Do Top-Tier Scientific Journal Publications Improve Corporate Innovations? *

Keng-Yu Ho
Department of Finance, National Taiwan University
Center for Research in Econometric Theory and Applications, National Taiwan University

Yanzhi Wang
Department of Finance, National Taiwan University
Center for Research in Econometric Theory and Applications, National Taiwan University

Chia-Wei Yeh
Department of Banking and Finance, National Chi Nan University

October 2022

Abstract

While colleges and universities emphasize publications in top-tier scientific journals, the study of a firm's high-impact publications and their effect on corporate innovation activities is scarce. To fill the gap, this paper explores the corporate innovations of firms with and without top-tier scientific journal publications. We conduct the analyses using detailed journal publication data that are collected manually from the Web of Science. Our empirical findings show that firms with top-tier journal publications receive more citations on their patent; these firms also generate greater patent generality and patent originality. The conclusion remains held under several identification strategies. Further analyses suggest that, instead of collaborating with top universities, geographic proximity between firms and universities is important regarding the effect of top-tier journal publications on corporate innovations.

Keywords: Corporate Innovation; Journal Publication; Patent

---

* We are grateful to Paul Brockman, Konan Chan, Ching-Hung Chang, Hsuan-Chi Chen, Yehning Chen, Wan-Chien Chiu, Axel Grossmann, Pei-Fang Hsieh, Po-Hsuan Hsu, Kenneth Huang, Chuan Yang Hwang, Woan-Lih Liang, Hung-Hua Pan, Jerry T. Parwada, Shu-Cing Peng, David Reeb, Johan Sulaiman, Huei-Wen Teng, Kevin Tseng, Meng-Fen Yen, Tze-Yu Yen, Chung-Ying Yeh, Min-The Yu, and Mu-Shu Yun for their valuable comments and suggestions. We also benefit from comments from the conference participants at the 2019 International Conference of Taiwan Finance Association as well as seminar participants at National Taiwan University, National Chiao Tung University, National Tsing Hua University, and National Chung Hsing University. Keng-Yu Ho and Yanzhi Wang appreciates the financial support by the Center for Research in Econometric Theory and Applications (Grant no. 107L900202, 108L900202) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education, Taiwan, and by the National Science and Technology Council, Taiwan (NSTC 105-2628-H-002-002-MY3; NSTC 107-2410-H-002-044-MY3; NSTC 107-2420-H-002-001; NSTC 107-3017-F-002-004).

Do Top-Tier Scientific Journal Publications Improve Corporate Innovations?

October 2022

Abstract

While colleges and universities emphasize publications in top-tier scientific journals, the study of a firm's high-impact publications and their effect on corporate innovation activities is scarce. To fill the gap, this paper explores the corporate innovations of firms with and without top-tier scientific journal publications. We conduct the analyses using detailed journal publication data that are collected manually from the Web of Science. Our empirical findings show that firms with top-tier journal publications receive more citations on their patent; these firms also generate greater patent generality and patent originality. The conclusion remains held under several identification strategies. Further analyses suggest that, instead of collaborating with top universities, geographic proximity between firms and universities is important regarding the effect of top-tier journal publications on corporate innovations.

Keywords: Corporate Innovation; Journal Publication; Patent

1. Introduction

Patents and papers in top-tier scientific journals are both important for the evaluation of scientific researchers' academic performance in many fields. Meanwhile, firms file patents to gain competitive advantage in the product market, yet do firms pursue scientific researches and publish papers in top-tier journals as scientific researchers? In fact, U.S. listed firms have published more than 8,000 articles in top-tier scientific journals including *Nature*, *Science*, and *Cell* since 1980. To date, we have known little about how researches published in top-tier scientific journals may affect the firm's innovation outputs and performance, and this paper fills this gap.

Papers in top-tier scientific journals are highly related to fundamental scientific research, which is the foundation of corporate innovations. University researchers generally have less incentive to keep their research as secret, and the research ideas generated by these talented researchers facilitate the process of corporate innovations in their neighborhood (Jaffe, 1989; Mansfield, 1991; 1998). Gittelman and Kogut (2003) suggest that firms could produce valuable innovations by establishing credible linkages through co-publications with the scientific communities such as prestigious universities and research institutes. Through the collaboration between university and industry, firms can gain increased access to new university research and discoveries, which help them better identify new research and development (R&D) projects (Cohen, Nelsen, and Walsh, 2002; Lee, 2000; Rosenberg and Nelson, 1994; Zucker, Darby, and Armstrong; 2002). Simeth and Cincera (2016) investigate the impact of scientific publications of firms on their own market values. They find positive impact of scientific publications on a firm's market value beyond the effects of in-house research and development and patenting activities.

In this paper, we explore the potential effect of firms' publications in top-tier scientific journals on innovation and firm performance. We focus on the publications in top-tier journals because it takes tremendous efforts to publish research in prestigious journals. Moreover, Furman, Jensen, Murray (2012) argue that the journals with high Journal Impact Factors, such as *Science* and *Nature*, maintain their quality and reputation by strict retraction system, ensuring the outcome of the scientific researcher implementable. Hence, these published papers in top-tier journals are expected to have significant contributions and high impacts on both industries and academics. Moreover, papers coauthored with the universities and research institutes can implicitly stand for the firm's R&D capability in additional to the in-house R&D. To publish papers in top-tier journals, firms should either have a solid research capability themselves or collaborate with scientists and experts from universities and research institutes. All these can benefit patenting activities of the firm.

Therefore, we hypothesize that firms with top-tier journal publications should generate better innovation quality, which is measured by patent citations, patent generality, and patent originality. In addition, firms could also improve their innovative efficiency from the publishing experience. Previous studies of Cohen, Diether, and Malloy (2013) and Hirshleifer, Hsu, and Li (2013) suggest that firms with better innovative efficiency tend to have better market valuation and operating performance, thus we also expect that firms with publications in top-tier journals experience better market valuation and operating performance.

We manually collect the detailed journal publication data from the Web of Science and explore the corporate innovations of public listed firms in the U.S. We search all articles and letters that are published in the top-tier journals with higher impact factor, including *Nature*, *Science*, *Cell*, *Journal of the American Chemical Society*, *Lancet*, *New England Journal of Medicine*, and *Proceedings of the National Academy of Sciences*. Corporate innovation data rely on patents of United States Patent and Trademark Office (USPTO) and their citation data. Our empirical results show that firms tend to generate patents with more forward citations if they publish academic papers top-tier scientific journals. We perform several robustness tests. First, we conduct the instrumental variable approach to deal with the potential endogenous problem of reverse causality, since firms with better patent innovation quality are likely to have higher chance to publish in top-tier journals. We also understand that our results might suffer from the self-selection bias. It is likely that firms succeed in publishing in top-tier journals have higher ability in research and innovation themselves, and thus would have better patent innovation quality. We account for the self-selection bias using the matched sample constructed by different matching procedures for robustness checks. We also conduct a placebo test and ensure that our findings are less likely to be driven by chance.

More importantly, we perform a difference-in-differences test by using the new submission policy for the *New England Journal of Medicine* (*NEJM*) as an exogenous shock. The International Committee of Medical Journal Editors (ICMJE) and its members journals require the authors to submit data for clinical trial registration in their initial submissions since 2005, increasing the difficulty of acceptance on medical journals as well as their experimental reproducibility. Therefore, firms with papers published in *NEJM* also have better chance and ability to implement their research works to commercialized patents and products. To perform the difference-in-differences analysis, we consider firms with published papers in *NEJM* as treated firms, and find their peer control firms with papers in other top journals rather than *NEJM* based on propensity score matching. We find that compared with control firms, treated firms tend to have better innovation performance, confirming the notation that firms generate better innovation output as their academic research works are published in a journal with tougher standard.

Next, we discuss how publishing papers in *Nature* or *Science* could differ from publishing papers in other top-tier journals. Our empirical findings suggest that the effect on patent citations is indeed stronger for firms with publications in top-tier journals, especially for Nature or Science. In addition, we examine the collaboration between firms and universities. Our findings show that instead of collaborating with coauthors from top schools in the U.S., firms perform better in innovation performance if the firms collaborate with coauthors from universities in the same state, meaning that companies gain more resources and facilitate themselves to make successful progress in corporate innovations. The results also implies that geographic proximity between firms and universities is important regarding the effect of top-tier journal publications on corporate innovations.

Finally, we examine the top-tier journal publication and firm performance. We measure the firm performance by Tobin's Q and profit margin and find that firms with publications in top-tier journals experience better Tobin's Q and operating performance. These results are aligned with abovementioned finding about the effect of top-tier journal publications on innovation performance.

Our study contributes to the existing literature on corporate innovations by exploring the effect of top-tier journal publications, which is rarely studied in the past literature. We note that Gittelman and Kogut (2003), a related study, shows mixed results that are difficult to explain, and those findings may due to the industry characteristics, since they only focus on one specific industry. Another possible reason for the mixed results could be due to those journals selected by Gittelman and Kogut (2003) are not restricted to top-tier journals. We believe that the impact of minor journal publications on improving corporate innovations is limited. Our primary results focus only on the publications in the top-tier journals, and our additional analyses also show that the effect of publications in non-top journals on improving corporate innovations is weaker. In addition, we use a much more comprehensive dataset which enables us to perform analyses among different industries and present empirical results from a more general perspective.

The remainder of the paper is organized as follows. Section 2 reviews related literature and develops the hypothesis to be tested. Section 3 describes data sources and definitions of the variables. Section 4 presents empirical results. Section 5 performs additional analyses. Section 6 concludes.

2. Literature Review and Hypothesis Development

Research papers and patents are both the outputs of scientific research and technological innovation, but they are different in many ways. Research papers are used to communicate the findings and the results of a scientific process with the relevant scientific communities and the general public, while patents are legal documents used to protect the described process or devices. Prior research has pointed out the importance of patents' role in intellectual property protection (Ginarte and Park, 1997; Lerner, 2002), and find a positive relationship between R&D activities and the patents granted (Czarnitzki and Toole, 2011; Ginarte and Park, 1997). Thus, it is important to study patents in corporate innovation, and the measures of patent quality suggested by Hall, Jaffe, and Trajtenberg (2001; 2005) are widely used to measure the research performance of companies in the related studies. However, the study of the role of research papers is scarce in corporate innovations. Although companies have less incentive to disclose their research findings and loss their competitive advantages in R&D activities, there are still many companies actively publish their findings in scientific journals and enjoy better performance in the market (Hicks, 1995; Simeth and Cincrea, 2016). Therefore, research papers also play an important role in corporate innovation.

The academic practice of companies is beneficial because it mediates links with other research organizations, and the research network can improve corporate innovation. Hicks (1995) explores the reasons why companies publish scientific and technical literature, and find companies have the need to link with other research organization. Therefore, companies will publish papers to signal the presence of knowledge and build technical reputation, and they will be able to attract more talented researchers or have higher opportunities to collaborate with universities. Adams, Black, Clemmons, and Stephan (2005) find an increasing trend in institutional collaborations and a growing trend in geographically dispersed scientific teams, both indicate the decline in cost of collaboration between universities and companies. They suggest the key determinant that universities collaborate with companies is due to the placement of graduate students, so companies can benefit from the key-talents movements by collaborating with universities, which is consistent with view of Bishop, D'Este, and Neely (2011). Lee and Miozzo (2019) find that the knowledge-intensive business services firms are active collaborators with universities for innovation. The findings in Bereskin, Campbell, and Hsu (2016) also suggest that companies tend to collaborate with other organizations when they want to develop new innovations outside their traditional area of expertise. The collaboration with research organizations is beneficial for corporate innovation, because when firms have better access to academic research, they are more likely to develop new product and process, and have better know-how to identify new R&D projects (Jong and Slavova, 2014; Lee, 2000; Rosenberg and Nelson, 1994; Zucker et al., 2002). Cohen et al. (2002) also suggest the key channels of public research on industrial R&D are through published papers, public conferences, informal information exchange, and consulting. We therefore expect the innovation performance will be enhanced if the company has research papers published in scientific journals, or collaborated with other research institutions to engage in doing scientific research. We thus develop the following hypothesis:

H1. Research papers published in scientific journals has a positive effect on innovation performance.

