Research Evaluation, 2020, 1–14
doi: 10.1093/reseval/rvaa029
Article

Interdisciplinary knowledge combinations and emerging technological topics: Implications for reducing uncertainties in research evaluation

Seokbeom Kwon [1,2,*], Jan Youtie [3,4] and Alan L. Porter[3,5]

[1]Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan, [2]Department of Technology Management for Innovation, The University of Tokyo, 7–3-1 Hongo, Bunkyo-ku, Tokyo 113-8654, Japan, [3]School of Public Policy, Georgia Institute of Technology, 685 Cherry St NW, Atlanta, GA 30332, USA, [4]Enterprise Innovation Institute, Georgia Institute of Technology, 75 5th St NW, Atlanta, GA 30308, USA and [5]Search Technology Inc., 6025 The Corners Pkwy NW # 202, Norcross, GA 30092, USA

*Corresponding author. Email: kwon.seokbeom@mail.u-tokyo.ac.jp.

Abstract

This article puts forth a new indicator of emerging technological topics as a tool for addressing challenges inherent in the evaluation of interdisciplinary research. We present this indicator and test its relationship with interdisciplinary and atypical research combinations. We perform this test by using metadata of scientific publications in three domains with different interdisciplinarity challenges: Nano-Enabled Drug Delivery, Synthetic Biology, and Autonomous Vehicles. Our analysis supports the connection between technological emergence and interdisciplinarity and atypicality in knowledge combinations. We further find that the contributions of interdisciplinary and atypical knowledge combinations to addressing emerging technological topics increase or stay constant over time. Implications for policymakers and contributions to the literature on interdisciplinarity and evaluation are provided.

Key words: technological emergence; knowledge combination; interdisciplinary research

1. Introduction

Improving information to evaluate and detect emerging technologies has been a significant interest of public and private R&D managers (Van Raan and Van der Velde 1991; Porter et al. 2002; Cozzens et al. 2010). This level of interest is due in part to the potential for emerging technologies to produce beneficial socio-economic impacts (Martin 1995).

The perceived importance of technological emergence has led to extensive debate about how to define it. For instance, Rotolo, Hicks and Martin (2015) conceptualized emerging technology by extracting five attributes from a comprehensive review of relevant literature. Other researchers contributed to defining specific emerging technological areas to explore their developmental trajectories. Efforts toward delineating nanotechnology (Mogoutov and Kahane 2007; Porter et al. 2008; Arora et al. 2013) and a recent study to explore its emerging subdomains (Wang et al. 2019) are examples.

Scholars have discussed what factors shape technological emergence. Among others, studies seem to broadly agree with the idea that science is the key (Martin 1995; Small, Boyack and Klavans 2014). Scientific progress becomes a driver of technological emergence, enabling exploration of solutions to emerging problems. At the same time, technological emergence itself becomes a crucial input for further scientific progress (Porter et al. 2002; Hung and Chu 2006; Archibugi 2017). In their study, Kwon et al. (2019) showed that research addressing emerging technological topics within a science domain generates greater and broader scientific impact, as measured by citations.

What scientific research contributes to addressing emerging technological problems? Literatures on philosophy of science and research evaluation suggest that research integrating knowledge in diverse fields in a creative way may particularly contribute to technological emergence. Combining (seemingly) distant knowledge generates new knowledge and tools for further research (among others, Schumpeter 1961; Jacobs 1969; March 1991). Such knowledge may enable creative solutions to advance emerging technological topics. This proposition helps to theoretically elaborate on how science and technological emergence relate, and yet, empirical evidence concerning interdisciplinary combinations and the emergence of technological topics has not been extensively explored. Such exploration is important to research evaluation because interdisciplinary knowledge combinations can introduce uncertainties less prevalent in traditional discipline-focused work that can make research evaluation more difficult.

The ability to reduce uncertainties in evaluating interdisciplinary research by measuring technological emergence underscores our research proposition. As discussed above, examining what contributes to technological emergence is a special interest of research evaluators and policymakers whose mission is identifying promising research projects to support. In addition, the potential of knowledge combination for solving new problems in science is related to the long-standing scholarly endeavor toward evaluating the interdisciplinarity and novelty in research projects (e.g. Uzzi et al. 2013).

In this study, we contribute to narrowing the gap in the literatures on atypical research combinations and interdisciplinarity and on technological emergence by employing a recently developed bibliometric indicator of technological emergence (Porter et al. 2019)—the Tech Emergence Score. We operationalize the theoretical proposition into the following two hypotheses:

1. research integrating interdisciplinary knowledge addresses emerging technological topics to a greater extent than discipline-focused research.
2. research combining knowledge with greater atypicality addresses emerging technological ideas more than research based on typical knowledge combinations.

We test these hypotheses by using publication metadata (obtained in Web of Science abstract record sets). To identify emerging technological problems (i.e. topics), we use the tech emergence score (Porter et al. 2019). This indicator enables identification of terms emerging across a corpus of research publications in a given domain and evaluation of the extent to which each term is emergent in the domain. After extracting terms with high emergence scores, we calculate a publication-level emergence score by aggregating the scores of emerging terms appearing in the abstract and title of the publication. The publication-level emergence score quantifies the extent to which the publication's research outcomes address emerging topics in the field.

By defining interdisciplinarity as the extent to which a body of research originates from the integration of knowledge in diverse science disciplines (e.g. Porter et al. 2006; Wagner et al. 2011), we measured the interdisciplinarity by using the integration score (Porter, Roessner, Heberger 2008). To measure the atypicality of knowledge combination, we employed the novelty measure used by Lee, Walsh and Wang (2015). Our multivariable regression analysis using 2013–2015 research publications in three selected subdomains: Nano-Enabled Drug Delivery (NEDD), Synthetic Biology (Synbio), and Autonomous Vehicle (AutoV) found evidence supporting both hypotheses. Our further investigation showed that in the three fields under analysis, interdisciplinary and atypical knowledge combinations increasingly or continuously contribute to address emerging technological topics in each field.

The contribution of this study is threefold. First, our finding suggests that public R&D managers may benefit from investigating the 'interdisciplinarity' of the research projects when trying to support research projects for cultivating a certain emerging technological domain. Second, the present study contributes to extending the strain of research evaluation studies on measuring and assessing interdisciplinary research. Past research evaluation approaches have addressed the challenges of investigating interdisciplinary domains by applying disciplinary methods to interdisciplinary research assessments (Laudel and Origgi 2006). Recent years have seen the rise of a new generation of methodologies for addressing the uncertainties and distances inherent in interdisciplinary fields. Examples of these methodologies and tools include new science mapping techniques (Degn, Mejlgaard and Schneider 2019), alternative translational frameworks (Molas-Gallart et al. 2016), and narratives (Bone et al. 2020). Our study contributes to these new methods and measures for conducting research evaluation in an interdisciplinary context by showing usefulness of a new indicator of technological emergence and testing its connections with interdisciplinarity. Third, in light of a long-standing discussion of the appearance of a systematic bias in funding decisions against novel research projects that integrate knowledge of diverse fields (Porter and Rossini 1985; Metzger and Zare 1999; Langfeldt 2006; Boudreau et al. 2016; Bromham, Dinnage and Hua 2016), our findings further emphasize the importance of continuing institutional efforts toward supporting such research projects.

The remainder of this article is structured as follows. In Section 2, we review the literature on (1) knowledge combination as a driver of exploring solutions to new technological and research problems, and (2) the nature of technological emergence and its association with knowledge combinations. By reconciling the two pillars of literature, we draw two hypotheses. Section 3 illustrates our empirical research design, and Section 4 presents the analysis and results. In Sections 5 and 6, the implications of the findings and conclusions are provided.

2. Literature review and hypotheses development

2.1. Knowledge combination for science

Science is comprised of creative activities to solve problems (Simon 1977; Simon, Langley and Bradshaw 1981; Klahr and Simon 1999). In the course of solving problems, research generates new knowledge that may detect and accumulate anomalies of a dominant scientific theory, thereby advancing a possible scientific paradigm shift (Kuhn 1962).

From whence does scientific creativity originate? Researchers seem to broadly agree on the idea that combination of existing knowledge is a source of creativity. In his book, Jacobs (1970) argued that 'adding new kinds of work to other kinds of older work' becomes the source of human creativity. March (1991) also argued that combination of existing knowledge helps with exploration of untested new approaches to addressing problems.

Knowledge combination is also discussed in innovation studies. Schumpeter (1961) explained that one of the drivers of technological innovation is a combination of existing technology/knowledge to bring new products into the market. From his comprehensive literature review, Desrochers (2001) concludes that the combination of seemingly distant knowledge drives the diffusion of knowledge and, thus, contributes to technological innovation.

