The Right Job and the Job Right: Novelty, Impact and Journal Stratification in Science

Nicolas Carayol  
University of Bordeaux  
GREThA, UMR CNRS 5113  
nicolas.carayol@u-bordeaux.fr

Oscar Llopis  
Rennes School of Business  
Rennes School of Business  
oscar.llopis-corcoles@rennes-sb.com

Agenor Lahatte  
HCERES  
Observatoire des Sciences et Techniques  
agenor.lahatte@hceres.fr

Abstract

We introduce a new measurement of novelty based on the frequencies of pairwise combinations of article keywords. On the set of all research articles published from 1999 to 2013 in the journals referenced by the WoS (more than ten million papers), we find no evidence of shrinking novelty in science over that period. Novel contributions are more often performed in larger teams that span more institutional boundaries and geographic areas. High novelty increases both citations and the odds of a "big hit" by more than forty percent but individual level returns to novelty may be low. In such a rapidly evolving environment, novelty is a lever for decreasing risk (not the reverse) as it increases the probability of addressing a problem that remains active in the future. We document that top journals play a very significant role in sustaining novelty in science.
1 Introduction

Looking backward, the old times of science are peopled with creative fellows such as the “renaissance man” (Jones, 2009), the “gentlemanly specialist” (Rudwick, 1985), the English “amateur scientist” (Shapin, 2008) or the French “savant”. These celebrated characters seem to have gone for ever. Today is a time for communities of professional scientists strongly subsidized by the states and organized through formal peer reviewed vetting procedures for recruitment, funding and publishing. A time of “big science” (Price, 1963) whose outcomes still double every ten to twenty years (Price, 1961; Olesen Larsen and von Ins, 2010), with increasing team size (Jones, Wuchty and Uzzi, 2008) and rising specialization and knowledge complexity (Jones, 2009). Though such a massification and professionalization phenomenon is still contemporaneous, do scientific communities maintain their standards of creativity and originality, which are crucial dimensions of the scientific ethos (Merton, 1942)? Who performs new research in today’s science, and does it still pay to support or to engage in newer research? Is it risky and which outlets are more likely to publish novel papers? The present paper provides answers to these core questions.

Many observers have recently expressed worries that originality and creativity could be under threat in science (Heinze et al., 2009). In particular, there is rising concern that the professional peer review procedures for funding and publishing are in fact not providing sufficient incentives to sustain innovative research. The peer review system is often criticized for its bias against groundbreaking and innovative research (Braben, 2004; Chubin and Hackett, 1990; Wesseley, 1998). A number of scholars have suggested that academic audiences reject novel contributions when they diverge too much from the dominant canon (Trapido, 2015; Shadish et al., 1995). Kolata (2009) quotes a past acting director of the NIH who, after noting that the review system for grant proposals “works over all pretty well, and is very good at ruling out bad things”, makes the point that the “system provides disincentives to funding really transformative research.”

In this article, we propose a new measurement of novelty based on the frequency of pairwise combinations of Author Keywords that we apply to the ten million research articles published over 1999-2013 by journals indexed in the Web of Science (WoS).  

\footnote{In fact, such tension between novelty and conservatism is not new, though as once famously stressed by Max Planck: “a new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it” (Planck, 1950).}

\footnote{Data come from the full WoS database provided by ISI Thomson Reuters, maintained and enriched in house at the OST-HCERES. We restrict the study to the post-1999 period which is fully controlled and normalized according to the most advanced scientometric standards.}
sustaining its standards of creativity and originality. We look at the relation between team characteristics and novelty. Through several dimensions, we investigate the returns of novel research in terms of forward citations. A specific attention is devoted to the role of journals in the publication of novel research.

2 Measuring Novelty in Science

This section introduces a new indicator of novelty based on author keyword combinations. In the first subsection, we discuss how novelty has been previously defined and empirically captured by the literature, and expose our forward looking notion of novelty and its associated index. In the second subsection, we study the behavior of the index on our data, comparing it to the one of neighboring indexes.

2.1 An approach based on Pairwise Author Keyword Combinations

From Atypical Combinations to New Research Questions Henri Poincaré, an exceptionally productive and creative mathematician, first introduced the idea that invention in mathematics proceeds from re-combinations of distinct pre-existing ideas (“mathematical entities”) in one’s mind (Poincaré, 1910). Weitzman (1998) proposes a mechanism for the growth of ideas in the economy that results from binary random combinations of existing ideas. Building on this idea, Uzzi et al. (2013) employs pairwise journal co-citations in reference lists of articles to identify recombinations of previous knowledge. The degree of “atypicality” or conventionality of those re-combinations are computed through their frequencies of occurrence over the whole period. The authors find that high-impact articles are more likely to combine infrequent pairwise combinations of journal references with conventional ones.

Atypical combinations are not all fruitful, however. Creativity, as Poincaré himself also argued, “consists precisely in not making useless combinations and in making those which are useful and which are only a small minority. Invention is discernment, choice” (Poincaré, 1910, p. 325). In this process, “the role of the preliminary conscious work [...] is evidently to mobilize certain of these [pre-existing mathematical entities], to unhook them from the wall and put them in swing... our will did not choose them at random; it pursued a perfectly determined aim” (Poincaré, 1910, p. 333-334). Even though serendipity,

3This view is also reminiscent of the concept of “abduction” developed by Charles S. Peirce, which is a form of inference capable of generating new ideas as it proceeds from the effects to the causes, in the
intuition or chance are obvious factors of breakthrough, fruitful combinations result from intentional exploration behaviors (in contrast to exploitation, March, 1991), a form of “tinkering” (a term popularized by Jacob, 1977), that allows researchers to intentionally address new scientific problems or questions.

To identify the scientific questions of research articles (the very problem they address) we use the keywords given by the authors themselves, the ones they freely chose to describe their contribution. We suggest that pairwise keyword combinations capture the different “angles” of a scientific paper, to employ a notion introduced by Jacob: “scientific advances often come from uncovering a hitherto unseen aspect of things as a result, not so much of using some new instrument, but rather of looking at objects from a different angle” (Jacob, 1977, p. 1161). We focus on the most infrequent pairwise combination of keywords of the paper as capturing its novelty, its most original “angle”. 4

**Pairwise Keyword Novelty Index** We consider all pairwise keywords combinations by papers published in a given year and research field. Research fields are identified by the subject categories, according to the classification scheme developed by ISI which assigns scientific journals to at least one of the 251 subject categories (and potentially to several ones). Keyword combination frequencies are calculated within subject categories since the degree of novelty of a publication is likely to be interpreted within a given field or community, not across fields. Moreover, some terms can be interpreted differently across communities.

Formally, the commonness of the combination of keywords $i$ and $j$, in subject category $c$ and year $t$ is computed as follows:

$$Com_{ijct} = \frac{N_{ijct}/N_{ct}}{N_{ict}/N_{ct} \times N_{jct}/N_{ct}} = \frac{N_{ijct} \times N_{ct}}{N_{ict} \times N_{jct}},$$

with $N_{ct}$ the number of (non-distinct) keyword combinations in papers published in subject category $c$ and year $t$. The terms $N_{ict}$, $N_{jct}$ and $N_{ijct}$ give the number of articles which use respectively keyword $i$, keyword $j$, and both keywords $i$ and $j$. 5 Equation (1) manifests itself simply as the share of keyword pairwise combinations that use $i$ and $j$ in the domain $c$, divided by the expected share of such pairs given the number of times keywords $i$ and $j$ are used in $c$.

---

4 In fact, we use (see below in this section) a combination of the ninetieth percentile (within subject categories) and of the maximum (across subject categories). However, we show that variations in a number of dimensions on the precise way of computing the novelty index makes no difference in the results.

5 Some author keywords may be misreported. Since most of the errors are very rare, we drop all keywords that appear in only one document, and thus all combinations implying at least one of these keywords.
When articles have keywords, they are likely to have more than two, which is the minimal possible value for pairwise keyword frequencies to be calculated. For a given article, let $K$ denote its set of distinct (unordered) pairs of keywords and $C$ its set of associated subject categories. We focus on the value of the 10th percentile of the distribution of pairwise commonness values, for each paper in subject category $c \in C$:

$$com_c = 10thPercentile(Com_{ij} | \forall ij \in K).$$

(2)

We use the tenth percentile because it avoids extreme value problems we would have encountered had we taken the minimum. But the underlying idea is very similar to taking the minimum. We want to select the less common combination of keywords, that is the most original “angle” of the paper, in each reference corpus (subject category). The reference to the publication year $t$ drops out as it is unambiguous and unnecessary for what follows.

We then use the inverse logarithmic transformation of commonness to obtain the novelty of a given paper in a given subject category $c$:

$$nov_c = -\log(com_c)$$

(3)

Since journals, and thus articles, may be attached to multiple subject categories, we attribute the maximal novelty over all associated subject categories:

$$nov = \max_{c \in C}(nov_c).$$

(4)

The field of research in which the novelty is to be considered is the one in which it is found to be the most original.

