



Short communication

Complexity and the uncertain impact of novel science[☆]Nicola Melluso^{ID}

Department of Business and Management, LUISS University, Viale Romania, 32, 00197 Roma, Italy

ARTICLE INFO

JEL classification:

I23
O31
O33
O38

Keywords:

Novelty
Complexity
Scientific impact

ABSTRACT

Novel scientific ideas drive progress, yet their outcomes are highly uneven. This paper examines whether problem complexity — the degree of interdependence among elements a study must coordinate — shapes the impact of novel research. I introduce a measure of complexity based on the NK framework and find that novel work addressing more complex problems is substantially more likely to achieve high impact, less likely to receive minimal citations, recognized earlier, and cited more broadly across fields. These patterns can be interpreted through the problem-solving view of science, the topology of search landscapes, and characteristics of the combinations themselves. The findings suggest that variation in the outcomes of novel research depends critically on the structure of the problems to which novelty is applied. Data on the complexity of papers are released at <https://doi.org/10.5281/zenodo.20813592>.

1. Introduction

Novel scientific ideas are widely recognized as a central engine of economic progress (Bush, 1945). A large body of empirical work shows that more novel scientific papers tend, on average, to receive more citations and are disproportionately associated with major breakthroughs (Uzzi et al., 2013; Arts et al., 2025; Azoulay et al., 2011). In this sense, novelty plays a central role in advancing scientific knowledge (Mokyr, 2002). At the same time, the realized impact of novel research is highly uneven. While some novel papers become highly influential, many other novel papers receive little recognition, particularly in the years immediately following publication¹ (Wang et al., 2017).

Prior research has documented a range of factors that help explain heterogeneity in scientific impact, including author reputation and institutional prestige (Merton, 1968; Azoulay et al., 2014), team size and composition (Wuchty et al., 2007; Singh and Fleming, 2010), journal placement and visibility (Wang et al., 2017), and timing relative to field trajectories (Van Raan, 2004). These explanations have substantially improved our understanding of why some contributions diffuse more widely than others. However, they share a common feature: they are largely external to the scientific content itself. They characterize who produces the research, where it appears, and when it is published,

rather than the nature of the scientific problem the research attempts to solve.

What remains poorly understood is how the structure of the scientific problems to which novelty is applied shapes with the impact of novel research. In other words, while novelty has been studied extensively as a property of ideas and their recombination, existing work offers limited insight into how differences in the problems being addressed condition the returns to novelty. This gap is striking given a long-standing view, rooted in the philosophy of science, that science is fundamentally a problem-solving activity (Laudan, 1978; Simon, 1977; Kuhn, 1970). From this perspective, scientific progress is not only about introducing new ideas, but about resolving problems that vary in their underlying structure and difficulty.

In a recombinatorial perspective, problems differ in the degree to which their constituent elements are interdependent. Some problems can be addressed by combining components that operate largely independently, whereas others require the coordinated alignment of multiple interacting elements. Work in combinatorial search and innovation formalizes this distinction through the concept of complexity, modeled as the degree of interdependence among elements in a solution space (Kauffman, 1993; Levinthal, 1997; Fleming, 2001). Taken together,

[☆] I am grateful to Stefano Bianchini, Bernardo Caldarola, Lee Fleming, Adam Jaffe, Erin Leahey, Conor O'Kane, Felix Poege, Emanuele Pugliese, Michael Rose, Valentina Tartari, Giovanni Valentini, and Reinhilde Veugelers for their valuable feedback. In particular, I am deeply thankful to Sam Arts for his dedicated support. I thank audience at the WOEPS24 Conference, DRUID24, ST&I2024 Leuven Summer School, LUISS, UNU-MERIT Maastricht and University of Strasbourg. Project funded from Piano Nazionale Di Riforma E Resilienza (PNRR) NextGeneration EU Progetto ID: SOE2024_0000059 CUP: I83C24001020006.

E-mail address: nmelluso@luiss.it.

¹ This coexistence of high average returns and substantial dispersion has shaped how novelty is understood and evaluated in science (Franzoni and Veugelers, 2025; Alberts et al., 2014; Packalen and Bhattacharya, 2020; Ayoubi et al., 2021; Veugelers et al., 2025; Franzoni et al., 2022). Because novelty is associated with a wide dispersion in realized influence, prior work has frequently interpreted novelty as entailing higher uncertainty and has used it as an empirical proxy for risk (Tripodi et al., 2025; Veugelers et al., 2025).

ther, the problem-solving view of science and the complexity literature suggest that novelty may not have uniform consequences across problems: introducing a novel solution to a simple problem may yield modest advances, while introducing novelty to a complex problem may be associated with resolving constraints that previously limited progress and with greater scientific value.²³ Despite these theoretical insights, we lack empirical evidence on whether and how problem complexity shapes the impact of scientific research, and in particular whether novel work is more influential when it tackles more complex problems.

Addressing this gap requires a setting in which both novelty and problem complexity are observable. Biomedical research is well suited for this purpose because papers combine identifiable chemical and biological substances—for example, compounds used to inhibit a disease pathway. These substances are the discrete components researchers use and report in their studies, making it possible to observe new combinations and to characterize how interdependent the components are.⁴ To measure novelty, I follow the well-established principle that novelty arises from new combinations of knowledge elements (Weitzman, 1998; Nelson and Winter, 1985; Schumpeter, 1964). Here, new combinations of substances capture the introduction of a novel idea. To measure complexity, I leverage the NK framework, a widely used model of combinatorial search (Kauffman, 1993; Sorenson et al., 2006; Fleming and Sorenson, 2001; Ganco, 2017; Zhang and Yang, 2022; Frenken, 2006; Levinthal, 1997; Billinger et al., 2014). In this framework, the substances used in a study represent the elements (N), and their documented interactions represent the degree of interdependence (K), where interdependence reflects the extent to which a change in one element requires coordinated changes in others. This approach allows me to quantify the complexity of the problem a paper addresses through the N and K values of the substances it employs. When a paper successfully uses a particular set of interdependent substances, it indicates — among published papers — that researchers have addressed a problem with the specific level of complexity defined by those interactions. Thus, a paper that introduces a new combination of substances contributes novelty to a problem whose complexity can be estimated through the interdependences among the substances involved.⁵

In this work, I analyze 9.6 million biomedical papers published between 1950 and 2023, measuring each paper's novelty and complexity, and examine how these dimensions relate to three established aspects of scientific impact: citation magnitude, the timing of recognition, and the breadth of diffusion across fields (Wang et al., 2017). Each of these aspects corresponds to a pattern that prior research has documented for novel research, and I examine whether problem

complexity conditions it. First, because novel papers are disproportionately represented among both the most highly cited papers and those that receive minimal recognition, I examine whether complexity shapes which novel papers reach the upper rather than the lower tail of the citation distribution. Second, because novel papers has shown to face delayed recognition, I examine whether complexity is associated with earlier recognition. Third, because novel papers attract citations from a wider range of fields beyond their original domain, I examine whether complexity broadens or narrows this diffusion.⁶ Among published papers with comparable levels of novelty, those addressing more complex problems are associated with higher citation impact, earlier recognition, and broader field diffusion, while the association between novelty and minimal recognition differs across levels of complexity. These patterns suggest that the uneven outcomes associated with novelty are not uniform, but depend on the structure of the problems to which novelty is applied.

