Scientific Relatedness Constrains Novelty but Strategic International Collaboration Buffers This Penalty in Global SDG Research

Robin Haunschild[1] & Meijun Liu[2]

[1]Max Planck Institute for Solid State Research, Stuttgart, Germany

[2]Institute of Global Public Policy, Fudan University, Shanghai, China

Mingze Zhang[1], Robin Haunschild[2], George Yunxiong Li[3], Meijun Liu[4,*]

1 Department of Information Resources Management, University of Chinese Academy of Sciences, Beijing, China

2 Max Planck Institute for Solid State Research, Stuttgart, Germany

3 ShanghaiTech University, School of Entrepreneurship and Management, Shanghai, China

4 Institute for Global Public Policy, Fudan University, Shanghai, China

Takeaways:

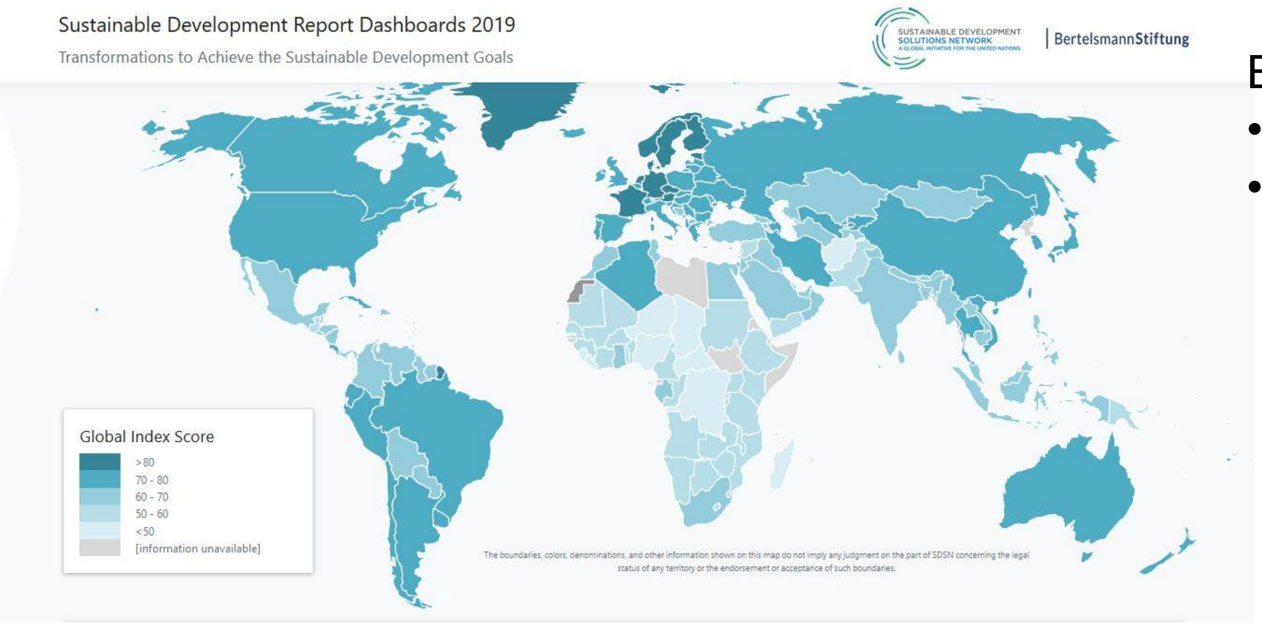
- Analyzing more than 40 million scientific publications from 165 countries ranging from 2000 to 2023, this study uncovers a dual effect of scientific relatedness on countries' future knowledge production.
- Scientific relatedness increases subsequent scientific productivity, but it simultaneously suppresses scientific novelty.
- This negative relatedness-novelty association is more pronounced in the Global South.
- We further evaluate strategies to buffer this novelty penalty, distinguishing between external and internal knowledge-boundary spanning:
  - Externally, enhancing international collaboration and maintaining brokerage roles in global research collaboration networks emerge as effective countermeasures.
  - The moderating role of international collaboration is asymmetric. Increased collaboration intensity helps the Global South, whereas for the Global North, only heterogeneous collaboration (i.e., collaboration with the Global South) proves effective.
  - Internally, strengthening interdisciplinarity of national-level knowledge structures are found to exacerbate, rather than mitigate, this novelty penalty.

Introduction

➢ ***Why study the SDG field?***

- The SDGs are a global agenda adopted by all 193 UN member states in 2015, aiming to be achieved by 2030.
- Noticeable disparity in the SDGs progress and SDG-related knowledge production persists across countries, particularly with the Global South countries lagging behind (Ma et al., 2025; Qi et al., 2024; Sciarra et al., 2021)

How can countries, especially those in the Global South, accelerate their progress toward the SDGs?



Existing approaches are focusing on

- policy implementation
- technology innovation

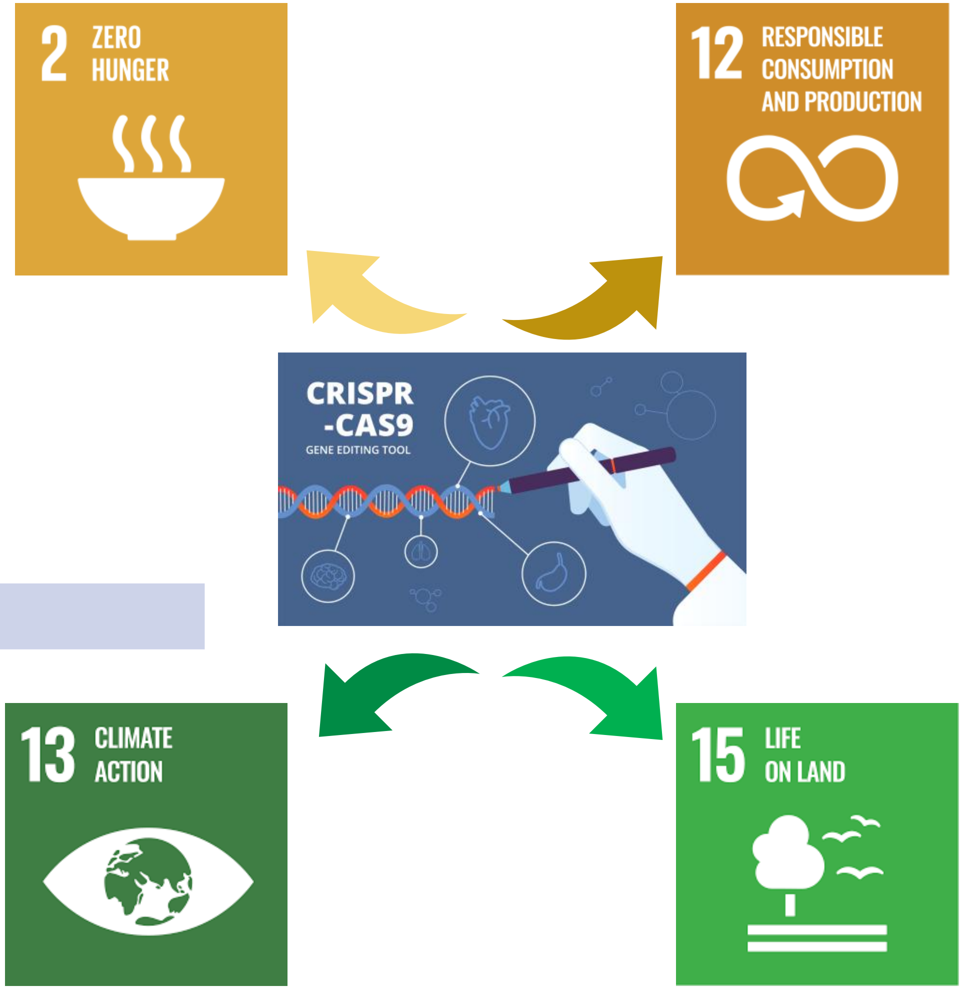
The current discourse has largely **overlooked the pivotal role of scientific novelty** - the generation of breakthrough knowledge - in driving sustainable solutions (Walsh et al., 2020).

