INSTITUTE
OF ECONOMICS

Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM
WORKING PAPER SERIES

The interdisciplinarity dilemma: public versus private interests

Magda Fontana [a,b]
Martina Iori [c]
Valerio Leone Sciabolazza [d]
Daniel Souza [b]

[a] Despina Big Data Lab, University of Turin, Italy.
[b] Department of Economics and Statistics, University of Turin, Italy.
[c] Institute of Economics & EMbeDS, Sant'Anna School of Advanced Studies, Pisa, Italy.
[d] Department of Business and Economics, University of Naples Parthenope, Italy.

2021/34 October 2021
ISSN(ONLINE) 2284-0400

The interdisciplinarity dilemma: public versus private interests

Magda Fontana[a,b], Martina Iori[c], Valerio Leone Sciabolazza[d], and Daniel Souza[b]

[a]Despina Big Data Lab, University of Turin, Lungo Dora Siena, 100A, Turin, 10153, Italy
[b]Department of Economics and Statistics, University of Turin, Lungo Dora Siena, 100A, Turin, 10153, Italy
[c]Institute of Economics & EMbeDS, Sant'Anna School of Advanced Studies, Piazza Martiri della Libertà, 33, 56127, Pisa, Italy
[d]Department of Business and Economics, University of Naples Parthenope, Via G. Parisi 13, 80132, Naples, Italy

Abstract

In this paper, we investigate how the choice to conduct interdisciplinary work affects a researcher's career. Using data on 23,926 articles published by 6,105 researchers affiliated with the University of Florida in the period 2008-2013, we show that synthesizing knowledge from diverse fields pays off in terms of reputation. However, if combining too-distant research fields, the impact of a work is penalized. Moreover, research conducted balancing the contribution of different scientific fields has a negative impact on the reputation of scientists in terms of the number of citations but a positive impact on the diffusion of knowledge across other disciplines. Our findings are robust to a number of controls, including individual, time, and field of study fixed effects, and they apply to all investigators regardless of their gender, collaboration behavior, performance, and affiliation. All in all, despite its public benefits, interdisciplinary research comes with a cost for a researcher's academic career. This trade-off poses challenging questions to policymakers.

Keywords: Interdisciplinarity; Research Policy; Academic Career; Generality; Incentives; Citations

JEL codes: I23; H5; 039

1 Introduction

Researchers often receive contrasting incentives when conducting their work. On the one hand, an interdisciplinary approach is required to produce effective team work and access to funding and research institutions. On the other hand, academic scholarships and evaluation mechanisms are still organized following the criteria of traditional disciplinary fields. As a result, investigators may not always find it profitable to pursue interdisciplinary research (IDR).

IDR represents a crucial driver for knowledge progress nowadays. In recent years, in fact, scientists have narrowed their expertise and specialization, and rely more and more on team-work joining different (sub)fields of specific knowledge to produce wide-ranging scientific advances (Cedrini and Fontana, 2018; Jones, 2009; Larsen and Ins, 2010). This behavior is the natural reaction to the increasing educational burden faced by recent generations of investigators. In present days, the soaring amount of knowledge accumulated in published articles is conveyed to prospective scientists through doctoral programs of longer duration and a larger post-doctoral training (Jones, 2010). The consequence is that a longer time is now required before first publishing (Conti and Liu, 2015), and the 'age at great achievement' (a proxy for educational attainment) has significantly risen among late trainee cohorts (Jones et al., 2014; The National Academies, 1998). In order to compensate the present burden of knowledge, scientists often seek to specialize in specific fields.

The growing importance of IDR is witnessed by the recent push by funding and research institutions for overcoming disciplinary barriers (Rylance, 2015), the reorganization of universities into interdisciplinary research centers (Biancani et al., 2018; Hackett et al., 2021), and the increasing trend of citation flows across disciplines in several fields of study (Angrist et al., 2020; Battiston et al., 2019). Despite this drive towards IDR, however, academic scholarship and its assessment mechanisms are still organized in separated disciplines or even in subfields. The specialization of journals (Stigler et al., 1995) together with the decreasing importance of generalist journals (Goel and Faria, 2007, p. 538) suggests that academic reputation tends to be built within niches. Moreover, the increasing importance of the rankings of field-specific journals – used to evaluate research performances of universities, departments, and individual scholars and then to assign funds and make hiring decisions (Cedrini and Fontana, 2018; Ritzberger, 2008) – render the interdisciplinary effort rather risky. Indeed, previous research shows that higher levels of interdisciplinarity are often associated with lower scientific impact – number of citations – and productivity (Leahey et al., 2017).

The consequences of adopting IDR on the career of a scholar remain largely unexplored, however. For this reason, in this paper we ask whether the undertaking of IDR, that is naturally aimed at merging and reaching diverse disciplines, is compatible with the goals of scholars in terms of career and reputation. If the answer is negative, then a crucial trade-off would emerge between the societal need of sustaining growth and innovation through IDR and the individual incentives to go interdisciplinary.

In order to tackle our research question, we present an econometric analysis on the impact that the decision to adopt an interdisciplinary approach in conducting research has on the scientific impact of a scholar. In doing so, we add to the extant literature by: 1) exploiting a novel and unique dataset of 6,105 researchers affiliated to the University of Florida (UF) along with their publication records (23,926 articles) and individual characteristics (gender, affiliation) over the period 2008-2013.[1] Albeit small in comparison with the samples used in other studies (Yegros-Yegros et al., 2015), our dataset has the unique feature of providing information about a panel of scholars operating in a wide range of scientific fields and affiliated to the same university. This allows sorting out a number of confounding factors often neglected by the literature, as the role played by institutional and national heterogeneity in determining the scientific impact of an investigator through an article, thereby increasing the accuracy of our analysis; 2) focusing on the effect of IDR on the scientific production of individual scholars. Because of the panel nature of our dataset, in fact, we are able to observe the effect of varying the degree of interdisciplinarity across articles by the same scholar.[2] With respect to extant literature (see, for instance, Yegros-Yegros et al., 2015), therefore, we are able to account for the investigators' individual characteristics that may play a crucial role in determining the impact of a scholar obtained with a published article. At the same time, by performing our analysis at the article level and comparing papers of the same researcher (through individual fixed effect), we avoid arbitrary aggregations of data at the researcher level (Leahey et al., 2017) and we test the individual's incentives in pursuing IDR; 3) accounting for a dimension, so far unexplored, of interdisciplinarity in science: the generality of knowledge. Interdisciplinarity has been intended only as a way to put together different sources of knowledge but, it is our conviction, that it also implies a wider circulation of such knowledge. As societies become more interconnected and grow in complexity, science needs to acquire knowledge from different domains in order to face

[1] The University of Florida is a large research university in the United States that comprises more than 5,000 researchers and 50,000 students. UF consistently ranks among the top ten public universities in the United States and is the flagship university in the state of Florida.

[2] In principle, also other dataset such as MAG may allow creating a longitudinal dataset about scholars using an identification code. However, such identification code are obtained through inferential methods, and they are not directly registered by the scholar or her/his institution. On the contrary, our information is more reliable since the association of articles to the same scholars is done by the UF, and there in no inference involved.

its challenges – for instance, environmental economics communicates with climatology, biology and ecology – but also shares its findings with these disciplines. Thus, it is strategic for the development of science that domain-specific knowledge could be exploited in other disciplines. We, therefore, account for a scholar contribution to the overall development of science by observing how the knowledge that she produces spills over domains others than hers.

When conducting our investigation, we measure the scientific impact of a scholar by looking at both the number of citations accrued by the papers written by the scholar, and their generality. Indeed, academic careers are characterized by their dependency on the community assessment and perception, thus accruing scholarly citations is a major mechanism of reputation signaling as well as of the relevance of the knowledge embedded in articles (Hamermesh and Pfann, 2012; Jones, 2021). In our perspective, citations across diverse disciplines also indicates the relevance of such knowledge beyond its field of origin, i.e. their generality. Generality is an index derived from the Hirschman-Herfindahl Index that measures the dispersion of citations across fields of study. The index is widely used to measure the range of inventions that derive from a patent by means of measuring the span of technology fields that cite the patent (Bresnahan and Trajtenberg, 1995; Squicciarini et al., 2013). Following Fontana et al. (2020) and Carley and Porter (2012), we apply it to the spreading of scientific knowledge and interpret it as a measure of the usefulness of the knowledge produced by a scholar – her impact – to other fields of study.

A more compelling task is to measure interdisciplinarity, that is the main independent variable of interest in our analysis. The literature on the measurement of interdisciplinarity is vast and variegated: while the concept is very intuitive,[3] its quantification can be focused on different aspects. In order to make our results more general and readable in comparison with the extant body of evidence, we highlight the distinct dimensions of IDR through three indicators (Porter and Rafols, 2009; Yegros-Yegros et al., 2015): the number of fields embedded in a paper (Variety); the evenness of their distribution (Balance), and the similarity between them (Disparity).[4] The use of multiple and distinct indicators allows capturing all the facets of a complex concept like interdisciplinarity.

[3]Interdisciplinarity 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" (The National Academies, 2005).

[4]The literature also uses the Integration Score (Stirling, 2007), an index that synthesizes the three dimensions. In addition to the loss of details, it has been shown (Fontana et al., 2020, Figure 10) that the Integration Score is highly correlated with Disparity, we therefore decided not to include it in our analysis.

Our identification strategy relies primarily on the use of individual, disciplinary-based citation norms, and year fixed effects, which allow registering the effect of a change in a specific dimension of the interdisciplinarity content of a study on the scientific impact of a researcher, while sorting out potential confounding factors and the influence of a change in other interdisciplinary dimensions. The information contained in our database, moreover, give us the chance to shed light on different sources of heterogeneity, and assess whether the impact of IDR differs across gender, collaboration types, research proficiency, and disciplinary affiliation. Importantly, our analysis cannot be biased by the potential role played by institutional factors on the scientific impact of an investigator, given that we consider a single university and researchers do not change departments in the observed period.

Our findings highlight the existence of a trade-off between private benefits in the form of greater academic reputation among peers and public returns arising from the diffusion of knowledge across disciplines when Balance is concerned. Indeed, a 10 percentage point increase in the evenness of the distribution of fields of study in the references contained in a study (i.e. thus increasing the balance of contributions from different scientific fields) results in a decrease of 35% in the number of citations received by a paper published by a researcher, but increases her extramural impact by 2%. This evidence suggests that, in addition to promoting funding and favoring the creation of synthesis centers, research policy should also act on individual incentives and rules of scholars' career assessment and hiring.

We also identify a trade-off in scholars' private benefits: synthesizing knowledge from diverse fields pays-off in terms of citations and generality, however, when combined fields are distant the effect of IDR on citations and generality is negative, implying that a scholar increases interdisciplinarity at the expenses of the impact and diffusion of her scientific contributions. Namely, the number of unique fields of studies in the references of an article – Variety – has a significant and positive impact in both total number of citations and generality, which shows that tapping into a diverse knowledge pool increases the research impact. However, when we consider the average dissimilarity between these fields – Disparity –, the effect of IDR on impact is negative.