This paper makes three contributions. First, it introduces and validates an empirical measure of problem complexity in science, based on the interdependence of substances used in biomedical research. Second, it documents how problem complexity relates to citation impact. Third, it shows that the association between novelty and citation impact varies with problem complexity, which helps explain why equally novel papers can have very different citation outcomes. The results are conditional on publication. They therefore describe how successful novel solutions to problems of a given complexity are recognized, not how often such attempts succeed or whether pursuing novelty on complex problems is less uncertain *ex ante*. I release the paper-level complexity data as open access at <https://doi.org/10.5281/zenodo.20813592>.

The paper proceeds abductively. Section 2 presents the empirical setting, and Section 3 reports the main results. Section 4 then develops an interpretation of the documented patterns by integrating complementary theoretical perspectives with descriptive evidence. Section 5 presents robustness checks, and Section 6 discusses limitations and directions for future research.

2. Empirical setting

2.1. Data and sources

I employ the sample of 25,097,834 journal and conference papers published between 1846 and 2023 indexed in both OpenAlex and PubMed. The sample derives from Arts et al. (2025), who provided a cleaned list of papers from the OpenAlex (January 2024 snapshot) by retaining only papers with an English title and abstract, complete publisher information, impact factors for their publication year, non-empty titles, at least one author, and no duplicated titles or abstracts.

2.2. Substances

My analysis focuses on the *substances* discussed in biomedical research papers. A substance represents any chemical compound, drug, enzyme, or biomolecule used within the experiment conducted in a study. These substances represent essential components of a research paper, as they constitute the specific elements that researchers employ within their experiments. By studying these substances, I can effectively characterize research papers based on their key experimental components (Foster et al., 2015).

PubMed indexes 251,758 unique substances across 12,371,478 papers.⁷ The National Library of Medicine (NLM) manages this content

² Specifically, Kuhn (1970) emphasizes that scientific progress consists in resolving “persistent puzzles” that resist existing approaches.

³ For example, developing a treatment for a routine bacterial infection often involves combining drug components that act largely independently, whereas addressing a multidrug-resistant infection requires coordinating multiple interdependent mechanisms for the therapy to be effective. In the latter case, the challenge lies not only in introducing something new but in solving a problem in which several elements must jointly align.

⁴ Biomedical research and its components — such as diseases, drugs, chemical compounds, and proteins — have been the focus of numerous studies. See, for instance, Azoulay et al. (2019b), Aslan et al. (2024), Hill and Stein (2025a,b), Foster et al. (2015).

⁵ I refer to *problem* complexity because the *problem* is the specific task of making a given set of substances work together in a study. Publication implies that this task has been achieved for that particular combination. A *problem* is more complex when the substances are more interdependent, so that the combination is harder to make work. In this study, problem complexity refers specifically to the coordination demands imposed by substance interdependence, which is distinct from the broader clinical challenge of treating a disease or the procedural difficulty of experimentation.

⁶ Resolving a problem whose elements are strongly interdependent may also relax constraints that limit progress in other fields, so the impact of complex papers could diffuse broadly rather than remain confined to its original field. As this study is the first to quantify complexity in science, neither this possibility nor whether novelty applied to a complex problem diffuses more widely than novelty applied to a simpler one has yet been examined.

through a controlled indexing process. Trained indexers examine each paper's title and abstract (and full text when available), and assign both MeSH (Medical Subject Headings) keywords and substances.⁸ Figure A1 shows two illustrative examples of papers in PubMed with their corresponding indexed substances. Table A1 presents the most common substances in the sample.

2.3. Measures

2.3.1. Complexity

I define the complexity of a paper as a function of the interdependence of its substances, where interdependence captures the extent to which the behavior or properties of one substance depend on and influence others (Barabási et al., 2011). The underlying assumption is that one substance cannot function properly without another (e.g., like insulin needs glucose to regulate blood sugar). A paper integrating highly interdependent substances is inherently more complex because the more interdependent the substances are, the higher the number of factors that have to be controlled for in conducting the experiments embedded in the study. This conceptualization of complexity aligns with the NK model's definition, widely applied in prior innovation literature (Levinthal, 1997; Fleming and Sorenson, 2001; Sorenson et al., 2006; Ganco, 2017).

To quantify paper complexity, I use Ganco's (2017) NK model approach. Paper complexity depends on two parameters: the number of substances (N) and their interdependence (K). I first calculate the interdependence K_i of each substance i as:

$$K_i \equiv \sum_{j \in P-i} \frac{1 + \text{count of papers with substance } i \text{ and } j}{1 + \text{count of papers with substance } i} \quad (1)$$

where j represents all other substances in paper p except i .⁹ Following Ganco (2017), I exclude the focal paper from this calculation, I exclude substances that appear only once and papers containing only one substance to prevent inflation of K .^{10,11} I then calculate the overall interdependence K_p of paper p as the average interdependence across all substances:

$$\text{Interdependence of paper } p \equiv K_p \equiv \frac{\sum_{i \in p} K_i}{\text{count of substances in paper } p} \quad (2)$$

Finally, following prior approaches (Fleming and Sorenson, 2001; Sorenson et al., 2006; Ganco, 2017), I define the complexity of paper p as the ratio of K and N .

⁷ To collect substances from each paper in PubMed I downloaded the PubMed citation records in XML and extract the attribute "Chemical" for each paper. Detailed documentation about this attribute available at this link: <https://dtd.nlm.nih.gov/ncbi/pubmed/doc/out/240101/el-Chemical.html> (last access: 2025/01/14).

⁸ In the MeSH tree structure, substances correspond specifically to Category D Descriptors, which include classes of chemicals (e.g., antibiotics, proteins), and Supplementary Concept Records (SCRs), which catalog individual compounds and drugs (e.g., paracetamol, uric acid). Unlike general MeSH terms, substances are linked to external databases via unique identifiers, such as CAS numbers (Chemical Abstracts Service), EC numbers (Enzyme Commission) for enzymes, or UNIs (Unique Ingredient Identifiers) from the Food and Drug Administration (FDA).

⁹ I add 1 to both denominator and numerator to avoid undefined measures when a substance appears for the first time. This adjustment still keep K in the correct scaling of $[0, N-1]$.

¹⁰ Because a substantial share of papers (20.90%) contain only one substance and are thus excluded, I conduct a robustness check, recalculating complexity including these papers and rerunning the analysis to ensure the validity of the results. Results shown and discussed in Appendix B and Table B6.

¹¹ I start counting from the first paper in the dataset (published in 1836). Papers before 1950 are excluded from the main analysis due to sparse coverage, but I include them when counting substance reuse.

To illustrate how the measure works, consider two papers dealing with the same infection (*Raoultella planticola*) but differing in the complexity of the problem they are addressing. Paper 1 deals with a non-multidrug-resistant *Raoultella planticola* infection, while Paper 2 investigates an infection caused by a multidrug-resistant strain. These differences in problem complexity are reflected in the distinct sets of substances employed, as shown in Table 1 and Figure A1. Paper 1 utilizes 4 substances, including piperacillin-tazobactam, a well-established treatment for common bacterial infections. The interdependence reflected in tazobactam's K value (1.757) arises because its effectiveness is linked to piperacillin's presence, a relationship well understood in standard clinical settings (Falagas et al., 2013). In contrast, Paper 2 confronts a more complex problem—a nosocomial pneumonia caused by multidrug-resistant *Raoultella planticola* carrying the blaNDM-1 gene. This multidrug resistance introduces significant complexity because standard treatments are less effective (Livermore, 2011). Paper 2 employs 6 substances, including tigecycline and levofloxacin, reflecting the need for a more elaborate strategy to combat this resistant strain. The higher interdependence values, exemplified by tigecycline ($K=2.124$) and beta-lactamase NDM-1 ($K=1.803$), signify the crucial interactions necessary to overcome the bacterial resistance mechanisms¹² (Livermore, 2011). Therefore, Paper 2, by addressing the inherently more complex problem of multidrug resistance, naturally exhibits a higher ratio of K and N and thus, higher complexity.