Economists agree with this conclusion. Nelson and Winter (1982) and Romer (1994) argued that scientific progress is an important driver of economic growth, and it originates from the combination of existing knowledge, materials, and arts. By modeling the knowledge creation process as the cumulative combination of existing knowledge (Azoulay, Graff Zivin and Manso 2011), Weitzman (1998) explicitly showed that the novel combinations of the old knowledge contribute to economic growth by becoming the crucial input for knowledge production.

Some studies attempted to extend these theoretical concepts through empirical analyses. For instance, by using patent data, Fleming (2001) quantified the extent to which an invention originated from combination of preexisting inventions in diverse technology fields. Strumsky and Lobo (2015) used a similar approach to explore which types of knowledge combinations are associated with technological novelty. Uzzi et al. (2013) suggested a new way of quantifying the novelty in knowledge combination in science. By using 17.9 million research articles indexed by Web of Science (WoS), the authors generated all the pair-wise combinations of the cited journals in these research articles. Then, they calculated how atypical the paired journals are based on joint citations. They used this metric as a measure of article-level novelty in knowledge combination. By modifying this method, Lee, Walsh and Wang (2015) examined the relationship between the characteristics of the research team and a resulting publication's novelty. Wang, Veugelers and Stephan (2017) subsequently showed that there is a great level of variability in the scientific impact of research that combines existing knowledge in atypical ways, while scholarly recognition of the value of such research is often delayed. More recently, Wagner, Whetsell and Mukherjee (2019) explored the relationship between international research collaboration and the novelty of resulting research outcomes.

If, as reviewed above, the novel combination of knowledge is a source of scientific creativity for solving new problems, how may it become relevant to addressing emerging technological issues?

2.2. Technological emergence and knowledge combination

Rotolo, Hicks and Martin (2015) characterized emerging technology as both radically novel (along with fast-growing, having coherence, and prominent impact) and at the same time, marked by uncertainty and ambiguity. According to this definition, it might not be too much of a stretch to interpret technological emergence as a phenomenon originating from the rise of a set of relevant new technological problems that increasingly draw the interest of a research community, but, as yet, lack a clear boundary or definition. This interpretation implies that technological emergence concerns new problems that need solutions or methods which have not been explored yet.

Given that emerging technological problems are new to the established research communities (Pistorius and Utterback 1997), and hence the solutions are not readily generated through conventional ways, addressing them may benefit from new scientific approaches, creatively applied. As we have reviewed previously, researchers repeatedly point to combining distant knowledge as a pathway toward this end. By combining knowledge that is seemingly distant, scientists may benefit in exploring possible solutions to real-world emerging technological issues (Belcher et al. 2016) that could not be addressed by applying single-domain knowledge. The derived from our literature review and consideration of the nature of technological emergence allows us to draw the following proposition:

> ***Proposition. Research that combines distant knowledge is more likely to contribute to addressing emerging technological issues than research combining proximate knowledge.***

The extent to which research combines distant knowledge can be measured by quantifying the extent of integration of knowledge in diverse disciplines (Mumford et al. 1991; Klein 2006; Mâsse et al. 2008) or how atypical the combination is in general (Uzzi et al. 2013; Lee, Walsh and Wang 2015; Wang, Veugelers and Stephan 2017). Studies conceptualized the former as *interdisciplinarity* or its variants (e.g. multidisciplinarity) in research (Porter et al. 2006; Wagner et al. 2011) while defining the later as *novelty* in knowledge combination. Accordingly, our proposition is distilled into the following two hypotheses:

> ***Hypothesis 1 (Interdisciplinarity): Research integrating multidisciplinary knowledge is more likely to address emerging technological topics than research involving disciplinary-focused knowledge.***
> ***Hypothesis 2 (Novelty): Research combining knowledge in novel ways is more likely to address emerging technological topics than is research involving conventional knowledge combinations.***

In the next section, we describe our research design to test the two hypotheses empirically.

3. Method

3.1. Overview of the research design

The empirical setting of this research is based on the study of Kwon et al. (2019). Because scientific publications document the original contributions of underlying research (Price 1963; Merton 1973), we consider an academic publication as a container of original research outcomes. We use the text description of the research in the abstracts and titles of publications as the primary information source that contains the essence of the research outcomes.

We propose a measure of the extent to which research outcomes address emerging topics in a field that can contribute to a reduction of evaluation research uncertainties by quantifying the degree to which an abstract or title in a scholarly publication mention emerging technological topics (terms) in the field. To this end, we extract emerging terms and calculate emergence scores by using the tech emergence score algorithm (Carley et al. 2018; Porter et al. 2019) from the corpus of publications published from 2003 to 2012 (10 years). The extracted terms represent the emerging topics in the field of interest over 10 time periods (years).

Next, we calculate the emergence score of each publication published in any of three consecutive years (i.e. 2013, 2014, 2015) by aggregating each of the extracted terms' emergence score appearing in the publication's abstract and title. In doing so, we quantify the degree to which the underlying research outcome in each publication addresses technological issues that evidence aggressive recent growth, while meeting emergence criteria concerning community, scope, novelty, and persistence.

We choose the 3 years considering the way publication-level emergence is calculated. The emerging terms are extracted from abstracts or titles of publications in the field of interest for the last

10 years. If we take too long a period of publication years for the analysis (for example from 2013 to 2019, rather than 2013 to 2015), the extracted emerging terms from 2003 to 2012 publications may less frequently appear in the newer publications as the technological emergence recognized presently is unlikely to be so after a few years. Taking too short a period (e.g. analysis of 2013 publications only) will be less useful because findings from the analysis can be susceptible to temporal patterns. Our choice of the 3 years is to accommodate these restrictions.

The unit of analysis is the individual publication. We provide further details on calculating the emergence score in Section 3.3.

3.2. Data

The ideal research design in bibliometric studies is often to use the metadata of the entire population of scientific publications. However, doing so is not preferable for this study because emerging topics are defined at the level of the technology domain. Instead, we start with several technology domains that are well-defined and distinct from one another.

We choose to analyze the publications in three research-driven domains: NEDD, Synbio, and AutoV. These three domains are selected because they represent potentially interdisciplinary fields with research that is oriented in novel combination of knowledge in diverse ways. NEDD represents more of a translational orientation to interdisciplinarity in the pharmaceutical research domain in the realm depicted by Molas-Gallart et al. (2016). AutoV represents an applied research field bringing together engineering and computer science fields for the automotive industry. Synbio bridges these two fields by merging biomedical and engineering and computer science but with economic potential for myriad industrial applications such as biofuels, medicine, and agricultural sectors among others. Testing our hypotheses about the extent of connection between interdisciplinary integration/atypical combination of knowledge and technological emergence across these domains (that represent diverse interdisciplinarity patterns) can offer insights into what research has contributed to the growth of these fields. Moreover, the dominant subdomains of relatively of these three fields encompass distinctive disciplines—Materials Science for NEDD, Biology for Synbio, and Information and Computer Science for AutoV—the disciplinary variation supports generalizability of findings.

Our data source is WoS, provided by Clarivate. We retrieved abstract records with metadata of NEDD publications from WoS by using the bibliometric definition formulated by Zhou et al. (2014). This definition yielded 92,514 publication records. Nearly 54,000 of these papers were published from 2003 to 2012, and 38,557 were 2013–2015 publications. For synthetic biology, we used the search strategy formulated by Shapira, Kwon and Youtie (2017) which yielded 7,377 publications. Of these, 4,041 were published in the 2003-2012 time period, and 3,336 were papers published from 2013 to 2015. We retrieved metadata of AutoV publications by using the keyword-based operational definition developed by Youtie et al. (2017). The search strategy resulting in 31,251 records of which 11,442 were published from 2013 to 2015.

We extracted emerging terms and calculated emergence scores from the 2003 to 2012 corpus by using *VantagePoint* (a text mining software—www.theVantagePoint.com) and relevant scripts provided by the Science, Technology, and Innovation Policy (STIP) program group at Georgia Tech. Then, we calculated a publication-level emergence score for each publication in the 2013–2015 corpus by aggregating the emergence scores of the emerging terms in its abstract and title.

3.3. Variables

3.3.1 Dependent variable

The dependent variable is the publication level emergence score (ES). We take a natural log transformation of ES, adding a value of 1 ($\boldsymbol{ln(ES+1)}$) to take into account the right-skewed distribution of the ES. To calculate the publication emergence score, we take the following steps as described in two prior studies (Carley et al. 2018; Porter et al. 2019).