Importantly, we find evidence that our results are confirmed even when considering scholars with different characteristics or affiliations. In other words, all scholars face the similar incentives and constraints in engaging in more interdisciplinary projects. Regardless of their characteristics or affiliations, the effects of IDR are the same for all research activities at the University of Florida.

The paper proceeds as follows. Section 2 presents the theoretical background and research hypotheses. In Sections 3 and 4, we describe our empirical strategy and data, respectively. Section 5 presents the results, and Section 6 concludes.

2 Interdisciplinarity research and researchers' incentives

While interdisciplinary research has been widely recognized as crucial to address the complex issues faced by modern societies, the current academic system, and especially its research evaluation and funding procedures (Geuna and Martin, 2003), might not provide the correct scholars' incentives to foster the adoption of IDR practices (Arnold et al., 2021).

The existing literature on IDR primarily focuses on scholars' scientific outcomes, rather than researchers themselves (Leahey and Barringer, 2020; Hackett et al., 2021). Several studies highlighted the mixed effect of the various aspects of interdisciplinarity on scientific impact, measured as the number of citations received by single articles (see, among others, Fontana et al., 2020; Yegros-Yegros et al., 2015).[5] Results vary across the dimensions of IDR and disciplines took into account. However, those studies commonly identified an inverted U-shaped relationship between the interdisciplinarity and impact of a single article. Moving from articles to research projects and grants, Bromham et al. (2016) suggested the existence of a bias against interdisciplinarity in funding evaluations.

Another stream of literature, instead, investigates the role of various forms of diversity in researchers' and inventors' team composition. Wu et al. (2019) showed that scientific and technological disruption is associated with smaller teams, while larger and more variegated teams are in charge of developing incremental steps starting from the existing literature. Anderson and Richards-Shubik (2021) highlighted the increasing importance of collaborations in Economics, suggesting the existence of a higher reward for larger teams in terms of citations. Concerning the team composition, recent studies highlighted that diversity, in different aspects (gender, education, ethnicity, culture), fosters innovations (Østergaard et al., 2011; Crescenzi et al., 2016). Moreover, migration and cultural diversity improve also the quality of the resulting innovation (Ferrucci and Lissoni, 2019).

[5] For a survey of the literature on interdisciplinarity see Wagner et al. (2011), for a review on the relationship between interdisciplinarity and impact see Zeng et al. (2017, section 6.1.1).

Although the effect of interdisciplinary and diversity on the knowledge production and scientific impact has been extensively studied in the literature, the impact of pursuing IDR on scholars' career is still underexplored. Leahey et al. (2017) provided one of the first study on potential scholars' costs and benefits associated with interdisciplinarity research. They collected 32,000 articles published by 854 researchers from a wide range of fields and universities. The authors computed researcher-level bibliometric indicators by considering scholars' publications in the entire period of analysis. Overall, they found that an increase in the average interdisciplinarity of scholars' work improves their visibility in the scientific community, measured as the cumulative number of citations, and decreases their productivity, as indicated by the number of articles published.

In what follows, instead, we propose a rigorous study of interdisciplinarity research and its impact by looking at potential trade-offs between public interest in fostering IDR and scholars' private benefit in the current academic system.

2.1 Measuring Interdisciplinarity

The indicators adopted in the paper measure interdisciplinarity as the Diversity of the combined knowledge, i.e. "the apportioning of elements or options in any system" (Stirling, 2007). Diversity consists of three independent components: Variety, Balance, and Disparity.[6] The three dimensions of Diversity have a specific meaning and autonomy, and refer respectively to the number of different disciplines involved in the making of the paper, their relative frequency, and their distance (Fontana et al., 2020; Porter and Rafols, 2009; Yegros-Yegros et al., 2015). We compute these indicators by using the disciplines of the papers listed in the references of the focal articles.

Variety is the basic form of interdisciplinarity: it returns the number of different disciplines that are referenced in the paper. Thus, we define Variety (V_j) as:

$$V_j \equiv \sum_{s \in F} 1, \tag{1}$$

where F is the set of disciplines s in references of a paper j. Variety provides *prima facie* evidence on the intensity of interdisciplinarity of an article, but gives no information on the relative importance of the involved disciplines.

[6] Diversity also includes a compound indicator, the Rao-Stirling diversity, that is more suitably computed when the distinct role of the IDR components is not relevant to the object of analysis.

Balance overcomes this drawback by building on Variety in order to quantify the distribution of disciplines in the references. Namely, it measures the evenness of the distribution of the disciplines in the references. We operationalize the Balance (B_j) as a normalized Shannon Entropy, defined as:

$$B_j \equiv \frac{1}{\log V_j} \sum_{s \in F} f_s \log f_s, \tag{2}$$

where V_j is the Variety measured as above and f_s is the frequency of discipline s in references of paper j. After normalization, this index assumes values from 0 to 1. Low values of Balance indicate that the focal paper references articles from a prevailing discipline. While, high values of Balance correspond to an even distribution of disciplines in the references.

Disparity measures a further dimension of Diversity: the proximity of the referenced disciplines in the knowledge space. The underlying idea is that disciplines that frequently, relative to all other occurrences, co-occur in references are closer than those that co-occur rarely with respect to all other occurrences. High values of Disparity signal that a paper references fields that are very distant – have a low proximity – in the knowledge space. This indicator is rather different from Variety and Balance in that it does not heavily depend on the system of data classification as they do: proximity is calculated over the entire sample of articles and, therefore, provides the effective relative distance between pairs of disciplines. Disparity (D_j) is defined as the normalized sum of proximity among fields:

$$D_j \equiv \frac{1}{V_j(V_j - 1)} \sum_{\substack{r,s \in F \\ r \neq s}} (1 - p_{rs}), \tag{3}$$

where p_{rs} is the proximity between disciplines r and s. The computation of proximity is usually based on the co-occurrence of disciplines in articles, normalized by the size of fields. A common indicator is cosine similarity, which measures the cosine between fields' vectors of co-occurrences in references. Disparity is defined for values between 0 and 1 and is independent of Variety and Balance. It is worth noting that Balance and Disparity are not defined for articles that cite only one discipline (i.e. when Variety is equal to one).

Figure 1 exemplifies the three measures of interdisciplinarity in the case of a paper that cites three unevenly-distributed disciplines, with different proximity to each other.

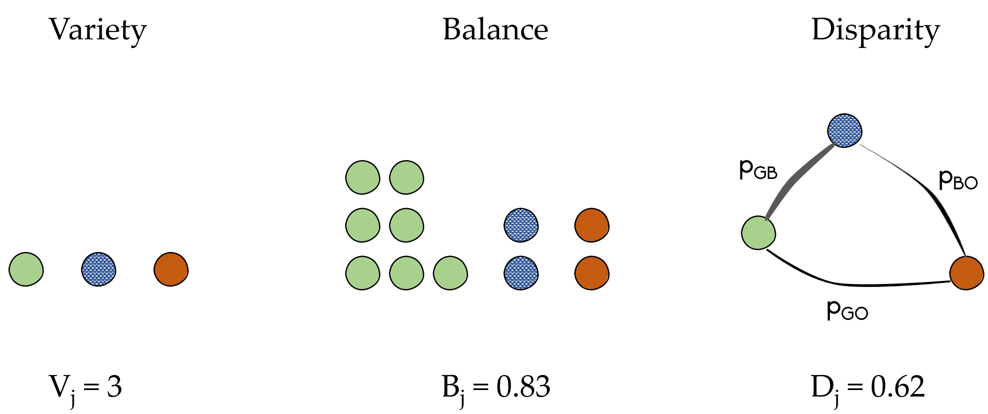

Figure 1: Example to illustrate the IDR measures. The example article cites three different disciplines (Green, Blue, Orange), with a prevalence of Green (7) over Blue (2) and Orange (2). In this example, Green and Blue are similar to each other (they are often cited together, i.e. they frequently co-occur in references), while Orange is more distant.

2.2 Interdisciplinarity and researchers' trade-offs: research hypotheses

In our empirical analysis, we explore the hypothesis that Diversity could influence the number of citations accrued by a scholar and the circulation of her papers across disciplines. Several mechanisms exist through which this might happen, and the existence and extent of the supposed trade-off between private and public benefits might also vary considerably across the dimensions of Diversity.

Firstly, there might exist a trade-off between the different dimensions of Diversity. For instance, increasing Variety implies that the pool of possible citing scholars increases and that, possibly, this impacts positively both on the number of citations and the diffusion of knowledge across fields. However, this might not hold when the referenced disciplines are very distant the one from the other or when the focal paper is hardly identifiable with a field of study and, likely, less useful for a wide range of disciplines. This results in a trade-off for the researcher, since increasing Variety will eventually end up in increasing Diversity. Combining these insights, we developed the second hypothesis that we will test in our empirical analysis:

Hypothesis 1 (HP1): *If IDR has an effect on the scholars' reputation and circulation of knowledge, this impact differs across the various dimensions of Diversity. A trade-off in scholars' private benefits exists: Variety and Disparity have an opposite effect on both indicators of scientific impact.*

Moreover, while the increase in Diversity is likely to positively affect the circulation of knowledge (public benefit), it might penalize the scholar prestige in a highly specialized academic environment (private benefit). This aspect might be particularly relevant for some dimensions of Diversity, like Balance: an even distribution of references to different disciplines will probably encourage the diffusion of the paper across a wide range of fields, but, at the same time, the paper will not have a target scientific community and will hardly be highly cited. We, therefore, test the following hypothesis:

Hypothesis 2 (HP2): *If IDR has an impact on the scholars' reputation and circulation of knowledge, the effect differs across these two indicators of scientific impact. A trade-off between public and private benefits in pursuing IDR exists: the increase of Balance in articles hampers receiving a high number of citations, while it favors knowledge diffusion.*

To test these hypotheses, we study, through a regression analysis (see below Section 3 for more details), the effect of the different dimensions of interdisciplinarity on researchers' academic prestige and the spread of knowledge across disciplines. Our unit of analysis is each article published in a specific year by an UF investigator in our time-window. We match each paper-researcher pair to the two outcome variables that measures researchers' reputation and knowledge diffusion.

We operationalized researchers' influence and prestige in academia as the total number of citations received in a five-year period after the publication date (Hamermesh and Pfann, 2012). It is described using the sum:

$$C_j \equiv \sum_{t=y_{pub}}^{y_{pub}+5} c_{jt}, \tag{4}$$

where y_{pub} is the article's publication year and c_{jt} represents the citations received by a paper j in year t. We count citations over a five-year time window to have a measure that is consistent between papers published in different years.

