FROM SINGLE TO SCALE: ESSAYS ON SOCIAL MEDIA,

ARTIFICIAL INTELLIGENCE AND START-UP SCALING

Xiaoning Wang

A DISSERTATION

in

Operations, Information and Decisions

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2023

Supervisor of Dissertation

Lynn Wu, Associate Professor of Operations, Information and Decisions

Graduate Group Chairperson

Nancy Zhang, Ge Li and Ning Zhao Professor,
Professor of Statistics and Data Science

Dissertation Committee

Lynn Wu, Associate Professor of Operations, Information and Decisions
Lorin Hitt, Professor of Operations, Information and Decisions
Serguei Netessine, Professor of Operations, Information and Decisions

FROM SINGLE TO SCALE: ESSAYS ON SOCIAL MEDIA, ARTIFICIAL INTELLIGENCE AND START-UP SCALING

ACKNOWLEDGEMENT

Time flies like an arrow, and in the blink of an eye, I have reached the moment of my doctoral graduation. I had prepared countless drafts in my mind for this acknowledgment, but when I finally put pen to paper, every word weighed heavily on me.

I have always considered myself fortunate on my academic journey. I have been guided by esteemed mentors, who have shone like stars and pointed the way forward for me. I have also received assistance from my helpful peers when I have been in difficulties or felt despondent. Of course, I cannot forget the understanding and support of my family and loved ones, who have been with me through it all. I do not possess the talent of my peers, but my greatest motivation to get to where I am today comes from those who are worthy of my lifelong gratitude.

Firstly, I would like to express my heartfelt thanks to each member of my dissertation committee, all of whom have provided invaluable guidance during my doctoral years and served as role models for me. I want to express my deepest gratitude to my advisor, Prof. Lynn Wu. It is your patient instruction that has transformed me from a state of complete academic ignorance to where I am today. Your sharp judgment on academic issues, rigorous academic attitude, and active learning spirit have been a source of inspiration for me throughout my doctoral years. I know I have been a troublemaker for you in the past years. I have taken every piece of your advice to heart and I am very grateful for your patience and support over the years.

I would also like to thank Prof. Lorin Hitt, who has always been there for me when I needed guidance. Your research seminar course not only broadened my research horizons, but also cultivated my research taste through discussions of numerous papers. I will never forget your words in the research seminar 5 years ago, "never forget why you come to Wharton", as well as your fatherly advice when I was working on my first paper revision. These teachings are my most precious spiritual wealth.

I am also immensely grateful to Serguei Netessine for agreeing to serve on my committee. Your personal charm has deeply attracted me, and in the future, I will continue to learn from you how to spice serious research with humor. You have always given me refreshing advice when I face major choices, helping me break free from the shackles of conventional thinking.

Beyond my committee, during my years of study at Wharton, I have been fortunate to have met an outstanding group of professors and peers (some of whom have already become professors). I am grateful to Prof. Sonny Tambe, Daniel Rock, Noah Gans, Santiago Gallino, Simone Marinesi, Sergei Savin, Hummy Song, Gad Allon, Joseph Simmons, Eric Clemons, Linda Zhao, Xu Cheng, Yuxin Chen, David Hsu, and other teachers for their help, which has been in the form of research assistance, career path choices, or life advice. I am also very grateful to my Ph.D. cohort, including Bowen Lou, Yi Liu, Atiye Cansu Erol, Xinyu Ma, Yuhao Fan, Zhongtian Chen, Christian Kaps, Arielle Anderer, Kan Xu, Anna Gao, Alex Miller, Amandeep Singh, Sirui Wang, Jiding Zhang, Dawson Kaaua, Berry Wang and Angel Chung. I cannot imagine surviving without your help.

Moreover, I want to take this opportunity to thank my family. Throughout my many years of education, my parents have always been my most solid support, providing me with the courage to continue pursuing higher academics and constantly exploring the unknown.

Finally, I want to give special thanks to my girlfriend, Chenxi. I am deeply indebted to you over the six years. Thank you for giving me confidence when I was disappointed and for giving me strength when I was lost. It is the nourishment from your love that has allowed me to thrive through these untrivial years, when there were pandemics, health issues, and other various challenges.

There are too many people I want to thank, and too many words of gratitude to express. Unfortunately, the space here is too small for me to write them all down. Thanks to these wonderful experiences, I will continue academia life after graduation, dedicating myself to scientific research and educating future generations.

As for the words left unwritten, I will leave it to the passing of time to complete the story.

ABSTRACT

FROM SINGLE TO SCALE: ESSAYS ON SOCIAL MEDIA,

ARTIFICIAL INTELLIGENCE AND START-UP SCALING

Xiaoning Wang

Lynn Wu

This dissertation, comprised of three essays, examines the influence of nascent information technologies including social media and artificial intelligence on facilitating the scaling process of entrepreneurial ventures. Through large-scale empirical analyses, this dissertation studies how social media and artificial intelligence can help start-ups improve financing outcomes, institutional environments and product performances, and whether such nascent information technologies will mitigate or exacerbate the existing disparities in the domain of entrepreneurship.

In the first essay, we examine whether social media use can improve funding outcomes for start-up firms, especially those founded by women and by other people also lacking connections to the investor network. Using Twitter data and start-up investment data from Crunchbase, we show that social media can mitigate some disparities in financing experienced by these firms through improving information access. We find this effect is stronger for first-time entrepreneurs than for experienced ones, stronger for attracting new investors than repeat ones, and stronger in more competitive markets.

In the second essay, using comprehensive start-up data in China, we find that social media activities can help companies that possess few political connections to obtain institutional resources. They are more likely to receive administrative approvals from the government to conduct their primary business activities. They are also more likely to receive positive media coverages. We find that social media can not only help these companies broadcast quality signals, but also direct public opinion and mobilize public support, all of which

are essential to establish public legitimacy.

In the third essay, we investigate the role of the lean start-up method (LSM) in explaining the impact of AI capability on product innovations in start-ups. We find that companies that adopt general AI capability create more innovative products every year, including both novel and incremental products. We also find that AI investments complement LSM in product innovations, and the complementarities depend on the type of product innovations and the type of AI used.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1: INTRODUCTION

This dissertation, comprised of three essays, examines the influence of nascent information technologies including social media and artificial intelligence on facilitating the scaling process of entrepreneurial ventures. Through large-scale empirical analyses, this dissertation studies how social media and artificial intelligence can help start-ups improve financing outcomes, institutional environments and product performances, and whether such nascent information technologies will mitigate or exacerbate the existing disparities in the domain of entrepreneurship.

Start-up scaling process represents the growth stage of entrepreneurial ventures between firm establishment and maturity, involving not only how to achieve replication at scale but also how to attend to the increased scope of activities that typically accompany growth (Chandler 1990, De Santola and Gulati 2017). While the effectiveness of information technologies in increasing firm productivity and profits is well documented, currently most literature focus on the mature stages of firms and few has examined the impact of information technologies in entrepreneurial scaling processes, especially for nascent information technologies such as social media and artificial intelligence. Such questions potentially include what are the benefits of social media and artificial intelligence, what extent to use and what kind of content to post, etc., and the strategic answers to these questions are contingent on different firm characteristics as what works for some firms may not be as effective for others.

Understanding this phenomenon is especially important not only due to a lack of large-sample empirical studies on entrepreneurial scaling strategies (De Santola and Gulati 2017), but also because existing literature has shown that both the adoption and returns to information and communication technologies are unequally distributed across firms (a “corporate digital divide”) (Bresnahan et al 2002, Forman 2005), and social media is also shown to have heterogenous levels of effectiveness on individuals and organizations (Wu and Kane 2021). Thus, it is critical to take firm-specific features and dynamics of development stages into consideration when evaluating corporate information technology strategies to uncover the nuance effects of how social media and artificial intelligence benefit the most. Moreover, a number of practitioner-oriented books and articles published in recent years unmask the anxieties entrepreneurs experience about the “black art” (Horowitz 2010) of managing their internal organizations through growth (Flamholtz and Randle 2015, Furr and Ahlstrom 2011). With the heightened technological and institutional disruptions globally in the recent and coming years (Wu et al. 2020), there is an urgent call for research in entrepreneurial scaling strategies from industry practitioners.

In the first essay, I examine whether social media use can improve financing outcomes for start-ups, especially those that are founded by women and that are constrained in the investor network, two groups of companies who traditionally have been less successful in securing VC financing. While the effect of social media on marketing outcomes is abundantly documented, it is unclear whether social media can also affect access to venture capital (VC) financing and if so, whether these effects can exacerbate or alleviate existing inequalities in financing. Examining this phenomenon is critical because the shift in the US economy away from traditional manufacturing and toward information technology (IT)-intensive services has increased the importance of innovation and entrepreneurship (Greenwood and Gopal 2015, Haltiwanger et al. 2013) which is reliant on venture financing.

Using social media activity data from Twitter and VC investment data of start-ups on Crunchbase, we show that social media can mitigate some disparities in financing experienced by these firms through improving information access. We find this effect is stronger for new entrepreneurs than for experienced ones, stronger for attracting new investors than existing ones, and stronger in competitive markets than in less competitive markets. Collectively, these results suggest that social media could primarily help women and less connected individuals obtain financing by alleviating information asymmetry between founders and investors.

In the second essay, we examine whether entrepreneurs can use social media technology to manage and improve their institutional environments. Institutional environments, including formal institutions such as regulatory regimes and informal institutions such as social norms, play an important role in entrepreneurial activities (Eesley et al. 2018). In the past few decades, while there have been many studies examining the influence of institutional environments on entrepreneurial behaviors (Eesley 2016), most of them regard the institutional environments as exogenous in business development. Accordingly, limited attention has been paid to how firms, especially start-ups, can conversely influence their external institutional environments.

Using comprehensive firm information and social media data from WeChat for about 2,000 Chinese start-ups from 2011 to 2020, we find that social media activities can help companies with few political connections to obtain institutional resources. They are more likely to receive administrative approvals from the government to conduct their primary business activities. They are also more likely to receive positive media coverages. We find that social media can not only help these companies broadcast quality signals, but also direct public opinions and mobilize public support in their favor, all of which are essential to establish public legitimacy. These effects are especially strong for firms that face greater institutional barriers and have scarce resources. Consistent with proposed mechanism, we find social media's effect is significantly reduced after implementing policies that lower institutional barriers, such as after the Mass Entrepreneurship and Innovation Policy.

In the third essay, we examine the impact of artificial intelligence on start-up product innovations. Despite the transformative potential AI has on business and innovation, many firms fail to reap any benefits from AI. We investigate one cause for why some firms successfully use AI to innovate and others fail to do so: the organizational support required to make AI perform well at two different tasks, novel and incremental innovation. Specifically, we examine the role of the lean start-up method (LSM) in explaining the impact of AI capability on product innovations at start-ups. We first identify each start-up's general AI capabilities based on the company's usage of AI. Using a dataset of some 2,000 start-ups in China from 2011 to 2020, we find that companies that adopt general AI capability create more innovative products every year, including both novel and incremental products. Next, we find that AI investments complement LSM in product innovations, and the complementarities depend on the type of product innovations and the type of AI used.

Specifically, we separate general AI capabilities into internally focused and externally focused, where internally focused AI refer to the ability to use AI to improve internal processes, and externally focused AI refer to the abilities to use AI to make sense of external market information. We find that LSM complements (1) externally focused AI in developing novel products but not incremental ones, and (2) internally focused AI in developing incremental products but not novel products. We show that the complementarities in 1 arise because AI can broaden search scope for market opportunities and LSM can validate these opportunities. Together they complement each other in alleviating market uncertainties in developing novel products. The complementarities in 2 come from leveraging LSM to generate experimental data that AI can then use to improve internal business processes. Together they speed up product iterations. These findings apply to companies developing software and to those developing hardware/physical products. Our research shows the importance of treating AI as a heterogeneous construct because different AI capabilities depend upon different organizational processes to succeed.

