# The Effect of Social Media on Corporate Innovation: Evidence from Seeking Alpha Coverage†

Qiyang He, Henry Leung, Buhui Qiu, and Zhou Zhou

This version: April 2024

**Abstract**

Seeking Alpha (SA) is the most popular crowdsourced social media platform specializing in the financial analysis of U.S. firms and attracts over 17 million visitors per month. We find that firm coverage initiation on SA significantly promotes corporate innovation activities. SA coverage promotes corporate innovation by disseminating innovation-related information about the covered firm to external investors, thereby alleviating the firm's financial constraints. Moreover, firms with higher information asymmetry, firms facing greater product market competition, and firms having lower levels of managerial short-termism reveal a greater increase in innovation outcomes following SA coverage initiation. Taken together, our findings suggest that third-party generated information on specialized social media mitigates the information asymmetry between firms and investors on corporate innovation activities and encourages corporate innovation.

***Keywords:*** Social Media; Seeking Alpha; Corporate Innovation; Information Asymmetry; Market Competition

***JEL:*** G30; G32

---

† He: qiyang.he@sydney.edu.au, University of Sydney Business School, the University of Sydney. Leung: henry.leung@sydney.edu.au, University of Sydney Business School, the University of Sydney. Qiu (corresponding author): buhui.qiu@sydney.edu.au, University of Sydney Business School, the University of Sydney. Zhou: littleboat365@gmail.com, the Commonwealth Bank of Australia. We are grateful to Bo Becker (Department Editor), an anonymous associate editor, and two anonymous referees for their very helpful and constructive comments and suggestions throughout the review process. We also thank Sabri Boubaker, Fengyu Li, Adrian Lee, Jane Luo, Hui Guo, Wanling Qiu, Gary Tian, Terry Pan, and seminar participants at the Dongbei University of Finance and Economics, Swansea University, Macquarie University, and the University of Sydney for helpful comments and suggestions. All errors are our own.

# The Effect of Social Media on Corporate Innovation: Evidence from Seeking Alpha Coverage

This version: April 2024


## Abstract

Seeking Alpha (SA) is the most popular crowdsourced social media platform specializing in the financial analysis of U.S. firms and attracts over 17 million visitors per month. We find that firm coverage initiation on SA significantly promotes corporate innovation activities. SA coverage promotes corporate innovation by disseminating innovation-related information about the covered firm to external investors, thereby alleviating the firm's financial constraints. Moreover, firms with higher information asymmetry, firms facing greater product market competition, and firms having lower levels of managerial short-termism reveal a greater increase in innovation outcomes following SA coverage initiation. Taken together, our findings suggest that third-party generated information on specialized social media mitigates the information asymmetry between firms and investors on corporate innovation activities and encourages corporate innovation.



***Keywords:*** Social Media; Seeking Alpha; Corporate Innovation; Information Asymmetry; Market Competition

***JEL:*** G30; G32

## 1. Introduction

About two-thirds of adults in the U.S. access information through social media (Matsa and Shearer, 2018). Thus, social media platforms play a pivotal role in acting as information intermediaries that analyze and disseminate firm-specific information. The literature shows that information from social media alters the trading behavior of investors (e.g., Lee, Hutton and Shu, 2015; Heimer, 2016) and predicts stock returns (e.g., Chen, De, Hu and Hwang, 2014; Campbell, DeAngelis and Moon, 2019). In addition, firms use social media to gather customer insights and promote the development of new products (Roberts, Piller and Lüttgens, 2016). Seeking Alpha (SA, hereafter) is the most popular crowdsourced social media platform specializing in the financial analysis of U.S. firms and attracts over 17 million visitors per month.[1] The literature on SA emphasizes its role in facilitating the incorporation of firm-specific information into the short-term performance of stocks (e.g., Chen et al., 2014; Campbell et al, 2019; Ding, Zhou and Li, 2020).

In contrast to more straightforward investment projects, such as capital expenditure, corporate innovation projects are often perceived as enigmatic or "black boxes." This is due to their intricate and multifaceted processes and decision-making structures, which can be difficult for outsiders to comprehend. Even those within the firm may struggle to understand the inner workings of innovation projects due to competitive pressures, intellectual property concerns, and the need to maintain a strategic advantage. The absence of transparency in innovation projects can foster uncertainty and skepticism among external investors, posing challenges in garnering support and building trust for these initiatives. The complex nature of innovation projects requires investors to have a deeper understanding of the risks and benefits involved, but this is often difficult to achieve when decision-making processes are opaque. The lack of transparency in innovation projects can be a significant obstacle to securing external financing and support, leading to substantial financial constraints in funding innovation projects. In this study, we show that *SA coverage can reduce financial constraints and encourage corporate innovation activities*.

Compared with traditional media, financial analyses provided by SA accelerate information dissemination by reaching a broader audience almost instantly. In contrast to sell-side

[1] Statistics in early 2021 shows that SA has over 10 million registered users. The one-year registration fee is $240, which unlocks 1 million articles, stock ratings, author ratings and quant ratings. The average visit duration of SA was 4 times more than that of *The Economist*, *Barron's*, and the *Wall Street Journal*. SA publishes over 7,000 articles and earnings transcripts covering more than 8,000 firms each month. For details, see https://static.seekingalpha.com/uploads/pdf_income/sa_media_kit_01.06.21.pdf.

equity research conducted by Wall Street brokerages, the firm analyses on SA are provided by buy-side analysts, many of whom are experienced individual and institutional investors. SA's contributors include industry experts, fund managers and analysts, some of whom have been featured on influencing TV programs such as CNBC, Bloomberg TV and Fox Business. They share their personal opinions and analyses in stock picking and portfolio management via SA. Moreover, authors of SA articles are required to disclose their personal positions in stocks and business relationship with the issuers of stocks analyzed in their articles. This required disclosure of stock position does not bias the financial analyses on SA (Campbell et al., 2019).

In addition, SA's incentive scheme encourages contributors to publish original and high-quality articles. SA financially rewards authors who contribute articles with high-quality analyses, especially high-quality articles with large page-views and/or small-cap research.[2] SA also grants active authors free access to its research library and exclusive access to firm executives. Firms pay close attention to what is written about them on SA, and some firm managers also contribute to SA via articles and comments. Therefore, writing for SA is a very efficient, visible way to demonstrate contributors' expertise to their peers, potential clients, and current and future employers, and the credibility of the information in SA articles is positively linked to authors' reputation. An editorial panel, acting as an independent party, will verify the quality of each article before publishing. Compared to other social media outlets such as Facebook and Twitter, SA articles do not have a word limit and SA's incentive scheme and verification process ensure the credibility and quality of its articles.

We find that contributing authors on SA often employ their specialized industry experience to guide investors towards a better understanding of the implications of innovation activities within firms. For example, the excerpt below from an SA article pertains to the analyses of corporate innovation. The article titled "Dextera Surgical: Innovation In Surgical Stapling Puts It On Path Towards Profitability" is authored by Terry Trover, who is a professional in medical equipment

[2] SA rewards $10 to $13 for every 1000 page view and up to $1000 for top ideas; it rewards $150 to $500 per article for selected small-cap research. For more details, see https://seekingalpha.com/article/2134803-how-much-does-seeking-alpha-pay-its-contributors and https://seekingalpha.com/page/payment-terms.

products.[3] This article explains the practical application and improvement of Dextera Surgical's innovation based on the author's industry experience. An excerpt of this article reads as follows:

> *"Dextera Surgical (OTC:DXTR) is a very small ($10.3M market cap) medical device manufacturer. … Dextera has reached an inflection point. Dextera has now invented and introduced a manual surgical stapling device called MicroCutter 5/80 whose features are said to make it a preferred approach for video-assisted lung surgery, called VATS. … Now, an even less invasive VATS procedure called "micro lobectomy" has been enabled by an innovation in surgical stapling technology. … Proprietary stapler enables patients to leave hospital days sooner than present standard of care."*

As explained in this article, DXTR's invention in surgical stapling is a breakthrough innovation which produces smaller incisions during the removal of lung tumors and shortens the recovery period of patients compared to traditional surgery. In this instance, the expertise of the author has helped to highlight the positive effects of the innovation activity, which investors may not have realized if an assessment of the performance outlook of DXTR was formed solely based on the losses incurred. The example above shows that SA authors who have specialized knowledge in a subject area or relevant industry experience may provide investors a deeper insight into the innovation activities of a firm.

The literature shows that interpersonal communication shapes investors' decision making in stock market (e.g., Shiller and Pound, 1989; Ivković and Weisbenner, 2007). In a similar vein, as shown by the above example, investors may gain a much better insight on corporate innovation activities through SA articles. In addition, the statistics in 2012 show that 54% of SA contributing authors are financial professionals, 13.8% of its audiences are C-suite executives, and 46% of its audiences follow technology industries.[4] The professionalism of SA authors, the involvement of firms' management teams, the SA audience's interest in technology development, and the financial rewards for high-quality, small-cap research create an effective external environment for firms to disseminate important information related to their innovation activities. Therefore, SA is an ideal

[3] For more details of article content, see https://seekingalpha.com/article/4046318-dextera-surgical-innovation-surgical-stapling-puts-path-towards-profitability. Terry Trover (https://www.linkedin.com/in/terrytrover/) has rich working and leadership experience in the medical equipment industry.

[4] See http://static.cdn-seekingalpha.com/uploads/2013/1/SA-MediaKit-download-2013.pdf.

platform to examine the impact of social media on corporate innovation. We find that firm coverage initiation on SA significantly promotes corporate innovation quantity and quality.

When investigating the effect of SA coverage on corporate innovation activities, endogeneity concerns (such as reverse causality and omitted variable bias) may arise because SA coverage is potentially non-random. For instance, SA authors may be more likely to cover innovative firms. To alleviate such endogeneity concerns, we use the initiation of SA coverage that is unrelated to a firm's innovation activities as an exogenous event. In particular, we employ a quasi-natural experiment that involves a sample of treatment firms with SA coverage initiation unrelated to the covered firm's innovation activities between 2011 and 2015, and a sample of industry-peer control firms that are not covered by SA and are carefully matched on pre-coverage firm characteristics (including pre-coverage innovation outputs) using the propensity score matching (PSM) approach. Considering that the coverage initiation of a particular firm by SA is one event, our study involves multiple events. A treatment firm and up to five matched industry-peer control firms related to an event constitute a cohort. We then stack all the cohorts together to form the sample. Importantly, we carefully exclude from the sample all SA coverage initiations that are due to either the covered firm's innovation activities or the author's business relationship with the firm. Therefore, the assignment of treatment (i.e., SA coverage initiation) in the final sample can be viewed as random when investigating the effect of SA coverage on corporate innovation. In particular, we show that ex-ante innovation performance and firm-level characteristics cannot predict SA coverage initiation in the matched sample.

Using a stacked-cohort Difference-in-Differences (DID) regression framework (e.g., Gormley and Matsa, 2011, 2016), we next include three years pre- and post-coverage-initiation and compare the innovation changes of treatment firms from pre- to post-coverage periods with those of control firms. We allow three years for firm innovation to take effect and thus examine the effect of SA coverage on three-year-ahead corporate innovation outcomes. By using a seven-year event window, our final sample covers the period from 2008 to 2018. We include cohort-firm and cohort-year fixed effects to control for firm and time heterogeneities within each cohort, respectively. Thus, our stacked-cohort DID specification exploits within-cohort-firm variation in innovation activities occasioned by SA coverage initiation.

We employ fixed-effects Poisson regressions for our DID analyses (Cohn, Liu, and Wardlaw, 2022; Chen and Roth, 2023). The findings reveal that the initiation of SA coverage results in significantly higher levels of patent quantity and quality. Subsequent to SA coverage initiation, treatment firms on average generate 88.3% more patents and increase patent's forward citations by 129.8% relative to control firms. Importantly, by using a dynamic DID regression framework and parallel trends plots, we find that the treatment effect of SA coverage on corporate innovation outcomes begins only in the year following the initiation of coverage and continues into future years. The fact that the treatment effect does not exist in any of the years before the coverage initiation suggests that the positive treatment effect of SA coverage on corporate innovation is unlikely to be driven by potential nonparallel innovation trends before the coverage initiation. That is, the positive effect of SA coverage on corporate innovation is likely causal.

We propose that SA coverage stimulates corporate innovation by disseminating innovation-related information about the covered firm to external investors, thereby alleviating the firm's financial constraints. Consequently, we investigate the impact of SA coverage initiation on the covered firm's financing activities and financial constraints. Consistent with the idea that equity is a preferred means of financing innovation activities over debt (Brown, Martinsson, and Petersen, 2013; Acharya and Xu, 2017), we observe that treatment firms, in comparison to matched control firms, experience a significant increase in raising new equity following SA coverage initiation. Notably, there is no corresponding increase in new debt. Our findings also indicate that the financial constraints of treatment firms indeed become more relaxed relative to the matched control firms after the initiation of SA coverage.

We further find that the positive impact of SA coverage initiation on corporate innovation is more pronounced in firms with higher information asymmetry before SA coverage initiation. Specifically, we find that the positive effect of SA coverage initiation on innovation is more pronounced in firms with ex-ante lower analyst following, higher abnormal accruals, higher idiosyncratic volatility, and/or higher effective spread. Since information asymmetry increases the firm's cost of capital (e.g., Diamond and Verrecchia, 1991; Easley and O'Hara, 2004), higher information asymmetry can make it much more costly to fund long-term innovation projects and thus impede corporate innovation. SA provides third-party generated and analyzed firm-specific information, which reduces information asymmetry between firms and investors regarding the firm's innovation activities.[5] Our findings thus highlight the informational role of specialized social media in encouraging corporate innovation.

Moreover, we find that the effect of SA coverage initiation on corporate innovation is more pronounced in firms with higher product market fluidity, lower product market concentration and lower market power before coverage initiation. Given that intense market competition can make innovation even more crucial for firms to survive the "creative destruction" process (Schumpeter, 1942; Aghion and Howitt, 1992), our results suggest that the initiation of SA coverage that helps convey information regarding corporate innovation to the market, incentivizes more innovation activities particularly if the covered firm's product market spaces are more competitive.

He and Tian (2013) show that firms followed by a greater number of sell-side financial analysts produce fewer and lower-quality patents. They interpret their findings as being consistent with financial analysts exerting excessive short-term pressures on firm managers to meet short-term earnings targets, thereby impeding firms' long-term investment in innovation. It is important to note that financial analyst coverage and SA coverage differ significantly in nature. Financial analysts focus on the earnings per share of firms and issue short-term earnings forecasts. They then make stock recommendations based on these short-term forecasts (e.g., Brennan, Jegadeesh, and Swaminathan, 1993; Hong, Lim, and Stein, 2000; He and Tian, 2013). The forecasts provided by financial analysts become targets for firm managers to achieve (e.g., Jensen and Fuller, 2002).[6]

In contrast, SA authors do not focus on the short-term earnings of firms. Importantly, they do not issue short-term earnings forecasts and, therefore, do not exert short-term pressures on firm managers. Many SA authors are industry experts, often employing their specialized industry experience to guide investors towards a better understanding of the implications of innovation activities within firms. Their articles provide insightful information about the firm's innovation 'black box' and reduce information asymmetry regarding corporate innovation activity. As information asymmetry around the innovation 'black box' decreases, firm financial constraints

[5] Ding et al. (2020) show that the articles from SA complement the formal disclosures from firms and have an incremental effect on unravelling firm-specific information to stock markets.

[6] Failing to meet analyst earnings forecasts can result in significantly negative consequences for firm managers, including stock price crashes, bonus cuts, and even forced managerial turnovers (e.g., Matsunaga and Park, 2001; Bartov, Givoly, and Hayn, 2002; Mergenthaler, Rajgopal, and Srinivasan, 2012; Ak, Rossi, Sloan, and Tracy, 2016). Thus, corporate managers have strong incentives to beat or meet short-term earnings targets set by financial analysts, and they are willing to cut long-term innovation projects to achieve these short-term targets (Graham, Harvey, and Rajgopal, 2005; Bhojraj, Hribar, Picconi, and McInnis, 2009).

relax. Consequently, more positive NPV innovation projects can be funded, leading to a greater quantity and higher quality of future innovation outputs.

We posit that the impact of reduced information asymmetry, resulting from the initiation of SA coverage, on corporate innovation may be less pronounced for firms characterized by elevated levels of managerial short-termism. In situations where managers exhibit heightened career concerns that lead to a more short-term focus, the diminished information asymmetry related to corporate innovation may not be sufficiently motivating for them to pursue long-term innovation projects. Consistent with this conjecture, our findings reveal that the positive effect of SA coverage initiation on the enhancement of corporate innovation outcomes is primarily observed in firms with low transient institutional ownership, high dedicated institutional ownership, high CEO age, and/or extensive CEO tenure before the initiation of SA coverage.[7]

Finally, we use the number of SA articles covering the firm in a year as a proxy for SA article intensity and use the average number of comments received per article on a firm in a year as a proxy for average impact per article, to investigate how SA article intensity and impact per article about a firm affects the firm's future innovation activities. We construct a firm-year panel dataset (all firms in this panel dataset are SA-covered firms) for this analysis. To address potential endogeneity issues (e.g., reverse causality), we employ two-stage least squares (2SLS) instrumental variable regressions and use the aggregate number of followers of unique SA authors writing about the firm in a year as the instrument. If SA authors are more popular (i.e., if they have a higher number of followers before writing about a firm), they are more likely to write more articles about the firms they cover, and their articles are likely to have a greater average impact per article (the relevance condition). However, the past accumulated personal popularity of SA authors (i.e., the number of followers they have before writing about a firm) should not directly affect the innovation outcomes of the focal firm, except through the SA articles they write about the firm and the average impact these articles generate (the exclusion restriction).

Consistent with the earlier findings from the DID analyses, we find that the instrumented article intensity and instrumented average comment intensity are both significantly and positively

[7] We implement a series of additional tests to verify the robustness of our main findings. We find that both innovation input measured by R&D expenses and SG&A expenses, and innovation output measured alternatively by innovation efficiency, patent value, efficiency value, patent originality and patent generality, increase in treatment firms relative to control firms after SA coverage initiation. Moreover, the results remain robust when we control for traditional news coverage such as the number of news articles and the tone of news. The results are also robust to controlling for sell-side analyst coverage, stock returns and return volatility.

related to future corporate innovation outcomes, suggesting that SA coverage indeed promotes corporate innovations. Moreover, we decompose the annual total number of SA articles written about a firm into innovation-focused versus non-innovation-focused numbers of articles. Our 2SLS results further reveal that it is the instrumented innovation-focused article intensities that significantly and positively affect future corporate innovation outcomes. Taken together, these findings based on panel dataset and 2SLS regressions strongly corroborate the earlier findings based on SA coverage initiation events and stacked-cohort DID analyses.

## 2. Literature and Hypotheses

Our study contributes to two strands of literature. First, we contribute to the literature on the effects of social media on firm performance. This literature emphasizes the informational role of social media and its instant effects on stock market and product market. For example, the literature shows that investor opinions and sentiments transmitted through social media predict stock returns and earnings surprises (e.g., Luo, Zhang and Duan, 2013; Chen et al., 2014; Bartov, Faurel and Mohanram, 2018). Moreover, social media coverage reduces stock return comovement especially for firms with higher financial reporting opacity (Ding et al., 2020). Disseminating news via social media reduces information asymmetry (Blankespoor, Miller and White, 2014). In addition, social media facilitates firms in marketing (Roberts et al., 2016) and knowledge transfer (Leonardi and Treem, 2012). Our study fills the gap in this literature by revealing the beneficial impact of social media on corporate innovation.