Furthermore, we used an index of generality of knowledge to measure knowledge diffusion: the spillover of knowledge outside disciplinary boundaries constitutes a measure of impact that complements the count of citations. A bit of knowledge that influences many, possibly distant, disciplines can be thought of as more impactful than one that is received only by few disciplines (Carley and Porter, 2012). This index captures the degree of applicability and influence of knowledge of paper on different fields of study. It is computed using the Hirschman-Herfindahl concentration index of citations across disciplines (Hall et al., 2001; Trajtenberg et al., 1997) and is defined as:

$$G_j \equiv 1 - \sum_{f=1}^{|F|} \frac{N_{jf}^2}{N_j^2}, \tag{5}$$

where N_{jf} is the number of forward citations received by a paper j from papers in the field of study f, while N_j is instead the total number of forward citations received by the paper. By definition, Generality is bounded between 0 and 1. Articles having their citations spread among many disciplines will have a high value of this indicator. This is our proxy for extramural impact. One shortcoming of this measure is that it is not defined in articles that did not receive any citations in the five-year windows. This may lead to selection bias concerns that are discussed in the following sections.

3 Empirical strategy

In our empirical analysis we examine the variation of citations received by papers written by the same scientist, within the same field of research, published in the same year, but with a different interdisciplinary content in one dimension, while accounting for changes in other interdisciplinary dimensions. To this purpose, we use the following model to run ordinary least square regressions of the form:[7]

$$Y_{ijft} = IDR_{ijft}\beta + X_{it}\gamma + K_{jf}\delta + \alpha_i + \phi_f + \theta_t + \epsilon_{ijft} \tag{6}$$

where Y_{ij} registers the scientific impact of paper j written by investigator i at time t in the field of study f, and it is measured alternatively by looking at the number of citations of paper j and its generality index (see Section 2.2 for their definitions); IDR_{ijft} measures the various interdisciplinarity dimensions of paper j as defined in Section 2.1 (i.e. Variety, Balance, and Disparity);

[7]We ran Poisson and Negative Binomial regressions for the specification that have the number of citations as dependent variable as a robustness check and obtained similar results. Results are available in Table B1 in Appendix B.

the variable X_{it} registers the h-index of investigator i at time t, that is an author-level metric that measures cumulative productivity and citation impact of the researcher; the variable K_{jf} includes a set of characteristics of the paper j, i.e. the number of authors, the presence or not of collaborators affiliated to an institution outside the United States, the adoption or not of a monodisciplinary approach;[8] α_i represents investigator fixed effects; ϕ_f includes a set of fixed effects controlling for the fields of study of paper j, and it accounts for the specific norms of citation in a specific academic context; and finally θ_t is a vector of year fixed effects.

In order to avoid over-weighting extreme values in our estimates, and correctly deal with the highly skewed nature of our continuous variables, these are all log-transformed. The descriptive statistics for these variables in our data are presented in Table 1.

4 Data

For our analysis, we construct a novel and unique dataset that includes detailed information about researchers and their publications: we study all the researchers affiliated to the University of Florida in the period 2008-2013. From the UF's registry office, we obtained information on researchers' gender, department affiliation, and publication record.[9] Then, by matching articles' title, authors, year, and journal with the traditional scholarly literature indexers Crossref and Scopus, we assign a unique identifier (DOI) to each publication to complement the data provided by the registry office with bibliometric information. More specifically, through the DOI, we retrieved citations and references from the Lens database, while we collected papers' fields of studies and international collaborations through the affiliations of coauthors from the Microsoft Academic Graph (MAG) database.[10] We exploit information about citations received by papers to compute the H-index for researchers, that we will use as a proxy for the quality of researchers.[11] The matching procedure is described in Appendix A.1.

[8] Please observe that in these cases Balance and Disparity are not defined.

[9] We focus on articles published in peer-reviewed journals, excluding books and other types of academic production from our analysis.

[10] The two databases used to complement our knowledge on articles by UF's researchers are becoming widespread for bibliometric analysis in the recent years. They are free for research and available at the following link: Microsoft Academic Graph and Lens.

[11] The h-index is an author-level metric that measures cumulative productivity and citation impact of each researcher. It takes into account the scholar's best cited papers and their number of citations. A researcher with n papers with at least n citations will have a h-index of n.

Table 1: Summary statistics

Variables	*Mean*	*SD*	*Min*	*50%*	*Max*	*Obs*
Panel A: Researcher-level Data						
Nb. papers/year	2.22	2.08	1	1.6	40.33	6,105
Nb. citations/year	17.65	42.37	0	5.0	955.50	6,105
H-index	2.37	2.88	0	1.5	35.83	6,105
Gender (Woman=1)	0.34	0.48	0	0	1.00	6,105
Panel B: Paper-level Data						
Nb. Citations	20.30	46.34	0	10	2,530	23,926
Generality	0.72	0.18	0	0.77	0.98	22,658
Variety	37.06	19.54	1	36	153	23,926
Balance	0.84	0.09	0	0.85	1	23,926
Disparity	0.68	0.07	0	0.70	0.94	23,926
Nb. References	40.21	33.01	1	34.00	926	23,926
Nb. of Authors	5.64	9.90	1	4	1,269	23,926
International Collab.	0.23	0.42	0	0	1	23,926

Notes: Panel A shows selected measures of productivity of 6105 researchers affiliated to the University of Florida from year 2008 to 2013. Gender is a dummy variable that assumes the value 1 when the researcher is a woman. Panel B shows descriptive statistics of the 23,926 articles published by these researchers in the time window 2008-2013. Nb. Citations is the total number of citations received in a 5 years time after the publication. The Generality captures the degree of applicability of the knowledge codified in a paper on different fields of study. It is worth noting that generality is not defined for papers with zero citations. International collaboration is a dummy variable that assumes the value 1 when at least one coauthor in the paper is affiliated to an institution outside the United States.

In the period 2008-2013, we observe 6,105 researchers at the UF with at least one article in a peer-reviewed journal, of which 34% are women. At the UF, researchers belongs to different colleges, which, in turns, are aggregated in academic units. Researchers in Medical Sciences, especially Medicine, prevail in our sample (more details in Table A1). Summary statistics about researchers are in Table 1 - Panel A.

The final database contains 23,926 articles, that made 646,280 references and received 366,024 citations in five years from the publication date. Each article belongs to one or more disciplines, and we rely on the classification scheme implemented by MAG to retrieve the field of studies associated to each paper. This scheme is a hierarchical classification that identifies 19 disciplines (first level) and 292 sub-disciplines (second level) at the first two levels of classification. The taxonomy uses state-of-the-art artificial intelligence methodologies to extract semantic content from documents, exploring natural language processing techniques and networks semantic reasoning to delineate disciplines (Sinha et al., 2015; Wang et al., 2019). There are several advantages of using this classification: it is based on concepts and language used at the paper-level, thus it avoids any bias that may arise from arbitrariness in the details of classifications that rely on human experts (Wang and Schneider, 2020);[12] it uses a heterogeneous network semantics analysis that exploits the context in which the publication's text is embedded, linking it to authors, affiliations, and locations (Wang et al., 2019); and it also mitigates the assignment errors that results from the loss of granularity when we adopt journal-based categorization. Moreover, journal-based taxonomies have difficulties in dealing with generalists journals like *Nature*, *Science*, and *PLoS ONE*. The number of publication by year is reported in Figure A1, while Table A2 summarizes the number of references and citations by fields of study.

To compute the interdisciplinarity measures defined in Section 2.1 we use the second hierarchical level of the fields of study classification provided by MAG (292 subfields). These categories are not exclusive, thus a paper can be assigned to multiple fields and subfields. Instead, we use the first level (19 categories) to control for differences in citational patterns across disciplines. The distribution of our articles over categories is described in Table 2.

To avoid biases due to the small number of papers in our sample and obtain a more reliable measure of similarity between disciplines (as a proxy of the easiness in combining different topics and techniques in a single research), we use an index of proximity among fields of studies computed over the universe of articles in MAG. This proximity measure is based on the Network Similarity Package, a series of processing functionalities for MAG that allow us to compare two fields of study and obtain a similarity score that represents how close these fields are based on the frequency they appear together in a same paper.[13] Based on this measure of similarity, we represent the networks of fields of studies, also known as knowledge space, in Figure 2. The graph connect disciplines whose co-occurrence is frequent in articles. Nodes represent fields of study at the second level of MAG classification, but, to ease the interpretation, their shapes and colors correspond to disciplines at the upper level of classification (conversion table is available in Appendix C). In the graph, sub-

[12]For example, the total number of categories of the two most frequently used system of classification, Web of Science (WoS) journal subject categories (SC) and the All Science Journal Classification (ASJC) from Scopus, varies drastically: there are 252 SCs and 330 ASJCs.

[13]For details on the Network Similarity package, see Microsoft Research (2020).

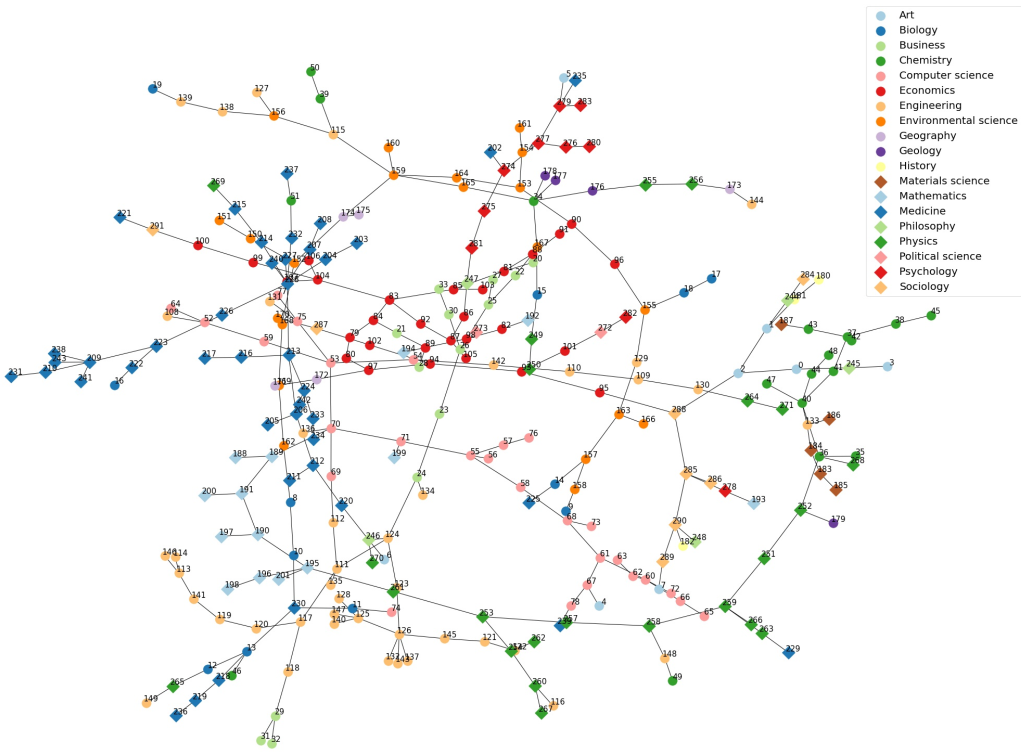

Figure 2: Knowledge space among fields of studies. The network shows the proximity between fields of study at the second level of the MAG classification (292 fields of studies). To ease the graph' interpretation, authors grouped fields of studies by discipline (the first level of MAG classification), which are represented by different colors and shapes, as reported in the plot legend. The conversion between the two levels as well as the field of studies corresponding to node IDs are reported in Table C1.

disciplines belonging to environmental science, medicine, and biology are on the left. At their right, we can observe the interconnection between economics and business. The bottom part of the network, instead, shows the connection between fields in mathematics (starting from the left), engineering, computer science, chemistry, physics, and material science. At the top of the figure, the interpenetration between art (included literature), psychology, sociology, history, and geography is evident.