In summary, using large-scale start-up data from multiple sources and leveraging various econometric methodologies, we study how nascent information technologies, such as social media and artificial intelligence can facilitate the start-up scaling process in multiple perspectives, including financing, institution and product innovations. This dissertation extends the information systems literature by showing how social media and artificial intelligence can alleviate "digital divide" and promote diversity, and also contributes in the innovation and entrepreneurship literature by highlighting the importance of information flow in facilitating the scaling process of early-stage companies.

CHAPTER 2: SOCIAL MEDIA ALLEVIATES VENTURE CAPITAL FUNDING INEQUALITY FOR WOMEN AND LESS CONNECTED ENTREPRENEURS

1. Introduction

With over 2.8 billion monthly active users on Facebook and 330 million on Twitter (as of mid-2021),[1] social media provide an indispensable channel for information transmission, offering an unprecedented ability to efficiently broadcast information about goods and services (Bharadwaj et al. 2013, Bond et al. 2019, Rishika and Ramaprasad, 2019). While this effect on marketing outcomes is abundantly documented, it is unclear whether social media can be used as a tool to increase venture capital (VC) financing and if so, whether these effects can exacerbate or alleviate existing disparities in financing. Examining this phenomenon is critical because the shift in the US economy away from traditional manufacturing and toward information technology (IT)-intensive services has increased the importance of innovation and entrepreneurship, activities reliant on venture financing (Breznitz et al. 2018, Greenwood and Gopal 2015, Haltiwanger et al. 2013). Public policies have also recognized the role that social media can play in obtaining venture financing. In June 2015, the US Securities and Exchange Commission (SEC) issued guidance that allows start-ups to tweet directly to investors about their business. Accordingly, research in this area has not only become important as some companies have used Twitter to dramatic effect (notably Tesla, among others), but because the SEC has signaled that social media use by companies is an important area of policy concern. Furthermore, existing literature has shown that both the adoption and returns to information and communication technologies are unequally distributed across firms (a "corporate digital divide") (Bresnahan et al. 2002, Forman 2005). As social media play a potentially important role in obtaining private financing, it is critical to examine who can benefit from using social media (Wu and Kane

---

[1] Data resource: www.statista.com. Accessed on September 29th, 2021.

2021) and the effectiveness of public policies to regulate their use to obtain financing.

Understanding how social media can help start-ups obtain VC financing is especially important as disparities in venture financing are well recognized (Kanze et al. 2018) especially for women, underrepresented minorities and other founders lacking connections to venture capital (VC) networks (Koellinger et al. 2013, Eddleston et al. 2016, Balachandra et al. 2019). If there is a technological solution that could help alleviate some of these disparities in venture financing, it could be a powerful tool in broadening venture funding and facilitating entrepreneurship. We hypothesize that social media can broaden information access and thereby alleviate investors' information asymmetries about start-ups founded by underrepresented groups or those who lack social capital. Conversely, it is also plausible that the same technology could exacerbate existing inequalities (Wu and Kane 2021) as increased information access creates information overload, discouraging broad searches (Jones et al. 2004, Feng et al. 2015) and can encourage decision makers to place greater reliance on pre-existing social relationships (Van Alstyne and Brynjolfsson 2005, Lee et al. 2014). We explore these countervailing forces to examine the effect of social media use on venture financing. Our goal is two-fold. First, similar to prior work, we document whether social media can help secure venture financing, and if so, what the mechanisms are through which social media facilitates VC funding. More importantly, we explore whether social media can alleviate or exacerbate existing disparities faced by start-ups founded by women and by others who lack social capital.

Our analysis relies on data from Twitter to capture social media adoption and usage, and VC investment data of early-stage firms from Crunchbase from 2007 to 2016. The Crunchbase data also enables the construction of a time-varying venture-investor-relationship network in which nodes are start-ups and links represent common investors. With this network, we can use standard network position measures such as network degree and Burt's network constraint ("structural holes") metrics which have been shown to identify the position of network participants in the information flow to be either better or worse (e.g., Burt

1992). They can influence performance outcomes such as success in obtaining financing. Specifically, we use the network constraint measure to identify firms and founders that are not well positioned in the investor network and therefore have poor social capital. To examine gender disparities, we obtain demographic information on founders from Crunchbase which enables us to identify start-ups where the founders are solely or predominantly women.

Examining the VC financing performance of start-ups before and after their adoption and use of social media, we find evidence that social media can alleviate the funding disparities for firms with poor network capital and for firms founded by women who are able to use social media effectively. We find that the effect primarily comes from reduced information asymmetry about a start-up's product and service quality and the effect is stronger when firms actively engage on Twitter. These findings are robust to multiple approaches for addressing causality. Interestingly, we also find evidence that Twitter ceases to provide an incremental benefit to women founders once their position in the VC funding network has improved. This suggests that some of the disadvantage experienced by women entrepreneurs in relation to VC funding could be ameliorated through improving their positions in the network. Lastly, consistent with the mechanism of reducing information asymmetry, we find this effect of Twitter is stronger in more financially competitive markets and for first-time entrepreneurs where the marginal benefits of improved information are likely to be especially high. Overall, our results suggest that social media can help disadvantaged start-ups secure venture financing by alleviating information asymmetries associated with a poor network position that limits investor interest, which, in turn, creates a bounded but persistent benefit as these firms expand their network connections and position.

Auxiliary analysis indicates that our results are driven by the benefit of improved information flow in both quality and volume, rather than the self-selection, unobserved heterogeneity or endogeneity of Twitter use. The social media benefits we find are associated with usage but not adoption, and stronger for business-focused content. These results are consistent with social media being used to disseminate useful information, rather than due to selection biases or other types of reverse causality (e.g., "better" start-ups choosing to use social media). We also find that the benefits are greatest in situations where the potential benefits of information improvement are larger: social media is more effective in attracting new investors (as opposed to repeat investors) and for first-time entrepreneurs (as opposed to experienced entrepreneurs). We also discuss how the potential CEO hyperbole and even fraud on social media impact our results in the conclusion.

Overall, our results are consistent with social media improving the flow of information between start-ups and investors. These benefits are especially strong for women founders and others that are less likely to start with strong network connections in the VC community. Start-ups can experience long-term information benefits by effectively using social media to engage with various stakeholders and improve network positions. More broadly, our results suggest that information technology can be used as a powerful tool to reduce the barriers faced by disadvantaged groups, problems that have been persistent and difficult to address.

2. Theoretical Background and Hypotheses

2.1 VC Investment and Information Asymmetry

Venture financing plays a crucial role in the economy, making solving access to its resources for disadvantaged groups all the more important. It is especially important to the IT industry (Bardhan et al. 2013, Kleis et al. 2012). According to the KMPG Venture Report, VC investment in the United States reached $136.5 billion in 2019, out of which about 40% is invested in information technology firms.[2]

---

[2] Data source: https://home.kpmg/xx/en/home/insights/2018/01/venture-pulse-global-analysis-of-venture-

Venture financing is also critical for firms in creating innovative products and services, attracting highly qualified labor, and reducing time from product concept to launch (Hellman and Puri 2000, Baron et al. 1996, Davila et al. 2003).

However, start-ups do not receive VC investments in part due to the information asymmetry between the start-up and investors, which hampers the ability to evaluate a start-up's worth (Cumming 2006). This information asymmetry arises both from the limited availability of information as well as the difficulty in assessing the veracity and the trustworthiness of the parties providing it. The inability to accurately evaluate start-ups is further compounded by the lack of traditional physical assets (e.g., assets that can be sold to recoup a failed investment) as well as the cashflows that can be a key part of valuation (Aldrich and Fiol 1994). Further, entrepreneurs are incentivized to inflate the perceived quality of their start-ups (Amit et al. 1990, Gompers 1995, Dessein 2005), making it even harder for investors to evaluate them accurately.

Historically, companies have managed information asymmetry in part by relying on social capital that allows them to monitor information flows in social networks (Granovetter 1993). The same has been true for the venture capital industry. It is well-established that social capital has been an important factor in alleviating information asymmetry between the VCs and start-ups (Ingram and Roberts 2000, Lerner and Malmendier 2013, Nanda and Sørensen 2010, Nicolaou and Birley 2003, Shane and Cable 2002). Founders and investors often form tight-knit communities to facilitate information exchange with a goal of reducing uncertainty (Bygrave 1988, Sorenson and Stuart 2001). The fact that these mechanisms may work better when founders and investors are geographically close to each other led to significant geographic concentration of entrepreneurial activity (Saxenian 1991, Hochberg et al. 2007). Information not only flows from entrepreneurs to investors, but also in reverse as network relationships facilitate mentoring start-ups

---

funding.html.

by experienced investors and entrepreneurs, raising the prospects of venture success and further increasing their attractiveness to investors (Greve and Salaff 2003, Hoang and Antoncic 2003).

Given the importance of network connections in the VC market, women, and those with scarce social capital face greater difficulties in attracting financing because they lack connections to the investor network. Even if they could provide information to potential VCs, it is difficult for the VCs to verify or trust the information (Guo et al. 2004). This effect is further exacerbated when VCs, inundated with financing opportunities, face information overload. Start-ups with connections to the investor network may further facilitate venture financing as investors increasingly rely on existing connections to help mitigate their overload.

As such, women and less connected entrepreneurs face greater information asymmetry in the VC market than their respective counterparts. They have more difficulty in getting investor's attention and demonstrate the potential of their start-ups to investors. Having a tool that can help them reduce information asymmetry would potentially help them get funded.

2.2 Gender Disparities in Financing Entrepreneurial Ventures

Gender disparities in VC financing have been widely documented. For instance, Brooks et al. (2014) show an "attractive male takes all" phenomenon arising from investors viewing physically attractive men (but not attractive women) as more persuasive. Other studies link gender disparities on a greater level of scrutiny of firms founded by women (Kanze et al. 2018, Brush et al. 2014, Cui et al 2022). More broadly, firms where women hold majority ownership are less likely to seek equity capital (Brush et al. 2001, Harrison and Mason 2007, Orser et al. 2006, Balachandra et al. 2019).

There are a range of possible factors involved in this disadvantage, including most notably sexism, a bias toward giving men funding and excluding women. But one factor might be particularly susceptible to remedy is via networking and perhaps by using social media in particular. Namely, it is plausible that part of the gender disparity in VC financing is due to women not being well positioned in the social networks of VCs. For instance, Gibson (2010) finds that many of the information disadvantages faced by women are a result of their homophilous networks, disadvantages that often disappear when their networks broaden. If lack of social capital is a significant hindrance for women entrepreneurs in receiving capital, having a tool that facilitates the acquisition of social capital could alleviate gender differences in VC financing markets.

2.3 Social Media as a Tool to Address Information-based Disparities

Social media as effective information channels have been well-documented. Reducing the cost of information dissemination by using social media platforms has been shown to be key to brand building, consumer engagement, and other long-term performance metrics (Angst et al. 2010, Aral and Walker 2011, Wu 2013). Similar information dissemination benefits may be present for entrepreneurs. Prior work has indicated that network connections in the financing market, facilitated by social media, can support information exchanges between investors and founders (Hochberg et al. 2007). Agrawal et al. (2011) show that peer-to-peer crowdfunding can facilitate lending by reducing costs related to information asymmetry. Aggarwal et al. (2012) examine how online reviews and judgments might affect the chances of receiving VC financing. Lin et al. (2013) have also shown that for lending platforms, friendships act as signals of creditworthiness and increase the probability of funding. Eesley and Wu (2020) use a randomized field experiment to show that social networks of mentors and advisors affect performance, especially when the firm is changing strategy. Jin et al. (2017) illustrate that social media activities are associated with an increase in funding and broadening the investor pool. Overall, these papers consistently show that social media could increase entrepreneurial success by facilitating information transfer, building stronger network connections, enabling founders to acquire the power and status necessary to signal quality, and facilitating access to critical resources such as knowledge and capital.

