Journal of Informetrics 18 (2024) 101480



Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi





Unveiling the impact and dual innovation of funded research

Alex J. Yang *

School of Information Management, Nanjing University, Nanjing, 210023, PR China
Jiangsu Key Laboratory of Data Engineering and Knowledge Service, Nanjing, 210023, PR China

ARTICLE INFO

Keywords:
Science of science
Funded research
Disruption
Novelty
Innovation

ABSTRACT

In the relentless pursuit of scientific advancement, comprehending the profound impact and innovation nature inherent in funded research projects assumes paramount significance. To illuminate this matter, I delve into the realm of research supported by the National Institutes of Health (NIH) and the National Science Foundation (NSF). The evaluative framework encompasses a spectrum of metrics, including citations by papers, patents, and Tweets, as markers of research impact. Moreover, I embrace ex-ante innovation (Novelty) and ex-post innovation (Disruption) as dual indispensable yardsticks for evaluating the innovative nature of research projects. Novelty denotes the manifestation of atypical combinations of existing knowledge, while Disruption signifies the extent of paradigm-shifting potential and the ability to exert a disruptive influence on future research endeavors. First, the analysis reveals that funded research projects manifest a conspicuously heightened impact in comparison to their non-funded counterparts. Second, I uncover a noteworthy finding: funded research demonstrates significantly higher levels of ex-ante innovation (Novelty). However, in a surprising twist, the impact of funding on ex-post innovation (Disruption) appears to be faint. Additionally, I undertake a meticulous scrutiny of the robustness of the research findings by scrutinizing patterns across years and fields. Despite the uneven distribution of NIH and NSF funded research and inconspicuous heterogeneity across fields, the patterns of the impact and dual innovation of funded research are consistent across almost all fields.

1. Introduction

In an era characterized by intense competition for research funding and a growing emphasis on scientific and societal impact (Azoulay et al., 2018; Fortunato et al., 2018; Gao et al., 2019; Jusup et al., 2022), understanding the dynamics and outcomes of funded research has become a pivotal concern for the scholarly community (Davies et al., 2022a; Shi et al., 2023; Yin et al., 2022). The allocation of financial resources towards research projects not only shapes the scientific landscape but also carries profound implications for scientific advancement, knowledge dissemination, and societal progress (Chu & Evans, 2021; Fleming et al., 2019; Packalen & Bhattacharya, 2020). As such, a comprehensive evaluation of the impact and innovation nature of funded research is crucial for guiding funding decisions, optimizing resource allocation, and fostering scientific excellence.

The examination of funded research represents a nexus where scientific inquiry intersects with financial support and societal expectations. Funding agencies, such as the National Institutes of Health (NIH) and the National Science Foundation (NSF), play a pivotal role in channeling resources to propel scientific discovery and innovation. Extensive research has investigated the mechanisms

* Correspondence at: School of Information Management, Nanjing University, Nanjing, 210023, PR China.
E-mail address: alexjieyang@outlook.com.

https://doi.org/10.1016/j.joi.2023.101480
Received 16 June 2023; Received in revised form 16 November 2023; Accepted 29 November 2023
Available online 4 December 2023
1751-1577/© 2023 Elsevier Ltd. All rights reserved.

and effects of funding on research outcomes, highlighting the role of funding in shaping the research landscape and influencing scientific productivity, knowledge dissemination, and societal relevance (Galkina Cleary et al., 2018; Li, 2017; McMahan & McFarland, 2021; Park et al., 2015).

The intrinsic unpredictability of scientific advancement, coupled with the complexities surrounding the assessment of impact and innovation in funded research, have ignited ongoing scholarly debates (Azoulay & Li, 2020). While extant literature has indeed furnished valuable insights into the potential impact and innovation embedded in funded research endeavors (Leydesdorff et al., 2019; Packalen & Bhattacharya, 2020), it is important to acknowledge that notable gaps and contentious issues persist, as underscored by recent work (Liu et al., 2023). Firstly, extant studies have predominantly fixated on either small-scale datasets or specific disciplinary domains (Azoulay et al., 2019; Shi et al., 2023), while the intricate and multifaceted nature of diverse disciplinary fields necessitates a more comprehensive approach. Secondly, previous investigations have largely confined themselves to assessing the scientific impact of funded research (Feliciani et al., 2022; Sun et al., 2021), neglecting the broader spectrum of technological and societal repercussions. Thirdly, while some studies have broached the subject of novelty in funded research (Fleming et al., 2019; Packalen & Bhattacharya, 2020), they have often limited their perspective to the ex-ante dimension. We are left with a conspicuous gap in our understanding of the disruptive potential inherent in funded research. The integration of concepts such as technological trajectories, disruptive innovations, and the theory of creative destruction within the discourse surrounding funded research remains an underexplored terrain.

Taken together, there exists a pressing need for a comprehensive, multidimensional analysis that delves into the intricacies of impact and dual innovation perspectives within the realm of funded research. Such an endeavor should consider diverse disciplinary contexts and employ advanced methodological approaches to provide a more nuanced and insightful understanding of this multifaceted landscape. This study compiles a large-scale dataset and develops various measures to assess the scientific, technological, and societal impact of research. Most notably, it constructs a dual innovation framework utilizing ex-ante innovation (Novelty) and ex-post innovation (Disruption) as indicators of the innovative nature of scientific research.

In light of the above considerations, this paper aims to address the following research questions:

(**RQ1**) What are the scientific, technological, and societal impacts of NIH and NSF funded papers compared to non-funded ones?
(**RQ2**) How does funding influence the ex-ante innovation (Novelty) and ex-post innovation (Disruption) of research outputs?
(**RQ3**) How do the impact and innovation of funded research evolve over time? What are the variations in the impact and innovation of funded research across different fields?

The objective of this paper is to provide a comprehensive and in-depth analysis of the impact and innovation potential of funded research. By examining a wide range of metrics, including total citation counts, 5-year citations, citation by patents, citation by Tweets, ex-ante innovation (Novelty), and ex-post innovation (Disruption), this study aims to illuminate the multifaceted outcomes of funded research projects. Furthermore, by considering disciplinary variations and temporal trends, this analysis seeks to offer nuanced insights into the role of funding in propelling scientific progress and knowledge generation across diverse fields.

2. Background

2.1. Dual innovation of science

In this study, we adopt a dual-dimensional approach to innovation, encompassing two essential facets: ex-ante innovation, referred to as Novelty, and ex-post innovation, termed Disruption.

Novelty, the first dimension of innovation, involves the introduction of new ideas or the recombination of existing fragments of knowledge (March, 1991; Nelson, 1985; Schumpeter, 1939). Scientific and technological advancements do not arise spontaneously but are derived from the existing corpus of knowledge (Arthur, 2009). Scientific breakthroughs are often likened to a journey through a vast expanse of combinatorial possibilities, leading to fresh insights and technological progress. Research studies have emphasized the crucial role of combinations in connecting innovation with scientific and technological impact (Hofstra et al., 2020; Trapido, 2015). Most efforts to model the creative process perceive it as an accumulative and interactive recombination of existing fragments of knowledge, merged in novel ways (Azoulay et al., 2011). Novel research endeavors, while potentially entailing higher risks, often strive to address challenging issues and projects sought after by policymakers (Stephan et al., 2017; Wang et al., 2018). Novel studies push the boundaries of existing knowledge and venture into unexplored domains of scientific inquiry (Xu et al., 2022).

Recent research has introduced various quantitative methodologies to assess the novelty of scientific papers (Shi & Evans, 2023; Uzzi et al., 2013; Wang et al., 2017). Among these approaches, the identification of atypical knowledge combinations (Uzzi et al., 2013) is widely acknowledged as the most suitable means of gauging the novelty of scientific research (Fontanaet al., 2020; Hofstra et al., 2020; Romero, 2017; Wagner et al., 2019). This approach, termed "atypicality," centers on identifying atypical combinations of journal pairs (Uzzi et al., 2013) or knowledge entities (Shi & Evans, 2023). While some other methods vary in their approaches to measuring novelty, such as identifying the first-time appearance of knowledge combinations (Wang et al., 2017), recent studies indicate that atypicality demonstrates higher convergent validity (Bornmann et al., 2019; Fontana et al., 2020). Hence, this method is employed in this study.

In the second dimension, Disruption, the concept pertains to the extent of a paradigm shift, as evidenced by the interconnections between former and future knowledge in science (Park et al., 2023). Scientific development mirrors a complex, dynamic system, where innovative science disrupts existing structures (Zeng et al., 2017). The significance of disruptive innovation lies in the profound transformation of paradigms, where the emergence of innovative entities results in marked divergences within the scientific system of

their respective fields (Hofstra et al., 2020). For example, Albert Einstein's theory of relativity (Einstein, 1916) fundamentally altered the conventional scientific framework, while the discovery of the DNA double helix structure by Watson and Crick reshaped our understanding of life's essence (Watson & Crick, 1953). These seminal scientific breakthroughs successfully upended traditional knowledge boundaries and gave rise to new paradigms. This illustrates the essence of innovative science. The concept of disruptive knowledge traces its roots to Thomas Kuhn's theory of paradigm shifts (Kuhn, 1962), suggesting that conventional science addresses unresolved issues within the current paradigm, while disruptive science, to a certain extent, liberates itself from the constraints of existing paradigms, creating entirely novel paradigms or fields. The paradigm shift theory aligns with the concepts of "creative destruction" articulated by Schumpeter (1912) and Bower and Christensen (1995), positing that the value of scientific innovation lies in its inherent potential to fundamentally reshape the structure of the scientific system and the developmental trajectory of the entire field.