Table 2: Distribution of focal papers by field of study

Field of Study	**Total**	**Avg. Nb. References**	**Avg. Nb. Citations**
Biology	7781	46.07	22.26
Medicine	6305	35.64	22.14
Chemistry	2628	41.60	19.86
Psychology	1703	46.51	17.31
Physics	1686	39.27	22.71
Mathematics	996	26.48	9.87
Materials science	785	34.34	21.68
Computer science	508	31.73	12.65
Geology	506	48.77	17.96
Economics	503	37.41	13.82
Engineering	404	27.78	13.71
Sociology	199	34.58	8.27
Environmental science	85	39.41	52.58
Geography	59	44.47	29.29
History	39	29.44	3.36
Political science	29	26.90	11.10
Business	26	57.19	24.00
Philosophy	20	31.75	2.15
Art	7	11.29	1.57

Notes: This table shows the distribution of focal articles per fields of study at the first level (19 categories). The average number of references relates to papers cited by our articles of interest and the average number of citations takes into account total number of citations within 5 years from the publication.

5 Results

5.1 Main results

In this section, we present the results we obtained from the estimation of equation (6) when considering both outcome variables.

In conducting our analysis, we are able to investigate two separated matters and test our hypotheses HP1 and HP2. First, we can observe whether different dimensions of interdisciplinarity have a similar impact on the same outcome variable (citations or extramural influence), or some of them are considered desirable and are rewarded by other researchers (e.g. with higher citations) while others are less desirable and thus penalized. If the latter case is verified, then we would find evidence for a trade-off across IDR dimensions on a specific outcome (HP1). Second, we can observe whether the same dimension of interdisciplinarity has the same impact when considering different outcome variables, or it is rewarded in some cases and penalized in other cases. If the latter scenario is detected, then evidence would be in favor of the presence of a trade-off between increasing citations and diffusing ideas across fields (HP2).

We begin our investigation by considering the number of citations as our dependent variable. Results are presented in Table 3. In column (1), we estimate the effects of the three dimensions of interdisciplinarity while accounting for monodisciplinarity, and individual, field of study, and year fixed effects. Our results are in favor of the existence of a trade-off across interdisciplinary dimensions, and we can confirm HP1 for what concerns the number of citations. In fact, only an increase in the Variety of the paper has a positive and statistically significant effect on the number of citations, whereas the other two measures (Balance and Disparity) have a negative and statistically significant effect. In column (2), we augment our model specification by including a control for the number of authors of the paper. The estimated coefficient of this variable indicates that increasing the number of authors has a positive and statistically significant impact on the number of citations accrued. This is consistent with the idea that the narrower expertise of researchers requires having larger teams to producing wide-spread research. Most importantly, all our previous findings on the existence of a trade-off across interdisciplinary dimensions are qualitatively unchanged. In column (3), we add to our model specification a dummy variable registering whether one of the co-authors of the paper is affiliated to an institution outside the U.S. We see that the presence of an international collaborator among the work team has a positive and statistically significant effect on the number of citations of the paper, hinting to the fact that working in an international team may expand the exposure of one's work. Again, all our results on the trade-off across interdisciplinary dimensions are left qualitatively unchanged. In column (4), we add a control for the H-index of the investigator, i.e. we estimate equation (6) with its entire set of controls. Perhaps unexpectedly, we find that having a higher H-index has a positive and statistically significant effect on the number of citations received. Most relevant to us, the results on the trade-off across interdisciplinary dimensions are still confirmed. In conclusion, then, we find that even when controlling for other relevant factors affecting the number of citations obtained by one's paper, interdisciplinarity has a large and significant impact on citations, and the direction of this effect depends on the dimension of interdisciplinarity considered.

We discuss the different effect of the considered interdisciplinarity dimensions by focusing on the estimates returned by our preferred model specification, i.e. that in column (4). We find that a 10% increase in the Variety adopted by a researcher when working on a research project increases her number of citations received within 5 years on a paper by 5.38%. This result is in line with Leahey et al. (2017), who finds the same positive effects on total number of citations. At the same time, we find the opposite effect for the other measures of interdisciplinarity: a 10% increase in the Balance decreases the citations accumulated by 35.20%, which supports the idea that an even distribution of the references among fields of study negatively impacts citations, i.e. researchers reward interdisciplinary articles when they build mainly on a specific field area, and more loosely refer to other fields. Finally, the same increase of 10% in the Disparity diminishes citations by 11.38%, suggesting that the integration of distant knowledge is negatively perceived by academic audiences and leads to penalization in terms of citations, in accordance with the previous literature (Yegros-Yegros et al., 2015). Overall, these results suggest that highly cited papers integrate knowledge from various, but not too distant, fields of studies while referring mainly to a specific discipline (and audience).

We now replicate our analysis by considering extramural influence as our dependent variable. Our results are presented in Table 4.[14] Across all model specifications, evidence is in favor of the existence of a trade-off across interdisciplinary dimensions, and we can confirm HP1 also for what concerns extramural influence. While an increase of Variety and Balance has a positive and statistically significant impact on extramural influence, the effect of Disparity is negative and statistically significant. Moreover, we observe that while a larger number of co-authors has a positive and statistically significant effect on the generality of the paper, the presence of international collaborators in the team, and the H-index of the researcher, have no statistically significant effect.

We again discuss the different effect of the considered interdisciplinarity dimensions by focusing on the estimates returned by our preferred model specification. Our results show that Variety has a modest but positive effect on extramural influence, with a 10% increase in the number of

[14] We report in this table the second stage of a two-step Heckman correction model to control for potential selection in our sample (i.e. the fact that some papers have zero citations). This exercise does not rely on the use of a specific exclusion restriction, and it only makes use of those variables included in the second stage of the model (i.e. our covariates). It is worth noting that, even when an exclusion restriction is not used, identification is formally achieved, though results may be less precise in terms of statistical significance. This should be not of any practical concern, however. Our aim is to test whether our results remain qualitatively unchanged even when controlling for the potential presence of selection issues. Reassuringly, the evidence produced by our exercise confirms all our model predictions. Results of the first stage are available in Table B2 of the Appendix B.

Table 3: Results OLS: Number of citations

	Dependent variable: log(Nb. of Citations+1)			
	(1)	(2)	(3)	(4)
log(Variety)	0.647^{***}	0.551^{***}	0.552^{***}	0.550^{***}
	(0.016)	(0.015)	(0.015)	(0.015)
log(Balance + 1)	-4.452^{***}	-4.578^{***}	-4.546^{***}	-4.552^{***}
	(0.201)	(0.191)	(0.191)	(0.191)
log(Disparity + 1)	-0.992^{***}	-1.295^{***}	-1.282^{***}	-1.268^{***}
	(0.246)	(0.229)	(0.229)	(0.229)
log(Number of Authors)		0.453^{***}	0.447^{***}	0.445^{***}
		(0.013)	(0.013)	(0.013)
International Collaboration			0.037^{**}	0.038^{**}
			(0.015)	(0.015)
log(H-index + 1)				0.127^{***}
				(0.020)
Variety = 1	YES	YES	YES	YES
Fields of Study Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES
Researcher Fixed Effects	YES	YES	YES	YES
Observations	46,159	46,159	46,159	46,159
R^2	0.442	0.479	0.479	0.480
Adjusted R^2	0.356	0.399	0.399	0.400

Note: S.E. clustered at researcher level *p<0.1; **p<0.05; ***p<0.01

unique fields of study in the paper's references leading to an increase of the Generality index by 0.36 percentage points. At the same time, Balance has a sizeable positive effect on extramural influence: a 10% increase in the Balance raise the generality index by 2.29 percentage points. Taken together, these two results suggest that some attributes of interdisciplinarity help the spread of ideas and concepts across multiple fields. On the contrary, however, we find that a 10% increase in the Disparity decreases the extramural influence by 1.90 percentage points. This result further supports the hypothesis that combinations of dissimilar knowledge are devalued not only in specific disciplines but also by the broader scientific community, given that this estimate implies that increasing the average distance of the disciplines referenced in each paper results in citations that originate from a more concentrated pool of fields of study.

By confronting the results from Table 3 and Table 4, we also obtain evidence from the presence of a trade-off between citation impact and extramural impact (HP2). In fact, the effect of the Balance

Table 4: Results OLS: Generality - second stage of the Heckman correction

	Dependent variable: log(Generality + 1)			
	(1)	(2)	(3)	(4)
log(Variety)	0.034***	0.038***	0.038***	0.038***
	(0.002)	(0.002)	(0.002)	(0.002)
log(Balance + 1)	0.287***	0.237***	0.238***	0.238***
	(0.024)	(0.025)	(0.025)	(0.025)
log(Disparity + 1)	−0.194***	−0.201***	−0.201***	−0.201***
	(0.038)	(0.037)	(0.037)	(0.037)
log(Number of Authors)		0.011***	0.011***	0.011***
		(0.001)	(0.001)	(0.001)
International Collaboration			0.001	0.001
			(0.001)	(0.001)
log(H-index + 1)				−0.001
				(0.002)
IMR	−0.117***	−0.069***	−0.069***	−0.069***
	(0.013)	(0.015)	(0.015)	(0.015)
Variety = 1	YES	YES	YES	YES
Fields of Study Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES
Researcher Fixed Effects	YES	YES	YES	YES
Observations	44,087	44,087	44,087	44,087
R^2	0.341	0.342	0.342	0.342
Adjusted R^2	0.238	0.239	0.239	0.239

Note: S.E. clustered at researcher level — *p<0.1; **p<0.05; ***p<0.01

over these two measures has opposite direction. This suggests that papers with lower Balance have more citations but reach a less diverse audience of academics. This qualitative difference of the effect of the Balance measure on the outcomes is consistent with our hypothesis that researchers face a trade-off between using some attributes of interdisciplinarity to accrue citations or to spread ideas and concepts across multiple fields. Therefore, we confirm our hypothesis HP2. A plausible explanation for this result is that papers that refer more evenly to the discipline pool have an ambiguous identity and thus are more difficult to be understood by an audience of specialists in a specific field; at the same time, they have a broader appeal because they bridge audiences that were previously separated, boosting the visibility of the work.