To validate the measure, I leverage drug-drug interactions (DDIs) from DrugBank, as they represent documented biological interdependencies where substances affect each other's functionality. While Ganco (2017) validates the effectiveness of the NK model in capturing problem complexity, a key limitation remains: it is unclear whether substance co-occurrences actually reflect true biological interactions. I address this limitation by using DDIs from DrugBank as a ground truth measure of biological interdependence.¹³ After matching drug names from DrugBank to substances, I identify 227,839 substance pairs that co-occur at least once in the same paper by the end of 2023.¹⁴ I then estimate a logit model predicting whether a substance pair represents a documented DDI, using the frequency of co-occurrence as the key predictor while controlling for the individual occurrence frequencies of both substances. While Table A4 shows the most common drug combinations, the results in Table A3 reveal that co-occurrence frequency positively and significantly predicts the likelihood of biological interaction, with an area under the curve (AUC) of 0.6011.¹⁵ This validation suggests that substance co-occurrences in papers capture biological interdependencies, supporting the measure's validity.

2.3.2. Novelty

I define the novelty of a paper as the first appearance of a new substance combination. Following prior studies that measure novelty as a new combination of existing knowledge (Azoulay et al., 2011; Wang et al., 2017; Arts et al., 2025; Lee et al., 2015; Uzzi et al., 2013; Shibayama et al., 2021), I quantify a paper's novelty as the sum of

¹² Notably, treating multidrug-resistant infections demands careful coordination and control of multiple factors during experiments due to the unpredictable nature of resistance and the interplay of drugs needed to bypass these mechanisms.

¹³ For example, given the drugs *Isotretinoin* and *Doxycycline*, DrugBank explicitly documents that "the risk or severity of pseudotumor cerebri can be increased when *Isotretinoin* is combined with *Doxycycline*".

¹⁴ I perform a lowercase string match of DrugBank drug names with the substances in my dataset, resulting in 4479 matches. Drugs represent approximately 3.03% of the unique substances indexed in PubMed for papers in my corpus. Focusing solely on drugs would significantly limit the sample size; therefore, I use them exclusively for validation and robustness purposes (see Section 5 and Appendix B).

¹⁵ AUC ranges from 0.5 (no predictive power) to 1 (perfect classification).

Table 1
Examples of paper complexity.

Paper 1: Cholangitis with septic shock caused by <i>Raoultella planticola</i> (Yokota et al., 2012)		
Substance	Interdependence (K_i)	Complexity
Anti-bacterial agents	0.005	0.168
Piperacillin/tazobactam	1.757	
Penicillanic acid	0.587	
Piperacillin	0.341	
Paper 2: Combination of Tigecycline and Levofloxacin for Successful Treatment of Nosocomial Pneumonia Caused by New Delhi Metallo- β -Lactamase-1-Producing <i>Raoultella planticola</i> (Pan et al., 2017)		
Substance	Interdependence (K_i)	Complexity
Anti-bacterial agents	0.072	0.187
Carbapenems	1.170	
Levofloxacin	0.622	
Tigecycline	2.124	
Beta-lactamase NDM-1	1.803	
Minocycline	0.937	

Notes: Illustrative example of two papers studying *raoultella planticola* infection but with different substances. K_i represents the interdependence for each substance. Paper complexity is calculated as the average of substance interdependence (K_p) divided by the number of substances (N).

its new substance combinations. For example, I identify the first paper¹⁶ introducing the combination of “cyclosporine” and “prednisone”, which revolutionized organ transplantation, or the first paper¹⁷ introducing the combination of “indinavir” and “zidovudine” which, together with “lamivudine”, demonstrated the efficacy of a novel drug regimen for treating HIV.

To validate the measure, I follow the case-control approach proposed by Arts et al. (2025). This approach tests whether a novelty metric can distinguish Nobel Prize-winning papers from a matched control sample of papers published in the same journal, year, and subfield. Using a sample of 112 Nobel Prize papers in Chemistry and Medicine published between 1950 and 2007 (Li et al., 2019) and 112 matched control papers, I employ a logit model with a binary indicator for Nobel Prize papers as the outcome variable, controlling for the number of substances in each paper, the number of references, and subfield and year fixed effects. Table A2 shows that the number of new substance combinations significantly distinguishes Nobel Prize papers from their matched controls, achieving an AUC of 0.7867, thus supporting the validity of the novelty measure.

2.3.3. Other covariates

To characterize each paper in the sample, I collect several additional measures from OpenAlex. First, I measure the *Impact Factor* of the paper’s publication venue (journal or conference), calculated as the mean number of citations received in the publication year for papers published in the preceding two years. Second, I collect the *Number of References* listed in each paper. Third, I obtain the *Number of Authors* of each paper. Fourth, I calculate the *Average Author Age* as the mean number of years between each author’s first publication and the focal paper. Fifth, I measure the *Average H-index* as the mean of all authors’ H-indices at the time of publication. Finally, after obtaining the primary *Field* and *Subfield*, I measure *Interdisciplinarity* as the count of different fields to which the paper is assigned.¹⁸

¹⁶ Starzl, T. E., Klintmalm, G. B., Porter, K. A., Iwatsuki, S., & Schröter, G. P. (1981). Liver transplantation with use of cyclosporin A and prednisone. *New England Journal of Medicine*, 305(5), 266.

¹⁷ Gulick, R. M., et al. (1997). “Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy”. *New England Journal of Medicine*, 337(11), 734–739.

¹⁸ OpenAlex assigns each paper to three topics using a machine learning model that analyzes its title, abstract, and co-citation patterns. Each topic receives a relevance score, with the highest-scoring one designated as the “primary” topic. These topics fall within a four-level hierarchy: 4 domains, 26 fields, 252 subfields, and 4516 topics.

Table 2
Summary statistics.

	Mean	Std	Min	p25	p50	p75	Max
Complexity	0.089	0.080	0.000	0.027	0.069	0.126	0.750
Novelty	1.717	5.253	0.000	0.000	0.000	1.000	786.000
N. of Substances	4.741	2.912	2.000	3.000	4.000	6.000	50.000
N. of Authors	5.568	6.736	1.000	3.000	5.000	7.000	7942.000
Avg Author Age	16.057	9.551	0.000	9.667	14.667	20.750	237.000
Avg H-index	25.917	34.819	0.000	7.500	16.833	32.500	4541.000
Impact Factor	4.187	4.296	0.000	2.038	3.153	4.802	363.134
Interdisciplinarity	1.857	0.691	1.000	1.000	2.000	2.000	3.000
N. of References	33.088	30.549	0.000	15.000	28.000	44.000	4129.000

Notes: Papers published between 1950 and 2023 with defined complexity (n=9,627,624).

2.3.4. Summary descriptive

The final sample consists of 9,627,624 papers published between 1950 and 2023 with defined complexity, with pre-1950 papers serving as a baseline due to partial substance data coverage. To analyze novelty and complexity systematically, I characterize each paper into discrete categories based on subfield-year distributions. For *Complexity*, each paper is placed into one of twenty ventiles (5-percentile intervals). For *Novelty*, papers below the 75th percentile are grouped into a single category, and the remainder are assigned to 5-percentile intervals.¹⁹ Subsequently, I collapse novelty and complexity ventiles into Low (below the 75th percentile within subfield-year) and High (at or above the 75th percentile) categories,²⁰ assigning each paper to exactly one of four mutually exclusive groups: “Low Novelty–Low Complexity”, “Low Novelty–High Complexity”, “High Novelty–Low Complexity”, and “High Novelty–High Complexity”. Table A6 illustrates 10 examples from Nobel Prize papers in Chemistry, showing their substances and their percentile categorization of *Novelty* and *Complexity*.²¹

The summary descriptive in Table 2 shows that *Novelty* exhibits a skewed distribution — consistent with prior findings (Uzzi et al., 2013;

¹⁹ Approximately 70% of papers introduce no new combinations of substances. These papers are assigned to a single baseline category (p0–p70), and in cases of overlapping values within a subfield-year, the middle interval is used (Arts et al., 2025).