To quantify disruption in science, citing links within the knowledge network serve as indicators of the forward and backward impact of papers and patents, contributing to the formation of an extensive network of knowledge flow (Newman, 2012). The disruptive nature of scientific knowledge can be quantified by analyzing the interlinkages within this network (Chen et al., 2021). Recently, the disruption index, introduced by Funk and Owen-Smith (2017), has provided a quantitative measure of the disruptive nature of technologies, specifically by examining the structural characteristics of the deep citation network. This index enables a precise evaluation of the extent to which a patent consolidates or disrupts its paradigm. Building upon this work, Wu et al (2019) and Park et al. (2023) have extended the application of disruption to the realm of scientific literature. Furthermore, the concept of disruption in science has been applied to diverse fields and levels, encompassing the infiltration and flow of disruptive knowledge entities (Wang et al., 2022), the disruptive nature of nodes in complex networks (Yang et al., 2023), the patterns of evolution of disruptive scientific achievements (Park et al., 2023), and the distribution of consolidating and disruptive impact in Nobel-winning papers (Wei et al., 2023) and scientists (Yang et al., 2023), among other areas of investigation.

2.2. Funded research

The investigation of funded research has long captivated the scholarly community, as it represents a pivotal nexus where scientific inquiry intersects with financial support and societal expectations (Davies et al., 2022b; Fortunato et al., 2018). In the current era of increasingly competitive research landscapes, the allocation of resources to research projects carries profound implications for scientific advancement and societal progress (Shi et al., 2023). Consequently, a substantial body of literature has emerged, aiming to elucidate the multifaceted dynamics of funded research, unravel its intricacies, and discern its impact on the scientific enterprise. Fundamental to this discourse is the recognition that funding plays a critical role in shaping the research landscape (Mulligan et al., 2022; Yin al., 2019). Scholars have extensively investigated the mechanisms by which funding agencies, such as the National Institutes of Health (NIH) and the National Science Foundation (NSF), allocate resources to propel scientific discovery (Galkina Cleary et al., 2018; Li, 2017; McMahan & McFarland, 2021; Park et al., 2015). Various theoretical frameworks have been proposed to explain the intricate interplay between funding mechanisms, research agendas, and scientific outcomes. These frameworks encompass concepts such as institutional isomorphism (Chu & Evans, 2021; Ma et al., 2015), path dependency (Davies et al., 2022a; Jones, 2009), and resource dependency (Ma et al., 2015; Packalen & Bhattacharya, 2020), providing lenses through which to understand the complex dynamics governing the allocation of financial resources.

An enduring theme within the literature on funded research revolves around the assessment of its impact. Researchers have sought to uncover the short-term and long-term effects of funded projects, examining factors such as scientific productivity (Shi et al., 2023), knowledge dissemination (Yin et al., 2022), and societal relevance (Leydesdorff et al., 2019). Moreover, the quest for innovation lies at the heart of investigations into funded research. Scholars have also explored the relationship between funding and innovation, seeking to elucidate how financial support influences the generation of novel ideas and the adoption of unconventional approaches (Fleming et al., 2019; Packalen & Bhattacharya, 2020).

However, despite the valuable insights provided by the existing literature on the impact and innovation potential of funded research, notable gaps and controversies persist (Azoulay & Li, 2020). The intricate nature of research outcomes across different fields, the inherent unpredictability of scientific progress over time, and the challenges associated with assessing scientific, technological and societal impact have fueled ongoing debates. Moreover, concepts such as technological trajectories, disruptive innovations, and creative destruction have yet to find their way into the discourse surrounding funded research. These concepts, widely recognized in the realm of innovation studies and technological change, have the potential to enhance the understanding of the dynamics at play within funded research projects. By exploring the role of these concepts in the evaluation of funded research projects, this paper aims to address these gaps and contribute to the existing literature, shedding light on the intricate relationship between funding, impact, and innovation in scientific research.

3. Data and methods

3.1. Data and distributions of funded papers

In this study, two extensive datasets, namely the Microsoft Academic Graph (MAG) and SciSciNet, have been incorporated. The MAG database is renowned for its comprehensive bibliometric coverage of scientific research (Wang et al., 2020). Conversely, SciSciNet is a relatively recent addition to the open-access dataset landscape, provided by Lin et al. (2023). Notably, SciSciNet enriches its dataset by offering supplementary information for each paper, including links to publicly accessible resources. These datasets, with

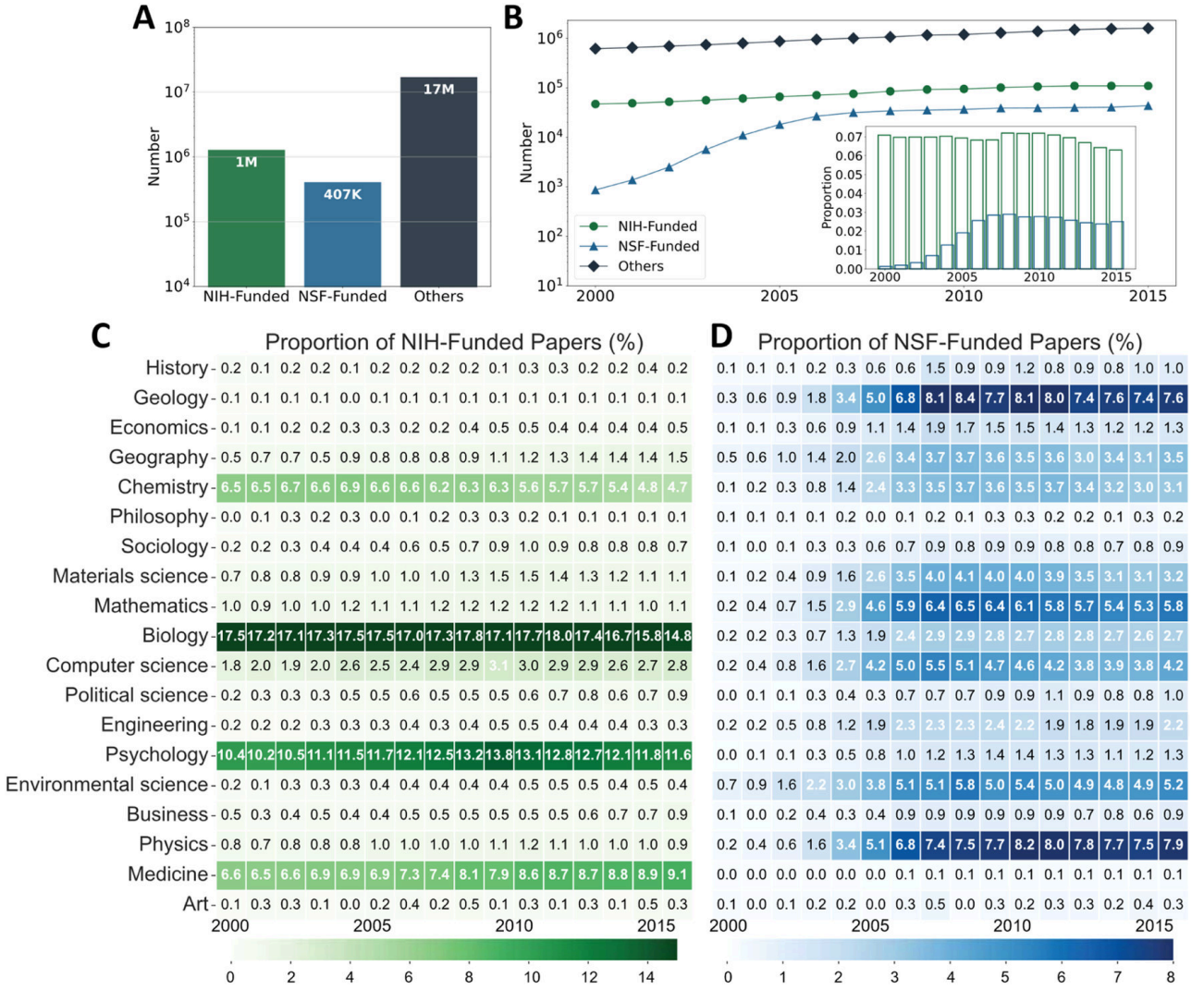


Fig. 1. Mapping the Distribution of NIH and NSF Funded Papers Across Fields and Years. (A) Number of NIH and NSF-funded papers compared to other papers. (B) Number and proportion of NIH and NSF-funded papers compared to other papers across years. (C-D) Proportion of NIH and NSF-funded papers compared to all papers in different years and fields.