The definition of any quantitative indicator imposes computational choices that balance various goals (robustness, clearness, simplicity,...) beyond effectively capturing the phenomena that it intends to measure. To further validate our indicator, we however consider a number of alternative specifications for Equations (1) to (4) of our indicator.$^6$

A first variant avoids making across subject category maximization of novelty. A second variant takes the minimum value of the distribution of keyword pairs frequencies in a given paper instead of the tenth percentile. Third, the computation of pairwise keyword frequencies is based only on the pool of papers that were published in the same year as the focal paper (publication year $t$). This way of proceeding is appealing as it is simple and easier to compute. Of course, novelty needs be defined with respect to what has been done in the past as well. Taking the past into consideration adds in complexity while it is not likely to add variation in the data. To prove this statement, we compute a variant of our indicator that considers a backward time window ($t - 2$ to $t$) to assess the frequency of pairs of keywords. We show below in the paper, that the mains results are robust to substituting any of those variants to the main pairwise keyword novelty indicator.

---

$^6$Detailed explanations of the variants specifications are presented in the Online Appendix.
2.2 Data, Pairwise Author Keywords Novelty Index and Benchmarks

Data and the Behavior of the Pairwise Keywords Novelty Index  Our dataset groups together all research articles published from 1999 to 2013 and indexed in the Web of Science (WoS). Therefore, review papers, letters and conference proceedings in particular are not considered. We use Author Keywords – this important choice is discussed in detail below in the present section. As our novelty indicator is based on keyword combinations, we exclude all papers with less than two Author Keywords. The sample comprises 10,229,644 research articles. These papers are classified in three major research areas: humanities and social sciences (7%), life sciences (47%) and hard sciences and engineering (46%).

We display, in Figure 1–Graph (a), the distribution of pairwise keyword novelty of the articles. The distribution is “well-shaped”, slightly asymmetric. Field differences are minor: the distribution is slightly sharper for the sub-sample of papers in humanities and social sciences and the one of life sciences. The novelty of Hard sciences and engineering papers is slightly both less concentrated and lower.

For a given paper, as the number of keywords grows, so does the number of possible pairwise keyword combinations for purely mathematical reasons. As our novelty indicator is defined as the 90th percentile novelty among all pairwise combinations of keywords (see Equation 2) which often coincides with the max, a positive correlation between novelty and the number of keywords is expected. This is verified on the data as the correlation coefficient between pairwise keyword combinations novelty and the number of keywords equals .37. All correlation coefficients are presented in Table 1. Figure ?? in the Appendix shows however that, if pairwise keyword novelty is positively related to the number of keywords from the first to the third quartile, above the third quartile, increasing further the number of keywords no longer correlates with changes in novelty. Fields differences are not behind such results as they extend to within-fields analyses.

Is there a relation between the degree of novelty of keywords themselves and the degree of novelty of pairwise keyword combinations? Phrased differently, are articles using newer keywords also more likely to use newer pairwise keyword combinations, are keyword novelty and pairwise keyword novelty substitutes, or are they unrelated? To answer this question, we create a keyword novelty indicator (see Appendix) as study its correlation with pairwise keyword novelty. We find a significant and negative correlation (−.23, Table ??) which applies only for articles having low levels of keyword novelty. This negative relation is much less clear with papers that have higher keyword novelty (after the fourth decile), and, when present, seems mainly driven by the hard sciences and engineering.
Figure 1 – Distributions of Pairwise Author Keywords Novelty, of Pairwise Keywords Plus Novelty, of Journal References Novelty, and of the number of Author Keywords Pairs and KeyWords Plus Pairs.

(a) Pairwise Author Keyword Novelty
(b) Pairwise KeyWord Plus Novelty
(c) Pairwise Journal Reference Novelty
(d) Author Keyword and KeyWord Plus Pairs

Notes: Kernel density plots. Based on the set of all research articles published in journals indexed by the WoS over period 1999-2013. The benchmark indicator of novelty is based on the atypicality of journal references combinations in each publication year. The precise computation of this indicator is presented in the Online Appendix.

These results support the idea that pairwise combinations of keywords provide a much more sophisticated indicator of novelty, distinct from keyword novelty.

We also wonder whether our novelty index could be correlated with the research field size. Two alternative proxies for field size are employed: i) the number of articles that were published in each subject category (Figure ??, graph a) and ii) the number of articles published in the same journal (Figure ??, graph b). In both cases, results suggest the existence of a slightly positive relation between field size and pairwise keywords novelty, which is however very limited and mainly explained by field-level differences.
Table 1 – Correlation Table of Various Novelty Indicators and Other Variables.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Pairwise Author Keyword Novelty</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Pairwise Author Keyword Novelty (top-10%)</td>
<td>.49</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Pairwise Author Keyword Novelty (top-5%)</td>
<td>.37</td>
<td>.69</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Pairwise Author Keyword Novelty (top-1%)</td>
<td>.19</td>
<td>.30</td>
<td>.44</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Pairwise KeyWord Plus Novelty</td>
<td>.22</td>
<td>.13</td>
<td>.10</td>
<td>.05</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Pairwise KeyWord Plus Novelty (top-10%)</td>
<td>-.10</td>
<td>-.04</td>
<td>-.03</td>
<td>-.01</td>
<td>.35</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Pairwise KeyWord Plus Novelty (top-5%)</td>
<td>-.13</td>
<td>-.07</td>
<td>-.05</td>
<td>-.02</td>
<td>.26</td>
<td>.91</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Pairwise KeyWord Plus Novelty (top-1%)</td>
<td>-.16</td>
<td>-.09</td>
<td>-.07</td>
<td>-.03</td>
<td>.14</td>
<td>.84</td>
<td>.92</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Pairwise Journal Reference Novelty</td>
<td>.12</td>
<td>.09</td>
<td>.07</td>
<td>.03</td>
<td>.16</td>
<td>-.20</td>
<td>-.24</td>
<td>-.29</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Pairwise Journal Reference Novelty (top-10%)</td>
<td>-.10</td>
<td>-.06</td>
<td>-.05</td>
<td>-.03</td>
<td>-.08</td>
<td>.26</td>
<td>.29</td>
<td>.31</td>
<td>.44</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Pairwise Journal Reference Novelty (top-5%)</td>
<td>-.11</td>
<td>-.07</td>
<td>-.06</td>
<td>-.03</td>
<td>-.10</td>
<td>.29</td>
<td>.33</td>
<td>.36</td>
<td>.35</td>
<td>.92</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 Pairwise Journal Reference Novelty (top-1%)</td>
<td>-.12</td>
<td>-.08</td>
<td>-.06</td>
<td>-.03</td>
<td>-.11</td>
<td>.33</td>
<td>.36</td>
<td>.40</td>
<td>.19</td>
<td>.85</td>
<td>.93</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 Author Keyword Novelty</td>
<td>-.23</td>
<td>-.17</td>
<td>-.14</td>
<td>-.09</td>
<td>.01</td>
<td>.02</td>
<td>.01</td>
<td>.00</td>
<td>.12</td>
<td>-.06</td>
<td>-.08</td>
<td>-.10</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 KeyWord Plus Novelty</td>
<td>.01</td>
<td>.02</td>
<td>.01</td>
<td>.01</td>
<td>-.04</td>
<td>-.07</td>
<td>-.09</td>
<td>-.06</td>
<td>.30</td>
<td>-.05</td>
<td>-.10</td>
<td>-.14</td>
<td>.41</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 Number of Author Keywords</td>
<td>.37</td>
<td>.09</td>
<td>.06</td>
<td>.01</td>
<td>-.03</td>
<td>-.11</td>
<td>-.12</td>
<td>-.12</td>
<td>.10</td>
<td>-.06</td>
<td>-.08</td>
<td>-.09</td>
<td>.11</td>
<td>.08</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 Field Size (Catcode)</td>
<td>.05</td>
<td>.07</td>
<td>.07</td>
<td>.04</td>
<td>.09</td>
<td>-.03</td>
<td>-.05</td>
<td>-.08</td>
<td>.14</td>
<td>-.03</td>
<td>-.05</td>
<td>-.06</td>
<td>.39</td>
<td>.35</td>
<td>.01</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>17 Field Size (Journal)</td>
<td>-.02</td>
<td>-.01</td>
<td>-.00</td>
<td>-.00</td>
<td>-.02</td>
<td>.10</td>
<td>.11</td>
<td>.11</td>
<td>-.00</td>
<td>.03</td>
<td>.04</td>
<td>.04</td>
<td>.09</td>
<td>.02</td>
<td>-.01</td>
<td>.08</td>
<td>1</td>
</tr>
</tbody>
</table>

Notes: All correlation coefficients are significant at \( p < 0.01 \). Calculated on the sample of observations having non-missing values in any of variables (4,397,918 observations).

**Keywords Chosen by the Authors vs. Assigned Keywords** The Web of Science provides two different types of keywords: KeyWords Plus and Author Keywords. KeyWords Plus are assigned to articles by a computer algorithm which extracts words from the titles of the reference list articles (Garfield and Sher, 1993). Author Keywords are given by the authors themselves when they submit their paper. Both types of keywords have been previously employed to identify research trends (Li et al., 2009). Azoulay, Graff Zivin and Manso (2011) and Boudreau et al. (2016) are reliant on a set of predefined keywords provided by Medline database to identify research themes and their novelty, namely the MESH Keywords of the National Library of Medicine’s controlled vocabulary thesaurus. Assigned keywords, such as KeyWords Plus or MESH Keywords, have the advantage of being less sensitive to strategic choices of the authors. We however refrain from using such externally defined categories as keywords provided by the authors themselves are more likely to precisely capture the original idea their paper carries to the community. Comparing the use of KeyWords Plus and Author Keywords a specific research field, Zhang et al. (2016) argue that KeyWords Plus more often describe methods and techniques whereas Author Keywords are more comprehensive in representing article content.