To summarize, our results show that interdisciplinarity has a statistically significant effect on citations and extramural influence. This finding is robust to all model specifications adopted. The direction of the effects of each dimension is different, though, hinting at the fact that researchers face a dilemma in how to approach IDR. In fact, despite the three dimensions are distinct, they are not completely independent. For instance, by increasing Variety (which has a positive effect on one's research impact) one will eventually increase Disparity (which has a negative effect instead). Moreover, we find evidence that one aspect of interdisciplinary (Balance) has strong but opposite effects on citations and extramural influence. This indicates that researchers face a trade-off between increasing their reputation and reaching out to other disciplines. The costs of IDR in terms of citations are important enough to negatively impact researchers academic careers, but the public benefits regarding the diffusion of knowledge are substantive and cannot be dismissed. This disconnection between private and public returns sets a challenge to research policy.

5.2 Heterogeneous effects

In this section, we explore whether the effects of interdisciplinarity may vary according to the characteristics of the investigators. To this purpose, we estimate equation (6) by considering only researchers with specific features.

Table 5 presents the estimates of the effects of IDR measures on citations using our main specification for different subgroups. We begin by testing, in columns (1) and (2), whether a gender difference exists in our estimates to assess if men and women face the same incentives to engage in IDR. Although the effects seems more pronounced for women, they are qualitatively the same for both citations and extramural influence. Thus, we conclude that man and women do not appear to face different costs and benefits to do interdisciplinary work. In the next step, we address the possibility that international collaboration may influence the heterogeneity of the team and the knowledge integration process, which may facilitate IDR. We report the results for the sub-samples corresponding to papers with international coauthors and those only with US-based researchers in columns (3) and (4). We did not find any qualitative difference on the effects on citations between the subgroups based on team diversity. Furthermore, we assess in specifications (5) and (6) whether our findings are driven by star researchers.[15] Once more the estimates are qualitatively similar. This evidence suggests that our baseline results are not driven by a small group of prolific researchers which may engage in high-risk, high-reward publication strategies are exposed to the same effects

[15] We define "superstars" as researchers in the upper 10th percentile of the h-index distribution within each year.

Table 5: Results OLS: Number of citations - gender, international collaboration, and superstar

	Dependent variable: log(Nb. of Citations+1)					
	Men	Women	Inter. Collab.	Only US	Superstar	Non-Superstar
	(1)	(2)	(3)	(4)	(5)	(6)
log(Variety)	0.548***	0.563***	0.650***	0.532***	0.594***	0.544***
	(0.017)	(0.031)	(0.041)	(0.016)	(0.044)	(0.015)
log(Balance + 1)	−4.540***	−4.630***	−4.085***	−4.643***	−5.328***	−4.446***
	(0.217)	(0.402)	(0.400)	(0.217)	(0.634)	(0.197)
log(Disparity + 1)	−1.101***	−2.013***	−2.914***	−0.965***	−1.263*	−1.255***
	(0.252)	(0.561)	(0.593)	(0.254)	(0.761)	(0.237)
log(Number of Authors)	0.444***	0.449***	0.568***	0.413***	0.527***	0.430***
	(0.014)	(0.026)	(0.033)	(0.014)	(0.042)	(0.012)
International Collaboration	0.035**	0.039			0.049	0.036**
	(0.017)	(0.028)			(0.043)	(0.016)
log(H-index + 1)	0.119***	0.136***	0.178***	0.101***	−0.171	0.153***
	(0.024)	(0.034)	(0.059)	(0.021)	(0.161)	(0.020)
Variety = 1	YES	YES	YES	YES	YES	YES
Fields of Study Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Researcher Fixed Effects	YES	YES	YES	YES	YES	YES
Observations	34,616	11,071	8,763	37,396	5,137	41,022
R^2	0.469	0.519	0.598	0.486	0.366	0.490
Adjusted R^2	0.400	0.404	0.425	0.391	0.346	0.401

Note: S.E. clustered at researcher level — *p<0.1; **p<0.05; ***p<0.01

of IDR than other researchers.

We further proceed with our analysis by estimating equation (6) when considering a sample of researchers affiliated to a specific academic unit division in which the researcher's college and department are included. Results are reported in Table 6. In the estimations presented in this table, we alternatively consider those affiliated to: the College of Liberal Arts and Science (CLAS), column (1); the College of Engineering (ENG), column (2); the Health Science Center (HSC), column (3); and the Institute of Food and Agricultural Sciences (IFAS).[16] Our results are qualitatively unchanged regardless of the affiliation considered, hinting to the fact that IDR has the same effect on the number of citations received across all academic environments.[17]

We replicate our exercise when considering extramural impact as our dependent variable. Our results are reported in Table 7 and Table 8. Also in this case, we find no evidence that IDR has

[16]The colleges included in each academic unit are reported in Table A1 in the Appendix A.2.

[17]We also estimate equation (6) using alternative disciplinary subdivisions based on researchers' paper fields of study, individual main field of publication (measured as the field where the researcher published most of her papers), and also using department-level affiliation. All our results are qualitatively unchanged. Results are available upon request.

Table 6: Results OLS: Number of citations - disciplines at the academic unit level

	Dependent variable: log(Nb. of Citations+1)			
	CLAS	ENG	HSC	IFAS
	(1)	(2)	(3)	(4)
log(Variety)	0.489***	0.493***	0.567***	0.555***
	(0.048)	(0.046)	(0.024)	(0.030)
log(Balance + 1)	−5.028***	−3.781***	−4.924***	−3.631***
	(0.573)	(0.771)	(0.283)	(0.374)
log(Disparity + 1)	−1.975***	−1.456**	−1.089***	−1.457**
	(0.637)	(0.564)	(0.338)	(0.658)
log(Number of Authors)	0.449***	0.317***	0.506***	0.343***
	(0.031)	(0.058)	(0.018)	(0.025)
International Collaboration	−0.048	0.023	0.071***	0.092***
	(0.040)	(0.041)	(0.026)	(0.030)
log(H-index + 1)	0.066	0.084	0.163***	0.040
	(0.070)	(0.073)	(0.030)	(0.035)
Variety = 1	YES	YES	YES	YES
Fields of Study Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES
Researcher Fixed Effects	YES	YES	YES	YES
Observations	5,267	4,946	18,263	9,825
R^2	0.503	0.360	0.474	0.478
Adjusted R^2	0.428	0.301	0.401	0.403

Note: S.E. clustered at researcher level — *p<0.1; **p<0.05; ***p<0.01

a heterogeneous impact across researchers with specific characteristics. At the same time, when considering researchers' affiliation, we observe that the effect of the Disparity is negative but no longer significant for the researchers in the College of Liberal Arts and Science (column (1)). This may suggest that researchers in social sciences, humanities, and hard sciences like physics are not penalized as much for combining dissimilar disciplines as more "applied" fields like engineering, health, and agricultural science. In all the other cases, however, we still find evidence that Variety and Balance have a positive and statistically significant impact on extramural impact for all considered subgroups, while Disparity has a negative and statistically significant impact. Therefore, our evidence still confirms our hypotheses and points to a trade-off across interdisciplinary dimensions with respect to extramural impact (HP1), and to a trade-off of the Balance across different outcome variables (HP2).

Taken together, our results suggest that all scholars face the similar incentives and constraints

Table 7: Results OLS: Generality - gender, international collaboration, and superstar

	Dependent variable: log(Generality+1)					
	Men	Women	Inter. Collab.	Only US	Superstar	Non-Superstar
	(1)	(2)	(3)	(4)	(5)	(6)
log(Variety)	0.037^{***}	0.045^{***}	0.028^{***}	0.041^{***}	0.038^{***}	0.040^{***}
	(0.003)	(0.005)	(0.005)	(0.003)	(0.005)	(0.003)
log(Balance + 1)	0.244^{***}	0.222^{***}	0.309^{***}	0.217^{***}	0.260^{***}	0.228^{***}
	(0.029)	(0.047)	(0.042)	(0.029)	(0.080)	(0.026)
log(Disparity + 1)	-0.149^{***}	-0.432^{***}	-0.147	-0.195^{***}	-0.280^{**}	-0.198^{***}
	(0.042)	(0.078)	(0.092)	(0.042)	(0.109)	(0.040)
log(Number of Authors)	0.011^{***}	0.012^{***}	0.008^{***}	0.013^{***}	0.013^{***}	0.011^{***}
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.001)
International Collaboration	0.0003	0.003			0.001	0.001
	(0.002)	(0.003)			(0.003)	(0.002)
log(H-index + 1)	-0.004	0.004	-0.010^{**}	-0.001	-0.021	-0.001
	(0.003)	(0.004)	(0.005)	(0.003)	(0.014)	(0.002)
IMR	-0.073^{***}	-0.064^{*}	-0.165^{***}	-0.048^{***}	-0.003	-0.067^{***}
	(0.016)	(0.034)	(0.048)	(0.017)	(0.041)	(0.016)
Variety = 1	YES	YES	YES	YES	YES	YES
Fields of Study Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Researcher Fixed Effects	YES	YES	YES	YES	YES	YES
Observations	33,009	10,619	8,563	35,524	4,974	39,113
R^2	0.326	0.397	0.558	0.347	0.206	0.355
Adjusted R^2	0.236	0.252	0.366	0.223	0.180	0.240

Note: S.E. clustered at researcher level — *p<0.1; **p<0.05; ***p<0.01

in engaging in more interdisciplinary projects. Regardless of their characteristics or affiliation, the effects of IDR are large and widespread, and affect in the same way all research activities at the University of Florida.