In this study, we especially focus on research papers that are published in the top-tier scientific journals, because the scientific findings reported in top-tier journals are considered to be more reliable. Some related surveys have point out the concern about published results in journal articles that cannot be reproduced (Baker, 2016; Ioannidis, 2005; Mullard, 2011; Prinz, Schlange, and Asadullah, 2011). Munafo and Flint (2010) suggest publication bias could be a possible factor that contribute to the presence of false results, and readers should put more emphasis on findings in mature literature. Previous literature also suggests that peer review can only mitigate the problem and improve the quality of papers to be published to some extent, and there are still major errors that cannot be detected by the reviewers (Baker, 2016; Schroter et al., 2008; Scott, 2007). Therefore, we expect the effect of publishing research papers in high-impact journals with higher standards and better reputations should be stronger than publishing research papers in journals with low impact factors. In addition to the reliability of high-impact journals, the signaling effect of publishing papers (Hicks, 1995) is also stronger for high-impact journals, because they attract more attention from the scientific organization. We suggest the following hypothesis:

H2. The effect of research papers on innovation performance is higher when companies published paper in high-impact journals.

The spillover of academic know-how from collaborated institutions to companies or the movement of talented researchers are important factors why companies benefit from having journal publications. Previous literature of Jaffe (1987) has suggested a geographical spillover effect from university research to corporate innovation, the empirical evidence in Mansfield (1991; 1998) also confirm the existence of the geographical spillover effect. Geographical proximity is also an important issue that influences the sources of knowledge available to organizations. Bishop et al. (2011) find companies are more likely to collaborate with geographically proximate universities, and it is crucial for them to get assistance from university scientists for problem solving. Bishop et al. (2011) and Mansfield and Lee (1996) also suggest that high-ranking universities have better research capacities, and collaborate with high-ranking universities is likely to influence the benefit that companies can obtain. We develop the following hypotheses:

H3. The effect of research papers on innovation performance is higher when companies published paper with geographically proximate universities.

H4. The effect of research papers on innovation performance is higher when companies published paper with high-ranking universities.

Finally, the firm performance will increase if the innovative efficiency improves accordingly. Previous studies of Cohen et al. (2013) and Hirshleifer et al. (2013) suggest that firms with better innovative efficiency tend to have better market valuation and operating performance. Simeth and Cincrea (2016) find a positive impact of scientific publications on companies' market value, so it is likely that firms with publications in top-tier journal also benefit from the improvement innovative efficiency. We suggest the following hypothesis:

H5. Research papers published in scientific journals has a positive effect on firm performance.

3. Data and Methodology

3.1. Data

We obtain financial data of our sample from the Compustat database. The detailed top-tier journal publication data are collected manually from the Web of Science. Base on the ranking of Google Scholar citations in 2016, we search all articles and letters that are published in the top-tier journals with higher impact factor, including Nature, Science, Cell, Journal of the American Chemical Society, Lancet, New England Journal of Medicine, and Proceedings of the National Academy of Sciences of the United States of America. Among these top journals, Nature, Science, and the Proceedings of the National Academy of Sciences of the United States of America are multidisciplinary scientific journals. In addition, Cell is the top-tier journal in the discipline of biology; the Journal of the American

Chemical Society is the top-tier journal in the discipline of chemistry; Lancet and New England Journal of Medicine are top-tier journals in the discipline of medicine. To avoid focusing our sample on certain specific academic disciplines, we also include other specialized journals of Nature Publishing Group as well as other family journals of Science.

We check the affiliation of all the authors of each publication and count the number of journal publications for each firm-year observation. The patent and citation data are collected from the latest edition of the European Patent Office Worldwide Patent Statistical Database (PATSTAT), which is a more comprehensive database that identifies the citations received by a patent and the citations made by a patent (Bena and Li, 2014).

To be included in our sample, firms are required to have financial data in the Compustat database in a given year. Financial firms and utilities firms are excluded from the sample. Since the latest patents tend to have fewer citations, we terminate our sample period at year 2013 to address for the patent truncation concern. Our sample construction process yields a firm-year panel of 82,547 observations from 1980 to 2013. The number of firms that have publications in top-tier journals in each year during our sample period ranges from 24 (in 1983) to 79 (in 2004). Overall, there are 1,674 firm-year observations that publish at least one paper in the top-tier journals.

3.2. The measures of innovation

The measures of patent innovation quality are subject to the truncation problem, because the latest patents tend to have fewer citations. Hence, we construct the patent innovation quality measures following the method suggested by Hall et al. (2001; 2005). First, a firm's total number of patent applications (*PatentCounts*) is adjusted by the median number of patent applications in the same International Patent Classification (IPC). In addition, a firm's total number of citations received on each patent applications (*Citations*) is the number of truncation-adjusted citations divided by the number of patent applications. We also use an alternative measure for a firm's total patent citations (*RelativeCitations*), which is the citations per patent divided by the median of citations per patent in the same IPC classification. We adopt both *Citations* and *RelativeCitations* as measures of patent innovation quality in our main analyses.

In addition to the patent citations measures, the measures of generality and originality are also widely used in literature to examine different aspects of patent innovations. The generality of a firm's patents i is defined as:

$$Generailty_i = 1 - \sum_{j}^{n_i} s_{ij}^2 \tag{1}$$

where s_{ij} represent the percentage of forward-citations received by patent i that belong to the IPC classification j, out of n_i technology classes. If a patent is cited by subsequent patents that spread on a wide range of technology classes (sum of squared term is lower), it will have higher value of this variable, which means the patent is more influential in different technology areas. Similarly, the originality of a firm's patents i is defined as:

$$Originailty_i = 1 - \sum_{j}^{n_i} t_{ij}^2 \tag{2}$$

where t_{ij} represent the percentage of backward-references made by patent i that belong to the IPC classification j, out of n_i patent technology classes. If a patent cites previous patents that belong to a narrow range of technology classes (sum of squared term is higher), it will have lower value of this variable, which means the patent is more rely on specific technology areas. Citing patents in a wide range of technology areas, by contrast, will increase patent originality. We adopt *Generality* and *Originality* as measures of patent innovation quality in additional analyses.

3.3. Summary statistics

Table 1 presents summary statistics of firm characteristics. *Publish* is an indicator variable, which equals to one if the firm-year observation has journal publications, and zero otherwise. *PublCounts* is the total number of journal publications for the firm-year observation. *Assets* and *Sales* are both in billion dollars. *TobinsQ* is the firm's market value divided by the book value of property, plant and equipment, where the firm's market value is computed using the market value of common equity plus the book value of debt minus the firm's current assets. *ProfitMargin* is the sum of income before extraordinary items, interest expenses, and total income taxes, divided by total sales. *Size* is the logarithm of total assets. *HHI* is Herfindahl-Hirschman Index, which is the squared sales-based market share upon 3-digit SIC industry. *AnalystCover* is the number of analysts following the firm in one month before earnings announcement, obtained from I/B/E/S. *R&DIntensity* is the R&D expenditures divided by total sales. *CashFlow* is the income before extraordinary items minus total accruals (changes in current assets plus changes in short-term debt, minus the sum of changes in cash, changes in current liabilities, and depreciation expenses), divided by average total assets. *CapitalLaborRatio* is the property, plant and equipment divided by the number of employees. All continuous variables are winsorized at the 1st and 99th percentiles.

The mean of *Publish* is 0.0203, which suggests that about 2% of firms in our sample have papers publish in top-tier scientific journals. An average firm has total assets of $2.99 billion and annual sales of $2.71 billion. On average, a sample firm publishes 0.10 papers each year, has about 2.95 patents granted each year, generates 67.45 patent citations and 2.62 relative citations. In addition, an average firm has a score of 0.07 for generality and 0.09 for originality. In regard to other variables, an average (median) firm has Tobin's Q of 0.04 (0.01), earns -$0.02 ($0.08) for each dollar of sales, has annual cash flow which is 6% (8%) of total assets, and has capital labor ratio of 0.17 (0.03). An average firm is also covered by 1 analyst and invests 9% of total sales to R&D activities.

[Insert Table 1 about here]

Table 2 provides univariate comparisons for our sample based on *Publish*. The preliminary results show that firms with top-tier scientific journal publications publish 5.02 papers in each year on average. In addition, firms with publications seem to have better innovation outputs and innovation quality. They tend to have higher *PatentCounts*, *Citations*, *RelativeCitations*, *Generality*, and *Originality*. Specifically, the mean and median of *PatentCounts* for firms with publications are 18.05 and 3.50, respectively; both are significantly higher than those for firms without publications, which are 2.63 and 0, respectively. The mean and median of *Citations* for firms with publications are 392.36 and 50.30, respectively; both are significantly higher than those for firms without publications. When we use *RelativeCitations* as an alternative measure, its mean and median are 15.68 and 2.04, respectively, for firms with publications; both are also significantly higher than those for firms without publications. The mean and median of *Generality* are 0.22 and 0.20, respectively, for firms with publications; both are significantly higher than firms without publications. Similarly, the mean and median of *Originality* for firms with publications are 0.27 and 0.31, respectively; both are also significantly higher. Firms with publications also have higher firm valuation, since their mean and median of *TobinsQ* are both significantly higher than that of firms without publications.

[Insert Table 2 about here]

4. Empirical Results

4.1. Main results

To examine the relationship between top-tier journal publication and innovation quality, we perform regression analyses based on the following models.

$$\begin{aligned} InnovationQuality_{i,t+1} = {} & \beta_0 + \beta_1 Publish_{i,t} + \beta_2 Size_{i,t} + \beta_3 HHI_{i,t} \\ & + \beta_4 Log(1{+}AnalystCover_{i,t}) + \beta_5 R\&DIntensity_{i,t} \\ & + \beta_6 TobinsQ_{i,t} + \beta_7 CashFlow_{i,t} + \beta_8 CapitalLaborRatio_{i,t} \\ & + IndustryFE + YearFE + \varepsilon_{it}. \end{aligned} \tag{3}$$

$$\begin{aligned} InnovationQuality_{i,t+1} = {} & \beta_0 + \beta_1 Log(1{+}PubliCounts_{i,t}) + \beta_2 Size_{i,t} + \beta_3 HHI_{i,t} \\ & + \beta_4 Log(1{+}AnalystCover_{i,t}) + \beta_5 R\&DIntensity_{i,t} \\ & + \beta_6 TobinsQ_{i,t} + \beta_7 CashFlow_{i,t} + \beta_8 CapitalLaborRatio_{i,t} \\ & + IndustryFE + YearFE + \varepsilon_{it}. \end{aligned} \tag{4}$$

The dependent variable, *InnovationQuality*, is measured by *Citations* and *RelativeCitations*. To capture to effect of top-tier journal publications on the innovation quality, *Publish* and *PublCounts* are included in the models in equations (3) and (4), respectively. We follow related literature and control for *Size*, *HHI*, *AnalystCoverage*, *R&DIntensity*, *TobinsQ*, *CashFlow*, and *CapitalLaborRatio* (Aghion, Bloom, Blundell, Griffith, and Howitt, 2005; Becker-Blease, 2011; Aghion, Van Reenen, and Zingales, 2013;

He and Tian, 2013; Fang, Tian, and Tice, 2014).[1] Furthermore, we also control for the factors that vary over time but not across industries and the factors that vary across industries but not over time, by including year indicators and industry indicators in the regression, where industries are based on the 2-digit SIC code.