Throughout this paper, we document that using keywords chosen by authors instead of predefined keywords is important to capture novelty in the sense of Poincaré and Jacob. For this purpose, all indicators are also computed using KeyWords Plus instead of Author Keywords. For now, in this section, we try to better understand how the two differentiate. In the next sections, we will show to what extent results are reliant on choosing Author Keywords.
Keywords instead of KeyWords Plus. Note first that author keyword novelty and KeyWord Plus novelty are positively correlated (.41, see Table 1). The originality of authors’ choices is more apparent in the way they associate keywords. Pairwise novelty indexes based on Author Keywords and on KeyWords Plus are positively correlated to a more limited extent (.22). Interestingly, this correlation reverses when focusing on top novel pairs of keywords. The dummy of top 10% pairwise KeyWords Plus novelty is negatively correlated with pairwise author keywords novelty (−.10) as well as with the top 10% dummy of this variable (−.04). The negative relation rises when increasingly focusing on top novel papers (top 5%, top 1%), for both types of keywords. This tells us that top novel papers relying on Author Keyword pairs frequencies and according to KeyWord Plus pairs frequencies are likely to NOT be the same.

How do the patterns of choosing and associating keywords by authors differ from the ones of the algorithm allocating KeyWords Plus. In a given year, there are in average .93 distinct Authors Keywords per document and 7.26 distinct pairs. This compares with .63 distinct KeyWords Plus and 14.88 distinct KeyWord Plus pairs per document. Authors’ choices of keywords are more idiosyncratic while their pairwise association of keywords is less heterogeneous. This can be explained by purely mathematical reasons, as the number of KeyWord Plus pairs per paper in much larger than the number of Author Keyword pairs. There are in average 27 pairs of KeyWords Plus per paper, vs. 8.8 Authors Keyword pairs per paper. Figure 1–Graph (c) exposes the distributions of articles with respect to their number of Author Keyword pairs and KeyWord Plus pairs. About forty percent of papers have (exactly) 45 pairs of KeyWord Plus (that is 10 KeyWords Plus), against 80% of articles having less than 10 pairs of Author Keywords (that is 5 Author Keywords). As KeyWords Plus pairs are more numerous on the average paper, they are more likely to be original (in mathematical terms, that is less frequent). A large number of keyword pairs may introduce noise and decrease precision as some of those pairs may be very original for very artificial reasons. The lack of precision is likely to be exacerbated by the shape of the distribution of pairwise KeyWord Plus novelty. Figure 1–Graph (b) shows that it is much more concentrated than pairwise author keyword novelty, with a particularly sharp decline just above the mode, thus rendering small differences decisive to be included in the sets of top novel articles. That may lead to an even noisier identification of top novel articles. Author Keywords are more original though they are less numerous in the average article. Their pairing is thus also much more precise and likely to better characterize the research direction of the paper.

**Pairwise Journal References Novelty Benchmark** We believe an approach based on the frequency of keyword combinations permits a more direct measurement of scientific novelty than one based on journal reference combinations and is more closely aligned
with neighboring concepts, such as creativity or knowledge exploration. However, as previous literature has mainly computed the novelty of articles using the frequency of journal reference combinations, we also compute a benchmark indicator of novelty based on pairwise journal reference combinations. This indicator is, like ours, time-variant, and close in spirit to the indicator developed by Lee, Walsh and Wang (2015) based on the atypicality of journal reference combinations in year $t$. The precise computation of this indicator is presented in the Online Appendix.

A positive correlation is expected between a pairwise keyword combinations indicator and alternative indicators based on pairwise journal co-citations (Uzzi et al., 2013; Lee, Walsh and Wang, 2015; Wang, Veugelers and Stephan, 2016; Stephan, Veugelers and Wang, 2017), because investigating new research questions may often require combining pre-existing pieces of knowledge in new ways. Surprisingly, the correlation between our novelty indicator and the indicator based on journal references appears to be positive but rather small (.12, see Table 1). Pairwise author keywords novelty increases with pairwise journal reference novelty essentially for the lowest deciles of that indicator. The relation is weak and non monotone for all other deciles. This implies that highly novel papers are often not the same depending on which indicator is used. This relationship is confirmed by the correlation analysis involving top journal pairwise novel articles. We find that all such variables (top 10%, 5%, 1%) are negatively correlated with pairwise author keywords novelty indexes. As for pairwise KeyWords Plus novelty, this indicator is sharply distributed around its mode as it is apparent from Figure ??–Graph (c). It is interesting to observe that top novel pairwise KeyWords Plus novelty and top pairwise journal reference novelty are positively correlated, which suggests these dummies capture close phenomena, which are quite distinct from pairwise author keywords novelty, as both are negatively correlated with this indicator.

3 The Evolution, Distribution and Contexts of Novelty in Science

Before examining the evolution of the novelty of research articles using Equations (1)–(4), we first look at the expansion of scientific inquiry through the evolution of distinct keyword combinations. We then investigate the relation between novelty and teams.

7Note that this indicator is subject to significant field differences that render it much less appealing.
3.1 The Expansion of “Scientific Inquiry”

There has been growing concern that, though scientific production is globally increasing, the degree of creativity embedded in published research may be actually shrinking. We document these dynamics by looking at the relation between the growth in the number of research articles and the growth in the explored “knowledge space”, approximated by the number of distinct keyword combinations that are used each year. In years 1999 and 2013, about five and fourteen millions distinct keyword combinations are respectively employed, while approximately fifteen times fewer distinct keywords are used and research papers are produced. As the orders of magnitude of the different variables are so different, year 1999 is taken as reference in the graphs of Figure 2 so that our analysis focuses on growth rates.

In the left plot we see that the number of distinct keyword combinations follows a very similar growth pattern to the number of research articles. Both increase steadily (in log scale) to reach overall growth of 290% over the thirteen years considered (which corresponds to an average 7.9% yearly growth rate). Note that as the vertical axis is in log scale, a straight line is consistent with a constant growth rate (exponential growth). These growth patterns contrast with the one of the number of keywords, which exhibits a decreasing growth rate over the period (less than linear growth in log scale).

The right plot of Figure 2 allows us to look at the expansion of scientific knowledge from a slightly different perspective, that is, with respect to the number of research articles – the horizontal axis. The number of keyword combinations used each year relative to year 1999 level very closely follows the straight line of the number of research articles (also relative to the number in year 1999). This confirms that, according to our indicator of novelty, the scientific community does not exhibit a decreasing rate of creativity, though the yearly production of knowledge increased exponentially over the thirteen years under investigation in our database of research articles.

This graph also traces the evolution of the number of possible pairwise combinations of keywords, calculated as half the square of the number of distinct keywords. Linear growth is also observed for the number of possible combinations. If we interpret the number of possible keyword combinations as prospective combinations that the scientific community could explore, we see that this exploration space increases at a rate nearly twice the growth rate in the combinations that are eventually explored. This is consistent with the idea that, as knowledge accumulates and as the frontier expands, the potentialities for exploration increase at a significantly greater rate.

The number of keywords is growing sub-linearly with the number of articles. This can be interpreted as a supplementary argument supporting the idea that pairwise keyword
combinations provide a better proxy to reflect the exploration of new scientific areas, as compared to the consideration of keywords only.

**Figure 2 – The Expansion of Scientific Knowledge**

(a) Number of Distinct Keyword Combinations, Keywords, and Research Articles over the period 1999-2013.

(b) Number of Distinct Possible Keyword Combinations, Keyword Combinations and Keywords with Respect to the Number of Research Articles.

Note: Based on the set of all research articles published in journals indexed by the WoS over period 1999-2013. The number of documents reached in year 2003 and year 2010 are given by the vertical red lines. The year 1999 is taken as reference for both graphs.

### 3.2 The Evolution of Novelty at the Article Level

Most article-level calculations use data for the 1999-2011 period because later we will need a sufficient time lag to collect three-year forward citations of publications and we want to remain fully consistent throughout the paper. This data set contains 7,896,301 articles which received more than 26 million citations.⁸

Figure 3a depicts the evolution of mean article novelty for the full sample of publications and by field of science. Mean novelty is pretty stable, very moderately increasing over the complete period for the full sample. This overall stability hides a slightly more pronounced increase in the field of life sciences. However, such variations still remain very limited (between 1% and 2% of variation over the whole period). It is interesting that, in times of major changes in the way scientists set up their research agendas, the degree of novelty remains stable. This is fully consistent with our results displayed above showing that keyword combinations and articles grow at the same rate.