As stated above, firms founded by women and firms otherwise lacking social capital in the investor networks face difficulties in signaling quality to potential investors. As such, they face greater information asymmetry in VC market than others. Social media could alleviate the information disadvantage by increasing VCs' knowledge about a start-up. The greater the information asymmetry, the greater the impact that social media can have in reducing it.

Start-ups can demonstrate products and service quality through social media, and the success of such efforts can be observed by investors, improving their network positions as well as their chance of securing financing. Although social media can help all firms build information channels, the relative benefit may be larger for firms that lack social capital because socially connected entrepreneurs already possess mechanisms to transfer information to potential investors using their pre-existing relationships. Entrepreneurs lacking these relationships can still increase their visibility and build relationships through social media. Thus, social media can become an important tool to secure investments for those lacking social capital. Better engagement on social media could help these entrepreneurs reach a broader audience including potential investors. Furthermore, social media can provide additional third-party information to help VCs better assess a start-up. For instance, the ability of a firm to build online engagement on social media can be an indicator of future firm success or market acceptance of its products. Thus, social media can provide quality signals that help reduce information asymmetry (Li and Wu 2018) and thereby increase the chances of securing funding.

Since information asymmetry is greatest amongst firms founded by women, first-time entrepreneurs and less connected entrepreneurs, these types of disadvantaged firms have the most to gain in terms of obtaining funding from using social media as a tool to reduce information asymmetry. On the other hand, firms that are founded by socially connected entrepreneurs and serial entrepreneurs, who are predominantly men, have a smaller information asymmetry in the VC market than others. Accordingly, these firms do not benefit from Twitter's ability to reduce it. Thus, we hypothesize that firms founded by women and other firms lacking social capital in the investor networks should gain more financial benefits from using social media than their better-positioned peers:

H1: Firms founded by women secure a greater amount of funding from using social media than do firms founded by men.

H2: Firms with less social capital in the investor network secure a greater amount of funding using social media than do firms with more social capital.

3. Data

3.1 Twitter activity data

Twitter is one of the most popular platforms used by start-ups. Tens of thousands of start-ups have registered official handles, and many firms post Tweets daily. We use Twitter activity as a proxy of start-up engagement on social media as it is the most widely adopted platform in our data, and entrepreneurs usually communicate with investors using Twitter. We identify a start-up's major official Twitter handle through Crunchbase. A key decision for a firm is whether to adopt Twitter at all, and if they do, what content to provide. Figure 1 shows the main page for Uber on Twitter and a Tweet posted by Uber about its new business plan on August 24, 2021. The distinctive aspect of social media, and of Twitter especially, is that they provide opportunities for any user to generate additional content for a firm by commenting, reposting ("retweet"), or liking ("likes") the Tweet. In 2016, about 70% of firms in our dataset had Twitter accounts. On average, companies tweet about 270 messages annually and receive 250 retweets, 500 comments and 800 likes per year.

We chose to analyze Twitter data over a decade from 2007, shortly after it started in 2006, through 2016.

For each start-up's official Twitter handle, we obtain the date when the firm started the Twitter account[3], the total number of Tweets posted in a year, and the Twitter interactions, measured by the total number of reposts, comments, and likes. We create the **Has Twitter** variable as a dummy variable to indicate whether the company has adopted Twitter or not in this year. The number of Tweets posted and Twitter interactions are strongly correlated, so we take the first principal component of the two variables to create a **Twitter Engagement** variable. We use this engagement variable to capture a start-up's overall Twitter activities. According to the data, one standard deviation change in Twitter Engagement is about 300 Tweets and 30,000 reposts, comments and likes per year.

3.2 Company founding and financing data

We focus on IT start-ups founded in the US as the IT industry is the most active sector in VC investments[4], and focusing on a single industry can eliminate industry-related heterogeneity. The financing data are obtained from Crunchbase, a database that curates information about start-ups, including prior venture participation and funding outcomes (rounds, amounts, and funding participants) and covers over 100,000 company-year observations our chosen time span from 2007 to 2016. For our analysis, we focus on early funding rounds since the average number of funding rounds in our data is about 1.1. We obtain the founding year from Crunchbase. To avoid simultaneity issues, we lead funding by one year relative to our start-up and network measures, and use the natural logarithm of a start-up's funding (plus $1) to account for a right-

[3] We refer to the date when they posted their initial Tweet, which might differ from the official registration date shown on their Twitter page, but we consider it more accurate to document the date when the company started to post on Twitter.

[4] In 2018, the IT industry accounted for over 60% of total VC investment in US. Source: https://www-cbinsights-com.proxy.library.upenn.edu/research/report/venture-capital-q4-2018/.

skewed distribution; this also implicitly recodes no funding years to zero.[5] We use this **Log Lead Funding**[6] measure for each company in each year since it has been founded and before it exited (e.g., through IPO, acquisition or company dissolution) as our dependent variable.[7]

Our primary independent variable, **Gender Diversity**, is binary, taking on a value of 1 if more than half of the founders are female and zero otherwise; 9% of the firms in our sample are gender diverse by this metric. Results are similar if we instead measure gender diversity as having *any* women on the founding team, in part because most of the firms with female founders have only one woman on the team.

Financial Competitiveness of industry sectors is based on a metric suggested by Campa and Kedia (2002) calculated as the fraction of companies within an industry sector that received funding in a given year (competitiveness is one minus this value suggesting financing is harder to find). We focus on six IT industry sectors based on the Global Industry Classification Standard: IT Services; Software; Communication Equipment; Technology Hardware, Storage & Peripherals; Electronic Equipment, Instruments & Components; and Semiconductors & Semiconductor Equipment. Figure 2 shows the financing competitiveness after 2006.

3.3 Investor network data

Using the investor relationship data provided by Crunchbase, we construct a firm-level investor network for each year. Most of the investors are venture capital firms. Each node in the network is a start-up and a

[5] Besides venture financing, some start-ups also seek crowdfunding, so we also collect crowdfunding data on Kickstarter.com. Our results remain substantially the same with Kickstarter data added as a control variable.

[6] Specific variable names are highlighted in **bold** text.

[7] Some start-ups in the sample were founded before 2007. We only keep their observations starting in 2007 because Twitter was not available before 2007.

link between two start-ups indicates the presence of an investor who invested in both start-ups at some point within a five-year period. These networks vary over time due to the emergence of new start-ups and the entry of new investors. More than 60% of the start-ups are isolated nodes in each year's network because most start-ups fail, with many never having received funding. The average firm survival time in our dataset is less than five years. These isolated nodes (singletons) are significantly different from those that are connected to the investor networks, not only because of their network structures, but also because they may differ in unobserved ways associated with their lack of funding. Thus, we separately estimate our models on isolated and connected companies. There are 15,462 isolated companies with 75,025 observations in our dataset; the set of connected companies, firms that were previously invested by VCs, comprises 6,640 firms with 40,959 observations. For connected companies, we measure degree, transitivity and the network constraint as defined below.

We use Burt's structural hole metric (Burt 2009), which measures the extent to which a node can act as a bridge connecting different disconnected groups. We choose network constraint as the main network measure in our context because nodes that bridge structural holes, or brokers, (identified as a low value in this measure) are shown to provide individuals with information benefits such as better and faster information access or the ability to better assess information based on knowledge of the source (Burt 1992, pp 13). Greater and faster access to information can directly reduce information asymmetry while the improved ability to screen information allows actors to make effective use of this information (Aral and Van Alstyn 2011). Specifically, constraint is calculated as:

$$constraint_i = \sum_j (p_{ij} + \sum_q p_{iq} p_{qj})^2 \, , q \neq i, j$$

Where $p_{mn} = 1$ if there is a connection between nodes m and n, otherwise $p_{mn} = 0$. The lower constraint value is often referred to as high network diversity or occupying "brokerage" positions (a node is critical to connecting other nodes in the network and occupying structural holes).

Based on Burt et. al. (2013), a network constraint (1) decreases with one's network size and (2) increases with one's network density. In our context a start-up with a high value on this measure implies either of the following two factors: (1) a start-up has fewer prior funding experiences (so that the network size is smaller); (2) investors who have invested in this start-up have managed smaller investment portfolios (so that the network density is high). Either factor suggests that the start-up lacks social capital in the investor-relationship networks and thus faces greater information asymmetry with the potential investors in the VC market.

Thus, we primarily use the network constraint to measure company network positions and use degree and transitivity as control variables when a firm is already a part of the VC network. Specifically, we measure **Degree** as the number of neighbors a start-up has in its network. Those start-ups that are funded by large and notable VC firms usually enjoy a higher network degree, since these venture capitalists have large investment portfolios. **Transitivity** is measured as the fraction of a node's neighbors that are also its neighbors' neighbors, sometimes known as the clustering coefficient. High investor network transitivity shows that the same group of investors has funded most of the companies in the triangular clique. Transitivity is thus another way to capture the clustering effect of investment.

All the network metrics we use are local metrics that do not require knowing the accurate global network, and are thus more robust to small changes in networks than global metrics. Sampling biases are less likely to affect these metrics because computing their values does not require the entire network.

3.4 Other control variables

Since we have a panel of data, we control time-invariant firm characteristics by adding **Company Fixed Effects** and control time trend effects by adding **Year Indicators**, which are dummy variables indicating each year in our sample, and a **Linear Time Trend** of Twitter adoption. We also add **Market-Time Fixed Effects** in the model, which are created by multiplying the time trend to the six market subcategories in our data (IT Services; Software; Communication Equipment; Technology Hardware, Storage & Peripherals; Electronic Equipment, Instruments & Components; and Semiconductors & Semiconductor Equipment). As companies in different developing stages usually seek different amounts of funding, we also control for company **Age** in each year, and **Total Funding Rounds**, which is the number of rounds of funding that a company has already received since its founding. Over 90% of observations are less than 15 years old and with three rounds of funding or fewer. We also control for advertising and other online activities using **Google Trend** search measures and **Website Rank** from Alexa[8] to prevent confounding Twitter activity with general online presence.

To control for the personal influence of individual founders, we also create a founder-alumni network in each year. Prior literature has shown that the educational level of entrepreneurs affects future venture performance (Hsu et al. 2007, Eesley et al. 2016). Although the educational backgrounds of founders are usually time invariant and captured by firm fixed effects, the alumni networks are dynamic and can also affect the venture financing outcomes (Lee 2017). We construct the founder-alumni network by connecting founders who went to the same university (undergraduate- or graduate-degree) in each year. We obtain all the LinkedIn profiles for all the founders of our sample through a LinkedIn API. All founders have at least a bachelor's degree. As new companies are founded each year and new founders enter the market, the founder-alumni network also grows dynamically. After obtaining the founder-alumni network, we calculate the network degree for each founder and calculate the maximum degree among founders in each start-up

---

[8] The Alexa Website rank captures the online traffic and offers a measure of the influential reach of a start-up's official website each year from 2008. This ranking lists the top 3 million popular websites worldwide, so we assign a lowest rank value of 3,000,000 for those unlisted companies.

each year and use this **Alumni Network Degree** as a control variable. Such alumni network metrics also help to rule out potential endogeneity concerns such as companies with low social capital being more likely to use Twitter (see Appendix A5).[9]

3.5 Summary statistics

The summary statistics and correlation table for the data are reported in Tables 1 and 2:

Our data include 115,984 observations for about 20,000 companies across the 10 years from 2007 to 2016. In the panel, only the gender diversity variable is fixed over time. The company network metrics, the website rank and Google Trend are measured at the end of the focal year. The Twitter activity variables measure the activities for the current year.