6 Conclusion

IDR is becoming more and more important to organize team work and produce wide-ranging scientific advances. However, despite the push from funding and research institutions for overcoming disciplinary barriers and promote IDR, academic scholarship and evaluation mechanisms are still organized in separated fields. As a result, researchers receive contrasting incentives in conducting their work using an interdisciplinary approach. In order to better understand the effect of IDR on researchers' career, we study the publication record of researchers from a large university in the United States to assess the effect of interdisciplinarity on the reputation of a scientist, measured in

Table 8: Results OLS: Generality - disciplines at the academic unit level

	Dependent variable: log(Generality+1)			
	CLAS	ENG	HSC	IFAS
	(1)	(2)	(3)	(4)
log(Variety)	0.026^{***}	0.037^{***}	0.037^{***}	0.051^{***}
	(0.005)	(0.008)	(0.004)	(0.006)
log(Balance + 1)	0.099^{**}	0.275^{***}	0.269^{***}	0.214^{***}
	(0.048)	(0.100)	(0.040)	(0.058)
log(Disparity + 1)	-0.101	-0.187^{*}	-0.139^{**}	-0.463^{***}
	(0.070)	(0.100)	(0.063)	(0.105)
log(Number of Authors)	0.014^{***}	0.014^{**}	0.010^{***}	0.006^{*}
	(0.003)	(0.006)	(0.002)	(0.003)
International Collaboration	-0.009^{**}	0.001	0.005^{***}	0.005
	(0.004)	(0.005)	(0.002)	(0.003)
log(H-index + 1)	-0.007	0.004	0.005	-0.015^{***}
	(0.007)	(0.008)	(0.003)	(0.005)
IMR	-0.038	-0.020	-0.088^{***}	-0.065^{*}
	(0.031)	(0.055)	(0.022)	(0.039)
Variety = 1	YES	YES	YES	YES
Fields of Study Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES
Researcher Fixed Effects	YES	YES	YES	YES
Observations	5,024	4,741	17,392	9,368
R^2	0.386	0.243	0.285	0.335
Adjusted R^2	0.294	0.171	0.183	0.237

Note: S.E. clustered at researcher level *p<0.1; **p<0.05; ***p<0.01

terms of citations accumulated and extramural influence.

We present evidence that the effects of different dimensions of interdisciplinarity conflict with one another in building the reputation of a scientist in her field, and across different fields. This is not the only trade-off faced by researchers, however. In fact, we also find that the adoption of an interdisciplinary approach may have different effects when considering different forms of reputation. Importantly, all our results are robust to a number of controls, including individual, time, field of study fixed effects, and they apply to all investigators regardless of their gender, collaboration behavior, performance, and affiliation.

Our main takeaway is that interdisciplinarity has a sizable and heterogeneous impact on scholarly performance, and this gives rise to trade-offs that must be faced by researchers who want to increase the interdisciplinarity of their works. These results are consistent with previous findings in the literature and have potentially important implications for the academic profession. They support the idea that the price paid by a researcher to increase interdisciplinarity is substantial, and that scientific evaluation based on citations metrics may influence the scientific process and hinder knowledge diffusion across fields.

Acknowledgements

This research uses data from the Bureau of Economics and Business Research (BEBR) of the University of Florida. We thank the staff of the BEBR and the director Christopher McCarty for their important contribution to this work. We are also grateful for helpful comments from seminar participants at the Regional Quantitative Analysis Group (AQR) seminar series of the University of Barcelona.

References

Anderson, K. A. and S. Richards-Shubik (2021). "Collaborative Production in Science: An Empirical Analysis of Coauthorships in Economics." *The Review of Economics and Statistics*, 1–45. 10.1162/rest_a_01025.

Angrist, J., P. Azoulay, G. Ellison, R. Hill, and S. F. Lu (2020). "Inside job or deep impact? Extramural citations and the influence of economic scholarship." *Journal of Economic Literature* 58(1), 3–52. 10.1257/jel.20181508.

Arnold, A. et al. (2021). "Perspective: Promoting and fostering multidisciplinary research in universities." *Research Policy* 50(9), 104334. 10.1016/j.respol.2021.104334.

Battiston, F. et al. (2019). "Taking census of physics." *Nature Reviews Physics* 1(1), 89–97. 10.1038/s42254-018-0005-3.

Biancani, S., L. Dahlander, D. A. McFarland, and S. Smith (2018). "Superstars in the making? The broad effects of interdisciplinary centers." *Research Policy* 47(3), 543–557. 10.1016/j.respol.2018.01.014.

Bresnahan, T. F. and M. Trajtenberg (1995). "General purpose technologies 'Engines of growth'?" *Journal of Econometrics* 65(1), 83–108. 10.1016/0304-4076(94)01598-T.

Bromham, L., R. Dinnage, and X. Hua (2016). “Interdisciplinary research has consistently lower funding success.” *Nature* 534, 684–687. 10.1038/nature18315.

Carley, S. and A. L. Porter (2012). “A forward diversity index.” *Scientometrics* 90(2), 407–427. 10.1007/s11192-011-0528-1.

Cedrini, M. and M. Fontana (2018). “Just another niche in the wall? How specialization is changing the face of mainstream economics.” *Cambridge Journal of Economics* 42(2), 427–451. 10.1093/cje/bex003.

Conti, A. and C. C. Liu (2015). “Bringing the lab back in: Personnel composition and scientific output at the MIT department of biology.” *Research Policy* 44(9), 1633–1644. 10.1016/j.respol.2015.01.001.

Crescenzi, R., M. Nathan, and A. Rodríguez-Pose (2016). “Do inventors talk to strangers? On proximity and collaborative knowledge creation.” *Research Policy* 45(1), 177–194.

Ferrucci, E. and F. Lissoni (2019). “Foreign inventors in Europe and the United States: diversity and patent quality.” *Research Policy* 48(9), 103774.

Fontana, M., M. Iori, F. Montobbio, and R. Sinatra (2020). “New and atypical combinations: An assessment of novelty and interdisciplinarity.” *Research Policy* 49(7), 104063. 10.1016/j.respol.2020.104063.

Geuna, A. and B. R. Martin (2003). “University research evaluation and funding: An international comparison.” *Minerva* 41(4), 277–304.

Goel, R. K. and J. R. Faria (2007). “Proliferation of academic journals: Effect on research quantity and quality.” *Metroeconomica* 58(4), 536–549. 10.1111/j.1467-999X.2007.00285.x.

Hackett, E. J. et al. (2021). “Do synthesis centers synthesize? A semantic analysis of topical diversity in research.” *Research policy* 50(1), 104069. 10.1016/j.respol.2020.104069.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001). “The NBER patent citation data file: Lessons, insights and methodological tools.” *NBER Working Paper* 8498. 10.3386/w8498.

Hamermesh, D. S. and G. A. Pfann (2012). “Reputation and earnings: The roles of quality and quantity in academe.” *Economic Inquiry* 50(1), 1–16. 10.1111/j.1465-7295.2011.00381.x.

Jones, B., E. J. Reedy, and B. A. Weinberg (2014). “Age and scientific genius.” *NBER Working Paper* w19866. 10.3386/w19866.

Jones, B. F. (2010). “Age and great invention.” *The Review of Economics and Statistics* 92(1), 1–14. 10.1162/rest.2009.11724.

— (2009). “The burden of knowledge and the “Death of the Renaissance Man”: Is innovation getting harder?” *The Review of Economic Studies* 76(1), 283–317. 10.1111/j.1467-937X.2008.00531.x.

Jones, B. F. (2021). “The rise of research teams: Benefits and costs in economics.” *Journal of Economic Perspectives* 35(2), 191–216. 10.1257/jep.35.2.191.

Larsen, P. O. and M. von Ins (2010). “The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index.” *Scientometrics* 84(3), 575–603. 10.1007/s11192-010-0202-z.

Leahey, E. and S. N. Barringer (2020). “Universities’ commitment to interdisciplinary research: To what end?” *Research Policy* 49(2), 103910. https://doi.org/10.1016/j.respol.2019.103910.

Leahey, E., C. M. Beckman, and T. L. Stanko (2017). “Prominent but less productive: The impact of interdisciplinarity on scientists’ research.” *Administrative Science Quarterly* 62(1), 105–139. 10.1177/0001839216665364.

Microsoft Research (2020). Multi-Sense network representation learning in Microsoft Academic Graph. Microsoft Research. www.microsoft.com/en-us/research/project/academic/articles/multi-sense-network-representation-learning-in-microsoft-academic-graph/ (visited on 10/10/2021).

Østergaard, C. R., B. Timmermans, and K. Kristinsson (2011). “Does a different view create something new? The effect of employee diversity on innovation.” *Research policy* 40(3), 500–509.

Porter, A. and I. Rafols (2009). “Is science becoming more interdisciplinary? Measuring and mapping six research fields over time.” *Scientometrics* 81(3), 719–745. 10.1007/s11192-008-2197-2.

Ritzberger, K. (2008). “A ranking of journals in economics and related fields.” *German Economic Review* 9(4), 402–430. 10.1111/j.1468-0475.2008.00447.x.

Rylance, R. (2015). “Grant giving: Global funders to focus on interdisciplinarity.” *Nature News* 525(7569), 313. 10.1038/525313a.

Sinha, A. et al. (2015). “An overview of Microsoft Academic Service (MAS) and applications.” In: *Proceedings of the 24th International Conference on World Wide Web*, 243–246. 10.1145/2740908.2742839.

Squicciarini, M., H. Dernis, and C. Criscuolo (2013). “Measuring patent quality: Indicators of technological and economic value.” *OECD Science, Technology and Industry Working Papers* 2013/03. 10.1787/18151965.

Stigler, G. J., S. M. Stigler, and C. Friedland (1995). “The journals of economics.” *Journal of Political Economy* 103(2), 331–359. 10.1086/261986.

Stirling, A. (2007). “A general framework for analysing diversity in science, technology and society.” *Journal of The Royal Society Interface*. 10.1098/rsif.2007.0213.

The National Academies (1998). *Trends in the Early Careers of Life Scientists*. Washington, DC: National Academies Press. 10.17226/6244.

— (2005). *Facilitating interdisciplinary research*. Washington, DC: National Academies Press. 10.17226/11153.

Trajtenberg, M., R. Henderson, and A. Jaffe (1997). "University Versus Corporate Patents: A Window On The Basicness Of Invention." *Economics of Innovation and New Technology* 5(1), 19–50. 10.1080/10438599700000006.

Wagner, C. S. et al. (2011). "Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature." *Journal of informetrics* 5(1), 14–26. 10.1016/j.joi.2010.06.004.

Wang, K. et al. (2019). "A review of Microsoft Academic Services for science of science studies." *Frontiers in Big Data* 2. 10.3389/fdata.2019.00045.

Wang, Q. and J. W. Schneider (2020). "Consistency and validity of interdisciplinarity measures." *Quantitative Science Studies* 1(1), 239–263. 10.1162/qss_a_00011.

Wu, L., D. Wang, and J. A. Evans (2019). "Large teams develop and small teams disrupt science and technology." *Nature* 566(7744), 378–382.

Yegros-Yegros, A., I. Rafols, and P. D'Este (2015). "Does interdisciplinary research lead to higher citation impact? The different effect of proximal and distal interdisciplinarity." *PLoS ONE* 10(8), e0135095. 10.1371/journal.pone.0135095.

Zeng, A. et al. (2017). "The science of science: From the perspective of complex systems." *Physics Reports* 714-715, 1–73. 10.1016/j.physrep.2017.10.001.

A Data appendix

In this appendix, we report the procedure we followed to construct the database used in the analysis and some additional descriptive statistics.

A.1 Data collection

From the data collected by the Bureau of Economics and Business Research (BEBR) of the University of Florida (UF), we retrieved information concerning publication records, department affiliation, and gender for the universe of UF's researchers in the period 2008-2013. Each researcher is identified by a unique code (UFID). Raw publication records provide information regarding 34,851 scholarly works including journal name, article title, and publication year. Based on UF registered publications' information, we retrieved the publications' Digital Object Identifiers (DOI) from Crossref and Scopus databases.[18] This procedure allows us to identify researchers' academic output that was indexed in the largest and most common scholarly works' databases.

