Social Media Analysts, Managerial Learning, and Corporate Innovation *

Jeroen Koenraadt †

Tim Martens ‡

Christoph J. Sextroh §

May 2024

Abstract

We study the role of non-traditional investment research as a source of information for managerial learning and corporate investment decisions. Using a comprehensive sample of social media analyst reports from Seeking Alpha and exogenous variation in social media analysts' coverage overlaps, we show that firms are more likely to invest into technologies similar to firms covered by the same analyst. The effect is incremental to coverage by professional sell-side analysts and varies with social media analysts' characteristics and differences in their contributed content that capture their unique information set. Overall, our results are consistent with non-traditional investment research enhancing firms' information environment as an additional source of information that may guide corporate investment decisions.

Keywords: Social media analysts; Seeking Alpha; Information intermediaries; Managerial learning; Information spillover; Corporate innovation; Patents

*We thank Oliver Binz, Frank Ecker, Charles Lee, Dawn Matsumoto, Phillip Quinn, Harm Schuett and Andrew Yim, as well as seminar participants at Bayes Business School, Bocconi University, Frankfurt School of Finance & Management, Georgetown University, Humboldt University Berlin, Johannes Gutenberg University Mainz, Stockholm School of Economics, University of Göttingen, University of Neuchâtel, University of Washington, University of Zurich, WHU Otto Beisheim School of Management, and the 2023 EASYS and the 2024 HARC for helpful comments and discussions. An earlier version of this paper was titled "*Do Firms Listen to Social Media Analysts? Evidence from Seeking Alpha*".

†London School of Economics and Political Science. Email: j.k.koenraadt@lse.ac.uk

‡Bocconi University. Email: tim.martens@unibocconi.it

§Carl von Ossietzky University of Oldenburg and Tilburg University. Email: christoph.sextroh@uni-oldenburg.de

1 Introduction

Social media have led to a significant increase in the quantity and accessibility of crowd-sourced non-traditional investment research. Social media platforms such as Seeking Alpha, StockTwits, and Estimize have emerged as large communities of individuals who share their thoughts on and expectations of different firms and industries (e.g., Cookson et al., 2024b). Evidence indicating that stock research published by these so-called "social media analysts" (SMAs) provides incremental information that affects market pricing is increasing (e.g., Bartov et al., 2018; Chen et al., 2014; Farrell et al., 2022; Jame et al., 2016). At the same time, anecdotal evidence also suggests that firms may increasingly review crowd-sourced information when seeking strategic business insights (e.g., Boudreau and Lakhani, 2013; DeWalt, 2021). Despite evidence on the informational role of social media analysts for capital markets, little is known about the specific role of information published by social media analysts in firms' decision making. We investigate this question in the context of corporate investment into innovation using a comprehensive sample of social media analyst reports from Seeking Alpha, a popular platform of crowd-sourced financial commentary.

It is well known that managers can extract valuable information for real decision making by observing market prices and interacting with financial intermediaries (e.g., Edmans et al., 2012). By monitoring crowd-sourced investment platforms and reading reports published by SMAs, firms may be able to gain insights into how investors perceive their brands, industry trends, competitor activities, and other potential risks and opportunities. Thus, firms may benefit from using crowd-sourced investment research for strategic decision making. At the same time, social media commentary typically includes considerable noise, and information is often difficult to verify or lacks credibility (Elliott et al., 2018; Kogan et al., 2021). In addition, whether research published by SMAs is incrementally relevant from a managerial learning perspective remains unclear, given that firms may have alternative and more valuable sources of competitive information. Firms trade off the relative costs and benefits and may thus be unaware of the existence of potentially useful crowd-sourced information or may be unwilling to spend the resources to extract it.

We explore the idea that firms learn from commentary published by SMAs by examining information spillovers between firms covered by the same SMA on the popular crowd-sourced investment platform Seeking Alpha. Prior studies generally find that commentary published on Seeking Alpha has incremental information content (e.g., Campbell et al., 2019; Chen et al., 2014; Drake et al., 2022). Anecdotal evidence suggests that content published on Seeking Alpha also contains information that may be relevant for corporate decision making. For example, Seeking Alpha articles not only include earnings news and investment analyses, but may also contain detailed information about a firms' patents and technologies (see Appendix A). Similarly, media discussion suggest that corporate employees may regularly follow the content published by SMAs.[1] Some SMAs even claim that they interact with corporate management and some managers are active as SMAs themselves.[2] Seeking Alpha thus provides a strong setting in which to test whether commentary published by SMAs is considered in corporate decision making.

Measuring managerial learning from information is challenging. It requires not only the identification of a flow of potentially relevant information, but also the ability to track subsequent managerial decision making back to this specific information. To address this issue, our research design exploits variation in SMA coverage overlaps and variations in interfirm patent citations over time (Gomes-Casseres et al., 2006). The general idea is that if crowd-based investment research provides useful information, firms covered by the same SMA should be more likely to learn about one anothers' activities and incorporate that information into their decision making. Thus, managerial learning should become visible as information spillovers between firms with shared SMA coverage. As a proxy for managerial learning, we rely on patent citations, which allows us to track information spillovers between

[1] https://www.vanityfair.com/news/2018/07/elon-musk-desperately-needs-a-hobby

[2] See also Campbell et al. (2019) for a related discussion on corporate attention towards content published on Seeking Alpha.

firms over time. It also allows us to construct our research design in a way that it explicitly rules out other factors that could possibly explain coverage overlaps and interfirm patent citations.

We first verify that non-traditional investment research from crowd-sourced investment platforms is a plausible source of additional information. Coverage portfolios of SMAs are indeed much more diverse and less focused on specific industries than those of traditional sell-side analysts. Also, the network of firms emerging from shared SMA coverage is not only larger but also fundamentally different than the coverage network emerging from traditional sell-side analyst coverage alone. Therefore, similar to prior studies that document incremental information content of user-generated content on social media platforms for capital markets (e.g., Antweiler and Frank, 2004; Chen et al., 2014; Hales et al., 2018; Huang et al., 2020; Jame et al., 2016), it is plausible that crowd-based investment research provides a distinct set of competitive information that may be incrementally informative for managerial learning and decision making.

To test whether firms extract information from SMA investment research, we exploit the directional nature of patent citations. Specifically, we estimate the relation between shared SMA coverage and patent citations for a time series of directional firm pairs. This structure enables us to incorporate directional fixed effects for each firm pair and separate time trends for both citing and cited firms. Put differently, identification rests only on variation within and across firm pairs over time. Alternative explanations related to structural differences in citation behavior between firm pairs, differences between individual firms, or changes within individual firms over time are subsumed by the fixed effect structure.

Our results suggest that firms with shared SMA coverage are indeed more likely to pursue similar follow-up investments as indicated by a higher likelihood of subsequent cross-citations than that of firm pairs without shared SMA coverage. The effect is incremental to common coverage by traditional sell-side analysts and robust to overlaps of other potential information providers (e.g., investors, inventors, board members, auditors). We also control for changes in business and technological similarity of firm pairs over time (Hoberg and Phillips, 2010, 2016; Jaffe and Trajtenberg, 1996). Hence, it is unlikely that our results can be attributed merely to changes in firms' business similarity and technological focus (affecting both analysts' coverage decisions and firms investment behavior) (see, e.g., Ali and Hirshleifer, 2020; Lee et al., 2024, 2019).

To address any remaining concerns about correlated omitted variables and reverse causality, we use exogenous variation in Seeking Alpha analyst coverage overlaps as an additional approach to identification. Specifically, we exploit a change in Seeking Alpha's payment system that affected analysts' incentive to cover small cap firms and, accordingly, also coverage overlaps between firms. Using the change as an instrument to implement a fuzzy difference-in-difference regression design including Lewbel instruments to further strengthen identification (Baum and Lewbel, 2019; Lewbel, 2012), we continue to find a significant positive relation between instrumented SMA coverage overlaps and interfirm patent citations.

We next explore the sources of managerial learning from social media analysts' investment research. On the one hand, the documented effect may simply emerge from the fact that firms are followed by 'more' information intermediaries which create a broader information environment for managerial learning. On the other hand, the nature of user-generated non-traditional investment research suggests that social media analysts provide a distinct set of competitive information that may be incrementally informative for managerial learning and decision making. To shed light on this question, we analyse social media analysts' coverage portfolio characteristics, differences in their contributed content, as well as their credibility and visibility.

Our results suggest that the value of SMAs' investment research may primarily emerge from incorporating a broader set of information and not necessarily from their experience or ability to process information. Effects are stronger for coverage overlaps of SMAs that do not specialize in specific industries, have larger coverage portfolios, focus on specific topics, provide more unique content, and focus on technology rather than earnings- or trading-related content. These results suggest that the value of SMAs may not simply emerge from an 'increased coverage'-effect due to more information intermediaries covering the firm but rather originates from access to a more diverse group of individuals who potentially collect and provide information that is not readily available from interactions with other information intermediaries.

To further contrast alternative explanations and address any remaining concern that the observed results merely emerge from correlated omitted variables, we replicate our main analysis using analyst coverage on *Estimize.com*, a platform which primarily collects and disseminates crowd-sourced earnings and revenue forecasts, but does not publish any accompanying investment analyses or commentary. We find that coverage overlaps emerging from Estimize fail to explain patent cross-citations across firm pairs. This further supports a plausibly causal interpretation that SMAs' investment research and commentary published on Seeking Alpha provides useful information for managerial decision making.

Finally, we investigate whether credibility and visibility of social media analysts affect firms' willingness or ability to extract information from social media analysts' investment research. We find results consistent with the notion that the credibility of the contributed content as well as firms' processing costs may indeed play a role for whether information provided by social media analysts is considered in managerial decision making.

Taken together, our results suggest that non-traditional investment research published on social media platforms not only provides information that is valuable for investors but also may enhance the set of information available to firms when making strategic business decisions. This effect does not seem to simply emerge from a general enhancement of firms' information environments through increased coverage by intermediaries or aggregate sentiment and attention effects. Instead our results indicate that the benefit of social media analysts' investment research for managerial learning also originates from access to a more diverse group of individuals who potentially collect and provide information that is not readily available other information sources and intermediaries, such as traditional sell-side analysts.

Our paper contributes to the literature on the role of social media in general and non-traditional investment commentary specifically in financial markets and firms' decision-making (see, e.g., Cookson et al., 2024b, for a recent review). A growing body of literature documents that user-generated content on social media platforms can include valuable information for capital markets despite concerns over quality and credibility (Antweiler and Frank, 2004; Bartov et al., 2018; Chen et al., 2014; Farrell et al., 2022; Gomez et al., 2022; Jame et al., 2016, 2022). Recent work begins to look into whether information disseminated through social media can also be a source of information for firms (e.g., Cookson et al., 2024c; He et al., 2024). While the literature so far has looked predominately into aggregate signals (e.g., sentiment) or the effect of increased coverage by intermediaries, little is known about the specific informational value of social media analysts and their investment commentary for managerial learning. Even if managers monitor social media platforms, it is not clear whether learning primarily emerges from the 'wisdom of the crowds' and aggregate signals about market trends and developments or whether also individual analysts' commentary includes valuable information for corporate decision making. This study documents the role of non-traditional investment research as a source of unique information for managerial decision making.

2 Social media analysts and managerial learning

2.1 The role of social media analysts in capital markets

User-generated online content has become increasingly prevalent in capital markets. Today's information environment is shaped by investors sharing their thoughts and opinions on companies, industries, and markets through general social media platforms such as Reddit, Twitter, or YouTube. Specialized investment platforms, such as Seeking Alpha, StockTwits, or Estimize, have emerged as large communities in which users regularly publish and exchange investment advice, company analyses, and expectations about the future (e.g., Blankespoor et al., 2020).

The increased quantity and accessibility of user-generated content through social media has two important implications for capital markets. First, the broad spectrum of (retail) investor beliefs and behavior becomes more visible and ultimately quantifiable. For example, investor discussions on social media platforms can be used to better understand investor disagreement (Cookson and Niessner, 2019). Tracking discussions, online information searches, and website visits allow for the identification of investors' information demands and attention (e.g., Antweiler and Frank, 2004; Drake et al., 2012, 2015; Lerman, 2020). Monitoring user commentary on stock message boards or social media can be used to capture investor sentiment or extract real-time trading signals (Cookson et al., 2024a). Furthermore — and equally important — user-generated content provides incremental information about the firm that is not readily available from other sources (e.g., Chen et al., 2014; Hales et al., 2018; Huang et al., 2020; Jame et al., 2016). As such, user-generated content on social media platforms not only allows to monitor the beliefs of market participants, but also enriches the corporate information environment as a source of additional firm-specific information.

Important contributors in this regard are social media analysts (SMAs). These users regularly publish firm-specific commentary and analyses on social media platforms, such as Seeking Alpha and StockTwits, that specialize in crowd-sourcing information about firms and investment advice from a broad range of users. Similar to traditional sell-side analysts, SMAs summarize, analyze, and disseminate existing information but may also synthesize and uncover new pieces of information.

Prior research on SMAs finds that they generally focus on what they find interesting and that their coverage decisions are much more independent compared to those of traditional sell-side analysts. While community building and reciprocity are presumably important factors to be active on social media platforms, SMAs also respond to monetary incentives (Chen et al., 2019; Gu et al., 2023; Koenraadt, 2023). Seeking Alpha, for example, provides monetary rewards for contributors based on a per-page-view model.[3] As a result, SMAs are more likely to cover firms with high retail ownership or high ESG ratings for which they can generate relatively more page views from retail-oriented Seeking Alpha readership (Chen et al., 2021; DeAngelis et al., 2021). Similarly, SMAs are more likely to cover firms in which they are personally invested (Campbell et al., 2019).

A fast-growing body of literature documents the benefits of SMAs for capital markets. SMA coverage is associated with improved capital market outcomes (e.g., Bartov et al., 2018; Chen et al., 2014; Farrell et al., 2022; Jame et al., 2016). SMAs have also been shown to make value-relevant earnings forecasts, improve responses to earnings announcements, and level the playing field among investors (e.g., Antweiler and Frank, 2004; Farrell et al., 2022; Gomez et al., 2022; Jame et al., 2016, 2022). However, concerns over the quality, credibility and anonymity of SMAs can overshadow their benefits (e.g., Campbell et al., 2019; Clarke et al., 2021; Dyer and Kim, 2021; Kogan et al., 2021; Mitts, 2020). SMAs are generally anonymous, but to varying degrees. Some share a real name with reference to official online profiles, others only include a generic user name. Commentary published by non-anonymous SMAs typically results in stronger market reactions, although anonymous SMAs seem to be are able to build reputation and credibility over time (Dyer and Kim, 2021). There is also concern that SMAs may even harm firms' information environment because they are less-informed due to not having access to the same sources of information as regular analysts and because they are less capable of extracting relevant information from the sources they do have available to them (Drake et al., 2017).

[3]See https://seekingalpha.com/article/2134803-how-much-does-seeking-alpha-pay-its-contributors for more information on how Seeking Alpha rewards contributors.