4. Methodology

To test whether Twitter has the effect of alleviating disadvantages that are subject of the study, we explore the interaction effect between Twitter usage and gender, and between Twitter usage and the network constraint measure. We expect the isolated companies (those that are new to the investor network) to behave very differently from connected companies (those that already have network connections in the investor network) and separately estimate our model in the two samples. The samples are separated by observation so a firm can be isolated and later become connected in our analysis. We first fit the following OLS model with company fixed effect (γ_i) and time dummies (τ_t) for all connected companies:

---

[9] We thank the anonymous editor and an anonymous reviewer for suggesting these tests.

$$
\begin{aligned}
log\_funding_{i,t+1} \\
&= \beta_0 + \beta_1 age_{i,t} + \beta_2 total\_funding\_rounds_{i,t} + \beta_3 website\_rank_{i,t} \\
&+ \beta_4 google\_trend_{i,t} + \beta_5 investor\_network\_transitivity_{i,t} \\
&+ \beta_6 investor\_network\_degree_{i,t} + \beta_7 investor\_network\_constraint_{i,t} \\
&+ \beta_8 has\_twitter_{i,t} + \beta_9 twitter\_engagement_{i,t} + \boldsymbol{\beta_{10} has\_twitter_{i,t}} \\
&\times \boldsymbol{investor\_network\_constraint_{i,t}} + \boldsymbol{\beta_{11} twitter\_engagement_{i,t}} \\
&\times \boldsymbol{investor\_network\_constraint_{i,t}} + \boldsymbol{\beta_{12} has\_twitter_{i,t}} \times \boldsymbol{gender\_diversity_i} \\
&+ \boldsymbol{\beta_{13} twitter\_engagement_{i,t}} \times \boldsymbol{gender\_diversity_i} + \tau_t + \gamma_i + \varepsilon_{i,t}
\end{aligned}
$$

4.1 Identification strategies

To estimate for isolated companies, we drop all the network-related variables because theirs is a network of only one. As most companies do not raise funds every year and some may never go to the VC market, there are many zeros in the dependent variable. This problem is partially addressed by our company fixed effect model, and we also use a Heckman Selection model and calculate the inverse Mills ratio to account for a firm's decision to raise funding. Following Campa and Kedia (2002), sector competitiveness can be an instrument for the decision to raise funds while not directly affecting the amount of funding raised. We have two ways to approximate market competitiveness: (1) product market competitiveness data (a measure of concentration) from Owler.com[10] in 2019 (which is cross-sectional) and (2) Google Trends search measures for each market which can measure the general level of interest for that market (which is longitudinal). Multiplying the two together provides a time-varying set of variables that can approximate sector competitiveness. We use this instrument in the Heckman selection model to account for the decision to raise funding. We note that we are using product market competitiveness as opposed to financing competitiveness for this instrument, as financing competitiveness would directly affect venture financing outcomes. For instance, other's financing outcomes would directly affect the focal firm's, but product

[10] Product market competitiveness is calculated as 1 minus the Herfindahl-Hirschman Index of market concentration: $\mathrm{Competitiveness_i} = 1 - \mathrm{HHI_i} = 1 - \sum_{j=1}^{n} s_j^2$, where s_js are the market shares of all the competing companies in the same domain, as provided by Owler.com.

market competitiveness should not have this type of mechanical correlations. We also estimate Oster's bound and models using coarsened exact matching to provide further robustness checks.

It is possible that both the Twitter activities and the funding outcomes are spuriously correlated due to unobservable characteristics. For instance, firms that are active on Twitter may simply be more active online generally, thus conflating the information benefits of Twitter with other information sources. To address this issue, we use the Alexa Website traffic ranking and Google Trends search measures to control for other online activities that are not captured in the Twitter data.

Unobserved firm quality could also bias our results. For example, while any firm can join Twitter, better-funded firms have more resources that may help build Twitter engagement. We can partially address this problem by separating adoption and engagement and using funding outcomes in future years as a dependent variable. Where possible, we use fixed effects to control both company and time-fixed effects. This will control for time-invariant characteristics (at least in the short run) such as the founders' quality. In addition, we use instrumental variables to address other forms of endogeneity associated with company-specific, time-varying characteristics (e.g., other dimensions of "quality") that may be simultaneously related to both social media activity and funding success.

First, we use the Twitter activities of firms in the same state, and in similar funding rounds (within 2 rounds of each other) to instrument firms' own Twitter activities. Start-ups of the same region are likely to be influenced by similar media factors, such as having similar types of Twitter followers. Previous literature has shown that the diffusion of Internet technologies and online interactions are similar within a geographical region (Forman et al. 2002, Forman et al. 2005, Kulshrestha et al. 2012, Kim et al. 2013). Accordingly, Twitter usage for firms located in the same region is likely to be influenced by the number of Twitter users in the region and users' propensity to interact with start-ups online. However, other firms' Twitter activities should not directly influence a start-up's own funding outcomes. Figure 3 shows the locations of the start-ups: 35% are in California, 10% are in New York, and the rest are somewhat more evenly distributed across the states. The distribution has enough variation to provide sufficient first stage power.

Second, we also use the network neighbors' Twitter adoption rate and Twitter engagement to instrument for the focal firm's Twitter activities. Network neighbors are firms with whom the focal firm has at least one investor in common. Firms with the same investors are likely to receive similar advice and consequently to formulate similar managerial and marketing strategies.[11] Accordingly, the focal firm's Twitter activities are likely correlated with those of its neighbors. At the same time, network neighbors vary substantially in their industry and type such that they may not face the same shocks faced by firms in the same industry or region. Thus, their financing outcomes are less likely to be correlated. One concern of this instrumental variable is that funding outcomes between network neighbors might be correlated since funding one start-up will leave an investor with fewer resources to fund another. However, the existence of multiple investors involved in each funding round (the average is three for each round) can somewhat alleviate the mechanical correlation. Nevertheless, we cannot use this instrument on isolated firms which have no network connections, which accounts for about 70% of all the companies and 60% of all the observations. These two sets of instrumental variables have been similarly used to address endogeneity related to the performance effect of IT in firms (Wu et al. 2019). For every 2SLS estimation in the following sections, we use the product of Twitter activities and network positions or gender diversity of geographical and network neighbors to instrument for all the interaction terms in the regression.

To check the validity of our instrument, we check whether the instrumental variables predict the future

[11] For 2007–2016, board member data is available for 18% of start-ups in our dataset; 60% of the neighboring firms share at least one board member in common for some length of time. The overlaps suggests similarities in managerial experience.

levels of funding in companies that have not adopted Twitter, following the procedures in Martin and Yurukoglu (2017). The logic is as follows: if the instrument is valid, then it should be correlated *only* with the focal company's funding through its effect on the focal company's Twitter activities. For companies without Twitter activities, we should not observe a significant correlation between the instrument and the outcome variable. To test this, we regress the dependent variable on the instrumental variables directly, using only data from observations that did or did not adopt Twitter. The results are shown in Appendix A1. We find the statistical relationship between the instruments and the dependent variable is significant using samples that have already adopted Twitter, while the relationship is not statistically significant using samples that did not adopt Twitter. These results provide further evidence for the validity of our instrumental variables.

Besides instrumental variables, we also conduct Oster's bound analysis for the main regressors (Oster 2019) to test our results against possible unobservable selection effects. This test indicates that our results are stable and robust to unobservables since Oster's delta is greater than 1 for all interaction terms in the regression models.

Lastly, we also conduct coarsened exact matching (Blackwell et al. 2009) on our observations to provide further evidence of identification. For testing H1, we match a comparable group of companies founded by males with another group of companies founded by females based on founder characteristics, start-up age, funding experiences, website ranking and network conditions. For testing H2, we match a comparable group of companies with above-median network conditions with another group of companies with below-median network conditions based on founder characteristics, start-up age, funding experiences and website ranking. The results are listed in Appendix (A3) and are consistent with our fixed-effect regression results.

5. Results

To explore how Twitter activities help alleviate gender inequalities, we first present an event study analysis that compares the effect of Twitter adoption on female-founder and male-founder start-ups using the coarsened exact matching samples of the data with fixed effects and control variables (Figure 4). Each data point on the graph shows the effect of Twitter adoption on funding for all the start-ups each year within a group (female vs. male founders) before or after the adoption event in a seven-year observation period. For both groups of companies, the effect of Twitter adoption on funding is positive at the first year, and the coefficient is greater for female-founder firms than male-founder ones, suggesting Twitter is beneficial generally, but more so for female entrepreneurs.

We first analyze the results for isolated companies in Table 3, Columns (1) and (2). These firms are the new entrants for whom no VCs have yet made an investment. Being a new entrant means that isolated firms have the worst network positions. Thus, we estimate these firms separately because they are very different from firms that are already a part of the VC network. Column (1) shows two-way fixed-effect regression results, and Column (2) uses two-stage least squares (2SLS). The dependent variable is the natural logarithm of funding in the next year. Both columns use a Heckman selection model with an inverse Mills ratio to correct for possible selection biases. After controlling for start-up age, total funding rounds, website rank, Google trends volume, and market-time fixed effects, the coefficients of "Has Twitter" and Twitter Engagement are all positive, suggesting Twitter use and engagement are positively correlated with attracting financing for firms that did not start out as a part of the investor network.

Next, we show the interactions between gender diversity and Twitter activities for the isolated firms. Table 3 Column (1) shows that Twitter use is correlated with attracting more funding for companies having a majority of female founders—about a 13% increase in funding outcomes for a one-standard-deviation increase in Twitter engagement, which is about 300 Tweets and 30,000 reposts, comments and likes per year. Since the average funding amount for isolated companies is about US$8 million annually, raising an additional 13% is substantial. The 2SLS results are directionally similar to the OLS results with first stage

F-statistics over 20.[12] These results collectively support H1 that women benefit more than men when using Twitter, especially when they do not start out as part of the investor network.

Columns (3) through (6) show Twitter's influence on connected companies, the incumbents that have received at least one round of financing. These firms also have time varying networks which allows us to use the variations on the network constraints in addition to gender diversity in our models. Columns (3) through (5) show the results for fixed effects OLS, while Column (6) shows the 2SLS estimations. All models adjust for inverse mills ratio added from a Heckman selection model. The coefficients of Twitter Engagement are all positive. Together with Columns (1) and (2), we show that Twitter engagement helps both isolated and connected firms to obtain VC financing. We controlled for start-up age, year fixed effects, market-time fixed effects, total funding rounds, website rank, Google trends, investor network degree and transitivity. The estimates for these control variables have the expected signs.

Next, we explore how gender and network position mediate the relationships between Twitter engagement and VC financing. Column (3) estimates the Twitter effects on female-founded firms. Compared to our results for isolated firms in Columns (1) and (2), where we find women benefit more from using Twitter when they are not a part of the investor network, we find Twitter use and adoption no longer affect female-founded firms once they become part of the investor networks. This suggests that the disadvantage in funding related to gender disparities is partly due to female entrepreneurs having constrained networks. We later show (in the mechanism analysis) that Twitter has a greater effect on improving network position for women than for men. This also suggests that once a female entrepreneur has improved her network position, the gender disadvantage from a lack of access to the VCs is reduced. Column (4) shows that Twitter

[12] Due to the significant increases in the coefficients in the 2SLS regressions we restrict our discussion to the sign of these coefficients, although these could suggest the magnitude of the effects we observe are in reality larger than what is suggested by the OLS results.

activities are correlated with securing more funding by firms with less social capital in the investor network. Specifically, a one-standard-deviation increase in Twitter engagement is associated with a 10% increase in the funding in the following year for a company that is one-standard-deviation higher than the average network constraint. Since the average funding amount is over US$17 million annually for connected companies, a 10% increase in funding is substantial. Column (5) has both interaction terms, and they similarly show that Twitter primarily helps firms with more constrained networks, and Twitter engagement is more important than Twitter adoption. The 2SLS results in Column (6) are directionally consistent with our OLS results with first stage F-statistics over 20. Collectively these results support H2.