The Panel A of Table 3 presents the empirical results of the ordinary least squares regressions, where models (1) and (2) employ the logarithm of one plus *Citations* as the dependent variable, while models (3) and (4) employ the logarithm of one plus *RelativeCitations* as the dependent variable. In models (1) and (3), the coefficients on *Publish* are 1.4224 and 0.7069, which are both significant at the 1% level, suggesting that publish in top-tier journals is positively related to patent citations. In models (2) and (4), the coefficients on the logarithm of one plus *PublCounts* are 0.7319 and 0.3840, which are also significant at the 1% level, suggesting that the number of publications in top-tier journals is also positively related to patent citations. Therefore, firms with scientific papers in top-tier journals generate better patents than firms without publications, which is consistent with *H1*.

In addition to the OLS regression, we estimate the regression models by using Tobit regression for robustness check, because the patent citations tend to be censored at 0 when we perform analyses on the Compustat universe. The results are shown in Panel B of Table 3, and the results are similar to those of OLS regression.

[Insert Table 3 about here]

4.2. Addressing endogeneity concerns

We recognize that our analyses might be subject to endogeneity concerns. First, we note that the chance for firms to publish in top-tier journals could be higher if they have better patent innovation quality. We employ the two-stage least squares (2SLS) instrumental variable approach to mitigate this concern

[1] Aghion, Bloom, Blundell, Griffith, and Howitt (2005) and Aghion, Van Reenen, and Zingales (2013) investigate the relationship between product market competition and innovations. Becker-Blease (2011) examines the relation between the corporate governance and corporate innovation; He and Tian (2013) and Fang, Tian, and Tice (2014) also investigate the relationship between institutional ownership and innovations. We use the Herfindahl-Hirschman Index (*HHI*) as the proxy for product market competition, and adopt *AnalystCoverage* as the proxy for corporate governance, because the analysts provide monitoring function from the outside of the corporation.

of reverse causality. Second, firms that succeed in publishing top-tier journal papers might have higher ability in research and innovation. Therefore, they are more likely to have better patent innovation quality as well. The particular characteristics, such as research ability or innovation ability, are omitted or difficult to observe and could also bias up the empirical results. To deal with this problem, we perform several different matching procedures. We also perform a difference-in-differences test by adopting an exogenous shock on publishing papers in the New England Journal of Medicine after it applies new submission policy after 2005.

4.2.1. Instrumental variable approach

In this subsection, we employ the 2SLS instrumental variable approach to address the endogeneity concern of reverse causality. In the first stage, we perform regressions of *Publish* or *PublCounts* on the instrumental variable. In the second stage, we use the predicted value of *Publish* or *PublCounts* in the regression models following equations (3) and (4). Specifically, we utilize the industry-year average value of the number of journal publications (*IndustryPublication*) as an instrument. This variable is exogenous because it could not be determined by any single firm in a particular industry.

Table 4 presents the empirical results of the 2SLS instrumental variable approach. The first-stage regression results are reported in Panel A of Table 4. *IndustryPublication* is significantly positively associated with *Publish* and *PublCounts*. The significant weak identification test statistics also suggest that the instrument is relevant and valid.

Panel B of Table 4 report the second-stage regression results. The coefficients on *PredictedPublish* and *PredictedPublCounts*, estimated from the first-stage regressions, are still significant and positively associated with the dependent variables as in Table 3. Therefore, the results are still consistent with our previous findings after we deal with the reverse causality concern.

[Insert Table 4 about here]

4.2.2. Analyses using matched samples

We understand that some unobservable variables associated a firm's top-tier publications also influence the patent quality. In this subsection, we create various matched samples to deal with this concern. We match each treated firm-year observation (*Publish* = 1) with a control firm-year observation (*Publish* = 0) in the same industry and year with similar firm characteristics. The matching procedures also balance the size of the observations with publications and the observations without publications.

We perform three different methods to assign control observations to the treated observations: (i) nearest-neighbor matching using the firm characteristics include *Size*, *CashFlow*, and *R&DIntensity*, which might capture the possibility for a firm to publish research papers in the top-tier journals; (ii) propensity score matching using the logit model of *Publish* on the firm characteristics as we used in the nearest-neighbor matching; (iii) radius matching, where treated observations are matched to five nearest control observations that have propensity scores within a specific radius.[2]

Table 5 presents the univariate comparisons for the matched samples. The results show that the differences of the control variables between the treated observations and the control observations are insignificant, which indicates that treated observations are indeed having similar firm characteristics to their matched peers. The univariate comparison results of *Citations* and *RelativeCitations* show that firms with publications in top-tier journals tend to have higher *Citations* and *RelativeCitations* than their matched control firms. For instance, under the matched samples constructed using the nearest-neighbor matching approach, *Citations* and *RelativeCitations* of firms with publication are 2.23 (=396.4247 ÷ 178.1413) and 2.14 (=15.7960 ÷ 7.3825) times higher than that of their matched counterparts, respectively.

[Insert Table 5 about here]

In Table 6, we rerun the regressions of Table 3 for the various matched samples. The coefficients on *Publish* and the logarithm of one plus *PublCounts* remain positive and statistically significant across all matching procedures. Therefore, the results of the matched samples are qualitatively similar to those shown in Table 6, which remain consistent with *H1*.

[2] We use a radius (or caliper) of 0.05, which enables us to use additional control observations having propensity scores that are close enough to the treated observations. The results are robust to alternative radiuses.

[Insert Table 6 about here]

4.2.3. Difference-in-differences test

In this subsection, we perform a difference-in-differences test by using the new submission policy for the New England Journal of Medicine (NEJM) as an exogenous shock. The International Committee of Medical Journal Editors (ICMJE) and its members journals decided to require paper submissions to submit data for clinical trial registration, therefore increases the difficulty of acceptance on medical journals. This policy applies after 2005, so we compare the citations per patent before 2004 and after 2006. We consider firms with published papers in NEJM as treated firms, and find their peer control firms with papers in other top journals rather than NEJM based on propensity score matching. Because the treated firms face higher standard to publish in NEJM, the papers in NEJM would have higher impact and credibility since papers have to step to clinical trial registration. But there is no such effect for other top journals.

Table 7 compares the changes in the average *Citations* and *RelativeCitations* before and after the submission policy is applied. The results show that treated firms have better performance in innovation than the control firms. The positive statistics for the difference-in differences test also suggest that the treated firms are having higher patent citations in comparison to the control firms.

[Insert Table 7 about here]

4.2.4. Placebo test

In this subsection, we conduct a placebo test in order to show that our findings are not driven by chance. We first replace each firm with top-tier journal publications by another randomly selected firm without top-tier journal publications and term it as the placebo. We then rerun regressions for *Citations* and *RelativeCitations* as in Table 3. We keep the estimation of the coefficients on *Publish* and the logarithm of one plus *PublCounts*. By repeating this procedure for 1,000 times, we collect 1,000 estimated coefficients. Since these placeboes are randomly assigned from firms without top-tier journal publications, we expect them to have no influence on the patent quality.

Figure 1 shows the distribution of estimations of coefficient on *Publish* (Panel A for *Citations* and

Panel C for *RelativeCitations*) and the logarithm of one plus *PublCounts* (Panel B for *Citations* and Panel D for *RelativeCitations*). As shown in Panel A, the mean and median of coefficients based on the top-tier journal publications placeboes are -0.0069 and -0.0084, respectively. Given that the regression coefficient on *Publish* for *Citations* in Table 3 is 1.4224, all 1,000 trials from the placebo test do not gain the effect as large as our primary results. We find similar results when we perform the regression analysis for other models in Table 3. Therefore, the results suggest that our findings of positive effect for top-tier scientific journal publications on innovation efficiency is a real one.

[Insert Figure 1 about here]

4.3. Effect of high-impact journals

In this subsection, we perform some comparisons to test the effect of publishing in high-impact journals. We compare between firms with publications in Nature or Science and firms with publications in other top-tier journals. Similar to the model following equation (3), we carry out regression analyses for *Citations* and *RelativeCitations*. We define four indicator variables. *Publish (Nature or Science)* equals to one if the firm-year observation has journal publications in Nature or Science, and zero otherwise. *Publish (Nature)* equals to one if the firm-year observation only has journal publications in Nature, and zero otherwise. *Publish (Science)* equals to one if the firm-year observation only has journal publications in Science, and zero otherwise. *Publish (Other)* equals to one if the firm-year observation only has journal publications in other top-tier journals, and zero otherwise.

Table 8 presents the regression results. In Panel A, firms that have publications in Nature or Science, and also have publications in other top-tier journals are excluded from the sample. Models (1) and (2) use *Citations* and *RelativeCitations* as the dependent variable, respectively. In model (1), the coefficients on *Publish (Nature or Science)* and *Publish (Other)*, are 1.7302 and 1.2807, respectively. Both of these coefficients are significant at 1% level. We further conduct the Wald test on the differences between the coefficients. The test statistics of the difference between *Publish (Nature or Science)* and *Publish (Other)* is 2.66 (*p*-value=0.10). The result shows that *Publish (Nature or Science)* has higher coefficient than *Publish (Other)*, but the difference is insignificant. In model (2), the coefficients on *Publish (Nature or Science)* and *Publish (Other)*, are 0.9104 and0.5782, respectively. Both of these coefficients are also significant at 1% level. We also conduct the Wald test on the differences between these coefficients. The test statistics of the difference between *Publish (Nature or Science)* and *Publish (Other)* is 4.15 (*p*-value=0.04). The result in model (2) shows that *Publish (Nature or Science)* has higher coefficient than *Publish (Other)*, and the difference is significant at 5% level. In Panel B, firms simultaneously publish in Nature or Science or other top-tier journals are excluded from the sample. Models (3) and (4) use *Citations* and *RelativeCitations* as the dependent variable, respectively. In model (3), the coefficients on *Publish (Nature)*, *Publish (Science)*, and *Publish (Other)*, are 1.5635, 1.8511, and 1.2801, respectively. All of these coefficients are significant at 1% level. We further conduct the Wald test on the differences between the coefficients. The test statistics of the difference between *Publish (Nature)* and *Publish (Science), Publish (Nature)* and *Publish (Other),* and *Publish (Science)* and *Publish (Other),* are 0.53 (*p*-value=0.47), 0.74 (*p*-value=0.39), and 2.77 (*p*-value=0.10), respectively. The results show that there is no significant difference between *Publish (Nature)* and *Publish (Science)*, and *Publish (Science)* has higher coefficient than *Publish (Other)*. In model (4), the coefficients on *Publish (Nature)*, *Publish (Science)*, and *Publish (Other)*, are 0.7666, 1.0374, and 0.5777, respectively. All of these coefficients are significant at 1% level. We also conduct the Wald test on the differences between these coefficients. The test statistics of the difference between *Publish (Nature)* and *Publish (Science), Publish (Nature)* and *Publish (Other),* and *Publish (Science)* and *Publish (Other),* are 1.24 (*p*-value=0.26), 0.88 (*p*-value=0.35), and 5.14 (*p*-value=0.02), respectively. The results in model (4) are similar to those in model (3). Overall, the results in Table 8 imply that compared with firms with publications in other top-tier journals, firms with publications in Nature or Science enhance the patent quality more. However, there is no significant difference in patent quality between firms with Nature publications and firms with Science publications.