Introduction

➢ ***Why study the SDGs from the perspective of science?***

- The generation and integration of novel knowledge into decision-making are critical for promoting SDGs progress (Allen et al., 2018; Wuelser & Pohl, 2016).

- Breakthroughs like CRISPR-Cas9 demonstrate how novel scientific discoveries can directly advance multiple SDGs.



A new research focus is needed. As emerging research highlights, access to frontier science is key to tackling sustainability challenges.

Understanding the national drivers of novel knowledge production is therefore essential for alleviating the global sustainability gap and accelerating progress toward the SDGs.

Introduction

Theoretical Foundation – Path Dependence in Knowledge Production

Evolutionary Perspective on Knowledge Production

A country's knowledge production is **path-dependent**

- Shaped by existing knowledge base, skills, and infrastructures
- Influences future output through **cumulativeness** and **interactive learning**
- *(Dosi, 1982; Nelson & Winter, 1985)*

At the regional level:

- Pre-existing local knowledge sets and activities configure present and future trajectories of science, technology, and industry *(Kogler, 2015)*

Key driver: Relatedness

- Defined as similarities between activities capturing the cognitive dimension of capabilities
- Drives regional branching in products, technologies, and industries *(Boschma, 2017)*

Research Gap – What Remains Unexplored?

This line of inquiry rarely touches a crucial issue:

What role does a national/regional knowledge base play in generating *truly novel* knowledge?

Introduction

Theoretical Extension – Introducing Scientific Relatedness

Adapting Relatedness to the Scientific Domain

Scientific Relatedness: The cognitive proximity between a country's existing knowledge structure and a specific scientific domain

A fundamental tension emerges:

Scenario	Effect
High scientific relatedness	Increases efficiency & productivity (familiar work → lower costs → higher volume)
High scientific relatedness	May cause **cognitive lock-in** → limits exploration of new ideas → dampens novelty

Introduction

The Core Argument – Dual Pressures of Scientific Relatedness

Exploration vs. Exploitation (March, 1991):
Scientific relatedness exerts dual, opposing pressures:

Mechanism	**Exploitation**	**Exploration**
Effect	Enhances efficiency and scientific productivity	Risky – may inhibit novelty due to cognitive lock-in

How to mitigate the "novelty penalty"? Two knowledge-boundary spanning mechanisms:

Mechanism	Key Insight
Internal: Interdisciplinary knowledge base	Mitigates lock-in by enabling atypical combinations from distant domains (Jacobs, 2016)
External: International collaboration	Integrates external knowledge and perspectives; enables cross-border spillovers (Wagner & Leydesdorff, 2005; Guan & Chen, 2012; Liu et al., 2022)

Why SDG research needs both:
Complexity & global interdependence of SDGs demand interdisciplinary + international efforts

Introduction

This study addresses the gaps through two research questions:

RQ1:
How does scientific relatedness of a country's existing knowledge base to an SDG domain shape the quantity (productivity) and scientific novelty of its future knowledge production?

RQ2:
How is the relationship between scientific relatedness and scientific novelty moderated by:
(a) the country's interdisciplinary knowledge structure, and
(b) international collaboration?

Research Hypothesis

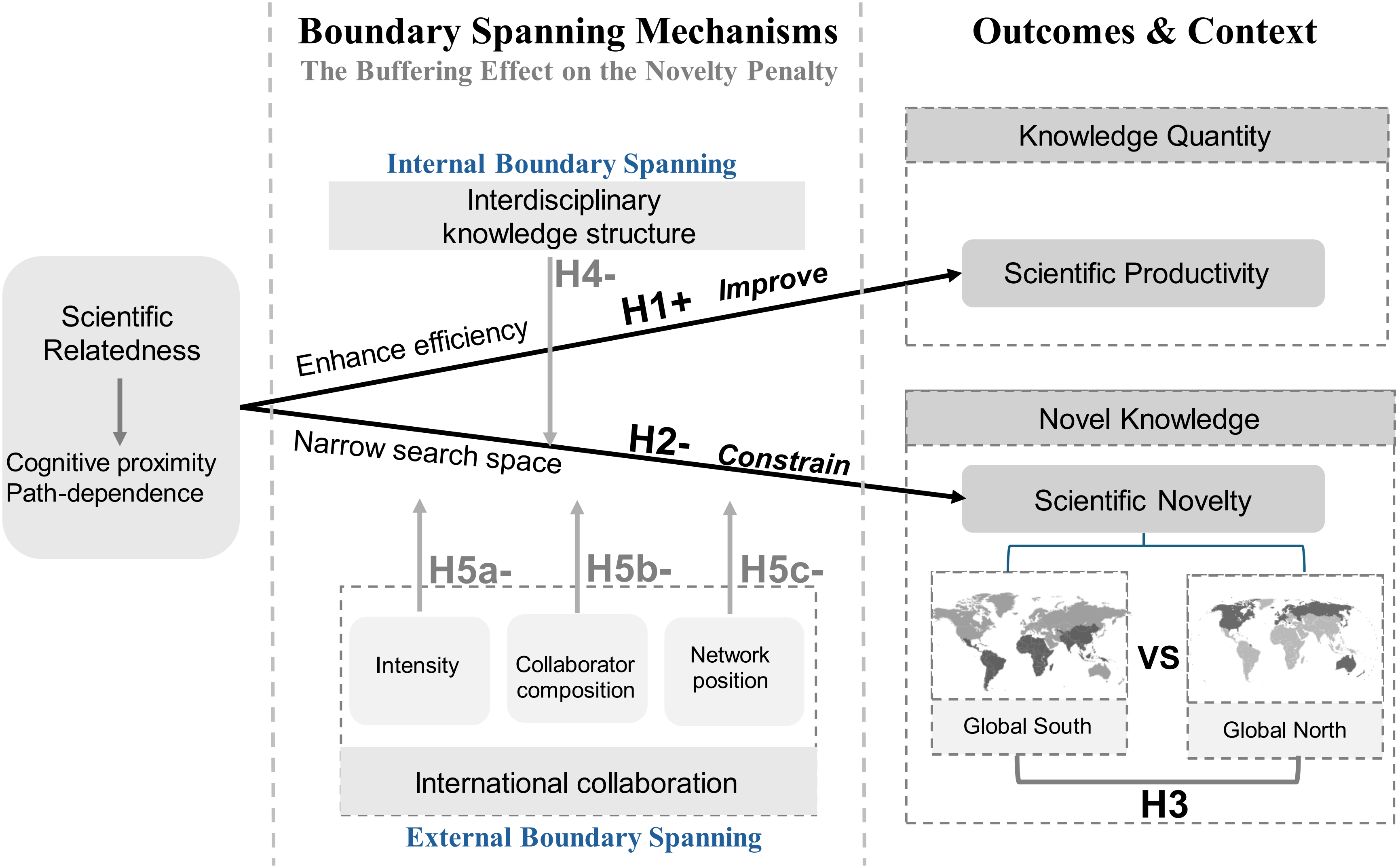


Figure 1. The dual effect of scientific relatedness on scientific productivity and novelty.