6. Mechanism and Robustness Checks

We identify mechanisms that shows how social media can reduce the high information asymmetry faced by women and network constrained founders and part of the social media effect comes from the resulting increase in the start-up's network diversity after using Twitter. First, we show that our results are consistent with the mechanism of increasing effective information dissemination to reduce information asymmetry. Specifically, we show (1) the effect of Twitter is stronger in more financially competitive markets than in less competitive ones, and (2) the effect of Twitter is stronger for attracting new investors than for repeat investors, and (3) the effect of Twitter is stronger for new entrants than for entrepreneurs already in the network from having previously funded by a VC. Together, these results show Twitter has the most effect when information asymmetry is high and there is greater marginal benefit of gathering more information.

We then explore how Twitter can help reduce information asymmetry through the content of their Twitter engagement as well as through improving network positions of the start-up. We explore what type of Twitter content is most conducive to reducing information asymmetry and receiving VC financing. We show that all content types which we classified can help isolated firms, but only certain content (product promotion and customer relations) seems to help connected firms.[13]

Next, we explore whether twitter might have helped disadvantaged entrepreneurs through improving their network positions in the VC network. We find that, after using Twitter, women founders' network position improves more than men's over time. The improved network position could also help women in obtaining venture financing.

Finally, we show our results are robust if we use the probability of being funded as opposed to using the amount of funding as the dependent variable. We also verify our results (Appendix A2) using a matching-estimator, a dynamic-panel model, and system GMM estimators (Arellano and Bond 1991, Blundell and Bond 1998). To ensure that our results are based on network capital from the VC investment network as opposed to general social capital, we replace the VC network with an alumni network. Appendix A4 shows that alumni networks have no impact in our setting.

6.1 Reducing information asymmetry

Financially competitive markets

Our theoretical discussion emphasizes the role of Twitter engagement in alleviating information asymmetry. In markets where getting VC-financing is competitive, we would expect alleviating information asymmetry would be more difficult given the plethora of options investors would need to eliminate to make an investment decision. As a result, investors may have less resources to collect information on each start-up they evaluate. Prior literature has shown that additional quality signals can be crucial for decision-making when there are many alternative choices and when outcomes are highly uncertain (Forman et al.

[13] We thank an anonymous reviewer for suggestions on these tests.

2008, Lu et al. 2019, Snir and Hitt 2003). If Twitter can provide additional low-cost signals to help investors reduce uncertainty, it should improve funding outcomes. In Table 4, we show that Twitter has a greater effect for markets where VC-financing is competitive than for markets where it is less competitive. Columns (1) to (3) show that the subsectors with high financing competitiveness are negatively associated with future funding. They also show that the interaction between Twitter activities and financing competitiveness is positive, suggesting that Twitter has an especially positive impact in financially competitive markets where it is relatively more difficult to engage investors. Column (4) shows the 2SLS results are directionally consistently with our OLS results with first-stage F-statistics over 10. Overall, these results suggest that Twitter has a larger effect in markets where reducing information asymmetry requires more resources.

Repeat vs. new investors

If Twitter is effective in alleviating the information asymmetry, it should have a larger effect in bringing in new investors than in bringing in investors that have already invested in the firm because repeat investors likely have more knowledge about the start-up as compared to new investors. Table 5 shows the effect of Twitter on attracting funding from new investors, and repeat investors' subsequent investment in the firm. Columns (1) and (2) in Table 5 show that the effect of Twitter Engagement on funding outcomes is greater for less-connected founders. However, Columns (3) and (4) show that Twitter has a limited influence on encouraging repeat investors to invest more, as the effects of Twitter Engagement are not economically and statistically significant for these investors. Comparing the columns, we find that those less connected in the VC networks benefit the most from using Twitter to attract funding from new investors, and the 2SLS results are directionally consistent with our OLS results and not attributable to weak instruments (first-stage F-statistics over 10). These results are consistent with the mechanism that Twitter can reduce information asymmetry between start-ups and potential investors, and the effect is greatest for less connected founders who face greater hurdles in overcoming information asymmetry.

Serial vs. first-time entrepreneurs

Prior literature has shown that serial entrepreneurs are better than novice entrepreneurs at building networks and managing organizations (Li et al. 2009). Investors usually have a strong preference for investing in serial entrepreneurs over first-time entrants because their prior founding experiences can help mitigate uncertainty (Wright et al. 1997, Parker 2013) and because the available information about first-time founders is more limited. If Twitter is effective in addressing disparities that arise from information asymmetry, this effect should be stronger for start-ups founded by first-time entrants than those founded by serial entrepreneurs. Table 6 shows the effect of Twitter on the two types of entrepreneurs. The coefficients of Twitter activities in Column (1) are significantly larger than in Column (2), suggesting Twitter is more effective in attracting funding for first-time entrants than for experienced entrepreneurs. This result parallels recent literature showing that Twitter has a stronger effect on the ability of new politicians to raise donations than it does on experienced ones (Petrova et al. 2020). For Columns (3) and (4), the dependent variable is network diversity (2 minus network constraint) in the following year. Column (3) shows Twitter's effect on the start-up's future network positions for first entrants and Column (4) shows the effect for serial entrepreneurs. The results also show that Twitter is more effective at increasing network diversity and improving network positions for first-time entrants than it is for serial entrepreneurs.

6.2 What information matters on Twitter

Finally, to check whether the contents of Tweets are effective in signaling quality and alleviating information asymmetry, we also conduct text analysis of the Tweets to see what kinds of content are associated with more funding. We first use a simple bag-of-words method by first removing the non-content words, punctuation and numbers, stemming the remaining words and then calculating the frequency of each word. Next, we run a latent Dirichlet allocation (LDA) test and analyze the effects of different Tweet content. The LDA procedure indicates that the top four topics start-ups discussed on Twitter can be broadly characterized as: Human Resource Management, Business Announcements, Product Promotions, and Customer Relations. Figure 5 shows the trend of different topics on Twitter.[14] We conduct regression analysis for different Twitter topics in Table 7 (for isolated firms) and 8 (for connected firms). We explore the mechanism of how Twitter reduces information asymmetry for isolated firms in Table 7. In Table 3, we show that isolated firms and especially firms that are founded by women benefit from all content types, with the strongest effect from content about human resources, business announcements and product promotions. Given the dearth of knowledge about new start-ups, most business-relevant information on Twitter is helpful. These results support our hypothesis that Twitter activities help alleviate the information asymmetry for those with the most serious information asymmetry problems in the VC market. Thus, any type of business-relevant Twitter activity is helpful as it provides additional information about the company.

For connected firms, Twitter engagements help firms that are already a part of the VC network, but the effect on funding is larger for firms that are network-constrained (Table 3). Table 8 explores this result further by analyzing the content of the Twitter engagement. We find weak evidence that business content helps these firms on average but firms with high network constraints seem to benefit from posting product promotions. These results have 2 implications: (1) For market incumbents, providing additional quality signals is only helpful for network-constrained start-ups to attract potential funding. This is similar to the effect we show for isolated firms, the most network-constrained firms. For well-connected companies, the effect is insignificant due to the relatively low information asymmetry between investors and start-ups; (2) Unlike new market entrants, the greatest benefit is achieved for posts associated with potential revenue

[14] Keyword dictionaries and example Tweets are shown in the Appendix A6.

growth. This may be because operational information (e.g., hiring, or business milestones) is already well known to investors, or because the relevant uncertainty for established firms is regarding their ability to continue growing, having already demonstrated their ability to establish a presence in the market. Such results support the mechanism that Twitter provide additional quality signals about the firm to investors where information asymmetry is significant. The results further alleviate the endogeneity concerns such as omitted variable biases and reverse causality as that would predict that all forms of communication would appear to be uniformly beneficial for all types of start-ups.

6.3 Improvements in network position

In addition to improving information asymmetry through directly providing content to potential investors, these Twitter interactions could improve a start-up's connections to VCs which can further reduce information asymmetry. First-time entrepreneurs, less connected, women, and others who face serious information asymmetry may benefit more than their counterparts. To examine this mechanism, we show whether their network position improved after Twitter was used over time. Comparing first-time and serial entrepreneurs, we show that Twitter improves the network diversity of first-time entrepreneurs more than it does for serial entrepreneurs (Table 6). This suggests that Twitter engagement helps forge connections to the VC networks for first time entrepreneurs.

We show a similar pattern for female entrepreneurs who tend to be less connected in the VC networks. We first show Twitter engagement can help improve the network position and whether the improvement is greater for women than for men (Figure 6). Using lead network diversity as the dependent variable, the event study shows that the coefficient of Twitter Engagement for women is larger than that for men both statistically and economically. We further show that the overall effect of Twitter on future financing is diminished once a start-up becomes more connected in the network (Table 3). In that table, we show that there is no incremental benefit for female founders in connected companies, while there is a strong positive effect on initial funding for isolated companies. Overall, these results show that Twitter can help women and first-time entrepreneurs to become more connected in the VC network, which in turn can help them with future funding. The effects are then diminished once they are a part of the VC network.

7. Discussion and Conclusion

By exploring how Twitter can narrow the gap between traditionally disadvantaged start-ups—start-ups with low social capital and women founders—and those that are already endowed with demographic and network advantages, we provide new evidence that gender and network related differences can be alleviated to some extent by reducing information asymmetry using Twitter. As discussed in the results section, we show that Twitter engagement can help start-up acquire VC financing, and the effect is particularly strong for start-ups founded by women and founders with low social capital in the VC network.

As many have documented, women and minorities face disparities in VC financing. Yet, closing this gap has been extremely difficult and the gap has not diminished over time. Our results show the potential of using technologies to address part of this challenging problem of closing the performance disparities in VC financing for women and those with low social capital in the VC network. Specifically, these founders received more VC funding from using Twitter than did others because they provided a much-needed mechanism of lowering the information barriers to reach investors. Consistent with this result, we find that the effect of Twitter is stronger when there is less ex-ante information (first round funding, fist-time entrepreneurs, new investors, those without an established network) and when the information disadvantage is more severe (in competitive markets).

Although no technology can eliminate all disparities, social media provide a much-needed channel for women and especially first-time women entrepreneurs to reach out to investors. Thus, it may be helpful to have policies that help women and those with low social capital to effectively engage on social media, so they are more likely to receive VC financing and grow their firms. By reducing information asymmetry between the start-ups and the investors, social media engagement will not only improve diversity in the entrepreneur ecosystems but also improve market efficiency. While the adoption, use, and effective engagement in social media are not without cost, it may be significantly less costly than building social networks for new entrepreneurs through traditional channels, and the underlying technologies are freely accessible regardless of one's ex-ante network position or personal relationships.

The strength of our measurement approach is limited by the likely existence of uninformative or inaccurate information transmitted through social media – for instance, entrepreneurs may exaggerate or spread inaccurate information on social media platforms distorting the signaling effect. While much of this will ultimately be addressed by the due diligence process applied by VCs prior to making an actual investment, we cannot distinguish between the 'real quality' and 'inflated quality' from online content in our setting. Despite these problems, which are inherent in the setting, we can detect an increase in VC investment in otherwise disadvantaged firms associated with social media use. As it is infeasible to conduct the same level of due diligence on all the firms in our sample, there is also the possibility that unobserved differences in start-up quality (possibly observable to investors but not to us) may bias the results. However, we try to minimize the impact of these types of endogeneity problems through various mechanism tests and econometric methods.