Second, we contribute to the literature on economic factors driving corporate innovation. Specifically, empirical evidence shows that investors' tolerance for failure (Tian and Wang, 2014), managerial compensation (Ederer and Manso, 2013), mergers and acquisitions (Bena and Li, 2014; Seru, 2014), stock liquidity (Fang, Tian and Tice, 2014), financial development (Hsu, Tian and Xu, 2014), banking deregulation (Chava, Oettl, Subramanian and Subramanian, 2013), financial analyst coverage (He and Tian, 2013; Guo, Pérez-Castrillo and Toldrà-Simats, 2019), institutional ownership (Aghion, Van Reenen and Zingales, 2013; Luong, Moshirian, Nguyen, Tian and Zhang, 2017; Brav, Jiang, Ma and Tian, 2018), corporate venture capital (Chemmanur, Loutskina and Tian, 2014), financial dependence (Lerner, Sorensen and Strömberg, 2011; Acharya and Xu, 2017), laws and regulations (Atanassov, 2013; Brown et al., 2013; Acharya, Baghai and Subramanian,

2014) and media reporting (Dai, Shen and Zhang, 2021), influence corporate innovation.[8] However, the extant literature does not study whether and how social media coverage affects corporate innovation activities. Our study addresses this gap by demonstrating that social media, such as SA, functions as a vital information intermediary between firms and investors and plays a pivotal role in promoting corporate innovation activities.

The literature suggests that SA provides reliable firm-specific information and facilitates the incorporation of firm-specific information into stock prices. For example, Chen et al. (2014) find that investors' opinions and sentiment transmitted through SA articles reliably predict stock returns and earnings surprises. Campbell et al. (2019) document significant two-day abnormal stock returns surround the SA articles' release dates, indicating that firm-specific analyses on SA are informative. Ding et al. (2020) document a negative relation between SA coverage and stock return comovement, and this relation is more pronounced for firms with higher financial reporting opacity. Thus, the extant empirical evidence in the literature reveals that SA plays an important role in reducing information asymmetry between corporate insiders and outside investors.

The corporate innovation process is often likened to a "black box," characterized by high levels of information asymmetry due to its complexity and the need to safeguard proprietary knowledge from competitors (Aghion and Tirole, 1994). Information asymmetry can significantly increase a firm's cost of capital, making it more difficult and costly to fund long-term, risky innovation projects (Diamond and Verrecchia, 1991; Easley and O'Hara, 2004), which can impede corporate innovation activities. SA coverage can play a crucial role in promoting corporate innovation by reducing information asymmetry, for the following reasons.

First, SA articles are paraphrased information, which enhances the readability of firm-specific information. Since the terminology in describing innovation might be hard to understand by the general public, outside investors may not realize the full potential value of a firm's innovation. SA coverage enables professional investors and industry experts to explain the firm's

[8] Dai et al. (2021) find media coverage impedes corporate innovation due to the media's role of pressuring managers to deliver short-term performance. In contrast, we find a positive relationship between SA coverage and corporate innovation because SA provides third-party generated and analyzed information on firm innovation, which mitigates information asymmetry between firms and outside investors. Our findings support the literature suggesting that the effects of traditional media and social media are very different. For example, the literature suggests that the opinions from social media communication are distinctive from those released in traditional media such as news (Ding et al., 2020). Moreover, Cookson and Niessner (2020) suggest that investors' sentiment dispersion in social media and firms' news coverage have different impacts on stock trading volume. Chernova (2014) reports that stock opinions on SA beats Wall Street analyst reports and financial news articles.

innovation in a way that is easy to understand by non-professional outsiders. After understanding the firm's innovation projects, existing shareholders may provide greater support for its innovation activities and have greater tolerance for the firm's temporary profit downturns, and potential investors may be attracted to invest in the firm. Hence, SA coverage helps reduce the level of information asymmetry and financial constraints on the covered firm's innovation activities, which should increase the firm's future innovation outputs in the long run.

Second, SA articles are verified and evaluated firm-specific information generated by the third party. SA users can exchange opinions by posting articles and discuss the article content by responding with comments. This interactive communication process facilitates the exchange of diverse perspectives, which can enhance the informativeness of original articles and amplify their impact. Together, SA articles and comments offer comprehensive and objective third-party analyses on firms' innovation activities. The interactive communication process enables outside investors to better understand the implications of corporate innovation projects, which reduces the information asymmetry about the firm's innovation activities.

Furthermore, innovation is essential in our knowledge-based economy. With intense competition in product markets, innovation has become increasingly vital for firms in today's corporate world. Corporate innovation helps firms to gain a competitive advantage over competitors (e.g., Fang et al., 2014; Chemmanur, et al., 2014; Acharya and Xu, 2017), and firms under intense market competition must keep innovating to survive the "creative destruction" process (Schumpeter, 1942; Aghion and Howitt, 1992). Correa and Ornaghi (2014) document a positive relation between competition and innovation measured in patents and citations. Kettler (1998) also argues that *competitiveness* is synonymous with *innovativeness* in R&D intensive industries because firms in such industries need to innovate to survive and grow. In competitive product market environments, innovation is thus crucial not only for a firm's long-term growth but also for its survival. Since innovation is particularly important for firms facing intense product market competition, social media that helps convey information on corporate innovation to outside investors may foster greater innovation activities within such firms.

As such, we hypothesize that SA coverage encourages corporate innovation by disseminating firm-specific, innovation-related information to external investors, thereby reducing the information asymmetry between inside managers and outside investors and relaxing the financial constraints on the firm's innovation activities. We also hypothesize that the beneficial impact of SA coverage on innovation should be more significant for firms facing greater information asymmetry and/or intense product market competition.

We also investigate whether the impact of reduced information asymmetry on corporate innovation outputs is influenced by managerial short-termism. When managers have heightened career concerns and exhibit greater myopia (i.e., are more short-term focused), the reduction in information asymmetry regarding corporate innovation activity may not sufficiently incentivize these managers to pursue long-term innovation projects. Thus, we further hypothesize that the effect of decreased information asymmetry on corporate innovation, through the initiation of SA coverage, should be weaker for firms facing higher levels of managerial short-termism.

## 3. Data and Methodology

### 3.1. Sample of Social Media Coverage

Our social media data consists of articles published on the SA online platform. Each SA article contains the title, content, stock name and ticker, publication date, author name and number of followers, and comments on the article. Table 1 presents the distribution of SA coverage from 2006 to 2017. The table reports the number of firms that are newly covered by SA and the number of SA articles in each year.

[Please insert Table 1 here]

In Table 1, the second and third column show that both number of articles and number of newly covered firms experience a jump in 2011. This jump may be attributed to two policies that SA implemented in 2011: a broader partnership with main-stream media companies and stock exchanges and a premium partnership program which rewards authors $10 for every 1,000 views per article.[9] These two policies attract more users and encourage a growing number of authors to post articles with high credibility, which has led to the rising popularity of SA since 2011.

### 3.2. Measuring Corporate Innovation

We obtain the updated patent data from Kogan, Papanikolaou, Seru and Stoffman (2017).[10] Following the innovation literature (e.g., He and Tian, 2013; Chemmanur et al., 2014; Balsmeier,

---

[9] See https://seekingalpha.com/instablog/229565-boaz-berkowitz/129214-bringing-you-more-exposure-msn-is-now-part-of-seeking-alphas-distribution-network and https://seekingalpha.com/article/2134803-how-much-does-seeking-alpha-pay-its-contributors.

[10] The data is available at https://host.kelley.iu.edu/nstoffma/. We thank Prof. Noah Stoffman and his coauthors for generously sharing this data with us.

Fleming and Manso, 2017; Bradley, Kim, and Tian, 2017; Guo et al., 2019), we mainly focus on patent number and citation number to measure the quantity and quality of a firm's innovation outputs, respectively. $PatentNum$ is the number of patents *applied* for (and eventually granted) in a firm-year. $CitedNum$ is the number of forward citations received by the patents that firms applied for (and eventually granted) in a year. We focus on patent applications because compared to patent grants, patent applications are more likely to reflect a firm's innovation activities in a given year.

Forward citations are subject to truncation bias because citations accrue over a long period of time (Hall, Jaffe and Trajtenberg, 2005; Lerner et al., 2011; Acharya and Xu, 2017). For example, since patents applied for in 2010 have longer time to receive forward citations than patents applied for in 2018, there is no surprise that the former patents receive more citations than the latter ones. To correct this truncation bias, when calculating $CitedNum$, we follow Hall, Jaffe and Trajtenberg (2001), Atanassov (2013) and Guo et al. (2019) and use the "time-technology class fixed effect" method, which divides the number of forward citations of a patent granted in a year by the average number of forward citations of the patents granted in the same technology class in the same year.

Besides patent number and citation number, in robustness analysis, we also use several alternative measures for corporate innovation activities. $Efficiency$ is the number of patents a firm has applied for (and eventually granted) in a given year per million R&D dollars, which reflects the firm's innovation efficiency in using R&D expenses to produce patents (Hirshleifer, Hsu and Li, 2013). $PatentValue$ is the total economic values of patents a firm has applied for (and eventually granted) in a given year scaled by the book value of total assets of that firm-year. The economic value of a patent is estimated based on the three-day stock returns following the patent's grant date (Kogan et al., 2017). Stock market rewards patents that are granted and highly cited (Hall et al., 2005; Kogan et al., 2017). Thus, patent value reflects the economic importance of patents and the private values of successful patents to shareholders. $EfficiencyValue$ is the total economic values of patents a firm has applied for (and eventually granted) in a given year per million R&D dollars. It captures the market rewards of R&D investments.

$Originality$ measures the diversity of technology classes a patent cites, and it is computed as one minus the Herfindahl index of backward citations for patent $p$ as $Originality_p = 1 - \sum_j F_{p,j}^2$ where $F_{p,j}$ is the portion of backward citations that belong to the technology class $j$, and

$Originality_p \in [0,1]$. Patents citing a broader array of technology classes have greater value of originality. $Generality$ measures the diversity of technology classes citing a patent, and it is calculated as one minus the Herfindahl index of forward citations for patent $p$ as $Generality_p = 1 - \sum_j G_{p,j}^2$ where $G_{p,j}$ is the portion of forward citations that belong to the technology class $j$, and $Generality_p \in [0,1]$. Patents receiving citations from a broader array of technology classes have a greater value of generality. A firm's originality and generality in a year are the average originality and generality values of the patents applied for (and eventually granted) by the firm in the year. Patent originality and generality capture patents' novelty based on the distribution of citations. Patents differ in their novelty because fundamental research tends to be risky and produce more influential innovation (Lerner et al., 2011; Bernstein, 2015; Acharya and Xu, 2017).

$R\&D$ is R&D expenses scaled by book value of total assets in a firm-year and $SG\&A$ is selling, general and administration expenses scaled by the book value of total assets in a firm-year.[11] These two measures capture firms' innovation inputs.

**3.3. Research Design and Sample Overview**

Data on firm characteristics are sourced from Compustat. Stock returns and liquidity data are sourced from the Center for Research in Security Prices (CRSP). Analyst data are from the Institutional Brokers' Estimate System (IBES) database, and news coverage data are from RavenPack. We require our sample firms to have at least one patent applied for (and eventually granted) during the 2008-2018 sample period. Following He and Tian (2013) and Fang et al. (2014), we exclude the firms in financial and utilities industries with Standard Industrial Classification (SIC) codes either between 4900 and 4999 or between 6000 and 6999.

Innovation activities of a firm may lead to more articles being written about the firm. This raises the concern of possible reverse causality between SA coverage and corporate innovation. Specifically, more innovative firms may attract SA coverage rather than SA coverage encouraging future corporate innovation activities. We address this endogeneity concern by focusing on the initiation of SA coverage of a firm that is *unrelated* to the firm's innovation activities and business relationship as an exogenous event. Further, we employ the DID approach to examine the treatment effect of SA coverage initiation on corporate innovation. In our DID setting, firms that

[11] SG&A expenses include R&D expenses (Compustat Online Manual). Many public firms do not separately report R&D but report their R&D expenses in SG&A.

experience SA coverage initiation are identified as treatment firms, and firms not covered by SA are identified as control firms. The coverage-initiation event is firm-specific, which happens in the publication year of a firm's first SA article. Hence, our sample contains multiple initiation events that occur at different times. We focus on the SA-coverage-initiation events from 2011 to 2015 because of the significant rise in recognition and popularity of SA among investors from 2011.[12]

To ensure that the initiations of SA coverage in our sample are not driven by either corporate innovation or business relationship, we filter the SA coverage initiations by the following steps. First, we initially classify coverage initiation articles as being innovation related if the content of an article includes more than two occurrences of a list of innovation-related keywords.[13] We then exclude all treatment firms with their SA articles' content being related to innovation in their coverage initiation years. Second, we further filter the remaining SA coverage initiations by retaining only the initiation articles whose authors declare no business relationships with the firms they write about. This step ascertains that the SA articles are not contributed by authors who have interest ties with the firm (including the employees of the firm). The first two steps yield 365 eligible initiation events of SA coverage. Third, given that the keywords we use in the first step may not be able to fully rule out all innovation-related articles, the content of *each* of the SA articles in the coverage-initiation years of the remaining 365 initiation events are manually and interpedently cross-checked by the authors and a research assistant to ensure that they are not related to corporate innovation.

After removing any potentially innovation-related articles that are not captured by the keywords used in the first step,[14] the final sample contains 175 eligible SA-coverage-initiation events from 2011 to 2015 that are *not* driven by either innovation-related corporate activities or authors having interest ties with the focal firm. Each of the 175 SA-coverage-initiation events has one unique treatment firm.

[12] We end the SA coverage initiation events in 2015. This is because we need three years of innovation data after an initiation event. We also want to minimize truncation bias, which affects data on patent applications due to the fact that it typically takes several years to process a patent application.

[13] The list of keywords are "research and development", "research & development", "R&D", "R & D", "patent", "cite", "citing", "cited", "citation", "licens", and "licenc".

[14] For example, some of the articles mention the new products of the covered firms, which could potentially be viewed as outcomes of successful corporate innovation. To ensure that the articles in our final treatment set is completely unrelated to corporate patenting and innovation activities, we have removed such articles.

Table A2 in the Appendix provides the coverage reason breakdown of the 175 non-innovation-related SA coverage initiation articles included in our final Difference-in-Differences regression analyses. Among the 175 articles, 157 are related to business fundamentals of the covered firms (e.g., industry conditions, business operations, profitability, etc.), 99 are related to stock prices and stock returns of the covered firms, 22 are related to the recent earnings announcements of the covered firms, 16 are related to M&A and corporate restructuring activities of the covered firms, 8 are related to corporate financing of the covered firms, 8 are related to corporate payouts of the covered firms, and 10 are related to corporate governance issues of the covered firms. Because some articles may be related to multiple of these reasons (e.g., an article can write about business fundamentals, stock prices and stock returns of the covered firm), the sum of article numbers across these reason categories exceeds 175.

There is a caveat regarding our identification strategy that is worth mentioning. While we strive to ensure that the content of SA coverage initiation articles is unrelated to corporate patenting and innovation activities, our identification strategy is not perfect. For example, when an SA article discusses the covered firm's stock returns, we cannot rule out the possibility that these returns somewhat reflect the recent innovation success of the firm. Similarly, when an SA article writes about M&A and restructuring activities of the covered firm, we cannot rule out the possibility that such activities may be motivated by corporate innovation. Nevertheless, we demonstrate below that in our propensity-score-matched sample, ex-ante patent quantity and quality, as well as other firm characteristics, fail to predict SA coverage initiation. Additionally, the increases in innovation quantity and quality for treatment firms relative to matched control firms occur only after the initiation of SA coverage.

We next apply the PSM approach to match up to five control firms for each treatment firm with replacement. Specifically, for each treatment firm, we require the matched control firms to: (1) be in the same two-digit SIC industry in year -1 (with the coverage-initiation event year being year 0); (2) *not* be covered by SA during the entire seven-year initiation event window; and (3) have a propensity score difference (between the control firm and the treatment firm) being smaller than 0.05. The propensity score is estimated using the predictors in Chemmanur et al. (2014) and Brav et al. (2018), including firm size (*Firm Size*), leverage (*Leverage*), Tobin's Q (*TobinQ*), return-on-asset (*ROA*), R&D expenses (*R&D*), asset tangibility (*Tangibility*), institutional ownership (*InstOwn (%)*), firm age (*Firm Age*), natural logarithm of one plus the number of patents a firm has applied for and later granted (*Log (1+PatentNum)*), and natural logarithm of one plus the number of forward citations received by patents that a firm has applied for and later granted (*Log (1+CitedNum)*), all measured in year -1. A description of each variable is presented in Table A1 in the Appendix. Each treatment firm is matched to its control firms to form an event cohort. In each cohort, we include firm-year observations from three years before to three years after the event year 0. We then stack all cohorts to form our DID regression sample. The choice of a seven-year window [-3, 3] reflects the trade-off between relevance and accuracy (He and Tian, 2013). Firms are not required to have available data for the full seven years around the coverage initiation. Since event year 0 spans from 2011 to 2015, event window [-3, 3] covers 2008 to 2018. The final DID regression sample comprises 175 cohorts and 5,362 cohort-firm-year observations.

Panel A of Table 2 reports the ex-ante firm characteristics of treatment and control firms in the year prior to the SA-coverage-initiation event (year -1). The second and third column are the mean values of firm characteristics for treatment and control firms, respectively. The last column is the *p*-values from *t*-tests of differences in mean values between the two groups of firms. Since *p*-values are all higher than 10%, we are unable to reject the null hypothesis that the control firms are ex-ante very similar to the treatment firms across all of the firm characteristics dimensions used to estimate propensity scores. In particular, the treatment and control firms have very similar innovation outputs in the year prior to the SA coverage initiation.

In Panel B of Table 2, we formally test whether a firm's pre-existing innovation performance can predict the SA coverage initiation or not in our matched DID regression sample following the approach in Beck, Levine, and Levkov (2010). Specifically, we run OLS regressions to predict one-year-ahead SA coverage initiation using ex-ante patent count and citation count and other firm-level control variables. We include the firm-year observations of the matched sample in these predictive regressions until the treatment event (i.e., SA coverage initiation) of a cohort occurred. We then drop all firm-year observations of the cohort after the SA coverage initiation year. We find that ex-ante innovation performance and firm-level controls cannot predict SA coverage initiation. Thus, the treatment assignment (i.e., SA coverage initiation) can be viewed as random.

[Please insert Table 2 here]

Table 3 provides summary statistics on the variables for the DID regression sample. All continuous variables are winsorized at the 1% and 99% levels. For each variable, we present mean

(Mean), 25 percentile (P25), median (Median), 75 percentile (P75), standard deviation (SD), skewness (Skewness), kurtosis (Kurtosis) and observation number (N). A description of each variable is presented in Table A1 in the Appendix. Table A3 in the Appendix presents the pairwise correlation matrix of the variables.