²⁰ The 75th percentile threshold is chosen due to the skewed distribution of novelty. This threshold ensures comparability across categories.

²¹ The median Nobel Prize paper in Chemistry falls within the 80th–85th percentile of both *Novelty* and *Complexity*. Table A6 also includes papers with relatively lower *Novelty* and *Complexity* to provide illustrative examples. On average, Nobel Prize-winning papers address more complex problems and introduce more novel combinations than typical papers in the sample.

Table 3
Scientific impact.

	N. Citations			Prob. of “Flop” (Citations ≤ p5)			Prob. of “Hit” (Citations > p95)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Complexity	0.392*** (0.014)		0.079*** (0.016)	−0.200*** (0.025)		0.085*** (0.027)	0.709*** (0.028)		0.153*** (0.033)
Novelty		0.003* (0.002)	−0.074*** (0.002)		0.050*** (0.004)	0.148*** (0.005)		0.020*** (0.003)	−0.111*** (0.005)
Complexity × Novelty			0.690*** (0.016)			−1.037*** (0.042)			1.178*** (0.031)
Pseudo R-squared	0.289	0.289	0.290	0.140	0.141	0.141	0.150	0.150	0.151

Notes: Table shows the coefficients from Poisson quasi-maximum likelihood models with 15-years citation as outcome (Model 1–3), logit models with a binary variable of whether the paper has number of citations below or equal the low 5th percentile of papers published in the same year and subfield (Model 4–6) and a binary variable of whether the paper has number of citations above the top 5th percentile of papers published in the same year and subfield. All models include controls for number of substances, number of references, interdisciplinarity, number of authors, average author age, and average h-index, with fixed effects for year and subfield. All variables, except for *Complexity*, are log-transformed (after adding 1 for zero values). Sample includes papers from 1950–2008 (n=4,235,254). Robust standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.10.

Arts et al., 2025; Wang et al., 2017; Foster et al., 2015) — whereas *Novelty* and *Complexity* themselves are not correlated (Table A5). This lack of correlation suggests that introducing new combinations of substances is not systematically tied to the number of interdependent components. As summarized in Table A7, the “High Novelty–High Complexity” category accounts for 3.5% of the papers, a pattern that is consistent with the intuition that novel contributions are harder to achieve when the underlying problem involves stronger interdependences (Kauffman, 1993; Ganco, 2017).

3. Results

3.1. Scientific impact and tail outcomes

I begin by examining how novelty and complexity jointly relate to scientific impact and to the likelihood that papers fall into the upper or lower tails of the citation distribution. To this end, I estimate three sets of models. First, I use a Poisson quasi-maximum likelihood²² regression to examine how *Novelty* and *Complexity* affect total citations accumulated over 15 years. Second, I implement a logit model to predict a paper’s likelihood of falling into the lower 5% (“flop”). Third, I run a logit model to predict whether a paper reaches the upper 5% (“hit”). Key explanatory variables include *Novelty*, *Complexity*, and their interaction, with controls for the number of substances, number of references, journal (or conference) impact factor, interdisciplinarity, number of authors, average authors’ age and average authors’ H-index. All models include year and subfield fixed effects. To avoid truncation bias, I restrict the sample to papers published by 2008, ensuring a complete 15-year citation window (Wang, 2013).

Table 3 reports regression estimates on the relationship between novelty, complexity, and scientific impact. The results show that both novelty and complexity are systematically associated with impact, but in distinct and interacting ways. Complexity is positively associated with overall citation counts and with a higher likelihood of extreme positive outcomes, while also being associated with a lower probability of minimal impact. Novelty, by contrast, is associated with greater dispersion in outcomes, as novel papers are more likely to appear in both

the upper and lower tails of the citation distribution. The interaction terms indicate that the association between novelty and impact differs across levels of complexity. At low levels of complexity, novel papers are more likely to appear in the lower tail of the citation distribution. At higher levels of complexity, novel papers are less likely to receive minimal attention and more likely to reach the upper tail. Taken together, the estimates indicate that differences in problem complexity account for systematic variation in how novelty is associated with both average impact and extreme citation outcomes.

To investigate the magnitude of the effects, to capture potential nonlinear relationships, and to visualize the results, I estimate two logit models where the main explanatory variables are sets of indicators for each novelty and complexity ventile (calculated within year-subfield) and their interactions. The dependent variables are indicators for minimal and exceptional impact. Fig. 1 presents two heatmaps showing the predicted probabilities for each novelty-complexity combination: Panel A shows the probability of belonging to the lower tail (“flop”), while Panel B shows the probability of belonging to the upper tail (“hit”).

The two heatmaps suggest that higher novelty increases the probability of both extreme outcomes, while complexity polarizes these probabilities in opposing directions. Panel A shows that at high novelty levels (p90–p95), the probability of minimal impact reaches its maximum when complexity is low, concentrating in the upper-left region of the heatmap. For these highly novel papers, the probability of minimal impact decreases from 10.97% at the lowest complexity level (p0–p5) to 4.61% at the highest complexity level (p95–p100)—a 2.37-fold reduction. Panel B shows the inverse relationship: high complexity increases novelty’s association with top citations, with the probability of reaching the upper tail rising from 3.38% to 8.76% for highly novel papers (p90–p95) as complexity increases from p0–p5 to p95–p100—a 2.59-fold increase. Notably, the heatmaps show no region with simultaneously high probabilities of both minimal and top citations. Instead, the peak in the probability of minimal impact occurs where the probability of top citations is comparatively low, and vice versa.

3.2. Delayed recognition

Prior research has established that novel papers often face initial resistance, manifesting in a *delayed recognition* pattern where impact emerges only after several years (Wang et al., 2017). To examine whether complexity influences this temporal pattern, I estimate separate logit models for each of the 15 years after publication. For each year t , I identify whether a paper ranks in the top 5% by *cumulative* citations among papers published in the same year and subfield. Each yearly model includes $\log(1 + \text{citations}_{t-1})$ to account for citation momentum, and controls for the number of substances, number of references, interdisciplinarity, number of authors, average author age,

²² Following Azoulay et al. (2019a), I employ Poisson quasi-maximum likelihood, which yields consistent estimates under relatively weak assumptions and remains robust to overdispersion, as long as the conditional mean is correctly specified (Wooldridge, 1999; Bertrand et al., 2004). This makes Poisson quasi-maximum likelihood suitable for the skewed nature of citation data. In the robustness I also use a Negative Binomial and an OLS model; results are reported in Appendix B (Table B5, Column (1)) and are consistent with the main findings.