[Insert Table 8 about here]

4.4. Effect of geographical proximity and research quality in collaboration

In this subsection, we investigate the effect of geographical proximity and research quality in collaboration. Previous literature has suggested that by collaborating with coauthors from prestigious universities or research institutes could help companies accessing new technologies and fostering their human resources for innovations (Cohen et al., 2002; Jong and Slavova, 2014 ; Lee, 2000; Rosenberg and Nelson, 1994; Zuckeret al., 2002). Collaborating with neighbor institutions should help companies better improve innovation because the geographical spillover effect is stronger (Bishop, D'Este, and Neely, 2011). In addition, high-ranking universities have better research capacities (Mansfield and Lee, 1996), collaborating with high-ranking universities could also help companies to improve innovation quality. To test these hypotheses, we focus on the sample of firms with journal publications in the given years, and further investigate the information of their coauthors. Base on the model following equation (3), we carry out regression analyses for *Citations* and *RelativeCitations*. We define an indicator variable, *Coauthor (Top25U)*, which equals to one if the firm-year observation has journal publications with coauthors from the top 25 universities in the given year, and zero otherwise. We follow the U.S. News Rankings to determine the top 25 colleges for each year, and investigate whether firms with publications in top-tier journals have coauthor from the top-ranking schools. Similarly, we define another indicator variable, *Coauthor (Same State)*, which equals to one if the firm-year observation has journal publications with coauthors from the top 50 universities in the same State, and zero otherwise. We incorporate both of these indicator variables in the regression model to examine whether it is more important to have coauthors from prestigious universities, or it is more critical to collaborate with the universities that might not be as good as those top-ranking universities, but are geographically closer to the company. Since we could only acquire the earliest ranking data of top 50 universities from 1996, our sample in this section will restrict to 1996-2013. We also control for *CoauthorCounts* in the regression, which is the average number of coauthors of journal publications for the firm-year observation.

Table 9 presents the regression results for the firms with papers in top-tier journals. In models (1) and (2), we use *Citations* and *RelativeCitations* as the dependent variable, respectively. In model (1), the coefficients on *Coauthor (Top25U)* and *Coauthor (Same State)* are -0.2939 and 0.7961, respectively. The coefficient on *Coauthor (Same State)* is significant at 10% level, while the coefficient on *Coauthor (Top25U)* is insignificant. This result indicates that collaboration with researchers from universities in the same State not only helps companies to produce good scientific publication but would also enhance their patent quality in the long run. However, such effect is less prominent when collaborating with researchers from famous Universities in other States. In model (2), the coefficients on *Coauthor (Top25U)* and *Coauthor (Same State)* are -0.1695 and 0.4356, respectively. Similar to the result in model (1), the coefficient on *Coauthor (Same State)* is statistically significant, while the coefficient on *Coauthor (Top25U)* is negative and insignificant. The results suggest that instead of collaborating with coauthors from top schools, it would be better for firms to collaborate with the universities that are not as famous as the top colleges but are geographically nearby. In addition, the coefficients on the interaction term of *Coauthor (Top25U)* and *Coauthor (Same State)* also suggest that the effect does not come from the top-ranking universities in the same State. Therefore, the results are consistent with *H3*.

[Insert Table 9 about here]

4.5. Firm performance

Our empirical findings show that firms with publications in top-tier journals improve the innovation quality. If their innovative efficiency improves accordingly as well, then firms with publications in top-tier journal could have higher firm valuation and better operating performance (Cohen et al., 2013; Hirshleifer et al., 2013). To examine whether firms with publications in top-tier journals experience higher firm valuation and operating performance, we adopt *TobinsQ* and *ProfitMargin* as the dependent variable. *TobinsQ* is defined as the firm's market value divided by its book value of property, plant and equipment, where the firm’s market value is computed using the market value of common equity plus the book value of debt minus the firm’s current assets. *ProfitMargin* is the sum of income before extraordinary items, interest expenses, and total income taxes, divided by total sales. We then regress *TobinsQ* and *ProfitMargin* following the models similar to equations (3) and (4).

Table 10 presents the regression analyses for firm performance. In models (1) and (2), we employ *TobinsQ* as the dependent variable; while in models (3) and (4), we employ *ProfitMargin* as the dependent variable. In models (1) and (2), the coefficients on *Publish* and the logarithm of one plus *PublCount* are 0.0126 and 0.0057, which are both significant, suggesting that firms with publications in top-tier journals have higher firm valuation. In models (3) and (4), the coefficients on *Publish* and the logarithm of one plus *PublCount* are 0.1189 and 0.0703, both are also significant at 1% level, suggesting that firms with publications in top-tier journals have better operating performance, which is consistent with *H5*.

[Insert Table 10 about here]

5. Additional Analyses

5.1. Robustness tests

It is possible that our previous findings can be dominated by the companies that frequently have publications in top journals. To mitigate this concern, we perform robustness test by removing firms from the sample after their first time to have publication in top-tier journals. The regression results are shown in Panel A of Table 11. The coefficients on *Publish* and the logarithm of one plus *PublCounts* still remain positive and statistically significant, which indicates the results are not driven by the firms that frequently have papers published in top-tier journals.

We also exclude the firms without any patent in the whole sample period from our sample, and the regression results are shown in Panel B of Table 11. We can find that the positive effect of paper publications on patent citations still hold for the firms with at least one patent in the whole sample period.

In addition, we use an alternative measure for product competition based on the 10K product descriptions following Hoberg and Phillips (2016), and the results are presented in Panel C of Table 11. The results are robust for the coefficients on *Publish* and the logarithm of one plus *PublCounts*, and the coefficients on *HHI (TNIC)* are negatively significant, which indicates that higher product market concentration will lead to worse patent performance.

We adopt other widely used measures for the innovation quality in Panel D of Table 11. Hall, Jaffe, and Trajtenberg (2001) propose two different measures to examine the patent innovations, which are generality and originality. *Generality* is defined as the sum of squared ratio of forward-citations received divided by the number of total forward-citations that belong to the same IPC classification. A greater generality of a patent, the patent is more influential in various technology areas. *Originality* is measured as the sum of squared ratio of backward-references made divided by the number of total backward-references that belong to the same IPC classification. Patent originality captures the diversification of technology knowledge background, and a patent gains higher originality if it does not rely on the knowledge pool of a specific area. Panel D of Table 11 presents the regression results for *Generality* and *Originality*. In models (1) and (2), we employ *Generality* as the dependent variable, while in models (3) and (4) we employ *Originality* as the dependent variable. In models (1) and (3), the coefficients on *Publish* are 0.0733 and 0.0938, which are both significant at the 1% level, suggesting that publish in top-tier journals is positively related to patent generality and patent originality. In models (2) and (4), the coefficients on the logarithm of one plus *PublCounts* are 0.0350 and 0.0398, which are a significant at the 5% level and 1% level, respectively. These results also suggesting that the number of publications in top-tier journals is also positively related to patent generality and patent originality. Therefore, we conclude that, in addition to patent citations, other innovation quality measures are also better for firms with scientific papers in top-tier journals than for firms without.

[Insert Table 11 about here]

5.2. Effect of previous publications

The R&D activities are long-term practices that take companies several years to have research outcomes. Therefore, previous publications in top-tier journals are also likely to have an influence on future innovation performance. We include the lagged terms of *Publish* and *PublCounts* in the regression model accordingly. The untabulated results suggest that previous publications in scientific journals indeed have positive effects on future innovation performance as well.

5.3. Results in different industries

In this subsection, we compare the results for the 12 industry groups based on Fama and French (1997). Since we exclude financial and utilities industries, there are 10 industry groups remain in our sample. We find the results are more pronounced in consumer durables industry, manufacturing industry, wholesale and retail industry, and healthcare, medical equipment, and drugs industry.

We also conduct a similar examination by further sorting the scientific journals into different categories. The untabulated results suggest that publishing papers in multidisciplinary science better influences innovation performance in consumer durables industry; publishing papers in biotechnology and applied microbiology, and physical chemistry, will have better effect on innovation performance in manufacturing industry; publishing papers in genetics and heredity, and multidisciplinary chemistry, will have better effect on innovation performance in wholesale and retail industry; and publishing papers in general and internal medicine, and cell biology, will have better effect on innovation performance in healthcare, medical equipment, and drugs industry.

6. Conclusion

This paper examines the effect of top-tier scientific journal publications on the innovation quality of firms. We collect the detailed journal publication data from the Web of Science to count the number of journal publications for each firm in the U.S. during the sample period from 1980 to 2013. We find that firms with publications in top-tier journals tend to have better innovation outputs measured by patent citations. The results are robust in the instrumental variable regressions as well as the regressions for the matched samples based on different matching procedures. We use the change in submission policy for NEJM as an exogenous shock, and we still find a positive effect of publications on patent citations. We also conduct a placebo test, and we do not find similar results in the pseudo samples that are randomly constructed, which indicates that our findings of positive effect for top-tier scientific journal publications on innovation efficiency is a real one.

In further analyses, we find that the positive effect of journal publications on patent citations is stronger for firms with publications in Nature or Science than for firms that only have publications in other top-tier journals. Because it requires a solid research capability or good collaborations with universities and scientific institutes to produce top-tier publications, these companies are more likely to produce innovation outputs with better quality.

Furthermore, our analyses of the collaboration with different kinds of coauthors show that instead of collaborating with coauthors from top schools, firms benefit more when they collaborate with universities nearby, which could facilitate them to recruit or train human capital for research and innovation. Further analyses also show that firms with top-tier journal publications experience better firm performance. Overall, our study complements the extant literature on corporate innovations by providing empirical evidence that publications in top-tier scientific journals benefit the firms.

References

[1] Adams, James D., Grant C. Black, J. Roger Clemmons, and Paula E. Stephan, 2005, Scientific teams and institutional collaborations: Evidence from U.S. universities, 1981-1999, *Research Policy* 34, 259-285.

[2] Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffth, and Peter Howitt, 2005, Competition and innovation: An inverted-u relationship, *Quarterly Journal of Economics* 120, 701-728.

[3] Aghion, Philippe, John Van Reenen, and Luigi Zingales, 2013, Innovation and institutional ownership, *American Economic Review* 103, 277-304.

[4] Baker, Monya, 2016, Is there a reproducibility crisis in science? *Nature* 533, 452-454.

[5] Becker-Blease, John R., 2011, Governance and innovation, *Journal of Corporate Finance* 17, 947-958.

[6] Bena, Jan and Kai Li, 2014, Corporate innovations and mergers and acquisitions, *Journal of Finance* 69, 1923-1960.

[7] Bereskin, Frederick L., Terry L. Campbell II, and Po-Hsuan Hsu, 2016, Corporate philanthropy, research networks, and collaborative innovation, *Financial Management* 45, 175–206.

[8] Bishop, Kate, Pablo D'Este, and Andy Neely, 2011, Gaining from interactions with universities: Multiple methods for nurturing absorptive capacity, *Research Policy* 40, 30-40.

[9] Cohen, Lauren, Karl Diether, and Christopher Malloy, 2013, Misvaluing innovation, *Review of Financial Studies* 26, 635-666.