⁸Note that all analyzes and regressions are also performed using a five-year window of forward citations that thus need to restrict to the 1999-2009 period (less than seven million papers).
Life sciences exhibit the highest degree of novelty, and hard sciences and engineering the lowest degree of novelty. This confirms the intuition that life sciences explore a higher and more diverse number of problems while hard sciences focus on a smaller number of problems.\footnote{Similar cross-field differences have been also found in recent studies on interdisciplinarity. For instance, Millar and Dillman (2012) found that life sciences have the highest proportion of interdisciplinary dissertations, and Leahey, Beckman and Stanko (2017) found that biotechnology and medicine were the most interdisciplinary fields.} When we look in more detail at novelty by discipline (Figure 3b) we find that there is a group of disciplines composed of medicine, humanities, fundamental biology, social sciences, chemistry and sciences of the universe that are characterized by a higher level of novelty. Then comes applied biology, followed by physics and engineering which are very close, and then math. Some disciplines are characterized by increasing novelty, such as medicine, chemistry or physics, while others follow a decreasing path such as the sciences of the universe or the humanities. Absolute differences across disciplines should be taken with caution as they may be influenced by social norms.

Figure 3 – The Evolution of Pairwise Keyword Novelty

(a) For All Fields and by Large Field of Science

(b) For the Ten Large Scientific Disciplines

Notes: Based on the set of all research articles published in journals indexed by the WoS, over period 1999-2011. Fractional polynomial estimates and 90\% confidence intervals.

The time evolution of the benchmark indicator depicts a very different pattern than what we have found for our novelty indicator. Reference novelty increases from 1999 to 2011 by 50\%, a sharp variation which seems hard to justify. Such variation concerns all large fields of science, and is more marked for social sciences and humanities, and engineering; and less marked for life sciences. Looking at the degree of novelty by scientific discipline (Figure ??b), we find significant differences from what we found for our indicator based on pairwise keyword novelty. Physics is for instance found to be significantly more novel than the sciences of the universe, the engineering sciences, the social sciences and
the humanities. This does not fit with the intuition that physics, as an older science, deals with a more limited number of problems and hence should exhibit lower levels of novelty.

3.3 Which Teams Produce More Novel Papers?

There has been an increasing interest in exploring the factors behind breakthrough scientific contributions. Heinze et al. (2009) looked in particular at the institutional factors. Jones, Wuchty and Uzzi (2008) and Jones (2009) connect radical scientific contributions to age. However, the relation between team composition and breakthrough science has been under-investigated (an exception is Lee, Walsh and Wang, 2015). With respect to scientific teams, we know that their size is increasing over time, as well as their institutional and geographical span (Adams et al., 2005; Wuchty, Jones and Uzzi, 2007; Jones, Wuchty and Uzzi, 2008; Adams, 2013). The literature mainly explains these evolutions by the falling costs of remote collaboration (thanks to the www and new information technologies) or as a response to knowledge specialization (Jones, Wuchty and Uzzi, 2008). Adams (2013) argues that there is a rising stratification in collaborations, in the sense that the best universities develop longer-range and higher-quality collaborations.

We wonder whether, to handle and solve newer problems, scientists are more likely to assemble in larger (vs. smaller) teams or to form cross-institutional (vs. within closed walls) collaborations. Our data show (see Figure 4) that novelty increases with team size (approximated by the number of co-authors). This is consistent across all scientific fields, with the exception of social sciences and humanities, where there is a decreasing relation after five co-authors. We also find similar results for the number of distinct institutions. Further, novelty increases when the teams involve members that are located in different world regions. All these results support the idea that “break things and think different” is more frequent in larger and more dispersed teams. This could be due either to a “diversity effect” or to a selection effect (better teams are more likely to span boundaries) that can not be easily disentangled – a natural experiment would be in order. We also find that novelty is higher when articles involve co-authors from North America and, to a lesser extent, Europe. Interestingly, the gap between the novelty of North American teams and that of European teams has widened since 2005.

We perform a series of regressions to consider whether the correlations of the different team characteristics with our novelty indicator remain controlling for a number of dimensions. The number of keywords, keyword novelty, year dummies as well as sub-domain dummies are included as controls in logistic regressions on top novel paper dummies (top 10%, top 5% and top 1% dummies of pairwise keyword novelty). As the different characteristics of the teams are highly correlated, we do not include the number of authors,
Notes: Based on the set of all research articles published in journals indexed by the WoS, over period 1999-2011. Fractional polynomial estimates and 90% confidence intervals.

the number of distinct institutions, and the number of word regions simultaneously in the regressions. These variables are included one at the time, together with with all controls and with world region dummies.

Results can be found on the Online Appendix. They confirm the positive effect of the number of distinct institutions on the odds of a paper to be highly novel. On average, an additional institution participating in a paper increases the odds of a paper being a top 10% highly novel by 5.5%. This remains positive and significant when we split our sample by scientific domains and use more restrictive dependent variables (top 5% and top 1%). The positive relation between the number of co-authors on the odds of a paper to be highly novel is preserved as well. The effect is larger though, as an additional author, on average, increases the odds of a top 10% novel paper by 10%. Again, results remain consistent by scientific domains and in the different specifications of top novelty. Turning
to the role of the number of geographical regions, our multivariate analysis comes from logit regressions excluding geographical locations dummies that are strongly correlated to this variable. We find that an additional world region in the team increases the odds of a top novel paper by more than 20%. This is robust to the various definitions of top novelty and more pronounced in the hard sciences and engineering.

4 Novelty and Scientific Impact

This section examines whether novel articles are more cited (Subsection 4.1), and to what extent novelty can be conceived as a predictor of impact and excellence (Subsection 4.2). We are also interested in the future conventionality of pairwise keywords as a predictor of scientific impact and its complementarity/substituability with novelty (Subsection 4.3). We further tackle the individual rewards of novelty by introducing more controls in Subsection 4.4 and conclude the section by analyzing how novelty relates to the notion of risk (Subsection ??).

4.1 Are Novel Articles more Cited?

To explore the relationship between novelty and academic impact, we rank all papers according to their pairwise keyword novelty. For each centile of novelty, we calculate the average number of forward citations received over a three-year period.\textsuperscript{10} We observe in Figure 5 that citations increase significantly with novelty. A paper in the last centile of pairwise keyword novelty receives, on average, two to three times more citations than a paper in the first centile. This applies for each field of science taken separately, though to a lesser extent in the social sciences and humanities.

We use the percentile-based approach (Waltman and Schreiber, 2013) to define “big hit” papers, namely the ones which are among the top 10% most cited in their subject category and publication year. Similar work is done for the top 5% most cited articles. This allows us to focus on the most cited papers, knowing that the distribution of citations is very skewed: half of our research articles receive –within a three-year time frame– fewer than 2 citations while the top 10% most cited articles receive on average 11 citations more than the “mean” article.

We find (see Figure 5) that the proportion of papers categorized as “big hits” rises with the centiles of novelty. While the proportion of “big hit” papers is slightly over 6% for the lowest centiles of novelty, this rate rises up to 14% for the highest centiles of novelty.

\textsuperscript{10}A 5-year citation window is also employed as a robustness check.
(again, two to three times more). This result is quite robust across scientific fields. Only social sciences and humanities exhibit a somewhat non-linear pattern so that the share of top 10% papers tends to decrease slightly above the seventh to eighth deciles of novelty. Very similar results are found when we define a “big hit” article as a paper in the top 5% of the citation distribution or when looking at 5-year forward citations in the Online Appendix.

Even if these first results clearly support the idea that keyword combination novelty strongly correlates with citations (and in particular with high impact), it is not yet clear whether novelty is a predictor or lever of academic impact. We address these questions more directly in the following subsection.

4.2 Novelty as a Predictor of Impact and Excellence

Let’s consider the following purely conceptual experiment: a scientific team is picked randomly in a given discipline, and one wants to calculate the probability that its ongoing research becomes a “big hit”, conditional on being published in some journal referenced by the WoS. To what extent does a supplementary conditioning on high novelty of the paper increases future academic impact?

To answer this question, we have performed a series of regressions whose results are summarized in Table 2. Generalized negative binomial regressions allow us to estimate
the impact of high novelty (top 10% highest novelty) on citations.\footnote{In this part, we only discuss the main coefficients of the generalized negative binomial estimation. The determinants of the dependent variable dispersion are discussed in Subsection ??} We find a 38% impact in the 3-year window. Moreover, logit regressions show that picking a highly novel article increases the chances of a “big hit” (top 10% most cited articles in a 3-year window) by 42%. That effect is slightly higher in the hard sciences and engineering and in the life sciences, and lower in the humanities and social sciences (25%). Similar results hold when considering an additional specification of “big hit” (top 5% most cited). A slight increase in the odds of a “big hit” is observed when citations are considered over a time window enlarged to five years, but the correlation with the number of citations does not rise when citations are recorded over this larger time scale.