Research Hypothesis

MAX PLANCK INSTITUTE FOR SOLID STATE RESEARCH

FUDAN UNIVERSITY

The Dual Effect of Scientific Relatedness – Productivity vs. Novelty Penalty

H1: Positive Effect on Productivity
Scientific relatedness to an SDG domain is **positively associated** with future scientific **productivity** (quantity) in that domain.
Mechanisms: Cognitive proximity + path dependence → lower cost, higher efficiency, increasing returns

H2: Negative Effect on Novelty
Scientific relatedness to an SDG domain is **negatively associated** with future **scientific novelty** in that domain.
Mechanisms: Cognitive lock-in, knowledge redundancy, restricted search space → "recombinant gap"

H3: Stronger Novelty Penalty in the Global South
The **negative** relatedness–novelty relationship is **stronger** for countries in the **Global South** than for those in the **Global North**.
Reasoning: Limited knowledge stocks, lower R&D, less diversified portfolios, peripheral network positions

Research Hypothesis

MAX PLANCK INSTITUTE FOR SOLID STATE RESEARCH

FUDAN UNIVERSITY

Buffering Mechanisms – How to Mitigate the Novelty Penalty

H4: Interdisciplinarity (Internal boundary spanning)

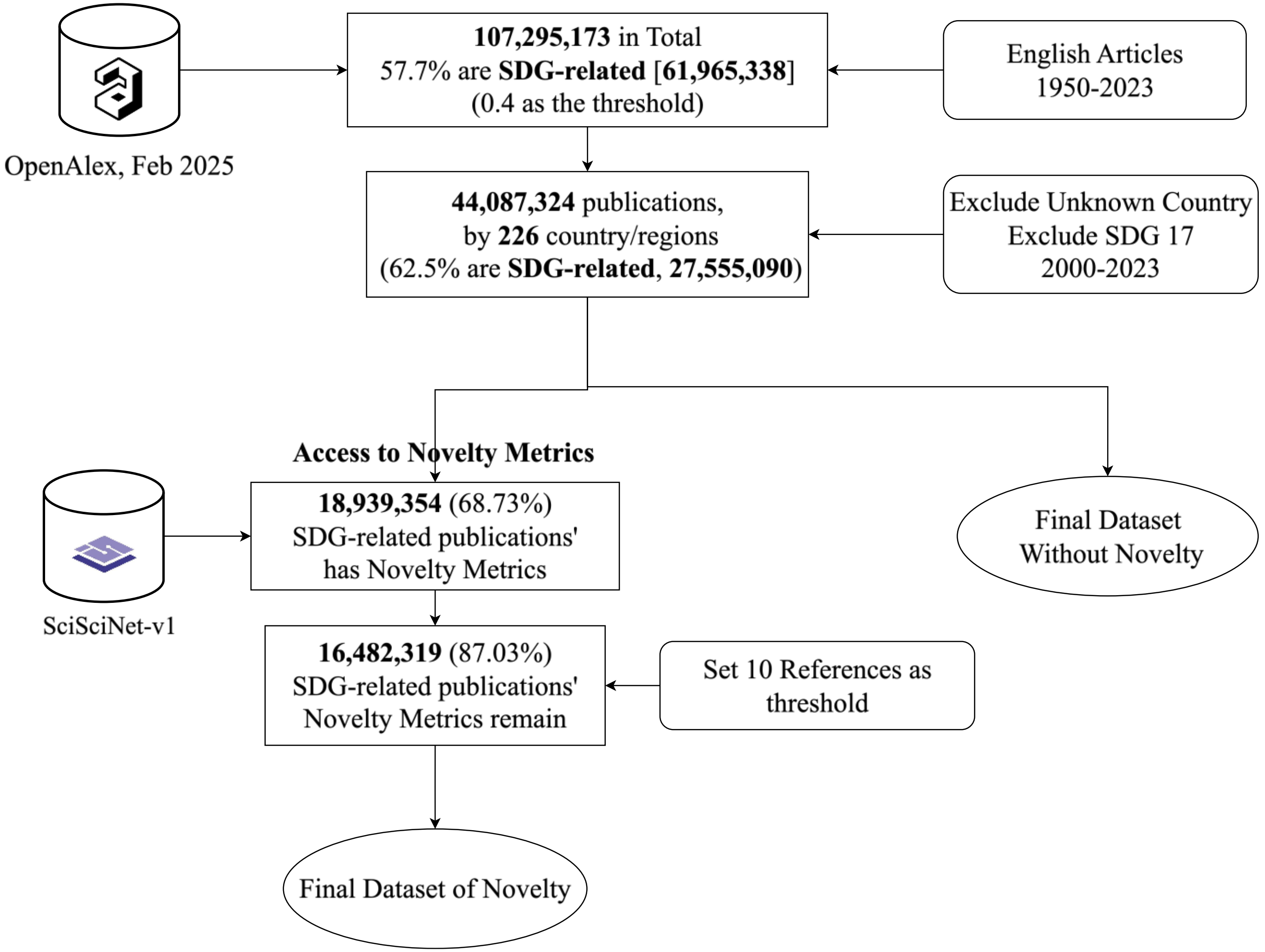
Interdisciplinarity of a country's existing knowledge structure **weakens** the negative relatedness–novelty relationship.

Mechanism: Cognitive flexibility, diverse recombinant material, expanded search scope

H5: International Collaboration (External boundary spanning)

Hypothesis	Statement
H5a	Stronger **intensity** of international collaboration **weakens** the negative relatedness–novelty relationship
H5b	**Heterogeneous collaboration** (Global North–Global South) **weakens** the negative relationship
H5b.1 (GN)	For **Global North** countries: higher share of **South collaboration** mitigates the penalty
H5b.2 (GS)	For **Global South** countries: higher share of **North collaboration** mitigates the penalty
H5c	**Network constraint** (structural disadvantage) **strengthens** the negative relatedness–novelty relationship (i.e., weakens the ability to buffer)

Methodology



- **Data Source**

Main data: **OpenAlex**

SciSciNet

Human Development Report Office

UNDP

SUSTAINABLE DEVELOPMENT REPORT

Global Macro Database

After this procedure, 43,507,884 publications from 165 countries spanning 2000-2023 remained. Among those publications, 62.5% (approximately 27 million publications) are assigned at least one SDG label (hereafter SDG-related publications).

Methodology

Independent Variable——Scientific Relatedness

We measured scientific relatedness by adapting the “product space” framework, and defined a country’s scientific relatedness to a specific SDG domain as the alignment between its existing research portfolio and the disciplinary knowledge that the domain requires in each year.

Step 1: Constructing a Knowledge Space for Each Country
First, we constructed a national knowledge space in each year, representing the country’s scientific capabilities as its RCA across 252 subfields.

Step 2: Constructing a Knowledge Space for Each SDG Domain
Second, we built a knowledge space for each SDG in each year, representing the distribution of disciplinary knowledge it demands.