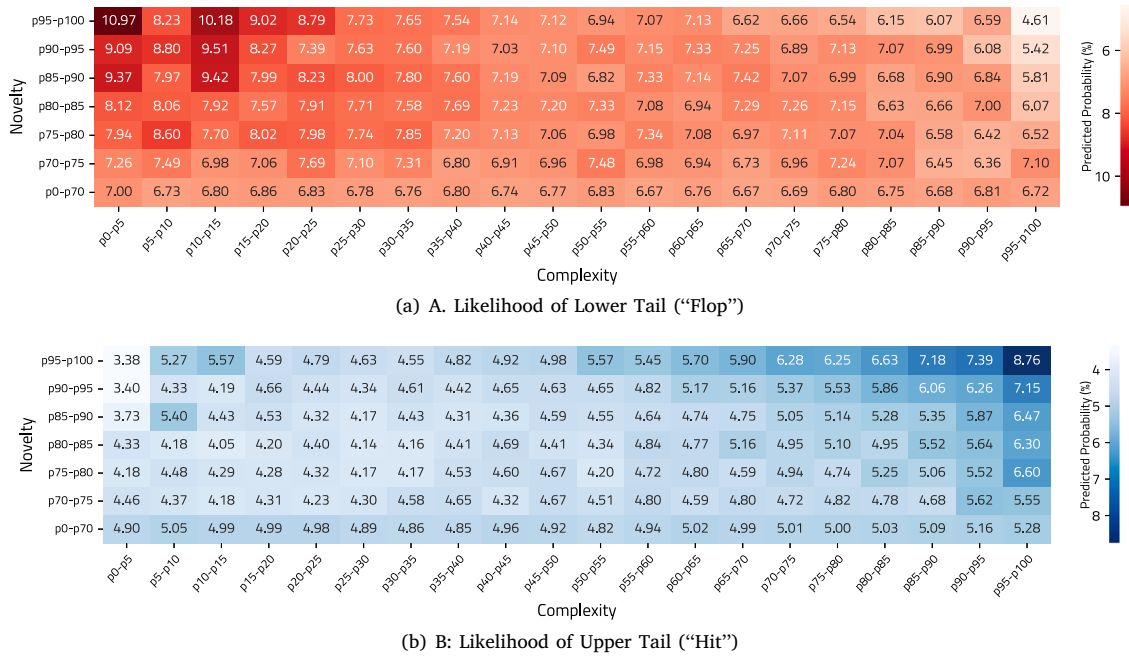


Fig. 1. Predicted Likelihood.

Notes: Heatmaps of predicted probabilities from logit models interacting ventile indicators of *Novelty* and *Complexity*. Panel A shows the probability of falling in the bottom 5% of citations within subfield-year. Panel B shows the probability of reaching the top 5% of citations within subfield-year. Reference categories are p45–p50 for *Complexity* and p0–p70 for *Novelty*. All papers below p70 for novelty form one category, as 70% of papers have no new combinations. All models include controls for number of substances, number of references, interdisciplinarity, number of authors, average author age, and average h-index, with fixed effects for year and subfield. All variables, except for categorical indicators, are log-transformed (after adding 1 for zero values). Sample includes papers from 1950–2008 ($n=4,235,254$).

average h-index, and the average age of substances, along with year and subfield fixed effects. I estimate three sets of models to facilitate clear visualization and interpretation of effect sizes. The first set uses a binary indicator for “High Novelty”, the second for “High Complexity”, and the third employs four categorical combinations of novelty and complexity (with “Low Novelty-Low Complexity” as the reference category). For each model, I calculate the predicted probabilities across the respective indicators.

Fig. 2 reveals that high-novelty, high-complexity papers do not exhibit the delayed-recognition pattern seen for high novelty alone (Panel A). While Panel A shows the expected delayed recognition for novel papers, Panel B illustrates that papers with both high novelty and complexity maintain substantially higher impact throughout all years. After 15 years, these papers reach a 6.31% probability of top citations — the highest among all categories — while those with high novelty but low complexity show the lowest probability at 4.86%. This pattern persists across time: papers with high novelty and complexity have every year, on average, a 31.86%²³ higher probability of achieving top 5% citations compared to novel papers with low complexity, with the advantage remaining stable from early years (32.02%²⁴ higher after 5 years) to later periods (29.75% after 15 years). These temporal

patterns suggest that novel papers avoid delayed recognition when coupled with high complexity.

3.3. Interdisciplinary impact

Beyond the magnitude and timing of impact, I examine its breadth. Prior research has established that novel papers generate broader interdisciplinary impact, attracting citations from a wider range of fields beyond their original domain (Wang et al., 2017). Whether complexity extends or restrains this diffusion is less evident: resolving a problem whose elements are strongly interdependent may relax constraints faced in other fields and broaden a paper’s reach, or it may keep the contribution tied to the field in which the problem arises. To assess complexity’s influence, I estimate a Poisson quasi-maximum likelihood model predicting the number of distinct citing fields (measuring citation breadth) and an OLS model predicting the share of citations from foreign fields.²⁵ Using the four categorical combinations of novelty and complexity (with “Low Novelty-Low Complexity” as the reference category), I estimate predicted values for each group. Both models include the same controls and fixed effects used in the scientific impact analysis, with the addition of the total citation counts.

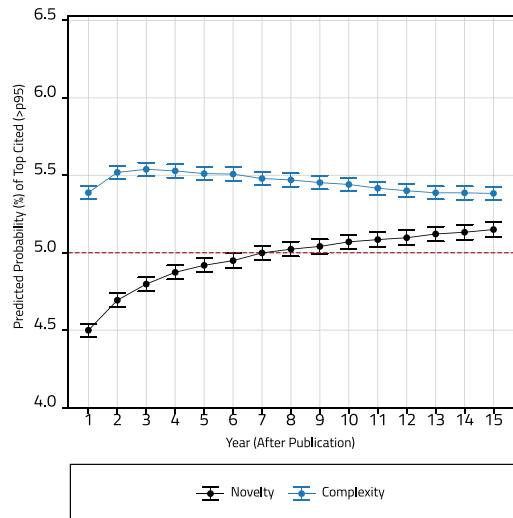
Table 4 reveals that papers with high novelty and complexity achieve the broadest interdisciplinary reach. These papers receive citations from 2.51% more distinct fields compared to novel papers with low complexity. The same pattern emerges for citations from foreign fields, where papers with high novelty and complexity show a 1.44% higher share compared to novel papers with low complexity. These patterns suggest that novel papers extend their interdisciplinary reach when coupled with high complexity.

²³ I calculate this value by computing the percentage difference in predicted probabilities between “High Novelty-High Complexity” and “High Novelty-Low Complexity” for each year, and then averaging across all 15 years.

²⁴ In year 5, the predicted probability of “High Novelty-Low Complexity” is 0.046265 and the predicted probability of “High Novelty-High Complexity” is 0.061080. Thus $\frac{0.061080-0.046265}{0.046265} = 0.3202$.

²⁵ Number of citations a paper receives from fields *other than* its primary assigned field, divided by the paper’s total number of citations.

A: Individual effects of *High Novelty* and *High Complexity*



B: Combined effects of *High Novelty* and *High Complexity*

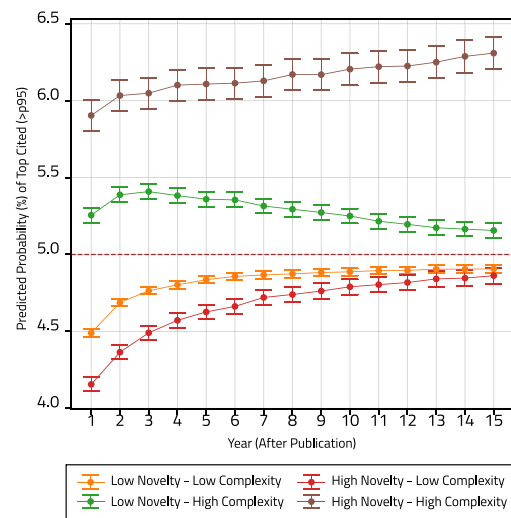


Fig. 2. Predicted Likelihood of “Hit” over Time.