Table 2 – Predicting Citations and “big hit” Probabilities (Pairwise Keyword Novelty)

<table>
<thead>
<tr>
<th></th>
<th>“big hit” (top 10%)</th>
<th>“big hit” (top 5%)</th>
<th>Gen. Neg. Bin. (mean)</th>
<th>Gen. Neg. Bin. (ln(\alpha))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3-year</td>
<td>5-year</td>
<td>3-year</td>
<td>5-year</td>
</tr>
<tr>
<td>Full sample</td>
<td></td>
<td></td>
<td>42%</td>
<td>45%</td>
</tr>
<tr>
<td>Humanities and social sciences</td>
<td>25%</td>
<td>28%</td>
<td>25%</td>
<td>29%</td>
</tr>
<tr>
<td>Hard sciences and engineering</td>
<td>45%</td>
<td>48%</td>
<td>42%</td>
<td>46%</td>
</tr>
<tr>
<td>Life sciences</td>
<td>45%</td>
<td>48%</td>
<td>46%</td>
<td>48%</td>
</tr>
</tbody>
</table>

Notes: Based on the set of all research articles published in journals indexed by the WoS, over period 1999-2011 (for 3-year citation window regressions) and 1999-2009 (for 5-year citation window regressions).

Obtained from exponentiated coefficients in generalized negative binomial estimations and logistic regressions. Dependent variable for negative binomial regressions: number of forward citations (3-year and 5-year).

Dependent variable for logistic regressions: dummy taking the value 1 if the paper is a “big hit” in its field (“top 10%” or “top 5%”). Control variables: number of keywords, publication year dummies and disciplines dummies. All results are significant at the 0.1% level. Detailed regression results can be found in the Online Appendix.

All results have been replicated by employing three alternative variants of pairwise keyword novelty. Overall, the obtained results remain consistent. A detailed explanation of the three variants can be found in the Online Appendix.

Highly novel papers means being among the top 10% most novel papers. Results have been replicated by defining highly novel papers with different thresholds (top 5% and top 1%).

4.3 Novelty and the Crowd

Uzzi et al. (2013) highlighted that citation impact is significantly enhanced when a publication simultaneously couples “high median conventionality” with “high tail novelty”. Rather looking at conventionality and novelty simultaneously, but on different dimensions as they do, the time-variant nature of our novelty indicator allows us to conceive articles’ novelty and conventionality in a sequential manner on the same dimension. Our contention is that those papers that anticipate further interest in the very dimension on which they innovate, will have greater impact. In other words, breakthrough contributions in directions that keep attracting interest from colleagues are likely to be much more cited. On the contrary, novelty without a sustained future interest may pay relatively little in terms of citations. It is often argued, based on anecdotal evidence, that science has some
common features with Keynes’ “beauty contest” idea or with finance, in the sense that it would pay more to address problems in new areas of science which very soon become trendy.

Testing these conjectures with our data implies creating a new variable on future commonness. We create a dummy that equals one if the same keyword combination employed to assess novelty in period $t$ is still used by papers published in periods $t+1$ and $t+2$. Otherwise, it takes the value 0. For a paper to be considered as common in the near future, its most original “angle” at the time of publication needs to remain active in the two following years after it was published. From now onwards, commonness always refers to the two years following the publication year without us systematically recalling this time lag.

We identify four different scenarios to categorize each article in terms of novelty and commonness: highly-novel-and-common, highly-novel-and-not-common, not-highly-novel-and-common and, not-highly-novel-and-not-common. As previously, we employ two different regression methods for estimation, depending on the sort of dependent variable: i) a generalized negative binomial regression for models with a count-based dependent variable (number of publications after 3 years), and ii) a logistic regression for models with a percentile-based dummy dependent variable (“big hits”). Figure 6 summarizes our results on the estimated coefficients. The left-side graph (a) reports coefficient estimates of each considered dummy from generalized negative binomial models, and the right-side graph (b) is based on similar estimates from logistic regressions.12

We find that the most successful papers are those that are highly novel at the time of publication ($t$) and which are published in a field still active later on. Novel-today-and-common-later-on papers receive, on average, 61.7% (see left-side graph) more citations than papers located at the reference category (not-highly-novel-and-not-common). When turning to “big hit” papers, estimates results show a similar effect, although even more pronounced. The odds ratios of a paper being a top 10% “big hit” are 74.3% greater for those highly-novel-and-common papers. This rises up to 81.7% on top 5% “big hit” papers.13

The dummies’ coefficients are the largest in the field of hard sciences and engineering. Publishing a highly novel paper within a knowledge area that remains active two years after publication receives 70.4% more citations, and has 76.4% and 84.9% more chances of becoming top 10% and top 5% “big hit” papers respectively.

12Detailed regression results can be found in Tables ?? and ??.
13Robustness checks employing a 5-year citation frame and additional definitions of top papers were also performed and can be found in the Online Appendix. No qualitatively significant differences were found.
The papers in the two “mixed” categories, highly-novel-and-not-common and not-highly-novel-and-common papers have similar performances in terms of citations (approaching 40% more than the baseline). However, in terms of likelihood of becoming a “big hit”, the latter category performs significantly better (57.6% vs. 41.1% for top 10% papers and 64.3% vs. 40.9% for top 5% papers).

Overall, we can conclude that novelty without commonness increases citations and the probability of becoming a “big hit” by 40%. This effect is enlarged by 50% (up to 100% for “big hits”) when coupled with future commonness. However, it is important to note that the anticipation of future interest is key for citations in general, not specifically for more novel articles. If there is a “beauty contest” reward dimension in science, it is not specific to highly novel papers. In other words, we do not find any complementarity between high novelty and commonness as, jointly, they do not increase citations more than they do separately. Indeed, the sum of the coefficients of highly-novel-and-not-common and of not-highly-novel-and-common is always more than the coefficient of highly-novel-and-common. In a nutshell, novelty and commonness appear to be more substitutes than complementary in rising citations.\footnote{As a robustness check, we also performed a set of regressions including novelty, future commonness and the interaction between those two variables. Results (Tables ??, ?? and ??) confirm this point (the odds ratios of the interaction term are always below unity).}

4.4 Does it Pay to Address more Novel Research Questions?

Merton (1957) and Dasgupta and David (1994) described well how priority is key for distributing credit in the Open Science community. Because highly novel papers are more likely to open up new knowledge areas, we expect those papers to more directly influence subsequent work and to more often become “citation classics”. As we know scientific credit and (monetary and non-monetary) rewards correlate with citations (Garfield, 1984; Gomez-Mejia and Balkin, 1992; Evans, 2008; Lynn, 2014), we are, from a normative point of view, very interested in novel articles being more cited as a sign of preserved incentives to engage in novel research.

We thus would like to better appreciate the direct individual rewards of performing novel research in terms of forward citations. A neat identification of such returns is clearly beyond the scope of this paper, mainly because article quality is not observable. Indeed, as article quality is likely to be positively correlated with novelty and to raise citations, the impact of novelty on citations is likely to be overestimated in a naive estimate. We can however include a number of additional covariates to partially capture article quality. A number of previous studies have documented a strong relationship between team com-
Figure 6 – Novel and Before the Crowd

(a) Dependent Variable: Forward Citations

(b) Dependent Variable: “big hit” Dummy

Notes: Citations are recorded in a 3-year window. “Big hit” papers are defined as top 10% articles which received the most citations in their subject category. Results are based on incidence rate ratios from generalized negative binomial regressions for the left graph. For the right graph, results are based on odds ratios from logistic regressions. The sample is the set of all research articles published in journals indexed by the WoS, over period 1999-2011. High novelty is a dummy equal to one if the paper is in the top 10% most novel papers. The High Commonness dummy is equal to one if the same keyword combination employed to assess novelty in period \( t \) is still used by papers published in periods \( t + 1 \) and \( t + 2 \). Otherwise, it takes the value 0. The baseline category is all articles that are non-novel in \( t \) and non-common in \( t + 1 \) and \( t + 2 \). Detailed regression results can be found in the Online Appendix.

position and citations (Adams et al., 2005; Wuchty, Jones and Uzzi, 2007; Jones, Wuchty and Uzzi, 2008; Adams, 2013). We thus include the number of co-authors, the number of distinct institutions and the geographical regions dummies as additional co-variates. We also include a number of other controls such as year and discipline dummies, the number of keywords, and keyword novelty which, as we have shown in Subsection 3.3, correlate with keyword combination novelty. Because those controls may jointly only imperfectly capture article quality, we are inclined to interpret the odds ratios of novelty on citations as upper-bounds. The results are synthesized in Table 3.

Our first evidence is that highly novel papers receive 32% more citations than other papers, a number which actually remains close to the one obtained in Section 4.2 (38%, see Table 2). The probability of becoming a “big hit” is more sensitive as the odds ratios drop to 25% (vs. 42% initially). These results are consistent across scientific fields. It is interesting to note that the odds ratios of a “big hit” are larger when citations are recorded over 5 years, indicating that time plays in favor of novelty. In short, the estimated returns on novelty are reduced as compared to the previous coefficients, but remain quite significant. However, as we are likely to interpret them as upper-bounds of the individual return, the latter could end up being rather small.
We have shown above that the success of a paper also very much depends on the knowledge area being active in the future. Moreover, we have seen that high novelty and future commonness correlate. Therefore it is likely that part of today’s novelty is in fact driven by the anticipation of future commonness, which in turn may affect citations. In other words, part of today’s novelty could be connected to those “hot topics” that may already be “in the air”. Researching in those fields could explain impact, biasing upward the odds ratios of novelty. In essence, we aim to appreciate the relation between novelty and impact, sorting out the effect, on citations, of following what already is “in the air”. In other words, we would like to be able to appreciate the relation between keyword novelty and academic impact when controlling for future commonness. As we have seen that novelty and commonness are positively correlated and are more substitutes than complementary in rising citations, we anticipate a reduced effect of novelty on citations when controlling for commonness.