Step 3: Calculating Scientific Relatedness
Finally, we calculated the alignment by summing the importance scores of all subfields relevant to the SDGs, but only for those in which the country holds an RCA.

$$Relatedness_{i,s,t} = \boldsymbol{v}_{i,t} \cdot \boldsymbol{v}_{s,t} = \sum_{i=1}^{n} v_{i,t}^{(f)} \, v_{s,t}^{(f)}$$

Methodology

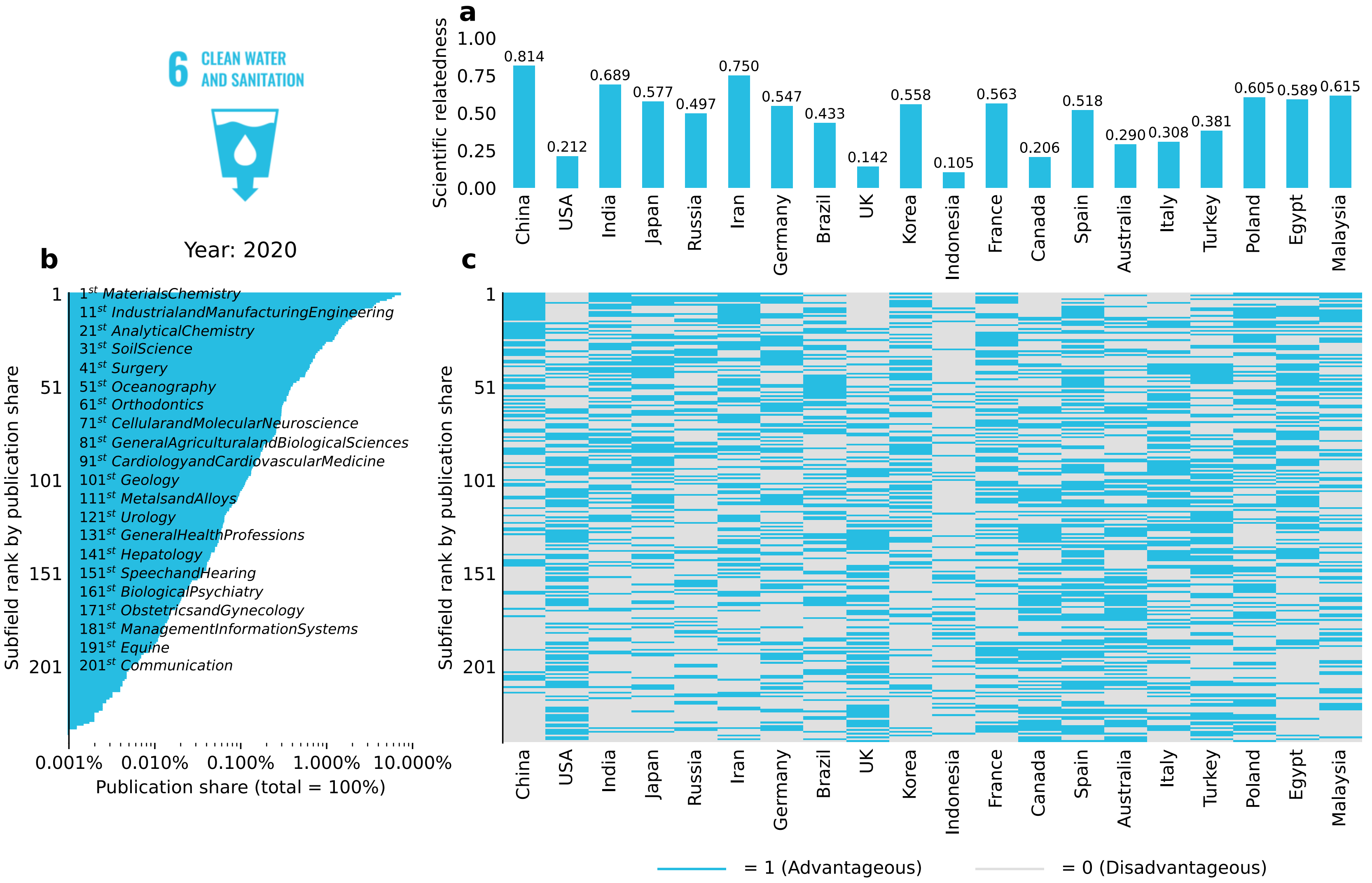


Figure 2. Example for scientific relatedness calculation for SDG 6 in 2020.

Methodology

Dependent Variable——Scientific Productivity and Novelty

Scientific productivity and scientific novelty are the key dependent variables.

Using a full counting method, we quantified scientific productivity of a country's future knowledge production ($LnSDGPub_{i,s,t+1}$), using the number of publications produced by country i within SDG s in year $t+1$.

To measure scientific novelty, we employed the widely validated recombinant novelty metric that is proposed by Uzzi et al. (2013).

Rooted in Schumpeterian innovation theory (Malerba & Orsenigo, 1996), this approach defines scientific novelty of a publication as an atypical combination of prior knowledge, where the observed frequency of any pair of cited journals is compared to the frequency of that pair occurring by chance.

We measured a country's future scientific novelty within an SDG domain using the variable $PrNovelty_{i,s,t+1}$:

$$PrNovelty_{i,s,t+1} = \frac{NP_{i,s,t+1}}{N_{i,s,t+1}}$$

Methodology

MAX PLANCK INSTITUTE FOR SOLID STATE RESEARCH

FUDAN UNIVERSITY

Moderating Variables

(1) Interdisciplinarity of existing knowledge structure

① Subfields similarity

$$\mathrm{S}_{\mathrm{mn,t}} = CosSim(V_{m,t}, V_{n,t})$$

② Proportions of disciplines applied in SDG

$$\mathrm{P}_{\mathrm{m,i,s,t}} = \frac{N_{m,i,s,t}}{N_{f,i,s,t}}, m \in f$$

$$\text{True Diversity}_{\mathrm{i,s,t}} = \frac{1}{\sum_{\mathrm{m \neq n}} \mathrm{s}_{\mathrm{mn,t}} \mathrm{P}_{\mathrm{m,i,s,t}} \mathrm{P}_{\mathrm{n,i,s,t}}}$$

(2) International collaboration

① **Intensity:** Proportion of its co-authored international publications

② **Network position**: Position in the global research network——structural hole

$$Network\ Constraint_{i,s,t} = \sum_{j \in N(i)} (p_{ij} + \sum_{q \in N(i)} p_{iq} p_{qj})^2$$

③ **Collaboration composition**: Collaboration with Global North and Collaboration with Global Sorth

Methodology

Variable	Definition
Dependent Var.	
$PrNovelty_{i,s,t+1}$	The proportion of novel publications published by country i in SDG s in year $t+1$
Independent Var.	
$Relatedness_{i,s,t}$	To degree to which country i's research strengths in year t align with the disciplinary knowledge structure required by SDG s
Control Var.	
Country-SDG level	
$lnSDGScore_{i,s,t}$	Country i's performance scores for achievement of SDG s in year t.
$lnSDGPubcount_{i,s,t}$	A Country's total number of publications in SDG s in year t.
$lnSDGCitation_{i,s,t}$	The average citations per article of publications published by country i in SDG s in year t.
Country-level	
$lnIC\_ratio_{i,s,t}$	The proportion of internationally collaborative publications published by country i in year t. An internationally collaborative publication is defined as the one where the authors are affiliated with institutions located in at least two countries.
$lnIC\_Constraint_{i,s,t}$	
$lnInterdisplinarity_{i,s,t}$	
$lnCitation_{i,t}$	Mean citations of publications that are published by country i in year t per article
$lnGDP_{i,t}$	Gross domestic capita of country i in year t
$lnPopulation_{i,t}$	Population size of country i in year t.
$HDI_{i,t}$	Human development index of country i in year t.

$$DV_{i,s,t+1} = \beta_0 + \beta_1 Relatedness_{i,s,t} + \gamma Z_{i,t} + \delta X_{i,s,t} + I_i + S_s + Y_t + \varepsilon_{i,s,t}$$