The potential informational value of SMAs' commentary originates from two distinct aspects. For one, a larger group of individuals who each provide information allows for the extraction of more precise and less biased aggregate signals compared to the opinions put forward by a few experts (e.g., Jame et al., 2016; Surowiecki, 2004). This 'wisdom-of-the-crowds' effect emphasizes the informational value of aggregate social media signals. At the same time, social media content originates from a diverse group of individuals. SMAs may thus also provide new information to the public that is not readily available from other sources. (e.g., Chen et al., 2014; Jame et al., 2016). While most of the literature focuses on aggregate signals from social media, few studies have zoomed into the specific informational value of investment commentary provided by SMAs. Koenraadt (2023), for example, finds that expertise and high quality commentary of SMAs can improve even weak information environments. This suggests that non-traditional investment research has value beyond capturing investor sentiment and drawing attention.

In this paper, we focus on investment research published on Seeking Alpha, the largest investment-related website in the U.S., with 17 (210) million monthly visitors (visits), 10 million registered users, and $17,000$ contributors as of 2021. Seeking Alpha is a platform on which user-contributed articles with investment analyses and commentary can be published, discussed, and read. Content published on Seeking Alpha follows an editorial process that ensures a certain quality level. Editors review submitted articles, decide whether to reject or to accept them, and may provide the author with suggestions for improving writing and structure. Moreover, statistics from Seeking Alpha suggest that roughly 5.5% of its contributors are company executives and C-suite managers themselves (see Campbell et al., 2019, Table 2, Panel A) making it a valuable source for business insights.

Prior research generally finds that articles published on Seeking Alpha have information content. Chen et al. (2014), for example, find that the negative tone of commentaries posted on Seeking Alpha predicts future abnormal returns and earnings surprises. Campbell et al. (2019) find short-window price responses to articles published on Seeking Alpha and that the stock positions of authors increase investors' perception of their articles' credibility. More recently, Drake et al. (2022) suggest that Seeking Alpha articles preempt information in sell-side analyst reports and contemporaneous work by Farrell et al. (2022) finds more profitable retail trading around Seeking Alpha articles. Overall, prior research suggests that Seeking Alpha is potentially an important information intermediary and that the content provided by SMAs may be a distinct and timely source of information.

While this is also true for other crowd-sourced investment platforms, Seeking Alpha content provides a more powerful setting for our analysis. First, quantitative platforms such as Estimize primarily collect and disseminate quantitative forecast information but provide no investment analysis or commentary. In addition, stock message boards, such as Twitter, Yahoo Finance, or StockTwits, allow for posting content without any quality control. As such, if firms incorporate information from social media analysts into their decision making, we should more likely observe this for reviewed investment-oriented content published on Seeking Alpha than content published on most other platforms. Taken together, Seeking Alpha provides a strong setting in which to test whether firms consider investment research published by SMAs in their decision making.

2.2 Managerial learning from social media analysts

It is well established that managers may gather useful information for their decision making from monitoring stock markets, interacting with information intermediaries, or consulting the traditional news media (e.g., Bae et al., 2022; Bond et al., 2012; Edmans et al., 2012). However, it is less clear whether managers also rely on information contained in user-generated investment commentary and analyses published on social media platforms.

Recent evidence suggests that firm management regularly engages in social media monitoring — not only to track hashtags, keywords, and mentions relevant to monitor consumer feedback, but also to engage with investors and obtain critical information about employees, competitors, and industry developments (e.g., Cision, 2017; Dube and Zhu, 2021; Flam et al., 2023). Social media investment websites give many firms the opportunity to track the discourse on their firm, and social media analysts the opportunity to engage with management.[4] Cookson et al. (2024c) are among the first to analyze whether managers actually learn from stock-related talk on social media. They investigate whether social media sentiment can predict merger withdrawals, an important type of corporate investment decision, and find that a decrease in abnormal sentiment on StockTwits increases the likelihood of a merger withdrawal. While this finding provides early evidence that investment-related talk on social media can shape corporate investment decisions, whether firms can also benefit from specific commentary and analyses provided by SMAs, in addition to sentiment, remains unclear.

On the one hand, if investment commentary and analyses published on investment platforms such as Seeking Alpha provide incrementally useful information to market participants, these platforms might also contain relevant information for firms' own decision making. First, if firms already obtain feedback from markets and professional information providers, monitoring SMAs may be a simple addition to this information set. In addition, user-generated content provides incremental information about the firm that is not readily available from other sources and as such may be a valuable source of additional firm-specific information. SMAs typically have a different coverage portfolio compared to other traditional information intermediaries (see, e.g., Jame et al., 2016, Table 2). For instance, SMAs are more likely to cover firms not covered by traditional sell-side analysts (Koenraadt, 2023) and smaller firms with high retail ownership (Chen et al., 2021). SMAs also make more conscious coverage decisions, which results in more diverse, albeit smaller coverage portfolios (Koenraadt, 2023).

Anecdotal evidence indeed suggests that articles published on Seeking Alpha attract corporate attention and that some analysts even interact with corporate management (see also Campbell et al., 2019). A user with the name *Where is the Yield?*, for example, posted an

[4]See https://blog.stocktwits.com/stocktwits-for-investor-relations-claim-your-ticker-today-d17fdba78eb6 and https://seekingalpha.com/article/2738495-seeking-alpha-expands-relationship-with-nasdaq.

article about a comment she received from an employee of a particular firm covered.[5] Other SMAs state that they met with management to discuss business developments.[6] While these examples illustrate that it is plausible that firms monitor social media platforms, we do not imply that direct monitoring is the only mechanism of how firms may extract useful information from crowd-sourced investment research. In fact, there are presumably several other direct and indirect mechanisms, such as social media monitoring services, that may allow firms to learn from crowd-sourced information.

On the other hand, firms may not utilize information provided by SMAs if they are not aware of it or doubt its credibility. In contrast to investors and professional analysts, firms typically do not regularly interact with SMAs and the cost of regularly monitoring the information provided by them may be considered too high (see, e.g., Blankespoor et al., 2020, for a detailed review of the role of awareness, acquisition, and integration costs). At the same time, firms may not be willing to obtain and utilize information from social media investment platforms if they doubt their credibility and informational value (e.g., Elliott et al., 2018; Kogan et al., 2021).

We investigate whether managers learn from commentary and analyses published on social media investment platforms by examining the link between SMA coverage overlaps and interfirm information flows.

3 Data and research design

3.1 Research design

Our research design follows Gomes-Casseres et al. (2006) and exploits the directional nature of patent citations, which enables us to incorporate a comprehensive fixed-effect structure to account for variations among individual firms and structural differences in citation

[5] See https://seekingalpha.com/article/22970-powershares-preferred-shares-etf-the-yield-just-isnt-there.

[6] See https://seekingalpha.com/article/3296885-qihoo-hardware-software-integration-a-challenge-mobile-search-ramping-up or https://seekingalpha.com/article/88904-is-soundbite-communications-back-on-track.

behavior within and between firm pairs over time. Due to the count-based and overdispersed nature of the outcome variable, we use a negative binomial regression model[7] as our main specification to estimate the relationship between shared analyst coverage and the number of citations within firm pairs:

$$\begin{aligned} Citations_{i,j,t} &= \beta_0 + \beta_1 Common\ SMAs_{i,j,t} + \sum Controls_{i,j,t} \\ &+ \sum Firm\ pair_{i,j} + \sum Citing\ Firm_i\ x\ Year_t + \sum Cited\ Firm_j\ x\ Year_t + \epsilon \qquad (1) \end{aligned}$$

where *Citations* is the number of citations from citing firm i to cited firm j in year t, and *Common SMAs* is the number of common SMAs of citing firm i and cited firm j in year t.

The model incorporates *directional* fixed effects for each firm pair (i.e., takes into account which firm in the pair is the citing and the cited firm) and separate time trends for both the citing and cited firms. This fixed effect structure allows us to effectively control for alternative explanations related to (1) differences between firm pairs, (2) differences between individual firms within pairs, and (3) changes within firms over time. For instance, firms with similar business models might be more likely to cite each other's patents, whereas SMAs might be more likely to cover firms operating in related or similar industries. Firm pair fixed effects address these structural similarities and differences between citing and cited firms across different firm pairs. Additionally, by including directional fixed effects, we explicitly consider potential supply-chain effects, i.e., within-firm-pair variations in citation flows that depend on whether a firm is considered the citing or cited firm.

To mitigate any potential concerns over identification arising from changes within firm pairs over time, we incorporate a set of time-varying firm-pair-specific controls into the regression model. These controls help explain any variations in the level of cross-citations within firm pairs over time and further strengthen the validity of our results.[8] First, we

[7]See, e.g., Cohn et al. (2022) for a discussion on different econometric approaches when working with count-based outcome variables.

[8]Although several other factors, such as patent examiners and citation norms, could account for cross-

include variables that serve as controls for other potential sources of industry or technological information spillovers at the firm pair level. These variables may be correlated with the variation in analyst coverage overlaps and information spillovers. Examples of such variables include common traditional sell-side analysts (Martens and Sextroh, 2021), strategic alliances (Gomes-Casseres et al., 2006), transfers of inventors between firms (Zacchia, 2020), and common investors (Reuer and Devarakonda, 2017).

In addition, we control for changes in firm pairs' business similarity and technological similarity (Hoberg and Phillips, 2010, 2016; Jaffe and Trajtenberg, 1996). In particular, technological similarity has been shown to be a significant factor that determines cross-firm patent citations and is expected to fluctuate over time (Jaffe and Trajtenberg, 1996). To address remaining concerns over changes within firm pairs over time, we also incorporate variables to capture firms' relative sizes, traditional sell-side analyst following, SMA following, patent stock, and citations for the citing and cited firms. These variables help account for potential changes in the relative information environment, investments, and innovative activities within firm pairs over time. Thus, our results unlikely merely reflect the relation between SMA coverage decisions and changes in firms' technological links and production complimentaries (Lee et al., 2024, 2019). Appendix B includes a list of all variable definitions. To account for extreme observations, we winsorize all continuous control variables at the top and bottom 1 percent. To control for potential correlations among the residuals, we calculate three-way clustered standard errors by firm pair, citing firm × year and cited firm × year (Abadie et al., 2023).

3.2 Patent citations as a measure of interfirm information spillovers

To measure the flow of (technological) knowledge between firms, we adopt an established approach from the fields of economics and finance by utilizing patent citations as a measure

citations of patents, our fixed-effect structure encompasses most of them. Furthermore, these broader factors are unlikely to have a systematic connection with SMA coverage overlaps.

of interfirm information spillovers (e.g., Agrawal et al., 2017; Belenzon and Schankerman, 2013; Gomes-Casseres et al., 2006; Jaffe et al., 1993).[9] Specifically, we use the number of patent citations from firm j (the cited firm) included in patents filed by firm i (the citing firm) in year t. Consistent with prior research, we interpret changes in citation flows as indicative of the underlying transfer of (technological) knowledge from the firm being cited to the firm making the citation.[10]

Using patent citations to determine knowledge flows between firms has several advantages. Most notably, the measure allows us to capture knowledge transfers between clearly defined firms, enabling us to compare knowledge flows across different firm pairs. Because relevant citations emerge from a distinct patent application, we can further exploit time-series variation in citation flows. Thus, we can consider not only the characteristics of firms and their patenting, including the total number of citations and the total stock of potentially citable patents in a given year, but also any idiosyncratic characteristics of each distinct firm pair. Other measures of knowledge flows either lack precise links between distinct firms or relate to one-off occurrences. Patent citations, on the other hand, serve as a well-established proxy for knowledge transfers between distinct firms and can plausibly identify the effects of overlaps in analyst coverage on interfirm information spillovers.[11]

To identify cross-citations, we use patent information from the Kogan et al. (2017) patent database, which records patent information for all patents issued in the United States

[9] See also Jaffe and Trajtenberg (2002) for a review of the literature that uses patent citations as a proxy for knowledge flows.

[10] Our analyses do not require a *direct* measurement of technology-related information flows through patent citations. Whether patent citations instead serve as an *indirect* gauge of market and competition-related information transfers between firms is of secondary importance. Regardless, these citations represent external information that firms incorporate into their innovations and are indicative of the inflow of knowledge toward the firm.

[11] One issue with using patent citations to measure knowledge flows is that their relevance can depend on the strategic objectives of the citing firm. Furthermore, patent citations may not always capture relevant knowledge flows due to factors such as firms not patenting all of their inventions or omitting citations for strategic reasons (Lampe, 2012). Despite these limitations, we believe that patent citations are a meaningful measure of interfirm knowledge flows for the types of relationships we are studying. While there may be some noise in the data due to strategic or extraneous citations, we do not expect these factors to systematically bias our results (see also Gomes-Casseres et al., 2006, for a related discussion).

between 1926 and 2020. Selecting the relevant patent pool and directional firm pairs is a critical decision when using cross-citations to measure interfirm information flows. Using all directional firm pairs that held a patent at any time results in many directional firm pairs with zero cross-citations, and not all patents in the patent pool are relevant for firms' decision making. To address these issues, we apply several sample selection criteria to construct the relevant patent pool. First, we only include a particular firm pair if the citing firm cites a patent of the cited firm at least once from 1926 through 2020. Second, we exclude firm-pair observations if the cited firm has no patents until and including year t. Third, we exclude firm-pair observations if the citing firm has not filed a patent application with the United States Patent and Trademark Office (USPTO) in year t (i.e., the firm will by definition not cite any patents in that year). These sample selection criteria ensure that the cited firm has relevant innovative capital and that the citing firm shows innovation. Finally, we exclude self-citations (i.e., a firm citing its own patents) from all firm pairs. Please note negative binomial regression requires excluding any group for which the outcome variable is zero for all observations because these groups contain no information about regression coefficients in this regression model (Cohn et al., 2022). Consequently, we adjust our sample accordingly to ensure that our outcome variable *Citations* contains sufficient variation in year t and t+1.[12]

We begin by identifying eligible firm pairs i-j and measuring the citation flow from cited firm j to citing firm i using the number of citations from citing firm i to cited firm j in year t based on the year of the patent application. In cases in which a given year has no cross-citations, we set the number of citations to zero. The resulting citation measure accounts for all citations included in citing firm i's patents applied for in year t to all past patents issued by cited firm j. In other words, the measure captures the degree to which past innovations of the cited firm are reflected in the innovations of the citing firm. To exploit within firm-pair variations in the direction of the information spillover, we focus on directional firm pairs (e.g.,

[12]Please also refer to Breuer and deHaan (2023) for a discussion on sufficient variation in the outcome variable.

citations of IBM's patents by Microsoft versus citations of Microsoft's patents by IBM). To avoid confusion, we use the terms *Citing* and *Cited* to indicate whether a variable refers to citing firm i or to cited firm j, respectively.

3.3 Measuring social media analyst coverage overlaps

We measure common SMA coverage (*Common SMAs*) based on articles published on the platform Seeking Alpha. By analyzing user-generated content, we identify SMAs that covered both citing firm i and cited firm j of a particular firm pair at any point during year t. We then consolidate the individual Seeking Alpha contributor observations and calculate the number of contributors who cover both citing firm i and cited firm j in year t for each unique firm pair. For example, we determine the number of contributors who cover both IBM and Microsoft during the same year. We use these contributors to measure common SMA coverage.

