 | 159,957 | 159,957 |
| $R^2$ | 0.091 | 0.089 | 0.090 |
| Adj. $R^2$ | 0.090 | 0.088 | 0.089 |
| $F$ | 331.691 | 337.921 | 342.150 |

*T*-values are shown in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Figure 3 provides a visualization of the regression coefficients. Firstly, when considering the results from both regressions with and without CEM, it is evident that the impact of conducting CEM on the coefficients of the independent variables is not substantial, which indicates that the results are essentially robust even when minimizing the potential influence of confounding factors. Secondly, from the field of medicine to humanities and arts, the 90% confidence intervals of the coefficients of the independent variables become progressively wider, and the probability of non-significant coefficients increases, which may be attributed to the reduction in sample size and resulting in an increase in the proportion of abnormal data (refer to Tables S4–S9 of the supplementary information for specific sample sizes for each field). Furthermore, it can be observed that variety has a promoting effect on policy citation across almost all academic fields from Fig. 3A. In other words, policy documents tend to cite scientific publications with diverse discipline categories. Figure 3B demonstrates that the effects of disparity on different fields are inconsistent, showing both positive and negative effects, as well as non-significant effects. Figure 3C indicates that balance positively influences policy citation in the fields of engineering and technology yet inhibits policy citations in natural science and humanities and arts. The impact of balance on the other two fields is not significant. This could be attributed to the fact that variety is more intuitive compared to the other two dimensions and is, therefore, more likely to be considered when determining interdisciplinarity based on a publication's reference list. Regarding the comprehensive interdisciplinarity indicators, Fig. 3D illustrates that, except for the natural science and humanities and arts, greater RS values lead to a greater probability of being cited by policy documents in all other fields. Furthermore, the results from Fig. 3E indicate that, except for the natural science and humanities and arts, there is a positive correlation between DIV

**Table 5 Regression results grounded on CEM as robustness tests. The threshold of judging interdisciplinary publications is median for Models 4, 5, and 6 and quartile for Models 7, 8, and 9.**

| | Median | | | Quartile | | |
|---|---|---|---|---|---|---|
| | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 | Model 9 |
| | Dependent variable: policy_cited | | | | | |
| variety | 0.856*** | | | 1.069*** | | |
| | (14.582) | | | (11.824) | | |
| disparity | 0.040*** | | | 0.009 | | |
| | (5.567) | | | (0.971) | | |
| balance | 0.023*** | | | 0.011 | | |
| | (3.633) | | | (0.974) | | |
| RS | | 0.096*** | | | 0.090*** | |
| | | (15.571) | | | (12.290) | |
| DIV | | | 1.734*** | | | 1.682*** |
| | | | (17.292) | | | (13.571) |
| science_citation | 0.000*** | 0.000*** | 0.000*** | 0.001*** | 0.001*** | 0.001*** |
| | (66.508) | (66.302) | (66.606) | (43.305) | (43.151) | (43.333) |
| reference_count | −0.000*** | 0.000 | −0.000*** | −0.000*** | 0.000*** | −0.000 |
| | (−6.285) | (0.552) | (−3.942) | (−3.821) | (2.786) | (−0.399) |
| team_size matched | YES | YES | YES | YES | YES | YES |
| journal_impact_factor matched | YES | YES | YES | YES | YES | YES |
| field matched | YES | YES | YES | YES | YES | YES |
| time matched | YES | YES | YES | YES | YES | YES |
| constant | 0.062*** | 0.076*** | 0.104*** | 0.077*** | 0.070*** | 0.091*** |
| | (8.648) | (23.178) | (65.106) | (6.767) | (20.107) | (41.338) |
| observations | 159,465 | 159,465 | 159,465 | 56,647 | 56,647 | 56,647 |
| $R^2$ | 0.029 | 0.029 | 0.029 | 0.037 | 0.037 | 0.037 |
| Adj. $R^2$ | 0.029 | 0.028 | 0.029 | 0.037 | 0.036 | 0.037 |
| $F$ | 948.993 | 1559.503 | 1578.877 | 439.283 | 715.852 | 727.286 |