Result

Global Disparities in Scientific Relatedness

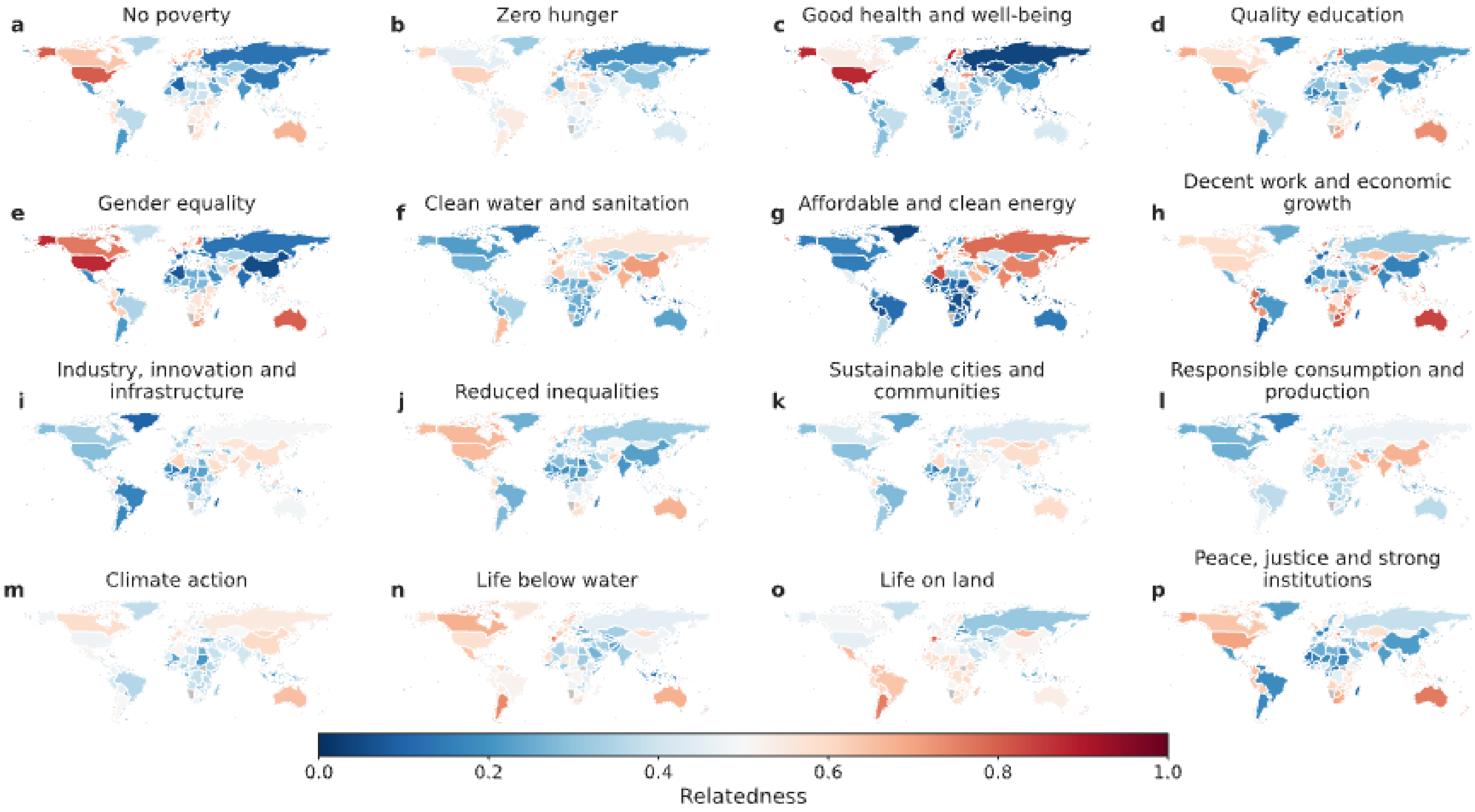


Fig. 3 reveals pronounced national disparities in scientific relatedness to SDGs, which seemly reflects distinct national priorities, capacities and specialization.

For instance, advanced economies exhibit higher relatedness to social-inclusion-oriented SDGs, such as "Peace, justice, and strong institutions" and "Reduced inequalities".

Several African countries show higher relatedness to "Zero hunger" and "Decent work and economic growth", likely mirroring their pressing local challenges.

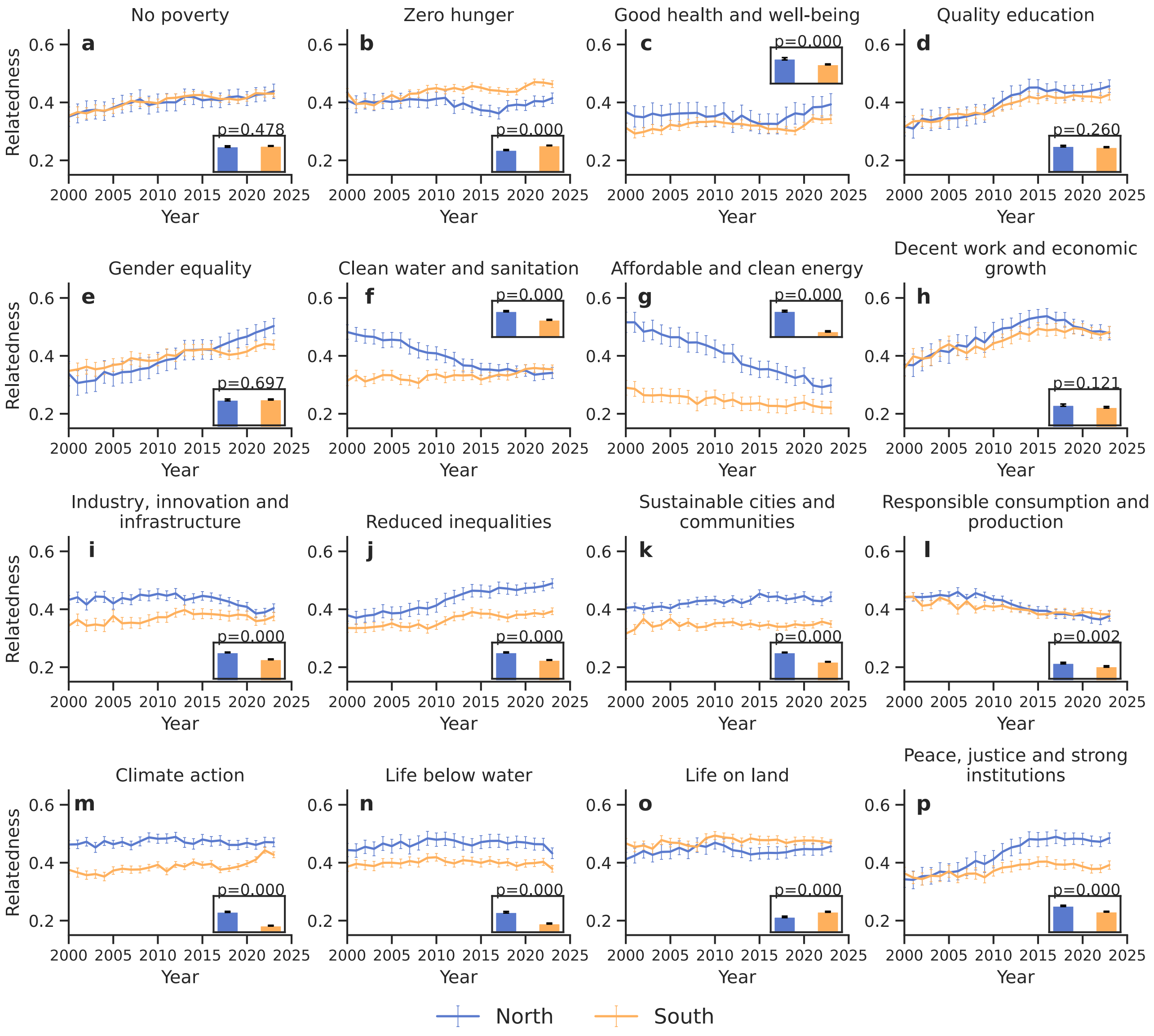
The United States demonstrates strong relatedness to "Good health and well-being", consistent with its robust research base in the life and medical sciences.