To create the ultimate set of directional firm-pair observations that have both citation and SMA data available, we match the sample of common SMA coverage overlaps (the independent variable) with the sample of cross-citations (the dependent variable) based on the year of the patent application.

The duration from the initial patent application to the ultimate patent grant can be significant, frequently stretching across multiple years. In addition, citations can be added at any stage during the application procedure.[13] As our citation metric relies on the citations listed in the ultimate patent grant, the dependent variable includes all citations regardless of when they were added during the application process, in addition to those made in the

[13]The applicant is obligated to inform the USPTO of any known prior work that is significant to patentability throughout the application process. This responsibility necessitates the filing of an *Information Disclosure Statement (IDS)* to notify the examiner of pertinent prior work references as soon as the applicant becomes aware of new ones. A reduced number of previous citations raises the possibility of patent invalidation, as it increases the likelihood that a cited patent could be utilized to challenge its validity (Allison and Lemley, 1998; Kesan, 2002).

application year.[14] Figure 1 visually represents the connection and temporal aspects of our measurement concerning overlaps and cross-citations in social media analyst coverage.

This research design is consistent with prior literature and has the advantage that it closely links the information available to the firm to the citations. (i.e., the citations in the years directly following the common SMA coverage). Any other matching that considers citations of patents filed later would link the common SMA coverage to citations that are in the distant future and hence make it more difficult to identify the effect. Consequently, we focus on citations of patents filed in year t but we repeat the analysis with citations of patents filed in year t+1 to make sure that our results are not driven by this research design choice.

[INSERT Figure 1]

Our final base sample ranges from 2006 to 2018 and contains 593,969 directional firm-pair-year observations based on 66,299 directional pairs of citing and cited firms. The final base sample is unbalanced, since, in some years, a firm pair might not fulfill the sample selection criteria (e.g., the citing firm has no innovations).

4 Social media analyst coverage and interfirm patent citations

4.1 Social media analysts' coverage portfolios

We first explore differences in coverage portfolios and firm connectivity based on shared coverage of social media analysts and traditional sell-side analysts to verify that user-generated, non-traditional investment research plausibly provides incremental information content compared to traditional sell-side research. Table 1 panel A presents general coverage statistics for social media analysts and traditional sell-side analysts for our sample of 53,678 firm-year observations. On average, firms enjoy more coverage from sell-side analysts (8.80)

[14]Due to data limitations, it is not feasible to pinpoint the precise timing of the citation.

than social media analysts (2.72). While this may be explained by fewer social media analysts than traditional sell-side analysts in the sample (10,129 vs. 14,855), we also find that social media analysts cover fewer unique firms per year (7.65 vs. 9.05). This suggests that the average SMA focuses on fewer firms, although we also observe much more variation in SMA coverage compared to sell-side analyst coverage. In general, these patterns are consistent with differences between SMAs' and traditional sell-side analysts' coverage decisions as documented by prior research (e.g., Campbell et al., 2019; Drake et al., 2022; Jame et al., 2016; Koenraadt, 2023).

To test whether SMAs' coverage portfolios are unique compared to those of traditional sell-side analysts, we examine the connectivity between firms emerging from their shared analyst coverage. We measure analysts' coverage portfolio connectivity as the percentage of firms with which firm i is connected through its analyst coverage.[15] Results in Table 1, panel B, suggest that connectivity emerging from SMAs is, on average, larger than connectivity emerging from sell-side analysts (2.23 percent vs. 1.32). Most of the connectivity emerging from shared SMA coverage is even incremental to traditional analyst-based connectivity. Despite covering fewer firms, SMAs thus seem to have a broader coverage portfolio that adds unique connections between firms that is not existent from traditional sell-side analysts, a prominent resource of industry and competitive information.

To further shed light on this result, Table 1, panel A, reports the top-20 firm-pair industries in terms of common social media analysts with at least 50 observations.[16] Consistent with panel B, only few firm-pair industries receive high coverage levels from both social media analysts (*Common SMAs*) and traditional sell-side analysts (*Common analysts*). For most of the top-20 firm pair industries in terms of social media analyst coverage, sell-side

[15]More specifically, common analysts-based connectivity is calculated as $F_{it} - 1/E_t - 1$, where $F_{it} - 1$ is the number of unique firms covered by analysts that also cover firm i in year t and E_t is the total number of unique firms covered by analysts in year t. We calculate the measure separately for sell-side analysts and social media analysts as well as both together.

[16]We limit the sample of this descriptive analysis to industry pairs that have at least 50 observations (non-directional) to ensure that firm pair industries with only few observations do not drive the results.

analyst coverage is relatively low. Seven firm-pair industries for which citing firm A and cited firm B have no common sell-side analyst coverage at all.[17] These patterns confirm that SMAs' coverage portfolios are somewhat different from those of traditional sell-side analysts.

Taken together, descriptive statistics suggest that SMAs focus on a broad array of different firms and industries, which results in coverage portfolios that are different from those of traditional sell-side analysts. User-generated investment research may therefore enhance firms' information environments with perspectives and insights not traditionally available from other information intermediaries.

[INSERT Table 1]

4.2 Baseline: Social media analyst coverage and interfirm patent citations

Table 2 presents the descriptive statistics for the 593,969 yearly firm-pair observations in our regression sample. On average, each firm pair makes 6.51 directional cross-citations per year and has 0.88 common social media analysts (*Common SMAs*) that cover both firms on Seeking Alpha. The number of cross-citations and SMA coverage overlaps varies considerably among the firm pairs in the sample, with a minimum (maximum) of 0 (37,222) cross-citations and 0 (121) common SMAs. All other descriptives are generally comparable to prior research with differences being largely due to the smaller sample and more recent sample period.

[INSERT Table 2]

Table 3 presents the estimates of equation [1] for different control specifications (see Whited et al. (2022) for a discussion). Columns (1) and (2) report coefficient estimates including firm-pair fixed effects and year fixed effects. Column (3) presents estimates for

[17] The statistics for the firm pairs firm A - firm B and firm B - firm A are not necessarily identical due to differences in the sample, e.g., when a firm did not issue a patent in a given year or had no patents to be cited.

the full specification including firm-pair fixed effects and firm-specific time trends instead of year fixed effects. An increase in common SMA coverage is significantly positively associated with an increase in cross-firm patent citations across all specifications.

Consistent with prior literature, we also find a significantly positive association for overlaps in sell-side analyst coverage. Judging from the size of the coefficient estimates in column (3), information spillovers from social media analysts are about half as likely as information spillovers from sell-side analysts. This is plausible considering that sell-side analysts are more likely to regularly interact with corporate management. At the same time, these patterns also suggest that firms may indeed be able to extract incrementally useful information from user-generated content on social media investment platforms such as Seeking Alpha.

We note that explanatory power in our main tests specification only slightly increases from 28.3% to 28.4% when adding additional firm-pair control variables (Table 3 column (2) compared to (1)). Adding firm-specific time trends, however, increases explanatory power to 34.8% (Table 3 column (3)). These patterns are reassuring of our fixed effect structure and suggest that it is unlikely that the remaining variation in cross-citations is the result of correlated omitted firm-pair characteristics that vary over time.

We run several additional tests for robustness (see Online Appendix). First, we re-estimate equation [1] based on OLS and logit specifications instead of negative binomial regressions. Second, to ensure that our results are not driven by few observations with extreme citation behavior, we run the main specifications excluding observations with citations that exceed the top 1% percentile. In both cases, results are comparable to the main results in terms of significance and effect size. Third, we replace the main dependent variable *Citations* in year t with the lead variable, *Citations* in year t + 1, to address potential timing issues in the matching of analyst coverage overlaps and interfirm information spillovers. The coefficient estimate for *Common SMAs* remains statistically significant, but decreases in size. In addition, the coefficient estimate on *Common analysts* becomes insignificant. These results not only support the information spillover explanation, but also the timing of our matching based on the application year. Finally, we investigate alternative means of clustering standard errors, i.e., by firm-pair, citing and cited firm, industry-pairs, as well as citing and cited industry. The coefficient estimate for *Common SMAs* remains statistically significant in all specifications. Overall, results seem robust to different model specifications, estimation approaches, and alternative clustering of standard errors.

[INSERT Table 3]

4.3 Approach to identification: Exogenous variation in SMAs' coverage incentives

Our main tests rely on exploiting variation in social media analysts' coverage overlaps and interfirm patent citations within firm-pairs. That is, due to including firm-pair fixed effects as well as firm-specific year fixed effects, we only exploit the residual variation within firm-pairs over time. This specification combined with control variables for changes over time in firm-pair-specific characteristics (e.g., business and technological similarity) already addresses many potential alternative explanations. However, one may still be concerned about some correlated omitted variables explaining the results in Table 3. To address this concern, we use plausibly exogenous variation in Seeking Alpha analyst coverage overlaps as an additional approach to identification.

On June 1, 2013, Seeking Alpha implemented a revised payment system aiming to increase coverage of small-cap stocks. Previously, contributors were compensated solely based on the page views their articles received, resulting in a bias toward larger, more popular companies. Under the new structure, contributors continued to receive payment per page view, but with additional bonuses for high-quality analyses of small-cap stocks. These bonuses ranged from a minimum of $150 for quality articles to $500. [18] This change in the

[18]See, for example, https://seekingalpha.com/article/2134803-how-much-does-seeking-alpha-pay-its-

payment system allows us to plausibly identify the effect of social media analyst coverage overlaps on interfirm patent citations.

Specifically, we follow Gu et al. (2023) and use this change in the payment system as an instrument to implement a fuzzy difference-in-difference regression design.[19] The change in the payment structure is a relevant instrument, since monetary incentives are strongly associated with SMA coverage (see e.g., Gu et al., 2023; Koenraadt, 2023) and thus coverage overlaps, especially when coverage of a specific group of firms (i.e., small cap stocks) is incentivized. Since these monetary incentives pertain to Seeking Alpha only and are exclusively awarded to SMAs, they can only affect intra-firm patent citations directly through changes in Seeking Alpha analysts coverage overlap. However, because the incentive is applicable only to small cap firms and does not depend on other firm characteristics (e.g., the intensity of prior coverage), it is also unlikely that the adjustment in the payment system is associated with any other firm characteristic that may affect intra-firm patent citations at the same time (see, e.g., Koenraadt, 2023).

We estimate the following two equations:

$$\begin{aligned} Common\ SMAs_{i,j,t} &= \beta_0 + \beta_1 Both\ small\ cap_{i,j,t}\ x\ Post_t + \beta_2 Both\ small\ cap_{i,j,t} \\ &+ \sum Lewbel\ instruments_{i,j,t} + \sum Controls_{i,j,t} + \sum Firm\ pair_{i,j} \\ &+ \sum Citing\ Firm_i\ x\ Year_t + \sum Cited\ Firm_j\ x\ Year_t + \epsilon \end{aligned} \tag{2}$$

$$\begin{aligned} Log(1 + Citations_{i,j,t}) &= \beta_0 + \beta_1 Common\ SMAs\ (instrumented)_{i,j,t} \\ &+ \sum Controls_{i,j,t} + \sum Firm\ pair_{i,j} + \sum Citing\ Firm_i\ x\ Year_t \\ &+ \sum Cited\ Firm_j\ x\ Year_t + \epsilon \end{aligned} \tag{3}$$

contributors and https://seekingalpha.com/article/1475331-why-were-boosting-payments-to-high-value-contributors.

[19]Please refer to De Chaisemartin and d'Haultfoeuille (2018) and Armstrong et al. (2022) for detailed discussions of this approach.

Together, both equations combine instrumental variables with a difference-in-difference approach to generate a fuzzy difference-in-difference estimator. Conceptually, this research design enables us to directly link citations for small cap firm-pairs with the increase in common special media analyst coverage on Seeking Alpha.

Equation [2] is the first stage of the fuzzy difference-in-difference research design where *Both small cap* is an indicator variable that takes the value of one if both firms' market capitalization is below one billion dollar in year t, and zero otherwise. *Post* is an indicator variable that takes the value of one in year 2014, the year after the change in the payment system, or later, and zero otherwise. Equation [2] does not separately include the base term for *Post* as it would be subsumed by the fixed effect structure. Besides the difference-in-difference specification, we also use all control variables to construct Lewbel instruments to strengthen our identification (see, e.g., Baum and Lewbel, 2019; Lewbel, 2012, for further explanation of this approach). Equation [3] is the second stage of the fuzzy difference-in-difference research design where *Common SMA* (*instrumented*) is the fitted value from estimating Equation [2].

Table 4 presents the estimates of this fuzzy difference-in-difference design. Column (1) presents estimates for the first stage, while column (2) presents the results for the second stage including the instrumented analyst coverage variable. First, we observe that the coefficient on the interaction of *Both small cap* x *Post* is positive and statistically significant. This suggests that the change in the payment structure and hence the incentives to cover small cap firms indeed caused more coverage of small cap firm pairs, and thus also a potential change in coverage overlaps for these firms. More importantly, however, we find that the coefficient on *Common SMAs* (*instrumented*) is positive and statistically significant, consistent with our main analysis. The fuzzy difference-in-difference approach thus confirms the interpretation that an increase in common social media analysts results in more cross-firm patent citations, which is consistent with managerial learning from user-generated analyses on social media.

[INSERT Table 4]

5 Determinants of managerial learning from social media analysts

We next explore the sources of managerial learning from social media analysts' investment research. On the one hand, the effect may simply emerge from the fact that firms are followed by 'more' information intermediaries which create a broader information environment for managerial learning. In this case, it does not matter whether the firm is followed by an additional social media analyst or any other information intermediary, such as an additional traditional sell-side analyst. On the other hand, the nature of crowd-sourced non-traditional investment research suggests that social media analysts provide a distinct set of competitive information that may be incrementally informative for managerial learning and decision making. In this case, social media analyst following would provide information that is not readily available from other information intermediaries or only at different cost. We explore this question by investigating social media analysts' coverage portfolio characteristics, differences in their contributed content, as well as their credibility and visibility.

5.1 Coverage portfolio characteristics

In our first set of tests we explore the relation between interfirm information spillovers and heterogeneity in the characteristics of common social media analysts (CSMAs). Following prior literature, we focus on experience, activity level, portfolio size, and industry specialization, as four core characteristics related to analysts' ability to provide meaningful information (e.g., Clement, 1999; Mikhail et al., 1997; Sonney, 2007). The results in Table 5, panel B, indicate some striking differences compared to prior findings for traditional sell-side analysts.

Experience is often considered a factor that positively affects analysts' ability to collect and process information. Prior research on traditional sell-side analysts documents that analysts' experience is also positively associated with managerial learning from analysts' research output. Interestingly, we only find weak evidence for the role of experience for SMAs. The coefficienty for high and regular experience CSMAs are hardly significantly different from each other, indicating that experience does not necessarily determine the information value of crowd-sourced investment research for corporate decision making (column (1)). At the same time, we find that the level of activity is positively associated with information spillovers (columns (2)). To avoid that activity merely captures mass-content providers, we split activity level into three categories (high, medium, low). Information spillovers seem to originate primarily from SMAs that show medium levels of activity. These are likely also those SMAs that engage in actual analyses and information search and do not only mass produce articles with little new information. This is consistent with the results for portfolio size, for which we again find no significant differences between large and regular portfolio CSMAs. These patterns suggest that it is neither analysts experience nor their mass dissemination of information that matters for managerial learning, but rather their collection, analysis, and perspective on a broader set of potentially relevant information.