*T*-values are shown in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

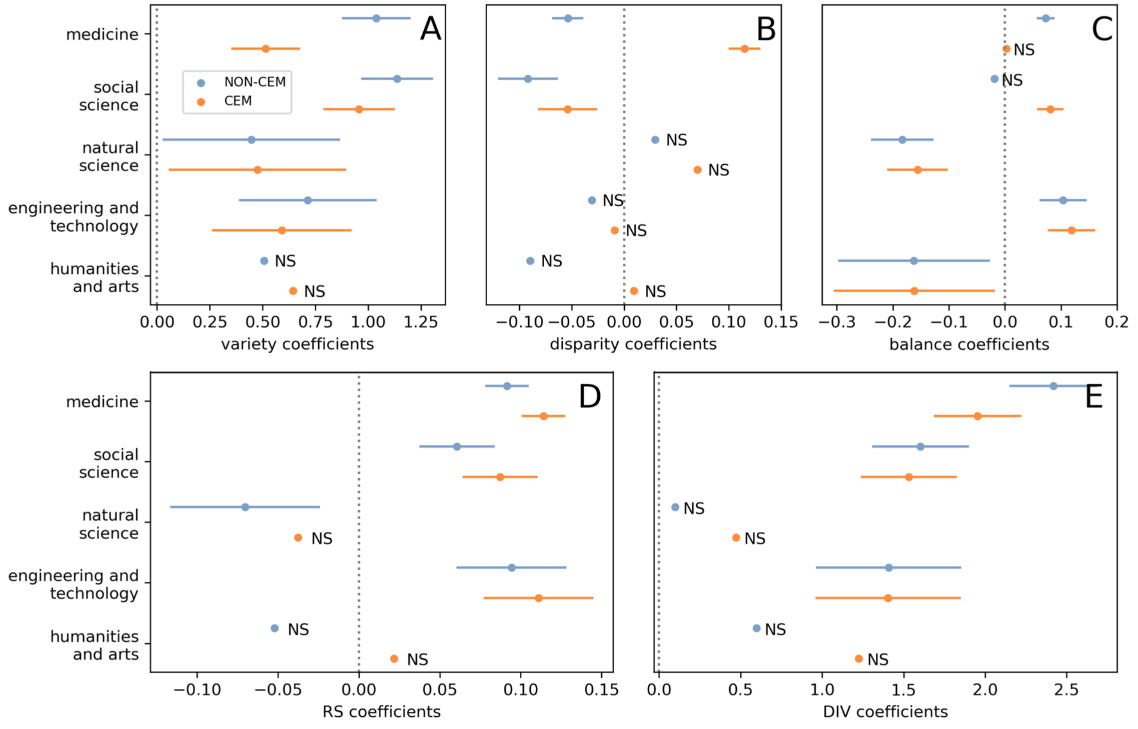


**Fig. 3 Coefficients of regressions with and without CEM for various fields.** Panels **A**, **B**, and **C** correspond to the three dimensions of interdisciplinarity, while panels **D** and **E** correspond to the comprehensive indicators RS and DIV, respectively. Whiskers represent 90% confidence intervals while "NS" indicates the coefficient is not significant. Colors indicate whether the regression is based on CEM, and the gray vertical dashed lines represent zero.

values and policy citation in all other fields, which implies that the stronger the interdisciplinarity of a scientific publication, the greater its tendency to attract attention from policy documents. In summary, despite distinguishing between different fields, there still exists varying degrees of positive correlation between interdisciplinarity and political attention.

## Discussion

In this study, we utilize metadata of scientific publications on the COVID-19 topic to explore the relationship between the interdisciplinarity of scientific research and its reception of attention from policy documents. Initially, we categorize the publications into major fields to ensure a reasonable sample size for each field. For each field, we divide the publications into bins based on interdisciplinarity indicators and conduct comparative analyses of how political attention varies with the changes in interdisciplinarity. Subsequently, we perform multiple linear regression analysis based on fixed effects for disciplines and time on the focal publications. To minimize the potential influence of confounding factors, coarsened exact matching and further regression analysis based on the matching results are conducted. Furthermore, we change the range of judging interdisciplinary publications as a robustness test. Finally, to explore more deeply into differences across disciplines, we perform direct regression and regression based on CEM for various fields and visualize the coefficients for better understanding.

Our findings indicate that there is a positive correlation between the interdisciplinarity of scientific publications and the attention they receive from policy documents in almost all fields. More specifically, the stronger the interdisciplinarity of scientific publications, the greater its ability to attract attention from policy documents. And among the three dimensions of interdisciplinarity, variety exhibits the most pronounced positive impact on political attention. That is to say, policymakers tend to cite scientific publications with diverse discipline categories. This study fills a previous research gap and provides insights for researchers and policymakers, highlighting that interdisciplinary research holds greater potential to impact policy formulation and implementation processes. In simpler terms, interdisciplinarity plays a role in facilitating the translation of scientific research into tangible policy outcomes. For researchers aiming to have their research cited in policy documents and thus exert a greater impact on policy-making, enhancing the interdisciplinarity of their research, such as referencing a wider range of publications, might be an effective strategy. On a deeper level, researchers may want to strive to engage with knowledge outside their research field and integrate it with existing knowledge to achieve innovation and thus gain more policy attention. Moreover, collaborating with scholars from fields beyond their own can also foster interdisciplinarity, ultimately leading to innovative outcomes. For policymakers, this research can enhance their understanding of the significance of interdisciplinary research with more empirical evidence showing the benefit of harnessing research findings from the scientific community to have a positive impact on society.