Result

Global Disparities in Scientific Relatedness

We further find a prominent divergence between the Global North and the Global South across most SDGs, with the Global North demonstrating a stronger scientific relatedness (Fig. 4).

This pattern holds across multiple years, with the Global North exhibiting higher relatedness in specific goals, such as “Good health and well-being”, “Quality education”, and “Industry, innovation and infrastructure”.



Result

The Dual Effect of Scientific Relatedness on Knowledge Production

Scientific relatedness increases the productivity of future knowledge production, while dampening its scientific novelty.

Increasing scientific relatedness ($\beta = 0.139$, $p < 0.01$) leads to a higher number of publications (Fig. 5a).

The negative coefficient of $Relatedness$ ($\beta = -0.235$, $p < 0.01$, Fig. 5b) suggest that a one-standard-deviation increase in $Relatedness$ corresponds to a 0.039-point (=0.235×0.165) decrease in the proportion of novel publications, equivalent to 9.35% of the mean of novelty probabilities (0.417).

The Global South experiences a stronger novelty penalty induced by scientific relatedness than the Global North.

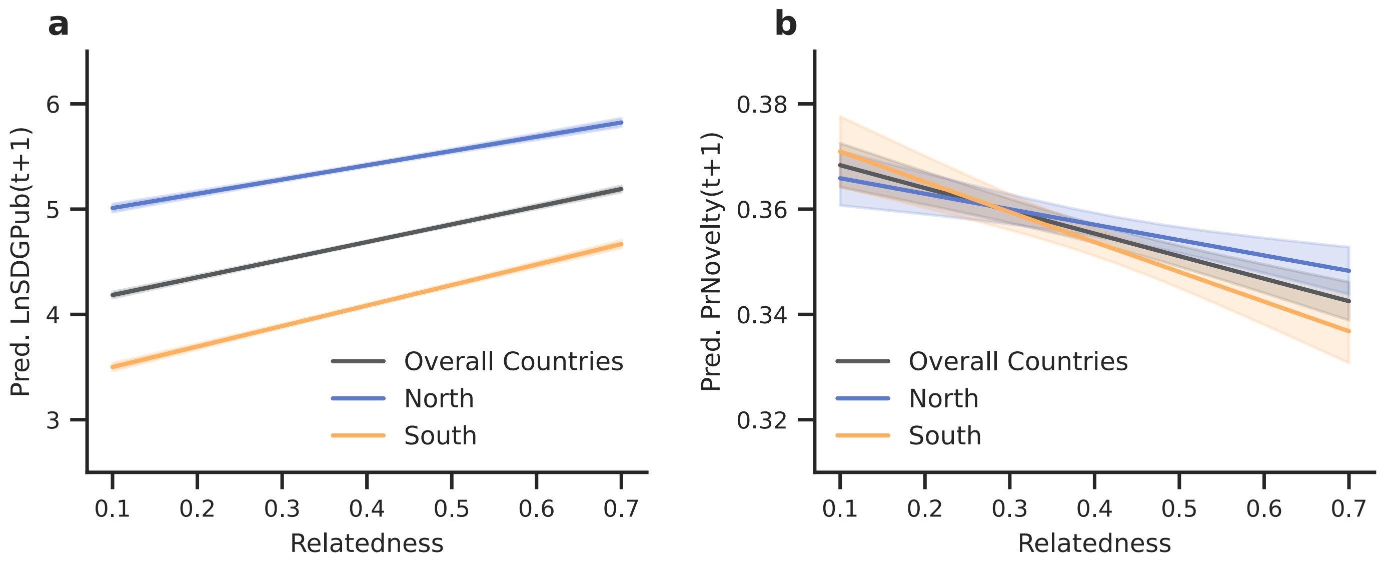


Figure 5. The estimated relationship between scientific relatedness, productivity, and novelty of future knowledge production.

Result

The Buffering Effect of Boundary Spanning on the Novelty Penalty: Internal Boundary Spanning

Increasing interdisciplinarity of a country's existing knowledge base in an SDG domain, amplifies, rather than mitigates, the negative relatedness-novelty relationship

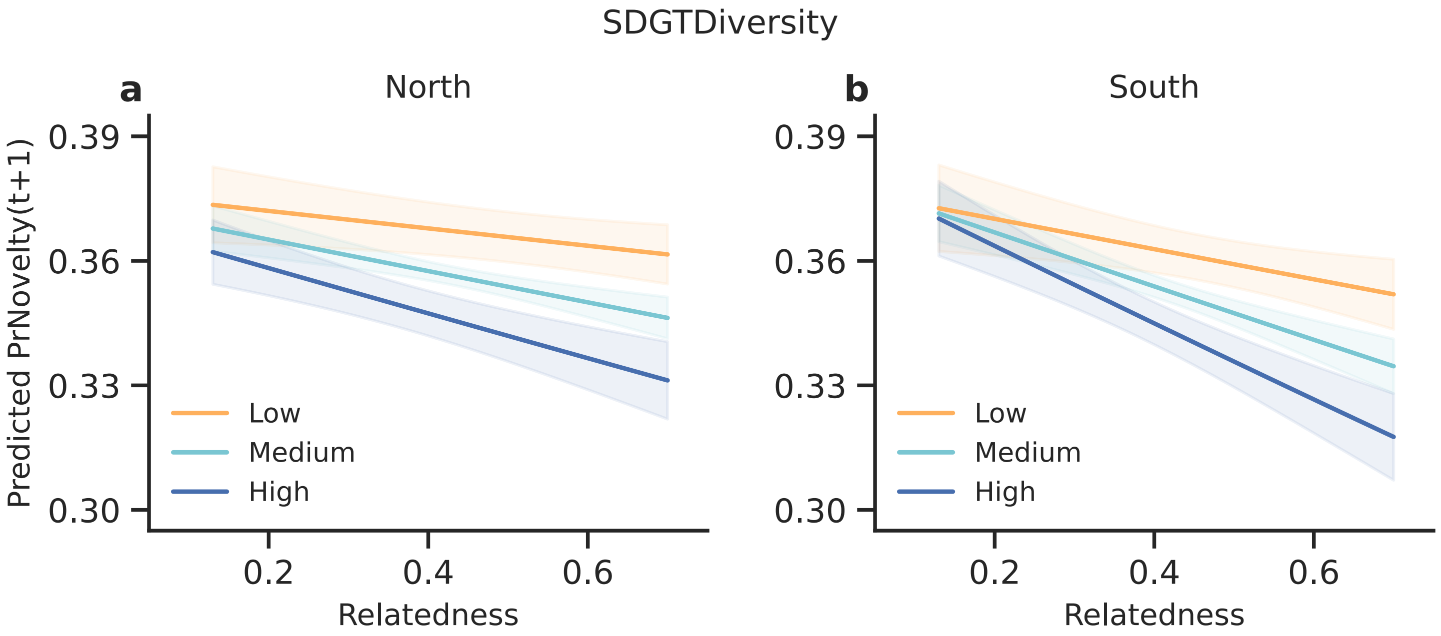


Figure 6. The estimated moderating effect of interdisciplinarity of a country's existing knowledge structure on the relatedness-novelty relationship.