This conjecture is further supported when testing for differences in industry specialization (column (4)). Knowledge of industry and market development is often considered the most valued attribute of traditional sell-side analysts (Bagnoli et al., 2008; Bradshaw, 2012; Kadan et al., 2012). Thus, one may expect that firms should be able to extract more useful information from research provided by highly specialized analysts. But while this may be true for traditional sell-side analysts (Martens and Sextroh, 2021), we find the opposite result for SMAs. High industry specialization is associated with fewer interfirm patent citations relative to those analysts that do not specialize. These results are again consistent with the notion that the informational benefits of SMAs for managerial learning emerges from their broader perspective: Specialized SMAs hardly add information to that already provided by other highly specialized information intermediaries, such as traditional sell-side analysts.

[INSERT Table 5]

5.2 Characteristics of contributed content

We further explore SMAs' unique information set as a plausible source for managerial learning from crowd-sourced information by analysing the characteristics of SMAs' contributed content. In particular, we further test the idea that the value of SMAs may be driven by their ability to collect and share additional decision-relevant information and perspectives not readily available from other sources. To do so, we focus on multiple characteristics of the content provided by SMAs: (1) the concentration of topics, (2) the similarity of content provided by a SMA with the content contributed by other SMAs, (3) the specificity of content provided, i.e., the relative frequency of references to specific entities, (4) coverage by technology-focused SMAs, (5) coverage by earnings-focused SMAs, and (6) coverage by trading-focused SMAs. The first three dimensions capture the extent to which the content is more unique and likely includes actionable information. The latter three dimensions capture to what extent SMAs provide information that is likely relevant for corporate decision making, with technology-focused content being conceptually directly related to our measure of (technological) information spillovers measured by patent citations.

The results presented in Table 6, panel B, suggest that managerial learning from crowd-sourced investment research may indeed depend on whether the content contains information not readily available from other information intermediaries. First, we use LDA topic modeling to identify and classify the most frequent 20 topics contained in the entire sample of Seeking Alpha articles. These topics range from investing and portfolio-focused content to earnings and technology-focused content (see Appendix B for an overview of the 20 identified article topics). Similar to industry specialization in Table 5, we then use the Herfindahl index to determine whether analysts concentrate on specific topics. Interestingly, and in contrast to industry specialization, we find that information spillovers are more likely if the firm shares coverage from analysts that specialize in topics. Second, we use cosine similarity to determine the similarity of all articles published on Seeking Alpha in a particular year. We then classify a SMA as a provider of more unique (i.e., less similar) content if the similarity of the analyst's reports to peer reports does not exceed the 10th percentile across all Seeking Alpha analysts. We find that the association of shared SMA coverage and interfirm information is larger for SMAs that publish more unique content on Seeking Alpha (i.e., have low content similarity with other SMAs in a given year, column (2)). This also alleviates concerns that our main results merely capture firms' and SMAs' reliance on the same third party information. If that is the case, we should expect higher effects for SMAs with more similar content as they merely pick up on the same pieces of information. Third, the effect is also larger for those more likely to include specific entity references in their publications and accordingly make it easier to process the information provided (column (3)). Taken together, these patterns are not only consistent with the informational role of social media analysts, but also suggest that their value for managerial learning primarily emerges from the provision of unique information.

Finally, we also specifically test whether information spillovers depend on the content provided. First, we use the topics identified to classify SMAs as technology-focused (earnings-focused) [trading-focused] if at least one of the most important topics in an analyst's articles is related to technology (earnings) [trading]. We find larger effects for SMAs that provide technology-related content, while the opposite is true for SMAs that mostly publish earnings- or trading-related articles. Put differently, managerial learning from SMAs is more likely if firms are covered by SMAs that provide information that is potentially relevant from a business strategy perspective, such as those related to technological developments.

[INSERT Table 6]

Second, we exploit an alternative source of SMA coverage that differs in the extent to which analysts provide detailed analyses about the firms they cover: *Estimize.com*. Estimize primarily publishes quantitative crowd-sourced earnings and revenue forecasts. Prior research suggests that these estimates provide incremental information and even have a disciplining effect on professional sell-side analysts (Da and Huang, 2020; Jame et al., 2016, 2022). Thus, while both Seeking Alpha and Estimize are useful sources of crowd-based information, Estimize does not provide any accompanying analysis or offers opportunity for users to publish additional information. Gomez et al. (2022), for example, use Estimize analysts to verify whether they have the same information processing effects for retail investors surrounding earnings announcements as do articles published on Seeking Alpha. They find no effect of Estimize coverage on information processing costs, consistent with the argument that Estimize primarily provides forecasts but no additional information.

We follow the idea of Gomez et al. (2022) and use analyst coverage on Estimize to test whether our results can plausibly be attributed to an information spillover effect or are merely driven by analysts' coverage decisions. Comparing the coverage statistics for Seeking Alpha analysts from Table 2 with those for Estimize analysts presented in Jame et al. (2016) suggests that both Estimize and Seeking Alpha users may be relatively similar in their coverage decisions. Comparing the effect of SeekingAlpha and Estimize thus also serves as an alternative approach to identification that overcomes the limitation that SMA coverage overlaps are endogenously determined and may covary with other unobserved firm-pair characteristics, which are potentially important determinants of our outcome variable. If our results are merely due to correlated omitted variables that explain both SMAs' coverage decisions and firms' similarity over time, we should find similar results for SMA coverage overlaps on Estimize. If, however, crowd-sourced investment research published on Seeking Alpha provides useful information for managerial decision making, no such effect should exist for coverage overlaps on Estimize (as these overlaps are not linked to any particular analysis or discussion piece).

In Table 7, we replicate our main analysis for the sample that includes the coverage of both Seeking Alpha and Estimize. The sample is smaller than that used for the main test, as Estimize was introduced later than Seeking Alpha. Consequently, the sample only ranges from 2012 to 2017. Column (1) presents the replication of the main result from Table 3 showing very similar results in terms of coefficient size and significance. The results in columns (2) and (3) reflect the addition of *Common Estimize analysts* to test for the (incremental) effect of coverage overlaps originating from analysts on Estimize. Coefficient estimates are insignificant and show virtually no association with subsequent patent citations. At the same time, the size and significance of the coefficient for *Common SMAs* remains the same. To the extent that coverage decisions on Seeking Alpha and Estimize are determined by similar factors, these results further support a plausibly causal interpretation of information spillovers between firms originating from investment research published by shared SMAs. In addition, the test directly contrasts general following of social media analysts (as proxied by coverage on Estimize) with their provision of information and analyses (as measured by coverage on SeekingAlpha).

Taken together, these results further alleviate concerns that SMA coverage overlaps merely capture similarities and differences between firms. Instead, our results suggest that the value of SMAs for managerial learning may not simply emerge from an 'increased coverage' effect due to more information intermediaries covering a firm, but rather originate from access to a more diverse group of individuals who collect and provide information that is not readily available from interactions with other information intermediaries.

[INSERT Table 7]

5.3 The role of social media analysts' credibility and visibility

In our final set of tests, we investigate characteristics of social media analysts that potentially affect firms willingness or ability to extract information. Specifically, we analyse differences in analysts' credibility and visibility. To capture the credibility of the information, we focus on whether SMAs publish their research anonymously or non-anonymously. We follow Dyer and Kim (2021) and classify an SMA as non-anonymous if the analyst discloses one of the following items on her profile: website URL, Twitter or LinkedIn link, geographic or university reference in her biography, or her email address. To capture an analysts' visibility, we focus on the number of followers on Seeking Alpha. Table 8 provides the results.

We find that information flows are more likely when firms are covered by non-anonymous SMAs (column (1)). In addition, information spillovers are more likely if firms are covered by analysts with a high following on Seeking Alpha (column (2)). These results suggest that the credibility of the contributed content as well as firms' processing costs may indeed play a role for whether the information provided by social media analysts is considered in managerial decision making.

[INSERT Table 8]

6 Conclusion

The rise of financial technology and social media has led to an increased amount and accessibility of user-generated investment research. While evidence shows that stock research published on social media platforms, such as Seeking Alpha, StockTwits, or Estimize, affects market pricing, the question remains whether non-traditional investment results published by SMAs also provides useful insights for firms' decision-making. Our results suggest that firms indeed collect information from monitoring user-generated investment research, as firms with shared SMA coverage are more likely to invest in similar technologies. This effect is incremental to shared coverage by traditional sell-side analysts and other potential sources of information. The value of SMAs may come from not simply an increased coverage effect but rather from access to a more diverse group of individuals who potentially collect and provide information not readily available from other intermediaries. In particular, results are stronger for SMAs that provide more unique content and discussions about technological developments rather than simple updates around corporate earnings announcements. In addition, consistent with firms learning from social media analysts' investment research and commentary, we do not find results for analysts that merely publish quantitative forecasts.

Overall, our results establish a robust association between social media analyst coverage and corporate investment decisions and therefore add to our understanding of the role of a nascent but very prevalent information intermediary. In particular, our findings suggest that SMAs not only enhance firms' information environment by providing useful information for capital markets, but also serve as an additional source of information to help guide corporate investment decisions.

References

Abadie, A., Athey, S., Imbens, G. W., Wooldridge, J. M., 2023. When should you adjust standard errors for clustering? Quarterly Journal of Economics 138, 1–35.

Agrawal, A., Galasso, A., Oettl, A., 2017. Roads and innovation. Review of Economics and Statistics 99, 417–434.

Ali, U., Hirshleifer, D., 2020. Shared analyst coverage: Unifying momentum spillover effects. Journal of Financial Economics 136, 649–675.

Allison, J. R., Lemley, M. A., 1998. Empirical evidence on the validity of litigated patents. AIPLA Quarterly Journal 26, 185–274.

Antweiler, W., Frank, M. Z., 2004. Is all that talk just noise? The information content of internet stock message boards. Journal of Finance 59, 1259–1294.

Armstrong, C. S., Glaeser, S. A., Huang, S., 2022. Contracting with controllable risk. The Accounting Review 97, 27–50.

Backus, M., Conlon, C., Sinkinson, M., 2021. Common ownership in America: 1980–2017. American Economic Journal: Microeconomics 13, 273–308.

Bae, J., Biddle, G. C., Park, C. W., 2022. Managerial learning from analyst feedback to voluntary capex guidance, investment efficiency, and firm performance. Management Science 68, 583–607.

Bagnoli, M., Watts, S. G., Zhang, Y., 2008. Reg-FD and the competitiveness of all-star analysts. Journal of Accounting and Public Policy 27, 295–316.

Bartov, E., Faurel, L., Mohanram, P. S., 2018. Can twitter help predict firm-level earnings and stock returns? The Accounting Review 93, 25–57.

Baum, C. F., Lewbel, A., 2019. Advice on using heteroskedasticity-based identification. The Stata Journal 19, 757–767.

Belenzon, S., Schankerman, M., 2013. Spreading the word: Geography, policy, and knowledge spillovers. Review of Economics and Statistics 95, 884–903.

Blankespoor, E., deHaan, E., Marinovic, I., 2020. Disclosure Processing Costs, Investors' Information Choice, and Equity Market Outcomes: A Review. Journal of Accounting and Economics 70, 1–46.

Bond, P., Edmans, A., Goldstein, I., 2012. The real effects of financial markets. Annual Review of Financial Economics 4, 339–360.

Boudreau, K. J., Lakhani, K. R., 2013. Using the crowd as an innovation partner. Harvard Business Review.

Bradshaw, M. T., 2012. Discussion of 'analysts' industry expertise'. Journal of Accounting and Economics 54, 121–131.

Breuer, M., deHaan, E., 2023. Using and interpreting fixed effects models. Journal of Accounting Research Forthcoming.

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, 731–779.

Chen, C., Ke, B., Li, D., Goh, K. Y., 2021. Individual Investors and Social Media Analysts' Stock Coverage Behavior. Working Paper.

Chen, H., Hu, Y. J., Huang, S., 2019. Monetary Incentive and Stock Opinions on Social Media. Journal of Management Information Systems 36, 391–417.

Chen, H., Prabuddha, D., Yu, J. H., Byoung-Hyoun, H., 2014. Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media. Review of Financial Studies 27, 1367–1403.

Cision, 2017. Social media monitoring: Is critical for investor relations. White Paper.

Clarke, J. C., Chen, H., Du, D., Hu, J., 2021. Fake News, Investor Attention, and Market Reaction. Information Systems Research 32, 35–52.

Clement, M. B., 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? Journal of Accounting and Economics 27, 285–303.

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

Cookson, J. A., Lu, R., Mullins, W., Niessner, M., 2024a. The Social Signal. Journal of Financial Economics Forthcoming.

Cookson, J. A., Mullins, W., Niessner, M., 2024b. Social media and finance. Working Paper .

Cookson, J. A., Niessner, M., 2019. Why Don't We Agree? Evidence from a Social Network of Investors. Journal of Finance 75, 173–228.

Cookson, J. A., Niessner, M., Schiller, C., 2024c. Can social media inform corporate decisions? Evidence from merger withdrawals. Journal of Finance Forthcoming.

Da, Z., Huang, X., 2020. Harnessing the wisdom of crowds. Management Science 66, 1783–2290.

De Chaisemartin, C., d'Haultfoeuille, X., 2018. Fuzzy differences-in-differences. The Review of Economic Studies 85, 999–1028.

DeAngelis, M. D., Mi, D., Moon, J., Wallace, D., 2021. Social media analyst coverage of socially responsible firms. Working Paper.

DeWalt, D., 2021. The next evolution of data and crowd sourcing: Crowd empowerment. Available online: `https://www.forbes.com/sites/forbesfinancecouncil/2021/07/20/the-next-evolution-of-data-and-crowdsourcing-crowd-empowerment/?sh=5560fb7b3a7b`, accessed: 2023-06-27.

Drake, M. S., Moon, J. R., Twedt, B. J., Warren, J. D., 2022. Social media analysts and sell-side analyst research. Review of Accounting Studies 28, 1–36.

Drake, M. S., Roulstone, D. T., Thornock, J. R., 2012. Investor information demand: Evidence from google searches around earnings announcements. Journal of Accounting Research 50, 1001–1040.

Drake, M. S., Roulstone, D. T., Thornock, J. R., 2015. The determinants and consequences of information acquisition via EDGAR. Contemporary Accounting Research 32, 1128–1161.

Drake, M. S., Thornock, J. R., Twedt, B. J., 2017. The internet as an information intermediary. Review of Accounting Studies 22, 543–576.

Dube, S., Zhu, C., 2021. The disciplinary effect of social media: Evidence from firms' responses to glassdoor reviews. Journal of Accounting Research 59, 1783–1825.

Dyer, T., Kim, E., 2021. Anonymous Equity Research. Journal of Accounting Research 59, 575–611.