[Please insert Table 3 here]

The average $PatentNum$ per firm-year in our DID sample is 34.776. The average $CitedNum$ per firm-year is 34.791. Patent number and citation number are both positively skewed with skewness at 4.660 and 4.441, respectively. The right skewed distribution of patents and citations is consistent with He and Tian (2013). The new equity (debt) raising as a ratio of lagged total assets per firm-year is, on average, 0.033 (0.020). We use the equity-focused financial constraint measures (i.e., $LW$ $(Full)$ and $LW$ $(Primitive)$) proposed by Linn and Weagley (2023) to examine whether the financial constraints of a treatment firm become relaxed upon the treatment.[15] We further augment their measures with the Hadlock and Pierce's (2010) size-age financial constraint index ($HP$). The average $LW(Full)$ index and $LW(Primitive)$ index are both -0.249, and the average $HP$ index is -3.706 per firm-year in the sample.

On average, firms in our sample have the natural logarithm of book value of assets 6.545, leverage ratio of 0.182, Tobin's Q of 0.534, ROA of -0.002, R&D scaled by book value of assets of 0.090, PPE scaled by book value of assets of 0.150, institutional share ownership of 54.092%. In addition, analyst following (i.e., the natural logarithm of one plus the number of analysts following a firm-year) is 2.085. The sample firms relate to an average of 185.7 news articles per firm-year based on the RavenPack database, among which the number of positive news is 16.853% more than that of negative news. The average stock returns and return volatilities per firm-year are 13.487% and 12.726%, respectively.

[15] The financial constraint measures of Linn and Weagley (2023) are machine-learning extensions of the text-based financial constraint measures developed in Hoberg and Maksimovic (2015), which identify direct statements indicating financial constraints. Linn and Weagley (2023) adopt a machine learning algorithm, random forests, to estimate a multi-dimensional mapping between firm-level accounting variables and financial constraints. The authors then estimate financial constraints both with a large set of accounting variables as predictors and, alternatively, with a small set of primitive, less endogenous variables (the "Exogenous" model), producing two different versions of financial constraint measures for most U.S. publicly-traded firms back to 1972. We use their equity-focused constraint measures estimated with both the full model and the "Exogenous" model. The data is available at https://www.danielweagley.com/research.html. We thank the authors for sharing the data with us.

## 4. Treatment Effects of SA Coverage Initiation on Corporate Innovation Activities

### 4.1. Baseline Analyses

We estimate the average treatment effect of SA coverage initiation across all coverage-initiation event cohorts using the following DID regression framework:

$$Innovation_{c,i,t} = \beta_1 Treat_{c,i} \times Post_{c,t} + Firm\ Controls_{c,i,t-1} + f_{c,i} + \lambda_{c,t} + \epsilon \qquad (1)$$

where $Innovation_{c,i,t}$ represents an innovation measure of firm $i$ in cohort $c$ at year $t$. The innovation measures include $PatentNum$ and $CitedNum$. $Treat_{c,i}$ is an indicator variable that equals one if firm $i$ in cohort $c$ is covered by SA (a treatment firm) and equals zero otherwise. $Post_{c,t}$ is an indicator variable that equals one if year $t$ is in the post-event period (i.e., year [1, 3]) and equals zero if it is in the pre-event period (i.e., year [-3, -1]). The regression coefficient of the DID term, $Treat_{c,i} \times Post_{c,t}$, captures the average treatment effects of SA coverage initiation on treatment firms' innovation activities across cohorts. We include both cohort-firm fixed effects ($f_{c,i}$) and cohort-year fixed effects ($\lambda_{c,t}$) to account for firm and year unobserved heterogeneities within each cohort. The $Treat$ indicator is absorbed by the cohort-firm fixed effects and the $Post$ indicator is absorbed by the cohort-year fixed effects. We use three years before and after the coverage-initiation event year to estimate equation (1) and exclude the coverage-initiation event year 0 from the estimation. We include lagged firm-level controls in some specifications, but leave such controls out in other specifications to alleviate concerns related to the 'endogenous control' problem (e.g., Angrist and Pischke, 2009; Gormley and Matsa, 2016). Standard errors are clustered at the cohort-firm level to account for potential within-firm autocorrelation.

Cohn et al. (2022) show that the common practice of using the log of one plus the patent or citation count as the dependent variable in OLS regressions can lead to biased estimates. They further show that using a fixed-effects Poisson model with count dependent variables can produce consistent and reasonably efficient estimates under more general conditions than commonly assumed. Similar conclusion is also reached by Chen and Roth (2023). Consequently, we employ fixed-effects Poisson regressions for our DID analyses.[16] Table 4 presents the baseline DID results.

[16] Our results remain qualitatively unchanged when we use OLS regressions and innovation count dependent variables (or inverse hyperbolic sine transformation or log one plus count transformation of the dependent variables).

We find that the coefficient estimates on $Treat \times Post$ are positive and statistically significant at the 1% level across all regression specifications.[17] In column 2, the coefficient estimate on $Treat \times Post$ is 0.633 ($t$-stat = 5.694) when the dependent variable is $PatentNum$. It indicates that patent number on average increases by 0.633 log points or 88.3% (i.e., exp(0.633)-1) for treatment firms relative to control firms after SA coverage initiation. Moreover, in column 4 for the dependent variable $CitedNum$, the DID coefficient estimate is 0.832 ($t$-stat = 6.705). Thus, forward citations per patent on average increase by 129.8% for treatment firms relative to control firms after SA coverage initiation. Overall, the results in Table 4 show that, relative to control firms, treatment firms exhibit significant increases in patent number and forward citations after SA coverage initiation.

[Please insert Table 4 here]

We further examine whether the positive treatment effects of SA coverage on corporate innovation outputs are driven by the potential nonparallel innovation trends prior to coverage initiation. We include the coverage-initiation event year 0 and employ the following dynamic DID regression framework to identify the exact timing of the treatment effects as follows:

$$Innovation_{c,i,t} = \beta_2 \sum_{n=-2}^{3} (Treat_{c,i} \times Yr(\Delta t = n)_{c,t}) + Firm\ Controls_{c,i,t-1} + f_{c,i} + \lambda_{c,t} + \epsilon \quad (2)$$

where $Yr(\Delta t = n)_{c,t}$ represents the year indicator for the $n^{th}$ year relative to the coverage initiation year in cohort $c$, where $\Delta t$ is the year difference between calendar year $t$ and the coverage initiation year in cohort $c$. $Yr(\Delta t = n)_{c,t}$ equals one if $\Delta t = n$ (where $n = \{-2, -1, 0, 1, 2, 3\}$) and equals zero otherwise. Table 5 reports the dynamic DID regression results from estimating equation (2).

The pre-trend coefficient estimates on $Treat \times Yr(\Delta t = -2)$ and $Treat \times Yr(\Delta t = -1)$ are statistically insignificant across all five models. The coefficient estimate of $Treat \times Yr(\Delta t = 0)$ is also insignificant. The coefficient estimates on $Treat \times Yr(\Delta t = 1)$ are significantly positive across the four models, suggesting that the effect of SA coverage on corporate innovation outcomes only starts to show up in the year after the coverage initiation year. Moreover, the

[17] It is worth noting that, due to the estimation process of Poisson regression, the observations in the regression sample differ slightly between columns 1 and 3 (as well as between columns 2 and 4). A firm with zero forward citation counts in all years is automatically dropped from the estimation process. Additionally, the control variable *FirmAge* is absorbed by cohort-firm fixed effects and cohort-year fixed effects and is therefore omitted from the table. The same arguments apply to the subsequent tables.

coefficient estimates on $Treat \times Yr(\Delta t = 2)$ and $Treat \times Yr(\Delta t = 3)$ are also significantly positive and large in magnitudes. These dynamic DID regression results show that the parallel trends assumption in the DID approach (that no significant difference in innovation patterns between the treatment and control firms before SA coverage initiation) is satisfied.

[Please insert Table 5 here]

To further visualize the parallel trends and treatment effects around SA coverage initiation, Figure 1 plots the coefficient estimates from columns 2 and 4 of Table 5 with 95% confidence intervals. The plotted coefficients reflect the change in innovation outputs for treatment firms relative to control firms. Figures 1A and 1B correspond to $PatentNum$ and $CitedNum$, respectively. In both figures, comparing to the post-event period (X-axis $> 0$), point estimates in the pre-event period (X-axis $\leq 0$) are close to zero. Thus, there is no difference in innovation activities between the treatment and control firms prior to SA coverage initiation. Treatment firms increase their innovation outputs relative to control firms only after SA coverage initiation. This increase begins in the year after the SA coverage-initiation year and continues for the next two years. The precise timing of the innovation increase shown in Figure 1 again suggests that the parallel trends assumption underlying our DID analyses is satisfied and the documented treatment effects of SA coverage initiation on corporate innovation activities are most likely causal. The finding suggests that the treatment effects of SA coverage initiation on corporate innovation activities tend to persist for at least several years after SA coverage initiation.

[Please insert Figure 1 here]

In summary, the results from our DID analyses suggest that SA coverage initiation significantly encourages corporate innovation outputs in terms of patent quantity and quality.

**4.2. The Effects of SA Coverage Initiation on Financing Activities and Financial Constraints**

We posit that SA promotes corporate innovation activities by providing innovation-related information of the covered firm to external investors and thus relaxing the firm's financial constraints. Hence, we next explore the effects of SA coverage initiation on the covered firm's financing activities and financial constraints.

Existing research suggests that equity financing may be preferable to debt financing for funding such activities (Brown et al., 2013; Acharya and Xu, 2017). This is because R&D investments are often intangible, have a high probability of failure, and offer little or no collateral value, making it difficult for firms to secure debt financing. However, access to stock market financing can be particularly beneficial for R&D investments, as R&D often limits firms' ability to use debt finance.[18] Following McLean and Zhao (2014), we measure new equity issuance ($EquityRaising$) as the change in book equity, plus the change in deferred taxes, minus the change in retained earnings, all scaled by lagged assets; and new debt issuance ($DebtRaising$) as the change in assets, minus the change in book equity, minus the change in deferred taxes, all scaled by lagged assets. We modify equation (1) by using $EquityRaising$ and $DebtRaising$ as the dependent variables and estimate the equation using OLS regressions. The results are reported in Panel A of Table 6.

We find that the coefficient estimate on $Treat \times Post$ is positive and significant at the 5% level in column 1 with the dependent variable being $EquityRaising$, but is negative and insignificant in column 3 with the dependent variable being $DebtRaising$. The results indicate that equity raising increases by 2.4% of total assets per firm-year for the treated firms relative to the matched control firms after SA coverage initiation. The findings remain qualitatively unchanged when we include lagged firm controls in columns 2 and 4. Overall, we find that treatment firms exhibit a significant increase in raising new equity relative to control firms after SA coverage initiation, which is consistent with the literature. These results indicate that equity issuance rather than debt issuance is a suitable source for innovation investment.

We further examine whether the financial constraints of a treatment firm become relaxed upon the treatment. As discussed earlier, we use the equity-focused financial constraint measures (i.e., $LW$ $(Full)$ and $LW$ $(Primitive)$) proposed by Linn and Weagley (2023) and Hadlock and Pierce's (2010) size-age financial constraint index ($HP$). As shown in Panel B of Table 6, the coefficient estimates on $Treat \times Post$ are negative and statistically significant across all regression specifications. Thus, the findings suggest that the financial constraints of the treatment firms indeed become relaxed relative to the matched control firms after the treatment of SA coverage initiation.

[Please insert Table 6 here]

[18] According to Acharya and Xu (2017), public firms in external financial dependent industries tend to invest more in R&D and generate a stronger patent portfolio compared to their private counterparts. This is because being listed on the stock market can lower the cost of capital, as investors' portfolios become more liquid and diversified.

Taken together, the results in this section support the notion that SA coverage initiation promotes corporate innovation activities by relaxing covered firms' financial constraints and allowing them to raise equity financing for their innovation investment.

### 4.3. Heterogeneity of the Treatment Effects

We next investigate the heterogeneity of the SA initiation treatment effects on corporate innovation activities. As discussed earlier in the hypothesis development (Section 2), we conjecture that the positive treatment effects of SA coverage initiation on corporate innovation activities should be more pronounced for firms with higher information asymmetry, firms facing greater product market competition, and firms having lower levels of managerial short-termism.

We follow Gormley and Matsa (2016) to test for the heterogeneity in the treatment effects of SA coverage initiation by partitioning our full sample into subsamples according to information asymmetry, product market competition and managerial short-termism, and estimating the baseline regression model separately for each subsample. Specifically, we first sort the full sample based on the median value of each proxy of information asymmetry, product market competition or managerial short-termism measured in year -1 (i.e., the year immediately before the initiation-event year), and split the full sample into the above-median (H) and below-median (L) subsamples. Thus, the partitioning variables are all constant based on pre-initiation value. By using pre-event measurements of the partitioning variables, we can reduce concerns related to endogeneity. If partitioning variables are time-varying, there is a risk that they are influenced by the treatment itself, leading to potentially biased results. We then re-estimate equation (1) for each subsample and further report the $p$-values from Chow tests for the differences in the DID coefficient estimates between the H and L subsamples.

#### 4.3.1. Information Asymmetry

We employ four proxies of information asymmetry to partition the sample: analyst following, abnormal accrual, idiosyncratic volatility, and effective spread. The first proxy is analyst following ($AnalystFollowing$), which is the natural logarithm of one plus the number of analysts following a firm in a year. Firms with greater analyst following have lower information asymmetry (e.g., Frankel and Li, 2004). The second proxy is abnormal accruals ($AbnormalAccrual$), which is defined as the median absolute value of discretionary accruals in the past five fiscal years (e.g., Francis, LaFond, Olsson and Schipper, 2005; Bhattacharya, Desai and Venkataraman, 2013; Guo and Qiu, 2016). The discretionary accruals are estimated based on the modified Jones model (Dechow, Sloan and Sweeney, 1995). Lee and Masulis (2009) show that firm financial disclosures are more opaque (i.e., higher information asymmetry) when abnormal accruals are larger.

The third proxy, effective spread ($IdiosyncraticVolatility$), is the standard deviation of the residuals from regressing daily individual stock returns on the Fama-French three-factors of a firm in a year. Greater idiosyncratic volatility is associated with higher information asymmetry (e.g., Jiang, Xu and Yao, 2009; Rajgopal and Venkatachalam, 2011). The fourth proxy, effective spread ($EffectiveSpread$), is the bid-ask spreads estimated using daily high and low stock prices of a firm following Corwin and Schultz (2012). Firms with greater bid-ask spreads have higher information asymmetry (e.g., Corwin, 2003; Karpoff, Lee and Masulis, 2013).

Table 7 reports the test results for treatment effect heterogeneity across the four proxies of information asymmetry. We find that the treatment effect of SA coverage initiation on promoting corporate innovation outputs is indeed more pronounced for firms with low analyst following, high abnormal accruals, high idiosyncratic volatility, and/or high effective spread. The *p*-values of the DID-term coefficient differences between high and low information-asymmetry subsamples suggest that the treatment effects are significantly different across subsamples. Firms with high information asymmetry are associated with a significantly greater increase in patent quantity and quality than firms with low information asymmetry after SA coverage initiation. The findings support the conjecture that SA coverage initiation incentivizes future corporate innovations by mitigating the information asymmetry between insider managers and outside investors. The communication via SA alleviates informational frictions by analyzing, explaining and discussing the meaning and implications of firms' complex information on innovation. If investors have a better understanding of the innovation projects in a firm, then they will provide greater support for corporate innovation activities and a higher tolerance for innovation project failure. Such firms are expected to have a greater incentive to engage in risky innovation projects.

[Please insert Table 7 here]

### 4.3.2. Market Competition

We next employ four proxy measures of product market competition to partition the sample: product market fluidity, TNIC HHI, HHI and Lerner index. The first proxy, product market fluidity ($Fluidity$), is constructed based on the business description section of firms' 10-K filings. $Fluidity$ captures changes in the word usage of rival firms in describing products that overlap with a firm's own product vocabulary. In essence, $Fluidity$ measures changes in the products of rival firms relative to the products of the focal firm. Hoberg, Phillips, and Prabhala (2014) have demonstrated that firms with a higher level of $Fluidity$ face greater product market competition.

The second proxy, TNIC HHI ($TNICHHI$), is the sales-based Herfindahl-Hirschman index of the firm's industry where industry is defined by text-based network industry classifications (TNIC). Firms are assigned to industries based on a clustering algorithm that maximizes firm pairwise product similarity within these industries (Hoberg and Phillips, 2010, 2016). Firms with a lower level of TNIC HHI face greater product market competition. We obtained data on product market fluidity and TNIC HHI from the Hoberg-Phillips Data Library.[19] The third proxy, HHI ($HHI$), is the sales-based Herfindahl-Hirschman index of the firm's industry based on 2-digit SIC industry classification. Similar to TNIC HHI, firms with a lower level of HHI face greater product market competition. The fourth proxy, Lerner index ($LernerIndex$), is defined as the median gross margin in the firm's two-digit SIC industry following Aghion et al. (2013). $Lerner$ being equal to one indicates perfect competition in an industry where sales price equals marginal cost at the median level. Lower values of $LernerIndex$ indicate lower market power and greater competition.

Table 8 reports the test results for treatment effect heterogeneity across the four proxies of product market competition. The results reveal that the treatment effect of SA coverage initiation on promoting corporate innovation outputs is more pronounced for firms with high product market fluidity, low TNIC HHI, low HHI and/or low Lerner index. The $p$-values of the DID-term coefficient differences between high and low product-market-competition subsamples suggest that the treatment effects are significantly different across subsamples. The results suggest that after SA coverage initiation, firms with higher product market fluidity, lower product market concentration and lower market power, yield higher levels of corporate innovation outputs than firms facing lower levels of product market competition. Because innovation is crucial for firms facing intense product market competition not only for their growth but also for survival, the initiation of SA coverage that helps convey analyzed information on corporate innovation to outside investors should be especially important in fostering innovation activities for such firms. The findings based on different proxies of product market competition hence support the

[19] See https://hobergphillips.tuck.dartmouth.edu/industryconcen.htm. We thank Professor Gerard Hoberg and Professor Gordon Phillips for generously sharing the data with us.

conjecture that the treatment effect of SA coverage initiation on corporate innovation is stronger in firms facing high levels of product market competition.

[Please insert Table 8 here]

#### 4.3.3. Managerial Short-termism

We further employ four proxy measures of managerial short-termism to partition the sample: transient institutional ownership ($TransientInstOwn$), dedicated institutional ownership ($DedicatedInstOwn$), CEO age ($CEOAge$) and CEO tenure ($CEOTenure$).

The literature suggests that short-term-focused institutional investors (i.e., those with short investment horizons) can exert significant pressures on firm managers to deliver expected earnings quickly. Following Bushee (1998, 2001), we categorize institutional investors into three types: transient, quasi-indexing, and dedicated institutions. Transient institutions are active investors with a short-term investment horizon and high portfolio turnover. Quasi-indexing institutions are passive investors with low turnover and high diversification. In contrast, dedicated institutions are long-term relational investors with low portfolio turnover. We use transient institutional ownership and dedicated institutional ownership as proxies for the short-term pressures on firm managers. A firm with high transient institutional ownership or low dedicated institutional ownership is considered more susceptible to investor pressures to deliver bottom-line earnings in the short run, reflecting greater managerial short-termism.

Additionally, we employ CEO age and CEO tenure as proxies for managerial career concerns and short-termism. Younger CEOs, with longer careers ahead, typically have greater career concerns and are therefore more prone to short-termism (e.g., Gibbons and Murphy, 1992). Similarly, CEO tenure is negatively related to career concerns and short-termism, as CEOs in the early years of their service typically have more significant career concerns compared to those in later years of service (e.g., Ali and Zhang, 2015).