Notes: Predicted probabilities with 95% confidence intervals from logit models estimating the probability of reaching the top 5% (baseline shown as dashed red line) of cumulative citations within subfield-year. Panel A shows results from logit models estimated separately for each year after publication, first with “High Novelty” as the main explanatory variable and then with “High Complexity” (both defined as papers above p75 in their subfield-year). Panel B shows results from logit models estimated separately for each year after publication using categorical combinations of novelty and complexity as the main explanatory variable, with “Low Novelty - Low Complexity” as reference category. All models include controls for the number of citations received in year $t - 1$, number of substances, number of references, interdisciplinarity, number of authors, average author age, and average h-index, with fixed effects for year and subfield. All variables, except for categorical indicators, are log-transformed (after adding 1 for zero values). Sample includes papers from 1950–2008 ($n=4,235,254$).

Table 4
Interdisciplinary impact.

	N. of citing fields (1)	Ratio foreign field citations (2)
Low Novelty - Low Complexity	4.385*** [4.382 – 4.387]	0.684*** [0.684 – 0.684]
Low Novelty - High Complexity	4.181*** [4.175 – 4.186]	0.669*** [0.668 – 0.669]
High Novelty - Low Complexity	4.503*** [4.497 – 4.510]	0.696*** [0.695 – 0.697]
High Novelty - High Complexity	4.616*** [4.604 – 4.628]	0.706*** [0.705 – 0.707]
(Pseudo) R-squared	0.061	0.172
Mean Outcome	4.373	0.684

Notes: Table shows the predicted values and 95% confidence intervals from a Poisson quasi-maximum likelihood model predicting the number of citing fields (Model 1) and an OLS model predicting the ratio of foreign field citations (Model 2). Fields are defined according to OpenAlex’s classification. All models use “Low Novelty - Low Complexity” as reference category and include controls for number of substances, number of references, number of total 15-years citations, number of authors, average author age, and average h-index, with fixed effects for year and subfield. All variables, except for categorical indicators, are log-transformed (after adding 1 for zero values). Sample includes papers from 1950–2008 ($n=4,235,254$). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

4. Interpretation

To interpret the observed results, I draw on complementary theoretical perspectives that suggest why the association between novelty and impact differs across levels of complexity.

First, solving complex problems carries greater scientific weight. In the problem-solving view of science, research communities organize around shared uncertainties, and progress is defined by the resolution of central puzzles and anomalies (Laudan, 1978; Kuhn, 1970). Problems involving many interdependent elements are more difficult to resolve because successful solutions require coordinating multiple components simultaneously (Simon, 1962). When a paper introduces a novel combination that resolves such a problem, it establishes that a demanding configuration can be made to work. Consistent with this view, highly

novel papers that address more complex problems are associated with greater attention and more citations.

Second, the structure of the search landscape shapes the distribution of outcomes for novel combinations at different levels of complexity. In the NK framework, scientists search across landscapes where different combinations of elements yield different levels of performance.²⁶ The degree of interdependence among elements determines the landscape’s topology. Low interdependence produces relatively smooth landscapes, while high interdependence produces rugged landscapes with substantial variation in peak heights (Kauffman, 1993; Levinthal, 1997;

²⁶ In this context, performance refers to how well a combination of substances works together to address the problem—for instance, therapeutic efficacy in treating a disease.

Fleming and Sorenson, 2001). On smooth landscapes, novel combinations tend to yield similar outcomes. On rugged landscapes, novel combinations vary widely—some correspond to low positions and remain unpublished, while others reach high peaks. Because publication requires a working configuration, published novel combinations on rugged landscapes tend to be high-performing and thus attract more attention and citations.

A further interpretation follows from what complexity entails for the nature of novel combinations themselves.

First, new combinations in high-complexity settings are more likely to be unexpected. Substances with high interdependence tend to appear with the same specific substances repeatedly. Substances with low interdependence, by contrast, appear with many different substances. When two high-interdependence substances are combined for the first time, neither was part of the other's repeated set of co-occurrences—otherwise, they would have already appeared together. Their prior co-occurrences are therefore dissimilar, meaning the two substances are distant in the knowledge space. When two low-interdependence substances are combined for the first time, they likely share overlapping co-occurrences, meaning they are closer in the knowledge space. High-complexity papers contain substances with higher average interdependence, so the new combinations they introduce tend to link more distant substances—and are therefore more unexpected. Prior work shows that such unexpected combinations are associated with higher impact because they bridge distant parts of the knowledge space (Foster et al., 2021; Shi and Evans, 2023).²⁷ Consistent with this reasoning, Table A8 shows that high-novelty–high-complexity papers introduce more unexpected combinations than high-novelty–low-complexity papers.²⁸

Second, new combinations in high-complexity settings tend to draw on younger knowledge. Substances with low interdependence can be used more independently—they do not require complementary substances to function effectively, which makes them easier to study and therefore more likely to appear early in the literature. Being among the first to be explored, such substances tend to be older. Substances with high interdependence, by contrast, cannot be used independently—they require complementary substances to function effectively. They can only be studied once those complements have been identified, and their inherent difficulty means they also require more advanced techniques to handle. Such substances therefore tend to enter the literature only as science progresses, making them younger on average. Because high-complexity papers contain substances with higher average interdependence, the knowledge they draw on tends to be younger. Viewed from this perspective, the younger knowledge in high-complexity papers is interdependent with knowledge from the past: it could enter the literature only once prior knowledge had accumulated sufficiently for its complements to be understood. This aligns with Jones (2009)'s burden-of-knowledge argument—in biomedicine, as the knowledge frontier expands, newer knowledge cannot be understood or used independently of knowledge accumulated earlier.²⁹

²⁷ According to Foster et al. (2021), Shi and Evans (2023), unexpected novelty can be impactful from two vantage points: *content* or *context*. In this paper, using substances, I test the content side—combinatorial novelty as the recombination of substances. However, the recombination of the *contexts* in which those substances are used is also a form of unexpectedness that can drive impact—for example, when a set of substances familiar in one subfield is studied in a problem or field whose readership rarely encounters them together. In that view, complex and high-novelty papers may gain impact because their contexts are unexpected, rather than because their substance combinations link distant co-occurrence patterns.

²⁸ I define paper-level unexpectedness as the mean unexpectedness across all new substance pairs introduced by the paper. I compute pair-level unexpectedness for each new pair (A, B) as $1 - \text{Jaccard}(N_A, N_B)$, where N_A and N_B denote the sets of substances that previously co-occurred with A and B , respectively. Higher values indicate that the two substances have more dissimilar prior co-occurrences.

Prior work shows that papers drawing on younger knowledge tend to achieve higher impact, as they address problems that have only recently become tractable once that prior knowledge is in place (Mukherjee et al., 2017). Consistent with this reasoning, Table A8 shows that high-novelty–high-complexity papers draw on younger substances than high-novelty–low-complexity papers.^{30,31}

5. Robustness

To ensure the validity of the findings, I conduct a series of robustness checks. A detailed description of these robustness checks is provided in Appendix B.