[10] Cohen, Wesley M., Richard R. Nelsen, and John P. Walsh, 2002, Links and impacts: The influence of public research on industrial R&D, *Management Science* 48, 1-23.

[11] Czarnitzki, Dirk, and Andrew A. Toole, 2011, Patent protection, market uncertainty, and R&D investment, *Review of Economics and Statistics* 93, 147-159.

[12] Fama, Eugene F., and Kenneth R. French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153-193.

[13] Fang, Vivian W., Xuan Tian, and Sheri Tice, 2014, Does stock liquidity enhance or impede firm innovation? *Journal of Finance* 69, 2085-2125.

[14] Furman, Jeffrey L., Kyle Jensen, and Fiona Murray, 2012, Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine. *Research Policy*, 41(2), 276-290.

[15] Ginarte, Juan C., and Walter G. Park, 1997, Determinants of patent rights: A cross-national study, *Research Policy* 26, 293-301.

[16] Gittelman, Michelle and Bruce Kogut, 2003, Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns, *Management Science* 49, 366-382.

[17] Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg, 2001, The NBER patent citation data file: Lessons, insights and methodological tools, NBER Working Papers w8498, National Bureau of Economic Research.

[18] Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg, 2005, Market value and patent citations, *RAND Journal of Economics* 36, 16-38.

[19] He, Jie and Xuan Tian, 2013, The dark side of analyst coverage: The case of innovation, *Journal of Financial Economics* 109, 856-878.

[20] Hicks, Diana, 1995, Published papers, tacit competencies and corporate management of the public/private character of knowledge, *Industrial and Corporate Change* 4, 401–424.

[21] Hirshleifer, David, Po-Hsuan Hsu, and Dongmei Li, 2013, Innovative efficiency and stock returns, *Journal of Financial Economics* 107, 632-654.

[22] Hoberg, Gerard, and Gordon Phillips, 2016, Text-based network industries and endogenous product differentiation, *Journal of Political Economy* 124, 1423-1465.

[23] Ioannidis, John P. A., 2005, Why most published research findings are false, *PLOS Medicine* 2, 696-701.

[24] Jaffe, Adam B., 1989, The real effects of academic research, *American Economic Review* 79, 957-970.

[25] Jong, Simcha, and Kremena Slavova, 2014, When publications lead to products: The open science conundrum in new product development, *Research Policy* 43, 645-654.

[26] Lee, Hsing-fen, and Marcela Miozzo, 2019, Which types of knowledge-intensive business services firms collaborate with universities for innovation? *Research Policy* 48, 1633-1646.

[27] Lee, Yong S., 2000, The sustainability of university-industry research collaboration: An empirical assessment, *Journal of Technology Transfer* 25, 111-133.

[28] Lerner, Josh, 2002, 150 years of patent protection, *American Economic Review* 92, 221-225.

[29] Mansfield, Edwin, 1991, Academic research and industrial innovation, *Research Policy* 20, 1-12.

[30] Mansfield, Edwin, and Jeong-Yeon Lee, 1996, The modern university: Contributor to industrial innovation and recipient of industrial R&D support, *Research Policy* 25, 1047-1058.

[31] Mansfield, Edwin, 1998, Academic research and industrial innovation: An update of empirical findings, *Research Policy* 26, 773-776.

[32] Mullard, Asher, 2011, Reliability of 'new drug target' claims called into question, *Nature Reviews Drug Discovery* 10, 643-644.

[33] Munafo, Marcus R., and Jonathan Flint, 2010, How reliable are scientific studies? *British Journal of Psychiatry* 197, 257-258.

[34] Prinz, Florian, Thomas Schlange, and Khusru Asadullah, 2011, Believe it or not: How much can we rely on published data on potential drug targets? *Nature Reviews Drug Discovery* 10, 712.

[35] Rosenberg, Nathan, and Richard R. Nelson, 1994, American universities and technical advance in industry, *Research Policy* 23, 323-348.

[36] Schroter, Sara, Nick Black, Stephen Evans, Fiona Godlee, Lyda Osorio, and Richard Smith, 2008, What errors do peer reviewers detect, and does training improve their ability to detect them? *Journal of the Royal Society of Medicine* 101, 507-514.

[37] Scott, Alister, 2007, Peer review and the relevance of science, *Futures* 39, 827-845.

[38] Simeth, Markus, and Michele Cincera, 2016, Corporate science, innovation, and firm value, *Management Science* 62, 1970-1981.

[39] Zucker, Lynne G., Michael R. Darby, and Jeff S. Armstrong, 2002, Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology, *Management Science* 48, 138-153.

Table 1. Summary Statistics

This table presents the summary statistics. The sample consists of the U.S. firms with financial data in the Compustat database, covering the period from 1980 to 2013. Refer to Table A1 for definitions of the variables. All continuous variables are winsorized at the 1st and 99th percentiles.

Variable	Mean	Standard Deviation	25th Percentile	Median	75th Percentile	N
Publish	0.0203	0.1410	0.0000	0.0000	0.0000	82,547
PublCounts	0.1019	1.3280	0.0000	0.0000	0.0000	82,547
PatentCounts	2.9473	11.6346	0.0000	0.0000	0.0000	82,547
Citations	67.4536	291.1851	0.0000	0.0000	0.0000	82,547
RelativeCitations	2.6156	10.6633	0.0000	0.0000	0.0000	82,547
Generality	0.0722	0.1682	0.0000	0.0000	0.0000	82,547
Originality	0.0871	0.1826	0.0000	0.0000	0.0000	82,547
Assets	2.9918	13.5677	0.0851	0.2911	1.2751	82,547
Sales	2.7114	12.9677	0.0840	0.3009	1.2622	82,547
TobinsQ	0.0427	0.1151	0.0052	0.0132	0.0368	82,547
ProfitMargin	-0.0231	0.7157	0.0235	0.0776	0.1407	82,547
Size	5.8808	1.9010	4.4437	5.6738	7.1508	82,547
HHI	0.1671	0.1404	0.0726	0.1239	0.2100	82,547
AnalystCover	1.1353	2.2756	0.0000	0.0000	1.0000	82,547
R&DIntensity	0.0908	0.4038	0.0000	0.0014	0.0498	82,547
CashFlow	0.0621	0.1438	0.0179	0.0821	0.1378	82,547
CapitalLaborRatio	0.1713	0.7261	0.0167	0.0319	0.0751	82,547

Table 2. Univariate Comparison

This table presents the univariate comparisons for our sample based on *Publish*. Refer to Table A1 for definitions of the variables. *, **, and *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

	Publish = 1		*Publish* = 0		Difference	
Variable	Mean	Median	Mean	Median	Mean	Median
PublCounts	5.0227	2.0000	0.0000	0.0000	5.0227***	2.0000***
PatentCounts	18.0471	3.5000	2.6348	0.0000	15.4123***	3.5000***
Citations	392.3579	50.3003	60.7284	0.0000	331.6295***	50.3003***
RelativeCitations	15.6788	2.0423	2.3452	0.0000	13.3336***	2.0423***
Generality	0.2195	0.2041	0.0691	0.0000	0.1504***	0.2041***
Originality	0.2662	0.3107	0.0834	0.0000	0.1828***	0.3107***
Assets	15.2747	2.5762	2.7376	0.2855	0.0519***	0.0315***
Sales	13.0722	1.9157	2.4970	0.2970	-0.5911***	0.0122***
TobinsQ	0.0936	0.0444	0.0417	0.0129	1.6467***	2.2000***
ProfitMargin	-0.6022	0.0897	-0.0111	0.0775	-0.0549***	-0.0665***
Size	7.4941	7.8541	5.8474	5.6541	1.7284***	2.0000***
HHI	0.1133	0.0585	0.1682	0.1250	0.5947***	0.1468***
AnalystCover	2.8286	2.0000	1.1002	0.0000	-0.0676***	-0.0046***
R&DIntensity	0.6734	0.1468	0.0787	0.0000	-0.0600***	0.0374***
CashFlow	-0.0041	0.0775	0.0635	0.0821	5.0227***	2.0000***
CapitalLaborRatio	0.1126	0.0687	0.1726	0.0313	15.4123***	3.5000***
N	1,674		80,873			

Table 3. Regression Analysis for the Relationship between Top-Tier Journal Publication and Patent Citation

This table presents regression results for the sample from 1980 to 2013. Observations are at the firm-year level. *Citations* and *RelativeCitations* are employed as dependent variables. *Citations* is the number of truncation-adjusted citations divided by the number of patent applications. *RelativeCitations* is the citations per patent divided by the median of citations per patent in the same IPC classification. Refer to Table A1 for definitions of all other variables. *t*-statistics based on standard errors robust to clustering by firms are reported in parentheses. *, **, and *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: OLS Regression				
Dependent Variable	$Log(1+Citations_{i,t+1})$		$Log(1+RelativeCitations_{i,t+1})$	
Model	(1)	(2)	(3)	(4)
Publish	1.4224***		0.7069***	
	(5.63)		(5.14)	
Log(1+PublCounts)		0.7319***		0.3840***
		(2.88)		(2.90)
Size	0.1518***	0.1535***	0.0857***	0.0862***
	(9.93)	(9.97)	(10.92)	(10.86)
HHI	0.0036	0.0053	0.0344	0.0356
	(0.02)	(0.03)	(0.45)	(0.46)
Log(1+AnalystCover)	0.3995***	0.4032***	0.1773***	0.1787***
	(14.57)	(14.61)	(12.70)	(12.70)
R&DIntensity	0.4519***	0.4919***	0.1629***	0.1820***
	(8.89)	(9.51)	(7.07)	(7.82)
TobinsQ	0.8729***	0.8886***	0.3010***	0.3083***
	(6.98)	(7.02)	(5.22)	(5.28)
CashFlow	0.2753***	0.2668***	0.1550***	0.1506***
	(3.03)	(2.93)	(3.76)	(3.65)
CapitalLaborRatio	-0.0399**	-0.0405**	-0.0200***	-0.0202***
	(-2.48)	(-2.51)	(-2.78)	(-2.80)
IndustryFE	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes
N	82,547	82,547	82,547	82,547
Adjusted R-squared	0.20	0.20	0.19	0.18

(continued on next page)

Table 3. ***(continued)***

Panel B: Tobit Regression				
Dependent Variable	*Log(1+Citations$_{i,t+1}$)*		*Log(1+RelativeCitations$_{i,t+1}$)*	
Model	(1)	(2)	(3)	(4)
Publish	2.5608***		1.1885***	
	(5.37)		(5.18)	
Log(1+PublCounts)		1.1243**		0.5429**
		(2.39)		(2.42)
Size	0.5848***	0.5961***	0.3019***	0.3065***
	(9.90)	(10.05)	(10.84)	(10.92)
HHI	-0.1749	-0.1782	-0.0252	-0.0257
	(-0.25)	(-0.25)	(-0.08)	(-0.08)
Log(1+AnalystCover)	1.5742***	1.5896***	0.6958***	0.7024***
	(16.84)	(16.90)	(16.12)	(16.16)
R&DIntensity	1.5037***	1.5878***	0.6324***	0.6709***
	(11.45)	(11.98)	(10.87)	(11.45)
TobinsQ	2.4392***	2.4782***	0.9743***	0.9919***
	(7.53)	(7.57)	(6.87)	(6.91)
CashFlow	0.5165	0.4884	0.2955*	0.2819
	(1.30)	(1.23)	(1.68)	(1.60)
CapitalLaborRatio	-1.2936***	-1.3005***	-0.5905***	-0.5935***
	(-2.70)	(-2.71)	(-2.88)	(-2.89)
IndustryFE	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes
N	82,547	82,547	82,547	82,547
Pseudo R-squared	0.11	0.11	0.14	0.14