Table 9 reports the test results for treatment effect heterogeneity across the four proxies of managerial short-termism. The results reveal that the treatment effect of SA coverage initiation on promoting corporate innovation outputs is more pronounced for firms with low transient institutional ownership, high dedicated institutional ownership, high CEO age and/or high CEO tenure. The *p*-values of the DID-term coefficient differences between high and low short-termism subsamples suggest that the treatment effects are significantly different across subsamples. The results suggest that the impact of reduced information asymmetry on corporate innovation is indeed contingent on managerial short-termism. The findings support the conjecture that the effect of decreased information asymmetry, stemming from the initiation of SA coverage, on corporate innovation outputs is weaker for firms facing higher levels of managerial short-termism. When managers have heightened career concerns and are thus more short-term focused, the reduction in information asymmetry regarding corporate innovation activities may not be sufficient to incentivize them to pursue long-term innovation projects.

[Please insert Table 9 here]

## 5. Robustness Tests

We conduct various robustness tests on the uncovered positive treatment effect of SA coverage initiation on future corporate innovation. Specifically, we use alternative measures of innovation and include additional controls.

### 5.1. Alternative Measures of Innovation

We further test the robustness of the baseline DID results using the aforementioned alternative measures of innovation, including innovation efficiency ($Efficiency$), patent value ($PatentValue$), efficiency value ($EfficiencyValue$), originality ($Originality$), generality ($Generality$), R&D ($R\&D$) and SG&A ($SG\&A$).

We re-estimate equation (1) using these alternative innovation measures ($Efficiency_t$, $PatentValue_t$, $EfficiencyValue_t$, $Originality_t$, $Generality_t$, $R\&D_t$, and $SG\&A_t$) as the dependent variables in OLS regressions. The results are reported in Table A4 in the Appendix. The coefficient estimates on $Treat \times Post$ are positive and significant at the 1% level across these alternative innovation measures. These results are consistent with the baseline results in Table 4.

### 5.2. Controlling for News Media Coverage, Analyst Coverage, Stock Returns and Return Volatility

Firm-specific analyses on SA are provided by buy-side analysts, many of whom experienced investors and industry experts. The literature suggests that news media coverage and sell-side financial analyst coverage may impede corporate innovation due to the increased short-term pressure on firm managers (i.e., He and Tian, 2013; Dai et al., 2021). In contrast, Guo et al. (2019) argue that sell-side financial analyst coverage increases patents, patent citations and patent novelty through encouraging firms to make more efficient innovation investments. Thus, we conduct robustness tests on the impact of SA coverage initiation on corporation innovation outcomes by controlling for news media coverage and sell-side financial analyst coverage. We augment equation (1) by adding $\#News$ and $NewsTone$ and/or $AnalystFollowing$ measured in year $t-1$. $\#News$ is the number of news articles based on the RavenPack database divided by 100. $NewsTone$ is the difference between the portion of positive and negative news articles. $AnalystFollowing$ is the natural logarithm of one plus the number of sell-side financial analysts following a firm in a year. Moreover, since stock returns and return volatility may correlate with both innovation outputs (e.g., Kogan et al., 2017) and SA coverage, we further include past stock returns ($StockReturn$) and return volatility ($StockVolatility$) measured in year $t-1$ as additional control variables in the DID regressions. The DID regression results are reported in Table A5 in the Appendix.

We find that the coefficient estimates on $AnalystFollowing$ and $\#News$ are significantly positive and those on $NewsTone$, $StockReturn$ and $StockVolatility$ are insignificant. The difference with He and Tian (2013) and Dai et al. (2021) is likely due to different research settings and different sample constructions. Importantly, with these additional controls, the coefficient estimates on $Treat \times Post$ remain positive and statistically significant at the 1% level for both patent quantity and quality.

The effects of SA coverage on future corporate innovation outcomes are distinct from those of news media and sell-side analyst coverage likely because, unlike news reporters and sell-side equity analysts, SA authors are buy-side analysts who usually do not impose short-term pressure on corporate managers (e.g., they do not issue earnings forecasts), thereby incentivizing firms to pursue long-term innovation projects. Moreover, compared with news reporters and sell-side analysts, SA authors are also more likely to be industry experts and thus play a more important role in reducing the information asymmetry on complex corporate innovation projects.

To summarize, the positive treatment effect of SA coverage initiation on corporate innovation continues to hold after controlling for news media coverage, analyst coverage, stock returns and return volatility.

## 6. SA Article Intensity, Comment Intensity, and Corporate Innovation Activities

In this section, instead of using SA-coverage-initiation events and a stacked-cohort DID regression setting, we construct a panel dataset on SA article intensity and article comment intensity to examine the effect of SA coverage on corporate innovation. We expect that higher intensity of SA articles, especially innovation-focused ones, on a firm should have a larger positive effect on future corporate innovation outcomes. Similarly, we expect that greater article impact, proxied by higher average number of comments received per article written on a firm in a year, should reveal a stronger positive effect on the firm's future innovation outcomes. As the relation between article intensity or comment intensity and corporate innovation may be confounded by endogeneity issues (e.g., the expectation of better corporate innovation outcomes can attract more SA articles and more comments on the articles), we use 2SLS instrumental variable regressions to extract the exogenous component of SA article intensity or article comment intensity, and relate the extracted component of article intensity or comment intensity to future corporate innovation.

Our proxy for the intensity of SA articles is $ArticleNum_{i,t}$, which is the total number of SA articles related to firm $i$ in year $t$. Our proxy for the intensity of SA article comments is $AvgComtNum_{i,t}$, which is defined as the average number of comments (excluding those comments made by authors themselves) per article related to firm $i$ in year $t$. We use an instrumental variable $Log(1 + FollowerNum)_{i,t}$, which is the natural logarithm of one plus the aggregate number of followers of all SA authors who publish articles on firm $i$ in year $t$. The aggregate number of followers is measured by adding up the past followers of these authors before they write on the firm. This instrument captures the past accumulated popularity of the SA authors who writes on firm $i$ in year $t$ among readers. More popular authors (i.e., those with higher numbers of followers) will likely produce more articles on the firms that they cover and generate greater average impact per article written (relevance condition). However, the past personal popularity of SA authors is unlikely to directly influence the focal firm's innovation outcomes other than through the articles they write on the firm and the average impact these articles generate (exclusion restriction).

We construct a panel dataset including firms covered by SA from 2011 to 2015. This sample consists of 5,142 firm-year observations. The 2SLS regression models include firm and year fixed effects, and firm control variables measured in year $t$.[20] Standard errors are clustered at the firm level.

[20] Control variables include firm size ($FirmSize$), firm age ($FirmAge$), leverage ($Leverage$), Tobin's Q ($TobinQ$), return on assets ($ROA$), R&D expenses ($R\&D$), asset tangibility ($Tangibility$), institutional ownership ($InstOwn$), stock returns ($StockReturn$), return volatility ($StockVolatility$), analyst following ($AnalystFollowing$), news coverage ($\#News$) and news tone ($NewsTone$).

Panel A of Table 10 reports the 2SLS regression results on the effects of SA article intensity. Since Poisson estimation cannot be used in 2SLS instrumental-variable regressions, we use the Inverse Hyperbolic Sine (IHS) transformation of the patent or citation count in year *t+1* as the second-stage dependent variable (Bellemare and Wichman, 2020) for the 2SLS regressions. Similarly, we also use the IHS transformation of $ArticleNum_{i,t}$ as the first-stage dependent variable.[21] Panel A reports the results on SA article intensity and future corporate innovation outcomes. The column under $IHS\ (ArticleNum)$ presents the first-stage regression results, where $IHS\ (ArticleNum)$ is regressed on the instrument $Log(1 + FollowerNum)$, control variables and fixed effects to obtain the fitted value of SA article intensity. The coefficient estimate on the instrument is positive and significant at the 1% level (*t*-stat = 20.236), which supports the relevance condition. The $p$-value of Cragg-Donald's Wald $F$ week-instrument test statistic is 0.000, which strongly rejects the null hypothesis that the instrument is weak (Cragg and Donald, 1993; Stock and Yogo, 2005). The rest of the columns present the second-stage regression results, where the innovation measures are regressed on the fitted value of $IHS\ (ArticleNum)$ and the same set of control variables as used in the first stage. We find that the coefficient estimates on the fitted component of $IHS\ (ArticleNum)$ are positive and statistically significant at the 5% level in both second-stage regressions. These results are consistent with those of the DID analyses based on SA-coverage-initiation events and indicate significantly positive effects of SA article intensity on corporate innovation outcomes.

We expect that SA coverage promotes future corporate innovation through the innovation-focused SA articles on firms. That is, higher intensity of innovation-focused SA articles on a firm should have a larger positive effect on the firm's future corporate innovation. Hence, we use $InnoArticleNum_{i,t}$ as a proxy of innovation-focused SA article intensity, which is the total number of SA innovation-focused articles on firm $i$ in year $t$. The innovation-related SA articles are selected based on the following keywords: "innovation", "patent", "R&D", "tech", "new products", "new drug" and "FDA".[22] The SA article is selected to be innovation focused if the

[21] It is worth noting that our results are insensitive to the IHS transformation of the dependent variables. We also use the raw count variables as the dependent variables in the 2SLS regressions and obtain qualitatively similar results, as shown in Tables A6 and A8 in the Appendix.

[22] The keywords of "new drug" and "FDA" are selected because innovation-related SA articles in pharmacy industry are highly likely to mention "FDA" and/or "new drug" rather than "patent" or "new product".

word count of any keyword exceeds the 99 percentile (exclusive) of the word-count distribution of the keyword in the sample.

We then repeat the 2SLS regressions in Panel A by replacing $IHS\ (ArticleNum)$ with $IHS\ (InnoArticleNum)$ and report the results in Panel B of Table 10. The column under $IHS\ (InnoArticleNum)$ presents the first-stage regression results. We find that the coefficient estimate on the instrument $Log(1 + FollowerNum)$ is positive and significant at the 1% level ($t$-stat = 7.065), which supports the relevance condition. The $p$-value of Cragg-Donald's Wald $F$ week-instrument test statistic is also 0.000, which rejects the null hypothesis that the instrument is weak. The rest of the columns present the second-stage regression results. We find that the coefficient estimates on the fitted component of $IHS\ (InnoArticleNum)$ are positive and statistically significant at the 5% level for future patent count and future citation count. Compared with the coefficient estimates of fitted $IHS\ (ArticleNum)$ in the second-stage regressions in Panel A, the coefficient estimates of fitted $IHS\ (InnoArticleNum)$ in Panel B are also much larger in magnitudes, indicating that firms with more innovation-focused SA articles are associated with even greater innovation outcomes.

We also examine the non-innovation-focused article intensity ($NonInnoArticleNum_{i,t}$). The results are reported in Panel C of Table 10. Although the coefficient estimates on fitted $IHS\ (NonInnoArticleNum)$ are positive and significant, they are much smaller in magnitudes relative to those of fitted $IHS\ (InnoArticleNum)$ in Panel B. These results complement the results in Panel B and suggest that the positive effect of SA coverage on corporate innovation is mainly through innovation-focused SA articles. Compared to non-innovation-focused articles, it is expected that innovation-focused SA articles should offer more detailed insights on corporate innovation and help reduce information asymmetry related to complex innovation. The stronger positive impact of higher intensity of innovation-focused articles on corporate innovation outcomes thus underscores the significant role that SA can play in disseminating valuable information and promoting innovation within firms.[23]

[Please insert Table 10 here]

[23] A potential concern is that the author-popularity-based instrument we use may not be clean enough, as a very well-followed SA author might be well-followed due to their extremely informative articles about corporate innovation. To address this concern, we further construct an instrument by only considering the total number of followers on the SA authors' past non-innovation-focused articles. The 2SLS regression results using this alternative instrument remain qualitatively similar and are shown in Table A7 in the Appendix.

Finally, we investigate the effects of article comment intensity on future corporate innovation outcomes. We conjecture that a greater average impact per article on a firm will have a stronger effect on the firm's future innovation. Table 11 reports the 2SLS regression results. Panel A of Table 11 reports the results on SA article comment intensity ($AvgComtNum$) and future corporate innovation outcomes. The first-stage regression results show that the coefficient estimate on the instrument $Log(1 + FollowerNum)$ is positive and significant at the 1% level (*t*-stat = 3.278), which supports the relevance condition and indicates that greater personal popularity of SA authors is related to higher average number of comments per article written on the firm. The second-stage results show that the coefficient estimates on fitted $IHS\,(AvgComtNum)$ are positive and significant in both second-stage regressions. In Panel B of Table 11, we alternatively measure SA article comment intensity, $MaxComtNum$, as the maximum number of comments (excluding those comments made by authors themselves) per article related to firm $i$ in year $t$. The results are qualitatively similar to those reported in Panel A. These findings indicate that a greater impact per article on a firm has a stronger positive effect on its future innovation outcomes.

[Please insert Table 11 here]

To summarize, the 2SLS regression results in this section suggest that both SA article intensity and article comment intensity encourage future innovation. The findings based on the panel dataset and 2SLS regressions hence corroborate the results from the DID analyses based on SA-coverage-initiation events and a stacked-cohort DID regression setting, indicating a positive effect of SA coverage on future corporate innovation activities.

## 7. Conclusion

In this paper, we study the effects of social media coverage on future corporate innovation using SA. SA provides third-party generated and analyzed firm-specific information, which mitigates the information asymmetry between firms and outside investors. We conjecture that after SA coverage initiation, innovation projects of firms covered by SA are more likely to be carefully evaluated and discussed among investors. The interactive communication process on SA bridges the information gap between inside managers and outside investors on complex innovation projects and permits existing and future investors to be more tolerant of covered firms' temporary profit downturns, thereby incentivizing these firms to engage in long-term innovation activities.

Using SA coverage initiation as a quasi-natural experiment, we document positive and potentially causal treatment effects of SA coverage initiation on corporate innovation activities. We further show that SA coverage stimulates corporate innovation activities by disseminating innovation-related information about the covered firm to external investors, thereby alleviating the firm's financial constraints. We further find that the positive treatment effect of SA coverage initiation on corporate innovation is more pronounced among firms with greater information asymmetry ex ante. Since information asymmetry significantly impedes corporate innovation (which is known to be like a "black box" to investors), our results emphasize the informational role of specialized social media such as SA in incentivizing corporate innovation. Moreover, we find the effect of SA coverage on innovation to be stronger for those firms facing greater product market competition ex ante. Intense product market competition can make innovation even more important for firms to survive the "creative destruction" process. Thus, our findings suggest that SA coverage initiation that helps convey innovation-related information to the market incentivizes more innovation activities when the firm's product market space is more competitive.

We also find that the impact of decreased information asymmetry, stemming from the initiation of SA coverage, on corporate innovation activities is less pronounced for firms characterized by heightened managerial short-termism. These findings imply that when managers harbor substantial career concerns, prompting a more short-term focus, the reduction in information asymmetry regarding corporate innovation may not be adequate to motivate them to undertake long-term innovation projects.

We conduct a battery of tests and confirm the robustness of our main findings. Importantly, corroborating the evidence from the stacked-cohort DID analyses based on SA coverage initiation, the results from our 2SLS panel regressions show that SA article intensity and average impact per article positively drive future corporate innovation outcomes. Our evidence emphasizes the value added by specialized social media in driving future corporate innovation activities. Our results complement the literature that suggests social media reduces information asymmetry and predicts firms' stock returns. The findings differentiate the role of specialized social media in corporate innovation from that of traditional news media.

Considering that SA stands as the world's largest social media platform dedicated to specialized financial information, our findings hold promise for generalizability. The study's implications may extend beyond SA to other specialized social media platforms providing expert information on innovation. Furthermore, since SA coverage encompasses a range of firm-specific non-innovation-related information, the implications of reduced information asymmetry may also extend to other facets of corporate activities, such as corporate governance events and M&A. Thus, the study opens avenues for further research into the real effects of social media on various corporate policies and activities.

## Reference

Acharya, V. V., Baghai, R. P., & Subramanian, K. V. (2014). Wrongful discharge laws and innovation. *Review of Financial Studies*, *27*(1), 301-346.

Acharya, V., & Xu, Z. (2017). Financial dependence and innovation: The case of public versus private firms. *Journal of Financial Economics*, *124*(2), 223-243.

Aghion, P., & Howitt. P. (1992). A model of growth through creative destruction. *Econometrica*, *60*(2), 323-351.

Aghion, P., & Tirole, J. (1994). Opening the black box of innovation. *European Economic Review*, *38*(3-4), 701-710.

Aghion, P., Van Reenen, J., & Zingales, L. (2013). Innovation and institutional ownership. *American Economic Review*, *103*(1), 277-304.

Ak, B. K., Rossi, S., Sloan, R., & Tracy, S. (2016). Navigating stock price crashes. *Journal of Portfolio Management*, *42*(4), 28-37.

Ali, A., & Zhang, W. (2015). CEO tenure and earnings management. *Journal of Accounting and Economics*, *59*(1), 60-79.

Angrist, J.D., and J.-S. Pischke, 2009. Mostly harmless econometrics. Princeton University Press, Princeton, NJ.

Atanassov, J. (2013). Do hostile takeovers stifle innovation? Evidence from antitakeover legislation and corporate patenting. *Journal of Finance*, *68*(3), 1097-1131.

Balsmeier, B., Fleming, L., & Manso, G. (2017). Independent boards and innovation. *Journal of Financial Economics*, *123*(3), 536-557.

Bartov, E., Faurel, L., & Mohanram, P. S. (2018). Can Twitter help predict firm-level earnings and stock returns?. *Accounting Review, 93*(3), 25-57.

Bartov, E., Givoly, D., & Hayn, C. (2002). The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics*, *33*(2), 173-204.

Beck, T., Levine, R., & Levkov, A. (2010). Big bad banks? The winners and losers from bank deregulation in the United States. *Journal of Finance*, *65*(5), 1637-1667.

Bellemare M. F., & Wichman C. J. (2020). Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*, *82*(1), 50-61.

Bena, J., & Li, K. (2014). Corporate innovations and mergers and acquisitions. *Journal of Finance*, *69*(5), 1923-1960.

Bernstein, S. (2015). Does going public affect innovation?. *Journal of Finance*, *70*(4), 1365-1403.

Bhattacharya, N., Desai, H., & Venkataraman, K. (2013). Does earnings quality affect information asymmetry? Evidence from trading costs. *Contemporary Accounting Research*, *30*(2), 482-516.

Bhojraj, S., Hribar, P., Picconi, M., & McInnis, J. (2009). Making sense of cents: An examination of firms that marginally miss or beat analyst forecasts. *Journal of Finance*, *64*(5), 2361-2388.

Blankespoor, E., Miller, G. S., & White, H. D. (2014). The role of dissemination in market liquidity: Evidence from firms' use of Twitter™. *Accounting Review*, *89*(1), 79-112.

Bradley, D., Kim, I., & Tian, X. (2017). Do unions affect innovation? *Management Science*, *63*(7), 2251-2271.

Brav, A., Jiang, W., Ma, S., & Tian, X. (2018). How does hedge fund activism reshape corporate innovation?. *Journal of Financial Economics*, *130*(2), 237-264.

Brennan, M. J., Jegadeesh, N., & Swaminathan, B. (1993). Investment analysis and the adjustment of stock prices to common information. *Review of Financial Studies*, 6(4), 799-824.

Brown, J. R., Martinsson, G., & Petersen, B. C. (2013). Law, stock markets, and innovation. *Journal of Finance*, *68*(4), 1517-1549.