First, a potential concern is that co-occurrence-based measures of interdependence may capture only complementary relationships among substances and thus fail to reflect the tradeoffs or negative interactions. To address this concern, I construct an alternative measure of complexity using only drug–drug interactions (DDIs) documented in DrugBank, many of which correspond to clinically adverse or antagonistic effects rather than complementarities. I restrict the sample to papers that use at least two DrugBank-listed drug and compute both novelty and complexity exclusively from the subset of drugs appearing in each paper, with interdependence defined by the documented DDIs among the drugs employed in the paper. I then replicate the main scientific impact analysis on this DDI-restricted sample (Table B1). The resulting estimates are consistent with the main findings. Second, a potential concern is that papers addressing new or highly burdensome diseases may receive more citations for reasons unrelated to novelty and complexity. To address this concern, I assign each paper a focal disease using disease tags from MeSH and PubChem and burden data from the Global Burden of Disease Study. Table B2 reports Poisson models with 15-year citation counts and the novelty–complexity interaction, adding *Disease Recency* (column 1), *Disease Burden* (column 2), and disease fixed effects (column 3). The resulting estimates are consistent with the main findings. Third, because this study is the first to quantify complexity in science, I also compute an alternative NK model-based complexity measure, following Fleming and Sorenson (2001), Sorenson et al. (2006). Table B4 shows that the results remain consistent under this alternative measure, reinforcing the reliability of the findings. Fourth, measuring complexity only at the time of publication may miss how scientific knowledge evolves over time, which could shape a paper's long-run impact (Kedrick et al., 2024). The interdependence of substances may change from year to year as new combinations (or substances themselves) emerge, so the influence on citations may not only depend on the initial complexity but also on how the interdependences change over time. To address this, I re-estimate the yearly post-publication regressions (in the same vein of results from Fig. 2) incorporating an additional control variable that captures how much a paper's complexity changes relative to its initial state. The results in

²⁹ Jones (2009) argues that as the knowledge frontier expands, newer knowledge presupposes an ever-larger stock of prior work. This paper aligns with that view by showing one mechanism through which this dependency manifests in biomedical research: interdependent substances—and the knowledge they embody—enter the literature only once the complementary knowledge required to study them is in place. As the frontier expands, knowledge components may become increasingly interdependent: each new component presupposes that prior components are already understood, and interdependent knowledge therefore tends to be younger.

³⁰ Table A8 uses the full descriptive sample (1950–2023, defined complexity), whereas the main impact regressions use papers published by 2008.

³¹ I define paper-level substance age as the mean age across all substances used in the paper. I measure age for each substance as the number of years since its first observed appearance in the dataset. The first year of observation is 1950. Substances that appear for the first time in the focal year are assigned age zero.

Table B3 show that the impact of novelty and complexity remains stable even when accounting for these shifts. Fifth, I assess whether specific modeling choices influence the findings, including the thresholds used to define flops” and hits” and other analytical specifications. To address this concern, I test alternative model specifications, vary the thresholds for extreme outcomes and include papers that contain only a single substance. The results remain stable across these variations, confirming the robustness of the findings.

6. Conclusion

Novel research is widely recognized as essential to economic progress, yet its outcomes are highly uneven—novel work is disproportionately represented among both breakthrough contributions and papers that receive minimal recognition. Despite extensive work on heterogeneity in research impact, prior literature has largely focused on who produces research, where it appears, and when it is published, offering limited insight into the role of problem complexity. This gap is particularly salient in light of long-standing theoretical perspectives — rooted in the philosophy of science and in formal models of search — that conceptualize scientific progress as problem solving and emphasize that problems differ in their underlying complexity, captured by the degree of interdependence among the elements that must work together. This paper examines whether the impact of novel research varies with the complexity of the problem it addresses.

Focusing on biomedical science, I find that novel work addressing more complex problems is substantially more likely to achieve high impact and less likely to receive minimal recognition, while novel work addressing less complex problems exhibits the opposite pattern. Beyond citation magnitude, novel research tackling complex problems is recognized earlier and cited across a broader range of fields. These patterns are consistent with theoretical perspectives suggesting that breakthroughs occur when complex problems are solved, and by the formal NK framework, where high complexity increases the likelihood of high-performing solutions. Indeed, given that we observe only published papers, such solutions have already proven capable of solving complex problems—and thus exhibit high performance and receive more citations. Additional explanations relate to the nature of the novel combinations and the age of knowledge components intrinsically involved in high-complexity papers: unexpected new combinations and younger knowledge components. These characteristics have been shown to drive greater impact and are more prevalent in complex papers.

However, this work has several limitations. First, the interdependence of substances, while relevant, may not fully represent complexity in science in general. Other factors may contribute to scientific complexity, such as the difficulty of experimental protocols, access to resources, or team dynamics, which are not captured here. Second, the analysis relies on a single operationalization of novelty: combinatorial content novelty, measured as the count of new substance combinations. This choice is grounded both theoretically and empirically—it follows the recombinant view of innovation and distinguishes papers with highly uneven impact (Uzzi et al., 2013; Arts et al., 2025). Yet the novelty construct itself is broader and still expanding; how alternative definitions relate to scientific impact, and to the role of problem complexity, remains to be established. Third, this analysis does not establish causality. Although I control for many potential confounding factors, unobserved elements — such as serendipitous discoveries and broader social dynamics within research communities — may also play a role (Azoulay et al., 2019a; Aslan et al., 2024). Fourth, while these findings are suggestive of broader dynamics within scientific research, they are estimated in the biomedical field. Biomedical research represents a substantial portion of scientific activity, but the specific dynamics explored here may not fully reflect those in other fields.

Importantly, the analysis is based on published research and therefore observes solutions that have already proven effective. Publication

signals that a research team has successfully addressed a problem with its chosen combination of substances. The findings thus characterize the impact of successful novel solutions to complex problems, not the likelihood of achieving such success. Researchers attempting to solve complex problems with novel approaches may face substantial challenges before arriving at a publishable result, and these efforts remain unobserved. Accordingly, the results should not be interpreted as evidence that pursuing novelty on complex problems is less uncertain *ex ante*.

Overall, the findings suggest that heterogeneity in the outcomes of novel research is partly rooted in the structure of the problems being solved. Incorporating problem complexity into studies of novelty may therefore shift the conversation from “how novel” a contribution is to “where novelty is applied” and why. Connecting the observed patterns to upstream stages — idea generation, project continuation, and resource allocation — would help clarify how problem structure shapes which novel approaches are worth pursuing, which succeed, and which ultimately diffuse.

CRediT authorship contribution statement

Nicola Melluso: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.respol.2026.105570>.

Data availability

I have shared the link to my data.

References

- Alberts, B., Kirschner, M.W., Tilghman, S., Varmus, H., 2014. Rescuing US biomedical research from its systemic flaws. *Proc. Natl. Acad. Sci.* 111 (16), 5773–5777.
- Arts, S., Melluso, N., Veugelers, R., 2025. Beyond citations: Measuring novel scientific ideas and their impact in publication text. *Rev. Econ. Stat.*
- Aslan, Y., Yaqub, O., Sampat, B.N., Rotolo, D., 2024. Unexpectedness in medical research. *Res. Policy* 53 (8), 105075.
- Ayoubi, C., Pezzoni, M., Visentin, F., 2021. Does it pay to do novel science? The selectivity patterns in science funding. *Sci. Public Policy* 48 (5), 635–648.
- Azoulay, P., Fons-Rosen, C., Graff-Zivin, J., 2019a. Does science advance one funeral at a time? *Am. Econ. Rev.* 109 (8), 2889–2920.
- Azoulay, P., Graff-Zivin, J., Li, D., Sampat, B., 2019b. Public R&D investments and private-sector patenting: evidence from nih funding rules. *Rev. Econ. Stud.* 86 (1), 117–152.
- Azoulay, P., Graff-Zivin, J., Manso, G., 2011. Incentives and creativity: evidence from the academic life sciences. *RAND J. Econ.* 42 (3), 527–554.
- Azoulay, P., Stuart, T., Wang, Y., 2014. Matthew: Effect or fable? *Manag. Sci.* 60 (1), 92–109.
- Barabási, A.-L., Gulbahce, N., Loscalzo, J., 2011. Network medicine: a network-based approach to human disease. *Nature Rev. Genet.* 12 (1), 56–68.
- Bertrand, M., Dufo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates? *Q. J. Econ.* 119 (1), 249–275.
- Billinger, S., Stieglitz, N., Schumacher, T.R., 2014. Search on rugged landscapes: An experimental study. *Organ. Sci.* 25 (1), 93–108.
- Bush, V., 1945. Science, the endless frontier: A report to the president.
- Falagas, M.E., Tansarli, G.S., Ikawa, K., Vardakas, K.Z., 2013. Clinical outcomes with extended or continuous versus short-term intravenous infusion of carbapenems and piperacillin/tazobactam: a systematic review and meta-analysis. *Clin. Infect. Dis.* 56 (2), 272–282.