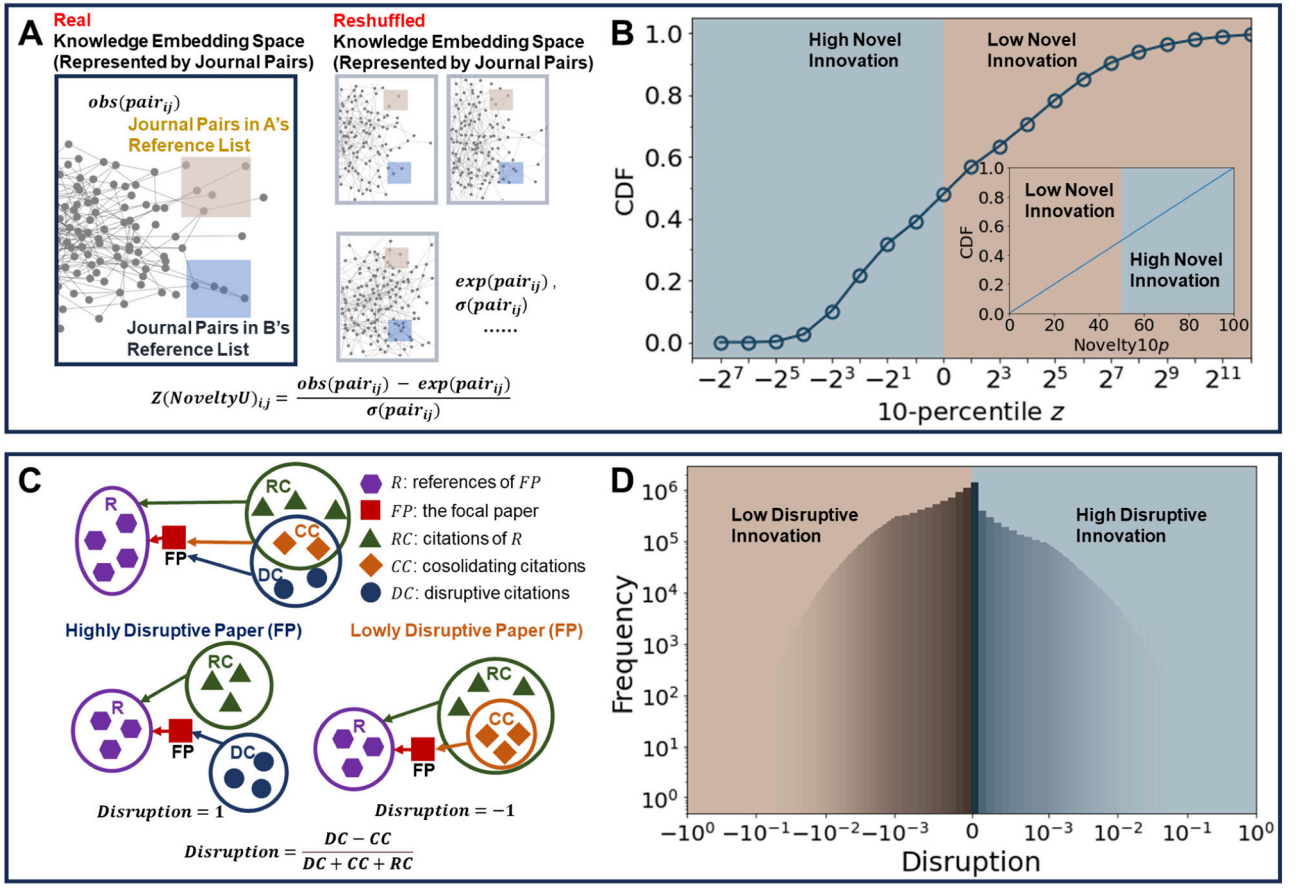


Fig. 2. Assessing Novelty and Disruption in Scientific Papers. (A) Novelty represented by atypical knowledge combinations. (B) Elevated 10-percentile z scores signifies lower novel innovation, so I convert them into $Novelty_{10p}$, which is positively correlated with novelty. (C) Disruption indicated by paradigm shifts in citation networks. (D) Higher Disruption index signifies greater disruptive innovation.

their rich and comprehensive content, serve as invaluable sources of citation data for scrutinizing the characteristics and trends associated with funded research. To ensure an appropriate citation window (Wang, 2013) for analyzing the impact and innovative patterns of funded research, this study focuses specifically on papers published in journals during the period from 2000 to 2015, constituting a total of 18,737,868 papers. The MAG and SciSciNet datasets offer a plethora of information for each paper, including details such as the publication year, paper DOI, paper types, journals, fields, author information, citation links, citation to patents, citation to Tweets, and a host of other properties. Additionally, crucial information pertaining to NIH and NSF funding associated with each paper is also accessible for analysis.

Before delving into the analysis, it is pertinent to present the distribution of NIH and NSF funded papers across fields of study and over time, as depicted in Fig. 1. It is evident from the data that only a relatively small sample of scientific papers receives funding from either the NIH or NSF. Although the number of funded papers has exhibited an upward trajectory over the years (as depicted in Fig. 1B), it is imperative to acknowledge that the average proportion of funded papers remains relatively low when juxtaposed with the overall number of journal papers. On average, approximately 7% of papers receive funding from NIH, while the average proportion of papers funded by the NSF ranges from 0% to 3% (as shown in Fig. 1B). It also reveals noteworthy variations in the proportion of funded papers across diverse fields (as illustrated in Fig. 2C and D). The results elucidate that NIH and NSF funding allocations are distributed unevenly across distinct fields, with each agency displaying a propensity for different domains. The NIH predominantly directs its funding towards research in fields such as Chemistry, Biology, Psychology, and Medicine, whereas the NSF primarily supports research in disciplines such as Geology, Mathematics, Computer Science, Environmental Science, and Physics. In contrast, fields like Philosophy and Art receive comparatively less funding from both the NIH and NSF, with proportions falling below 1%. These findings underscore the complexity and diversity of funding patterns within the research landscape.

3.2. Independent variables

In this study, several variables have been employed to assess various facets of the impact and innovation associated with scientific research papers. These variables are instrumental in providing a holistic evaluation of research output, encompassing academic recognition, real-world applications, and societal implications. Below, I elucidate these variables and their significance:

(1) Citation (C) captures the number of citations received by scientific papers. It serves as a fundamental and widely accepted metric for gauging the academic influence and significance of research papers (Fortunato et al., 2018; Waltman, 2016). By quantifying the extent to which a paper is referenced in the scholarly literature, C offers insights into its academic impact and contribution to the advancement of knowledge.
(2) 5-Year Citation (C5) evaluates the number of citations a paper receives within five years of its publication. It provides a contemporaneous assessment of a paper's influence (Wang, 2013). Focusing on this specific time frame offers a snapshot of the paper's impact during a critical period, providing an indication of its reception within the scientific community and its relevance within a relatively short timeframe.
(3) Citation by patents (CP) offers insights into the practical application and societal implications of scientific research. It quantifies how frequently a scientific publication is cited in patents, establishing a tangible link between scientific knowledge and technological advancements (Marx & Fuegi, 2020, 2022). CP delves into the real-world utilization of research findings, providing a measure of research impact on innovation and highlighting its contribution to societal progress. It sheds light on the significance of research in driving technological development and practical solutions.
(4) Citation by Tweets (CT) provides insights into the societal impact of scientific research, specifically its presence and engagement on social media platforms. CT establishes connections between scientific research and the public, offering a means to investigate discussions and interactions surrounding science on social media platforms. It has become an essential facet of research evaluation as it examines how science is portrayed and discussed in the realm of social media (Hmielowski et al., 2013; Kreps & Kriner, 2020).

These variables collectively provide a comprehensive framework for evaluating the multifaceted dimensions of research impact, encompassing academic recognition, practical applications, and societal resonance. By scrutinizing these dimensions, this paper gains valuable insights into the overall impact of scientific research within these spheres. In addition to measuring the impact, the study also employs other variables to assess the level of innovation, contributing to a nuanced understanding of the research landscape.

(1) To quantify Novelty of papers, I follow the work of Uzzi et al. (2013), and use the codes provided by Gates and Barabási (2023). Specifically, I first conduct ten Monte Carlo simulations (Shor et al., 2007) to generate reshuffled networks by randomly reassigning edges while preserving the temporal and distributional characteristics of the original citation network, as shown in Fig. 2a. Based on the real network and reshuffled network, I examine the distribution of journal pairs found within a paper's reference list. For each referenced journal pair, I convert them into z-scores, a standardized representation. Through Each paper exhibits a distinct distribution of journal pairings, with certain pairings being more or less prevalent than expected by chance. The calculation of $Z(NoveltyU)_{i,j}$ for each journal pair is expressed as:

$$Z(NoveltyU)_{i,j} = \frac{obs(pair_{ij}) - exp(pair_{ij})}{\sigma(pair_{ij})} \quad (1)$$

where $obs(pair_{ij})$ represents the observed frequency of the journal pair in the actual dataset, $exp(pair_{ij})$ denotes the mean, and $\sigma(pair_{ij})$ represents the standard deviation of the number of journal pairs obtained from 10 randomized simulations of the reshuffled network. The distribution of novel knowledge combinations is denoted as a set $\{Z(NoveltyU)_{i,j} \mid i,j \in J\}$, where J represents all journals in the reference list. To encapsulate the information contained within the distribution, the study employs the 10th percentile of the $Z(NoveltyU)_{i,j}$ set as a summary statistic, a methodology established by Uzzi et al. (2013). This summary statistic is then transformed into percentile values, denoted as $Novelty_{10p}$, offering a quantifiable measure of the degree of Novelty exhibited by each paper, as shown in Fig. 2b. By employing this variable, the study gains valuable insights into the unique combinations of knowledge presented within papers.

(1) The disruption index (Funk & Owen-Smith, 2017; Park et al., 2023; Wu et al., 2019) offers a quantitative measure for evaluating the Disruption in scientific papers. As shown in Fig. 2c, the disruption index is defined as:

$$Disruption = \frac{DC - CC}{DC + CC + RC} \tag{4}$$

where CC represents the count of publications citing the focal paper (FP) and its references, DC represents the count of publications citing FP without referencing its cited sources, and RC represents the count of publications citing FP's references without citing FP itself. The disruption index considers the interplay among these citation patterns to gauge the level of Disruption. As shown in Fig. 2c and d, lower disruption indicates that the focal paper cites references that are also cited by a substantial number of papers referencing FP. This suggests the consolidation of the existing paradigm, as the focal paper aligns with and reinforces established knowledge. Conversely, higher disruption values characterize papers as highly disruptive, signifying their capacity to induce paradigm shifts. The perspective of disruption gains insights into the disruptive nature of scientific research, shedding light on patterns of knowledge dissemination and paradigm shifts in science.