Bushee, B. (1998). The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review*, *73*(3), 305-333.

Bushee, B. J. (2001). Do institutional investors prefer near-term earnings over long-run value? *Contemporary accounting research*, *18(2)*, 207-246.

Campbell, J. L., DeAngelis, M. D., & Moon, J. R. (2019). Skin in the game: Personal stock holdings and investors’ response to stock analysis on social media. *Review of Accounting Studies*, *24*(3), 731-779.

Chava, S., Oettl, A., Subramanian, A., & Subramanian, K. V. (2013). Banking deregulation and innovation. *Journal of Financial Economics*, *109*(3), 759-774.

Chemmanur, T. J., Loutskina, E., & Tian, X. (2014). Corporate venture capital, value creation, and innovation. *Review of Financial Studies*, *27*(8), 2434-2473.

Chen, H., De, P., Hu, Y. J., & Hwang, B. H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies*, *27*(5), 1367-1403.

Chen, J., & Roth, J. (2023). Logs with zeros? Some problems and solutions. *Quarterly Journal of Economics*, forthcoming.

Chernova, Y. (2014). Study: Crowdsourced stock opinions beat analysts. *Dow Jones Institutional News*.

Cohn, J. B., Liu, Z., & Wardlaw, M. I. (2022). Count (and count-like) data in finance. *Journal of Financial Economics*, *146*(2), 529-551.

Cookson, J. A., & Niessner, M. (2020). Why don't we agree? Evidence from a social network of investors. *Journal of Finance*, *75*(1), 173-228.

Correa, J. A., & Ornaghi, C. (2014). Competition & innovation: Evidence from US patent and productivity data. *Journal of Industrial Economics*, *62*(2), 258-285.

Corwin, S. A., & Schultz, P. (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *The Journal of Finance, 67(2)*, 719-760.

Cragg, J. G., & Donald, S. G. (1993). Testing identifiability and specification in instrumental variable models. *Econometric Theory*, *9*(2), 222-240.

Dai, L., Shen, R., & Zhang, B. (2021). Does the media spotlight burn or spur innovation?. *Review of Accounting Studies*, *26*(1), 343-390.

Diamond, D. W., & Verrecchia. R. E. (1991). Disclosure, liquidity, and the cost of capital. *Journal of Finance*, *46*(4), 1325-1359.

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1995). Detecting earnings management. *Accounting Review*, *70*(2), 193-225.

Ding, R., Zhou, H., & Li, Y. (2020). Social media, financial reporting opacity, and return comovement: Evidence from Seeking Alpha. *Journal of Financial Markets*, *50*, 1-25.

Easley, D., & O'Hara, M. (2004). Information and the cost of capital. *Journal of Finance*, *59*(4), 1553-1583.

Ederer, F., & Manso, G. (2013). Is pay for performance detrimental to innovation?. *Management Science*, *59*(7), 1496-1513.

Fang, V. W., Tian, X., & Tice, S. (2014). Does stock liquidity enhance or impede firm innovation?. *Journal of Finance*, *69*(5), 2085-2125.

Frankel, R., & Li, X. (2004). Characteristics of a firm's information environment and the information asymmetry between insiders and outsiders. *Journal of Accounting and Economics*, *37*(2), 229-259.

Francis, J., LaFond, R., Olsson, P., & Schipper, K. (2005). The market pricing of accruals quality. *Journal of Accounting and Economics*, *39*(2), 295-327.

Gibbons, R., & Murphy, K. J. (1992). Optimal incentive contracts in the presence of career concerns: Theory and evidence. *Journal of Political Economy*, *100*(3), 468-505.

Gormley, T.A., & Matsa, D.A., (2011). Growing out of trouble? Corporate responses to liability risk. *Review of Financial Studies 24*(8), 2781-2821.

Gormley, T. A., & Matsa, D. A. (2016). Playing it safe? Managerial preferences, risk, and agency conflicts. *Journal of Financial Economics*, *122*(3), 431-455.

Graham, J. R., Harvey, C. R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics*, *40*(1-3), 3-73.

Guo, B., Pérez-Castrillo, D., & Toldrà-Simats, A. (2019). Firms' innovation strategy under the shadow of analyst coverage. *Journal of Financial Economics*, *131*(2), 456-483.

Guo, H., & Qiu, B. (2016). A better measure of institutional informed trading. *Contemporary Accounting Research*, *33*(2), 815-850.

Hadlock, C. J., & Pierce, J. R. (2010). New evidence on measuring financial constraints: Moving beyond the KZ index. *The Review of Financial Studies, 23(5)*, 1909-1940.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools. *National Bureau of Economic Research*.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics 36*(1), 16-38.

He, J. J., & Tian, X. (2013). The dark side of analyst coverage: The case of innovation. *Journal of Financial Economics*, *109*(3), 856-878.

Heimer, Rawley Z. 2016. Peer pressure: Social interaction and the disposition effect. *Review of Financial Studies*, *29*(11), 3177–3209.

Hirshleifer, D., Hsu, P. H., & Li, D. (2013). Innovative efficiency and stock returns. *Journal of Financial Economics*, *107*(3), 632-654.

Hoberg, G., & Maksimovic, V. (2015). Redefining financial constraints: A text-based analysis. *Review of Financial Studies*, 28(5), 1312–1352.

Hoberg, G., & Phillips, G. (2010). Real and financial industry booms and busts. *Journal of Finance*, *65*(1), 45-86.

Hoberg, G., & Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, *124*(5), 1423-1465.

Hoberg, G., Phillips, G., & Prabhala, N. (2014). Product market threats, payouts, and financial flexibility. *Journal of Finance*, *69*(1), 293-324.

Hong, H., Lim, T., & Stein, J. (2000). Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance*, *55*(1), 265-295.

Hsu, P. H., Tian, X., & Xu, Y. (2014). Financial development and innovation: Cross-country evidence. *Journal of Financial Economics*, *112*(1), 116-135.

Ivković, Z., & Weisbenner, S. (2007). Information diffusion effects in individual investors' common stock purchases: Covet thy neighbours' investment choices. *Review of Financial Studies*, *20*(4), 1327-1357.

Jensen, M. C., & Fuller, J. (2002). Just say no to Wall Street. *Journal of Applied Corporate Finance, 14*(4), 41-46.

Jiang, G. J., Xu, D., & Yao, T. (2009). The information content of idiosyncratic volatility. *Journal of Financial and Quantitative Analysis, 44*(1), 1-28.

Karpoff, J. M., Lee, G., & Masulis, R. W. (2013). Contracting under asymmetric information: Evidence from lockup agreements in seasoned equity offerings. *Journal of Financial Economics*, *110*(3), 607-626.

Kettler, H. (1998). Competition through innovation, innovation through competition. Monograph 000434, Office of Health Economics.

Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics*, *132*(2), 665-712.

Lee, G., & Masulis, R. W. (2009). Seasoned equity offerings: Quality of accounting information and expected flotation costs. *Journal of Financial Economics*, *92*(3), 443-469.

Lee, L. F., Hutton, A. P., & Shu, S. (2015). The role of social media in the capital market: Evidence from consumer product recalls. *Journal of Accounting Research*, *53*(2), 367-404.

Leonardi, P. M., & Treem, J. W. (2012). Knowledge management technology as a stage for strategic self-presentation: Implications for knowledge sharing in organizations. *Information and Organization*, *22*(1), 37-59.

Lerner, J., Sorensen, M., & Strömberg, P. (2011). Private equity and long-run investment: The case of innovation. *Journal of Finance*, *66*(2), 445-477.

Linn, M., & Weagley, D. (2023). *Uncovering Financial Constraints. Journal of Financial and Quantitative Analysis*, forthcoming.

Luo, X., Zhang, J., & Duan, W. (2013). "Social media and firm equity value". *Information Systems Research*, *24*(1), 146-163.

Luong, H., Moshirian, F., Nguyen, L., Tian, X., & Zhang, B. (2017). How do foreign institutional investors enhance firm innovation?. *Journal of Financial and Quantitative Analysis*, *52*(4), 1449-1490.

Matsa, K. E., & Shearer, E. (2018). News use across social media platforms 2018. *Pew Research Center*, *10*.

Matsunaga, S. R., & Park, C. W. (2001). The effect of missing a quarterly earnings benchmark on the CEO's annual bonus. *The Accounting Review*, *76*(3), 313-332.

McLean, R. D., & Zhao, M. (2014). The business cycle, investor sentiment, and costly external finance. *The Journal of Finance*, *69*(3), 1377-1409.

Mergenthaler, R., Rajgopal, S., & Srinivasan, S., (2012). CEO and CFO career penalties to missing quarterly analyst forecasts. Working paper. Pennsylvania State University.

Rajgopal, S., & Venkatachalam, M. (2011). Financial reporting quality and idiosyncratic return volatility. *Journal of Accounting and Economics*, *51*(1-2), 1-20.

Roberts, D. L., Piller, F. T., & Lüttgens, D. (2016). Mapping the impact of social media for innovation: The role of social media in explaining innovation performance in the PDMA comparative performance assessment study. *Journal of Product Innovation Management*, *33*, 117-135.

Schumpeter, J. A. (1942). Capitalism, socialism and democracy. New York: Harper and Brothers.

Seru, A. (2014). Firm boundaries matter: Evidence from conglomerates and R&D activity. *Journal of Financial Economics*, *111*(2), 381-405.

Shiller, R. J. and Pound, J. (1989). Survey evidence on diffusion of interest and information among investors. *Journal of Economic Behavior & Organization*, *12*(1), 47-66.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, Cambridge University Press, 80-108.

Tian, X., & Wang, T. Y. (2014). Tolerance for failure and corporate innovation. *Review of Financial Studies*, *27*(1), 211-255.

**Figure 1. Parallel Trends between Treatment and Control Firms**

This figure shows the dynamic effects of SA coverage initiation on corporate innovation, which are estimated using the following dynamic difference-in-differences Poisson regression:

$$Innovation_{c,i,t} = \beta_2 \sum_{n=-2}^{3} (Treat_{c,i} \times Yr(\Delta t = n)_{c,t}) + f_{c,i} + \lambda_{c,t} + \epsilon$$

where $Innovation_{c,i,t}$ represents two innovation measures: $PatentNum$ is the number of patents firm $i$ of cohort c applied for (and eventually granted) in year $t$, and $CitedNum$ is the number of forward citations received by the patents that firm $i$ of cohort $c$ has applied for (and eventually granted) in year $t$. $Treat$ is an indicator variable that equals one if a cohort-firm is covered by SA (a treatment firm) and equals zero (a control firm) otherwise. $Yr(\Delta t = n)_{c,t}$ represents the year indicator for the $n^{th}$ year relative to the coverage initiation year in cohort $c$, where $\Delta t$ is the year difference between calendar year $t$ and the coverage initiation year in cohort $c$. $Yr(\Delta t = n)_{c,t}$ equals one if $\Delta t = n$ (where $n = \{-2, -1, 0, 1, 2, 3\}$) and equals zero otherwise. We further include cohort-firm fixed effects ($f_{c,i}$) and cohort-year fixed effects ($\lambda_{c,t}$) to account for firm and year unobserved heterogeneities within each cohort. The x-axis shows the years relative to SA coverage initiation ($\Delta t$) or the time difference between calendar year $t$ and the coverage-initiation year of a cohort. The y-axis shows the dynamic treatment effect of SA coverage initiation on corporate innovation measures. The bars represent 95 percent confidence intervals. Standard errors are clustered at cohort-firm level. Table A1 in the Appendix provides detailed variable definitions.

Figure 1A

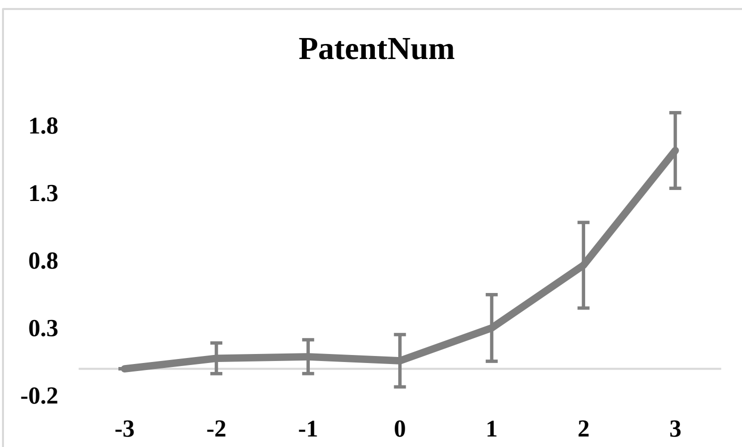


Figure 1B

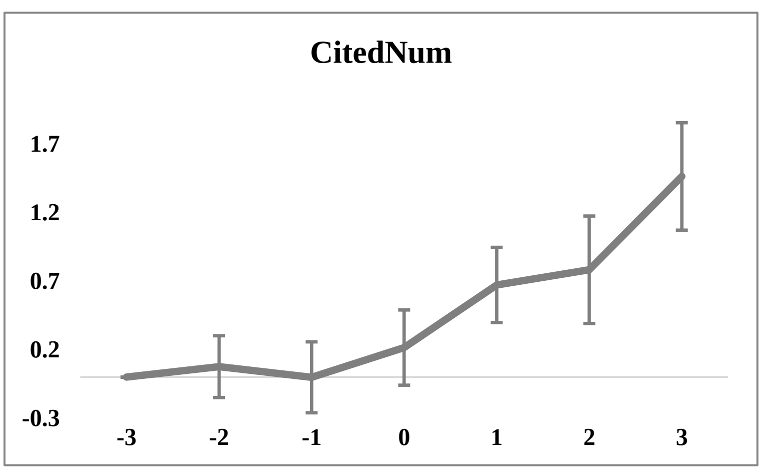

**Table 1. Seeking Alpha Coverage Distribution**

This table presents the number of firms newly covered by SA (*No. firms*) and number of SA articles (*No. articles*) in each year from 2006 to 2017.

| Year | No. firms | No. articles |
|---|---|---|
| 2006 | 139 | 284 |
| 2007 | 392 | 1,139 |
| 2008 | 421 | 3,061 |
| 2009 | 372 | 3,862 |
| 2010 | 373 | 4,001 |
| 2011 | 583 | 5,722 |
| 2012 | 661 | 13,976 |
| 2013 | 1,009 | 15,885 |
| 2014 | 744 | 19,570 |
| 2015 | 521 | 18,904 |
| 2016 | 312 | 15,538 |
| 2017 | 315 | 14,476 |

## Table 2. Ex-ante Firm Characteristics before Seeking Alpha Coverage Initiation

This table compares the firm characteristics between treatment and control firms before the SA coverage initiation events and tests whether ex-ante corporate innovation measures can predict the events. Panel A reports the mean values of firm characteristics between treatment firms and matched control firms in the year prior to SA coverage initiation (i.e., year -1) and the *p*-values from *t*-tests of the differences between treatment and matched control firms. The treatment firms are those firms that SA initiated coverage on between 2011 and 2015. Each treatment firm is matched with up-to-five control firms (which are not covered by SA during the entire seven-year initiation event window) in the same year and the same industry (two-digit SIC). Propensity score is estimated using firm size (*FirmSize*), leverage (*Leverage*), Tobin's Q (*TobinQ*), return-on-asset (*ROA*), R&D expenses (*R&D*), asset tangibility (*Tangibility*), institutional ownership (*InstOwn (%)*), firm age (*FirmAge*), natural logarithm of one plus the number of patents a firm has applied for and later granted (*Log (1+PatentNum)*), and natural logarithm of one plus the number of forward citations received by patents that a firm has applied for and later granted (*Log (1+CitedNum)*), all measured in year -1. In Panel B, we investigate whether the two corporate innovation measures, *Log (1+PatentNum)* and *Log (1+CitedNum)*, can predict the SA coverage initiation event. We include the firm-year observations of the matched sample in these predictive regressions until the treatment event (i.e., SA coverage initiation) of a cohort occurred. We then drop all firm-year observations of the cohort after the SA coverage initiation year. The dependent variable $I\ (SA\ Coverage\ Initiation)$ is an indicator that equals one if a firm is covered by SA in year *t*, and otherwise equals zero. Column 1 does not include any fixed effect or firm controls. Column 2 adds year and firm fixed effects, and column 3 further includes a battery of firm characteristics measured in year *t-1*. Table A1 in the Appendix provides detailed variable definitions. *t*-statistics, calculated based on standard errors clustered at the cohort-firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

***Panel A. Ex-ante Firm Characteristics between Treatment and Control Firms***

| Dependent Var. | Treatment | Control | *p*-value of difference |
|---|---|---|---|
| FirmSize | 6.430 | 6.540 | 0.632 |
| FirmAge | 24.470 | 26.753 | 0.244 |
| Leverage | 0.179 | 0.160 | 0.351 |
| TobinQ | 0.493 | 0.541 | 0.419 |
| ROA | 0.027 | 0.010 | 0.299 |
| R&D | 0.074 | 0.089 | 0.188 |
| Tangibility | 0.160 | 0.146 | 0.337 |
| InstOwn (%) | 52.592 | 57.396 | 0.268 |
| Log (1+PatentNum) | 1.741 | 1.709 | 0.881 |
| Log (1+CitedNum) | 1.541 | 1.591 | 0.816 |
| PatentNum | 40.993 | 40.475 | 0.971 |
| CitedNum | 34.311 | 41.768 | 0.581 |

***Panel B. Predicting SA Coverage Initiation***

| | (1) | (2) | (2) |
|---|---|---|---|
| Dependent Var. | | $I\ (SA\ Coverage\ Initiation)_{t}$ | |
| Log (1+PatentNum) $_{t-1}$ | 0.004 | 0.009 | 0.009 |
| | (0.514) | (0.836) | (0.748) |
| Log (1+CitedNum) $_{t-1}$ | -0.008 | 0.006 | 0.004 |
| | (1.153) | (0.608) | (0.360) |
| FirmSize $_{t-1}$ | | | 0.026 |
| | | | (0.644) |
| Leverage $_{t-1}$ | | | 0.026 |
| | | | (0.312) |
| TobinQ $_{t-1}$ | | | 0.020 |
| | | | (0.869) |
| ROA $_{t-1}$ | | | 0.034 |
| | | | (0.622) |
| R&D $_{t-1}$ | | | -0.197 |
| | | | (0.976) |
| Tangibility $_{t-1}$ | | | 0.357 |
| | | | (1.216) |
| InstOwn $_{t-1}$ | | | 0.000 |
| | | | (0.746) |
| Model | OLS | OLS | OLS |
| Year FE | NO | YES | YES |
| Firm FE | NO | YES | YES |
| Obs. | 2,908 | 2,908 | 2,908 |
| Adj. R2 | 0.001 | 0.210 | 0.211 |

## Table 3. Summary Statistics

This table presents the summary statistics of all variables used in our main difference-in-differences regression sample. The summary statistics include mean, median, quartile (25% and 75%), standard deviation (SD), skewness, kurtosis and number of observations for each variable. The sample period covers 2008 to 2018. Table A1 in the Appendix provides detailed variable definitions.