Table 4. Instrumental Variables Approach for the Relationship between Top-Tier Journal Publication and Patent Citation

This table presents two-stage least-squares regression results for the sample from 1980 to 2013. Observations are at the firm-year level. In the first-stage regression, *Publish* and *PublCounts* are employed as dependent variables. *Publish* is an indicator variable, which equals to one if the firm-year observation has journal publications, and zero otherwise. *PublCounts* is the total number of journal publications for the firm-year observation. *IndustryPublication* is the average number of journal publications for the firms in the same 2-digit SIC industry. In the second-stage, *Citations* and *RelativeCitations* are employed as dependent variables. *Citations* is the number of truncation-adjusted citations divided by the number of patent applications. *RelativeCitations* is the citations per patent divided by the median of citations per patent in the same IPC classification. Refer to Table A1 for definitions of all other variables. *t*-statistics based on standard errors robust to clustering by firms are reported in parentheses. Kleibergen-Paap rk Wald *F*-statistics for the weak identification test of the instrument variable are reported for the first-stage regression. *, **, and *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: First-Stage Regression Result		
Dependent Variable	*Publish*	*Log(1+PublCounts)*
Model	(1)	(2)
IndustryPublication	0.0794***	0.1722***
	(4.42)	(4.94)
Size	0.0105***	0.0181***
	(6.38)	(4.78)
HHI	-0.0082	-0.0184
	(-0.70)	(-0.89)
Log(1+AnalystCover)	0.0135***	0.0212***
	(6.86)	(5.27)
R&DIntensity	0.0507***	0.0443***
	(7.37)	(4.87)
TobinsQ	0.0244**	0.0260*
	(2.55)	(1.80)
CashFlow	-0.0021	0.0077
	(-0.28)	(0.63)
CapitalLaborRatio	-0.0023***	-0.0036***
	(-3.30)	(-2.98)
IndustryFE	Yes	Yes
YearFE	Yes	Yes
N	82,547	82,547
Adjusted R-squared	0.13	0.11
Weak identification test	19.57***	24.42***

(continued on next page)

Table 4. ***(continued)***

Panel B: Second-Stage Regression Result				
Dependent Variable	*Log(1+Citations$_{i,t+1}$)*		*Log(1+RelativeCitations$_{i,t+1}$)*	
Model	(1)	(2)	(3)	(4)
PredictedPublish	7.0989***		3.2759***	
	(3.44)		(3.33)	
Log(1+PredictedPublCounts)		3.2721***		1.5099***
		(3.63)		(3.55)
Size	0.0919***	0.1071***	0.0586***	0.0656***
	(3.30)	(4.51)	(4.35)	(5.66)
HHI	0.0460	0.0482	0.0535	0.0545
	(0.30)	(0.30)	(0.72)	(0.71)
Log(1+AnalystCover)	0.3218***	0.3485***	0.1421***	0.1544***
	(8.46)	(10.72)	(7.66)	(9.63)
R&DIntensity	0.1743	0.3892***	0.0372	0.1364***
	(1.45)	(5.72)	(0.65)	(4.30)
TobinsQ	0.7337***	0.8220***	0.2380***	0.2787***
	(5.30)	(6.28)	(3.72)	(4.63)
CashFlow	0.2877***	0.2479***	0.1606***	0.1422***
	(2.93)	(2.61)	(3.62)	(3.32)
CapitalLaborRatio	-0.0280*	-0.0320**	-0.0146**	-0.0165**
	(-1.72)	(-2.00)	(-2.02)	(-2.32)
IndustryFE	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes
N	82,547	82,547	82,547	82,547
Adjusted R-squared	0.19	0.19	0.18	0.18

Table 5. Univariate Comparison in the Matched Samples

This table presents the univariate comparison results for the matched samples. We match each treated firm-year observation (*Publish* = 1) with a control firm-year observation (*Publish* = 0) in the same industry and year using three different matching methods. Each matched sample includes treated observations and their matched peers with similar *Size*, *CashFlow*, and *R&DIntensity*. Refer to Table A1 for definitions of all other variables. *, **, and *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Nearest-Neighbor Matching			
Variable	Treated (*Publish* = 1)	Control (*Publish* = 0)	Difference
Size	7.6613	7.5426	0.1187
CashFlow	0.0158	0.0138	0.0020
R&DIntensity	0.5755	0.5578	0.0177
Citations	396.4247	178.1413	218.2835***
RelativeCitations	15.7960	7.3825	8.4135***
N	1,504	1,504	
Panel B: Propensity Score Based on Logit			
Variable	Treated (*Publish* = 1)	Control (*Publish* = 0)	Difference
Size	7.4941	7.5237	-0.0296
CashFlow	-0.0041	-0.0002	-0.0040
R&DIntensity	0.6734	0.5991	0.0743
Citations	392.3579	145.3841	246.9738***
RelativeCitations	15.6787	6.0196	9.6591***
N	1,674	1,674	
Panel C: Radius Matching			
Variable	Treated (*Publish* = 1)	Control (*Publish* = 0)	Difference
Size	7.4405	7.4044	0.0361
CashFlow	-0.0065	-0.0095	0.0030
R&DIntensity	0.6746	0.6261	0.0485
Citations	387.8476	148.8485	238.9991***
RelativeCitations	15.6062	6.2840	9.3222***
N	1,621	4,878	

Table 6. Regression Analysis for the Matched Samples

This table presents regression results for the matched samples. Observations are at the firm-year level. *Citations* and *RelativeCitations* are employed as dependent variables. *Citations* is the number of truncation-adjusted citations divided by the number of patent applications. *RelativeCitations* is the citations per patent divided by the median of citations per patent in the same IPC classification. Refer to Table A1 for definitions of all other variables. *t*-statistics based on standard errors robust to clustering by firms are reported in parentheses. *, **, and *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Nearest-Neighbor Matching				
Dependent Variable	*Log(1+Citations$_{i,t+1}$)*		*Log(1+RelativeCitations$_{i,t+1}$)*	
Model	(1)	(2)	(3)	(4)
Publish	1.2136***		0.5600***	
	(4.66)		(3.96)	
Log(1+PublCounts)		0.5124*		0.2485*
		(1.89)		(1.73)
Size	0.0110	-0.0403	0.0579	0.0335
	(0.13)	(-0.51)	(1.29)	(0.79)
HHI	0.5967	0.9610	0.7344	0.9013
	(0.33)	(0.53)	(0.75)	(0.91)
Log(1+AnalystCover)	0.6581***	0.7378***	0.3129***	0.3457***
	(4.88)	(5.39)	(4.27)	(4.60)
R&DIntensity	0.0806	0.0894	0.0174	0.0214
	(0.88)	(0.96)	(0.38)	(0.46)
TobinsQ	1.3036**	1.3628**	0.7546***	0.7807***
	(2.58)	(2.53)	(2.74)	(2.69)
CashFlow	-0.2139	-0.1498	-0.1266	-0.0944
	(-0.31)	(-0.22)	(-0.35)	(-0.26)
CapitalLaborRatio	-0.2777	-0.2948	-0.1385	-0.1467
	(-0.87)	(-1.01)	(-0.84)	(-0.96)
IndustryFE	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes
N	3,008	3,008	3,008	3,008
Adjusted R-squared	0.18	0.16	0.19	0.17

(continued on next page)

Table 6. ***(continued)***

Panel B: Propensity Score Based on Logit				
Dependent Variable	*Log(1+Citations$_{i,t+1}$)*		*Log(1+RelativeCitations$_{i,t+1}$)*	
Model	(1)	(2)	(3)	(4)
Publish	1.4861***		0.6978***	
	(5.88)		(5.04)	
Log(1+PublCounts)		0.6597**		0.3277**
		(2.50)		(2.35)
Size	0.0249	-0.0434	0.0579	0.0246
	(0.33)	(-0.63)	(1.44)	(0.68)
HHI	0.0126	0.0390	0.4199	0.4458
	(0.01)	(0.03)	(0.51)	(0.52)
Log(1+AnalystCover)	0.5696***	0.6386***	0.2694***	0.2966***
	(4.92)	(5.46)	(4.26)	(4.60)
R&DIntensity	0.2448***	0.2450***	0.0920**	0.0919**
	(3.28)	(3.13)	(2.45)	(2.36)
TobinsQ	1.1706**	1.3554**	0.6693**	0.7511**
	(2.40)	(2.58)	(2.42)	(2.57)
CashFlow	0.3194	0.4176	0.1673	0.2147
	(0.61)	(0.79)	(0.62)	(0.79)
CapitalLaborRatio	-0.2835	-0.3551	-0.1012	-0.1333
	(-1.19)	(-1.62)	(-0.82)	(-1.18)
IndustryFE	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes
N	3,348	3,348	3,348	3,348
Adjusted R-squared	0.21	0.19	0.20	0.18

(continued on next page)

Table 6. ***(continued)***

Panel C: Radius Matching				
Dependent Variable	*Log(1+Citations$_{i,t+1}$)*		*Log(1+RelativeCitations$_{i,t+1}$)*	
Model	(1)	(2)	(3)	(4)
Publish	1.2238***		0.5584***	
	(4.75)		(4.06)	
Log(1+PublCounts)		0.6110**		0.2918**
		(2.38)		(2.18)
Size	0.1220**	0.1187**	0.1002***	0.0980***
	(2.56)	(2.49)	(3.96)	(3.86)
HHI	-0.6075	-0.6019	-0.2141	-0.2081
	(-0.92)	(-0.90)	(-0.61)	(-0.58)
Log(1+AnalystCover)	0.8552***	0.8675***	0.4106***	0.4153***
	(9.81)	(9.88)	(8.58)	(8.61)
R&DIntensity	0.2219***	0.2360***	0.0904***	0.0963***
	(3.88)	(4.06)	(3.40)	(3.56)
TobinsQ	0.4560	0.5059	0.1604	0.1816
	(1.21)	(1.32)	(0.87)	(0.97)
CashFlow	-0.7816**	-0.7837**	-0.3631**	-0.3629**
	(-2.43)	(-2.43)	(-2.33)	(-2.33)
CapitalLaborRatio	-0.2571**	-0.2616**	-0.1354**	-0.1372***
	(-2.29)	(-2.35)	(-2.55)	(-2.60)
IndustryFE	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes
N	3,242	3,242	3,242	3,242
Adjusted R-squared	0.19	0.19	0.19	0.19

Table 7. Difference-in-Differences Test on New England Journal of Medicine Sample before 2004 and after 2006

This table compares the changes in *Citations* and *RelativeCitations* between firms with publications in New England Journal of Medicine (Treated) and firms without (Control). We match each treated firm with a control firm in the same industry and year based on propensity score matching method. The clinical trial registration is required for New England Journal of Medicine after 2005, so we compare the average *Citations* (or *RelativeCitations*) before 2004 and after 2006. *Citations* is the number of truncation-adjusted citations divided by the number of patent applications. *RelativeCitations* is the citations per patent divided by the median of citations per patent in the same IPC classification. *, **, and *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: *Citations*			
	Before Clinical Trial Registration Required	After Clinical Trial Registration Required	Difference
Treated	3.5715	2.6133	-0.9582***
Control	4.0481	1.8559	-2.1922***
Difference	-0.4766*	0.7574**	1.2340***
N	216	128	
Panel B: *RelativeCitations*			
	Before FDA Trial Registration Required	After FDA Trial Registration Required	Difference
Treated	1.7047	1.0384	-0.6663***
Control	1.8783	0.6842	-1.1941***
Difference	-0.1736	0.3542*	0.5278**
N	216	128	