- Extract all terms from the abstract and title of a corpus of publications published from 2003 to 2012 (a 10-year period) via *VantagePoint's* NLP (Natural Language Processing) to extract phrases, then consolidated using its 'RefineNLP' routine (entailing fuzzy matching of term variants and application of thesauri to remove noisy terms).
- Select the terms that pass the following thresholds:
  - *Growth*: The growth rates in frequency of the terms are at least 1.5 times greater than the growth rate of the overall publication records in the corpus.
  - *Community*: There are at least two organizations that have publications containing the term in question in the 10-year corpus.
  - *Scope*: Calculate the inverse document frequency (IDF)-value of each term based on a corpus of randomly retrieved publication records from WoS. If the calculated IDF-value of a term within the corpus of the technology domain of interest is greater than the IDF-value of the random publications, screen out this term.
  - *Novelty*: The term appeared more than $x\%$ (benchmark = 15%) of the publications from 2003 to 2005.
  - *Persistence*: The term appeared in at least seven records, in at least 3 years (to avoid 'one-hit wonders').
- Calculate an emergence score for each term: Calculate the following three metrics first—active trend, recent trend, and slope. The active trend measures the change in the extent of publications containing the term of interest between the period of the 4th–6th years and the 8th–10th publication years. The recent trend quantifies the change in a more recent period (9th–10th years versus 7th–8th years), and the slope takes the average year-growth rate of the share of publications containing the term by calculating the difference in the extent of publications containing the terms at the 7th and 10th publication years. The emergence score is calculated by aggregating the three variables. All terms that have lower emergence scores than a certain threshold value (set, based on empirical testing, at the square root of π, 1.77) are removed to clear out the terms that may be too weak to consider as a term representing an emerging technological idea.

Finally, the publication-level emergence score is calculated by summing up the emergence score of the terms that appeared in the abstract and title of the publication in question. This variable quantifies the extent to which the research outcome in the focal publication contains the terms related to technical emergence in the target domain. The higher the emergence score, the greater the extent to which the publication contains emerging technological terms and addresses cutting edge technological problems in the domain.

When publication does not contain the emerging technical terms in the domain in abstract or title, the emergence score for the publication is calculated as 0. If a publication has a positive emergence score, this positive score indicates that the publication has at least one emerging technical term.

Note that there are several pre-determined parameters employed in calculating the emergence score. The recent study by Liu and Porter (2020) tested the sensitivity of the emergence scoring algorithm to those parameters. According to their study, the selected parameters in the present study are in a relatively stable range in terms of sensitivity.

3.3.2 Independent variables

Following a definition of interdisciplinary research (i.e. research that integrates knowledge arising in diverse science disciplines) (e.g. Porter et al. 2006; Wagner et al. 2011), we operationalize interdisciplinarity as an integration score (***iScore***) that essentially measures the disciplinary diversity of the knowledge base (Stirling 2007) in the cited reference list, as the independent variable for testing hypothesis 1. We calculate ***iScore*** by using the following formula.

$$iScore_p = 1 - \frac{\sum_i \sum_j f_i f_j \cos(i,j)}{\sum_i \sum_j f_i f_j} \in [0,1]$$

where f_i is the share of a subject category i in the cited reference of publication p, cos(i, j) is the cosine similarity between subject categories i and j calculated based on the co-citation pattern by papers in category i and j.

The scientific discipline is proxied by WoS subject categories (WoS SCs) (Porter, Roessner, Heberger 2008). The greater the value of the ***iScore***, the more diverse the disciplines of cited papers, and, thus, the higher the interdisciplinary knowledge integration. When ***iScore*** takes a value of 0, this means that all the cited papers in the focal publication belong to a single sub-discipline. When ***iScore*** equals 1, it indicates that the disciplines of the cited references are fully distributed. To transform the ***iScore*** so that it approximates a normal distribution, we take the natural log transformation of iScore + 1 (***ln(iScore + 1)***). Note that the ***iScore*** is defined only if there is at least one cited WoS SC in the references. Hence, publications that have no valid cited WoS SCs information take a missing value.

To test hypothesis 2, we use the ***Novelty*** measure as the independent variable. Uzzi et al. (2013) suggested a way of measuring the extent to which a body of research originated from atypical combinations of existing knowledge by quantifying how rare the cited sources (i.e. journals) by a publication jointly appear. This method creates an imaginary counterpart publication by assigning the same number of randomly selected cited papers as the focal publication's cited reference list has. Then, the relative rarity of the combination of the cited sources in the focal paper compared to the counterpart is calculated. Uzzi's method was later modified by Lee, Walsh and Wang (2015) to make it computationally less intensive. In our study, we choose to use the method proposed by Lee et al for its computational benefit. Here we illustrate the details for calculating ***Novelty***.

- Retrieve a corpus of publications published in time t, in the technology domain of interest.
- For each publication in the corpus, generate pairs of all the cited sources.
- Pool the generated pairs together. The resulting list of the source pairs is *Ut*. The size of the *Ut* set is *Nt*.

The extent to which the sources in each pair (source i and j) are jointly cited in year t ($Commoness_{ijt}$) is calculated by the following formula.

$$Commoness_{ijt} = \frac{N_{ijt}}{\frac{N_{it}}{N_t} \times \frac{N_{jt}}{N_t} \times N_t}$$

Where N_{it} is the number of instances of source i in U_t. N_{ijt} is the number of instances of the pair of sources i and j in U_t. $\frac{N_{it}}{N_t}$ is the empirical probability that source i appears in U_t.

As a result, each publication comes to have a list of 'commonness' values for each pair of cited sources. The top 10th-percentile of commonness values is designated as the threshold for the *commonness* measure at the publication-level. Publication level novelty is calculated by taking on the inverse of the natural log of publication-level commonness. The novelty measure takes a continuous value without lower and upper limits. The greater the value of the ***Novelty***, the greater the 'unconventionality' or 'atypicality' in the combinations of the cited sources among the publications in the same field as the focal publication.

3.3.3 Control variables

To estimate the direct correlation between the independent and dependent variables, we introduce several control variables into the regression analysis.

First, we control for source-level fixed effects (***Source FE***) to take into account the probable heterogeneity in the correlation based on the venue where papers are published. For example, some journals explicitly target interdisciplinary research, while others are more in favor of discipline-oriented research. It is also plausible that some journals can be more active in publishing research on emerging technological topics, while others are not.

Second, we introduce a set of dummy variables for each of the publication years (i.e. 2013, 2014, and 2015) to take into account the probable time trend (***Pub Year FE***).

The third and fourth control variables are associated with research team characteristics. Wagner, Whetsell and Mukherjee (2019) showed that international collaborations for research tend to have lower novelty in terms of knowledge combinations because of the substantial transaction costs involved in coordinating interactions between researchers in different countries. According to their study, the transaction cost may suppress the research team's activity in exploring new combinations of knowledge. To consider this aspect, we introduce ***Int Collabo*** as a control variable. This variable takes the value of 1 if there are two or more countries listed in the authors' country information.

Besides, studies have shown that research team size is associated with research performance, while research collaboration is associated with scientific creativity (Cohen and Bailey 1997; Wuchty, Jones and Uzzi 2007; Vogel et al. 2013; Lee, Walsh and Wang 2015). To control for team size effects, we take into account the research team size as the fourth control variable by using the number of authors of the publication (***Team Size***).

It is plausible that the members of the interdisciplinary research team (or center) are more active in the integration of knowledge in diverse fields. Such interdisciplinary nature in the human capital of the research team may make the team more active in finding better solutions to emerging technological domain issues (Bishop et al. 2014). Yet, we argue that team-level interdisciplinarity is an alternative measure of the ***iScore***, rather than a compounding factor that needs to be controlled (Aydinoglu, Allard and Mitchell 2016) because the research team members' idea/information exchange and their knowledge integration are reflected in the disciplinary diversity of the cited references in the resulting research publication.

Fifth, a binary variable indicating whether the publication has a funding acknowledgment (***Funding***) is controlled. This variable is included to capture the fact that some grants are designed to support research on specific emerging technological areas while it has been argued that research funding allocation has been systematically biased against interdisciplinarity (Porter and Rossini 1985; Metzger and Zare 1999; Bromham, Dinnage and Hua 2016) or novel research (Boudreau et al. 2016).

Finally, we control for the first authors' country fixed effects (***Country FE***). This is to consider the difference in the research practice or resources by the country where the leading author is located.

For the regression analysis, we fit our data to an OLS regression model using robust standard errors to take into account probable violation of the homoskedasticity assumption in estimation.

If integration of interdisciplinary knowledge and combination of prior knowledge in atypical ways positively affect the extent to which research addresses emerging technological topics, the coefficients of ***ln(iScore + 1)*** and ***Novelty*** both are anticipated to take positive values.