3.3. Control variables and fixed effects

To address potential confounding factors that might exert influence on the levels of impact and innovation within funded research, this analysis incorporates a set of control variables.

Firstly, collaboration in scientific research can significantly impact research outcomes. This study controls for collaboration through several dimensions: (1) The number of authors involved in a research paper (Wuchty et al., 2007), which accounts for the collaborative nature of the research effort. (2) The unique number of institutions engaged in a study, signifying cross-disciplinary collaboration (Jones et al., 2008), acknowledging that diverse institutional contributions can influence research outcomes. (3) The unique number of countries participating in a study, indicating international collaborations (Wagner et al., 2019), recognizing the global reach and potential impact of such collaborations.

Secondly, the number of references in a scientific paper is considered a control variable as it reflects the breadth of the literature reviewed and incorporated into the research (Leydesdorff et al., 2019). This variable is indicative of the extent to which a paper draws upon existing knowledge, which can influence its impact and innovation potential.

Thirdly, the analysis incorporates the Rao-Stirling index (Stirling, 2007) as a measure of interdisciplinarity. It is defined as

$$Interdisciplinarity(RS) = 2\sum_{i \neq j} d_{ij} p_i p_j$$

where p_i denotes the fraction of references of a paper in discipline i, and d_{ij} is the distance between discipline i and j. An interdisciplinarity of 0 reflects a paper with no diversity, while an interdisciplinarity of 1 denotes maximal diversity. This study uses 292 s-level disciplines in MAG to calculate the interdisciplinarity. The interdisciplinarity takes into account the balance, variety, and disparity of papers, reflecting the diversity of knowledge sources utilized in the research (Leydesdorff, 2018). It aids in controlling for the interdisciplinary nature of research papers, ensuring that this aspect is adequately considered.

Fourthly, the h-index of the first author is included as a control variable to account for the ability and academic status of authors (Hirsch, 2005; Yang et al., 2023). This measure provides insights into the scholarly impact and productivity of the primary author, which can exert an influence on the overall impact of the research.

Furthermore, to address potential heterogeneity across scientific fields, journals, and temporal variations, three-level fixed effects are incorporated into the regression models: (1) Second-level fields in the MAG dataset (totaling 292) are utilized to control for variations among different scientific domains, acknowledging that research impact and innovation may differ across disciplines. (2) The inclusion of journal fixed effects helps account for heterogeneity related to the publication outlets. A total of 49,066 distinct journals are included in the journal fixed effects. (3) Year fixed effects are also included to capture temporal variations and trends over time, ensuring that any changes in impact and innovation are adequately considered within the analysis.

3.4. Descriptive statistics

Table 1 offers insights into the characteristics of NIH and NSF funded papers in comparison to non-funded counterparts. Notably, the average impact of NIH and NSF funded papers substantially surpasses that of non-funded ones. For example, the average citation count for NIH-funded papers stands at 74.19, while NSF-funded papers receive an average of 64.06 citations, in stark contrast to the average of 33.03 citations for non-funded papers. This disparity underscores the pronounced influence and recognition garnered by funded research endeavors within the scholarly community. Additionally, the average level of Novelty exhibited by funded papers surpasses that of non-funded papers. However, it is noteworthy that the average level of Disruption, whether for funded or non-funded papers, remains virtually identical, with both groups registering an average score of 0.00 in this regard.

Furthermore, NIH and NSF funded papers appear to feature larger and more diverse research teams, greater interdisciplinarity, and elevated first-author h-index values. These observations align with findings from previous studies (Azoulay & Li, 2020).

4. Results

4.1. RQ1: scientific impact of funded research

In pursuit of an in-depth understanding of the scientific impact of funded research, I conduct Ordinary Least Squares (OLS) regression analyses, the outcomes of which are detailed in Table 2. The results unveil a significant and positive effect of both NIH and NSF funding on the scientific impact of research endeavors. When controlling for all relevant variables, NIH funding demonstrates the capacity to augment research impact, leading to an increase of 15.1% in the citation count and 12.8% in the 5-year citation count. Similarly, NSF funding exerts a notable influence, culminating in an 11.5% increase in the citation count and a 10.4% increase in the 5-year citation count. This compelling finding not only underscores the tangible benefits of funding in enhancing research impact but also suggests that funding mechanisms wield a more pronounced influence on long-term impact as compared to short-term impact.

These results resplendently illuminate the pivotal and enduring role played by NIH and NSF in fortifying the scientific impact of research, thereby furnishing empirical substantiation for the enduring significance of funding agencies in advancing the dissemination of scientific knowledge and the attainment of scholarly recognition.

Table 1
Descriptive statistics of all papers.

		Sample Size	Mean	Std	Min	Median	Max
Papers funded by NIH	**C**	1,287,998	74.19	192.33	1.00	37.00	33,462.00
	C5	1,287,998	34.21	71.45	0.00	19.00	24,550.00
	CP	1,287,998	1.83	11.68	0.00	0.00	2617.00
	CT	1,287,998	0.26	24.88	0.00	0.00	22,604.00
	$Novelty_{10p}$	1,287,998	61.19	27.51	0.00	67.25	100.00
	Disruption	1,287,998	0.00	0.01	−0.97	0.00	0.99
	#Authors	1,287,998	5.88	7.52	1.00	5.00	2918.00
	#Institutions	1,287,998	2.00	2.09	1.00	1.00	463.00
	#Countries	1,287,998	1.26	0.72	1.00	1.00	60.00
	#References	1,287,998	46.02	34.61	2.00	40.00	1000.00
	Interdisciplinarity	1,287,705	0.19	0.06	0.00	0.20	0.41
	Author h-index	1,287,998	39.24	31.32	1.00	32.00	306.00
Papers funded by NSF	**C**	407,317	64.06	204.30	1.00	28.00	31,597.00
	C5	407,317	30.57	84.81	0.00	16.00	22,466.00
	CP	407,317	0.82	10.46	0.00	0.00	3210.00
	CT	407,317	0.34	15.24	0.00	0.00	5096.00
	$Novelty_{10p}$	407,317	54.46	31.45	0.00	57.20	100.00
	Disruption	407,317	0.00	0.01	−0.97	0.00	0.88
	#Authors	407,317	8.43	95.59	1.00	4.00	5104.00
	#Institutions	407,317	2.35	6.83	1.00	2.00	314.00
	#Countries	407,317	1.43	1.62	1.00	1.00	53.00
	#References	407,317	41.87	33.90	2.00	35.00	1149.00
	Interdisciplinarity	406,570	0.20	0.06	0.00	0.20	0.40
	Author h-index	407,317	32.33	25.52	1.00	27.00	241.00
Other papers	**C**	17,085,049	33.03	108.25	1.00	14.00	106,719.00
	C5	17,085,049	14.49	36.45	0.00	7.00	37,405.00
	CP	17,085,049	0.54	8.47	0.00	0.00	5712.00
	CT	17,085,049	0.14	12.13	0.00	0.00	16,414.00
	$Novelty_{10p}$	17,085,049	49.09	28.73	0.00	48.72	100.00
	Disruption	17,085,049	0.00	0.02	−1.00	0.00	1.00
	#Authors	17,085,049	4.22	8.05	1.00	4.00	3144.00
	#Institutions	17,085,049	1.49	1.19	1.00	1.00	295.00
	#Countries	17,085,049	1.19	0.56	1.00	1.00	69.00
	#References	17,085,049	29.45	27.30	2.00	24.00	1361.00
	Interdisciplinarity	16,782,364	0.18	0.07	0.00	0.19	0.44
	Author h-index	17,085,049	21.31	21.29	1.00	15.00	306.00

Table 2
Scientific impact of funded papers.

Models	Citation (C)				5-year Citation (C5)			
	(1) log(C)	(2) log(C)	(3) log(C)	(4) log(C)	(5) log(C5)	(6) log(C5)	(7) log(C5)	(8) log(C5)
NIH	0.3062***	0.1510***			0.2597***	0.1284***		
	(0.0055)	(0.0037)			(0.0045)	(0.0028)		
NSF			0.2309***	0.1153***			0.2009***	0.1038***
			(0.0099)	(0.0073)			(0.0082)	(0.0063)
log(#Au)		0.2029***		0.2032***		0.1674***		0.1677***
		(0.0043)		(0.0043)		(0.0031)		(0.0031)
log(#Ins)		0.0148***		0.0208***		0.0147***		0.0197***
		(0.0008)		(0.0009)		(0.0009)		(0.0010)
log(#Coun)		0.0532***		0.0460***		0.0512***		0.0452***
		(0.0034)		(0.0036)		(0.0029)		(0.0032)
log(#Ref)		0.5549***		0.5569***		0.4705***		0.4722***
		(0.0041)		(0.0042)		(0.0035)		(0.0035)
Interdisciplinarity		−1.216***		−1.216***		−1.485***		−1.486***
		(0.1016)		(0.1013)		(0.0401)		(0.0399)
Author h-index		0.0053***		0.0054***		0.0047***		0.0048***
		(0.0002)		(0.0002)		(0.0001)		(0.0001)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Journal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	18,576,248	18,434,146	18,576,248	18,434,146	18,576,248	18,434,146	18,576,248	18,434,146
R^2	0.20294	0.27653	0.20227	0.27643	0.03903	0.04981	0.03902	0.04981

Note: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$. All models are based on OLS regression. Clustered standard errors are in parentheses.