Table 8. Comparison between Publishing in Nature, Science, and Other Top-Tier Journals

This table presents regression results for the sample from 1980 to 2013. Observations are at the firm-year level. *Citations* and *RelativeCitations* are employed as dependent variables. *Citations* is the number of truncation-adjusted citations divided by the number of patent applications. *RelativeCitations* is the citations per patent divided by the median of citations per patent in the same IPC classification. In Panel A, firms that have publications in Nature or Science, and also have publications in other top-tier journals are excluded from the sample. *Publish (Nature or Science)* is an indicator variable, which equals to one if the firm-year observation only has journal publications in Nature or Science, and zero otherwise. *Publish (Other)* is an indicator variable, which equals to one if the firm-year observation only has journal publications in other top-tier journals, and zero otherwise. In Panel B, firms simultaneously publish in Nature or Science or other top-tier journals are excluded from the sample. *Publish (Nature)* is an indicator variable, which equals to one if the firm-year observation only has journal publications in Nature, and zero otherwise. *Publish (Science)* is an indicator variable, which equals to one if the firm-year observation only has journal publications in Science, and zero otherwise. *Publish (Other)* is an indicator variable, which equals to one if the firm-year observation only has journal publications in other top-tier journals, and zero otherwise. Refer to Table A1 for definitions of all other variables. *t*-statistics based on standard errors robust to clustering by firms are reported in parentheses. *F*-statistics for the test for differences between regression coefficients are reported. *, **, and *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Firms with publications in Nature or Science, and also publish in other top-tier journals are excluded		
Dependent Variable	$Log(1+Citations_{i,t+1})$	$Log(1+RelativeCitations_{i,t+1})$
Model	(1)	(2)
Publish (Nature or Science)	1.7302***	0.9104***
	(5.98)	(5.31)
Publish (Other)	1.2807***	0.5782***
	(6.83)	(5.41)
Size	0.1544***	0.0863***
	(10.21)	(11.15)
HHI	0.0026	0.0307
	(0.02)	(0.42)
Log(1+AnalystCover)	0.4055***	0.1798***
	(14.78)	(12.88)
R&DIntensity	0.4493***	0.1639***
	(9.06)	(7.46)
TobinsQ	0.8456***	0.2860***
	(6.82)	(5.02)
CashFlow	0.2849***	0.1581***
	(3.17)	(3.91)
CapitalLaborRatio	-0.0412**	-0.0208***
	(-2.56)	(-2.89)
IndustryFE	Yes	Yes
YearFE	Yes	Yes
N	81,968	81,968
Adjusted R-squared	0.19	0.18
Test: *Publish (Nature or Science)=Publish (Other)*	2.66	4.15**

(continued on next page)

Table 8. ***(continued)***

Panel B: Firms simultaneously publish in Nature or Science or other top-tier journals are excluded		
Dependent Variable	*Log(1+Citations$_{i,t+1}$)*	*Log(1+RelativeCitations$_{i,t+1}$)*
Model	(1)	(2)
Publish (Nature)	1.5635***	0.7666***
	(4.56)	(3.65)
Publish (Science)	1.8511***	1.0374***
	(5.21)	(4.94)
Publish (Other)	1.2801***	0.5777***
	(6.82)	(5.41)
Size	0.1545***	0.0863***
	(10.22)	(11.16)
HHI	-0.0006	0.0280
	(-0.00)	(0.38)
Log(1+AnalystCover)	0.4059***	0.1800***
	(14.79)	(12.89)
R&DIntensity	0.4489***	0.1636***
	(9.05)	(7.45)
TobinsQ	0.8451***	0.2859***
	(6.81)	(5.02)
CashFlow	0.2846***	0.1576***
	(3.17)	(3.90)
CapitalLaborRatio	-0.0413**	-0.0208***
	(-2.56)	(-2.90)
IndustryFE	Yes	Yes
YearFE	Yes	Yes
N	81,944	81,944
Adjusted R-squared	0.19	0.18
Test: *Publish (Nature)= Publish (Science)*	0.53	1.24
Test: *Publish (Nature)=Publish (Other)*	0.74	0.88
Test: *Publish (Science)=Publish (Other)*	2.77*	5.14**

Table 9. Regression Analysis for the Relationship between Collaboration with Different Coauthors and Patent Citation

This table presents regression results for the sample from 1996 to 2013. Observations are at the firm-year level. *Citations* and *RelativeCitations* are employed as dependent variables. *Citations* is the number of truncation-adjusted citations divided by the number of patent applications. *RelativeCitations* is the citations per patent divided by the median of citations per patent in the same IPC classification. *Coauthor (Top25U)* is an indicator variable, which equals to one if the firm-year observation has journal publications with coauthors from top 25 universities, and zero otherwise. *Coauthor (Same State)* is an indicator variable, which equals to one if the firm-year observation has journal publications with coauthors from top 50 universities in the same State, and zero otherwise. *CoauthorCounts* is the average number of coauthors of journal publications for the firm-year observation. Refer to Table A1 for definitions of all other variables. *t*-statistics based on standard errors robust to clustering by firms are reported in parentheses. *, **, and *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

Dependent Variable	$Log(1+Citations_{i,t+1})$	$Log(1+RelativeCitations_{i,t+1})$
Model	(1)	(2)
Coauthor (Top25U)	-0.2939	-0.1695
	(-1.08)	(-1.21)
Coauthor (Same State)	0.7961*	0.4356*
	(1.74)	(1.73)
Coauthor (Top25U)×Coauthor (Same State)	0.2636	0.1657
	(0.46)	(0.53)
Size	0.1974	0.1592**
	(1.49)	(2.25)
HHI	0.8923	1.1577
	(0.31)	(0.71)
Log(1+AnalystCover)	-0.0527	-0.0265
	(-0.31)	(-0.28)
R&DIntensity	0.2080*	0.0703
	(1.96)	(1.22)
TobinsQ	2.0815***	1.3256***
	(3.74)	(4.61)
CashFlow	0.3975	0.1533
	(0.52)	(0.38)
CapitalLaborRatio	0.4149	0.2543
	(0.59)	(0.67)
Log(1+CoauthorCounts)	0.0933	0.0124
	(0.56)	(0.14)
IndustryFE	Yes	Yes
YearFE	Yes	Yes
N	1,068	1,068
Adjusted R-squared	0.23	0.27

Table 10. Regression Analysis for the Relationship between Top-Tier Journal Publication and Firm Performance

This table presents regression results for the sample from 1980 to 2013. Observations are at the firm-year level. *TobinsQ* and *ProfitMargin* are employed as dependent variables. *TobinsQ* is the firm's market value divided by the book value of property, plant and equipment, where the firm's market value is computed using the market value of common equity plus the book value of debt minus the firm's current assets. *ProfitMargin* is the sum of income before extraordinary items, interest expenses, and total income taxes, divided by total sales. Refer to Table A1 for definitions of all other variables. *t*-statistics based on standard errors robust to clustering by firms are reported in parentheses. *, **, and *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

Dependent Variable	$TobinsQ_{i,t+1}$		$ProfitMargin_{i,t+1}$	
Model	(1)	(2)	(3)	(4)
Publish	0.0126***		0.1189***	
	(3.11)		(3.81)	
Log(1+PublCounts)		0.0057**		0.0703***
		(2.42)		(4.22)
Size	-0.0012***	-0.0012***	0.0129***	0.0129***
	(-5.94)	(-5.83)	(8.01)	(8.03)
HHI	-0.0115***	-0.0115***	-0.0654***	-0.0651***
	(-5.35)	(-5.35)	(-4.40)	(-4.38)
Log(1+AnalystCover)	0.0030***	0.0030***	0.0027	0.0028
	(5.57)	(5.64)	(0.87)	(0.92)
R&DIntensity	0.0143***	0.0147***	-1.0592***	-1.0562***
	(7.85)	(8.12)	(-42.31)	(-42.44)
TobinsQ	0.4396***	0.4397***	-0.1566***	-0.1555***
	(23.37)	(23.37)	(-3.15)	(-3.12)
CashFlow	0.0256***	0.0256***	0.7859***	0.7851***
	(6.25)	(6.23)	(20.84)	(20.81)
CapitalLaborRatio	-0.0028***	-0.0028***	-0.0092	-0.0093
	(-10.91)	(-10.92)	(-1.55)	(-1.55)
IndustryFE	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes
N	82,547	82,547	82,547	82,547
Adjusted R-squared	0.40	0.40	0.55	0.55

Table 11. Robustness Test for the Relationship between Top-Tier Journal Publication and Patent Citation

This table presents regression results for the sample from 1980 to 2013. In Panel A, we exclude the firms after their first time to get on top-journal from the sample. In Panel B, we exclude the firms without any patent in the whole sample period from the sample. In Panel C, we use an alternative measure for *HHI* based on the product descriptions in 10K reports following Hoberg and Phillips (2016). In Panel D, we use *Generality* and *Originality* as alternative measures for innovation quality. Refer to Table A1 for definitions of all other variables. *t*-statistics based on standard errors robust to clustering by firms are reported in parentheses. *, **, and *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Exclude Firms after First Publishing				
Dependent Variable	*Log(1+Citations$_{i,t+1}$)*		*Log(1+RelativeCitations$_{i,t+1}$)*	
Model	(1)	(2)	(3)	(4)
Publish	1.5152***		0.7189***	
	(10.39)		(9.61)	
Log(1+PublCounts)		1.4245***		0.6851***
		(9.08)		(8.39)
ControlVariables	Yes	Yes	Yes	Yes
IndustryFE	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes
N	79,218	79,218	79,218	79,218
Adjusted R-squared	0.18	0.18	0.16	0.16

Panel B: Exclude Firms without Patent				
Dependent Variable	*Log(1+Citations$_{i,t+1}$)*		*Log(1+RelativeCitations$_{i,t+1}$)*	
Model	(1)	(2)	(3)	(4)
Publish	0.7045***		0.3853***	
	(2.94)		(2.80)	
Log(1+PublCounts)		0.4397*		0.2589*
		(1.77)		(1.89)
ControlVariables	Yes	Yes	Yes	Yes
IndustryFE	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes
N	34,739	34,739	34,739	34,739
Adjusted R-squared	0.21	0.21	0.22	0.22

(continued on next page)

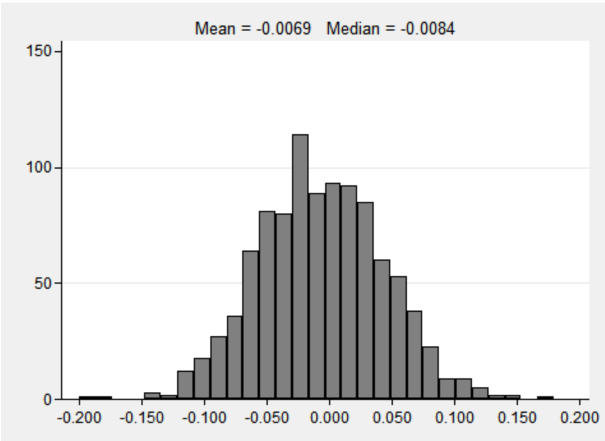
Table 11. ***(continued)***

Panel C: Alternative *HHI* Measure				
Dependent Variable	*Log(1+Citations$_{i,t+1}$)*		*Log(1+RelativeCitations$_{i,t+1}$)*	
Model	(1)	(2)	(3)	(4)
Publish	1.3875***		0.6880***	
	(6.23)		(5.60)	
Log(1+PublCounts)		0.7623***		0.3980***
		(3.31)		(3.28)
ControlVariables	Yes	Yes	Yes	Yes
IndustryFE	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes
N	44,437	44,437	44,437	44,437
Adjusted R-squared	0.23	0.22	0.21	0.21