This paper has important managerial implications for multiple stakeholders in the entrepreneurship and VC market. For start-ups and entrepreneurs, information tools like social media offer a lower cost channel to signal qualities and alleviate information asymmetry, which can in turn help reduce the disparities in VC financing for women and less connected entrepreneurs. These groups usually face greater barriers in accessing financing resources and it is largely unknown how these barriers can be reduced. We show that social media are a plausible mechanism to reduce some of these barriers. This not only benefits these entrepreneurs themselves but also the greater entrepreneurship and innovation environment. For venture capitalists, social media provides a source of information to better evaluate the quality of a start-up, enhancing their ability to improve their investment portfolios. For policy makers, it is important to think about bringing technology as a part of the tool kit to combat disparities. Our research shows that providing more communications between investors and start-up companies helps alleviate some disparities for women and less connected founders and potentially benefit diversity in entrepreneurship while simultaneously improving investment outcomes.

This paper contributes to a broader understanding of the social impact of information and communication technologies. Over the past few decades there has been a substantial body of research on how the benefits of information technology may or may not be evenly distributed across society. At least in this setting, we show that social media can be used to broaden participation in entrepreneurship. This paper also contributes to understanding the economic impact of social media and specifically to the small but growing literature focused specifically on start-up financing.

Future studies could consider other types of disparities, such as racial disparities in VC financing. Minorities has long faced difficulty in obtaining VC financing. Understanding how technologies could be used as a tool to mitigate the racial disparities would be a fruitful future direction. It is also important to examine the long-term impact of social media on venture financing, especially after the publication of new policies by the SEC regarding investment-related communications. Our results suggest that early-stage firms should not only establish a presence on Twitter, but also effectively engage with their target audience through them, especially for those where their potential investors are not well-informed about their firms.

Figure 1: Twitter Account of Uber (Left) and A Tweet posted by Uber (Right)

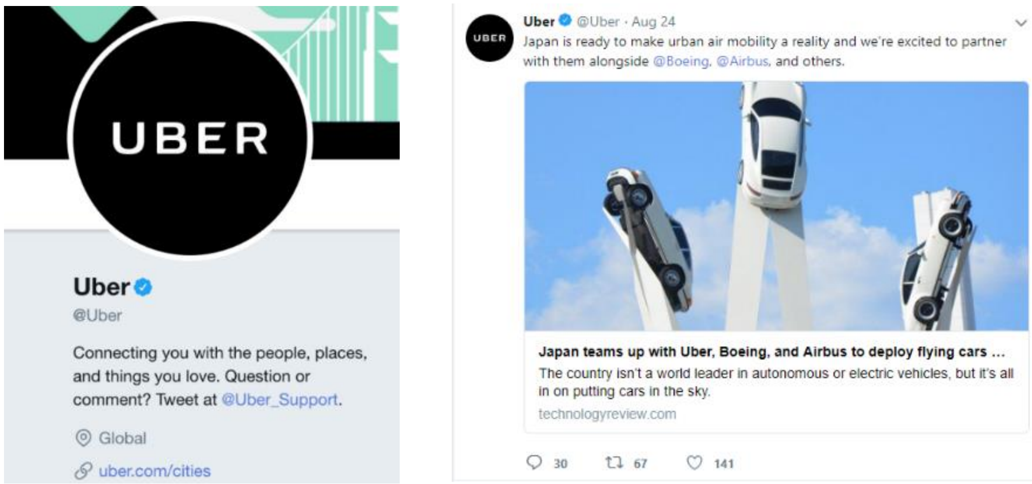


Figure 2: Financing Competitiveness in Subsectors

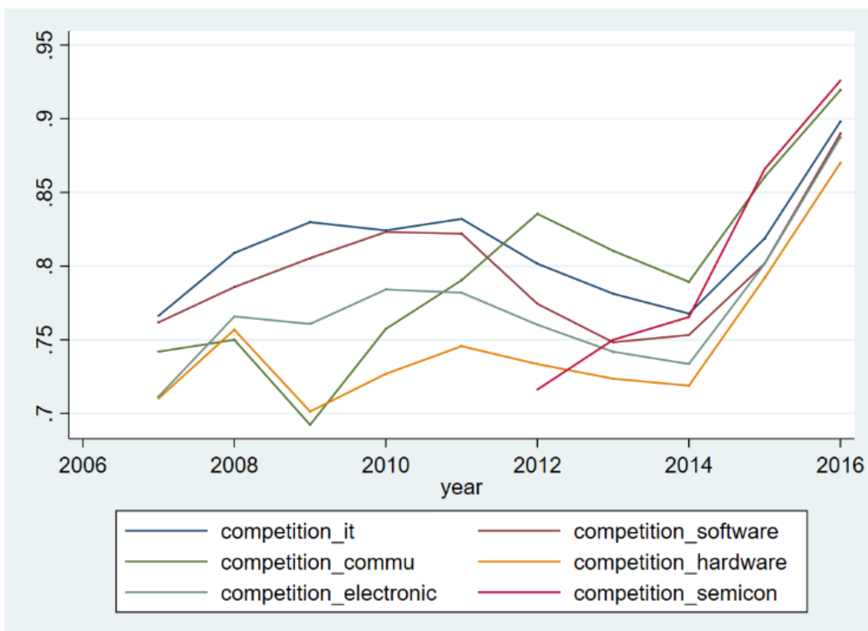


Figure 3: Geographic Distribution of Sample Start-ups

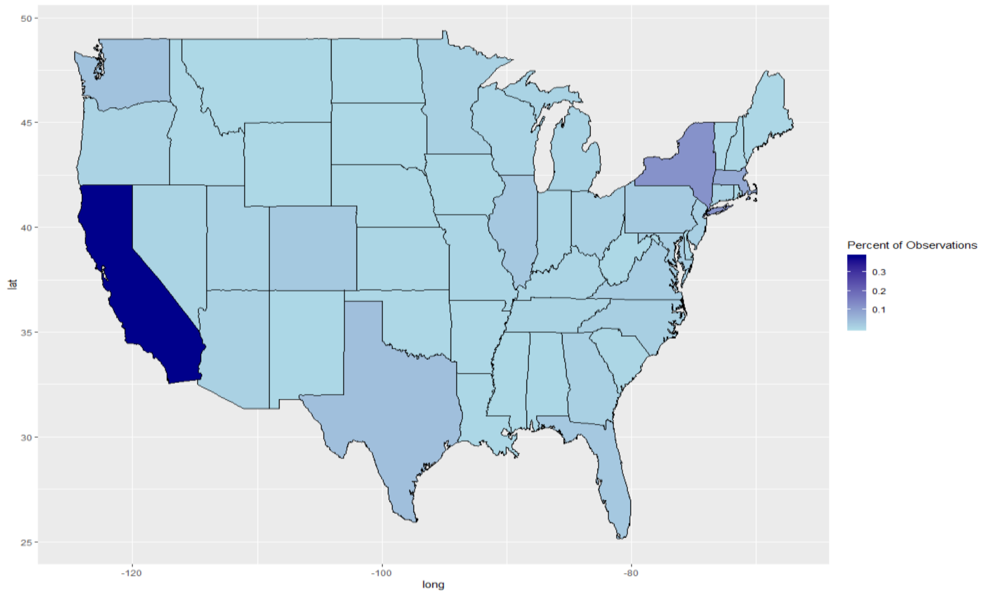
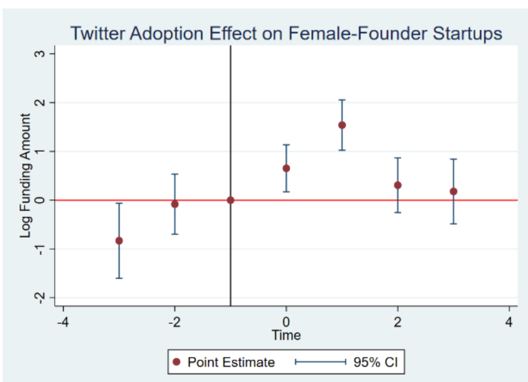
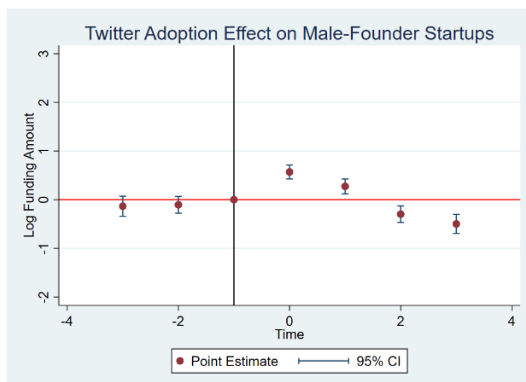


Figure 4: Event Study: Twitter Adoption Effect on Funding Performance



(a) Female-Founder Start-ups



(b) Male-Founder Start-ups

Notes: Each point on the graph shows the average effect (with 95% confidence interval) of Twitter adoption on the funding amount in the year relative to Twitter adoption. From this graph, it shows that the funding outcome goes up after the adoption year (X=0), and the effect is stronger for female-founded firms than male-founded ones. The plots are created using *eventdd* package in stata after controlling firm fixed effects, year fixed effects and market-time fixed effects.

Figure 5: The Average Number of Tweets in a Year by Topics

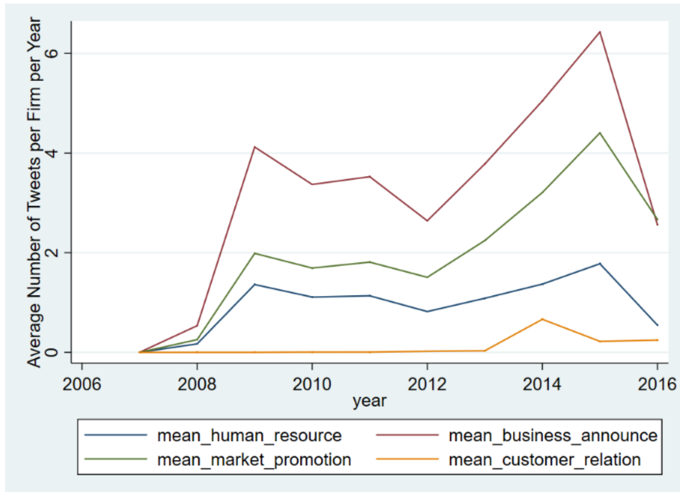
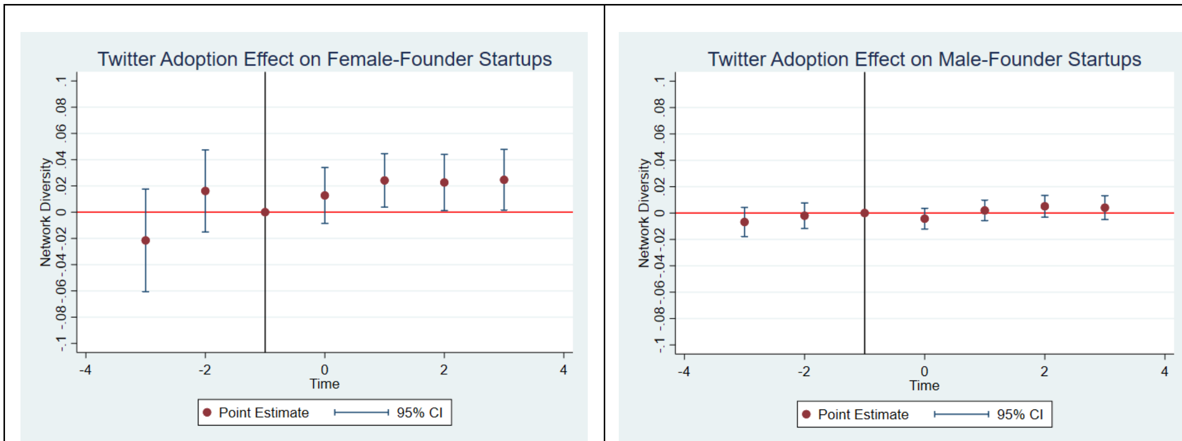


Figure 6: Event Study: Twitter Adoption Effect on Network Positions



(a) Female-Founder Start-ups

(a) Male-Founder Start-ups

Notes: Each point on the graph shows the average effect (with 95% confidence interval) of Twitter adoption on the network position (network diversity = 2 – network constraint) in the year relative to Twitter adoption. From this graph, it shows that the network position goes up for female-founded firms and is statistically significant at 95% level. For male-founded start-ups, the effect is not statistically significant. The plots are created using *eventdd* package in stata after controlling firm fixed effects, year fixed effects and market-time fixed effects.