4.2. RQ1: technological and societal impact of funded research

Table 3 presents the estimated effects of NIH and NSF funding on the technological and societal impact of research, shedding light on the multifaceted influence of funding sources. The results reveal that both NIH and NSF funding significantly enhance the technological and societal impact of research endeavors.

Upon meticulous control for all pertinent variables, NIH funding distinctly showcases its capacity to amplify research impact within the technological domain, culminating in a substantial 12.1% increase in citation by patents. Furthermore, although the impact on social engagement, as measured by citation by Tweets, appears relatively modest, it remains statistically significant, with a 0.4% increase attributable to NIH funding. Similarly, NSF funding exerts a notable influence, yielding a 4.5% augmentation in citation by patents and a 2.1% elevation in citation by Tweets. It is imperative to note that while the coefficients associated with these effects attain statistical significance, the magnitudes of the effects may seem relatively diminutive. This observation can be attributed to the

Table 3
Technological and societal impact of funded papers.

Models	Citation by Patents (CP)				Citation by Tweets (CT)			
	(1) log(CP)	(2) log(CP)	(3) log(CP)	(4) log(CP)	(5) log(CT)	(6) log(CT)	(7) log(CT)	(8) log(CT)
NIH	0.1627***	0.1212***			0.0133***	0.0040**		
	(0.0135)	(0.0090)			(0.0024)	(0.0013)		
NSF			0.0794***	0.0447***			0.0288***	0.0207***
			(0.0076)	(0.0040)			(0.0033)	(0.0023)
log(#Au)		0.0413***		0.0425***		−0.0027***		−0.0027***
		(0.0039)		(0.0040)		(0.0004)		(0.0004)
log(#Ins)		−0.0164***		−0.0097***		0.0101***		0.0099***
		(0.0013)		(0.0012)		(0.0017)		(0.0018)
log(#Coun)		−0.0039		−0.0118***		0.0032***		0.0032***
		(0.0022)		(0.0028)		(0.0006)		(0.0006)
log(#Ref)		0.0599***		0.0626***		0.0097***		0.0096***
		(0.0068)		(0.0072)		(0.0012)		(0.0012)
Interdisciplinarity		−0.3992***		−0.4056***		−0.0031		−0.0034
		(0.0433)		(0.0444)		(0.0043)		(0.0044)
Author h-index		0.0014***		0.0016***		0.0003***		0.0003***
		(0.0001)		(0.0001)		(3.78e-5)		(3.86e-5)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Journal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	18,576,248	18,434,146	18,576,248	18,434,146	18,576,248	18,434,146	18,576,248	18,434,146
R^2	0.20294	0.27653	0.20227	0.27643	0.03903	0.04981	0.03902	0.04981

Note: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$. All models are based on OLS regression. Clustered standard errors are in parentheses.

relatively infrequent occurrence of citation by patents and Tweets in research papers, as elucidated in Supplementary Material Note 3.

Notwithstanding the relatively modest effect sizes, these results underscore the unequivocal and positive impact of funding sources on both the technological and social dimensions of research impact. The infrequency of such links notwithstanding, these findings affirm the constructive influence of NIH and NSF funding in advancing the broader societal outreach and technological relevance of research endeavors.

4.3. RQ2: dual innovation of funded research

The exploration of dual dimensions of innovation, encompassing ex-ante innovation (Novelty) and ex-post innovation (Disruption) within funded researches, unveils intriguing insights into the influence of NIH and NSF funding on the innovative character of research output, as delineated in Table 4.

It becomes evident that both NIH and NSF funding exert a significant and positive influence on the Novelty of research papers. Even after meticulous control for confounding factors, these funding sources contribute notably to the enhancement of the Novelty of papers. NIH funding leads to a 2.1 unit increase, while NSF funding results in a more substantial 2.6 unit increase. This observation underscores the pivotal role of funding agencies in fostering research endeavors that manifest novel combinations of knowledge, thereby advancing the innovative landscape of scientific output.

Conversely, the effect of NIH and NSF funding on Disruption appears to exhibit a more nuanced pattern. When considering the controlled variables, the impact on Disruption remains positive but of a notably smaller magnitude. NIH funding is associated with a minute 0.0002 increase, while NSF funding results in a slightly larger 0.0003 increase in Disruption. These effects, while statistically significant, appear quite modest in magnitude. Intriguingly, when not controlling for these variables, the relationship between funding and Disruption takes a surprising turn, suggesting a negative association between NIH and NSF funding and the disruptive nature of research papers. Based on this finding, it is inconclusive whether NIH and NSF funding can significantly enhance the Disruption of papers.

These findings illuminate the intricate interplay between funding and innovation within the research landscape. While fundings significantly enhance the ex-ante innovation (Novelty) of papers, their impact on the ex-post innovation (Disruption) of research appears to be faint.

4.4. RQ3: robustness check and heterogeneity analysis

Given the extensive purview of this research, predicated on a substantial dataset, a meticulous examination of the robustness of the findings becomes imperative. To appraise the stability of the results, I initiate a discerning comparison of impact and dual innovation dimensions across NIH and NSF funded papers vis-à-vis other papers. This comprehensive analysis, detailed in Supplementary Material Note 3, elucidates that NIH and NSF funded papers consistently manifest elevated impact and ex-ante innovation across diverse temporal and disciplinary domains.

To further fortify the reliability of the findings, I execute a split-sample regression for distinct time periods spanning from 2000 to

Table 4
Ex-ante (Novelty) and ex-post (Disruption) innovation of funded papers.

Models	Novelty				Disruption			
	(1) $Novelty_{10p}$	(2) $Novelty_{10p}$	(3) $Novelty_{10p}$	(4) $Novelty_{10p}$	(5) Disruption	(6) Disruption	(7) Disruption	(8) Disruption
NIH	4.356***	2.054***			−0.0004***	0.0002***		
	(0.1095)	(0.0942)			(4.8e-5)	(2.79e-5)		
NSF			4.375***	2.565***			−0.0002.	0.0003*
			(0.0752)	(0.1033)			(0.0001)	(0.0001)
log(#Au)		1.891***		1.894***		−0.0004***		−0.0004***
		(0.0441)		(0.0432)		(3.01e-5)		(3.02e-5)
log(#Ins)		0.9846***		1.047***		0.0002***		0.0003***
		(0.0407)		(0.0409)		(1.01e-5)		(9.7e-6)
log(#Coun)		−1.614***		−1.698***		−0.0002***		−0.0002***
		(0.0636)		(0.0638)		(3.23e-5)		(3.25e-5)
log(#Ref)		10.53***		10.55***		−0.0031***		−0.0031***
		(0.1878)		(0.1875)		(0.0003)		(0.0003)
Interdisciplinarity		37.63***		37.62***		−0.0094***		−0.0094***
		(1.039)		(1.037)		(0.0011)		(0.0011)
Author h-index		0.0095***		0.0109***		−8.12e-6***		−7.96e-6***
		(0.0021)		(0.0020)		(4.44e-7)		(4.47e-7)
Field FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Journal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	18,576,248	18,434,146	18,576,248	18,434,146	18,576,248	18,434,146	18,576,248	18,434,146
$\mathbf{R^2}$	0.20294	0.27653	0.20227	0.27643	0.03903	0.04981	0.03902	0.04981

Note: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$. All models are based on OLS regression. Clustered standard errors are in parentheses.