- Fleming, L., 2001. Recombinant uncertainty in technological search. *Manag. Sci.* 47 (1), 117–132, Publisher: INFORMS.
- Fleming, L., Sorenson, O., 2001. Technology as a complex adaptive system: evidence from patent data. *Res. Policy* 30 (7), 1019–1039.
- Foster, J.G., Rzhetsky, A., Evans, J.A., 2015. Tradition and innovation in scientists' research strategies. *Am. Sociol. Rev.* 80 (5), 875–908.
- Foster, J.G., Shi, F., Evans, J.A., 2021. Surprise! measuring novelty as expectation violation. <http://dx.doi.org/10.31219/osf.io/2t46f>.
- Franzoni, C., Stephan, P., Veugelers, R., 2022. Funding risky research. *Entrep. Innov. Policy* 1 (1), 103–133.
- Franzoni, C., Veugelers, R., 2025. Funding breakthrough research: An update and research agenda. In: MacGarvie, M., Veugelers, R. (Eds.), *Economics of Science*. University of Chicago Press.
- Frenken, K., 2006. Technological innovation and complexity theory. *Econ. Innov. New Technol.* 15 (2), 137–155.
- Ganco, M., 2017. NK model as a representation of innovative search. *Res. Policy* 46 (10), 1783–1800.
- Hill, R., Stein, C., 2025a. Race to the bottom: Competition and quality in science. *Q. J. Econ.*
- Hill, R., Stein, C., 2025b. Scooped! estimating rewards for priority in science. *J. Political Econ.* 133 (3).
- Jones, B.F., 2009. The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *Rev. Econ. Stud.* 76 (1), 283–317.
- Kauffman, S., 1993. *The Origins of Order*. Oxford University Press.
- Kedrick, K., Levitskaya, E., Funk, R.J., 2024. Conceptual structure and the growth of scientific knowledge. *Nat. Hum. Behav.* 8 (10), 1915–1923.
- Kuhn, T.S., 1970. *The structure of scientific revolutions*, third ed. University of Chicago Press, Chicago.
- Laudan, L., 1978. *Progress and its problems: Towards a theory of scientific growth*, vol. 282, Univ of California Press.
- Lee, Y.-N., Walsh, J.P., Wang, J., 2015. Creativity in scientific teams: Unpacking novelty and impact. *Res. Policy* 44 (3), 684–697.
- Levinthal, D., 1997. Adaptation on rugged landscapes. *Manag. Sci.* 43 (7), 934–950.
- Li, J., Yin, Y., Fortunato, S., Wang, D., 2019. A dataset of publication records for nobel laureates. *Sci. Data* 6 (1), 33.
- Livermore, D.M., 2011. Discovery research: the scientific challenge of finding new antibiotics. *J. Antimicrob. Chemother.* 66 (9), 1941–1944.
- Merton, R.K., 1968. *Social theory and social structure*. Simon and Schuster.
- Mokyr, J., 2002. *The gifts of Athena: Historical origins of the knowledge economy*. Princeton University Press, Princeton and Oxford.
- Mukherjee, S., Romero, D.M., Jones, B., Uzzi, B., 2017. The nearly universal link between the age of past knowledge and tomorrow's breakthroughs in science and technology: The hotspot. *Sci. Adv.* 3 (4), e1601315.
- Nelson, R., Winter, S., 1985. *An evolutionary theory of economic change*. Harvard University Press.
- Packalen, M., Bhattacharya, J., 2020. NIH funding and the pursuit of edge science. *Proc. Natl. Acad. Sci.* 117 (22), 12011–12016.
- Pan, Z., Liu, R., Zhang, P., Zhou, H., Fu, Y., Zhou, J., 2017. Combination of tigecycline and levofloxacin for successful treatment of nosocomial pneumonia caused by new Delhi metallo- β -lactamase-1-producing *raoultella planticola*. *Microb. Drug Resist.* 23 (1), 127–131.
- Schumpeter, J.A., 1964. *Business cycles: A theoretical, historical and statistical analysis of the capitalist process*, 1939. Acessado Em 4.
- Shi, F., Evans, J., 2023. Surprising combinations of research contents and contexts are related to impact and emerge with scientific outsiders from distant disciplines. *Nat. Commun.* 14 (1), 1641.
- Shibayama, S., Yin, D., Matsumoto, K., 2021. Measuring novelty in science with word embedding. *PLoS One* 16 (7), e0254034.
- Simon, H.A., 1962. The architecture of complexity. *Proc. Am. Phil. Soc.* 106 (6), 467–482.
- Simon, H.A., 1977. Scientific discovery and the psychology of problem solving. In: *Models of Discovery: And Other Topics in the Methods of Science*. Springer, pp. 286–303.
- Singh, J., Fleming, L., 2010. Lone inventors as sources of breakthroughs: Myth or reality? *Manag. Sci.* 56 (1), 41–56.
- Sorenson, O., Rivkin, J.W., Fleming, L., 2006. Complexity, networks and knowledge flow. *Res. Policy* 35 (7), 994–1017.
- Tripodi, G., Zheng, X., Qian, Y., Murray, D., Jones, B.F., Ni, C., Wang, D., 2025. Tenure and research trajectories. *Proc. Natl. Acad. Sci.* 122 (30), e2500322122.
- Uzzi, B., Mukherjee, S., Stringer, M., Jones, B., 2013. Atypical combinations and scientific impact. *Science* 342 (6157), 468–472.
- Van Raan, A.F., 2004. Sleeping beauties in science. *Scientometrics* 59 (3), 467–472.
- Veugelers, R., Wang, J., Stephan, P., 2025. Do funding agencies select and enable novel research: evidence from ERC. *Econ. Innov. New Technol.* 1–18.
- Wang, J., 2013. Citation time window choice for research impact evaluation. *Scientometrics* 94 (3), 851–872.
- Wang, J., Veugelers, R., Stephan, P., 2017. Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Res. Policy* 46 (8), 1416–1436.
- Weitzman, M.L., 1998. Recombinant growth*. *Q. J. Econ.* 113 (2), 331–360.
- Wooldridge, J.M., 1999. Quasi-likelihood methods for count data. *Handb. Appl. Econ.* Vol. 2: Microeconomics 321–368.
- Wuchty, S., Jones, B.F., Uzzi, B., 2007. The increasing dominance of teams in production of knowledge. *Science* 316 (5827), 1036–1039.
- Yokota, K., Gomi, H., Miura, Y., Sugano, K., Morisawa, Y., 2012. Cholangitis with septic shock caused by *raoultella planticola*. *J. Med. Microbiol.* 61 (3), 446–449.
- Zhang, Y., Yang, W., 2022. Breakthrough invention and problem complexity: Evidence from a quasi-experiment. *Strat. Manag. J.* 43 (12), 2510–2544.