Edmans, A., Goldstein, I., Jiang, W., 2012. The real effects of financial markets: The impact of prices on takeovers. Journal of Finance 67, 933–971.

Elliott, W. B., Gale, B., Grant, S. M., 2018. Navigating through the crowd: How do investors assess contributor credibility and make investment judgments on social media platforms? Working Paper.

Farrell, M., Clifton Green, T., Jame, R., Markov, S., 2022. The Democratization of Investment Research and the Informativeness of Retail Investor Trading. Journal of Financial Economics 145, 616–641.

Flam, R., Shafron, E., Sharp, N., Twedt, B., 2023. Managing the media: Corporate media relations officers and the evolving media landscape. Working Paper.

Gomes-Casseres, B., Hagedoorn, J., Jaffe, A. B., 2006. Do alliances promote knowledge flows? Journal of Financial Economics 80, 5–33.

Gomez, E., Heflin, F., Moon, J., Warren, J., 2022. Financial analysis on social media and disclosure processing costs: Evidence from Seeking Alpha. Working Paper.

Gu, P., Yost, B., Zou, Y., 2023. Firm visibility and acquisition likelihood: Evidence from seeking alpha coverage. Working Paper.

Hales, J., Moon, J. R., Swenson, L. A., 2018. A new era of voluntary disclosure? Empirical evidence on how employee postings on social media relate to future corporate disclosures. Accounting, Organizations and Society 68-69, 88–108.

He, Q., Leung, H., Qiu, B., Zhou, Z., 2024. The effect of social media on corporate innovation: Evidence from seeking alpha coverage. Management Science Forthcoming.

Hoberg, G., Phillips, G., 2010. Product market synergies and competition in mergers and acquisitions: A text-based analysis. Review of Financial Studies 23, 3773–3811.

Hoberg, G., Phillips, G., 2016. Text-based network industries and endogenous product differentiation. Journal of Political Economy 124, 1423–1465.

Huang, K., Li, M., Markov, S., 2020. What Do Employees Know? Evidence from a Social Media Platform. The Accounting Review 95, 199–226.

Jaffe, A. B., Trajtenberg, M., 1996. Flows of knowledge from universities and federal laboratories: Modeling the flow of patent citations over time and across institutional and geographic boundaries. Proceedings of the National Academy of Sciences 93, 12671–12677.

Jaffe, A. B., Trajtenberg, M., 2002. Patents, citations, and innovations: A window on the knowledge economy. MIT Press.

Jaffe, A. B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics 108, 577–598.

Jame, R., Johnston, S., Markov, S., Wolfe, M. C., 2016. The Value of Crowdsourced Earnings Forecasts. Journal of Accounting Research 54, 1077–1110.

Jame, R., Markov, S., Wolfe, M., 2022. Can FinTech competition improve sell-side research quality? The Accounting Review 97, 287–316.

Kadan, O., Madureira, L., Wang, R., Zach, T., 2012. Analysts' industry expertise. Journal of Accounting and Economics 54, 95–120.

Kesan, J. P., 2002. Carrots and stickes to create a better patent system. Berkeley Technology Law Journal 17, 763–797.

Koenraadt, J., 2023. Monetary Incentives of Social Media Analysts. Working Paper.

Kogan, L., Papanikolaou, D., Seru, A., Stoffman, N., 2017. Technological innovation, resource allocation, and growth. Quarterly Journal of Economics 132, 665–712.

Kogan, S., Moskowitz, T. J., Niessner, M. T., 2021. Social Media and Financial News Manipulation. Working Paper .

Lampe, R., 2012. Strategic citation. Review of Economics and Statistics 94, 320–333.

Lee, C. M., Shi, T., Teng Sun, S., Zhang, R., 2024. Production complimentarity and information transmission across industries. Journal of Financial Economics 155, 103812.

Lee, C. M., Teng Sun, S., Wang, R., Zhang, R., 2019. Technological links and predictable returns. Journal of Financial Economics 132, 76–96.

Lerman, A., 2020. Individual investors' attention to accounting information: Evidence from online financial communities. Contemporary Accounting Research 37, 2020–2057.

Lewbel, A., 2012. Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. Journal of Business & Economic Statistics 30, 67–80.

Martens, T., Sextroh, C. J., 2021. Analyst coverage overlaps and interfirm information spillovers. Journal of Accounting Research 59, 1425–1480.

Mikhail, M. B., Walther, B. R., Willis, R. H., 1997. Do security analysts improve their performance with experience? Journal of Accounting Research 35, 131–157.

Mitts, J., 2020. Short and Distort. Journal of Legal Studies 49, 287–334.

Reuer, J. J., Devarakonda, R., 2017. Partner selection in R&D collaborations: Effects of affiliations with venture capitalists. Organization Science 28, 574–595.

Sonney, F., 2007. Financial analysts' performance: Sector versus country specialization. Review of Financial Studies 22, 2087–2131.

Surowiecki, J., 2004. The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations. Anchor Books, New York.

Whited, R. L., Swanquist, Q. T., Shipman, J. E., Moon Jr, J. R., 2022. Out of control: The (over) use of controls in accounting research. The Accounting Review 97, 395–413.

Zacchia, P., 2020. Knowledge spillovers through networks of scientists. Review of Economic Studies 87, 1989–2018.

Appendix A

Excerpt from an article published on Seeking Alpha

Seeking Alpha^α^ CREATE FREE ACCOUNT LOG IN

Home > Stock Ideas > Long Ideas

The War In Cloud Computing And SaaS Is Heating Up For VMware And Others

Aug. 21, 2013 5:04 PM ET | **VMware, Inc. (VMW) Stock** | CRM, CTXS, MSFT... | 12 Comments

Tech Alpha
124 Followers

Follow

[...]

VMware has to find a good response to this formidable new competitor's threat ASAP, but it is uncertain whether VMware will be able to do so. As one can see from VMware's financial reports and its management's last forecast, the company's growth rate has dropped and, going forward, will probably drop further. As a thousand-pound gorilla, the company no longer has the ability to come up with major innovations frequently, drastically overhauling its technologies/framework to meet the market's ever-evolving needs. Moreover, in the case of Sphere 3D, because it has filed patents for its proprietary technologies, VMware may not be able to put good features of Sphere 3D's solutions into its own products even if its R&D team is able to mimic that company's features. Therefore, VMware may be forced to use less effective and more costly workarounds to avoid violating Sphere 3D's patents. One alternative for VMware to diffuse this threat is to adopt Bill Gates' famous strategy: "If you cannot beat them, buy them." VMware itself is famous for this strategy as it has acquired numerous companies in the past. Of course, it is unknown whether Sphere 3D's founders are willing to consider selling the company right now. Assuming they are willing to give away their years of hard work for several hundred million dollars, it may be a money well spent for VMware to further strengthen its technologies and product offerings to spare a lot of trouble and risk down the road.

Source: https://seekingalpha.com/article/1649942-the-war-in-cloud-computing-and-saas-is-heating-up-for-vmware-and-others

Appendix B

Variable Definitions

Patent-related variables	
Citations	The number of citations from citing firm i to cited firm j in year t. Source: Kogan et al. (2017) Patent Data updated 2020
Cited/cited patent stock ratio	Citing patent stock plus one divided by cited patent stock plus one in year t in thousands. Source: Kogan et al. (2017) Patent Data updated 2020
Citing/cited citation ratio	Citing total citations plus one divided by cited total citations plus one in year t in thousands. Source: Kogan et al. (2017) Patent Data updated 2020
Social media analyst-specific variables	
Common SMAs	The number of common Seeking Alpha analysts of citing firm i and cited firm j in year t. Source: Seeking Alpha Data
High experience CSMAs	The number of common Seeking Alpha analysts that are high experience Seeking Alpha analysts in year t. We classify a Seeking Alpha analyst as high experience Seeking Alpha analyst if the Seeking Alpha analyst's years of experience in year t exceed the 90th percentile across all Seeking Alpha analysts. Source: Seeking Alpha Data
Regular experience CSMAs	The number of common Seeking Alpha analysts that are regular experience Seeking Alpha analysts in year t. We classify a Seeking Alpha analyst as high experience Seeking Alpha analyst if the Seeking Alpha analyst's years of experience in year t do not exceed the 90th percentile across all Seeking Alpha analysts. Source: Seeking Alpha Data
Large portfolio CSMAs	The number of common Seeking Alpha analysts that are large portfolio Seeking Alpha analysts in year t. We classify Seeking Alpha analysts to cover a large portfolio of firms if the number of firms in the Seeking Alpha analyst's portfolio in year t exceeds the 90th percentile across all Seeking Alpha analyst portfolios. Source: Seeking Alpha Data
Regular portfolio CSMAs	The number of common Seeking Alpha analysts that are regular portfolio Seeking Alpha analysts in year t. We classify Seeking Alpha analysts to cover a regular portfolio of firms if the number of firms in the Seeking Alpha analyst's portfolio in year t does not exceed the 90th percentile across all Seeking Alpha analyst portfolios. Source: Seeking Alpha Data
High specialisation CSMAs	The number of common Seeking Alpha analysts that are high specialisation Seeking Alpha analysts in year t. To determine Seeking Alpha analyst's specialization we use the Herfindahl index and measure the industry concentration of Seeking Alpha analysts' coverage portfolios in year t on the two-digit SIC level (similar to Sonney (2007)). High specialisation Seeking Alpha analysts are Seeking Alpha analysts with a Herfindahl index exceeding 0.9 in year t. Source: Seeking Alpha Data

(Continued on next page)

Appendix B
(continued)

Regular specialisation CSMAs	The number of common Seeking Alpha analysts that are regular specialisation Seeking Alpha analysts in year t. To determine Seeking Alpha analyst's specialization we use the Herfindahl index and measure the industry concentration of Seeking Alpha analysts' coverage portfolios in year t on the two-digit SIC level (similar to Sonney (2007)). Regular specialisation Seeking Alpha analysts are Seeking Alpha analysts with a Herfindahl index not exceeding 0.9 in year t. Source: Seeking Alpha Data
Non-anonymous CSMAs	The number of common Seeking Alpha analysts that are non-anonymous Seeking Alpha analysts in year t. We classify a Seeking Alpha analyst as non-anonymous Seeking Alpha analyst if the Seeking Alpha analyst has at least one of the following information on their profile: website, twitter/linkedin, geography, university, email. Source: Seeking Alpha Data
Anonymous CSMAs	The number of common Seeking Alpha analysts that are anonymous Seeking Alpha analysts in year t. We classify a Seeking Alpha analyst as anonymous Seeking Alpha analyst if the Seeking Alpha analyst has none of the following information on their profile: website, twitter/linkedin, geography, university, email. Source: Seeking Alpha Data
High content similarity CSMAs	The number of common Seeking Alpha analysts that are high similarity Seeking Alpha analysts in year t. We classify a Seeking Alpha analyst as a high similarity Seeking Alpha analyst if the Seeking Alpha analyst's reports similarity to peer reports exceed the 10th percentile across all Seeking Alpha analysts. Source: Seeking Alpha Data
Low content similarity CSMAs	The number of common Seeking Alpha analysts that are low similarity Seeking Alpha analysts in year t. We classify a Seeking Alpha analyst as a low similarity Seeking Alpha analyst if the Seeking Alpha analyst's reports similarity to peer reports does not exceed the 10th percentile across all Seeking Alpha analysts. Source: Seeking Alpha Data
High content specificity CSMAs	The number of common Seeking Alpha analysts that are high specificity Seeking Alpha analysts in year t. We classify a Seeking Alpha analyst as a high specificity Seeking Alpha analyst if the ratio of entities to total words in the Seeking Alpha analyst's reports exceed the 90th percentile across all Seeking Alpha analysts. Source: Seeking Alpha Data
Low content specificity CSMAs	The number of common Seeking Alpha analysts that are low specificity Seeking Alpha analysts in year t. We classify a Seeking Alpha analyst as a low specificity Seeking Alpha analyst if the ratio of entities to total words in the Seeking Alpha analyst's reports does not exceed the 90th percentile across all Seeking Alpha analysts. Source: Seeking Alpha Data

(Continued on next page)

Appendix B

(continued)

Technology-focused CSMAs	The number of common Seeking Alpha analysts that are technology analysts in year t. We classify a Seeking Alpha analysts as a technology analyst if at least one of the analyst's reports most important topic is one of the technology topics 4 and 17. All 20 article topics are obtained from an LDA on the full sample of Seeking Alpha articles and presented in Appendix C. Source: Seeking Alpha Data
Non-technology-focused CSMAs	The number of common Seeking Alpha analysts that are non-technology analysts in year t. We classify a Seeking Alpha analysts as a non-technology analyst if none of the analyst's reports most important topic is one of the technology topics 4 and 17. All 20 article topics are obtained from an LDA on the full sample of Seeking Alpha articles and presented in Appendix C. Source: Seeking Alpha Data
Earnings-focused CSMAs	The number of common Seeking Alpha analysts that are earnings analysts in year t. We classify a Seeking Alpha analysts as an earnings analyst if at least one of the analyst's reports most important topic is the earnings topic 10. All 20 article topics are obtained from an LDA on the full sample of Seeking Alpha articles and presented in Appendix C. Source: Seeking Alpha Data
Non-earnings-focused CSMAs	The number of common Seeking Alpha analysts that are non-earnings analysts in year t. We classify a Seeking Alpha analysts as a non-earnings analyst if none of the analyst's reports most important topic is the earnings topic 10. All 20 article topics are obtained from an LDA on the full sample of Seeking Alpha articles and presented in Appendix C. Source: Seeking Alpha Data
Trading-focused SMAs	The number of common Seeking Alpha analysts that are trading analysts in year t. We classify a Seeking Alpha analyst as a trading analyst if at least one of the analyst's reports most important topic is either trading topics 2 or 12. All 20 article topics are obtained from an LDA on the full sample of Seeking Alpha articles and presented in Appendix C. Source: Seeking Alpha Data
Non-trading-focused CSMAs	The number of common Seeking Alpha analysts that are non-trading analysts in year t. We classify a Seeking Alpha analysts as a non-trading analyst if none of the analyst's reports most important topic is either trading topics 2 or 12. All 20 article topics are obtained from an LDA on the full sample of Seeking Alpha articles and presented in Appendix C. Source: Seeking Alpha Data
High following CSMAs	The number of common Seeking Alpha analysts that are high following Seeking Alpha analysts. We classify a Seeking Alpha analyst as a high following Seeking Alpha analyst if the number of followers exceed the 50th percentile across all Seeking Alpha analysts. The number of followers is measured at the end of the sample period. Source: Seeking Alpha Data

(Continued on next page)

Appendix B

(continued)