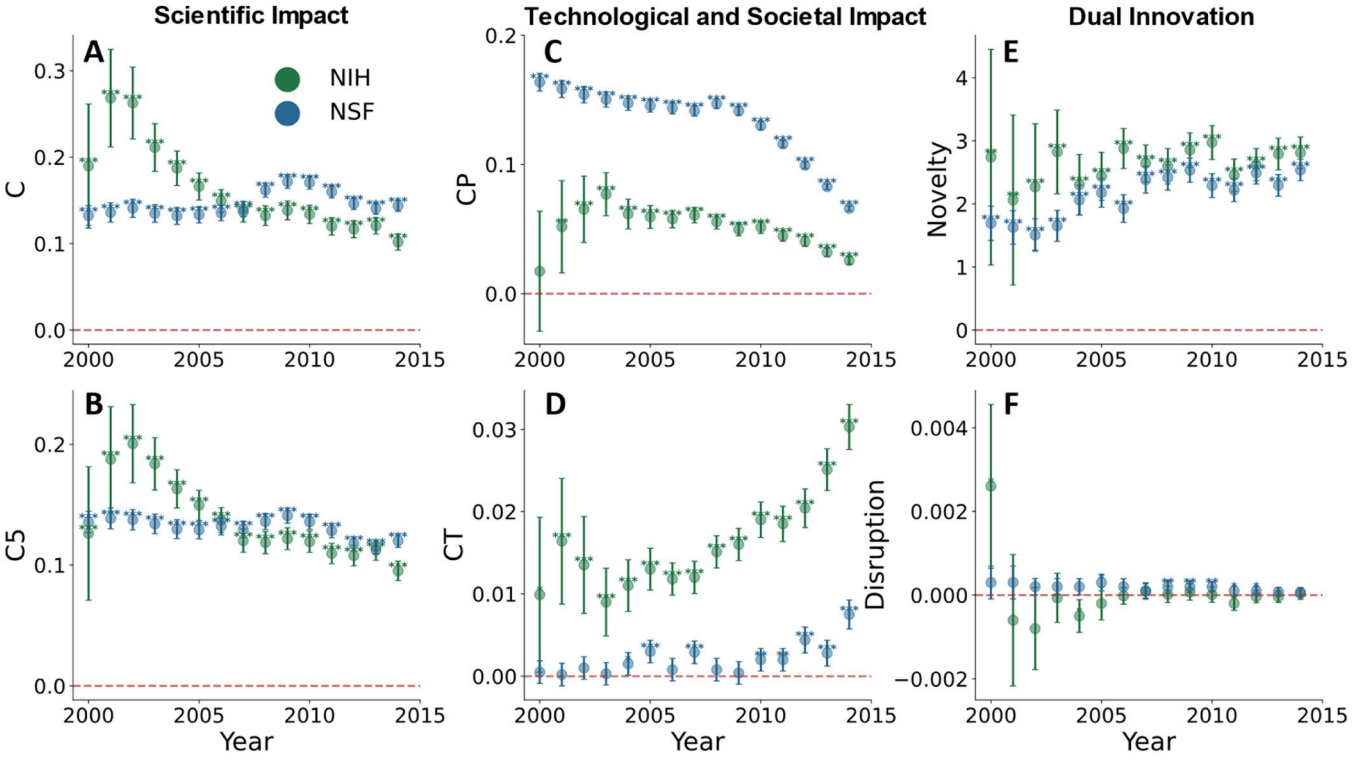


Fig. 3. Temporal Trajectories of Impact and Dual Innovation Effects in NIH and NSF Fundings: A Split-Sample Regression Analysis (2000–2015). (A-F) The effects of (A) Citation, (B) 5-Year Citation, (C) Technological Impact (CP), (D) Societal Impact (CT), (E) Novelty and (F) Disruption for NIH and NSF-funded papers over the examined time period. The error bars represent the upper and lower bounds of 95% confidence intervals. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

2015. Incorporating all relevant control variables and fixed effects as detailed in Tables 2-4, the outcomes are depicted in Fig. 3 and Supplementary Material Note 5. The discernible patterns unveil the persistent role of funding in amplifying scientific impact, technological and societal impact, as well as ex-ante innovation (Novelty) of research papers. Intriguingly, however, the discerned findings indicate a muted effect of funding in elevating ex-post innovation (Disruption). Moreover, a nuanced temporal evolution is discernible in the impact of funding on technological and societal dimensions. Specifically, the effect of funding in augmenting technological impact diminishes over time, while its influence on societal impact exhibits an escalating trajectory.

Furthermore, when discerning the comparative impacts of NIH and NSF, a temporal bifurcation emerges. The effect of NIH in enhancing scientific impact surpasses that of NSF before 2007, subsequently yielding to a reversal in which NIH demonstrates a lower impact compared to NSF after that point. Concurrently, NIH exhibits a more pronounced impact in elevating societal impact, while NSF attains greater influence in amplifying technological impact. This temporal disjunction underscores the necessity for additional research to delve more profoundly into the intricate mechanisms and multifaceted factors steering these temporal patterns.

Secondly, this investigation undertakes a nuanced exploration of the robustness of its findings by conducting comparisons across diverse academic fields. Employing a split-sample regression for discrete time periods across 19 first-level fields and integrating all pertinent control variables and fixed effects, including those associated with the 292 s-level fields, the outcomes are encapsulated in Fig. 4. The results illuminate significant heterogeneity discernible across disparate fields. Notably, in Art, Philosophy, and History, the standard errors exhibit considerable magnitude, and the coefficients fail to attain statistical significance in numerous instances. This underscores the intricate nature of these disciplines, where the impact of funding sources may be less straightforward and is often subject to greater variability. It is noteworthy that the effects of NIH and NSF fundings emerge as largely significant in fields aligned with their intended focus, specifically the STEM subjects or natural sciences, as expounded in Section 3.1 and delineated in Fig. 1. This alignment underscores the targeted impact of these fundings within domains where their design and objectives are most effectively realized.

Further scrutiny is devoted to papers funded by both NIH and NSF, yielding results consistent with the main findings. However, it also reveals that papers funded by singular sources of grants may not exhibit significant advantages in terms of impact and Novelty (see Supplementary Material Note 4). Collectively, these additional analyses confirm the stability and consistency of the results, reinforcing

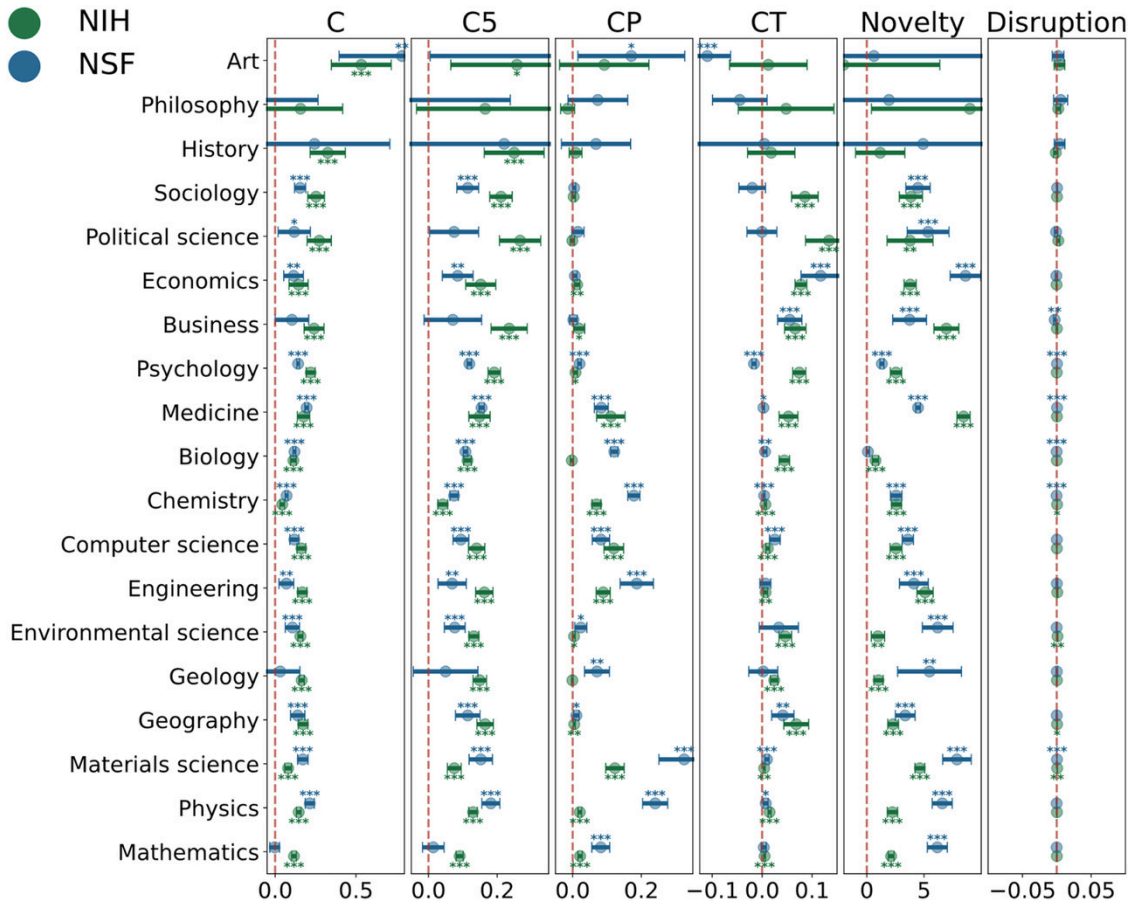


Fig. 4. Field Heterogeneity: Impact and Dual Innovation Effects in NIH and NSF Fundings - A Split-Sample Regression Analysis across 19 Fields. The error bars represent the upper and lower bounds of 95% confidence intervals. Significance levels are denoted as *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$.

the argument that NIH and NSF funding positively influence impact and ex-ante innovation (Novelty) across a broad spectrum of scientific disciplines, while their effect on ex-post innovation (Disruption) appears to be less pronounced.

5. Discussion

5.1. Theoretical and practical implications

This research underscores the pivotal role of funding agencies in fostering scientific impact. The findings contribute to the growing body of literature emphasizing the significance of financial support in shaping research outcomes (Azoulay & Li, 2020; Davies et al., 2022a; Shi et al., 2023; Sun et al., 2021; Yin et al., 2022). By empirically demonstrating the positive influence of these funding sources on scientific, technological and societal impact, this study reinforces the theoretical underpinning that research funding is a crucial driver of scientific progress (Mulligan et al., 2022; Packalen & Bhattacharya, 2020). Most importantly, this paper reveals a nuanced relationship between funding and different dimensions of innovation (Park et al., 2023; Uzzi et al., 2013; Wu et al., 2019). While NIH and NSF funding significantly enhance ex-ante innovation (Novelty), their effects on ex-post innovation (Disruption) are more subdued. This finding introduces a new layer of complexity to the theory of funding and innovation, highlighting the need for a more comprehensive understanding of how funding sources shape research outcomes across various dimensions.

The findings of this study have several implications for funding agencies, policymakers, and researchers. The analysis of the impact and innovation potential of funded research provides valuable guidance for funding decisions and resource allocation. The findings provide insights into the importance of funding agencies as critical drivers of scientific progress, highlighting their role in promoting impactful research and advancing knowledge within various fields (Galkina Cleary et al., 2018; McMahan & McFarland, 2021). Understanding these patterns can help funding agencies identify projects with the potential for significant scientific and societal contributions. This knowledge can inform the development of funding strategies and policies that foster scientific excellence and maximize the return on investment in research (Azoulay et al., 2018; Yin et al., 2022).