For that purpose, we add the commonness dummy introduced in the previous section among the regressors. This way of proceeding amounts to assuming that future commonness in the area is exogenous. In particular, the number of subsequent papers that employ the same pair of keywords is assumed not to be due to the focal paper. Though we believe this is often a reasonable assumption, it may not always be true and thus we
are now inclined to consider the regression results obtained as lower-bounds.

Table 3 – Individual Decision Making: Predicting Citations and “big hit” Probabilities with more Controls.

<table>
<thead>
<tr>
<th></th>
<th>big hit (top 10%)</th>
<th>big hit (top 5%)</th>
<th>Gen. Neg. Bin. (coefficient)</th>
<th>Gen. Neg. Bin. (ln(alpha))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3-year</td>
<td>5-year</td>
<td>3-year</td>
<td>5-year</td>
</tr>
<tr>
<td>Full Sample</td>
<td>25%</td>
<td>27%</td>
<td>24%</td>
<td>26%</td>
</tr>
<tr>
<td>Humanities and social sciences</td>
<td>23%</td>
<td>28%</td>
<td>26%</td>
<td>30%</td>
</tr>
<tr>
<td>Hard sciences and engineering</td>
<td>22%</td>
<td>25%</td>
<td>20%</td>
<td>23%</td>
</tr>
<tr>
<td>Life sciences</td>
<td>27%</td>
<td>29%</td>
<td>28%</td>
<td>30%</td>
</tr>
</tbody>
</table>

|                                | big hit (top 10%) | big hit (top 5%) | Gen. Neg. Bin. (mean) | Gen. Neg. Bin. (ln(alpha)) |
|                                | 3-year | 5-year | 3-year | 5-year | 3-year | 5-year | 3-year | 5-year |
| Full Sample                    | 5%     | 6%     | 3%     | 5%     | 17%    | 16%    | -7%    | -7%    |
| Humanities and social sciences | 5%     | 9%     | 7%     | 9%     | 16%    | 17%    | -12%   | -12%   |
| Hard sciences and engineering  | 3%     | 5%     | 0%     | 2%     | 6%     | 6%     | -7%    | -7%    |
| Life sciences                  | 7%     | 9%     | 6%     | 9%     | 19%    | 19%    | -5%    | -4%    |

Notes: Based on the set of all research articles published in journals indexed by the WoS, over period 1999-2011 (for 3-year citation window regressions) and 1999-2009 (for 5-year citation window regressions).

Obtained from exponentiated coefficients in generalized negative binomial estimations and logistic regressions.

Dependent variable for negative binomial regressions: number of forward citations (3-year and 5-year). Dependent variable for logistic regressions: dummy taking the value 1 if the paper is a “big hit” in its field (“top 10%” “top 5%”).

The explaining variable is high novelty, a dummy equal to one if the article is in the top 10% most novel articles. All results are significant at the 0.1% level.

Original tables are presented in the Online Appendix.

All results have been replicated by employing three alternative specifications of pairwise keyword novelty. Overall, the obtained results remain consistent. A detailed explanation of the five variants can be found in the Online Appendix.

† Control variables: number of keywords, keyword novelty, publication year, discipline dummies (10 fields), geographical dummies (Europe, USA-Canada, America (other), Asia-Oceania, Other), number of institutions and number of authors.

‡ Same controls as for the above regressions plus the commonness dummy defined in Subsection 4.3.

Highly novel papers means being among the top 10% most novel papers. Results have been replicated by defining highly novel papers with different thresholds (top 5% and top 1%), and can be found in Tables ?? - ??.

Regression results presented in the bottom part of Table 3 indicate that both novelty and commonness remain positively and significantly related to the number of citations received by a paper three years after publication. However, when controlling for future commonness, highly novel papers now receive 17% more citations, and have only 5% more chances of becoming a “big hit”. The latter effect even drops to 3% when “big hit” papers are defined as top 5% most cited articles (instead of the top 10% most cited).15

Even though we are inclined to consider these odds ratios as lower-bounds, they are rather small, in particular for explaining “big hit” papers. It is therefore not clear at all whether the costs of undertaking novel research would actually be less than their individual returns.

15In fact, the odds ratios are much more sensitive to the commonness dummy which increases the number of citations by 31% and the chances of being a big hit article by 44%, and up to 50% depending on the specifications (see detailed Tables in the Online Appendix).
4.5 Variants and Benchmarks

Variants of Pairwise Author Keyword Novelty  We assess the insensitivity of our results to alternative forms to compute pairwise author keyword novelty. As indicated in Subsection 2.1, we construct three distinct variants of pairwise author keyword novelty whose details are reported in the Online Appendix. All our findings are robust, and estimates are similar in magnitude to our findings employing our main pairwise author keywords novelty indicator.

Pairwise KeyWords Plus Novelty  As discussed in Subsection 2.2, an important choice we made concerns the use of Author Keywords over KeyWords Plus. To benchmark both alternatives, we have replicated all our main results employing KeyWords Plus. Table 4 provides a summary of the results concerning the prediction of citations (basically the analogue of Table 2). A highly novel paper based on KeyWord Plus pairs frequencies receives, on average, 8% more citations. This is significantly less than the 38% rise when employing Author Keywords. Even more surprising is the negative relation with top citations, which is consistent across scientific fields and using different specifications of a “big hit” (top 10% and top 5%). Pairs of Author Keywords, which we believe are much better suited to capture the original angle of papers, turn out to be key for predicting articles future success, and in particular “big hits”. Besides, nothing really differentiates the two forms of novelty in their relation with risk.

Table 4 – Predicting Citations and “big hit” Probabilities (Pairwise KeyWords Plus Novelty)

<table>
<thead>
<tr>
<th></th>
<th>“big hit” (top 10%)</th>
<th>“big hit” (top 5%)</th>
<th>Gen. Neg. Bin. (mean)</th>
<th>Gen. Neg. Bin. (ln(alpha))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3-year</td>
<td>5-year</td>
<td>3-year</td>
<td>5-year</td>
</tr>
<tr>
<td>Full sample</td>
<td>-8%</td>
<td>-7%</td>
<td>-10%</td>
<td>-8%</td>
</tr>
<tr>
<td>Humanities and social sciences</td>
<td>-12%</td>
<td>-9%</td>
<td>-15%</td>
<td>-9%</td>
</tr>
<tr>
<td>Hard sciences and engineering</td>
<td>0%</td>
<td>0%</td>
<td>-3%</td>
<td>-2%</td>
</tr>
<tr>
<td>Life sciences</td>
<td>-10%</td>
<td>-9%</td>
<td>-10%</td>
<td>-8%</td>
</tr>
</tbody>
</table>

Notes: Obtained from exponentiated coefficients in generalized negative binomial estimations and logistic regressions. Dependent variable for negative binomial regressions: number of forward citations (3-year and 5-year). We have employed the same sample of papers that we have used for the models based on Pairwise Author Keywords Novelty. Dependent variable for logistic regressions: dummy taking the value 1 if the paper is a “big hit” in its field (“top 10%” or “top 5%”). Control variables: number of keywords, publication year and disciplines dummies. Detailed regression results can be found in Tables ??- ??.

Pairwise Journal Reference Combinations Novelty  A second benchmark study concerns novelty based on pairwise journal reference combinations.

We study how the proportion of papers categorized as “big hits” across the centiles of
novelty. It turns out that the relationship between journal references novelty and academic impact is much less clear than for pairwise author keywords novelty. The correlation is globally positive, but nonlinearities are observed for high levels of reference combination novelty. Moreover, the slope is much less pronounced as compared to the observed slope in pairwise keywords novelty: when in the top decile of journal references novelty, the probability of being a top 5% cited article is “only” 25% higher than on the lowest decile (to be compared with the 200 to 300% for keyword combination novelty).

We next substitute the indicator based on the novelty of journal references pairs (see Subsection 2.2) to the standard pairwise author keyword novelty. Table 5 provides a summary of the main results. Pairwise journal references novelty is a positive predictor of “big hits”. However, incidence ratio rates are much smaller than the ones obtained for pairwise author keywords novelty. For instance, high pairwise journal references novelty rises the odds of a “big hit” by only 12 to 16%, to be compared with 42 to 46% obtained for pairwise author keyword novelty. Note that the results obtained for sample of papers from humanities and social sciences behaves differently however as high pairwise journal references novelty increases the odds of a top 10% paper by 37%, while pairwise author keyword novelty does so by 25% only. Interestingly, the benchmark journal references novelty indicator is a better predictor of impact for the sub-sample of humanities and social sciences only. This could reflect the fact that scientific innovation in those disciplines is often due to the importation of techniques and concepts from other fields (that would be captured by infrequent combinations of journal references).

Table 5 – Predicting Citations and “big hit” Probabilities (Pairwise Journal Reference Novelty).