More specifically, we used an automated script to extract bibliographic metadata of UF publications available in the original dataset through Scopus Database API Interface and Crossref REST API.[19] The three main steps of this procedure are the following:

1. Get articles partial metadata based on publication title: From titles of publications in the UF records, the script – through queries to Scopus and Crossref APIs – collects publications matching our list of articles' titles and retrieves their metadata (DOI, journal name, publication title, publication year). We collect the first ten results of the queries for each title and store them in a new database.

2. Cleaning and processing article's title: The article titles in the raw data and in the data retrieved by API queries are cleaned and then processed. Cleaning consists in eliminating spaces, special characters, and punctuation. Processing consists in coercing characters to lowercase and comparing the raw (original) and newly extracted titles.

3. Title-DOI matching procedure: Matches are determined according to a fuzzy matching algo-

[18] The databases are available at the following webpage: Crossref and Scopus.
[19] Data collection using Scopus and Crossref occurred in 2018.

rithm implemented in the *fuzzywuzzy* text similarity package in Python.[20] The script considers a match if titles have a higher than 90% similarity ratio and the matching is unique. Matched publications and its respective metadata are assigned to the associated researcher. Unique matches with more than one DOI were manually checked and disambiguated. Publications without a unique match are dropped.

With this procedure, we were able to identify the DOIs of 28,239 publications of our original database. Using these DOIs, we collect the full metadata through Lens and Microsoft Academic Graph (MAG) databases.[21] Metadata from Lens API platform includes: IDs (Lens articles ID, Microsoft Academic Graph ID); publication type (journal article, book, working paper); list of citations; list of references; fields of study (computed by the MAG algorithm as described in Section 4); and authors' affiliations. We decided to focus on Lens database to collect citations and references data because it also provides their disciplines based on the natural language processing algorithms used by MAG. Furthermore, the Lens' scholarly citation data, contrary to Microsoft Academic Graph, indexes only publications of selected document types (journal article, book, working paper).[22] Publications missing references or missing fields of study are dropped. In addition, we restrict our sample to only journal articles. Our final database consists of 23,926 articles and their full metadata.

In the last data collection phase, we extracted from MAG a proximity measure between the fields of study using the functionality Network Similarity Package, as described in Section 4.[23] We collected similarity scores for all possible combinations between the 19 fields (first level of classification) and 292 subfields (second level of classification).

A.2 Additional descriptive statistics

Figure A1 shows the evolution of the total number of publications in our database (in the period 2008–2013). Table A1 reports the distribution of researchers across academic units and colleges, while A2 shows the distribution of citations by field of study. Table A3 reports the correlation between variables used in our regression analysis.

[20] Documentation about *fuzzywuzzy* is available here: fuzzywuzzy.
[21] These databases are available at the following link: Microsoft Academic Graph and Lens.
[22] Data collection using the Lens occurred in 2019.
[23] Data collection Microsoft Academic Graph occurred in 2020.

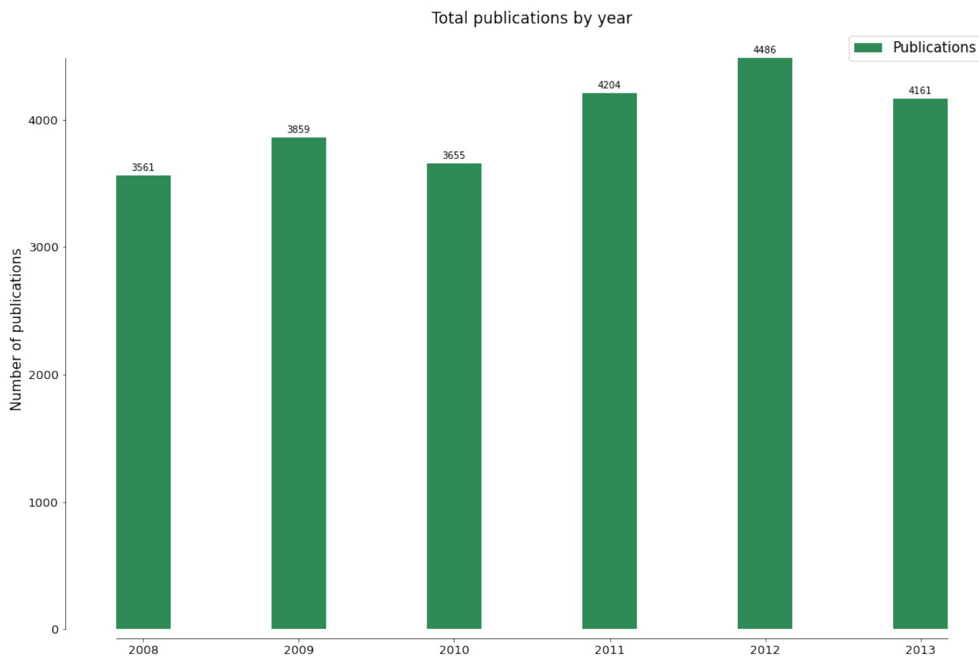

Figure A1: Total publications by year

Table A1: Number of researchers by academic units and colleges

Academic Units	**Colleges**	**Researchers**
Liberal Arts and Sciences	College of Liberal Arts and Sciences	665
Engineering	College of Engineering	389
Health Sciences	Medicine	1545
	Medicine-Jacksonville	175
	Public Health and Health Professions	166
	Pharmacy	140
	Dentistry	139
	Nursing	36
	Health Affairs	14
Food and Agricultural Sciences	Agricultural And Life Sciences	978
	Veterinary Medicine	218
	Institute of Food and Agricultural Sciences	2
Other	UF Students	946
	Uncategorized Departments	687

Notes: This table shows the distribution researchers affiliated to the academic units and colleges at the University of Florida (UF) to 2008 to 2013. The total number of researchers with a college affiliation is 5130. Researchers that are not affiliated to any specific academic unit are counted in the category “Other”. Researchers classified as students in the UF registry office are counted in “UF Students” and faculty affiliated to departments not belonging to any college are counted in “Uncategorized Departments”.

Table A2: Distribution of citations by field of study

Field of Study	**References**	**Citations**
Art	5490	1393
Biology	1181592	359734
Business	21831	6804
Chemistry	329228	101792
Computer science	88408	23706
Economics	64521	15417
Engineering	79673	28339
Environmental science	40883	14674
Geography	27317	6791
Geology	89619	25773
History	8581	1026
Materials science	54980	29450
Mathematics	101585	19532
Medicine	1071812	378978
Philosophy	10873	2082
Physics	225173	78280
Political science	10374	2585
Psychology	233354	62381
Sociology	38379	7401

Notes: This table shows the distribution of the documents of each field in the focal papers' references and which cited our focal paper (citations). The total number of documents referenced is 646,280 and the total number of citations is 366,024

Table A3: Correlation table between variables used in regressions

	Nb. Citations	Generality	Variety	Balance	Disparity	Nb. References	Nb. of Authors	Inter. Collab.
Nb. Citations	1.00	0.18	0.18	-0.12	0.09	0.33	0.45	0.11
Generality	0.18	1.00	0.23	0.12	0.12	0.16	0.06	0.03
Variety	0.18	0.23	1.00	-0.05	0.57	0.67	0.07	0.04
Balance	-0.12	0.12	-0.05	1.00	0.10	-0.41	-0.08	-0.11
Disparity	0.09	0.12	0.57	0.10	1.00	0.28	0.06	0.02
Nb. References	0.33	0.16	0.67	-0.41	0.28	1.00	0.20	0.09
Nb. of Authors	0.45	0.06	0.07	-0.08	0.06	0.20	1.00	0.16
Inter. Collab.	0.11	0.03	0.04	-0.11	0.02	0.09	0.16	1.00

B Robustness checks and first stage of Heckman correction

Table B1 allows comparing the estimations obtained through the use of OLS with those resulting from Poisson and negative binomial. As evident in the table, our results are robust to the different estimation approaches.

Table B1: Results using OLS, Poisson, and negative binomial to estimate the effect of IDR con the number of citations

Dependent Variables:	log(Nb. of Citations+1)	Nb. of Citations	Nb. of Citations
	(1)	(2)	(3)
	OLS	*Poisson*	*Neg. Bin.*
log(Variety)	0.5499***	0.6920***	0.6211***
	(0.0146)	(0.0350)	(0.0226)
log(Balance + 1)	-4.552***	-4.410***	-4.526***
	(0.1910)	(0.4070)	(0.2432)
log(Disparity + 1)	-1.268***	-2.296***	-1.518***
	(0.2287)	(0.5492)	(0.3545)
log(Number of Authors)	0.4454***	0.5343***	0.4793***
	(0.0126)	(0.0351)	(0.0192)
International Collaboration	0.0376**	0.1133***	0.0615***
	(0.0149)	(0.0360)	(0.0227)
log(H-index + 1)	0.1266***	0.2715***	0.1506***
	(0.0196)	(0.0328)	(0.0233)
Variety = 1	YES	YES	YES
Fields of Study Fixed Effects	YES	YES	YES
Year Fixed Effects	YES	YES	YES
Researcher Fixed Effects	YES	YES	YES
Observations	46,159	45,974	45,974
Squared Correlation	0.47950	0.39767	0.31271
Pseudo R^2	0.21431	0.44458	0.09057
BIC	176,539.3	954,572.6	402,511.9
Over-dispersion			1.6172

Note: S.E. clustered at researcher level — *p<0.1; **p<0.05; ***p<0.01

Table B2, instead, reports the first stage of Heckman correction used to estimate the effect of IDR on Generality. Generality is, indeed, only defined for articles that receive at least one citation.

Table B2: First stage of the Heckman correction

	Dependent variable: Cited Paper			
	probit			
	(1)	(2)	(3)	(4)
log(Variety)	0.762^{***}	0.653^{***}	0.653^{***}	0.645^{***}
	(0.023)	(0.023)	(0.023)	(0.023)
log(Balance + 1)	-6.193^{***}	-6.054^{***}	-6.023^{***}	-5.916^{***}
	(0.327)	(0.331)	(0.332)	(0.332)
log(Disparity + 1)	-1.479^{***}	-1.797^{***}	-1.782^{***}	-1.769^{***}
	(0.315)	(0.324)	(0.324)	(0.324)
log(Number of Authors)		0.496^{***}	0.491^{***}	0.479^{***}
		(0.021)	(0.021)	(0.021)
International Collaboration			0.055	0.040
			(0.037)	(0.037)
log(H-index + 1)				0.115^{***}
				(0.016)
Constant	3.836^{***}	3.571^{***}	3.545^{***}	3.424^{***}
	(0.238)	(0.243)	(0.243)	(0.243)
Variety = 1	YES	YES	YES	YES
Fields of Study Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES
Researcher Fixed Effects	NO	NO	NO	NO
Observations	46,176	46,159	46,159	46,159
Log Likelihood	-6,555.701	-6,247.986	-6,246.882	-6,219.928
Akaike Inf. Crit.	13,195.400	12,581.970	12,581.760	12,529.850

Note: S.E. clustered at researcher level *p<0.1; **p<0.05; ***p<0.01

C Fields of study classification

Table C1 reports the conversion table, made by authors, between the first and the second level of fields of studies, as classified by MAG. The first level classify articles in 19 disciplines, while the second one has 292 possible values, corresponding to sub-disciplines. The table also include the ID used to represent fields of studies at the second level in the knowledge space (Figure 2).