From a broader perspective, the findings of this study contribute to the ongoing discussions surrounding research evaluation and the assessment of scientific impact and innovation. By examining a wide range of metrics, this study provides a more comprehensive approach to evaluating research outcomes. This research also expands the existing theoretical frameworks by incorporating concepts such as technological trajectories, disruptive innovations, and creative destruction into the discourse surrounding funded research (Bower & Christensen, 1995; Wu et al., 2019; Yang et al., 2023). This holistic perspective can inform the development of more robust and nuanced evaluation frameworks that capture the multidimensional nature of research impact and innovation (Zeng et al., 2017). Such frameworks can facilitate fair and accurate assessments of research quality and enable more informed decision-making in research funding, hiring, and promotion processes.

5.2. Limitations and outlook

Despite the comprehensive analysis conducted in this paper, it is important to acknowledge several limitations that warrant consideration. Firstly, the observed paradoxical relationship between ex-ante innovation (Novelty) and ex-post innovation (Disruption) in funded research calls for further exploration. While this study identifies differences in their impact, the underlying mechanisms and contextual factors contributing to these contrasting patterns of innovation remain unclear. Future research should delve into these complexities to provide a deeper understanding of the dynamics at play.

Secondly, while this analysis reveals subtle heterogeneity across specific fields and years, a more detailed examination of these variations was not undertaken. Additionally, the focus on NIH and NSF funding sources, while significant, overlooks the multitude of funding agencies worldwide, including international agencies. Investigating this heterogeneity within a broader funding landscape could yield valuable insights into the factors influencing the impact and innovation of funded research across diverse contexts.

Thirdly, this study employs fixed-effect regression methods to analyze the role of funding in impact and innovation. However, it is important to recognize that this approach may be limited by potential endogeneity issues, and not all confounding factors can be controlled for. Future studies should explore quasi-experimental (Angrist & Pischke, 2010) or real-world randomized experiments (Azoulay, 2012) to establish causal relationships and better understand the true impact of funding, requiring endeavors for both researchers and funding agencies (Azoulay et al., 2018).

Lastly, this study primarily focuses on examining the impact and innovation of funded research without delving into the underlying mechanisms. While the results suggest associations between funding and research outcomes, it is crucial to conduct additional investigations to uncover the mechanisms through which funding influences impact and innovation. Qualitative and mixed-methods approaches (Liu et al., 2023) could be employed to explore specific mechanisms and contextual factors mediating the relationship between funding and research outcomes, enriching our understanding of these complex dynamics.

By addressing these limitations and embarking on future research endeavors, we can further enhance our comprehension of the impact and innovation dynamics in funded research. This knowledge will not only guide funding decisions and resource allocation strategies but also foster scientific excellence and contribute to societal progress.

CRediT authorship contribution statement

Alex J. Yang: Writing – original draft, Conceptualization, Methodology, Software, Visualization, Data curation, Resources, Writing – review & editing.

Declaration of Competing Interest

I declare no competing interest relevant to this article.

Data availability

The data linkage associated with SciSciNet can be accessed through the following link: https://doi.org/10.6084/m9.figshare.c.6076908.v1 . This platform offers unrestricted access to downloadable data files in zip or tsv format, without any limitations or restrictions.

Acknowledgments

The author acknowledges that no funding support was received for this research paper. The study was conducted independently, and the author would like to express gratitude to three anonymous reviewers for their valuable insights.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.joi.2023.101480.

References

Angrist, J. D., & Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives, 24*(2), 3–30. https://doi.org/10.1257/jep.24.2.3

Arthur, W. B. (2009). *The nature of technology: What it is and how it evolves*. Simon and Schuster.

Azoulay, P. (2012). Turn the scientific method on ourselves. *Nature, 484*(7392), 31–32. https://doi.org/10.1038/484031a

Azoulay, P., Fuchs, E. R. H., Goldstein, A. P., & Kearney, M. (2018). Funding breakthrough research: promises and challenges of the "ARPA Model. *Innovation Policy and the Economy, 19*, 69–96.

Azoulay, P., Graff-Zivin, J., Uzzi, B., Wang, D., Williams, H., Evans, J. A., Jin, G. Z., Lu, S. F., Jones, B. F., Börner, K., Lakhani, K. R., Boudreau, K. J., & Guinan, E. C. (2018). Toward a more scientific science. *Science, 361*(6408), 1194–1197. https://doi.org/10.1126/science.aav2484

Azoulay, P., Graff Zivin, J. S., Li, D., & Sampat, B. N. (2019). Public R&D investments and private-sector patenting: evidence from NIH funding rules. *The Review of Economic Studies, 86*(1), 117–152. https://doi.org/10.1093/restud/rdy034

Azoulay, P., Graff Zivin, J. S., & Manso, G. (2011). Incentives and creativity: evidence from the academic life sciences. *The RAND Journal of Economics, 42*(3), 527–554. https://doi.org/10.1111/j.1756-2171.2011.00140.x

Azoulay, P., & Li, D. (2020). *National Bureau of Economic Research Working Paper Series, No. 26889*. https://doi.org/10.3386/w26889

Bornmann, L., Tekles, A., Zhang, H. H., & Ye, F. Y. (2019). Do we measure novelty when we analyze unusual combinations of cited references? A validation study of bibliometric novelty indicators based on F1000Prime data. *Journal of Informetrics, 13*(4), Article 100979. https://doi.org/10.1016/j.joi.2019.100979

Bower, J. L., & Christensen, C. M. (1995). Disruptive Technologies: Catching the Wave. *Journal of Product Innovation Management, 1*, 75–76.

Chen, J., Shao, D., & Fan, S. (2021). Destabilization and consolidation: Conceptualizing, measuring, and validating the dual characteristics of technology. *Research Policy, 50*(1), Article 104115. https://doi.org/10.1016/j.respol.2020.104115

Chu, J. S. G., & Evans, J. A. (2021). Slowed canonical progress in large fields of science. *Proceedings of the National Academy of Sciences, 118*(41), Article e2021636118. https://doi.org/10.1073/pnas.2021636118

Davies, B., Gush, J., Hendy, S. C., & Jaffe, A. B. (2022a). Research funding and collaboration [Article]. *Research Policy, 51*(2), Article 104421. https://doi.org/10.1016/j.respol.2021.104421

Davies, B., Gush, J., Hendy, S. C., & Jaffe, A. B. (2022b). Research funding and collaboration. *Research Policy, 51*(2), Article 104421. https://doi.org/10.1016/j.respol.2021.104421

Einstein, A. (1916). The foundation of the general theory of relativity. *Annalen Phys, 49*(7), 769–822.

Feliciani, T., Morreau, M., Luo, J., Lucas, P., & Shankar, K. (2022). Designing grant-review panels for better funding decisions: Lessons from an empirically calibrated simulation model. *Research Policy, 51*(4), Article 104467. https://doi.org/10.1016/j.respol.2021.104467

Fleming, L., Greene, H., Li, G., Marx, M., & Yao, D. (2019). Government-funded research increasingly fuels innovation. *Science, 364*(6446), 1139–1141. https://doi.org/10.1126/science.aaw2373

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

Fortunato, S., Bergstrom, C. T., Boerner, K., Evans, J. A., Helbing, D., Milojevic, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., Vespignani, A., Waltman, L., Wang, D., & Barabasi, A.-L. (2018). Science of science. *Science, 359*(6379), eaao0185. https://doi.org/10.1126/science.aao0185

Funk, R. J., & Owen-Smith, J. (2017). A dynamic network measure of technological change. *Management Science, 63*(3), 791–817. https://doi.org/10.1287/mnsc.2015.2366

Galkina Cleary, E., Beierlein, J. M., Khanuja, N. S., McNamee, L. M., & Ledley, F. D (2018). Contribution of NIH funding to new drug approvals 2010–2016. *Proceedings of the National Academy of Sciences, 115*(10), 2329–2334. https://doi.org/10.1073/pnas.1715368115

Gao, J., Zhang, Y.-C., & Zhou, T. (2019). Computational socioeconomics. *Physics Reports, 817*, 1–104. https://doi.org/10.1016/j.physrep.2019.05.002

Gates, A. J., & Barabási, A.-L. (2023). Reproducible science of science at scale: pySciSci. *Quantitative Science Studies*, 1–11. https://doi.org/10.1162/qss_a_00260

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences, 102*(46), 16569–16572. https://doi.org/10.1073/pnas.0507655102

Hmielowski, J. D., Feldman, L., Myers, T. A., Leiserowitz, A., & Maibach, E. (2013). An attack on science? Media use, trust in scientists, and perceptions of global warming. *Public Understanding of Science, 23*(7), 866–883. https://doi.org/10.1177/0963662513480091

Hofstra, B., Kulkarni, V. V., Munoz-Najar Galvez, S., He, B., Jurafsky, D., & McFarland, D. A (2020). The diversity-innovation paradox in science. *Proceedings of the National Academy of Sciences, 117*(17), 9284–9291. https://doi.org/10.1073/pnas.1915378117

Jones, B. F. (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. https://doi.org/10.1111/j.1467-937X.2008.00531.x