4. Results

4.1. Descriptive analysis

Table 1 presents correlations between the key variables. All the correlation values are below 0.4, suggesting no critical issue in multicollinearity in the regression analysis. Note that some of the datapoints drop for the ***ln(iScore + 1)***. This is because the ***iScore*** for the publications that have no valid information about the cited WoS SCs cannot be calculated.

Figures 1 and 2 present simple correlations between the independent and dependent variables. Although there is a difference in the size of the slope, the ***ln(iScore + 1)*** is positively associated with the ***ln(ES + 1)*** across all three domains. Similarly, ***Novelty*** is positively associated with ***ln(ES + 1)*** in all three fields. These observations indicate that the greater the research interdisciplinarity, the more the extent to which the research addresses emerging technological topics within the field. The same pattern holds when it comes to the relations between ***Novelty*** and ***ln(ES + 1)***, implying that the greater the atypicality in knowledge combination, the more the research addresses emerging technological topics in the field.

4.2. Regression analysis

In this section, we report and interpret the regression results. Table 2 presents the main regression table.

The first four columns contain regression results using the ***ln(iScore + 1)*** as the independent variable. In the first column, we report the regression results using the entire dataset with the technology domain dummy. The coefficient of ***ln(iScore + 1)*** is 0.75 and statistically significant at the 0.01 significance level. When the integration score increases by 1%, the publication-level emergence score increases by 0.75%, on average, holding the other variables constant. The regression results for NEDD, Synbio, and AutoV publications are reported from the second to fourth columns, respectively. According to the result, a 1% increase in the integration score is associated with 1.43% and 0.63% increases in the publication-level emergence score, respectively, in NEDD and Synbio. Although the coefficient of ***ln(iScore + 1)*** for AutoV (see, the fourth column) is statistically insignificant at the 0.1 significance level, it remains positive (0.035, 0.035% change).

Overall, the regression results support Hypothesis 1. The greater the extent to which research integrates interdisciplinary knowledge, the more the research addresses emerging technological topics in the field. The exception to this finding is the AutoV regression. Although the association between iScore and emergence in the AutoV regression was positive, we cannot reject the null hypothesis stating that the two variables are not correlated. We argue that this is because of the low-coverage of WoS SCs in the cited references in AutoV publications. We provide a more detailed discussion in Section 4.3.

The fifth to the last column contains the regression results using ***Novelty*** as the independent variable. The coefficients of ***Novelty*** are all positive and statistically significant at the 0.05 significance level. According to the estimation result, a one-unit increase in ***Novelty*** is associated with a 10.2% increase in the publication-level emergence score, on average. For NEDD, Synbio, and AutoV publications, one unit increase in the ***Novelty*** variable results in 61.4%, 13%, and 1.4% increases in the publication-level emergence score, respectively. The positive and statistically significant coefficients of ***Novelty*** indicate that the higher the atypicality in research knowledge combination, the more the research addresses emerging topics in the field, which support Hypothesis 2.

As an alternative model, we also run Tobit regression. Our Tobit regression results reported in Supplementary Appendix B-1 show consistent findings with the OLS regression analysis.[1] For further robustness checking, we perform a regression analysis that controls for the number of cited references (***nRef***) while checking the sensitivity of the findings against the parameter that was used in calculating ***Novelty***. The findings are consistent with those from the main regression. We report these robustness check results in Supplementary Appendices B-2 and B-3, respectively.

4.3. Inconsistent findings: The case of autonomous vehicles

Although the series of regression analyses found evidence supporting Hypothesis 1, by and large, the correlation between ***ln(iScore + 1)*** and ***ln(ES + 1)*** was not statistically significant in the case of AutoV (see the fourth column of Table 2). We argue that this result does not necessarily undermine the validity of our findings. When it comes to publications on AutoV, many of their cited references were technical reports, conference proceedings, or working papers that were not indexed in WoS. Hence, most of the cited references by AutoV publications lack WoS SCs in the data, and this low number of WoS SCs might not produce enough variations in estimating the association between two variables of interest. Indeed, our data show that, on average, only 6.5% of the cited reference sources in an AutoV publication were indexed by WoS.

Our argument is further supported by the finding that the association between ***Novelty*** and ***ln(ES + 1)*** was positive and statistically significant for AutoV publications. In this analysis, the novelty measure was calculated by using data from 'sources of cited references' instead of the WoS SCs. This enables full utilization of the

Table 1. Correlation and summary statistics

NEDD	ln(ES + 1)	ln(iScore + 1)	Novelty	Pub year	Funding	Int Collabo	Team size
ln(ES + 1)	1.00						
ln(iScore + 1)	0.18	1.00					
	[0.00]						
Novelty	0.27	0.37	1.00				
	[0.00]	[0.00]					
Pub year	0.04	0.04	0.02	1.00			
	[0.00]	[0.00]	[0.00]				
Funding	0.10	0.08	0.11	0.01	1.00		
	[0.00]	[0.00]	[0.00]	[0.06]			
Int Collabo	−0.02	0.04	0.02	0.02	0.07	1.00	
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]		
Team size	−0.06	−0.03	0.00	0.03	0.19	0.22	1.00
	[0.00]	[0.00]	[0.70]	[0.00]	[0.00]	[0.00]	
Obs	38,549	37,454	37,543	38,549	38,549	38,549	38,549
Mean	2.71	0.43	0.91	2014.07	0.79	0.22	5.94
SD	1.55	0.10	0.40	0.81	0.41	0.41	3.27
Min	0	0	−14.88	2013	0	0	1
Max	5.49	0.69	2.32	2015	1	1	80
ln(ES + 1)	1.00						
ln(iScore + 1)	0.11	1.00					
	[0.00]						
Novelty	0.10	0.15	1.00				
	[0.00]	[0.00]					
Pub year	−0.01	0.05	0.08	1.00			
	[0.52]	[0.01]	[0.00]				
Funding	0.11	0.17	0.23	0.05	1.00		
	[0.00]	[0.00]	[0.00]	[0.01]			
Int Collabo	0.02	0.07	−0.02	0.03	0.11	1.00	
	[0.24]	[0.00]	[0.27]	[0.12]	[0.00]		
Team size	−0.07	−0.01	0.07	0.02	0.23	0.25	1.00
	[0.00]	[0.58]	[0.00]	[0.21]	[0.00]	[0.00]	
Obs	3,336	3,095	3,116	3,336	3,336	3,336	3,336
Mean	1.14	0.40	0.58	2014.07	0.76	0.22	4.66
SD	1.14	0.14	0.86	0.81	0.43	0.41	3.47
Min	0	0	−11.09	2013	0	0	1
Max	3.98	0.69	1.58	2015	1	1	58
AutoV	ln(ES + 1)	ln(iScore + 1)	Novelty	Pub Year	Funding	Int Collabo	Team Size
ln(ES + 1)	1.00						
ln(iScore + 1)	0.02	1.00					
	[0.11]						
Novelty	0.03	0.15	1.00				
	[0.00]	[0.00]					
Pub Year	0.00	−0.02	−0.03	1.00			
	[0.99]	[0.24]	[0.00]				
Funding	0.05	0.12	0.13	−0.09	1.00		
	[0.00]	[0.00]	[0.00]	[0.00]			
Int Collabo	0.01	0.04	0.07	0.03	0.16	1.00	
	[0.23]	[0.00]	[0.00]	[0.00]	[0.00]		
Team Size	0.02	−0.03	−0.02	0.03	0.12	0.20	1.00
	[0.01]	[0.04]	[0.06]	[0.00]	[0.00]	[0.00]	
Obs	11,440	6,021	11,139	11,440	11,440	11,440	11,440
Mean	0.53	0.35	−0.63	2014.09	0.28	0.16	3.61
SD	0.80	0.24	2.01	0.82	0.45	0.37	1.99
Min	0	0	−12.28	2013	0	0	1
Max	3.14	0.69	2.14	2015	1	1	74

Note: P-values in [].