Low following CSMAs	The number of common Seeking Alpha analysts that are low following Seeking Alpha analysts. We classify a Seeking Alpha analyst as a low following Seeking Alpha analyst if the number of followers does not exceed the 50th percentile across all Seeking Alpha analysts. The number of followers is measured at the end of the sample period. Source: Seeking Alpha Data
High activity CSMAs	The number of common Seeking Alpha analysts that are high activity Seeking Alpha analysts for the firm pair in year t. We classify an Seeking Alpha analyst as a high activity Seeking Alpha analyst for citing firm i (cited firm j) in year t if the number of days at which the Seeking Alpha analyst exhibits forecast activity for citing firm i (cited firm j) in year t exceeds the 50th percentile. We then label an Seeking Alpha analyst as a high activity Seeking Alpha analyst for the firm pair if the Seeking Alpha analyst is a high activity Seeking Alpha analyst for both firms in year t. Source: Seeking Alpha Data
Medium activity CSMAs	The number of common Seeking Alpha analysts that are medium activity Seeking Alpha analysts for the firm pair in year t. We classify an Seeking Alpha analyst as a high activity Seeking Alpha analyst for citing firm i (cited firm j) in year t if the number of days at which the Seeking Alpha analyst exhibits forecast activity for citing firm i (cited firm j) in year t exceeds the 50th percentile. We classify an Seeking Alpha analyst as a low activity Seeking Alpha analyst for citing firm i (cited firm j) in year t if the number of days at which the Seeking Alpha analyst exhibits forecast activity for citing firm i (cited firm j) in year t does not exceed the 50th percentile. We then label an Seeking Alpha analyst as a medium activity Seeking Alpha analyst for the firm pair if the Seeking Alpha analyst is a high activity Seeking Alpha analyst for citing firm i (cited firm j) and a low activity Seeking Alpha analyst for the cited firm j (citing firm i) in year t. Source: Seeking Alpha Data
Low activity CSMAs	The number of common Seeking Alpha analysts that are low activity Seeking Alpha analysts for the firm pair in year t. We classify an Seeking Alpha analyst as a low activity Seeking Alpha analyst for citing firm i (cited firm j) in year t if the number of days at which the Seeking Alpha analyst exhibits forecast activity for citing firm i (cited firm j) in year t does not exceed the 50th percentile. We then label an Seeking Alpha analyst as a low activity Seeking Alpha analyst for the firm pair if the Seeking Alpha analyst is a low activity Seeking Alpha analyst for both firms in year t. Source: Seeking Alpha Data

(Continued on next page)

Appendix B
(continued)

High topic concentration CSMAs	The number of common Seeking Alpha analysts that are high topic concentration Seeking Alpha analysts in year t. To determine Seeking Alpha analyst's topic concentration we use the Herfindahl index and measure the topic concentration of Seeking Alpha analysts' coverage using the topic probabilities of coverage in year t. All 20 article topics are obtained from an LDA on the full sample of Seeking Alpha articles and presented in Appendix C. We classify Seeking Alpha analysts as high topic concentration Seeking Alpha analysts if the topic concentration exceeds the 50th percentile across all Seeking Alpha analysts. Source: Seeking Alpha Data
Low topic concentration CSMAs	The number of common Seeking Alpha analysts that are low topic concentration Seeking Alpha analysts in year t. To determine Seeking Alpha analyst's topic concentration we use the Herfindahl index and measure the topic concentration of Seeking Alpha analysts' coverage using the topic probabilities of coverage in year t. All 20 article topics are obtained from an LDA on the full sample of Seeking Alpha articles and presented in Appendix C. We classify Seeking Alpha analysts as low topic concentration Seeking Alpha analysts if the topic concentration does not exceed the 50th percentile across all Seeking Alpha analysts. Source: Seeking Alpha Data
Citing/cited SMAs ratio	Citing total Seeking Alpha analysts plus one divided by cited total Seeking Alpha analysts plus one in year t in thousands. Source: Seeking Alpha Data
Other firm-specific variables	
Common analysts	The number of common analysts of citing firm i and cited firm j in year t. Source: I/B/E/S Detail History
Common inventors	The number of inventors that moved from the cited firm to the citing firm over the last three years and year t. Source: Patentsview Data
Common alliance	An indicator variable that takes the value of one if the firm pair has initiated a strategic alliance over the last three years and year t, and zero otherwise. Source: SDC Platinum Data
Common board member	An indicator variable that takes the value of one if the firm pair shares at least one board member in year t. Source: Boardex data
Common auditor	An indicator variable that takes the value of one if the firm pair shares the same auditor in year t. Source: Audit analytics data
Common investors	The average of the Backus et al. (2021) measure of common ownership in year t. Source: Thomson/Refinitiv Institutional (13f) Holdings Data
Common Estimize analysts	The number of common Estimize analysts of citing firm i and cited firm j in year t. Source: Estimize Data

(Continued on next page)

Appendix B

(continued)

Business similarity	The Hoberg and Phillips (2010, 2016) similarity measure at the two-digit SIC level. Source: Hoberg and Phillips (2010, 2016) Industry Data
Technology similarity	The Jaffe and Trajtenberg (1996) similarity measure based on the share of patent portfolios that fall in the same technological classes. Technology similarity is calculated as follows: $$\frac{\sum_{c=1}^{C} P_{ict}P_{jct}}{\sqrt{(\sum_{c=1}^{C} P_{ict}^2)(\sum_{c=1}^{C} P_{jct}^2)}}$$ where P_{ict} is the number of patents held by firm i in class c in year t, and P_{jct} is the number of patents held by firm j in class c in year t and C is the total number of technological classes. Source: Kogan et al. (2017) Patent Data
Citing/cited asset ratio	Citing total assets divided by cited total assets in year t in thousands. Source: CRSP Compustat Merged Data
Citing/cited analysts ratio	Citing total analysts plus one divided by cited total analysts plus one in year t in thousands. Source: I/B/E/S Detail History Data
Both small cap	An indicator variable if the citing firm and the cited firm are classified as small cap firms in year t. We classify a firm as a small cap firm if the market capitalization in year t does not exceed 1 Billion Dollars. Source: CRSP Compustat Merged Data

Appendix C

Topic Overview

#	Topic Name	Top-5 Associated Words				
1	Idea	Get	Think	Say	Big	Thing
2	Trading 1: Portfolio	Investment	Fund	Buy	Portfolio	Risk
3	Media	Game	Content	Subscriber	Service	Stream
4	Technology 1: Biotech	Drug	Treatment	Approval	Trial	Study
5	Automotive	Car	Demand	Vehicle	Production	Fuel
6	Macrofinance	Financial	State	Government	Public	Information
7	Commodities 1: Oil & Gas	Oil	Production	Gas	Barrel	Crude
8	Investment in PPE	Contract	Construction	Development	Facility	Rig
9	Social Media	User	Mobile	Platform	Ad	Search
10	Earnings	Income	Expense	Margin	Loss	Cost
11	Macroeconomy	Percent	Economy	Economic	Rise	Demand
12	Trading 2: Return	Dividend	Yield	Return	Rate	Valuation
13	Sell-side Analysts	Estimate	Analyst	Guidance	Expectation	Target
14	Smart Devices	Device	Phone	Product	Smartphone	Tablet
15	Real Estate	Bank	Rate	Mortgage	Interest	REIT
16	Commodities 2: Metals	Gold	Production	Metal	Mine	Ounce
17	Technology 2: Software	Technology	Software	Customer	System	Service
18	Retail	Store	Brand	Retail	Consumer	Customer
19	Capital & Financing	Debt	Asset	Management	Deal	Capital
20	Growth	Grow	Margin	Product	Industry	Segment

Notes: This table shows the twenty topics commonly discussed in our sample of Seeking Alpha articles. The twenty topics are identified with Latent Dirichlet allocation (LDA) using the full sample of articles, and named based on the five most common words associated with the topic.

Figure 1

Illustration of the Main Specification

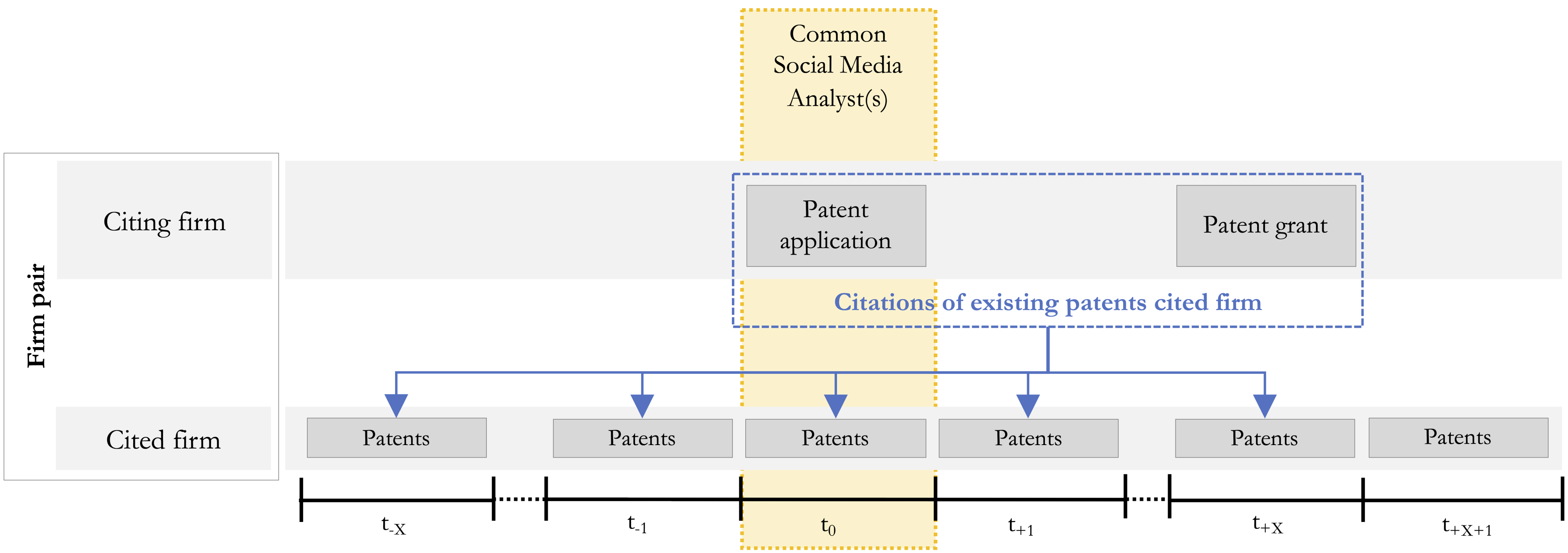

Table 1

Descriptive Statistics for Portfolio Structures of Social Media Analysts and Sell-side Analysts

Panel A: General coverage descriptives

	N	Mean	S.D.	Min.	1rst	25th	Med.	75th	99th	Max.
Firm-year analyst coverage										
Social media analysts	53,678	2.72	8.49	0	0	0	1	2	35	349
Sell-side analysts	53,678	8.80	9.00	0	0	0	6	13	39	70
Social media analysts										
Unique firms covered	10,129	11.58	35.57	1	1	1	3	8	145	1,294
Unique firms covered per year	20,440	7.65	19.29	1	1	1	2	6	88	573
Number of articles	10,129	26.06	119.33	1	1	1	3	12	439	4,654
Number of articles per year	20,440	12,91	39.56	1	1	1	3	8	188	1,450
Sell-side analysts										
Unique firms covered	14,855	11.52	14.95	1	1	2	5	16	64	253
Unique firms covered per year	64,833	9.05	8.74	1	1	2	6	15	35	157

Panel B: Analysts' coverage portfolio connectivity

	N	Mean	Std	Min.	1rst	25th	Med.	75th	99th	Max.
Common SMAs-based connectivity	53,678	2.23	4.00	0.00	0.00	0.00	0.14	2.83	17.18	38.14
Common analysts-based connectivity	53,678	1.32	1.10	0.00	0.00	0.42	1.22	1.99	4.32	7.62
Common SMAs- and *Common analysts*-based connectivity	53,678	3.42	4.10	0.02	0.04	0.93	1.87	4.01	18.73	39.47
Incremental *Common SMAs*-based connectivity	53,678	2.10	3.70	0.00	0.00	0.00	0.12	2.65	16.10	35.71

(Continued on next page)

Table 1

[continued]

Panel C: Distribution of common analysts across firm-pair industries (Top-20)									
Firm A		Firm B		Firm A citing Firm B			Firm B citing Firm A		
SIC	Industry	SIC	Industry	N	*Common SMAs*	*Common analysts*	N	*Common SMAs*	*Common analysts*
78	Motion Pictures	60	Depository Institutions	26	13.27	0.00	26	13.27	0.00
53	General Merchandise Stores	59	Miscellaneous Retail	52	13.27	3.23	52	11.44	2.77
59	Miscellaneous Retail	29	Petroleum & Coal Products	26	13.12	0.00	26	13.12	0.00
29	Petroleum & Coal Products	60	Depository Institutions	24	11.67	0.00	50	11.14	0.00
29	Petroleum & Coal Products	48	Communications	42	9.93	0.00	50	7.24	0.00
53	General Merchandise Stores	99	Non-Classifiable Establishments	53	9.83	0.06	78	6.90	0.04
53	General Merchandise Stores	60	Depository Institutions	86	9.22	0.05	115	9.04	0.05
36	Electronic & Other Electric Equipment	78	Motion Pictures	107	8.75	0.51	166	7.70	0.38
59	Miscellaneous Retail	60	Depository Institutions	118	8.62	0.01	139	7.90	0.01
59	Miscellaneous Retail	99	Non-Classifiable Establishments	59	8.46	0.03	67	7.34	0.03
99	Non-Classifiable Establishments	60	Depository Institutions	94	8.40	0.05	148	7.95	0.05
73	Business Services	78	Motion Pictures	194	8.13	2.28	304	5.75	1.67
29	Petroleum & Coal Products	99	Non-Classifiable Establishments	78	8.00	0.04	101	5.81	0.12
99	Non-Classifiable Establishments	21	Tobacco Products	26	7.92	0.23	26	7.92	0.23
58	Eating & Drinking Places	36	Electronic & Other Electric Equipment	31	7.68	0.19	48	1.12	0.12
73	Business Services	21	Tobacco Products	42	7.43	0.00	70	4.60	0.00
23	Apparel & Other Textile Products	73	Business Services	45	7.16	0.00	25	5.00	0.00
20	Food & Kindred Products	58	Eating & Drinking Places	47	6.57	1.32	53	4.89	1.26
29	Petroleum & Coal Products	29	Petroleum & Coal Products	112	6.21	7.78	112	6.21	7.78
78	Motion Pictures	48	Communications	81	5.98	2.86	99	5.75	2.08
57	Furniture & Homefurnishings Stores	37	Transportation Equipment	36	5.72	0.00	49	2.24	0.02
73	Business Services	53	General Merchandise Stores	438	5.58	0.05	609	4.13	0.06
53	General Merchandise Stores	48	Communications	134	5.24	0.01	212	3.86	0.02
60	Depository Institutions	60	Depository Institutions	586	5.00	7.96	586	5.00	7.96
57	Furniture & Homefurnishings Stores	35	Industrial Machinery & Equipment	55	4.85	0.18	152	0.89	0.02

Notes: Table 1, panel A, reports the mean number of *Common SMAs* and *Common analysts* by citing industry and cited industry. Entries represent the 20 industry pairs with the largest mean in terms of common social media analysts coverage in our sample. Panel B reports descriptive statistics on the firm-pair, the SMA and the sell-side analyst level. Panel C reports descriptive statistics for analysts' coverage portfolio connectivity, i.e., the percentage of firms with analyst coverage to which firm i is connected through its own analyst coverage. Please refer to Appendix A for a full description of all variables.