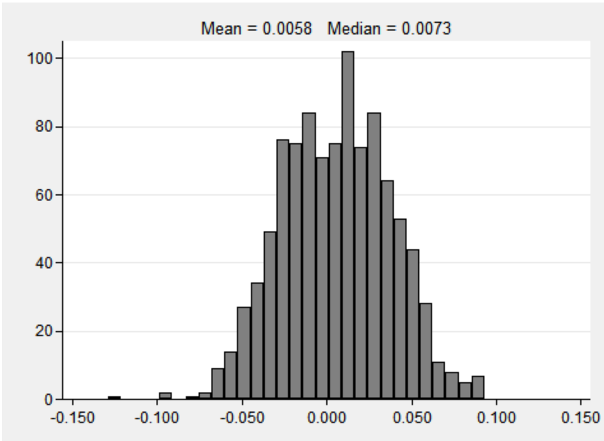
Panel D: Alternative Innovation Quality Measures				
Dependent Variable	*Generality$_{i,t+1}$*		*Originality$_{i,t+1}$*	
Model	(1)	(2)	(3)	(4)
Publish	0.0733***		0.0938***	
	(5.31)		(5.84)	
Log(1+PublCounts)		0.0350**		0.0398***
		(2.56)		(2.58)
ControlVariables	Yes	Yes	Yes	Yes
IndustryFE	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes
N	82,547	82,547	82,547	82,547
Adjusted R-squared	0.16	0.16	0.15	0.14

Figure 1. Placebo Test for the Relationship between Top-Tier Journal Publication and Patent Citation

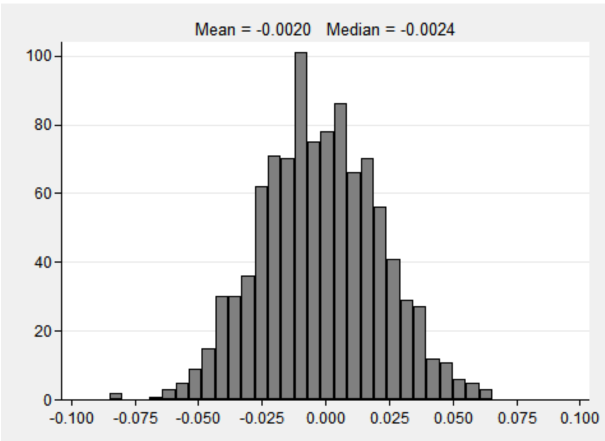
This figure displays the distributions for coefficient estimates from the regressions in Table 3. We perform the regression analysis for 1,000 trials upon top-tier journal publications placebos. *Citations* and *RelativeCitations* are employed as dependent variables. We replace each firm with top-tier journal publications by another randomly selected firm without top-tier journal publications, which we term it as the placebo. We perform regression analyses for *Citations* (Panel A and Panel B) and *RelativeCitations* (Panel C and Panel D) with top-tier journal publications placebos. The figures show the distribution of coefficient estimates of the *Publish* and *Log(1+PublCounts)* for 1,000 trials.



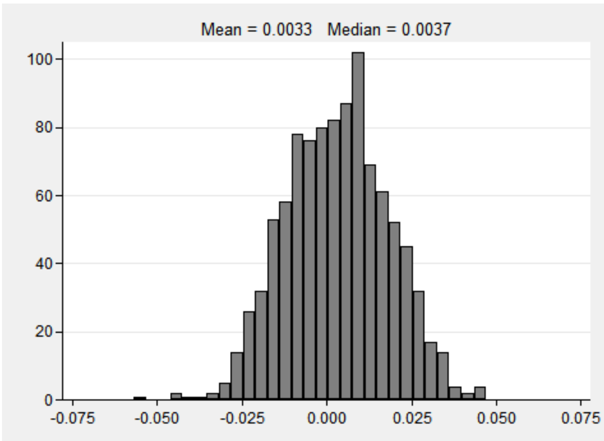
Panel A. Distribution of coefficients on Publish



Panel B. Distribution of coefficients on Log(1+PublCounts)



Panel C. Distribution of coefficients on Publish



Panel D. Distribution of coefficients on Log(1+PublCounts)

Appendix

Table A1. Variable Definitions

This table presents the definitions of the variables.

Variable	Definition
Publish	Dummy variable, which equals to one if the firm-year observation has journal publications, and zero otherwise.
PublCounts	The total number of journal publications for the firm-year observation.
PatentCounts	The total number of patent applications for the firm-year observation, which is adjusted by the median number of patent applications in the same International Patent Classification (IPC).
Citations	The number of truncation-adjusted citations divided by the number of patent applications.
RelativeCitations	The citations per patent divided by the median of citations per patent in the same IPC classification.
Generality	The sum of squared ratio of forward-citations received divided by the number of total forward-citations that belong to the same IPC classification.
Originality	The sum of squared ratio of backward-references made divided by the number of total backward-references that belong to the same IPC classification.
Assets	Total assets in billion.
Sales	Total sales in billion.
TobinsQ	The firm's market value divided by the book value of property, plant and equipment, where the firm's market value is computed using the market value of common equity plus the book value of debt minus the firm's current assets.
ProfitMargin	The sum of income before extraordinary items, interest expenses, and total income taxes, divided by total sales.
Size	The logarithm of total assets.
HHI	Herfindahl-Hirschman Index, which is the squared sales-based market share upon three-digit SIC industry.
HHI (TNIC)	Herfindahl-Hirschman Index, which is based on the product descriptions in 10K reports following Hoberg and Phillips (2016).
AnalystCover	The number of analysts following the firm in one month before earnings announcement, obtained from I/B/E/S.
R&DIntensity	The R&D expenditures divided by total sales.
CashFlow	The income before extraordinary items minus total accruals (changes in current assets plus changes in short-term debt, minus the sum of changes in cash, changes in current liabilities, and depreciation expenses), divided by average total assets.
CapitalLaborRatio	The property, plant and equipment divided by the number of employees.
IndustryPublication	The average number of journal publications for the firms in the same two-digit SIC industry.
Coauthor (Top25U)	Dummy variable, which equals to one if the firm-year observation has journal publications with coauthors from top 25 universities, and zero otherwise.
Coauthor (Same State)	Dummy variable, which equals to one if the firm-year observation has journal publications with coauthors from top 50 universities in the same State, and zero otherwise.
CoauthorCounts	The average number of coauthors of journal publications for the firm-year observation.

Table A2. Number of Firm-Year Observations Publishing in Each Journal

This table presents the number of firm-year observations that have published research papers in each journal.

Journal	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2009	2010-2013	Overall
Nature	42	80	85	101	75	65	50	498
Science	83	93	94	126	96	52	45	589
Cell	7	19	39	46	25	18	12	166
J. Am. Chem. Soc.	51	47	65	86	118	73	48	488
Lancet	33	37	44	34	33	33	31	245
New. Engl. J. Med.	23	46	51	51	58	71	74	374
P. Natl. Acad. Sci. USA.	0	98	129	171	157	136	94	785
Nat. Biotechnol.	0	0	0	41	43	50	48	182
Nat. Cell. Biol.	0	0	0	1	15	4	5	25
Nat. Chem.	0	0	0	0	0	0	5	5
Nat. Chem. Biol.	0	0	0	0	0	5	9	14
Nat. Clim. Change.	0	0	0	0	0	0	1	1
Nat. Commun.	0	0	0	0	0	0	20	20
Nat. Genet.	0	0	6	36	34	29	26	131
Nat. Geosci.	0	0	0	0	0	0	1	1
Nat. Immunol.	0	0	0	0	21	18	8	47
Nat. Mater.	0	0	0	0	8	7	10	25
Nat. Med.	0	0	0	42	44	31	19	136
Nat. Methods.	0	0	0	0	4	15	12	31
Nat. Nanotechnol.	0	0	0	0	0	6	9	15
Nat. Neurosci.	0	0	0	3	12	3	5	23
Nat. Photonics.	0	0	0	0	0	7	9	16
Nat. Phys.	0	0	0	0	0	6	5	11
Nat. Protoc.	0	0	0	0	0	13	5	18
Nat. Rev. Cancer.	0	0	0	0	1	2	0	3
Nat. Rev. Clin. Oncol.	0	0	0	0	0	0	4	4
Nat. Rev. Drug. Discov.	0	0	0	0	3	14	12	29
Nat. Rev. Genet.	0	0	0	0	0	0	1	1
Nat. Rev. Immunol.	0	0	0	0	0	0	2	2
Nat. Rev. Microbiol.	0	0	0	0	0	0	1	1
Sci. Signal.	0	0	0	0	0	4	7	11
Sci. Transl. Med.	0	0	0	0	0	2	31	33
Any Journal Listed Above	128	167	218	293	336	285	247	1,674

Table A3. Number of Publications in Each Journal

This table presents the total number of publications for the sample firms in each journal.

Journal	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2009	2010-2013	Overall
Nature	104	187	194	215	160	102	74	1,036
Science	154	254	278	280	185	105	63	1,319
Cell	12	41	63	78	27	28	16	265
J. Am. Chem. Soc.	131	178	238	255	273	175	103	1,353
Lancet	49	76	83	64	54	70	61	457
New. Engl. J. Med.	28	79	92	107	124	162	135	727
P. Natl. Acad. Sci. USA.	0	362	440	473	411	298	180	2,164
Nat. Biotechnol.	0	0	0	55	50	69	73	247
Nat. Cell. Biol.	0	0	0	1	20	4	5	30
Nat. Chem.	0	0	0	0	0	0	6	6
Nat. Chem. Biol.	0	0	0	0	0	5	12	17
Nat. Clim. Change.	0	0	0	0	0	0	1	1
Nat. Commun.	0	0	0	0	0	0	32	32
Nat. Genet.	0	0	8	53	45	44	40	190
Nat. Geosci.	0	0	0	0	0	0	1	1
Nat. Immunol.	0	0	0	0	36	25	10	71
Nat. Mater.	0	0	0	0	14	9	10	33
Nat. Med.	0	0	0	58	67	38	25	188
Nat. Methods.	0	0	0	0	5	19	14	38
Nat. Nanotechnol.	0	0	0	0	0	12	20	32
Nat. Neurosci.	0	0	0	4	16	3	5	28
Nat. Photonics.	0	0	0	0	0	8	13	21
Nat. Phys.	0	0	0	0	0	13	9	22
Nat. Protoc.	0	0	0	0	0	13	5	18
Nat. Rev. Cancer.	0	0	0	0	1	2	0	3
Nat. Rev. Clin. Oncol.	0	0	0	0	0	0	4	4
Nat. Rev. Drug. Discov.	0	0	0	0	4	16	21	41
Nat. Rev. Genet.	0	0	0	0	0	0	1	1
Nat. Rev. Immunol.	0	0	0	0	0	0	2	2
Nat. Rev. Microbiol.	0	0	0	0	0	0	1	1
Sci. Signal.	0	0	0	0	0	4	10	14
Sci. Transl. Med.	0	0	0	0	0	2	44	46
Total	478	1,177	1,396	1,643	1,492	1,226	996	8,408