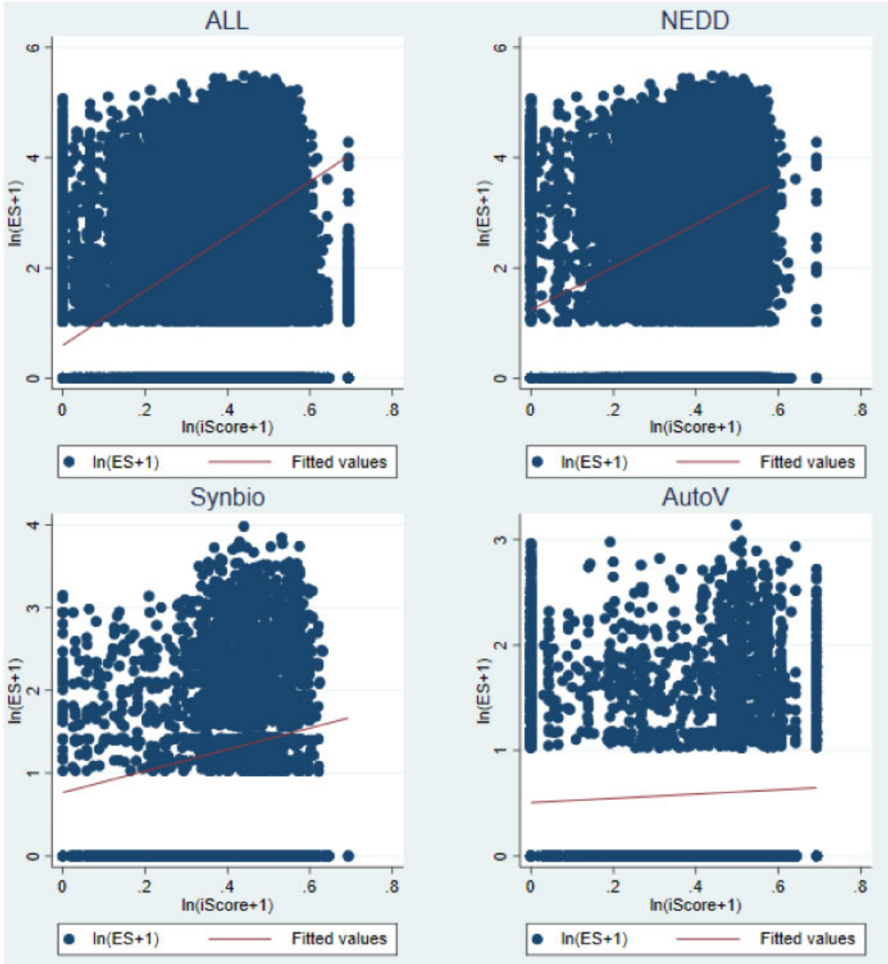

Figure 1. Scatter Plot of ln(iScore+1) vs. ln(ES+1)

information in the cited reference for estimating the association between the two variables of interest. Accordingly, we argue that the statistically insignificant association between ***ln(iScore + 1)*** and ***ln(ES + 1)*** for the AutoV publications was caused by the missing cited WoS SC information in many of the AutoV publications.

5. Change of the association over time

To what extent do interdisciplinarity and atypical knowledge combination offer utility in addressing emerging problems in a science domain? In this section, we explore an empirical answer through an additional analysis.

Toward this end, we examine the change in the marginal effect of the ***ln(iScore + 1)*** and ***Novelty*** on ***ln(ES + 1)*** over time. For our empirical setting, we use all the publications in the data (i.e. from 2003 to 2015) and create dummy variables corresponding to each publication year. Then, we generate the interaction terms between the publication year dummy variables and the two independent variables, respectively.

The coefficients of the interaction terms present the difference in the size of the marginal effect of the independent variable on the dependent variables. The estimated correlation between the independent and dependent variables using publications in 2003 becomes the reference. If the coefficients of the interaction terms take positive values, this indicates an increasing contribution of knowledge combination to addressing emerging technological topics from 2003. In contrast, if the coefficients are negative, it implies a diminishing contribution of the combination of distant knowledge to addressing emerging topics.

Figure 3 visualizes the regression result (the full regression table is reported in Supplementary Appendix C). Note that the upper bounds of all the 95% confidence intervals of the estimated coefficients never go below zero. This implies that, although there are differences in the pattern by domain, the marginal effect of atypical knowledge combination and integration of interdisciplinary knowledge on the extent to which research outcomes address emerging topics in the three domains seem to increase or stay constant over time. From this analysis, we find no evidence showing that the marginal contribution of combining distant knowledge to addressing emerging topics decreases over time in the three domains.

6. Discussion

In this study, we have examined whether research that combines distant knowledge contributes more to addressing emerging technological issues. We derived two hypotheses: (1) research integrating

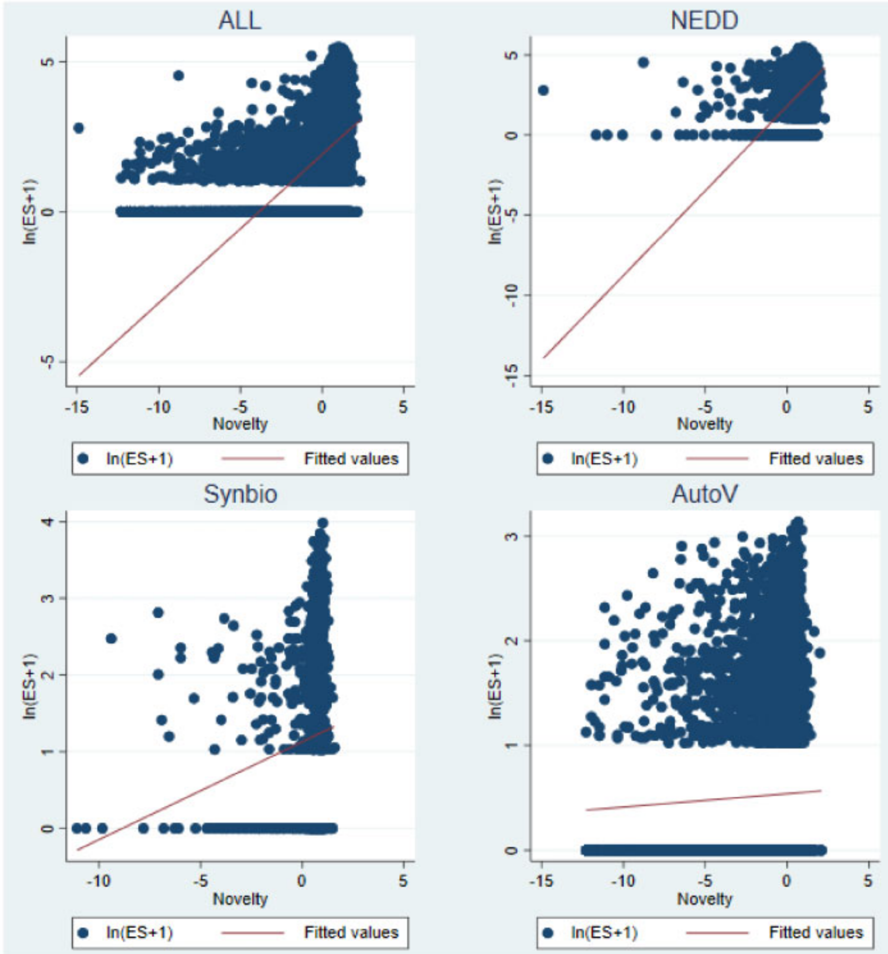

Figure 2. Scatter Plot of Novelty vs. ln(ES+1)

knowledge from diverse disciplines addressed more emerging technological topics, and (2) research combining prior knowledge in atypical ways addresses emerging technological topics better in the field.

We tested the two hypotheses by analyzing the text in titles and abstracts of WoS-indexed publications in NEDD, Synbio, and AutoV. We measured the extent to which a body of research addresses emerging technological topics by using publication-level 'emergence scores'.

Our analysis found consistent evidence supporting both hypotheses. The results indicated that the higher the integration score and the novelty of publication, the greater the publication-level emergence score. Our findings imply that research outcomes with greater interdisciplinarity and novelty in knowledge combinations address more emerging technological topics within the three domains we analyzed.

Do our findings imply that combining knowledge from diverse fields in an atypical (i.e. novel) way will necessarily make the research outcomes better in addressing emerging technological topics? Our research does not provide a definitive answer. First, it would be reasonable to interpret the findings such that researchers who try to address emerging topics, in the beginning, might tend to seek to combine knowledge in different disciplines. In the course of the searching process, researchers may try to combine knowledge in diverse disciplines in novel ways. This interpretation implies that encouraging researchers to search for knowledge in various disciplines and incorporate the knowledge into their research in an atypical way may not necessarily guarantee the creation of research outcomes that actually address emerging technological topics in the field.