| Variables | Mean | SD | P25 | Median | P75 | Skewness | Kurtosis | N |
|---|---|---|---|---|---|---|---|---|
| ***Dependent Variables*** | | | | | | | | |
| PatentNum | 34.776 | 113.796 | 0.000 | 1.000 | 9.000 | 4.660 | 25.938 | 5,362 |
| CitedNum | 34.791 | 110.208 | 0.000 | 0.676 | 10.121 | 4.441 | 23.431 | 5,029 |
| EquityRaising | 0.033 | 0.152 | 0.000 | 0.000 | 0.023 | 5.480 | 36.658 | 3,808 |
| DebtRaising | 0.020 | 0.125 | -0.013 | 0.000 | 0.029 | 3.684 | 24.458 | 3,808 |
| LW (Full) | -0.249 | 0.556 | -0.588 | -0.290 | 0.037 | 0.415 | 3.496 | 3,567 |
| LW (Primitive) | -0.249 | 0.558 | -0.633 | -0.314 | 0.081 | 0.539 | 3.277 | 3,567 |
| HP | -3.706 | 0.724 | -4.127 | -3.656 | -3.230 | -0.213 | 2.699 | 3,564 |
| ***Firm-Level Controls*** | | | | | | | | |
| FirmSize | 6.545 | 2.009 | 5.104 | 6.529 | 7.827 | 0.279 | 2.778 | 4,069 |
| Leverage | 0.182 | 0.203 | 0.000 | 0.133 | 0.313 | 1.038 | 3.317 | 4,069 |
| TobinQ | 0.534 | 0.516 | 0.166 | 0.472 | 0.798 | 0.791 | 3.975 | 4,069 |
| ROA | -0.002 | 0.159 | -0.021 | 0.037 | 0.072 | -3.690 | 27.139 | 4,069 |
| R&D | 0.090 | 0.089 | 0.030 | 0.067 | 0.118 | 2.747 | 17.920 | 4,069 |
| Tangibility | 0.150 | 0.117 | 0.064 | 0.120 | 0.205 | 1.504 | 5.825 | 4,069 |
| InstOwn (%) | 54.092 | 36.704 | 14.803 | 70.159 | 87.083 | -0.419 | 1.549 | 4,069 |
| AnalystFollowing | 2.085 | 1.177 | 1.609 | 2.398 | 2.890 | -0.638 | 2.369 | 2,773 |
| #News | 1.857 | 1.504 | 0.840 | 1.370 | 2.470 | 1.699 | 6.201 | 2,773 |
| NewsTone (%) | 16.853 | 16.577 | 6.173 | 16.895 | 28.352 | -0.227 | 3.634 | 2,773 |
| StockReturn (%) | 13.487 | 53.160 | -19.400 | 5.890 | 36.390 | 1.304 | 6.237 | 2,773 |
| StockVolatility (%) | 12.726 | 6.053 | 8.600 | 11.350 | 15.270 | 1.976 | 10.182 | 2,773 |

## Table 4. Seeking Alpha Coverage Initiation and Corporate Innovation

This table reports the treatment effects of SA coverage initiation on corporate innovation, which is estimated using the following difference-in-differences Poisson regression:

$$Innovation_{c,i,t} = \beta_1 Treat_{c,i} \times Post_{c,t} + Firm\ Controls_{c,i,t-1} + f_{c,i} + \lambda_{c,t} + \epsilon$$

where $Innovation_{c,i,t}$ represents two innovation measures: $PatentNum$ is the number of patents firm *i* of cohort c applied for (and eventually granted) in year *t*, and $CitedNum$ is the number of forward citations received by the patents that firm *i* of cohort *c* has applied for (and eventually granted) in year *t*. $Treat_{c,i}$ is an indicator variable that equals one if firm $i$ in cohort $c$ is covered by SA (a treatment firm) and equals zero otherwise. $Post_{c,t}$ is an indicator variable that equals one if year $t$ is in the post-event period (i.e., year [1, 3]) and equals zero if it is in the pre-event period (i.e., year [-3, -1]). All specifications include cohort-firm fixed effects ($f_{c,i}$) and cohort-year fixed effects ($\lambda_{c,t}$). Columns 2 and 4 further control for lagged firm characteristics. Table A1 in the Appendix provides detailed variable definitions. $t$-statistics, calculated based on standard errors clustered at the cohort-firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| $Treat \times Post$ | 0.722*** | 0.633*** | 0.779*** | 0.832*** |
| | (4.675) | (5.694) | (3.800) | (6.705) |
| $FirmSize_{t-1}$ | | 0.986*** | | 1.222*** |
| | | (4.216) | | (4.752) |
| $Leverage_{t-1}$ | | 0.753*** | | 0.709** |
| | | (2.871) | | (2.450) |
| $TobinQ_{t-1}$ | | 0.583*** | | 0.828*** |
| | | (6.105) | | (5.643) |
| $ROA_{t-1}$ | | -0.636** | | -0.480* |
| | | (2.198) | | (1.712) |
| $R\&D_{t-1}$ | | 6.015*** | | 7.349*** |
| | | (3.494) | | (3.732) |
| $Tangibility_{t-1}$ | | 0.306 | | 1.708 |
| | | (0.266) | | (1.510) |
| $InstOwn_{t-1}$ | | 0.004*** | | 0.005*** |
| | | (3.092) | | (5.905) |
| Model | POISSON | POISSON | POISSON | POISSON |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 4,543 | 3,370 | 4,196 | 3,190 |
| Pseudo R2 | 0.952 | 0.960 | 0.945 | 0.958 |

## Table 5. Dynamics of Corporate Innovation around Seeking Alpha Coverage Initiation

This figure shows the dynamic effects of SA coverage initiation on corporate innovation, which are estimated using the following dynamic difference-in-differences Poisson regression:

$$Innovation_{c,i,t} = \beta_2 \sum_{n=-2}^{3} (Treat_{c,i} \times Yr(\Delta t = n)_{c,t}) + Firm\ Controls_{c,i,t-1} + f_{c,i} + \lambda_{c,t} + \epsilon$$

where $Innovation_{c,i,t}$ represents two innovation measures: $PatentNum$ is the number of patents firm $i$ of cohort c applied for (and eventually granted) in year $t$, and $CitedNum$ is the number of forward citations received by the patents that firm $i$ of cohort $c$ has applied for (and eventually granted) in year $t$. $Treat$ is an indicator variable that equals one if a cohort-firm is covered by SA (a treatment firm) and equals zero (a control firm) otherwise. $Yr(\Delta t = n)_{c,t}$ represents the year indicator for the $n^{th}$ year relative to the coverage initiation year in cohort $c$, where $\Delta t$ is the year difference between calendar year $t$ and the coverage initiation year in cohort $c$. $Yr(\Delta t = n)_{c,t}$ equals one if $\Delta t = n$ (where $n = \{-2, -1, 0, 1, 2, 3\}$) and equals zero otherwise. All specifications include cohort-firm fixed effects ($f_{c,i}$) and cohort-year fixed effects ($\lambda_{c,t}$). Columns 2 and 4 further control for lagged firm characteristics. Table A1 in the Appendix provides detailed variable definitions. $t$-statistics, calculated based on standard errors clustered at the cohort-firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| $Treat \times Yr(\Delta t = -2)$ | 0.036 | 0.078 | 0.025 | 0.076 |
| | (0.754) | (1.339) | (0.244) | (0.658) |
| $Treat \times Yr(\Delta t = -1)$ | 0.076 | 0.090 | -0.072 | -0.002 |
| | (1.108) | (1.402) | (0.560) | (0.014) |
| $Treat \times Yr(\Delta t = 0)$ | 0.094 | 0.060 | 0.180 | 0.215 |
| | (0.930) | (0.613) | (1.247) | (1.537) |
| $Treat \times Yr(\Delta t = 1)$ | 0.289* | 0.303** | 0.533*** | 0.672*** |
| | (1.955) | (2.413) | (3.024) | (4.783) |
| $Treat \times Yr(\Delta t = 2)$ | 0.742*** | 0.768*** | 0.670*** | 0.783*** |
| | (3.728) | (4.744) | (2.644) | (3.921) |
| $Treat \times Yr(\Delta t = 3)$ | 1.852*** | 1.619*** | 1.371*** | 1.464*** |
| | (7.735) | (11.328) | (3.773) | (7.329) |
| Model | POISSON | POISSON | POISSON | POISSON |
| Firm Controls | NO | YES | NO | YES |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 5,362 | 4,069 | 5,029 | 3,907 |
| Pseudo R2 | 0.956 | 0.962 | 0.947 | 0.957 |

## Table 6. Seeking Alpha Coverage Initiation, Financing Activities, and Financial Constraint

This table compares corporate financing activities and financial constraints for treatment and control firms three years before and after the SA coverage initiation. $EquityRaising$ is measured as the change in book equity, plus the change in deferred taxes, minus the change in retained earnings, all scaled by lagged assets. $DebtRaising$ is measured as the change in assets, minus the change in book equity, minus the change in deferred taxes, all scaled by lagged assets. *LW (Full)* is the financial constraint index of a firm-year estimated by Linn and Weagley (2023) using the full model (based on a wide range of accounting variables). *LW (Primitive)* is the financial constraint index of a firm-year estimated by Linn and Weagley (2023) using the primitive model (based on four primitive accounting variables). $HP$ is the financial constraint index of a firm-year estimated by Hadlock and Pierce (2010). $Treat$ is an indicator variable that equals one if a cohort-firm is covered by SA (a treatment firm) or zero (a control firm) otherwise. $Post$ is an indicator variable that equals one for the post-event period (year [1, 3]) or zero for the pre-event period (year [-3, -1]). All specifications include cohort-firm and cohort-year fixed effects. Columns 2 and 4 further control for lagged firm characteristics. Table A1 in the Appendix provides detailed variable definitions. $t$-statistics, calculated based on standard errors clustered at the cohort-firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

***Panel A. Financing Activities***

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $EquityRaising_t$ | | $DebtRaising_t$ | |
| $Treat \times Post$ | 0.024** | 0.023** | -0.010 | -0.020* |
| | (2.155) | (2.129) | (0.980) | (1.955) |
| $FirmSize_{t-1}$ | | -0.064*** | | -0.086*** |
| | | (3.530) | | (4.312) |
| $Leverage_{t-1}$ | | -0.051 | | -0.134*** |
| | | (1.272) | | (3.038) |
| $TobinQ_{t-1}$ | | 0.013 | | 0.019 |
| | | (0.834) | | (0.958) |
| $ROA_{t-1}$ | | -0.012 | | 0.013 |
| | | (0.178) | | (0.287) |
| $R\&D_{t-1}$ | | -0.083 | | -0.596*** |
| | | (0.438) | | (3.245) |
| $Tangibility_{t-1}$ | | -0.139 | | 0.381*** |
| | | (1.456) | | (3.756) |
| $InstOwn_{t-1}$ | | 0.000 | | 0.000* |
| | | (0.290) | | (1.888) |
| Model | OLS | OLS | OLS | OLS |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 3,808 | 2,904 | 3,808 | 2,904 |
| Adj. R2 | 0.412 | 0.342 | -0.064 | 0.018 |

***Panel B. Financial Constraints***

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Var. | *LW (Full)* $_t$ | | *LW (Primitive)* $_t$ | | *HP* $_t$ | |
| $Treat \times Post$ | -0.159*** | -0.153*** | -0.107*** | -0.098** | -0.028** | -0.012* |
| | (4.352) | (4.436) | (2.610) | (2.430) | (2.382) | (1.664) |
| FirmSize $_{t-1}$ | | -0.207*** | | -0.255*** | | -0.215*** |
| | | (3.926) | | (4.517) | | (21.337) |
| Leverage $_{t-1}$ | | 0.213* | | 0.268** | | 0.134*** |
| | | (1.727) | | (2.028) | | (5.216) |
| TobinQ $_{t-1}$ | | 0.044 | | 0.051 | | -0.038*** |
| | | (1.372) | | (1.505) | | (5.182) |
| ROA $_{t-1}$ | | -0.228*** | | -0.189** | | -0.056* |
| | | (3.365) | | (2.477) | | (1.915) |
| R&D $_{t-1}$ | | -0.136 | | -0.304 | | -0.129* |
| | | (0.592) | | (1.144) | | (1.654) |
| Tangibility $_{t-1}$ | | -1.389*** | | -1.069*** | | -0.167** |
| | | (5.383) | | (3.277) | | (1.992) |
| InstOwn $_{t-1}$ | | -0.001** | | -0.001** | | 0.000*** |
| | | (2.001) | | (2.093) | | (3.109) |
| Model | OLS | OLS | OLS | OLS | OLS | OLS |
| Cohort-Firm FE | YES | YES | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES | YES | YES |
| Obs. | 3,567 | 3,217 | 3,567 | 3,217 | 3,654 | 3,543 |
| Adj. R2 | 0.728 | 0.749 | 0.661 | 0.665 | 0.986 | 0.993 |

## Table 7. Treatment Effect Heterogeneity: Information Asymmetry

This table presents the test results for treatment effect heterogeneity based on four proxies of information asymmetry. Analyst following ($AnalystFollowing$) is the number of analysts following of a firm in a year. Abnormal accruals ($AbnormalAccrual$) is the median absolute value of discretionary accruals in the past five fiscal years, where discretionary accruals are estimated based on the modified Jones model. Idiosyncratic volatility ($IdiosyncraticVolatility$) is the standard deviation of the residuals from regressing daily individual stock returns on the Fama-French three-factors of a firm in a year. Effective spread ($EffectiveSpread$) is the bid-ask spreads estimated using daily high and low stock prices of a firm following Corwin and Schultz (2012). Based on the median value of each proxy measured in year -1, the sample is split into above-median (H) and below-median (L) subsamples. The dependent variables $PatentNum$ is the number of patents firm $i$ of cohort $c$ applied for (and eventually granted) in year $t$; $CitedNum$ is the number of forward citations received by the patents that firm *i* of cohort *c* has applied for (and eventually granted) in year $t$. $Treat$ is an indicator variable that equals one if a cohort-firm is covered by SA (a treatment firm) and equals zero (a control firm) otherwise. $Post$ is an indicator variable that equals one for the post-event period (year [1, 3]) and equals zero for the pre-event period (year [-3, -1]). We also report *p*-values of Chow tests on whether the coefficient estimates on $Treat \times Post$ for any pair of H and L subsamples are the same. All specifications include cohort-firm and cohort-year fixed effects. Table A1 in the Appendix provides detailed variable definitions. *t*-statistics, calculated based on standard errors clustered at the cohort-firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

***Panel A. Analyst Following***

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| | **H** | **L** | **H** | **L** |
| $Treat \times Post$ | 0.754*** | 1.423*** | 1.111*** | 1.430*** |
| | (5.024) | (10.237) | (8.723) | (8.290) |
| *p*-value of Difference | 0.001 | | 0.136 | |
| Model | POISSON | POISSON | POISSON | POISSON |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 1,698 | 1,979 | 1,580 | 1,840 |
| Pseudo R2 | 0.952 | 0.952 | 0.946 | 0.946 |

***Panel B. Abnormal Accruals***

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| | **H** | **L** | **H** | **L** |
| $Treat \times Post$ | 1.287*** | 0.617*** | 1.272*** | 0.658*** |
| | (4.917) | (3.598) | (5.174) | (2.891) |
| *p*-value of Difference | 0.032 | | 0.066 | |
| Model | POISSON | POISSON | POISSON | POISSON |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 2,319 | 2,176 | 2,184 | 1,974 |
| Pseudo R2 | 0.955 | 0.955 | 0.943 | 0.943 |

***Panel C. Idiosyncratic Volatility***

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| | **H** | **L** | **H** | **L** |
| $Treat \times Post$ | 1.450*** | 0.485* | 1.555*** | 0.582* |
| | (4.044) | (1.928) | (6.694) | (1.879) |
| *p*-value of Difference | 0.028 | | 0.012 | |
| Model | POISSON | POISSON | POISSON | POISSON |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 1,806 | 1,839 | 1,665 | 1,682 |
| Pseudo R2 | 0.952 | 0.952 | 0.944 | 0.944 |

***Panel D. Effective Spread***

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| | **H** | **L** | **H** | **L** |
| $Treat \times Post$ | 1.523*** | 0.478* | 1.450*** | 0.585* |
| | (5.933) | (1.907) | (6.147) | (1.859) |
| *p*-value of Difference | 0.004 | | 0.028 | |
| Model | POISSON | POISSON | POISSON | POISSON |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 1,836 | 1,809 | 1,678 | 1,669 |
| Pseudo R2 | 0.951 | 0.951 | 0.945 | 0.945 |

## Table 8. Treatment Effect Heterogeneity: Market Competition

This table presents the test results for treatment effect heterogeneity based on four proxies of market competition. Product market fluidity ($Fluidity$) is a text-based measure of competitive threats faced by a firm in the product market, which captures changes in rival firms' product description relative to the firm's in 10-Ks. TNIC HHI ($TNICHHI$) is the Herfindahl-Hirschman index based on text-based network industry classification. HHI ($HHI$) is the Herfindahl-Hirschman index based on 2-digit SIC industry classification. Lerner index ($LernerIndex$) is the median gross margin in the firm's two-digit SIC industry. Based on the median value of each proxy measured in year -1, the sample is split into above-median (H) and below-median (L) subsamples. The dependent variables $PatentNum$ is the number of patents firm $i$ of cohort c applied for (and eventually granted) in year $t$; $CitedNum$ is the number of forward citations received by the patents that firm $i$ of cohort $c$ has applied for (and eventually granted) in year $t$. $Treat$ is an indicator variable that equals one if a cohort-firm is covered by SA (a treatment firm) and equals zero (a control firm) otherwise. $Post$ is an indicator variable that equals one for the post-event period (year [1, 3]) and equals zero for the pre-event period (year [-3, -1]). We also report $p$-values of Chow tests on whether the coefficient estimates on $Treat \times Post$ for any pair of H and L subsamples are the same. All specifications include cohort-firm and cohort-year fixed effects. Table A1 in the Appendix provides detailed variable definitions. $t$-statistics, calculated based on standard errors clustered at the cohort-firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

***Panel A. Product Market Fluidity***

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| | **H** | **L** | **H** | **L** |
| $Treat \times Post$ | 0.998*** | 0.409 | 1.367*** | 0.392 |
| | (6.218) | (1.273) | (8.103) | (1.191) |
| $p$-value of Difference | 0.100 | | 0.008 | |
| Model | POISSON | POISSON | POISSON | POISSON |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 1,983 | 1,878 | 1,880 | 1,685 |
| Pseudo R2 | 0.952 | 0.952 | 0.951 | 0.951 |

***Panel B. TNIC HHI***

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| | **H** | **L** | **H** | **L** |
| $Treat \times Post$ | 0.187 | 1.083*** | 0.156 | 1.379*** |
| | (0.604) | (6.477) | (0.539) | (8.589) |
| $p$-value of Difference | 0.011 | | 0.000 | |
| Model | POISSON | POISSON | POISSON | POISSON |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 1,856 | 2,065 | 1,728 | 1,889 |
| Pseudo R2 | 0.946 | 0.946 | 0.943 | 0.943 |

***Panel C. HHI***

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| | **H** | **L** | **H** | **L** |
| $Treat \times Post$ | 0.303 | 0.944*** | 0.280 | 1.198*** |
| | (1.224) | (6.091) | (1.013) | (10.598) |
| *p*-value of Difference | 0.028 | | 0.002 | |
| Model | POISSON | POISSON | POISSON | POISSON |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 1,939 | 2,604 | 1,766 | 2,430 |
| Pseudo R2 | 0.959 | 0.959 | 0.934 | 0.934 |