There are undoubtedly certain limitations in this study. Firstly, our dataset is confined to the COVID-19 topic, which might lack generalizability to other areas of research. Secondly, our method of measuring whether scientific publications receive policy attention by checking if they are cited by policy documents is relatively coarse. Future research could consider using the number of citations from policy documents as the dependent variable in regression analysis, for a more fine-grained exploration. Additionally, the measurement of interdisciplinarity currently lacks a universally accepted method, and the RS and DIV indicators we used are only relatively common. Finally, due to data constraints, we are going to add more control variables in our regression models that would enhance the model's reliability in the near future.

## Data availability

The main datasets we adopted are OpenAlex and Overton. OpenAlex is an open-available datasets that can be fully downloadable from https://openalex.org/. Overton offers a paid subscription service with purchase. The datasets generated during and analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 4 November 2023; Accepted: 26 February 2024;
Published online: 11 March 2024

## References

Bornmann L (2013) What is societal impact of research and how can it be assessed? A literature survey. J Am Soc Inf Sci Technol 64(2):217–233. https://doi.org/10.1002/asi.22803

Bornmann L, Haunschild R, Marx W (2016) Policy documents as sources for measuring societal impact: how often is climate change research mentioned in policy-related documents? Scientometrics 109(3):1477–1495. https://doi.org/10.1007/s11192-016-2115-y

Bozeman B, Sarewitz D (2011) Public value mapping and science policy evaluation. Minerva 49(1):1–23. https://doi.org/10.1007/s11024-011-9161-7

Bu Y, Li M, Gu W, Huang W (2021) Topic diversity: a discipline scheme-free diversity measurement for journals. J Assoc Inf Sci Technol 72(5):523–539

Carayol N, Thuc Uyen NT(2005) Why do academic scientists engage in interdisciplinary research? Res Eval 14(1):70–79. https://doi.org/10.3152/147154405781776355

Chavarro D, Tang P, Rafols I (2014) Interdisciplinarity and research on local issues: evidence from a developing country. Res Eval 23(3):195–209. https://doi.org/10.1093/reseval/rvu012

Choi BCK, Pak AWP (2006) Multidisciplinarity, interdisciplinarity and transdisciplinarity in health research, services, education and policy: 1. Definitions, objectives, and evidence of effectiveness. Clin Investig Med Med Clin Et Exp 29(6):351–364

Fontana M, Iori M, Montobbio F, Sinatra R (2020) New and atypical combinations: an assessment of novelty and interdisciplinarity. Res Policy 49(7):104063. https://doi.org/10.1016/j.respol.2020.104063

Frodeman R, Mitcham C (2007) New directions in interdisciplinarity: broad, deep, and critical. Bull Sci Technol Soc 27(6):506–514. https://doi.org/10.1177/0270467607308284

Gunn W (2013) Social signals reflect academic impact: what it means when a scholar adds a paper to mendeley. Inf Stand Q 25(2):33. https://doi.org/10.3789/isqv25no2.2013.06

Ho DE, Imai K, King G, Stuart EA (2007) Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Political Anal 15(3):199–236. https://doi.org/10.1093/pan/mpl013

Iacus SM, King G, Porro G (2011) Multivariate matching methods that are monotonic imbalance bounding. J Am Stat Assoc 106(493):345–361. https://doi.org/10.1198/jasa.2011.tm09599

Iacus SM, King G, Porro G (2012) Causal inference without balance checking: coarsened exact matching. Political Anal 20(1):1–24. https://doi.org/10.1093/pan/mpr013

Larivière V, Haustein S, Börner K (2015) Long-distance interdisciplinarity leads to higher scientific impact. PLoS ONE 10(3):e0122565. https://doi.org/10.1371/journal.pone.0122565

Levitt JM, Thelwall M (2008) Is multidisciplinary research more highly cited? A macrolevel study. J Am Soc Inf Sci Technol 59(12):1973–1984. https://doi.org/10.1002/asi.20914

Lewison G, Sullivan R (2008) The impact of cancer research: how publications influence UK cancer clinical guidelines. Br J Cancer 98(12):1944–1950. https://doi.org/10.1038/sj.bjc.6604405

Leydesdorff L, Rafols I (2011) Indicators of the interdisciplinarity of journals: diversity, centrality, and citations. J Informetr 5(1):87–100. https://doi.org/10.1016/j.joi.2010.09.002