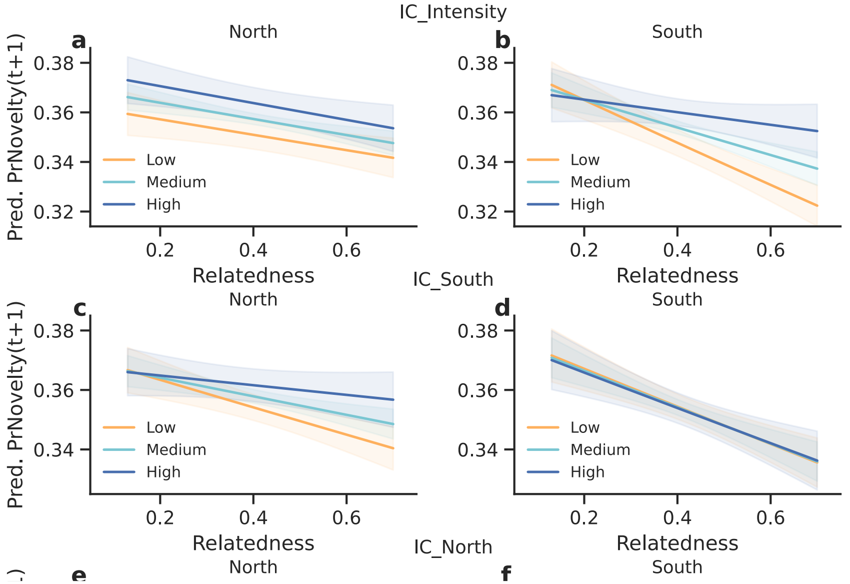
Result

The Buffering Effect of Boundary Spanning on the Novelty Penalty: External Boundary Spanning

Increased collaboration with the Global South, i.e., heterogenous collaboration, effectively buffers the novelty penalty for the Global North (Fig.7c), while increased homogenous collaboration intensifies this penalty (Fig.7e).

For the Global South, however, neither heterogeneous nor homogeneous collaboration shows a significant moderating effect (Figs. 7d and 7f).

A stronger network constraint in the global research collaboration network amplifies the negative relatedness-novelty relationship



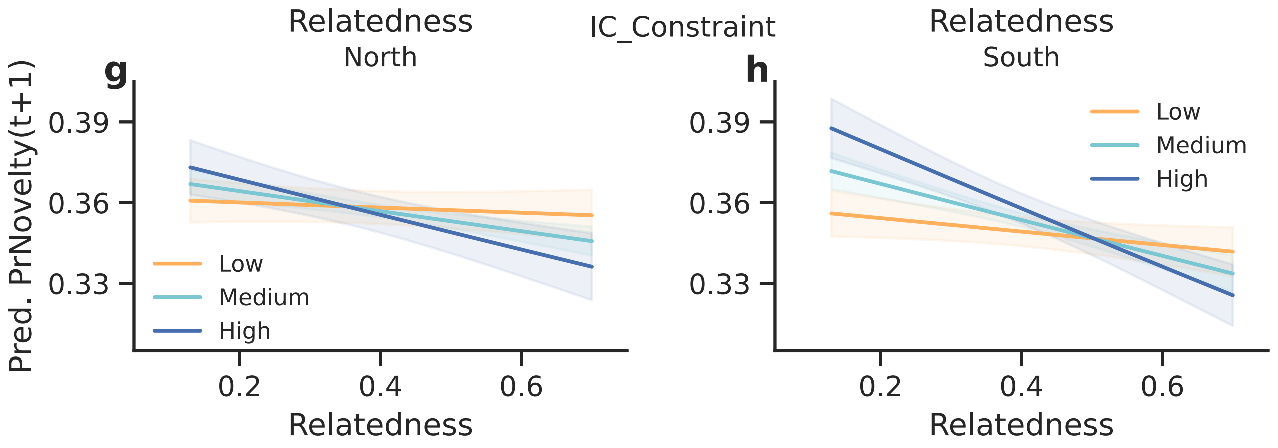


Figure 7. The estimated moderating effect of international collaboration variables on the relatedness-novelty relationship.

Result

Robustness Checks

First, we incorporated countries' annual research and development expenditure as an additional control variable to our regression formula. We subsequently extended the inherent time required for existing knowledge stock to translate into new research outputs from one year to three years (Makri et al., 2010).

To mitigate potential estimation artifacts arising from countries with small annual productivity in specific SDG domains, we applied a Bayesian smoothing procedure to the novelty proportion measures (Cressie, 1995). Moreover, we replicated the analysis using an alternative, text-based novelty measure (Arts et al., 2025).

Then, we applied a stricter criterion to identify SDG-related publications, by increasing the threshold from 0.4 to 0.6.

As the final robustness check, we restricted the sample to publications where the focal country holds a leadership author position (first or last author) (Ioannidis, 2025; Mingers & Xu, 2010). All robustness checks yielded results consistent with our main findings, confirming the stability and reliability of the reported relationships.

Result

Discussion and Conclusions

- Our findings uncover a persistent novelty penalty associated with scientific relatedness. This penalty hinders novel knowledge production in SDG-related research at the national level, with the Global South being disproportionately affected.

- The identified mitigating effect of international collaboration, particularly its capacity to buffer this penalty in both the Global North and the Global South contexts, therefore carries crucial policy implications.

- Our findings further challenge the conventional wisdom by revealing that interdisciplinarity does not uniformly facilitate novel knowledge production, and may, in fact, hinder it within SDG-related domains.

- Safeguarding and strengthening global research partnerships is not merely a scientific preference but a strategic imperative. For both the Global North and the Global South, international collaboration represents a vital mechanism to overcome structural disadvantages in the global knowledge system.

MAX PLANCK INSTITUTE FOR SOLID STATE RESEARCH

復旦大學 FUDAN UNIVERSITY

ICPP

THANKS

References

Allen, C., Metternicht, G., & Wiedmann, T. (2018). Initial progress in implementing the Sustainable Development Goals (SDGs): a review of evidence from countries. *Sustainability Science*, *13*(5), 1453-1467. https://doi.org/10.1007/s11625-018-0572-3

Antonietti, R., & Montresor, S. (2021). Going beyond Relatedness: Regional Diversification Trajectories and Key Enabling Technologies (KETs) in Italian Regions. *Economic Geography*, *97*(2), 187-207. https://doi.org/10.1080/00130095.2021.1920390

Apa, R., De Noni, I., Orsi, L., & Sedita, S. R. (2018). Knowledge space oddity: How to increase the intensity and relevance of the technological progress of European regions. *Research policy*, *47*(9), 1700-1712.

Balland, P. A., Boschma, R., & Frenken, K. (2015). Proximity and Innovation: From Statics to Dynamics. *Regional Studies*, *49*(6), 907-920. https://doi.org/10.1080/00343404.2014.883598

Balland, P. A., & Rigby, D. (2017). The Geography of Complex Knowledge. *Economic Geography*, *93*(1), 1-23. https://doi.org/10.1080/00130095.2016.1205947

Blicharska, M., Smithers, R. J., Kuchler, M., Agrawal, G. K., Gutiérrez, J. M., Hassanali, A.,…Mikusinski, G. (2017). Steps to overcome the North-South divide in research relevant to climate change policy and practice. *Nature Climate Change*, *7*(1), 21-27. https://doi.org/10.1038/Nclimate3163

Boschma, R. (2017). Relatedness as driver of regional diversification: a research agenda. *Regional Studies*, *51*(3), 351-364. https://doi.org/10.1080/00343404.2016.1254767

References

Boschma, R., Heimeriks, G., & Balland, P. A. (2014). Scientific knowledge dynamics and relatedness in biotech cities. *Research Policy*, *43*(1), 107-114. https://doi.org/10.1016/j.respol.2013.07.009

Boschma, R. A. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, *39*(1), 61-74. https://doi.org/10.1080/0034340052000320887

Confraria, H., Ciarli, T., & Noyons, E. (2024). Countries' research priorities in relation to the Sustainable Development Goals [Article]. *Research Policy*, *53*(3), 20, Article 104950. https://doi.org/10.1016/j.respol.2023.104950

Corradini, C. (2019). Location determinants of green technological entry: evidence from European regions. *Small Business Economics*, *52*, 845-858.