***Panel D. Lerner Index***

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| | **H** | **L** | **H** | **L** |
| $Treat \times Post$ | 0.251 | 0.994*** | 0.283 | 1.239*** |
| | (1.182) | (5.995) | (1.091) | (10.060) |
| *p*-value of Difference | 0.006 | | 0.001 | |
| Model | POISSON | POISSON | POISSON | POISSON |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 2,255 | 2,288 | 2,089 | 2,107 |
| Pseudo R2 | 0.959 | 0.959 | 0.941 | 0.941 |

## Table 9. Treatment Effect Heterogeneity: Short-termism

This table presents the test results for treatment effect heterogeneity based on four proxies of short-termism. Transient institutional ownership ($TransientInstOwn$) is the percentage of shares of a firm owned by transient institutional investors in a year following Bushee (2001). Dedicated institutional ownership ($DedicatedInstOwn$) is the percentage of shares of a firm owned by dedicated institutional investors in a year following Bushee (2001). CEO age ($CEOAge$) is the age of the CEO in a firm-year. CEO tenure ($CeoTenure$) is the tenure year of the CEO in a firm-year. Based on the median value of each proxy measured in year -1, the sample is split into above-median (H) and below-median (L) subsamples. The dependent variables $PatentNum$ is the number of patents firm *i* of cohort c applied for (and eventually granted) in year *t*; $CitedNum$ is the number of forward citations received by the patents that firm *i* of cohort *c* has applied for (and eventually granted) in year *t*. $Treat$ is an indicator variable that equals one if a cohort-firm is covered by SA (a treatment firm) and equals zero (a control firm) otherwise. $Post$ is an indicator variable that equals one for the post-event period (year [1, 3]) and equals zero for the pre-event period (year [-3, -1]). We also report *p*-values of Chow tests on whether the coefficient estimates on $Treat \times Post$ for any pair of H and L subsamples are the same. All specifications include cohort-firm and cohort-year fixed effects. Table A1 in the Appendix provides detailed variable definitions. *t*-statistics, calculated based on standard errors clustered at the cohort-firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

***Panel A. Transient Institutional Ownership***

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| | **H** | **L** | **H** | **L** |
| $Treat \times Post$ | 0.390 | 0.894*** | 0.477 | 1.043*** |
| | (1.516) | (5.393) | (1.576) | (5.314) |
| *p*-value of Difference | 0.099 | | 0.117 | |
| Model | POISSON | POISSON | POISSON | POISSON |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 2,268 | 2,275 | 2,153 | 2,043 |
| Pseudo R2 | 0.956 | 0.956 | 0.949 | 0.949 |

***Panel B. Dedicated Institutional Ownership***

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| | **H** | **L** | **H** | **L** |
| $Treat \times Post$ | 2.413*** | 0.616*** | 2.392*** | 0.596*** |
| | (6.955) | (3.949) | (8.709) | (3.050) |
| *p*-value of Difference | 0.000 | | 0.000 | |
| Model | POISSON | POISSON | POISSON | POISSON |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 2,333 | 2,210 | 2,128 | 2,068 |
| Pseudo R2 | 0.953 | 0.953 | 0.947 | 0.947 |

***Panel C. CEO Age***

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| | **H** | **L** | **H** | **L** |
| $Treat \times Post$ | 1.094*** | 0.300 | 1.260*** | 0.453 |
| | (6.206) | (1.006) | (7.986) | (1.284) |
| *p*-value of Difference | 0.022 | | 0.037 | |
| Model | POISSON | POISSON | POISSON | POISSON |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 1,206 | 1,194 | 1,112 | 1,086 |
| Pseudo R2 | 0.950 | 0.950 | 0.950 | 0.950 |

***Panel D. CEO Tenure***

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent Var. | $PatentNum_t$ | | $CitedNum_t$ | |
| | **H** | **L** | **H** | **L** |
| $Treat \times Post$ | 1.278*** | 0.219 | 1.430*** | 0.378 |
| | (5.194) | (0.854) | (8.434) | (1.180) |
| *p*-value of Difference | 0.003 | | 0.004 | |
| Model | POISSON | POISSON | POISSON | POISSON |
| Cohort-Firm FE | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES |
| Obs. | 1,197 | 1,203 | 1,104 | 1,094 |
| Pseudo R2 | 0.945 | 0.945 | 0.951 | 0.951 |

## Table 10. Seeking Alpha Article Intensity and Corporate Innovation: Two-stage Least Squares Regressions

This table presents the two-stage least squares regression results to address endogeneity concerns. The instrument variable in the first stage is $Log\ (1 + FollowerNum)$, the natural logarithm of one plus the aggregated number of followers of all SA authors who publish articles on a firm in year *t*, which is measured by summing up the past follower numbers of the authors before they write on the firm. Panel A reports the results on total SA articles, where $IHS\ (ArticleNum)$ is the inversed hyperbolic sine of the total number of SA articles related to a firm in year *t*. Panel B reports the results on innovation-focused SA articles, where $IHS\ (InnoArticleNum)$ is the inversed hyperbolic sine of the number of innovation-focused SA articles on a firm in year *t*. Panel C reports the results on non-innovation-focused SA articles, where $IHS\ (NonInnoArticleNum)$ is the inversed hyperbolic sine of the number of non-innovation-focused SA articles on a firm in year *t*. In all panels, the dependent variables in the second stage are innovation outputs measured in year *t+1*, including the inversed hyperbolic sine of the number of patents a firm has applied for and later granted ($IHS\ (PatentNum)$), and the inversed hyperbolic sine of the number of forward citations received by the patents that a firm has applied for and later granted ($IHS\ (CitedNum)$). Control variables include firm size (*FirmSize*), firm age (*FirmAge*), leverage (*Leverage*), Tobin's Q (*TobinQ*), return-on-asset (*ROA*), R&D (*R&D*), asset tangibility (*Tangibility*), institutional ownership (*InstOwn*), last 12 months' stock returns (*StockReturn*), return volatility (*StockVolatility*), analyst following (*AnalystFollowing*), news coverage (*#News*), and news tone (*NewsTone*), all estimated in year *t*. All specifications include firm and year fixed effects. We omit the coefficients of control variables in Panel B and C for brevity. Table A1 in the Appendix provides detailed variable definitions. *t*-statistics, calculated based on standard errors clustered at the firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

***Panel A. Total SA Articles***

| | First Stage | Second Stage | |
|---|---|---|---|
| Dependent Var. | $IHS\ (ArticleNum)_t$ | $IHS\ (PatentNum)_{t+1}$ | $IHS\ (CitedNum)_{t+1}$ |
| $\text{Log}\ (1 + \text{FollowerNum})_t$ | 0.106*** | | |
| | (20.236) | | |
| $\text{IHS (ArticleNum)}_t$ | | 0.117** | 0.117** |
| | | (2.383) | (2.054) |
| FirmSize $_t$ | 0.094*** | 0.093** | 0.059 |
| | (2.958) | (2.081) | (1.296) |
| FirmAge $_t$ | -0.209** | -0.270** | -0.271* |
| | (2.305) | (2.215) | (1.895) |
| Leverage $_t$ | -0.042 | -0.064 | -0.023 |
| | (0.538) | (0.672) | (0.197) |
| TobinQ $_t$ | 0.078** | 0.031 | 0.059 |
| | (2.063) | (0.695) | (1.072) |
| ROA $_t$ | -0.077 | -0.025 | -0.038 |
| | (1.309) | (0.364) | (0.508) |
| R&D $_t$ | -0.086 | -0.065 | -0.243 |
| | (0.504) | (0.319) | (1.219) |
| Tangibility $_t$ | -0.129 | 0.275 | 0.281 |
| | (0.716) | (1.496) | (1.345) |
| InstOwn $_t$ | -0.003*** | 0.003** | 0.003** |
| | (3.534) | (2.316) | (2.015) |
| StockReturn $_t$ | -0.027 | 0.002 | 0.004 |
| | (1.275) | (0.107) | (0.148) |
| StockVolatility $_t$ | 0.532** | -0.069 | 0.015 |

| | | | |
|---|---|---|---|
| | (2.472) | (0.315) | (0.060) |
| AnalystFollowing $_t$ | 0.073** | -0.058 | -0.107*** |
| | (2.285) | (1.643) | (2.607) |
| #News $_t$ | 0.086*** | -0.016 | -0.001 |
| | (5.217) | (0.659) | (0.052) |
| NewsTone $_t$ | -0.002*** | 0.001 | 0.001 |
| | (3.598) | (1.475) | (1.527) |
| Model | OLS | OLS | OLS |
| Year FE | YES | YES | YES |
| Firm FE | YES | YES | YES |
| Obs. | 5,142 | 5,142 | 5,142 |
| Adj. R2 | 0.624 | 0.942 | 0.919 |

***Panel B. Innovation-focused SA Articles***

| | First Stage | Second Stage | |
|---|---|---|---|
| Dependent Var. | $IHS\ (InnoArticleNum)_t$ | $IHS\ (PatentNum)_{t+1}$ | $IHS\ (CitedNum)_{t+1}$ |
| Log (1 + FollowerNum) $_t$ | 0.022*** | | |
| | (7.065) | | |
| IHS (InnoArticleNum) $_t$ | | 0.434** | 0.404* |
| | | (2.212) | (1.771) |
| Model | OLS | OLS | OLS |
| Firm Controls | YES | YES | YES |
| Year FE | YES | YES | YES |
| Firm FE | YES | YES | YES |
| Obs. | 5,142 | 5,142 | 5,142 |
| Adj. R2 | 0.706 | 0.942 | 0.919 |

***Panel C. Non-innovation-focused SA Articles***

| | First Stage | Second Stage | |
|---|---|---|---|
| Dependent Var. | $IHS\ (NonInnoArticleNum)_t$ | $IHS\ (PatentNum)_{t+1}$ | $IHS\ (CitedNum)_{t+1}$ |
| Log (1 + FollowerNum) $_t$ | 0.084*** | | |
| | (19.039) | | |
| IHS (NonInnoArticleNum) $_t$ | | 0.113** | 0.105* |
| | | (2.283) | (1.819) |
| Model | OLS | OLS | OLS |
| Firm Controls | YES | YES | YES |
| Year FE | YES | YES | YES |
| Firm FE | YES | YES | YES |
| Obs. | 5,142 | 5,142 | 5,142 |
| Adj. R2 | 0.734 | 0.942 | 0.919 |

## Table 11. Seeking Alpha Article Comments and Corporate Innovation: Two-stage Least Squares Regressions

This table presents the two-stage least squares regression results to address endogeneity concerns. The instrument variable in the first stage is $Log\ (1 + FollowerNum)$, the natural logarithm of one plus the aggregated number of followers of all SA authors who publish articles on a firm in year *t*, which is measured by summing up the past follower numbers of the authors before they write on the firm. Panel A reports the results on average number of comments per SA article on a firm, where $IHS\ (AvgComtNum)$ is the inversed hyperbolic sine of the average number of comments (excluding those made by authors) per article related to a firm in year *t*. Panel B reports the results on maximum number of comments per SA article on a firm, where $IHS\ (MaxComtNum)$ is the inversed hyperbolic sine of the maximum number of comments (excluding those made by authors) per article related to a firm in year *t*. In both panels, the dependent variables in the second stage are innovation outputs measured in year *t+1*, including the inversed hyperbolic sine of the number of patents a firm has applied for and later granted ($IHS\ (PatentNum)$), and the inversed hyperbolic sine of the number of forward citations received by the patents that a firm has applied for and later granted ($IHS\ (CitedNum)$). Control variables include firm size (*FirmSize*), firm age (*FirmAge*), leverage (*Leverage*), Tobin's Q (*TobinQ*), return-on-asset (*ROA*), R&D (*R&D*), asset tangibility (*Tangibility*), institutional ownership (*InstOwn*), last 12 months' stock returns (*StockReturn*), return volatility (*StockVolatility*), analyst following (*AnalystFollowing*), news coverage *(#News*), and news tone (*NewsTone*), all estimated in year *t*. All specifications include firm and year fixed effects. We omit the coefficients of control variables in Panel B for brevity. Table A1 in the Appendix provides detailed variable definitions. *t*-statistics, calculated based on standard errors clustered at the firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

***Panel A. Average Number of Comments per SA article on a Firm***

| | First Stage | Second Stage | |
|---|---|---|---|
| Dependent Var. | $IHS\ (AvgComtNum)_t$ | $IHS\ (PatentNum)_{t+1}$ | $IHS\ (CitedNum)_{t+1}$ |
| $\text{Log}\ (1 + \text{FollowerNum})_t$ | 0.012*** | | |
| | (3.278) | | |
| $\text{IHS (AvgComtNum)}_t$ | | 1.035* | 1.041* |
| | | (1.944) | (1.741) |
| $\text{FirmSize}_t$ | 0.032** | 0.049 | 0.019 |
| | (2.152) | (1.118) | (0.388) |
| $\text{FirmAge}_t$ | -0.004 | -0.232** | -0.231* |
| | (0.099) | (2.030) | (1.649) |
| $\text{Leverage}_t$ | 0.016 | -0.077 | -0.043 |
| | (0.429) | (0.897) | (0.391) |
| $\text{TobinQ}_t$ | 0.024 | 0.007 | 0.032 |
| | (1.370) | (0.152) | (0.598) |
| $\text{ROA}_t$ | 0.022 | -0.055 | -0.062 |
| | (0.653) | (0.792) | (0.818) |
| $\text{R\&D}_t$ | 0.053 | -0.116 | -0.276 |
| | (0.655) | (0.637) | (1.400) |
| $\text{Tangibility}_t$ | -0.035 | 0.275 | 0.268 |
| | (0.395) | (1.507) | (1.308) |
| $\text{InstOwn}_t$ | -0.001*** | 0.003** | 0.003** |
| | (2.605) | (2.506) | (2.182) |
| $\text{StockReturn}_t$ | -0.020* | 0.019 | 0.023 |
| | (1.930) | (0.822) | (0.793) |
| $\text{StockVolatility}_t$ | 0.071 | -0.068 | 0.001 |
| | (0.744) | (0.323) | (0.004) |

| | | | |
|---|---|---|---|
| AnalystFollowing $_t$ | 0.021 | -0.064* | -0.102** |
| | (1.388) | (1.740) | (2.417) |
| #News $_t$ | -0.004 | 0.003 | 0.015 |
| | (0.560) | (0.122) | (0.592) |
| NewsTone $_t$ | -0.001 | 0.001* | 0.002* |
| | (1.540) | (1.722) | (1.771) |
| Model | OLS | OLS | OLS |
| Year FE | YES | YES | YES |
| Firm FE | YES | YES | YES |
| Obs. | 5,142 | 5,142 | 5,142 |
| Adj. R2 | 0.178 | 0.942 | 0.919 |

***Panel B. Maximum Number of Comments per SA Article on a Firm***

| | First Stage | Second Stage | |
|---|---|---|---|
| Dependent Var. | $IHS\ (MaxComtNum)_t$ | $IHS\ (PatentNum)_{t+1}$ | $IHS\ (CitedNum)_{t+1}$ |
| $\text{Log}\ (1 + \text{FollowerNum})_t$ | 0.040*** | | |
| | (9.505) | | |
| $\text{IHS}\ (\text{MaxComtNum})_t$ | | 0.312** | 0.313** |
| | | (2.335) | (2.011) |
| Model | OLS | OLS | OLS |
| Firm Controls | YES | YES | YES |
| Year FE | YES | YES | YES |
| Firm FE | YES | YES | YES |
| Obs. | 5,142 | 5,142 | 5,142 |
| Adj. R2 | 0.371 | 0.942 | 0.919 |

## Appendix

### Table A1. Variable Descriptions

| Variable | Definition |
|---|---|
| *Panel A. Dependent variables* | |
| PatentNum | The number of patents a firm has applied for (and eventually granted) in a year (Source: Kogan et al. (2017)). |
| CitedNum | The number of forward citations received by the patents that a firm has applied for (and eventually granted) in a year (Source: Kogan et al. (2017)). |
| Efficiency | The number of patents a firm has applied (and eventually granted) in a year per million R&D dollars (Source: Kogan et al. (2017) and Compustat). |
| PatentValue | The economic values of patents a firm has applied (and eventually granted) in a year scaled by the book value of total assets of that firm-year (Source: Kogan et al. (2017)). |
| EfficiencyValue | The economic values of patents a firm has applied (and eventually granted) in a year per million R&D dollars (Source: Kogan et al. (2017) and Compustat). |
| Originality | One minus the average Herfindahl index of backward citations of patents applied in a firm-year (Source: Kogan et al. (2017) and USPTO Bulk Downloads). |
| Generality | One minus the average Herfindahl index of forward citations of patents applied in a firm-year (Source: Kogan et al. (2017) and USPTO Bulk Downloads). |
| EquityRaising | The change in book equity, plus the change in deferred taxes, minus the change in retained earnings, all scaled by lagged assets (Source: Compustat: $(\Delta CEQ + \Delta TXDB - \Delta RE)/\ lagged\ AT$). |
| DebtRaising | The change in assets, minus the change in book equity, minus the change in deferred taxes, all scaled by lagged assets (Source: Compustat: $(\Delta \mathrm{AT} - \Delta CEQ - \Delta TXDB)/\ lagged\ AT$). |
| LW (Full) | The financial constraint index of a firm-year estimated by Linn and Weagley (2023) using the full model (based on a wide range of accounting variables). |
| LW (Primitive) | The financial constraint index of a firm-year estimated by Linn and Weagley (2023) using the primitive model (based on four primitive accounting variables). |
| HP | The financial constraint index of a firm-year estimated by Hadlock and Pierce (2010). |
| Log (1+FollowerNum) | The natural logarithm of one plus the aggregated number of followers of all SA authors who publish articles on a firm in a year, which is measured by summing up the past follower numbers of the authors before they write on the firm (Source: Seeking Alpha platform). |
| IHS (ArticleNum) | The inversed hyperbolic sine of the total number of SA articles related to a firm in a year (Source: Seeking Alpha platform). |
| IHS (InnoArticleNum) | The inversed hyperbolic sine of the number of SA articles with innovation content on a firm in a year (Source: Seeking Alpha platform). |
| IHS (NonInnoArticleNum) | The inversed hyperbolic sine of the number of SA articles without innovation content on a firm in a year (Source: Seeking Alpha platform). |
| IHS (AvgComtNum) | The inversed hyperbolic sine of the average number of comments (excluding those made by authors) per article related to a firm in a year (Source: Seeking Alpha platform). |
| IHS (MaxComtNum) | The inversed hyperbolic sine of the maximum number of comments (excluding those made by authors) per article related to a firm in a year (Source: Seeking Alpha platform). |
| IHS (PatentNum) | The inversed hyperbolic sine of the number of patents a firm has applied for (and later granted) in a year. |
| IHS (CitedNum) | The inversed hyperbolic sine of the number of forward citations received by the patents that a firm has applied for (and later granted) in a year. |
| | |
| *Panel B. Control variables* | |
| FirmSize | The natural logarithm of book value of total assets in millions (U.S. dollars) (Source: Compustat). |
| FirmAge | Firm age in years (Source: Compustat). |
| Leverage | Sum of current and long-term debt divided by market value of total asset (Source: Compustat). |

| | |
|---|---|
| TobinQ | The natural logarithm of the sum of book value of total asset and market value of common equity minus book value of common equity, divided by book value of total asset (Source: Compustat). |
| ROA | Net income divided by book value of total assets (Source: Compustat). |
| R&D | Annual R&D expenses scaled by book value of total assets (Source: Compustat). |
| Tangibility | Property, plant and equipment expense scaled by book value total assets (Source: Compustat). |
| InstOwn | The percentage of shares of a firm owned by institutional investors in a year (Source: Thomson Reuters) |
| StockReturn | Last 12 months' stock returns (Source: CRSP). |
| StockVolatility | Annualized total stock return volatility (Source: CRSP). |
| #News | The number of news articles divided by 100 in a firm-year (Source: RavenPack). |
| NewsTone | The portion of positive news minus the portion of negative news (in percentage) in a firm-year (Source: RavenPack). |
| | |
| *Panel C. Proxy measures of information asymmetry, short-termism, and market competition* | |
| AnalystFollowing | The natural logarithm of one plus the number of analysts following a firm in a year (Source: IBES). |
| AbnormalAccrual | The median absolute value of discretionary accruals in the past five fiscal years where the discretionary accruals are estimated based on the modified Jones model in Dechow, Sloan and Sweeney (1995) (Source: Compustat). |
| IdiosyncraticVolatility | The standard deviation of the residuals from regressing daily individual stock returns on the Fama-French three-factors of a firm in a year (Source: CRSP) |
| EffectiveSpread | The bid-ask spreads estimated using daily high and low stock prices of a firm following Corwin and Schultz (2012). |
| TransientInstOwn | The percentage of shares of a firm owned by transient institutional investors in a year following Bushee (2001). |
| DedicatedInstOwn | The percentage of shares of a firm owned by dedicated institutional investors in a year following Bushee (2001). |
| CEOAge | The age of the CEO in a firm-year (Source: Execucomp). |
| CEOTenure | The tenure of the CEO in a firm-year (Source: Execucomp). |
| Fluidity | Text-based measure of competitive threats faced by a firm in the product market that captures changes in rival firms' product description in 10-Ks relative to a firm's (Source: Hoberg, Phillips and Prabhala (2014)). |
| TNICHHI | The sales-based Herfindahl-Hirschman index of a firm's industry based on text-based network industry classification (Source: Hoberg and Phillips (2016)). |
| HHI | The sales-based Herfindahl-Hirschman index of a firm's industry based on 2-digit SIC industry classification (Source: Compustat). |
| LernerIndex | The median gross margin of a firm's two-digit SIC industry (Source: Compustat). |