Second, before reaching any conclusion, it is necessary to properly take into account the fact that research teams are likely to face extensive transaction costs when they integrate knowledge from diverse disciplines in a novel way (Wagner, Whetsell and Mukherjee 2019). Such integration entails the extra cost of searching for knowledge outside one's field and assimilating the information into the research. Furthermore, as many prior studies have highlighted, research combining prior knowledge in atypical ways may run the risk of delayed recognition of its scientific contributions (Garfield 1980; Van Raan 2004; Stephan, Veugelers and Wang 2017; Wang, Veugelers and Stephan 2017). These research team dynamics do not necessarily function negatively for the process of interdisciplinary/atypical knowledge combination. Studies suggest that if a research team consists of members with a diverse knowledge base, the transaction cost in interdisciplinary or novel knowledge integration can be mitigated (Falk-Krzesinski et al. 2011; Basner et al. 2013). This suggests that the desirability and feasibility for research teams to combine knowledge from diverse disciplines in novel ways will partly depend on the research team's collective capacity of orchestrating

Table 2. Main regression results

Variable	All three ln(ES + 1)	NEDD ln(ES + 1)	SynBio ln(ES + 1)	AutoV ln(ES + 1)	All three ln(ES + 1)	NEDD ln(ES + 1)	SynBio ln(ES + 1)	AutoV ln(ES + 1)
ln(iScore + 1)	**0.748***** **(0.0510)**	**1.427***** **(0.0864)**	**0.631***** **(0.206)**	0.0351 (0.0609)				
Novelty					**0.102***** **(0.00640)**	**0.614***** **(0.0463)**	**0.130**** **(0.0596)**	**0.0136**** **(0.00597)**
Funding	0.0477** (0.0200)	0.0417* (0.0226)	0.171* (0.0888)	0.00958 (0.0437)	0.0333* (0.0187)	0.0139 (0.0224)	0.199** (0.0882)	−0.00739 (0.0309)
Int Collabo	−0.0181 (0.0163)	−0.0193 (0.0182)	0.0530 (0.0660)	−0.0250 (0.0381)	−0.0190 (0.0155)	−0.0202 (0.0180)	0.0641 (0.0656)	−0.00522 (0.0273)
Team Size	−0.000238 (0.00239)	0.00139 (0.00265)	−0.0237*** (0.00703)	0.00804 (0.00784)	0.000311 (0.00232)	0.0000975 (0.00261)	−0.0231*** (0.00708)	0.00784 (0.00552)
AutoV	−1.688*** (0.0764)				−1.626*** (0.0684)			
SynBio	−0.930*** (0.0385)				−0.930*** (0.0382)			
Constant	2.131*** (0.405)	2.079*** (0.472)	1.052*** (0.141)	−0.735** (0.370)	2.554*** (0.351)	2.680*** (0.460)	1.212*** (0.127)	0.0192 (0.640)
R^2	0.508	0.379	0.374	0.349	0.556	0.388	0.376	0.286
Adjusted R^2	0.440	0.313	0.122	0.031	0.492	0.322	0.121	0.032
Pub Year FE	All	NEDD	SynBio	AutoV	All	NEDD	SynBio	AutoV
Source FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
1st author country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Publication period	(13–15)	(13–15)	(13–15)	(13–15)	(13–15)	(13–15)	(13–15)	(13–15)
Observations	46,461	37,396	3,077	5,988	51,650	37,493	3,101	11,056

Robust standard errors in parentheses.
*P < 0.1, **P < 0.05, ***P < 0.01.

the interdisciplinary/atypical knowledge integration process. We believe that empirically testing this proposition is an intriguing research question for future studies.

One may question *how* possibly research combining knowledge in diverse fields in a novel way contributes to scientific progress. Our findings in this study, in conjunction with the conclusions of the study by Kwon et al. (2019), suggest an answer. They have shown that research addressing emerging technological topics had greater and broader citation-based impacts on subsequent research. Incorporating this conclusion into our findings leads us to make an argument that research with interdisciplinary and novel combinations of knowledge contributes to addressing emerging technological topics, and it will contribute to generating new knowledge that has a greater and broader scientific impact. That is, papers that cite more diverse research tend to be more apt to address cutting edge (emerging) topics, and, eventually, to be more widely cited themselves. Whether the effect exists and how large the effect is are intriguing questions for future research.

7. Conclusions

Our study provides broad implications for policymakers and the research evaluation community. First, an indicator of technological emergence can reduce evaluation uncertainties by highlighting which research topics are more likely to persist in the future and be novel, grow, and have a community around them. In evaluation of emerging science and technology research, identifying and measuring the extent to which a research outcome addresses emerging technological ideas has been a substantial challenge because technological emergence is accompanied by uncertainties as well as ambiguities in its definition and operationalization. Our research contributes to addressing this difficulty by offering a method of calculating a research outcome-level emergence score. This score identifies which pathways are more likely to be taken up in an emerging science and technology domain in future years (Porter et al. 2019), which can help in developing evaluation designs for these new research areas.

We show that research combining interdisciplinary knowledge is more likely than disciplinary-focused research to address emerging technological ideas. Policymakers who seek to support research projects on cutting edge topics may benefit from this finding. For example, when trying to support research projects for cultivating a certain emerging technological domain, public R&D managers may need to evaluate not only whether the research project explicitly targets emerging technological topics in the field, but also the research team's capability of combining knowledge from diverse disciplines.

Our finding also emphasizes the necessity of continuing support for interdisciplinary research through science policy measures. Researchers have found evidence that, although interdisciplinary research can generate scientifically impactful knowledge (Kwon et al. 2017) and science has become increasingly interdisciplinary (Porter and Rafols 2009), studies have shown that research funding schemes have been biased against interdisciplinary teams because reviewers of the research proposals often favor discipline-oriented research (Porter and Rossini 1985; Metzger and Zare 1999; Bromham, Dinnage and Hua 2016). Our article contributes to these studies by showing that interdisciplinary research can distinctively contribute to addressing emerging topics. This suggests that research metrics gauging degree of interdisciplinarity or novelty in knowledge combination warrant research evaluation attention.

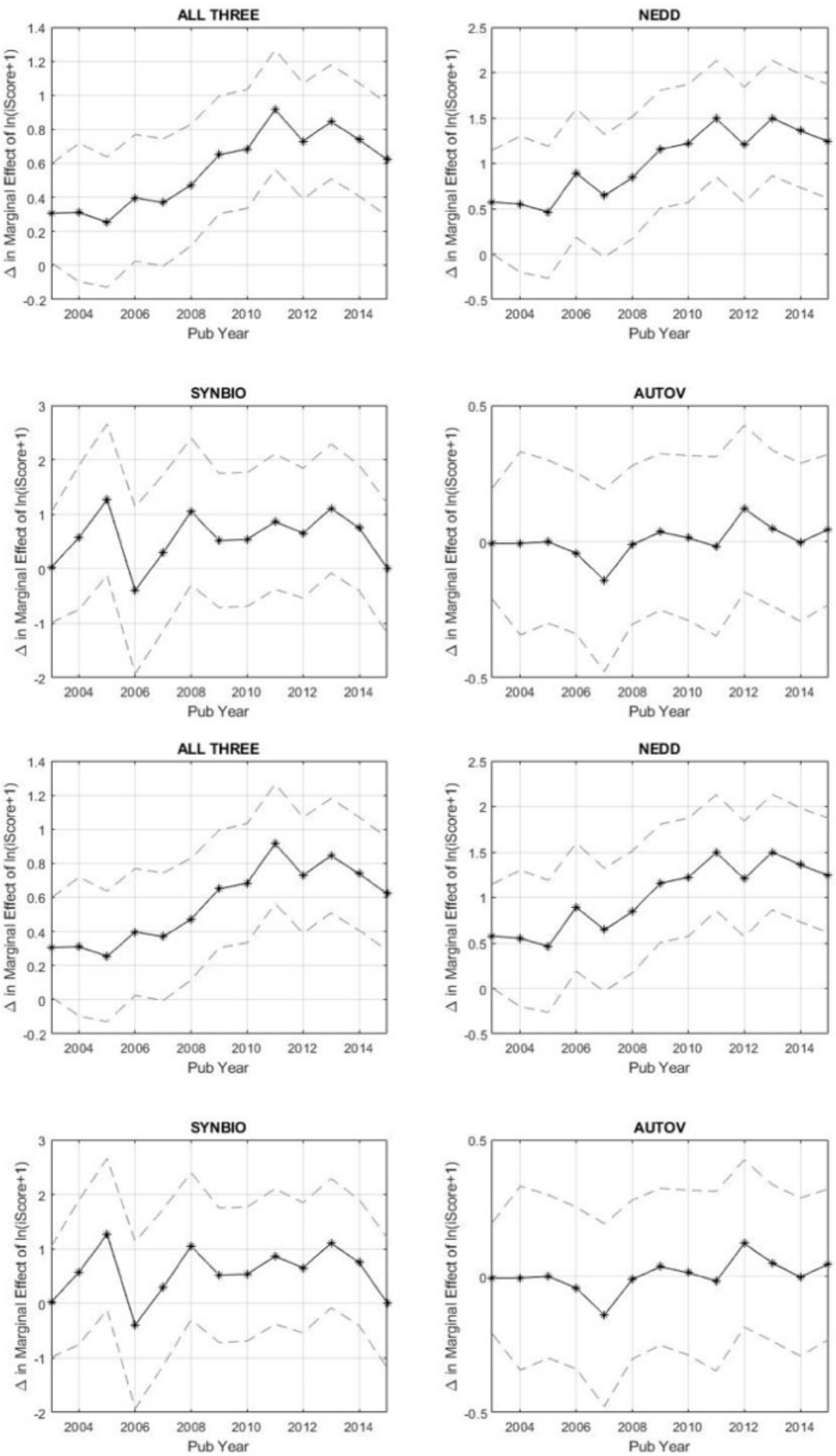

Figure 3. Change in the Marginal Effect of ln(iScore+1) and Novelty over time Black solid: estimated coefficient; Gray dashed: 95% confidence interval (using roust standard error)