<table>
<thead>
<tr>
<th></th>
<th>big hit(top 10%)</th>
<th>big hit(top 5%)</th>
<th>Gen. Neg. Bin. (mean)</th>
<th>Gen. Neg. Bin. (ln(alpha))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3-year</td>
<td>5-year</td>
<td>3-year</td>
<td>5-year</td>
</tr>
<tr>
<td>Full Sample</td>
<td>12%</td>
<td>12%</td>
<td>14%</td>
<td>16%</td>
</tr>
<tr>
<td></td>
<td>13%</td>
<td>10%</td>
<td>-7%</td>
<td>-3%</td>
</tr>
<tr>
<td>Humanities and social sciences</td>
<td>37%</td>
<td>31%</td>
<td>44%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>22%</td>
<td>20%</td>
<td>0%</td>
<td>4%</td>
</tr>
<tr>
<td>Hard sciences and engineering</td>
<td>9%</td>
<td>10%</td>
<td>12%</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>14%</td>
<td>-6%</td>
<td>-1%</td>
</tr>
<tr>
<td>Life sciences</td>
<td>12%</td>
<td>12%</td>
<td>13%</td>
<td>14%</td>
</tr>
<tr>
<td></td>
<td>9%</td>
<td>6%</td>
<td>-8%</td>
<td>-4%</td>
</tr>
</tbody>
</table>

Notes: Obtained from exponentiated coefficients in generalized negative binomial estimations and logistic regressions. Dependent variable for negative binomial regressions: number of forward citations (3-year and 5-year). Dependent variable for logistic regressions: dummy taking the value 1 if the paper is a “big hit” in its field (“top 10%” or “top 5%”). Control variables: number of keywords, publication year and disciplines dummies. All results are significant at the 0.1% level. Detailed regression results can be found in Tables ??- ??.

Novelty and Long-Term Impact  Recent empirical evidence (Wang, Veugelers and Stephan, 2017; Stephan, Veugelers and Wang, 2017) point to the idea that more novel papers may suffer from a delayed recognition. In other words, the greater impact of highly novel research takes time to materialize. To explore the extent to which this phenomena
occurs with our Pairwise Author Keywords Novelty indicator, we estimate the expected number of forward citations received by highly novel papers for citation windows ranging from 1 to 10 years. We restrict our analysis to the sub-sample of papers that were published within the period 1999 to 2001 to take advantage of a longer time period for citations to be recorded. As an additional robustness check, we replicate the same analysis with the pairwise KeyWords Plus novelty and with pairwise journal reference novelty. Table 6 summarizes the main results.

Table 6 – Novelty and Citation Impact in the short to the long term.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pairwise Author Keyword Novelty</td>
<td>42%</td>
<td>49%</td>
<td>49%</td>
<td>48%</td>
<td>46%</td>
<td>45%</td>
<td>44%</td>
<td>43%</td>
<td>42%</td>
<td>41%</td>
</tr>
<tr>
<td>Pairwise KeyWord Plus Novelty</td>
<td>5%</td>
<td>9%</td>
<td>10%</td>
<td>9%</td>
<td>8%</td>
<td>8%</td>
<td>7%</td>
<td>6%</td>
<td>6%</td>
<td>5%</td>
</tr>
<tr>
<td>Pairwise Journal Reference Novelty</td>
<td>4%</td>
<td>15%</td>
<td>19%</td>
<td>19%</td>
<td>19%</td>
<td>19%</td>
<td>18%</td>
<td>19%</td>
<td>19%</td>
<td>19%</td>
</tr>
<tr>
<td><strong>Ln(alpha)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pairwise Author Keyword Novelty</td>
<td>-15%</td>
<td>-13%</td>
<td>-11%</td>
<td>-10%</td>
<td>-9%</td>
<td>-9%</td>
<td>-9%</td>
<td>-9%</td>
<td>-9%</td>
<td>-9%</td>
</tr>
<tr>
<td>Pairwise KeyWord Plus Novelty</td>
<td>1%</td>
<td>0%</td>
<td>-2%</td>
<td>-2%</td>
<td>-1%</td>
<td>-1%</td>
<td>-1%</td>
<td>-1%</td>
<td>-1%</td>
<td>-1%</td>
</tr>
<tr>
<td>Pairwise Journal Reference Novelty</td>
<td>0%</td>
<td>-11%</td>
<td>-9%</td>
<td>-7%</td>
<td>-5%</td>
<td>-4%</td>
<td>-3%</td>
<td>-2%</td>
<td>-2%</td>
<td>-1%</td>
</tr>
</tbody>
</table>

Estimates are obtained from exponentiated coefficients of generalized negative binomial models. Ln(alpha) are obtained from the dispersion estimates of generalized negative binomial models.

Dependent variable: number of forward citations after 1 to 10 years. Novelty variables are dummies indicating if the paper is among the top 10% most novel.

Control variables in all Pairwise Author Keywords Novelty models: number of author keywords, publication year dummies, discipline dummies. Control variables in all Pairwise Keywords Plus Novelty models: number of Keywords Plus, publication year dummies, discipline dummies. Control variables in all Pairwise Journal References Novelty models: number of references, publication year dummies, discipline dummies. All detailed regression results are available upon request.

Our findings clear contradict the idea that highly novel papers suffer from a delayed recognition. Instead, the estimated number of forward citations are stable for alternative citation windows. For instance, the forward citations one year after publication are already greater by 42% for highly novel papers. The largest incidence ratio rates are obtained for three year citations (49%). They decrease slightly after that, down to 41% more citations over ten years after publication. This stability suggest that the form of novelty captured by our indicator leads to an immediate citation advantage which remains stable over time. Unreported regression results show that adding more controls (as for the regressions reported in Table 3) do not alter this result.

Such invariance is not specific to our way of calculating novelty as the same analysis performed on our two main benchmark indicators (pairwise KeyWords Plus novelty and pairwise journal reference novelty) do not evidence any delayed recognition effect either. Therefore, our indicator based on Author Keywords much better predicts forward citations, for all considered citation windows. The results on the dispersion of citations (risk) are also invariant to the variation in the length of the citation window considered.
Alternative Definitions of Highly Novel Papers Most of the analyzes presented above rely on a certain definition of highly novel papers. A paper is highly novel if it is in the top 10%. As a robustness check, we use alternative thresholds to define highly novel papers: top 1% and top 5% of the novelty distribution. All the regressions using the top novelty dummy are replicated afresh using top 1% and top 5% dummies. All results are reported in the Online Appendix. Essentially, our results are consistent when using alternative thresholds. Moreover, as expected, the effect of pairwise author keywords novelty on forward citations is accentuated for most of the models when considering a more restrictive threshold (larger incidence ratio rates for top 1% and top 5% dummies on impact).

5 Journals, Journal Stratification and Novelty

Publishing in recognized journals has become one of the key signals as it enhances recruitment and promotion opportunities. Serious concerns have been raised in the literature on the peer review system being potentially biased against novelty due to some form of conservatism (Braben, 2004; Chubin and Hackett, 1990; Wesseley, 1998). Rather than being perceived as valuable, divergent contributions may instill confusion and even irritation among evaluators, leading to potential publication penalties (Leahey, Beckman and Stanko, 2017). Resch, Ernst and Garrow (2000) demonstrated that studies supporting unorthodox medical treatments receive lower ratings even though the supporting data are equally strong. Luukkonen (2012) argues that frontier research is less likely to appear as rigorous while Boudreau et al. (2016) show that high levels of novelty are associated with lower evaluations by experts.

To our knowledge, the literature on peer reviews has only considered specific scientific journals or funding programs. Our work adopts a distributional approach to explore the potentially varying role played by academic journals on novelty with respect to their stratification. In the first subsection (5.1), we document that higher impact journals significantly publish more novel articles. Subsection 5.2 shows that high impact journals publish more highly novel articles than their average article quality suggests. Subsection 5.3 provides a natural explanation for the previous result: highly novel articles benefit more from being published in high impact journals.
5.1 Which Journals Publish More Novel and Highly Novel Papers

The question we address in here is: are highly novel papers more likely to be published by large-audience academic journals, namely journals with the highest impact factor, for instance because their editors can pick path-breaking exploratory contributions? Or are highly novel articles more likely to be published in peripheral journals as alternative views would be more easily accepted in “niches”? Based on our previous results in this article (more novel articles are more cited), we are inclined to find that the former hypothesis is likely to be true.

We have arranged all articles according to the impact factor of the journals which published them and have calculated the average novelty for each centile of journal impact factor (see graph a of Figure 7). We find that larger-audience journals are more likely to publish novel papers. This effect is observed at any point of the impact factor distribution and for all fields of science.

The positive relation between the centiles of journal impact factors and novelty is particularly strong when one focuses on the top novel articles (graphs b, c and d of Figure 7). Graph (b) shows in particular that the average article in the highest impact factor centile is six times more likely to be a highly novel paper (top 10%) than the average paper in the first centiles of impact factor. A similar multiplicative factor applies for top 5% and top 1% novel articles (see graphs c and d), and for any field of science.