## Table A2. Reasons for SA Coverage Initiation

This table provides the coverage reason breakdown of the 175 non-innovation-related SA coverage initiation articles included in our final Difference-in-Differences regression analyses. We report the number of articles related to a certain reason (*# Articles*) and the percentage of articles related to a certain reason *(% Articles*), respectively. Because some articles may be related to multiple of these reasons (e.g., an article can write about business fundamentals, stock prices and stock returns of the covered firm), the sum of article numbers across these reason categories exceeds 175.

| **Reasons for SA Coverage Initiation** | **# Articles** | **% Articles** |
|---|---|---|
| 1: Business Fundamentals | 157 | 89.714% |
| 2: Stock Prices/Returns | 99 | 56.571% |
| 3: Earnings Announcements | 22 | 12.571% |
| 4: M&A/Joint Venture/Alliance/Spinoff | 16 | 9.143% |
| 5: Corporate Financing | 8 | 4.571% |
| 6: Corporate Payouts | 8 | 4.571% |
| 7: Corporate Governance | 10 | 5.714% |
| Total Number of Articles | 175 | |

## Table A3. Pairwise Correlation

This table presents pairwise correlations of the variables in our main analysis, with * indicating statistical significance at the 5% level or less. Definitions of the variables are provided in Appendix Table A1.

| | PatentNum | CitedNum | FirmSize | Leverage | TobinQ | ROA | R&D | Tangibility | InstOwn | AnalystFollowing | #News | NewsTone | StockReturn | StockVolatility |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PatentNum | 1 | | | | | | | | | | | | | |
| CitedNum | 0.848* | 1 | | | | | | | | | | | | |
| FirmSize | 0.486* | 0.475* | 1 | | | | | | | | | | | |
| Leverage | 0.127* | 0.094* | 0.426* | 1 | | | | | | | | | | |
| TobinQ | -0.019 | 0.012 | -0.086* | -0.148* | 1 | | | | | | | | | |
| ROA | 0.087* | 0.083* | 0.384* | -0.014 | 0.086* | 1 | | | | | | | | |
| R&D | 0.039* | 0.007 | -0.342* | -0.241* | 0.265* | -0.520* | 1 | | | | | | | |
| Tangibility | -0.045* | -0.090* | 0.150* | 0.147* | -0.171* | 0.090* | -0.224* | 1 | | | | | | |
| InstOwn | 0.015 | 0.079* | 0.258* | -0.029 | 0.125* | 0.302* | -0.128* | 0.071* | 1 | | | | | |
| AnalystFollowing | 0.236* | 0.261* | 0.466* | 0.039* | 0.199* | 0.321* | -0.043* | 0.038* | 0.534* | 1 | | | | |
| #News | 0.529* | 0.549* | 0.694* | 0.262* | 0.117* | 0.180* | -0.060* | -0.029 | 0.226* | 0.590* | 1 | | | |
| NewsTone | 0.095* | 0.084* | 0.171* | 0.151* | 0.023 | 0.080* | -0.058* | 0.008 | -0.115* | 0.016 | 0.093* | 1 | | |
| StockReturn | 0.001 | 0.026 | 0.076* | -0.011 | 0.374* | 0.233* | -0.064* | -0.009 | 0.103* | 0.070* | 0.057* | 0.089* | 1 | |
| StockVolatility | -0.115* | -0.120* | -0.482* | -0.095* | -0.01 | -0.350* | 0.224* | -0.031 | -0.310* | -0.276* | -0.334* | 0.011 | 0.143* | 1 |

## Table A4. Seeking Alpha Coverage Initiation and Alternative Innovation Measures

This table reports the results from difference-in-differences regressions of equation (1) for treatment and control firms three years before and after SA coverage initiation. The dependent variables $Efficiency$ is the number of patents a firm has applied (and eventually granted) in year *t* per million R&D dollars. $PatentValue$ is the economic values of patents a firm has applied (and eventually granted) in year *t* scaled by the book value of total assets of that firm-year. $EfficiencyValue$ is the economic values of patents a firm has applied (and eventually granted) in year *t* per million R&D dollars. $Originality$ is one minus the average Herfindahl index of backward citations of patents applied in a firm-year. $Generality$ is one minus the average Herfindahl index of forward citations of patents applied in a firm-year. $R\&D$ is the research and development expenses divided by the total asset of a firm in year $t$. $SG\&A$ is the selling, general and administration expenses divided by the total asset of a firm in year $t$. $Treat$ is an indicator variable that equals one if a cohort-firm is covered by SA (a treatment firm) and equals zero (a control firm) otherwise. $Post$ is an indicator variable that equals one for the post-event period (year [1, 3]) and equals zero for the pre-event period (year [-3, -1]). All specifications include cohort-firm and cohort-year fixed effects. Table A1 in the Appendix provides detailed variable definitions. *t*-statistics, calculated based on standard errors clustered at the cohort-firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Dependent Var. | $Efficiency_t$ | $PatentValue_t$ | $EfficiencyValue_t$ | $Originality_t$ | $Generality_t$ | $R\&D_t$ | $SG\&A_t$ |
| $Treat \times Post$ | 0.102*** | 0.055*** | 0.832*** | 0.201*** | 0.243*** | 0.025*** | 0.078*** |
| | (2.761) | (4.252) | (4.221) | (8.043) | (8.514) | (4.060) | (4.936) |
| Model | OLS | OLS | OLS | OLS | OLS | OLS | OLS |
| Cohort-Firm FE | YES | YES | YES | YES | YES | YES | YES |
| Cohort-Year FE | YES | YES | YES | YES | YES | YES | YES |
| Obs. | 4,315 | 4,543 | 4,315 | 4,543 | 4,543 | 4,543 | 4,543 |
| Adj. R2 | 0.525 | 0.626 | 0.668 | 0.382 | 0.405 | 0.581 | 0.603 |

## Table A5. Robustness Checks: Additional Controls

This table conducts two robustness tests. In panel A, we add additional control variables, including last 12 months' stock returns ($StockReturn$), return volatility ($StockVolatility$), analyst following ($AnalystFollowing$), number of news coverage ($\#News$), and percentage difference between positive and negative news ($NewsTone$). All additional controls are measured in year $t$-1. The dependent variables $PatentNum$ is the number of patents firm $i$ of cohort $c$ applied for (and eventually granted) in year $t$, and $CitedNum$ is the number of forward citations received by the patents that firm $i$ of cohort $c$ has applied for (and eventually granted) in year $t$. $Treat$ is an indicator variable that equals one if firm $i$ in cohort $c$ is covered by SA (a treatment firm) and equals zero otherwise. $Post$ is an indicator variable that equals one if year $t$ is in the post-event period (i.e., year [1, 3]) and equals zero if it is in the pre-event period (i.e., year [-3, -1]). Both specifications include cohort-firm, cohort-year fixed effects and other basic firm controls presented in the baseline models. Table A1 in the Appendix provides detailed variable definitions. $t$-statistics, calculated based on standard errors clustered at the cohort-firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

| | (1) | (2) |
|---|---|---|
| Dependent Var. | $PatentNum_t$ | $CitedNum_t$ |
| $Treat \times Post$ | 0.434*** | 0.811*** |
| | (3.603) | (7.614) |
| $StockReturn_{t-1}$ | 0.000 | -0.000 |
| | (0.185) | (0.146) |
| $StockVolatility_{t-1}$ | 0.005 | -0.003 |
| | (0.768) | (0.417) |
| $AnalystFollowing_{t-1}$ | 0.349*** | 0.480*** |
| | (3.345) | (3.967) |
| $\#News_{t-1}$ | 0.079*** | 0.083*** |
| | (3.302) | (3.448) |
| $NewsTone_{t-1}$ | 0.002 | 0.002 |
| | (1.444) | (1.032) |
| Model | POISSON | POISSON |
| Other Firm Controls | YES | YES |
| Cohort-Firm FE | YES | YES |
| Cohort-Year FE | YES | YES |
| Obs. | 2,773 | 2,616 |
| Pseudo R2 | 0.961 | 0.960 |

## Table A6. Seeking Alpha Article Intensity and Corporate Innovation: Two-stage Least Squares Regressions (Raw Number)

This table presents the two-stage least squares regression results to address endogeneity concerns. The instrument variable in the first stage is $Log\ (1 + FollowerNum)$, the natural logarithm of one plus the aggregated number of followers of all SA authors who publish articles on a firm in year *t*, which is measured by summing up the past follower numbers of the authors before they write on the firm. Panel A reports the results on total SA articles, where $ArticleNum$ is raw total number of SA articles related to a firm in year *t*. Panel B reports the results on innovation-focused SA articles, where $InnoArticleNum$ is the raw number of innovation-focused SA articles on a firm in year *t*. Panel C reports the results on non-innovation-focused SA articles, where $NonInnoArticleNum$ is the raw number of non-innovation-focused SA articles on a firm in year *t*. In both panels, the dependent variables in the second stage are innovation outputs measured in year *t+1*, including the raw number of patents a firm has applied for and later granted ($PatentNum$), and the raw number of forward citations received by the patents that a firm has applied for and later granted ($CitedNum$). Control variables are included as in Table 10 but omitted for brevity. All specifications include firm and year fixed effects. Table A1 in the Appendix provides detailed variable definitions. *t*-statistics, calculated based on standard errors clustered at the firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

***Panel A. Total SA Articles***

| Dependent Var. | (1)<br>$ArticleNum_{t}$ | (2)<br>$PatentNum_{t+1}$ | (3)<br>$CitedNum_{t+1}$ |
|---|---|---|---|
| $\text{Log}\ (1 + \text{FollowerNum})_{t}$ | 0.318*** | | |
| | (16.798) | | |
| $\text{ArticleNum}_{t}$ | | 1.872*** | 3.363*** |
| | | (3.046) | (3.949) |
| Firm Controls | YES | YES | YES |
| Model | OLS | OLS | OLS |
| Year FE | YES | YES | YES |
| Firm FE | YES | YES | YES |
| Obs. | 5,142 | 5,142 | 5,142 |
| Adj. R2 | 0.622 | 0.951 | 0.904 |

***Panel B. Innovation-focused SA Articles***

| Dependent Var. | (1)<br>$InnoArticleNum_{t}$ | (2)<br>$PatentNum_{t+1}$ | (3)<br>$CitedNum_{t+1}$ |
|---|---|---|---|
| $\text{Log}\ (1 + \text{FollowerNum})_{t}$ | 0.054*** | | |
| | (7.279) | | |
| $\text{InnoArticleNum}_{t}$ | | 5.971** | 9.871*** |
| | | (2.516) | (3.084) |
| Model | OLS | OLS | OLS |
| Firm Controls | YES | YES | YES |
| Year FE | YES | YES | YES |
| Firm FE | YES | YES | YES |
| Obs. | 5,142 | 5,142 | 5,142 |
| Adj. R2 | 0.611 | 0.951 | 0.903 |

***Panel C. Non-innovation-focused SA Articles***

| Dependent Var. | (1)<br>$NonInnoArticleNum_{t}$ | (2)<br>$PatentNum_{t+1}$ | (3)<br>$CitedNum_{t+1}$ |
|---|---|---|---|
| $\text{Log}\ (1 + \text{FollowerNum})_{t}$ | 0.258*** | | |
| | (8.828) | | |
| $\text{NonInnoArticleNum}_{t}$ | | 1.246** | 2.060*** |
| | | (2.456) | (2.991) |
| Model | OLS | OLS | OLS |
| Firm Controls | YES | YES | YES |
| Year FE | YES | YES | YES |
| Firm FE | YES | YES | YES |
| Obs. | 5,142 | 5,142 | 5,142 |
| Adj. R2 | 0.736 | 0.951 | 0.903 |

## Table A7. Robustness Check: Total Seeking Alpha Articles Instrumented by the Aggregated Number of Followers of Unique Authors on Non-innovation-focused Articles

This table presents the two-stage least squares regression results to address endogeneity concerns. The instrument variable in the first stage is $Log\ (1 + FollowerNum)$, the natural logarithm of one plus the aggregated number of followers of all SA authors who publish non-innovation-focused articles on a firm in year $t$. Panel A reports the results where the first-stage and second-stage dependent variables are transformed with inversed hyperbolic sine: $IHS\ (ArticleNum)$ is the inversed hyperbolic sine of the total number of SA articles related to a firm in year $t$; $IHS\ (PatentNum)$ is the inversed hyperbolic sine of the number of patents a firm has applied for in year $t+1$; $IHS\ (CitedNum)$ is the inversed hyperbolic sine of the number of forward citations received by the patents that a firm has applied for in year $t+1$. Panel B reports the results where the first-stage and second-stage dependent variables are raw number: $ArticleNum$ is the raw total number of SA articles related to a firm in year $t$; $PatentNum$ is the raw number of patents a firm has applied for in year $t+1$; $CitedNum$ is the raw number of forward citations received by the patents that a firm has applied for in year $t+1$. Control variables are included as in Table 10 but omitted for brevity. All specifications include firm and year fixed effects. Table A1 in the Appendix provides detailed variable definitions. $t$-statistics, calculated based on standard errors clustered at the firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

***Panel A. Inversed Hyperbolic Sine***

| | First Stage | Second Stage | |
|---|---|---|---|
| Dependent Var. | $IHS\ (ArticleNum)_t$ | $IHS\ (PatentNum)_{t+1}$ | $IHS\ (CitedNum)_{t+1}$ |
| Log (1 + FollowerNum) $_t$ | 0.083*** | | |
| | (21.542) | | |
| IHS (ArticleNum) $_t$ | | 0.110** | 0.104* |
| | | (2.252) | (1.811) |
| Model | OLS | OLS | OLS |
| Firm Controls | YES | YES | YES |
| Year FE | YES | YES | YES |
| Firm FE | YES | YES | YES |
| Obs. | 5,142 | 5,142 | 5,142 |
| Adj. R2 | 0.623 | 0.942 | 0.919 |

***Panel B. Raw Number***

| | First Stage | Second Stage | |
|---|---|---|---|
| Dependent Var. | $ArticleNum_t$ | $PatentNum_{t+1}$ | $CitedNum_{t+1}$ |
| Log (1 + FollowerNum) $_t$ | 0.246*** | | |
| | (17.949) | | |
| ArticleNum $_t$ | | 1.251** | 2.165*** |
| | | (2.561) | (3.350) |
| Model | OLS | OLS | OLS |
| Firm Controls | YES | YES | YES |
| Year FE | YES | YES | YES |
| Firm FE | YES | YES | YES |
| Obs. | 5,142 | 5,142 | 5,142 |
| Adj. R2 | 0.621 | 0.951 | 0.903 |

## Table A8. Seeking Alpha Article Comments and Corporate Innovation: Two-stage Least Squares Regressions (Raw Number)

This table presents the two-stage least squares regression results to address endogeneity concerns. The instrument variable in the first stage is $Log\ (1 + FollowerNum)$, the natural logarithm of one plus the aggregated number of followers of all SA authors who publish articles on a firm in year *t*, which is measured by summing up the past follower numbers of the authors before they write on the firm. Panel A reports the results on average number of comments per SA article on a firm, where $AvgComtNum$ is the raw average number of comments (excluding those made by authors) per article related to a firm in year *t*. Panel B reports the results on maximum number of comments per SA article on a firm, where $MaxComtNum$ is the raw maximum number of comments (excluding those made by authors) pre article related to a firm in year *t*. In both panels, the dependent variables in the second stage are innovation outputs measured in year *t+1*, including the raw number of patents a firm has applied for and later granted ($PatentNum$), and the raw number of forward citations received by the patents that a firm has applied for and later granted ($CitedNum$). Control variables are included as in Table 11 but omitted for brevity. All specifications include firm and year fixed effects. Table A1 in the Appendix provides detailed variable definitions. *t*-statistics, calculated based on standard errors clustered at the firm level, are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

***Panel A. Average Number of Comments per SA article on a Firm***

| | First Stage | Second Stage | |
|---|---|---|---|
| Dependent Var. | $AvgComtNum_{t}$ | $PatentNum_{t+1}$ | $CitedNum_{t+1}$ |
| $\text{Log}\ (1 + \text{FollowerNum})_{t}$ | 0.014** | | |
| | (2.527) | | |
| $\text{AvgComtNum}_{t}$ | | 41.794** | 75.097** |
| | | (1.962) | (2.151) |
| Firm Controls | YES | YES | YES |
| Model | OLS | OLS | OLS |
| Year FE | YES | YES | YES |
| Firm FE | YES | YES | YES |
| Obs. | 5,142 | 5,142 | 5,142 |
| Adj. R2 | 0.159 | 0.951 | 0.904 |

***Panel B. Maximum Number of Comments per SA Article on a Firm***

| | First Stage | Second Stage | |
|---|---|---|---|
| Dependent Var. | $MaxComtNum_{t}$ | $PatentNum_{t+1}$ | $CitedNum_{t+1}$ |
| $\text{Log}\ (1 + \text{FollowerNum})_{t}$ | 0.073*** | | |
| | (9.374) | | |
| $\text{MaxComtNum}_{t}$ | | 8.133*** | 14.614*** |
| | | (2.951) | (3.725) |
| Model | OLS | OLS | OLS |
| Firm Controls | YES | YES | YES |
| Year FE | YES | YES | YES |
| Firm FE | YES | YES | YES |
| Obs. | 5,142 | 5,142 | 5,142 |
| Adj. R2 | 0.378 | 0.951 | 0.904 |