As a related issue, research evaluators have repeatedly raised concerns that contemporary research funding allocation practices favor less-risky research projects (Azoulay, Graff Zivin and Manso 2011; Petsko 2012) and are somewhat biased against novel research (Boudreau et al. 2016). Novel research may have delayed recognition of its contribution to science (Garfield 1980; Van Raan 2004; Wang, Veugelers and Stephan 2017) and the optimal incentive for innovation is tolerating early failures while rewarding long-term successes (Manso 2011). Thus, our findings highlight the importance of institutionalizing support for 'novel research' that may particularly contribute to emerging technological topics within a domain. That implies value in evaluating proposal novelty based on measures of the reach of cited research across disciplines.

Aside from the issue of funding allocation for interdisciplinary research as discussed above, our research contributes to the scholarly effort to address considerable challenge in evaluating interdisciplinary research. Interdisciplinarity poses uncertainties because the research is not anchored in journals for conventional fields where knowledge is more well defined (Degn, Mejlgaard and Schneider 2019). Interdisciplinarity also exacerbates distances between researchers in different disciplines, with Molas-Gallart et al. (2016) and Bone et al. (2020) reflecting these distances in their versions of the proximity framework of Boschma (2005) as geographic, cognitive, social, organizational, and institutional distances. New methods and tools have appeared in recent years to address the interdisciplinarity research evaluation challenges. Degn, Mejlgaard and Schneider (2019) have used co-nomination alongside traditional bibliometric methods to develop maps of science for interdisciplinary social science and humanities fields. Molas-Gallart et al. (2016) put forth an alternative research evaluation framework to the linear research continuum in which translational gaps are placed at the end of the continuum with a socio-economic orientation. The alternative framework proposes application of multiple methods—such as geographic information systems, science maps, and social network analysis—to assess the types of gaps between researchers in different medically related disciplines. Bone et al. (2020) promote a Diversity Approach to Research Evaluation (DARE) method which brings narratives of participants together with science maps and indicators of diversity and cohesiveness. Our approach suggests the addition of an indicator of technological emergence to these methods. The concept of technological emergence has been developed in the literature to reflect the multiple dimensions underlying the combination of different disciplines to create new knowledge. The ability to measure and incorporate technological emergence into research evaluation is important to reducing interdisciplinarity-related uncertainties and ambiguities.

The present study also contributes to advancing the broad literature on scientific creativity and novelty. Ever since the seminal work by Uzzi et al. (2013), there has been a substantial amount of subsequent work on elucidating various dynamics associated with creativity and novelty in science (e.g. Lee, Walsh and Wang 2015; Wang, Veugelers and Stephan 2017; Wagner, Whetsell and Mukherjee 2019). These studies have contributed to improving the understanding of novelty in research, the nature of novel research, the impacts of novelty on science, and various associated dynamics such as the characteristics of research teams and research collaboration. In addition to these contributions, we reveal another important pathway concerning how novelty in knowledge combinations contributes to science by showing that atypicality in knowledge combinations are positively associated with addressing emerging technological issues. These findings affirmatively answer the question of whether the combination of distant knowledge is one of the drivers of scientific progress.

The present study has several limitations, which we hope that future research can address. First, we used the novelty measure that was developed by Lee, Walsh and Wang (2015). However, as Wang, Veugelers and Stephan (2017) and Wagner, Whetsell and Mukherjee (2019) indicate, various ways of calculating novelty exist. Whether other ways of operationalizing novel knowledge combinations produce different or consistent findings with the present study is an empirical question. Second, for the purpose of the present study, we have measured 'interdisciplinarity' by using the cited reference information. However, as prior studies have shown, there are various dimensions of interdisciplinarity, including how future research is affected by interdisciplinary studies (Carley and Porter 2012) or more subtle subdimensions of the notion of interdisciplinarity itself (e.g. Stirling 2007). Future studies can examine whether these other dimensions of, or ways of measuring, interdisciplinarity play a similar role in technological emergence to the findings in this study. Third, our analysis using the metadata of scientific publications leaves open the question as to whether our findings would also hold when analyzing patents. Because patents contain detailed information about a technological idea, we believe that replicating our analysis using patent information could extend the conclusions of the present study.

Supplementary data

Supplementary data are available at *Research Evaluation Journal* online.

Notes

1. Note that in the Tobit regression, the source FE is not controlled because the currently available statistical package is not capable of handling a large number of dummy variables. In our data, over 12,000 sources were appeared, which implies that more than 12,000 dummy variables should be introduced into the regression for Source FE. This could be possible in OLS regression with the *areg* with *absorb* option in Stata. However, the similar function is currently not available for the Tobit regression.

Acknowledgements

We acknowledge support from the US National Science Foundation (Award #1759960—'Indicators of Technological Emergence') to Search Technology, Inc., and Georgia Tech. Seokbeom Kwon was also supported by the grant from Japan Society for the Promotion of Science KAKENHI (Grant # 20K22116). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the National Science Foundation or Japan Society for the Promotion of Science.

References

Archibugi, D. (2017) 'Blade Runner Economics: Will Innovation Lead the Economic Recovery?', *Research Policy*, 46, 535–43.

Arora, S. K. et al. (2013) 'Capturing New Developments in an Emerging Technology: An Updated Search Strategy for Identifying Nanotechnology Research Outputs', *Scientometrics*, 95, 351–70.

Aydinoglu, A. U., Allard, S., and Mitchell, C. (2016) 'Measuring Diversity in Disciplinary Collaboration in Research Teams: An Ecological Perspective', *Research Evaluation*, 25, 18–36.

Azoulay, P., Graff Zivin, J. S., and Manso, G. (2011) 'Incentives and Creativity: Evidence from the Academic Life Sciences', *The Rand Journal of Economics*, 42, 527–54.

Basner, J. E. et al. (2013) 'Measuring the Evolution and Output of Cross-Disciplinary Collaborations within the NCI Physical Sciences–Oncology Centers Network', *Research Evaluation*, 22, 285–97.

Belcher, B. M. et al. (2016) 'Defining and Assessing Research Quality in a Transdisciplinary Context', *Research Evaluation*, 25, 1–17.

Bishop, P. R. et al. (2014) 'Impacts of an Interdisciplinary Research Center on Participant Publication and Collaboration Patterns: A Case Study of the National Institute for Mathematical and Biological Synthesis', *Research Evaluation*, 23, 327–40.

Bone, F. et al. (2020) 'DARE to Be Different? a Novel Approach for Analysing Diversity in Collaborative Research Projects', *Research Evaluation*, 29, 300–15.

Boschma, R. (2005) 'Proximity and Innovation: A Critical Assessment', *Regional Studies*, 39, 61–74.

Boudreau, K. J. et al. (2016) 'Looking across and Looking beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science', *Management Science*, 62, 2765–83.

Bromham, L., Dinnage, R., and Hua, X. (2016) 'Interdisciplinary Research Has Consistently Lower Funding Success', *Nature*, 534, 684.

Carley, S. and Porter, A. L. (2012) 'A Forward Diversity Index', *Scientometrics*, 90, 407–27.

Carley, S. F. et al. (2018) 'An Indicator of Technical Emergence', *Scientometrics*, 115, 35–49.

Cohen, S. G. and Bailey, D. E. (1997) 'What Makes Teams Work: Group Effectiveness Research from the Shop Floor to the Executive Suite', *Journal of Management*, 23, 239–90.

Cozzens, S. et al. (2010) 'Emerging Technologies: Quantitative Identification and Measurement', *Technology Analysis & Strategic Management*, 22, 361–76.

Degn, L., Mejlgaard, N., and Schneider, J. W. (2019) 'Using Mixed Methods to Map Vaguely Defined Research Areas', *Research Evaluation*, 28, 394–404.

Desrochers, P. (2001) 'Local Diversity, Human Creativity, and Technological Innovation', *Growth and Change*, 32, 369–94.