Leydesdorff L, Wagner CS, Bornmann L (2019) Interdisciplinarity as diversity in citation patterns among journals: rao-Stirling diversity, relative variety, and the Gini coefficient. J Informetr 13(1):255–269. https://doi.org/10.1016/j.joi.2018.12.006

Liu S, Chen H, Bu Y (2022) Comparing different perspectives of characterizing interdisciplinarity of scientific publications: author vs. publication perspectives. Proceedings of the 22nd ACM/IEEE Joint Conference on Digital Libraries, 1–10. https://doi.org/10.1145/3529372.3530914

OECD (2015) Scientific advice for policy making: the role and responsibility of expert bodies and individual scientists (OECD Science, Technology and Industry Policy Papers 21; OECD Science, Technology and Industry Policy Papers, vol. 21). https://doi.org/10.1787/5js33l1jcpwb-en

Petersen AM, Fortunato S, Pan RK, Kaski K, Penner O, Rungi A, Riccaboni M, Stanley HE, Pammolli F (2014) Reputation and impact in academic careers. Proc Natl Acad Sci USA 111(43):15316–15321. https://doi.org/10.1073/pnas.1323111111

Piwowar H (2013) Value all research products. Nature 493(7431):159–159. https://doi.org/10.1038/493159a

Priem J, Piwowar H, Orr R (2022) OpenAlex: a fully-open index of scholarly works, authors, venues, institutions, and concepts. http://arxiv.org/abs/2205.01833

Rafols I, Meyer M (2010) Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience. Scientometrics 82(2):263–287. https://doi.org/10.1007/s11192-009-0041-y

Rao CR (1982) Its measurement, decomposition, apportionment and analysis. Indian J Stat Ser A (1961-2002) 44(1):1–22

Salter A, Salandra R, Walker J (2017) Exploring preferences for impact versus publications among UK business and management academics. Res Policy 46(10):1769–1782. https://doi.org/10.1016/j.respol.2017.08.005

Stirling A (2007) A general framework for analysing diversity in science, technology and society. J R Soc Interface 4(15):707–719. https://doi.org/10.1098/rsif.2007.0213

Szomszor M, Adie E (2022) Overton: a bibliometric database of policy document citations. Quant Sci Stud 3(3):624–650. https://doi.org/10.1162/qss_a_00204

The National Academies (2004) Facilitating interdisciplinary research (p. 11153). National Academies Press. https://doi.org/10.17226/11153

Uzzi B, Mukherjee S, Stringer M, Jones B (2013) Atypical combinations and scientific impact. Science 342(6157):468–472. https://doi.org/10.1126/science.1240474

Wagner CS, Roessner JD, Bobb K, Klein JT, Boyack KW, Keyton J, Rafols I, Börner K (2011) Approaches to understanding and measuring interdisciplinary scientific research (IDR): a review of the literature. J Informetr 5(1):14–26. https://doi.org/10.1016/j.joi.2010.06.004

Wang J, Thijs B, Glänzel W (2015) Interdisciplinarity and impact: distinct effects of variety, balance, and disparity. PLoS ONE 10(5):e0127298. https://doi.org/10.1371/journal.pone.0127298

Wang K, Shen Z, Huang C, Wu C-H, Dong Y, Kanakia A (2020) Microsoft academic graph: when experts are not enough. Quant Sci Stud 1(1):396–413. https://doi.org/10.1162/qss_a_00021

Xie J, Gong K, Li J, Ke Q, Kang H, Cheng Y (2019) A probe into 66 factors which are possibly associated with the number of citations an article received. Scientometrics 119(3):1429–1454. https://doi.org/10.1007/s11192-019-03094-z

Zhang L, Rousseau R, Glänzel W (2016) Diversity of references as an indicator of the interdisciplinarity of journals: taking similarity between subject fields into account. J Assoc Inf Sci Technol 67(5):1257–1265

Zhou H, Guns R, Engels TC (2021) Are social sciences becoming more interdisciplinary? Evidence from publications 1960–2014. J Assoc Inf Sci Technol 73(9):1201–1221

## Acknowledgements

The authors acknowledge the financial support by the National Natural Science Foundation of China (No. 72104007 and No. 72174016) and by the Fundamental Research Funds for the Central Universities, Peking University. The authors are grateful to Jialin Liu and other members of the Knowledge Discovery lab, at Peking University. The authors wish to thank Overton for offering access to the database.

## Author contributions

Conceptualization: LH and YB; methodology: LH and YB; formal analysis: LH; writing (original draft preparation): LH; writing (review and editing): WH and YB; visualization: LH; supervision: WH and YB. All authors have read and agreed to the published version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

## Informed consent

This article does not contain any studies with human participants performed by any of the authors.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1057/s41599-024-02915-8.

**Correspondence** and requests for materials should be addressed to Yi Bu.