ID	2^{nd} level	1^{st} level
0	Visual arts	Art
1	Classics	Art
2	Art history	Art
3	Literature	Art
4	Linguistics	Art
5	Communication	Art
6	Library science	Art
7	Humanities	Art
8	Zoology	Biology
9	Botany	Biology
10	Evolutionary biology	Biology
11	Computational biology	Biology
12	Cell biology	Biology
13	Molecular biology	Biology
14	Animal science	Biology
15	Astrobiology	Biology
16	Microbiology	Biology
17	Food science	Biology
18	Biotechnology	Biology
19	Biological system	Biology
20	Economic system	Business
21	Financial system	Business
22	Commerce	Business
23	Knowledge management	Business
24	Process management	Business
25	Marketing	Business
26	Public relations	Business
27	Advertising	Business
28	Accounting	Business
29	Operations research	Business
30	Management	Business
31	Operations management	Business
32	Management science	Business
33	Business administration	Business
34	Geochemistry	Chemistry
35	Computational chemistry	Chemistry
36	Physical chemistry	Chemistry
37	Organic chemistry	Chemistry
38	Stereochemistry	Chemistry
39	Environmental chemistry	Chemistry
40	Inorganic chemistry	Chemistry
41	Photochemistry	Chemistry
42	Combinatorial chemistry	Chemistry
43	Polymer chemistry	Chemistry
44	Analytical chemistry	Chemistry

ID	2^{nd} level	1^{st} level
45	Medicinal chemistry	Chemistry
46	Biochemistry	Chemistry
47	Nuclear chemistry	Chemistry
48	Chromatography	Chemistry
49	Radiochemistry	Chemistry
50	Toxicology	Chemistry
51	Pharmacology	Chemistry
52	Embedded system	Computer science
53	Distributed computing	Computer science
54	Computer network	Computer science
55	Artificial intelligence	Computer science
56	Pattern recognition	Computer science
57	Computer vision	Computer science
58	Machine learning	Computer science
59	Real-time computing	Computer science
60	World Wide Web	Computer science
61	Information retrieval	Computer science
62	Internet privacy	Computer science
63	Computer security	Computer science
64	Operating system	Computer science
65	Human–computer interaction	Computer science
66	Multimedia	Computer science
67	Natural language processing	Computer science
68	Data mining	Computer science
69	Programming language	Computer science
70	Theoretical computer science	Computer science
71	Algorithm	Computer science
72	Data science	Computer science
73	Database	Computer science
74	Bioinformatics	Computer science
75	Parallel computing	Computer science
76	Computer graphics (images)	Computer science
77	Computational science	Computer science
78	Speech recognition	Computer science
79	International economics	Economics
80	International trade	Economics
81	Market economy	Economics
82	Econometrics	Economics
83	Macroeconomics	Economics
84	Monetary economics	Economics
85	Economic policy	Economics
86	Positive economics	Economics
87	Neoclassical economics	Economics
88	Industrial organization	Economics
89	Finance	Economics
90	Natural resource economics	Economics
91	Environmental economics	Economics
92	Keynesian economics	Economics
93	Political economy	Economics
94	Development economics	Economics
95	Economic history	Economics
96	Agricultural economics	Economics
97	Economy	Economics
98	Financial economics	Economics
99	Labour economics	Economics
100	Demographic economics	Economics
101	Law and economics	Economics

ID	2^{nd} level	1^{st} level
102	Economic growth	Economics
103	Public economics	Economics
104	Microeconomics	Economics
105	Classical economics	Economics
106	Mathematical economics	Economics
107	Welfare economics	Economics
108	Computer hardware	Engineering
109	Electronic engineering	Engineering
110	Electrical engineering	Engineering
111	Systems engineering	Engineering
112	Software engineering	Engineering
113	Control engineering	Engineering
114	Control theory	Engineering
115	Environmental engineering	Engineering
116	Mechanics	Engineering
117	Manufacturing engineering	Engineering
118	Industrial engineering	Engineering
119	Mechanical engineering	Engineering
120	Engineering drawing	Engineering
121	Aerospace engineering	Engineering
122	Aeronautics	Engineering
123	Construction engineering	Engineering
124	Engineering management	Engineering
125	Geotechnical engineering	Engineering
126	Civil engineering	Engineering
127	Pulp and paper industry	Engineering
128	Structural engineering	Engineering
129	Agricultural engineering	Engineering
130	Optoelectronics	Engineering
131	Computer architecture	Engineering
132	Architectural engineering	Engineering
133	Chemical engineering	Engineering
134	Risk analysis (engineering)	Engineering
135	Reliability engineering	Engineering
136	Computer engineering	Engineering
137	Transport engineering	Engineering
138	Process engineering	Engineering
139	Biochemical engineering	Engineering
140	Petroleum engineering	Engineering
141	Automotive engineering	Engineering
142	Telecommunications	Engineering
143	Forensic engineering	Engineering
144	Remote sensing	Engineering
145	Marine engineering	Engineering
146	Simulation	Engineering
147	Mining engineering	Engineering
148	Nuclear engineering	Engineering
149	Biomedical engineering	Engineering
150	Atmospheric sciences	Environmental science
151	Meteorology	Environmental science
152	Climatology	Environmental science
153	Environmental resource management	Environmental science
154	Environmental planning	Environmental science
155	Agricultural science	Environmental science
156	Waste management	Environmental science
157	Agronomy	Environmental science
158	Horticulture	Environmental science

ID	2^{nd} level	1^{st} level
159	Hydrology	Environmental science
160	Soil science	Environmental science
161	Environmental protection	Environmental science
162	Ecology	Environmental science
163	Agroforestry	Environmental science
164	Water resource management	Environmental science
165	Geomorphology	Environmental science
166	Forestry	Environmental science
167	Earth science	Environmental science
168	Oceanography	Environmental science
169	Fishery	Environmental science
170	Environmental health	Environmental science
171	Regional science	Geography
172	Economic geography	Geography
173	Geodesy	Geography
174	Physical geography	Geography
175	Cartography	Geography
176	Petrology	Geology
177	Mineralogy	Geology
178	Paleontology	Geology
179	Crystallography	Geology
180	Archaeology	History
181	Ancient history	History
182	Genealogy	History
183	Metallurgy	Materials science
184	Composite material	Materials science
185	Ceramic materials	Materials science
186	Nanotechnology	Materials science
187	Polymer science	Materials science
188	Combinatorics	Mathematics
189	Discrete mathematics	Mathematics
190	Pure mathematics	Mathematics
191	Algebra	Mathematics
192	Statistics	Mathematics
193	Mathematics education	Mathematics
194	Actuarial science	Mathematics
195	Mathematical analysis	Mathematics
196	Applied mathematics	Mathematics
197	Topology	Mathematics
198	Calculus	Mathematics
199	Mathematical optimization	Mathematics
200	Arithmetic	Mathematics
201	Geometry	Mathematics
202	Psychiatry	Medicine
203	Orthodontics	Medicine
204	Dentistry	Medicine
205	Medical emergency	Medicine
206	Emergency medicine	Medicine
207	Ophthalmology	Medicine
208	Optometry	Medicine
209	Endocrinology	Medicine
210	Internal medicine	Medicine
211	Nursing	Medicine
212	Family medicine	Medicine
213	Intensive care medicine	Medicine
214	Radiology	Medicine
215	Nuclear medicine	Medicine

ID	2^{nd} level	1^{st} level
216	Physical therapy	Medicine
217	Physical medicine and rehabilitation	Medicine
218	Cancer research	Medicine
219	Oncology	Medicine
220	Medical education	Medicine
221	Gerontology	Medicine
222	Virology	Medicine
223	Immunology	Medicine
224	Pediatrics	Medicine
225	Veterinary medicine	Medicine
226	Pathology	Medicine
227	General surgery	Medicine
228	Surgery	Medicine
229	Nuclear magnetic resonance	Medicine
230	Genetics	Medicine
231	Cardiology	Medicine
232	Anesthesia	Medicine
233	Obstetrics	Medicine
234	Gynecology	Medicine
235	Neuroscience	Medicine
236	Gastroenterology	Medicine
237	Traditional medicine	Medicine
238	Physiology	Medicine
239	Audiology	Medicine
240	Urology	Medicine
241	Andrology	Medicine
242	Dermatology	Medicine
243	Anatomy	Medicine
244	Theology	Philosophy
245	Aesthetics	Philosophy
246	Engineering ethics	Philosophy
247	Epistemology	Philosophy
248	Environmental ethics	Philosophy
249	Astronomy	Physics
250	Astrophysics	Physics
251	Molecular physics	Physics
252	Chemical physics	Physics
253	Quantum electrodynamics	Physics
254	Quantum mechanics	Physics
255	Seismology	Physics
256	Geophysics	Physics
257	Particle physics	Physics
258	Nuclear physics	Physics
259	Atomic physics	Physics
260	Classical mechanics	Physics
261	Mathematical physics	Physics
262	Theoretical physics	Physics
263	Condensed matter physics	Physics
264	Optics	Physics
265	Biophysics	Physics
266	Computational physics	Physics
267	Statistical physics	Physics
268	Thermodynamics	Physics
269	Medical physics	Physics
270	Engineering physics	Physics
271	Acoustics	Physics
272	Law	Political science

ID	2^{nd} level	1^{st} level
273	Public administration	Political science
274	Clinical psychology	Psychology
275	Psychotherapist	Psychology
276	Social psychology	Psychology
277	Developmental psychology	Psychology
278	Pedagogy	Psychology
279	Cognitive psychology	Psychology
280	Applied psychology	Psychology
281	Psychoanalysis	Psychology
282	Criminology	Psychology
283	Cognitive science	Psychology
284	Religious studies	Sociology
285	Social science	Sociology
286	Gender studies	Sociology
287	Socioeconomics	Sociology
288	Media studies	Sociology
289	Ethnology	Sociology
290	Anthropology	Sociology
291	Demography	Sociology

Table C1: Conversion table between the second level (292 sub-disciplines) and the first level (19 disciplines) of fields of study. The table also reports node IDs used in the knowledge space (Figure 2).