Falk-Krzesinski, H. J. et al. (2011) 'Mapping a Research Agenda for the Science of Team Science', *Research Evaluation*, 20, 145–58.

Fleming, L. (2001) 'Recombinant Uncertainty in Technological Search', *Management Science*, 47, 117–32.

Garfield, E. (1980) 'Premature Discovery or Delayed recognition-Why', *Current Contents*, 21, 5–10.

Hung, S.-C. and Chu, Y.-Y. (2006) 'Stimulating New Industries from Emerging Technologies: Challenges for the Public Sector', *Technovation*, 26, 104–10.

Jacobs, J. (1969) *The Economy of Cities*. New York: Vintage Books.

Klahr, D. and Simon, H. A. (1999) 'Studies of Scientific Discovery: Complementary Approaches and Convergent Findings', *Psychological Bulletin*, 125, 524.

Klein, J. T. (2006) 'Afterword: The Emergent Literature on Interdisciplinary and Transdisciplinary Research Evaluation', *Research Evaluation*, 15, 75–80.

Kuhn, T. S. (1962) *The Structure of Scientific Revolutions*. Chicago: University of Chicago press.

Kwon, S. et al. (2017) 'A Measure of Knowledge Flow between Specific Fields: Implications of Interdisciplinarity for Impact and Funding', *PLoS One*, 12, e0185583.

Kwon, S. et al. (2019) 'Research Addressing Emerging Technological Ideas Has Greater Scientific Impact', *Research Policy*, 48, 103834.

Langfeldt, L. (2006) 'The Policy Challenges of Peer Review: Managing Bias, Conflict of Interests and Interdisciplinary Assessments', *Research Evaluation*, 15, 31–41.

Laudel, G. and Origgi, G. (2006) 'Introduction to a Special Issue on the Assessment of Interdisciplinary Research', *Research Evaluation*, 15, 2–4.

Lee, Y.-N., Walsh, J. P., and Wang, J. (2015) 'Creativity in Scientific Teams: Unpacking Novelty and Impact', *Research Policy*, 44, 684–97.

Liu, X. and Porter, A. L. (2020) 'A 3-Dimensional Analysis for Evaluating Technology Emergence Indicators', *Scientometrics*, 124, 27–55.

Manso, G. (2011) 'Motivating Innovation', *The Journal of Finance*, 66, 1823–60.

March, J. G. (1991) 'Exploration and Exploitation in Organizational Learning', *Organization Science*, 2, 71–87.

Martin, B. R. (1995) 'Foresight in Science and Technology', *Technology Analysis & Strategic Management*, 7, 139–68.

Mâsse, L. C. et al. (2008) 'Measuring Collaboration and Transdisciplinary Integration in Team Science', *American Journal of Preventive Medicine*, 35, S151–60.

Merton, R. K. (1973) *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago press.

Metzger, N. and Zare, R. N. (1999) 'Interdisciplinary Research: From Belief to Reality', *American Association for the Advancement of Science, 283, 642–43.*

Mogoutov, A. and Kahane, B. (2007) 'Data Search Strategy for Science and Technology Emergence: A Scalable and Evolutionary Query for Nanotechnology Tracking', *Research Policy*, 36, 893–903.

Molas-Gallart, J. et al. (2016) 'Towards an Alternative Framework for the Evaluation of Translational Research Initiatives', *Research Evaluation*, 25, 235–43.

Mumford, M. D. et al. (1991) 'Process Analytic Models of Creative Capacities', *Creativity Research Journal*, 4, 91–122.

Nelson, R. R. and Winter, S. G. (1982) *An Evolutionary Theory of Economic Change*. Cambridge: Harvard University Press.

Petsko, G. A. (2012) 'Goodbye, Columbus', *Genome Biology*, 13, 155.

Pistorius, C. W. and Utterback, J. M. (1997) 'Multi-Mode Interaction among Technologies', *Research Policy*, 26, 67–84.

Porter, A. and Rafols, I. (2009) 'Is Science Becoming More Interdisciplinary? Measuring and Mapping Six Research Fields over Time', *Scientometrics*, 81, 719–45.

Porter, A. L. and Rossini, F. A. (1985) 'Peer Review of Interdisciplinary Research Proposals', *Science, Technology, & Human Values*, 10, 33–8.

Porter, A. L. et al. (2002) 'Measuring National 'Emerging Technology'capabilities', *Science and Public Policy*, 29, 189–200.

Porter, A. L. et al. (2006) 'Interdisciplinary Research: Meaning, Metrics and Nurture', *Research Evaluation*, 15, 187–95.

Porter, A. L. et al. (2008) 'Refining Search Terms for Nanotechnology', *Journal of Nanoparticle Research*, 10, 715–28.

Porter, A. L. et al. (2019) 'Emergence Scoring to Identify Frontier R&D Topics and Key Players', *Technological Forecasting and Social Change*, 146, 628–43.

Porter, A. L., Roessner, D. J., and Heberger, A. E. (2008) 'How Interdisciplinary is a Given Body of Research?', *Research Evaluation*, 17, 273–82.

Price, D. J. (1963) *Little Science, Big Science... and beyond*. New York: Columbia University Press.

Romer, P. M. (1994) 'The Origins of Endogenous Growth', *Journal of Economic Perspectives*, 8, 3–22.

Rotolo, D., Hicks, D., and Martin, B. R. (2015) 'What is an Emerging Technology? ', *Research Policy*, 44, 1827–43.

Schumpeter, J. A. (1961) *Theory of Economic Development*. New York: 3rd Edition, Oxford University Press.

Shapira, P., Kwon, S., and Youtie, J. (2017) 'Tracking the Emergence of Synthetic Biology', *Scientometrics*, 112, 1439–69.

Simon, H. A. (1977) '*Scientific Discovery and the Psychology of Problem Solving', in Models of Discovery,* 286–303. Dordrecht: Springer.

Simon, H. A., Langley, P. W., and Bradshaw, G. L. (1981) 'Scientific Discovery as Problem Solving', *Synthese*, 47, 1–27.

Small, H., Boyack, K. W., and Klavans, R. (2014) 'Identifying Emerging Topics in Science and Technology', *Research Policy*, 43, 1450–67.

Stephan, P., Veugelers, R., and Wang, J. (2017) 'Reviewers Are Blinkered by Bibliometrics', *Nature News*, 544, 411.

Stirling, A. (2007) ‘A General Framework for Analysing Diversity in Science, Technology and Society’, *Journal of the Royal Society Interface*, 4, 707–19.
Strumsky, D. and Lobo, J. (2015) ‘Identifying the Sources of Technological Novelty in the Process of Invention’, *Research Policy*, 44, 1445–61.
Uzzi, B. et al. (2013) ‘Atypical Combinations and Scientific Impact’, *Science*, 342, 468–72.
Van Raan, A. F. (2004) ‘Sleeping Beauties in Science’, *Scientometrics*, 59, 467–72.
Van Raan, A. F. and Van der Velde, J. (1991) ‘Bibliometric Mapping and Early Warning for Emerging Topics: Exploring Developments in Materials Science’, *Research Evaluation*, 1, 115–27.
Vogel, A. L. et al. (2013) ‘The Team Science Toolkit: Enhancing Research Collaboration through Online Knowledge Sharing’, *American Journal of Preventive Medicine*, 45, 787–9.
Wagner, C. S. et al. (2011) ‘Approaches to Understanding and Measuring Interdisciplinary Scientific Research (IDR): a Review of the Literature’, *Journal of Informetrics*, 5, 14–26.
Wagner, C. S., Whetsell, T. A., and Mukherjee, S. (2019) ‘International Research Collaboration: Novelty, Conventionality, and Atypicality in Knowledge Recombination’, *Research Policy*, 48, 1260–70.
Wang, J., Veugelers, R., and Stephan, P. (2017) ‘Bias against Novelty in Science: A Cautionary Tale for Users of Bibliometric Indicators’, *Research Policy*, 46, 1416–36.
Wang, Z. et al. (2019) ‘Updating a Search Strategy to Track Emerging Nanotechnologies’, *Journal of Nanoparticle Research*, 21, 199.
Weitzman, M. L. (1998) ‘Recombinant Growth’, *The Quarterly Journal of Economics*, 113, 331–60.
Wuchty, S., Jones, B. F., and Uzzi, B. (2007) ‘The Increasing Dominance of Teams in Production of Knowledge’, *Science*, 316, 1036–9.
Youtie, J. et al. (2017) *Autonomous Systems: A Bibliometric and Patent Analysis (No. 14-2018)*. Studien zum deutschen Innovationssystem.
Zhou, X. et al. (2014) *Nano-Enabled Drug Delivery: A Research Profile*.Nanomedicine: Nanotechnology, Biology and Medicine 10, pp. e889–96.