Table 2

Descriptive statistics

	N	Mean	Std.	Min	1rst	25th	Median	75th	99th	Max
Citations	593,969	6.51	93.44	0.00	0.00	0.00	0.00	2.00	102.00	37,222
Common SMAs	593,969	0.88	3.03	0.00	0.00	0.00	0.00	1.00	14.00	121.00
Control variables										
Common analysts	593,969	0.72	2.19	0.00	0.00	0.00	0.00	0.00	13.00	13.00
Common investors	593,969	0.36	0.35	0.00	0.00	0.00	0.31	0.59	1.34	1.34
Common inventors	593,969	0.10	0.37	0.00	0.00	0.00	0.00	0.00	2.00	2.00
Common alliance	593,969	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Common board member	593,969	0.01	0.10	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Common auditor	593,969	0.20	0.40	0.00	0.00	0.00	0.00	0.00	1.00	1.00
Business similarity	593,969	0.05	0.05	0.00	0.00	0.01	0.04	0.08	0.23	0.23
Technology similarity	593,969	0.23	0.27	0.00	0.00	0.03	0.12	0.35	0.97	0.97
Citing/cited analysts ratio	593,969	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.03
Citing/cited SMAs ratio	593,969	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.09	0.09
Citing/cited total assets ratio	593,969	0.09	0.43	0.00	0.00	0.00	0.00	0.01	2.95	3.79
Citing/cited patent stock ratio	593,969	0.07	0.32	0.00	0.00	0.00	0.00	0.01	1.80	3.24
Citing/cited citation ratio	593,969	0.60	2.60	0.00	0.00	0.00	0.00	0.04	18.82	20.44

Notes: Please refer to Appendix A for a full description of all variables.

Table 3

Common Social Media Analysts and Interfirm Patent Citations

	Citations		
	(1)	(2)	(3)
Common SMAs	0.0089***	0.0089***	0.0026**
	(0.0023)	(0.0023)	(0.0012)
Common analysts		0.0084***	0.0048**
		(0.0030)	(0.0022)
Other control variables			
Common investors		0.0564	0.0070
		(0.0396)	(0.0263)
Common inventors		0.2094***	0.0559***
		(0.0105)	(0.0062)
Common alliance		0.1737***	0.0703**
		(0.0425)	(0.0307)
Common board member		0.0291	0.0267
		(0.0397)	(0.0341)
Common auditor		0.0112	-0.0203
		(0.0234)	(0.0182)
Business similarity		1.238***	0.4865***
		(0.2330)	(0.1557)
Technology similarity		0.8821***	0.2584***
		(0.0947)	(0.0727)
Citing/cited analysts ratio		5.089*	2.060
		(3.019)	(3.966)
Citing/cited SMAs ratio		2.136***	0.5431
		(0.8023)	(0.4783)
Citing/cited total assets ratio		-0.0359	-0.0260
		(0.0312)	(0.0299)
Citing/cited patent stock ratio		-0.5809***	-0.0537
		(0.0458)	(0.0351)
Citing/cited citations ratio		0.0273***	-0.0021
		(0.0031)	(0.0025)
Firm-pair fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	No
Citing-firm × Year fixed effects	No	No	Yes
Cited-firm × Year fixed effects	No	No	Yes
N	593,969	593,969	593,969
Pseudo R^2	0.28209	0.28365	0.34237

Notes: Table 3 presents results for the relation between the number of cross-citations and *Common SMAs* based on negative binomial regressions. Three-way clustered standard errors by firm-pair, citing firm x year and cited firm x year in parentheses. $*$, $**$, and $***$ represent significance at the 10 percent, 5 percent, and 1 percent level, respectively. Please refer to Appendix A for a full description of all variables.

Table 4

Approach to Identification: Incentives to Cover Small Cap Firms

	First stage	Second stage
	Common SMAs	*Log(1 + Citations)*
	(1)	(2)
Common SMAs (instrumented)		0.0053***
		(0.0014)
Both small cap	-0.2322***	
	(0.0380)	
Both small cap × Post	0.1904**	
	(0.0744)	
Other control variables		
Common analysts	0.0015	0.0116***
	(0.0043)	(0.0016)
Common investors	-0.1785***	-0.0062
	(0.0347)	(0.0102)
Common inventors	0.1515***	0.0969***
	(0.0158)	(0.0053)
Common alliance	-0.0068	0.0760***
	(0.0821)	(0.0258)
Common board member	-0.0095	0.0168
	(0.0484)	(0.0157)
Common auditor	0.0162	-0.0074
	(0.0182)	(0.0065)
Business similarity	-0.1514	0.3374***
	(0.1753)	(0.0701)
Technology similarity	0.1328**	0.2237***
	(0.0602)	(0.0302)
Citing/cited analysts ratio	-11.91	-1.014
	(8.535)	(1.329)
Citing/cited SMAs ratio	-63.26***	0.3330
	(3.154)	(0.2255)
Citing/cited total assets ratio	-0.0499	-0.0188**
	(0.0447)	(0.0094)
Citing/cited patent stock ratio	-0.0575	-0.0757***
	(0.0362)	(0.0150)
Citing/cited citations ratio	-0.0154***	-0.0077***
	(0.0040)	(0.0011)
Lewbel instruments	Yes	No
Firm-pair fixed effects	Yes	Yes
Citing-firm × Year fixed effects	Yes	Yes
Cited-firm × Year fixed effects	Yes	Yes
N	593,969	593,969
Adjusted R^2	0.88394	0.74617
F-test	34,566	

Notes: Table 4 presents results for the relation between the number of cross-citations and *Common SMAs (instrumented)* based on an OLS regression. Three-way clustered standard errors by firm-pair, citing firm x year and cited firm x year in parentheses. $*$, $**$, and $***$ represent significance at the 10 percent, 5 percent, and 1 percent level, respectively. Please refer to Appendix A for a full description of all variables.

Table 5

Characteristics of Social Media Analysts and Interfirm Patent Citations

Panel A: Descriptive statistics

	N	Mean	Std.
Experience			
High experience CSMAs	593,969	0.06	0.36
Regular experience CSMAs	593,969	0.82	2.85
Activity level			
High activity CSMAs	593,969	0.26	1.19
Medium activity CSMAs	593,969	0.35	1.26
Low activity CSMAs	593,969	0.26	0.92
Portfolio size			
Large portfolio CSMAs	593,969	0.82	2.67
Regular portfolio CSMAs	593,969	0.06	0.54
Industry specialization			
High specialisation CSMAs	593,969	0.00	0.09
Low specialisation CSMAs	593,969	0.87	3.02

Panel B: Regression analysis

	Citations			
	(1)	(2)	(3)	(4)
Experience				
High experience CSMAs	0.0067			
	(0.0071)			
Regular experience CSMAs	0.0023*			
	(0.0013)			
Activity level				
High activity CSMAs		0.0021		
		(0.0024)		
Medium activity CSMAs		0.0051**		
		(0.0023)		
Low activity CSMAs		-0.0003		
		(0.0026)		
Portfolio size				
Large portfolio CSMAs			0.0025*	
			(0.0015)	
Regular portfolio CSMAs			0.0030	
			(0.0039)	
Industry specialization				
High specialisation CSMAs				-0.0600*
				(0.0323)
Low specialisation CSMAs				0.0028**
				(0.0012)
Wald test 1. coef. = 2. coef. [p-value]	*[0.0974]**	*[0.0411]***	*[n.s.]*	*[0.0152]***
Wald test 1. coef. = 3. coef. [p-value]		*[n.s.]*		
Wald test 2. coef. = 3. coef. [p-value]		*[0.0943]**		

(Continued on next page)

Table 5
[continued]

	Citations			
	(1)	(2)	(3)	(4)
Control variables				
Common analysts	0.0049**	0.0048**	0.0048**	0.0045**
	(0.0022)	(0.0022)	(0.0022)	(0.0022)
Common investors	0.0071	0.0069	0.0070	0.0071
	(0.0263)	(0.0263)	(0.0263)	(0.0263)
Common inventors	0.0559***	0.0559***	0.0559***	0.0559***
	(0.0062)	(0.0062)	(0.0062)	(0.0062)
Common alliance	0.0700**	0.0701**	0.0703**	0.0702**
	(0.0307)	(0.0307)	(0.0307)	(0.0306)
Common board member	0.0268	0.0267	0.0267	0.0271
	(0.0341)	(0.0341)	(0.0341)	(0.0341)
Common auditor	-0.0203	-0.0202	-0.0203	-0.0202
	(0.0182)	(0.0182)	(0.0182)	(0.0182)
Business similarity	0.4852***	0.4858***	0.4864***	0.4898***
	(0.1558)	(0.1557)	(0.1557)	(0.1556)
Technology similarity	0.2579***	0.2585***	0.2583***	0.2596***
	(0.0728)	(0.0727)	(0.0727)	(0.0727)
Citing/cited analysts ratio	1.965	2.026	2.054	2.107
	(3.970)	(3.957)	(3.968)	(3.958)
Citing/cited SMAs ratio	0.5447	0.5508	0.5405	0.5477
	(0.4782)	(0.4790)	(0.4802)	(0.4783)
Citing/cited total assets ratio	-0.0259	-0.0260	-0.0260	-0.0262
	(0.0300)	(0.0299)	(0.0300)	(0.0299)
Citing/cited patent stock ratio	-0.0537	-0.0535	-0.0537	-0.0539
	(0.0351)	(0.0351)	(0.0351)	(0.0350)
Citing/cited citations ratio	-0.0020	-0.0021	-0.0021	-0.0021
	(0.0025)	(0.0025)	(0.0025)	(0.0025)
Firm-pair fixed effects	Yes	Yes	Yes	Yes
Citing-firm × Year fixed effects	Yes	Yes	Yes	Yes
Cited-firm × Year fixed effects	Yes	Yes	Yes	Yes
N	593,969	593,969	593,969	593,969
Pseudo R^2	0.34237	0.34237	0.34237	0.34237

Notes: Table 5 presents results for the relation between the number of cross-citations and various characteristics of common social media analysts (CSMAs) based on negative binomial regressions. Three-way clustered standard errors by firm-pair, citing firm x year and cited firm x year in parentheses. $*$, $**$, and $***$ represent significance at the 10 percent, 5 percent, and 1 percent level, respectively. Please refer to Appendix A for a full description of all variables.

Table 6

Characteristics of Contributed Content and Interfirm Patent Citations

Panel A: Descriptive statistics

	N	Mean	Std.
Topic concentration			
High topic concentration CSMAs	593,969	0.64	2.13
Low topic concentration CSMAs	593,969	0.24	1.09
Content similarity			
High content similarity CSMAs	593,969	0.69	2.81
Low content similarity CSMAs	593,969	0.19	0.73
Content specificity			
High content specificity CSMAs	593,969	0.27	0.79
Low content specificity CSMAs	593,969	0.61	2.55
Technology-focus			
Technology-focused CSMAs	593,969	0.79	2.64
Non-technology-focused CSMAs	593,969	0.09	0.53
Earnings-focus			
Earnings-focused CSMAs	593,969	0.59	1.99
Non-earnings-focused CSMAs	593,969	0.29	1.23
Trading-focus			
Trading-focused CSMAs	593,969	0.74	2.47
Non-trading-focused CSMAs	593,969	0.14	0.71

(Continued on next page)

Table 6

[continued]

Panel B: Regression analysis						
	Citations					
	(1)	(2)	(3)	(4)	(5)	(6)
Topic concentration						
High topic concentration CSMAs	0.0051***					
	(0.0019)					
Low topic concentration CSMAs	-0.0018					
	(0.0028)					
Content similarity						
High content similarity CSMAs		0.0029**				
		(0.0012)				
Low content similarity CSMAs		0.0185***				
		(0.0047)				
Content specificity						
High content specificity CSMAs			0.0109**			
			(0.0045)			
Low content specificity CSMAs			0.0015			
			(0.0014)			
Technology-focus						
Technology-focused CSMAs				0.0036**		
				(0.0015)		
Non-technology-focused CSMAs				-0.0022		
				(0.0040)		
Earnings-focus						
Earnings-focused CSMAs					0.0011	
					(0.0017)	
Non-earnings-focused CSMAs					0.0052**	
					(0.0025)	
Trading-focus						
Trading-focused CSMAs						0.0016
						(0.0015)
Non-trading-focused CSMAs						0.0062*
						(0.0035)
Wald test 1. coef. = 2. coef. [p-value]	*[0.0211]***	*[< 0.001]****	*[0.0124]***	*[0.0486]***	*[0.053]**	*[0.0536]**
Control variables						
Common analysts	0.0046**	0.0047**	0.0047**	0.0048**	0.0048**	0.0048**
	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0022)
Common investors	0.0064	0.0055	0.0064	0.0069	0.0069	0.0069
	(0.0263)	(0.0263)	(0.0263)	(0.0263)	(0.0263)	(0.0263)
Common inventors	0.0560***	0.0562***	0.0561***	0.0559***	0.0560***	0.0560***
	(0.0062)	(0.0063)	(0.0062)	(0.0062)	(0.0062)	(0.0062)
Common alliance	0.0711**	0.0718**	0.0715**	0.0705**	0.0704**	0.0705**
	(0.0307)	(0.0306)	(0.0306)	(0.0307)	(0.0307)	(0.0307)
Common board member	0.0266	0.0260	0.0266	0.0270	0.0267	0.0265
	(0.0341)	(0.0341)	(0.0341)	(0.0341)	(0.0341)	(0.0341)

(Continued on next page)

Table 6

[continued]

	Citations					
	(1)	(2)	(3)	(4)	(5	(6)
Common auditor	-0.0202	-0.0197	-0.0200	-0.0202	-0.0202	-0.0203
	(0.0182)	(0.0182)	(0.0182)	(0.0182)	(0.0182)	(0.0182)
Business similarity	0.4865***	0.4835***	0.4854***	0.4881***	0.4853***	0.4838***
	(0.1556)	(0.1554)	(0.1557)	(0.1557)	(0.1557)	(0.1558)
Technology similarity	0.2591***	0.2616***	0.2595***	0.2585***	0.2587***	0.2583***
	(0.0728)	(0.0729)	(0.0728)	(0.0727)	(0.0727)	(0.0727)
Citing/cited analysts ratio	2.246	2.414	2.333	2.132	2.130	2.057
	(3.953)	(3.937)	(3.956)	(3.966)	(3.971)	(3.969)
Citing/cited SMAs ratio	0.5426	0.5698	0.5569	0.5608	0.5416	0.5428
	(0.4783)	(0.4777)	(0.4786)	(0.4797)	(0.4781)	(0.4780)
Citing/cited total assets ratio	-0.0263	-0.0275	-0.0266	-0.0260	-0.0261	-0.0260
	(0.0299)	(0.0299)	(0.0299)	(0.0299)	(0.0299)	(0.0300)
Citing/cited patent stock ratio	-0.0537	-0.0547	-0.0538	-0.0536	-0.0538	-0.0536
	(0.0351)	(0.0351)	(0.0351)	(0.0351)	(0.0351)	(0.0351)
Citing/cited citations ratio	-0.0021	-0.0023	-0.0022	-0.0021	-0.0021	-0.0021
	(0.0025)	(0.0025)	(0.0025)	(0.0025)	(0.0025)	(0.0025)
Firm-pair fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Citing-firm × Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Cited-firm × Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	593,969	593,969	593,969	593,969	593,969	593,969
Pseudo R^2	0.34237	0.34238	0.34237	0.34237	0.34237	0.34237

Notes: Table 6 presents results for the relation between the number of cross-citations and various SMA-characteristics based on negative binomial regressions. Three-way clustered standard errors by firm-pair, citing firm x year and cited firm x year in parentheses. $*$, $**$, and $***$ represent significance at the 10 percent, 5 percent, and 1 percent level, respectively. Please refer to Appendix A for a full description of all variables.