Table 1: Summary Statistics

Variables	Obs.	Mean	Std. Dev.	Min	Max
Company Information					
Log lead funding	98,196	1.951	4.218	0	21.976
Age	115,984	6.226	6.456	0	37
Gender diversity	115,984	0.087	0.282	0	1
Total funding rounds	115,984	1.153	1.523	0	14
Domain competitiveness	115,984	0.470	0.396	0	0.919
Network Metrics					
Investor network transitivity	115,984	0.108	0.230	0	1
Investor network degree	115,984	46.507	147.723	0	3,524
Investor network constraint	115,984	0.721	0.412	0.005	1.389
Twitter Activity					
Has twitter	115,984	0.457	0.498	0	1
Twitter engagement	115,984	0.019	0.665	-0.383	7.362
Human Resource Tweets	115,917	1.075	12.140	0	565
Business Announcement Tweets	115,915	3.775	38.427	0	1,710
Product Promotion Tweets	115,917	2.450	18.395	0	847
Customer Relation Tweets	115,984	0.180	9.549	0	1,035
Controlling Variables					
Website Rank	115,984	1,919,084	1,410,115	0	3,000,000
Google Trend	115,984	14.413	19.681	0	95

Table 2: Correlations between Main Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) Log lead funding	1.000															
(2) Years old	-0.095	1.000														
(3) Gender diversity	0.012	-0.094	1.000													
(4) Total funding rounds	0.008	0.264	-0.051	1.000												
(5) Domain competitiveness	0.099	-0.030	0.018	0.023	1.000											
(6) Investor network transitivity	-0.009	0.081	0.003	0.137	-0.015	1.000										
(7) Investor network degree	0.057	0.038	0.003	0.415	0.048	0.058	1.000									
(8) Investor network constraint	-0.054	-0.151	0.014	-0.523	0.121	-0.550	-0.535	1.000								
(9) Has twitter	-0.054	0.035	0.019	0.038	0.095	0.023	0.066	-0.014	1.000							
(10) Twitter engagement	-0.008	0.001	0.041	0.030	0.081	0.031	0.097	-0.040	0.681	1.000						
(11) Human Resource Tweets	-0.013	0.033	0.060	0.044	-0.017	0.017	0.014	-0.026	0.107	0.179	1.000					
(12) Business Announcement Tweets	-0.012	0.029	0.095	0.049	-0.013	0.028	0.028	-0.043	0.117	0.204	0.983	1.000				
(13) Product Promotion Tweets	-0.009	0.031	0.069	0.073	-0.005	0.050	0.061	-0.076	0.149	0.279	0.972	0.967	1.000			
(14) Customer Relation Tweets	0.002	-0.001	0.000	-0.004	0.004	-0.003	-0.001	0.004	0.015	0.068	0.003	0.003	0.006	1.000		
(15) Website rank	-0.059	0.036	0.025	0.035	-0.039	0.004	-0.026	0.054	0.198	0.115	-	-	0.000	0.009	1.000	
(16) Google trend	-0.040	0.115	-0.004	0.030	0.002	0.008	0.067	-0.022	0.128	0.145	0.010	0.016	0.031	0.011	0.375	1.000

Table 3: Twitter Effect on Start-up Funding Performance

DV: Log Lead Funding	Isolated Companies		Connected Companies			
Model	FE	FE+2SLS	FE	FE	FE	FE+2SLS
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Inverse mills ratio (isolated)	-0.012	-0.013				
	(0.192)	(0.191)				
Inverse mills ratio (connected)			-0.117	-0.112	-0.112	-0.099
			(0.123)	(0.122)	(0.123)	(0.121)
Investor network constraint			-0.504***	-0.628***	-0.629***	-0.293
			(0.166)	(0.200)	(0.200)	(0.274)
Has twitter	0.278***	0.283***	-0.458***	-0.469***	-0.469***	-0.492**
	(0.084)	(0.084)	(0.123)	(0.123)	(0.123)	(0.122)
Twitter engagement	0.366***	0.357***	0.180***	0.191***	0.191***	0.224***
	(0.041)	(0.041)	(0.053)	(0.052)	(0.053)	(0.055)
Gender diversity * Has twitter	0.262***	0.077	0.004		0.004	0.000
	(0.069)	(0.140)	(0.098)		(0.098)	(0.097)
Gender diversity * Twitter engagement	0.131***	0.309**	-0.008		-0.004	0.003
	(0.050)	(0.131)	(0.042)		(0.041)	(0.041)
Investor network constraint * Has twitter				0.139*	0.139*	-0.082
				(0.073)	(0.073)	(0.143)
Investor network constraint * Twitter engagement				0.095**	0.095**	0.335**
				(0.046)	(0.046)	(0.142)
Other Controls	Year FE, Market-Year FE, Age, Funding Rounds, Website Rank, Google Trend, Alumni Network Degree, Twitter Adoption Linear Time Trend		Year FE, Market-Year FE, Age, Funding Rounds, Website Rank, Google Trend, Alumni Network Degree, Investor Network Degree, Investor Network Transitivity, Twitter Adoption Linear Time Trend			
Observations	64,076	63,198	34,070	34,070	34,070	33,372
R-squared	0.470		0.492	0.492	0.492	

Notes: 1. Dependent variable is the log funding in the next year. Twitter engagement is standardized, where one-standard-deviation change is associated with about 300 Tweets and 30,000 reposts, comments and favorites posted each year. All columns use a Heckman selection model with inverse mills ratio.

2. Columns (1) to (2) report the results for isolated companies. Columns (3) to (6) report the results for connected companies. Two sets of instruments: (a) Twitter adoption and engagement of companies in the same state within two-round-of-funding distance; (b) Twitter adoption and engagement of companies that share a same investor in the past 5 years.

3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Twitter Effect and Domain Competition

DV: Log Lead Funding	Company Fixed Effect			FE+2SLS
VARIABLES	(1)	(2)	(3)	(4)
Domain competitiveness	-1.034	-0.734	-0.965	-0.385
	(0.910)	(0.911)	(0.911)	(0.922)
Has twitter	0.337***	0.345***	0.340***	0.361***
	(0.058)	(0.058)	(0.058)	(0.058)
Twitter engagement	0.317***	0.303***	0.313***	0.279***
	(0.032)	(0.032)	(0.032)	(0.033)
Has twitter * Domain competitiveness	0.587***		0.527***	0.027
	(0.040)		(0.050)	(0.136)
Twitter engagement * Domain competitiveness		0.295***	0.062**	0.574***
		(0.025)	(0.031)	(0.136)
Other Controls	Year Indicators, Firm Age, Total Funding Rounds, Website Rank, Google Trends			
Observations	98,196	98,196	98,196	97,812
R-squared	0.374	0.373	0.374	

Notes:

1. Dependent variable is the log funding in the next year. Twitter engagement is standardized, where one-standard-deviation change is associated with about 300 Tweets and 30,000 reposts, comments and favorites each year.

2. All columns report the OLS results with company fixed effect. Both the isolated companies and the connected companies are pooled together. Two sets of instruments: (a) Twitter adoption and engagement of companies in the same state within two-round-of-funding distance; (b) Twitter adoption and engagement of companies that share a same investor in the past 5 years.

3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Twitter Effect for Attracting New versus Repeat Investors

Dependent Variable	Log Lead Funding from New Investors		Log Lead Funding from Old Investors	
Model	FE	FE+2SLS	FE	FE+2SLS
VARIABLES	(1)	(2)	(3)	(4)
Has Twitter	0.124	-0.588**	0.023	0.043
	(0.157)	(0.295)	(0.048)	(0.081)
Twitter engagement	0.105**	0.150***	0.008	0.012
	(0.052)	(0.055)	(0.015)	(0.015)
Investor network constraint	-0.460***	-0.475***	0.093*	0.094*
	(0.160)	(0.158)	(0.052)	(0.052)
Investor network constraint * Has twitter	0.108**	-0.143	0.006	0.014
	(0.053)	(0.141)	(0.016)	(0.030)
Investor network constraint * Twitter engagement	0.119***	0.498***	0.004	-0.008
	(0.046)	(0.145)	(0.016)	(0.041)
Gender diversity * Has twitter	-0.014	-0.143	-0.011	0.071
	(0.099)	(0.141)	(0.036)	(0.069)
Gender diversity * Twitter engagement	0.012	0.130	-0.009	-0.079
	(0.042)	(0.103)	(0.013)	(0.051)
Other Controls	Year Indicators, Firm Age, Total Funding Rounds, Website Rank, Google Trends			
Observations	34,120	33,372	34,120	33,372
R-squared	0.473		0.377	

Notes:

1. All columns report the OLS results with company fixed effect. Only connected companies are included because companies need existing investors as reference. Twitter engagement is standardized, where one-standard-deviation change is associated with about 300 Tweets and 30,000 favorites each year.

2. Columns (1) to (2) use the log amount of funding from new investors (as opposed to those that have invested the focal company before) in the next year as the dependent variable. Columns (3) and (4) use the log amount of funding from old investors (those who have made investments for the focal company before) in the next year.

3. Columns (2) and (4) use instrumental variable 2SLS. Two sets of instruments: (a) Twitter adoption and engagement of companies in the same state within two-round-of-funding distance; (b) Twitter adoption and engagement of companies that share a same investor in the past 5 years.

4. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Twitter Effect for First-Time Entrants vs. Serial Entrepreneurs

Dependent Variable	Log Lead Funding		Lead Network Diversity	
Type of Founders	First-Time Entrants	Serial Entrepreneurs	First-Time Entrants	Serial Entrepreneurs
VARIABLES	(1)	(2)	(3)	(4)
Has Twitter	0.343***	0.235	0.098***	0.024**
	(0.061)	(0.185)	(0.004)	(0.010)
Twitter engagement	0.334***	0.266**	0.012***	0.010*
	(0.034)	(0.116)	(0.002)	(0.006)
Other Controls	Year Indicators, Firm Age, Total Funding Rounds, Website Rank, Google Trends			
Observations	89,716	7,339	89,716	7,339
R-squared	0.375	0.329	0.780	0.768

Notes:

1. For column (1) and column (2), the dependent variable is the log funding in the next year; for column (3) and column (4), the dependent variable is the network diversity (i.e., 2 minus investor network constraint) in the next year. For isolated companies, network diversity is assigned a value of 0. Twitter engagement is standardized, where one-standard-deviation change is associated with about 300 Tweets and 30,000 reposts, comments and favorites each year.