Jones, B. F., Wuchty, S., & Uzzi, B. (2008). Multi-university research teams: shifting impact, geography, and stratification in science. *Science, 322*(5905), 1259–1262. https://doi.org/10.1126/science.1158357
Jusup, M., Holme, P., Kanazawa, K., Takayasu, M., Romić, I., Wang, Z., Geček, S., Lipić, T., Podobnik, B., Wang, L., Luo, W., Klanjšček, T., Fan, J., Boccaletti, S., & Perc, M. (2022). Social physics. *Physics Reports, 948*, 1–148. https://doi.org/10.1016/j.physrep.2021.10.005
Kreps, S. E., & Kriner, D. L. (2020). Model uncertainty, political contestation, and public trust in science: Evidence from the COVID-19 pandemic. *Science Advances, 6*(43), eabd4563. https://doi.org/10.1126/sciadv.abd4563
Kuhn, T. S. (1962). Historical structure of scientific discovery. *Science, 136*(3518), 760–764.
Leydesdorff, L. (2018). Diversity and interdisciplinarity: how can one distinguish and recombine disparity, variety, and balance? *Scientometrics, 116*(3), 2113–2121. https://doi.org/10.1007/s11192-018-2810-y
Leydesdorff, L., Bornmann, L., & Wagner, C. S. (2019). The relative influences of government funding and international collaboration on citation impact. *Journal of the Association for Information Science and Technology, 70*(2), 198–201. https://doi.org/10.1002/asi.24109
Leydesdorff, L., Wagner, C. S., & Bornmann, L. (2019). Interdisciplinarity as diversity in citation patterns among journals: Rao-Stirling diversity, relative variety, and the Gini coefficient. *Journal of Informetrics, 13*(1), 255–269. https://doi.org/10.1016/j.joi.2018.12.006
Li, D. (2017). Expertise versus bias in evaluation: Evidence from the NIH. *American Economic Journal: Applied Economics, 9*(2), 60–92. https://doi.org/10.1257/app.20150421
Lin, Z., Yin, Y., Liu, L., & Wang, D. (2023). SciSciNet: A large-scale open data lake for the science of science research. *Scientific Data, 10*(1), 315. https://doi.org/10.1038/s41597-023-02198-9
Liu, L., Jones, B. F., Uzzi, B., & Wang, D. (2023). Data, measurement and empirical methods in the science of science. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-023-01562-4
Ma, A., Mondragón, R. J., & Latora, V. (2015). Anatomy of funded research in science. *Proceedings of the National Academy of Sciences, 112*(48), 14760–14765. https://doi.org/10.1073/pnas.1513651112
March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science, 2*(1), 71–87. https://doi.org/10.1287/orsc.2.1.71
Marx, M., & Fuegi, A. (2020). Reliance on science: Worldwide front-page patent citations to scientific articles. *Strategic Management Journal, 41*(9), 1572–1594. https://doi.org/10.1002/smj.3145
Marx, M., & Fuegi, A. (2022). Reliance on science by inventors: Hybrid extraction of in-text patent-to-article citations. *Journal of Economics & Management Strategy, 31*(2), 369–392. https://doi.org/10.1111/jems.12455
McMahan, P., & McFarland, D. A. (2021). Creative destruction: The structural consequences of scientific curation. *American Sociological Review, 86*(2), 341–376. https://doi.org/10.1177/0003122421996323
Mulligan, K., Lenihan, H., Doran, J., & Roper, S. (2022). Harnessing the science base: Results from a national programme using publicly-funded research centres to reshape firms' R&D. *Research Policy, 51*(4), Article 104468. https://doi.org/10.1016/j.respol.2021.104468
Nelson, R. R. (1985). *An evolutionary theory of economic change*. Harvard University Press.
Newman, M. E. J. (2012). Communities, modules and large-scale structure in networks. *Nature Physics, 8*(1), 25–31. https://doi.org/10.1038/nphys2162
Packalen, M., & Bhattacharya, J. (2020). NIH funding and the pursuit of edge science. *Proceedings of the National Academy of Sciences, 117*(22), 12011–12016. https://doi.org/10.1073/pnas.1910160117
Park, H., Lee, J., & Kim, B.-C. (2015). Project selection in NIH: A natural experiment from ARRA. *Research Policy, 44*(6), 1145–1159. https://doi.org/10.1016/j.respol.2015.03.004
Park, M., Leahey, E., & Funk, R. J. (2023). Papers and patents are becoming less disruptive over time. *Nature, 613*(7942), 138–144. https://doi.org/10.1038/s41586-022-05543-x
Romero, F. (2017). Novelty versus replicability: Virtues and vices in the reward system of science. *Philosophy of Science, 84*(5), 1031–1043. https://doi.org/10.1086/694005
Schumpeter, J. A. (1912). *The theory of economic development*. Harvard University Press.
Schumpeter, J. A. (1939). *Business cycles, 1*. Mcgraw-Hill.
Shi, D., Liu, W., & Wang, Y. (2023). Has China's Young Thousand Talents program been successful in recruiting and nurturing top-caliber scientists? *Science, 379*(6627), 62–65. https://doi.org/10.1126/science.abq1218
Shi, F., & Evans, J. (2023). Surprising combinations of research contents and contexts are related to impact and emerge with scientific outsiders from distant disciplines. *Nature Communications, 14*(1), 1641. https://doi.org/10.1038/s41467-023-36741-4
Shor, B., Bafumi, J., Keele, L., & Park, D. (2007). A bayesian multilevel modeling approach to time-series cross-sectional data. *Political Analysis, 15*(2), 165–181. https://doi.org/10.1093/pan/mpm006
Stephan, P., Veugelers, R., & Wang, J. (2017). Reviewers are blinkered by bibliometrics. *Nature, 544*(7651), 411–412. https://doi.org/10.1038/544411a
Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of The Royal Society Interface, 4*(15), 707–719. https://doi.org/10.1098/rsif.2007.0213
Sun, Y., Livan, G., Ma, A., & Latora, V. (2021). Interdisciplinary researchers attain better long-term funding performance. *Communications Physics, 4*(1), 263. https://doi.org/10.1038/s42005-021-00769-z
Trapido, D. (2015). How novelty in knowledge earns recognition: The role of consistent identities. *Research Policy, 44*(8), 1488–1500. https://doi.org/10.1016/j.respol.2015.05.007
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, C. S., Whetsell, T. A., & Mukherjee, S. (2019). International research collaboration: Novelty, conventionality, and atypicality in knowledge recombination. *Research Policy, 48*(5), 1260–1270. https://doi.org/10.1016/j.respol.2019.01.002
Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics, 10*(2), 365–391. https://doi.org/10.1016/j.joi.2016.02.007
Wang, J. (2013). Citation time window choice for research impact evaluation. *Scientometrics, 94*(3), 851–872. https://doi.org/10.1007/s11192-012-0775-9
Wang, J., Lee, Y.-N., & Walsh, J. P. (2018). Funding model and creativity in science: Competitive versus block funding and status contingency effects. *Research Policy, 47*(6), 1070–1083. https://doi.org/10.1016/j.respol.2018.03.014
Wang, J., Veugelers, R., & Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy, 46*(8), 1416–1436. https://doi.org/10.1016/j.respol.2017.06.006
Wang, K., Shen, Z., Huang, C., Wu, C.-H., Dong, Y., & Kanakia, A. (2020). Microsoft academic graph: When experts are not enough. *Quantitative Science Studies, 1*(1), 396–413. https://doi.org/10.1162/qss_a_00021
Wang, S., Ma, Y., Mao, J., Bai, Y., Liang, Z., & Li, G. (2022). Quantifying scientific breakthroughs by a novel disruption indicator based on knowledge entities. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.24719
Watson, J. D., & Crick, F. H. C. (1953). Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature, 171*(4356), 737–738. https://doi.org/10.1038/171737a0
Wei, C., Li, J., & Shi, D. (2023). Quantifying revolutionary discoveries: Evidence from Nobel prize-winning papers. *Information Processing & Management, 60*(3), Article 103252. https://doi.org/10.1016/j.ipm.2022.103252
Wu, L. F., Wang, D. S., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature, 566*(7744), 378. https://doi.org/10.1038/s41586-019-0941-9. -+.
Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science, 316*(5827), 1036–1039. https://doi.org/10.1126/science.1136099
Xu, F., Wu, L., & Evans, J. (2022). Flat teams drive scientific innovation. *Proceedings of the National Academy of Sciences, 119*(23), Article e2200927119. https://doi.org/10.1073/pnas.2200927119

Yang, A. J., Deng, S., Wang, H., Zhang, Y., & Yang, W. (2023). Disruptive coefficient and 2-step disruptive coefficient: Novel measures for identifying vital nodes in complex networks. *Journal of Informetrics, 17*(3), Article 101411. https://doi.org/10.1016/j.joi.2023.101411

Yang, A. J., Hu, H., Zhao, Y., Wang, H., & Deng, S. (2023). From consolidation to disruption: A novel way to measure the impact of scientists and identify laureates. *Information Processing & Management, 60*(5), Article 103420. https://doi.org/10.1016/j.ipm.2023.103420

Yang, A. J., Wu, L., Zhang, Q., Wang, H., & Deng, S. (2023). The k-step h-index in citation networks at the paper, author, and institution levels. *Journal of Informetrics, 17*(4), Article 101456. https://doi.org/10.1016/j.joi.2023.101456

Yin, Y., Dong, Y., Wang, K., Wang, D., & Jones, B. F. (2022). Public use and public funding of science. *Nature Human Behaviour, 6*(10), 1344–1350. https://doi.org/10.1038/s41562-022-01397-5

Yin, Y., Wang, Y., Evans, J. A., & Wang, D. (2019). Quantifying the dynamics of failure across science, startups and security. *Nature, 575*(7781), 190–194. https://doi.org/10.1038/s41586-019-1725-y

Zeng, A., Shen, Z. S., Zhou, J. L., Wu, J. S., Fan, Y., Wang, Y. G., & Stanley, H. E. (2017). The science of science: From the perspective of complex systems. *Physics Reports, 714*, 1–73. https://doi.org/10.1016/j.physrep.2017.10.001