Cottafava, D., Ascione, G. S., Corazza, L., & Dhir, A. (2022). Sustainable development goals research in higher education institutions: An interdisciplinarity assessment through an entropy-based indicator. *Journal of Business Research*, *151*, 138-155. https://doi.org/10.1016/j.jbusres.2022.06.050

Dosi, G. (1982). Technological Paradigms and Technological Trajectories - a Suggested Interpretation of the Determinants and Directions of Technical Change. *Research Policy*, *11*(3), 147-162. https://doi.org/10.1016/0048-7333(82)90016-6

Fontana, M., Iori, M., Montobbio, F., & Sinatra, R. (2020). New and atypical combinations: An assessment of novelty and interdisciplinarity. *Research policy*, *49*(7), 104063.

References

Haunschild, R., & Bornmann, L. (2025). Mapping National Research That Targets Sustainable Development Goals: The Responsible Visualization of Openalex Data for Societal Impact Measurements of Research.

Heimeriks, G., & Boschma, R. (2014). The path- and place-dependent nature of scientific knowledge production in biotech 1986-2008. *Journal of Economic Geography*, *14*(2), 339-364. https://doi.org/10.1093/jeg/lbs052

Hidalgo, C. A., Klinger, B., Barabasi, A. L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science*, *317*(5837), 482-487. https://doi.org/10.1126/science.1144581

Kogler, D. F. (2015). Evolutionary economic geography–theoretical and empirical progress. In (Vol. 49, pp. 705-711): Taylor & Francis.

Li, L., Zhao, Y., & Shen, Z. (2024). Comparison of Sustainable Development Goals Labeling Systems based on Topic Coverage. *arXiv preprint arXiv:2408.13455*.

Lin, Z., Yin, Y., Liu, L., & Wang, D. (2023). SciSciNet: A large-scale open data lake for the science of science research. *Sci Data*, *10*(1), 315. https://doi.org/10.1038/s41597-023-02198-9

Liu, M., Bu, Y., Chen, C., Xu, J., Li, D., Leng, Y.,…Ding, Y. (2022). Pandemics are catalysts of scientific novelty: Evidence from COVID-19. *J Assoc Inf Sci Technol*, *73*(8), 1065-1078. https://doi.org/10.1002/asi.24612

Malekpour, S., Allen, C., Sagar, A., Scholz, I., Persson, A., Miranda, J. J.,…Al-Ghanim, K. (2023). What scientists need to do to accelerate progress on the SDGs. *Nature*, *621*(7978), 250-254. https://doi.org/10.1038/d41586-023-02808-x

References

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

Martin, R. (2010). Roepke Lecture in Economic Geography-Rethinking Regional Path Dependence: Beyond Lock-in to Evolution. *Economic Geography*, *86*(1), 1-27. https://doi.org/10.1111/j.1944-8287.2009.01056.x

Miao, L., Murray, D., Jung, W. S., Lariviere, V., Sugimoto, C. R., & Ahn, Y. Y. (2022). The latent structure of global scientific development. *Nat Hum Behav*, *6*(9), 1206-1217. https://doi.org/10.1038/s41562-022-01367-x

Neffke, F., Henning, M., & Boschma, R. (2011). How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. *Economic geography*, *87*(3), 237-265. https://doi.org/10.1111/j.1944-8287.2011.01121.x

Nelson, R. R., & Winter, S. G. (1985). *An evolutionary theory of economic change*. harvard university press.

Obamba, M. O., & Mwema, J. K. (2009). Symmetry and asymmetry: New contours, paradigms, and politics in African academic partnerships. *Higher Education Policy*, *22*(3), 349-371. https://doi.org/10.1057/hep.2009.12

Papke, L. E., & Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, *145*(1-2), 121-133. https://doi.org/10.1016/j.jeconom.2008.05.009

Perianes-Rodriguez, A., Waltman, L., & van Eck, N. J. (2016). Constructing bibliometric networks: A comparison between full and fractional counting. *Journal of Informetrics*, *10*(4), 1178-1195. https://doi.org/10.1016/j.joi.2016.10.006

References

Prashad, V. (2013). *The poorer nations: A possible history of the global south*. Verso.
Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv preprint arXiv:2205.01833*.
Santoalha, A., Consoli, D., & Castellacci, F. (2021). Digital skills, relatedness and green diversification: A study of European regions. *Research Policy*, *50*(9), 104340. https://doi.org/10.1016/j.respol.2021.104340
Schmidt, F., & Vanderfeesten, M. (2021). Evaluation on accuracy of mapping science to the United Nations' Sustainable Development Goals (SDGs) of the Aurora SDG queries. *Zenodo*.
Sweileh, W. M. (2020). Bibliometric analysis of scientific publications on "sustainable development goals" with emphasis on "good health and well-being" goal (2015-2019) [Article]. *Globalization and Health*, *16*(1), 13, Article 68. https://doi.org/10.1186/s12992-020-00602-2
Turba, R., Thoré, E. S., Bertram, M. G., Bridg, H., Sabet, S. S., Gamboa, M.,…Sampaio, A. C. (2025). Global North-South science inequalities due to language and funding barriers.
Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical Combinations and Scientific Impact [Article]. *Science*, *342*(6157), 468-472. https://doi.org/10.1126/science.1240474
van Knippenberg, D., De Dreu, C. K., & Homan, A. C. (2004). Work group diversity and group performance: an integrative model and research agenda. *J Appl Psychol*, *89*(6), 1008-1022. https://doi.org/10.1037/0021-9010.89.6.1008

MAX PLANCK INSTITUTE FOR SOLID STATE RESEARCH

FUDAN UNIVERSITY

References

Vanderfeesten, M., Jaworek, R., & Keßler, L. (2022). AI for mapping multi-lingual academic papers to the United Nations' Sustainable Development Goals (SDGs). *Zenodo*.

Wagner, C. S., & Leydesdorff, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. *Research Policy*, *34*(10), 1608-1618. https://doi.org/10.1016/j.respol.2005.08.002

Wang, S. Y., Ma, Y. X., Mao, J., Bai, Y., Liang, Z. T., & Li, G. (2023). Quantifying scientific breakthroughs by a novel disruption indicator based on knowledge entities [Article; Early Access]. *Journal of the Association for Information Science and Technology*, 18. https://doi.org/10.1002/asi.24719

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wuelser, G., & Pohl, C. (2016). How researchers frame scientific contributions to sustainable development: a typology based on grounded theory. *Sustainability Science*, *11*(5), 789-800. https://doi.org/10.1007/s11625-016-0363-7

Wulff, D. U., Meier, D. S., & Mata, R. (2024). Using novel data and ensemble models to improve automated labeling of Sustainable Development Goals. *Sustain Sci*, *19*(5), 1773-1787. https://doi.org/10.1007/s11625-024-01516-3

Yang, Y., Tian, T. Y., Woodruff, T. K., Jones, B. F., & Uzzi, B. (2022). Gender-diverse teams produce more novel and higher-impact scientific ideas [Article]. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(36), 8, Article e2200841119. https://doi.org/10.1073/pnas.2200841119

Zhang, M., Wang, L., & Li, Z. (2025). The impact of team compositions on disruptive and novel research in large-scale research infrastructures. *Scientometrics*, *130*(5), 2987–3011. https://doi.org/10.1007/s11192-025-05319-w