What is really striking is that the relation between the centiles of impact factor and high novelty is convex and that convexity accentuates when we focus on the top 5% and top 1% most novel papers (see graphs c and d). All the variation in the probability of a paper being in the top 1% most novel articles is observed above the 60th centile of impact factor. The average article published in a journal whose impact factor is median has about the same chance of being highly novel (top 1%) as the average paper published in the lowest impact factor journals. But at the same time, it is six times less likely to be highly novel than the average article published in top impact journals. This “publication premium” for highly novel papers is amplified for hard sciences and engineering, so that the average article published in a top impact journal in this field is fourteen times more likely to be highly novel than the average article in the lowest impact factor journals.
5.2 Disentangling the Quality from the Novelty Effect

The fact that top journals tend to publish more novel papers than lower tier journals seems to provide empirical grounds for the idea that top journals play a crucial role in selecting articles that uncover new research questions, or new dimensions of old questions. However, as we have seen previously, more novel articles are also more cited on average. Therefore, as citations are correlated with unobserved article quality and as top journal editors have their own signals on the quality of submitted articles (e.g. their own reading, reviewers’ comments, authors’ reputation or prestige of the host institutions...), they may end up selecting novel papers more often, not because they are more novel, but only because higher quality papers are also more novel on average. Therefore it would not be surprising that higher impact-factor journal end up publishing more novel papers.
Disentangling precisely such “quality effect” from the “novelty effect” may be very difficult. However, if the quality effect alone explained the sorting of novel articles in journals according to their impact factors, then the relation between the number of citations and novelty should be very similar to the relation between impact factor and novelty. In other words, the average article in a journal of a given centile of impact factor distribution should be as novel as the articles in the same centile of citation distribution. Journal stratification would then be “neutral”, that is the correlation between journal impact factor and novelty would only reflect the underlying relation between citations and novelty.

To test this idea, we have calculated the probability of an article of being a top novel one all along the distributions of both journal impact factors and forward citations. The left graph of Figure 8 shows that, up to the 80th centile of these distributions, the proportion of top 1% novel articles is lower in the impact factor distribution than in the citation distribution. Above the 80th percentile of these distributions, the likelihood of being top novel is greater in the impact factor distribution. Top-tier journals offer more slots to newer research than their quality standard suggests. In other words, their editors and reviewers allocate a publication premium to highly novel contributions. This finding suggests that top journals play a key role in publishing highly novel papers, to a significantly larger extent than their average quality would “naturally deliver”. Similar results are obtained with top 10% and top 5% highly novel articles.

Figure 8 – Pairwise Keyword Novelty with Respect to Citations and Impact Factor Centiles

(a) Impact Factor Citation Centiles and Proportion of Highly Novel Papers

(b) Citations-to-Impact-Factor Ratio Across Novelty Centiles

Notes: Based on the set of all research articles published in journals indexed by the WoS, over period 1999-2011. Fractional polynomial estimates and 90% confidence intervals. Citations are recorded in a 3-year window.
5.3 Why do High Impact Journals Overly Publish More Novel Papers?

To address the ‘why’ question with our data, we have calculated, for each article, its number-of-citations-to-impact-factor ratio. A ratio higher than the unity means that the article has been cited more often than the average article published the same year by the same journal. We plot the average number-of-citations-to-impact-factor ratio for each centile of article novelty in the right hand (b) graph of Figure 9. We find an increasing relation between novelty centiles and the ratio. This indicates that the more novel the papers, the more they perform as compared to the average paper in their journal and year. In other words, articles are increasingly “over-cited” when they are more novel. Below a certain level of novelty (about the 40th centile), articles have a rate lower than one. That is, those articles are on average published by journals that have a higher impact factor than their own citation score. On the other side of the aisle, most novel papers have a 10% premium over the average paper in their journal. Those papers benefit more from the “quality-certifying” role of passing the selection process of top journals. This could explain why top journal editors may want to attract highly novel papers and why the authors of novel papers may be (even) more willing to be published there.

As we have shown in the previous sections that novelty correlates with future commonness, we wonder whether the role of top journals in publishing novel papers could be partly due to their capacity to promote articles on “hot topics”. To assess this, we have computed, for each centile of journal impact factor, the proportion of high-novelty-and-common articles (see Section 4.3), which we have divided by the overall proportion of such articles. We have performed similar calculations for the other three categories of articles: not-high-novelty-and-common, high-novelty-and-not-common and not-high-novelty-and-not-common papers. A ratio greater than one means that the corresponding centile of impact factor is characterized by a greater than average frequency of the considered category of papers. Figure 8 presents the resulting ratios when high novelty is defined as the top 10% or as top 1% most novel articles. We find that the curves of high-novelty-and-common and of high-novelty-and-not-common articles have very similar patterns. They are both increasing from .5 up to 2 or 3 (depending on the definition of high novelty) with impact factor centiles. Surprisingly, we find that among highly novel articles, the non-common articles are even relatively more frequent than the common ones in the top journals in the (b) graph, that is, when high novelty is defined as the top

\footnote{This is equivalent to the relative citation rate (RCR) introduced by Schubert and Braun (1986) because our Impact Factor is calculated according to exactly the same period length, citing sources and starting date as citations. Note this definition of Impact Factor is different from the traditional one. It is calculated in-house as an “expected citation score” which averages the citations received (in the future) by the articles published by a given journal in a given year.}
1% most novel articles. In fact, commonness instead makes a big difference among the non-highly-novel articles in gaining access to high impact factor journals. The not-high-novelty-and-not-common are less frequent in higher impact journals while the reverse holds for not-high-novelty-and-common papers. This is consistent with the idea that novelty and commonness are substitutes for getting published in higher impact journals.

Figure 9 – Novelty, commonness and journal impact factor

Notes: Based on the set of all research articles published in journals indexed by the WoS, over period 1999-2011. Fractional polynomial estimates and 90% confidence intervals. Citations are recorded in a 3-year window. High novelty is a dummy equal to one if the paper is in the top 10% most novel papers for the left graph, and in the top 10% most novel papers for the right graph. Commonness dummy is equal to one if the same keyword combination employed to assess novelty in period \( t \) is still used by papers published in periods \( t + 1 \) and \( t + 2 \). Otherwise, it takes the value 0. Relative concentration is computed as the share of each considered category in the centile, divided by the share of that category over all centiles.

If we use the number of citations as a proxy for the “scientific value” of articles (Gottfredson, 1978; Campanario, 1995; Bornmann and Daniel, 2008, 2009; Bornmann et al., 2011) and consider the impact factor scales as their “market value”, this graph highlights that more novel articles face a negative bias as they should be published in higher impact scientific outlets.

6 Conclusion

In this article we propose a new measurement of the novelty of scientific articles based on keyword pairwise combination frequencies which we compute on the set of all research articles in the WoS (about ten million articles) that have at least two keywords over fifteen
years (from 1999 to 2013). We find that novelty is not declining in that period and that more novel articles are not likely to be performed with in closed walls, but are more often produced in larger teams that span more institutions and geographical regions.

It turns out that novelty is a good predictor of citations, as it increases the probability of being a “big hit” by 42% and the number of citations by 37%. As this correlation is not counterbalanced by a higher risk, that is, a higher variance of citational outcomes of novel articles (provided that papers get published in a referenced journal), these results provide a systematic empirical grounding for agencies to fund research projects that are more novel and disruptive.

We prove that science has common traits with finance and “beauty contests”, as publishing an article whose most novel dimension is still active in the following years significantly increases citations. A paper which has an “angle” that is new at the time of publication which itself is still active in the following years has its probability of becoming a citational “big hit” increased by 67% – and up to 72% in the hard sciences. However, there is no complementarity between present novelty and future commonness, which are in fact positively correlated with each other and are partial substitutes in rising citations.

When we hold constant a number of co-variates, we obtain estimates that become closer to the citational returns on novelty at the individual level. We have found that the impact of novelty on citations remains, but decreases significantly. In particular, when we control for future commonness –as, in fact, it is positively correlated with both present novelty and citations– we find that the odds ratios of high novelty on “big hit” probability become pretty low. This suggests that, though the returns on novelty at the collective level are large, they may actually be very limited at the individual level. As our study does not consider the projects that did not lead to research articles published in referenced journals, and as at the same time, more novel articles are probably much more difficult and costly, thus the rewards for novelty net of their costs may be rather low, and potentially negative.

This raises major concerns about the implicit reward system of science to provide sufficient incentives for undertaking new projects. It also raises questions on why academic researchers engage in novel projects. In fact, we have discovered that engaging in novel research does not correspond to risk taking. The reverse rather holds: avoiding novelty increases risk, in particular the portion of risk which is caused by others no longer being interested in your work –and thus not citing it. In a rapidly changing environment like science, avoiding novelty puts agents at risk, not novelty. Novelty is reducing risk, not increasing it.
References


Olesen Larsen, Peder and Markus von Ins. 2010. “The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index.” *Scientometrics* 84:575–603.


Zhang, Juan, Qi Yu, Fashan Zheng, Chao Long, Zuxun Lu and Zhiguang Duan. 2016. “Comparing keywords plus of WOS and author keywords: A case study of patient adherence research: Comparing Keywords Plus of WOS and Author Keywords.” *Journal of the Association for Information Science and Technology* 67(4):967–972.