Table 7

Information Spillovers versus Coverage Decisions:
Common Social Media Analysts Controlling for Common Estimize Coverage

	Citations		
	(1)	(2)	(3)
Common SMAs	0.0026*		0.0026*
	(0.0014)		(0.0014)
Common Estimize analysts		5.57×10^{-7}	-1.48×10^{-5}
		(0.0001)	(0.0001)
Control variables			
Common analysts	0.0041*	0.0040*	0.0041*
	(0.0024)	(0.0024)	(0.0024)
Common investors	-0.0559	-0.0581	-0.0560
	(0.0401)	(0.0398)	(0.0401)
Common inventors	0.0328***	0.0328***	0.0328***
	(0.0078)	(0.0078)	(0.0078)
Common alliance	0.0982*	0.0924*	0.0981*
	(0.0519)	(0.0522)	(0.0519)
Common board member	-0.0463	-0.0452	-0.0463
	(0.0462)	(0.0463)	(0.0462)
Common auditor	-0.0119	-0.0121	-0.0119
	(0.0245)	(0.0246)	(0.0245)
Business similarity	0.2564	0.2632	0.2568
	(0.2118)	(0.2125)	(0.2118)
Technology similarity	0.2453**	0.2437**	0.2452**
	(0.1139)	(0.1139)	(0.1138)
Citing/cited analysts ratio	7.468*	7.518*	7.472*
	(4.481)	(4.536)	(4.483)
Citing/cited SMAs ratio	0.6653	0.3962	0.6654
	(0.5635)	(0.5578)	(0.5635)
Citing/cited total assets ratio	-0.0356	-0.0338	-0.0357
	(0.0456)	(0.0456)	(0.0456)
Citing/cited patent stock ratio	-0.1067	-0.1075	-0.1067
	(0.0688)	(0.0688)	(0.0688)
Citing/cited citations ratio	-0.0025	-0.0023	-0.0025
	(0.0035)	(0.0035)	(0.0035)
Firm-pair fixed effects	Yes	Yes	Yes
Citing-firm × Year fixed effects	Yes	Yes	Yes
Cited-firm × Year fixed effects	Yes	Yes	Yes
N	206,335	206,335	206,335
Pseudo R^2	0.36979	0.36979	0.36979

Notes: Table 7 presents results for the relation between the number of cross-citations and *Common Estimize* based on negative binomial regressions. Three-way clustered standard errors by firm-pair, citing firm x year and cited firm x year in parentheses. $*$, $**$, and $***$ represent significance at the 10 percent, 5 percent, and 1 percent level, respectively. Please refer to Appendix A for a full description of all variables.

Table 8

The Role of Social Media Analysts' Credibility and Visibility for Information Spillovers

Panel A: Descriptive statistics

	N	Mean	Std.
Anonymity			
Non-anonymous CSMAs	593,969	0.69	2.25
Anonymous CSMAs	593,969	0.19	0.93
Followers			
High following CSMAs	593,969	0.32	1.35
Low following CSMAs	593,969	0.56	1.83

Panel B: Regression analysis

	Citations	
	(1)	(2)
Anonymity		
Non-anonymous CSMAs	0.0059***	
	(0.0019)	
Anonymous CSMAs	-0.0042	
	(0.0033)	
Followers		
High following CSMAs		0.0055**
		(0.0025)
Low following CSMAs		-0.0002
		(0.0023)
Wald test 1. coef. = 2. coef. [p-value]	*[0.0064]****	*[0.0502]***
Control variables		
Common analysts	0.0046**	0.0048**
	(0.0022)	(0.0022)
Common investors	0.0066	0.0071
	(0.0263)	(0.0263)
Common inventors	0.0560***	0.0560***
	(0.0062)	(0.0062)
Common alliance	0.0707**	0.0697**
	(0.0306)	(0.0307)
Common board member	0.0267	0.0273
	(0.0341)	(0.0341)
Common auditor	-0.0200	-0.0203
	(0.0182)	(0.0182)
Business similarity	0.4872***	0.4871***
	(0.1556)	(0.1558)
Technology similarity	0.2598***	0.2581***
	(0.0728)	(0.0727)

(Continued on next page)

Table 8
[continued]

	Citations	
	(1)	(2)
Citing/cited analysts ratio	2.230	2.033
	(3.955)	(3.967)
Citing/cited SMAs ratio	0.5703	0.5228
	(0.4782)	(0.4788)
Citing/cited total assets ratio	-0.0265	-0.0258
	(0.0299)	(0.0300)
Citing/cited patent stock ratio	-0.0535	-0.0536
	(0.0351)	(0.0351)
Citing/cited citations ratio	-0.0021	-0.0020
	(0.0025)	(0.0025)
Firm-pair fixed effects	Yes	Yes
Citing-firm × Year fixed effects	Yes	Yes
Cited-firm × Year fixed effects	Yes	Yes
N	593,969	593,969
Pseudo R^2	0.34237	0.34237

Notes: Table 8 presents results for the relation between the number of cross-citations and various SMA-characteristics based on negative binomial regressions. Three-way clustered standard errors by firm-pair, citing firm x year and cited firm x year in parentheses. $*$, $**$, and $***$ represent significance at the 10 percent, 5 percent, and 1 percent level, respectively. Please refer to Appendix A for a full description of all variables.

Social Media Analysts, Managerial Learning, and Corporate Innovation

ONLINE APPENDIX

May 2024

Table OA.1

Common Social Media Analysts and Interfirm Patent Citations: Alternative Model Specifications

	OLS		*Logit*
	Log(1 + Citations)	*Citation indicator*	*Citation indicator*
	(1)	(2)	(3)
Common SMAs	0.0058***	0.0008**	0.0088**
	(0.0010)	(0.0004)	(0.0043)
Control variables			
Common analysts	0.0116***	0.0012*	0.0183***
	(0.0016)	(0.0007)	(0.0064)
Common investors	-0.0061	-0.0052	-0.0641
	(0.0102)	(0.0063)	(0.0545)
Common inventors	0.0969***	0.0114***	0.1466***
	(0.0053)	(0.0021)	(0.0208)
Common alliance	0.0760***	0.0138	0.3284***
	(0.0258)	(0.0104)	(0.1270)
Common board member	0.0168	-0.0039	-0.0163
	(0.0157)	(0.0083)	(0.0736)
Common auditor	-0.0074	-0.0031	-0.0348
	(0.0065)	(0.0040)	(0.0346)
Business similarity	0.3375***	0.1739***	1.494***
	(0.0701)	(0.0393)	(0.3426)
Technology similarity	0.2237***	0.0643***	0.5836***
	(0.0302)	(0.0139)	(0.1348)
Citing/cited analysts ratio	-0.9981	1.158	8.336
	(1.327)	(0.9209)	(7.572)
Citing/cited SMAs ratio	0.3764*	0.2669**	2.728***
	(0.2104)	(0.1207)	(1.025)
Citing/cited total assets ratio	-0.0188**	-0.0018	-0.0222
	(0.0094)	(0.0055)	(0.0476)
Citing/cited patent stock ratio	-0.0757***	-0.0213***	-0.0720
	(0.0150)	(0.0064)	(0.0688)
Citing/cited citations ratio	-0.0077***	0.0003	0.0041
	(0.0011)	(0.0005)	(0.0042)
Firm-pair fixed effects	Yes	Yes	Yes
Citing-firm × Year fixed effects	Yes	Yes	Yes
Cited-firm × Year fixed effects	Yes	Yes	Yes
N	593,969	593,969	532,825
Adjusted R^2	0.74617	0.38346	
Pseudo R^2	0.52472	0.45612	0.34708

Notes: Table OA.1 presents results for the relation between the number of cross-citations and *Common SMAs* based on OLS and Logit regressions. Three-way clustered standard errors by firm-pair, citing firm x year and cited firm x year in parentheses. $*$, $**$, and $***$ represent significance at the 10 percent, 5 percent, and 1 percent level, respectively. Please refer to Appendix A for a full description of all variables.

Table OA.2

Common Social Media Analysts and Interfirm Patent Citations: Dropping Observations with a Large Number of Citations

	Citations
	(1)
Common SMAs	0.0030**
	(0.0015)
Control variables	
Common analysts	0.0047**
	(0.0023)
Common investors	0.0198
	(0.0271)
Common inventors	0.0575***
	(0.0064)
Common alliance	0.0705**
	(0.0332)
Common board member	0.0311
	(0.0348)
Common auditor	-0.0165
	(0.0182)
Business similarity	0.5141***
	(0.1634)
Technology similarity	0.2664***
	(0.0718)
Citing/cited analysts ratio	0.9193
	(3.895)
Citing/cited SMAs ratio	0.4445
	(0.4817)
Citing/cited total assets ratio	-0.0343
	(0.0289)
Citing/cited patent stock ratio	-0.0554
	(0.0350)
Citing/cited citations ratio	0.0003
	(0.0024)
Firm-pair fixed effects	Yes
Citing-firm × Year fixed effects	Yes
Cited-firm × Year fixed effects	Yes
N	588,041
Pseudo R^2	0.30583

Notes: Table OA.2 presents results for the relation between the number of cross-citations and *Common SMAs* based on negative binomial regressions. Three-way clustered standard errors by firm-pair, citing firm x year and cited firm x year in parentheses. $*$, $**$, and $***$ represent significance at the 10 percent, 5 percent, and 1 percent level, respectively. Please refer to Appendix A for a full description of all variables.

Table OA.3

Common Social Media Analysts and Interfirm Patent Citations: Alternative Specification with Lead Citations

	$Citations_{t+1}$		
	(1)	(2)	(3)
Common SMAs	0.0146***	0.0151***	0.0020*
	(0.0026)	(0.0027)	(0.0012)
Control variables			
Common analysts		0.0064*	0.0007
		(0.0035)	(0.0023)
Common investors		0.0912**	-0.0052
		(0.0413)	(0.0268)
Common inventors		0.1562***	0.0367***
		(0.0106)	(0.0060)
Common alliance		0.1034**	0.0342
		(0.0415)	(0.0298)
Common board member		0.0456	0.0267
		(0.0375)	(0.0324)
Common auditor		0.0089	-0.0219
		(0.0239)	(0.0187)
Business similarity		1.181***	0.5435***
		(0.2489)	(0.1608)
Technology similarity		0.5353***	0.0429
		(0.0947)	(0.0752)
Citing/cited analysts ratio		0.9912	-5.558
		(2.728)	(3.815)
Citing/cited SMAs ratio		3.476***	0.5393
		(0.8683)	(0.4938)
Citing/cited total assets ratio		-0.0423	-0.0082
		(0.0321)	(0.0315)
Citing/cited patent stock ratio		-0.5210***	-0.0708**
		(0.0433)	(0.0361)
Citing/cited citations ratio		0.0127***	-0.0004
		(0.0028)	(0.0025)
Firm-pair fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	No
Citing-firm × Year fixed effects	No	No	Yes
Cited-firm × Year fixed effects	No	No	Yes
N	593,969	593,969	593,969
Pseudo R^2	0.28405	0.28497	0.34884

Notes: Table OA.3 presents results for the relation between the number of lead cross-citations and *Common SMAs* based on negative binomial regressions. Three-way clustered standard errors by firm-pair, citing firm x year and cited firm x year in parentheses. $*$, $**$, and $***$ represent significance at the 10 percent, 5 percent, and 1 percent level, respectively. Please refer to Appendix A for a full description of all variables.

Table OA.4

Common Social Media Analysts and Interfirm Patent Citations: Alternative Clustering of Standard Errors

	Citations			
	Clustered by			
	Firm Pair	Citing Firm & Cited Firm	SIC2 Pair	Citing SIC2 & Cited SIC2
	(1)	(2)	(3)	(4)
Common SMAs	0.0026**	0.0026*	0.0026**	0.0026***
	(0.0011)	(0.0013)	(0.0012)	(0.0009)
Control variables				
Common analysts	0.0048***	0.0048*	0.0048**	0.0048**
	(0.0018)	(0.0024)	(0.0019)	(0.0020)
Common investors	0.0070	0.0070	0.0070	0.0070
	(0.0253)	(0.0272)	(0.0240)	(0.0126)
Common inventors	0.0559***	0.0559***	0.0559***	0.0559***
	(0.0059)	(0.0073)	(0.0064)	(0.0075)
Common alliance	0.0703**	0.0703**	0.0703**	0.0703**
	(0.0300)	(0.0355)	(0.0335)	(0.0337)
Common board member	0.0267	0.0267	0.0267	0.0267
	(0.0342)	(0.0352)	(0.0372)	(0.0377)
Common auditor	-0.0203	-0.0203	-0.0203	-0.0203
	(0.0178)	(0.0202)	(0.0214)	(0.0194)
Business similarity	0.4865***	0.4865***	0.4865***	0.4865***
	(0.1420)	(0.1834)	(0.1694)	(0.1380)
Technology similarity	0.2584***	0.2584**	0.2584***	0.2584**
	(0.0581)	(0.1043)	(0.0912)	(0.1035)
Citing/cited analysts ratio	2.060	2.060	2.060	2.060
	(3.450)	(4.252)	(3.802)	(3.745)
Citing/cited SMAs ratio	0.5431	0.5431	0.5431	0.5431
	(0.4737)	(0.3684)	(0.5262)	(0.5247)
Citing/cited total assets ratio	-0.0260	-0.0260	-0.0260	-0.0260**
	(0.0292)	(0.0283)	(0.0303)	(0.0121)
Citing/cited patent stock ratio	-0.0537	-0.0537	-0.0537	-0.0537
	(0.0334)	(0.0406)	(0.0405)	(0.0363)
Citing/cited citations ratio	-0.0021	-0.0021	-0.0021	-0.0021
	(0.0024)	(0.0028)	(0.0029)	(0.0030)
Firm-pair fixed effects	Yes	Yes	Yes	Yes
Citing-firm × Year fixed effects	Yes	Yes	Yes	Yes
Cited-firm × Year fixed effects	Yes	Yes	Yes	Yes
N	593,969	593,969	593,969	593,969
Pseudo R^2	0.34237	0.34237	0.34237	0.34237

Notes: Table OA.4 presents results for the relation between the number of cross-citations and *Common SMAs* based on negative binomial regressions with standard errors clustered in different ways. $*$, $**$, and $***$ represent significance at the 10 percent, 5 percent, and 1 percent level, respectively. Please refer to Appendix A for a full description of all variables.