2. All columns report the OLS results with company fixed effect. Column (1) and (3) show the results of start-ups founded by first-time entrants, and column (2) and (4) show the results of start-ups founded by serial entrepreneurs.

3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Twitter Content Effect on Isolated Companies

DV: Log lead funding	Company Fixed Effects			
VARIABLES	(1)	(2)	(3)	(4)
Inverse mills ratio (isolated)	-0.005	-0.005	-0.005	-0.001
	(0.196)	(0.196)	(0.196)	(0.197)
Has Twitter	0.636***	0.636***	0.636***	0.678***
	(0.075)	(0.075)	(0.075)	(0.074)
Gender diversity * Has Twitter	0.280***	0.280***	0.280***	0.428***
	(0.061)	(0.061)	(0.061)	(0.055)
Human Resource	0.035*			
	(0.018)			
Gender diversity * Human Resource	0.185***			
	(0.045)			
Business Announcement		0.036**		
		(0.018)		
Gender diversity * Business Announcement		0.187***		
		(0.046)		
Product Promotion			0.035*	
			(0.018)	
Gender diversity * Product Promotion			0.185***	
			(0.045)	
Customer Relations				0.017**
				(0.007)
Gender diversity * Customer Relations				0.064***
				(0.007)
Other Controls	Year Indicators, Market-Year Fixed Effect, Firm Age, Total Funding Rounds, Website Rank, Google Trends, Twitter Adoption Linear Time Trend			
Observations	64,039	64,039	64,039	64,039
R-squared	0.466	0.466	0.466	0.466

Notes:

1. Dependent variable is the log funding in the next year.

2. All columns report the OLS results with company fixed effect. The sample only include the isolated companies.

3. Human Resource, Business Announcement, Product Promotion and Customer Relations stand for the standardized value of the number of Tweets in each specific topic posted by each start-up in each year.

4. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Twitter Content Effect on Connected Companies

DV: Log lead funding	Company Fixed Effects			
VARIABLES	(1)	(2)	(3)	(4)
Inverse mills ratio (connected)	-0.113	-0.111	-0.112	-0.110
	(0.122)	(0.122)	(0.122)	(0.122)
Has Twitter	-0.257**	-0.255**	-0.254**	-0.262**
	(0.104)	(0.104)	(0.105)	(0.103)
Network constraint * Has Twitter	0.217***	0.215***	0.207***	0.228***
	(0.061)	(0.061)	(0.062)	(0.061)
Human Resource	-0.020			
	(0.014)			
Network constraint * Human Resource	0.028*			
	(0.016)			
Business Announcement		-0.019		
		(0.015)		
Network constraint * Business Announcement		0.032*		
		(0.017)		
Product Promotion			-0.015	
			(0.015)	
Network constraint * Product Promotion			0.034**	
			(0.015)	
Customer Relations				0.050**
				(0.025)
Network constraint * Customer Relations				0.077*
				(0.045)
Other Controls	Year Indicators, Market-Year Fixed Effect, Firm Age, Total Funding Rounds, Website Rank, Google Trends, Network Degree, Network Transitivity, Network Constraint, Twitter Adoption Linear Time Trend			
Observations	34,040	34,040	34,040	34,040
R-squared	0.492	0.492	0.492	0.492

Notes
1. Dependent variable is the log funding in the next year.
2. All columns report the OLS results with company fixed effect. The sample only include the connected companies.
3. Human Resource, Business Announcement, Product Promotion and Customer Relations stand for the standardized value of the number of Tweets in each specific topic posted by each start-up in each year.
4. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

CHAPTER 3: CAN START-UPS' SOCIAL MEDIA ACTIVITIES HELP THEM OVERCOME INSTITUTIONAL BARRIERS? EVIDENCE FROM CHINA

1. Introduction

Institutional environments, including formal institutions such as regulatory regimes and informal institutions such as social norms, play an important role in determining entrepreneurial success regardless of national contexts (Eesley et al. 2018, Eesley et al. 2016). Regulatory institutions directly determine whether start-ups can legally operate in many markets, and non-regulatory institutions such as social norms shape the influence of regulations on entrepreneurial outcomes. While there have been many studies examining the influence of institutional environments on entrepreneurial behaviors (Aldrich and Fiol 1994, Sine et al. 2005, Eesley 2016, Eberhart et al. 2017), most of them regard the institutional environments as exogenous to business development. Accordingly, limited attention has been paid to how firms, especially start-ups, can meaningfully influence their external institutional environments and improve them. Understanding and enabling this process is critical for propelling business success due to the heightened technological and institutional disruptions around the world in the recent years (Wu et al. 2020). For instance, when Uber and other firms in the sharing economy first entered the market, they encountered strict scrutiny from the governments and widespread resistance from industry associations. Similarly, the debate about blacklisting TikTok in the U.S. has raged on for several years. More recently, the awareness of personal data privacy issues has spawned data privacy laws, such as the General Data Protection Regulation (GDPR), now being implemented by governments around the world. As a result, business organizations are becoming more involved with regulatory changes that can directly affect their operations. Developing strategies to better manage and improve institutional environments is instrumental to success in these dynamic environments.

However, changing the institutional environment is notoriously difficult because it often involves changing cultural norms and appeasing existing political interests. In this study, we examine whether entrepreneurs can use social media to improve their standing in their institutional environment. Even if technology could alleviate even a small part of the disparities that start-ups face, it would be extremely helpful because social media and social media strategies are available to every firm, while the ability to amass the political resources to improve their standings in the institutional environment is available to relatively few.

For instance, firms can use social media to not only signal their social values and manage public opinion, but also mobilize public support that can result in preferential treatment. Li Ziqi, a typical e-shop owner in rural China, gained significant attention through her social media videos about her shop's agricultural products. Because her videos helped spread traditional Chinese culture, she quickly attracted the attention and support from the Chinese government and became the local cultural heritage promotion ambassador.[15] This governmental recognition has since brought many benefits to her business. For instance, she was allowed to post her videos on YouTube, which is banned by the Chinese government. As a result, she garnered traffic from around the world, and in 2021 broke the Guinness World Records for the most subscribers to a Chinese language channel on YouTube.[16]

Although these signals can help firms receive preferential treatment from government agencies, it is unclear who can benefit from social media because the effectiveness of these signals may depend on firms' existing political and social capital (Faccio et al. 2006, Peng et al. 2008, Angulo-Ruiz et al. 2019, Wu and Kane 2021) and their ability to effectively deploy the capital (Oliver and Holzinger 2008, Wang et al. 2012). On the one hand, social media may amplify existing political connections of politically connected firms by

[15] For example, see https://news.cgtn.com/news/2021-11-30/VHJhbnNjcmlwdDYwNzU5/index.html.

[16] See: https://www.guinnessworldrecords.com/news/2021/2/li-ziqi-breaks-youtube-subscribers-record-for-chinese-language-channel-647202. https://www.youtube.com/channel/UCoC47do520os_4DBMEFGg4A.

broadening a firm's access to even more political connections, creating a virtuous cycle of acquiring and converting capital to business advantages (Dorobantu et al. 2017, Tarafdar and Kajal Ray 2021). On the other hand, social media could mitigate the weakness of less politically connected firms who have limited channels available to connect to government agencies. Using social media, these firms can directly provide quality signals to governmental agencies (Wang and Wu 2019) without political connections. To critically examine what type of firms would benefit from using social media, we explore the two countervailing forces that could either exacerbate or alleviate disparities between those that are politically connected and those that are not.

Thus, in this study our goal is two-fold: First, we want to document whether social media can help firms improve their standing in the institutional environments they face, and if so, to discover the mechanisms through which social media facilitates this improvement. Second, we seek to determine whether social media can alleviate or exacerbate political disparities in institutional environments, a factor that could contribute to the understanding and narrowing of the digital divide in entrepreneurship (Bresnahan et al 2002, Forman 2005).

We use China as our research context because China is a major economy in the world with one of the fastest-changing institutional environments for business (Eesley et al. 2016, Eesley 2016). For example, the *Data Security Administrative Measure* in 2021 took less than two years from drafting to implementation. Coupled with wide-ranging economic activities and fast technological development, China provides an ideal setting to study the effect of institutions on entrepreneurship and the role of technology in mediating their effects. China's rapidly evolving institutional environment could be used to argue that our results do not apply elsewhere. But every country imposes institutional constraints, suggesting that discovering tools that help start-ups survive and circumvent them in one context could be useful in many contexts (Eesley et al. 2018).

We construct a novel dataset of 1,800 Chinese start-ups founded between 2011 and 2020 in the information technology sector. Examining the number of administrative approvals and news stories for each start-up before and after its adoption and use of social media, we find evidence that social media can alleviate institutional barriers as adopters are getting more administrative approvals and attracting more positive media coverage than peers that do not employ social media. In China, administrative approvals not only allow a company to legally operate in an industry, but also grant the company a quasi-monopoly market power in that industry. Furthermore, we find the social media's effect we uncover is stronger for companies with fewer political resources, suggesting that social media can mitigate the lack of corporate political connections in managing institutional environments.

To explore the underlying mechanisms of our findings, we use both the volume and the quality of social media content to examine whether our results are tied to information being transmitted on social media platforms. We show that social media benefits are associated with engagement but not mere adoption and the benefits are greater in the long term than the short term. Applying natural language processing to social media content, we find that corporate social responsibility, or CSR appear to have stronger effects on a company's standing in its institutional environments than others such as posts on marketing or customer relations. Collectively, these results show that social media content helps firms signal social value and establish public legitimacy, and these effects are from information transferred on social media as opposed to selection biases or reverse causality (e.g., that better start-ups choose to use social media).

To identify how social media affects the standing of a firm in the institutional environment, we use two sets of instrumental variables and an exogenous change in the institutional environments. Mass Entrepreneurship and Innovation Policy (MEIP), implemented in 2014, has significantly lowered institutional barriers for businesses in China. If social media activities are beneficial in alleviating companies' institutional barriers, we should observe such benefits to be higher prior to MEIP's implementation than after. We show that after MEIP, the benefits of social media have declined over the years. Additionally, we find using social media is particularly beneficial for start-ups in competitive industries where institutional barriers to reaching governmental agencies are high. We also find the benefit of social media is higher in the first two years after the start-up receives an administrative punishment or experienced a negative news event, both of which immediately raise the institutional barriers a company faces. To sum up, our results are twofold. They suggest that social media can help start-ups obtain institutional resources by providing quality information to the public and to government agencies. And, secondly, these benefits are greater for firms that face greater institutional barriers.

2. Literature and Theory

In this section we first describe why institutional environments are critical for start-ups and the existing ways start-ups use to mitigate their institutional challenges. Then we explain why social media can help address these challenges, and the kind of social media content that is most helpful.

2.1 Institutions Matter in Entrepreneurship

Institutional environments' effect on business can come from both formal and informal institutions (Eesley et al. 2018). Formal institutions enact formal sanctions, such as government regulations, licenses, permits, and fines (Scott 1995, Aldrich and Fiol 1994) while informal institutions are less visible but similarly powerful. They relate to socially accepted beliefs, values, and norms (Scott 1995, Aldrich and Baker 2001, Aldrich and Yang 2012). Prior research has established theoretical and empirical foundations to understand how institutional environments affect entrepreneurial activities (Suchman 1995, Sine et al. 2005, Eesley 2016) at both firm and industry levels. For individual firms, lowering institutional constraints often fosters entrepreneurship. For instance, Meek et al. (2010) find that the founding rate of solar companies increased significantly after the government lowered institutional barriers to entry. Sine et al. (2005) show that more energy firms were founded after a new policy opened the market for the sale of electricity. Eesley (2016) find that lowering institutional constraints can facilitate entrepreneurship in China. However, in a few scenarios, lowering institutional constraints can also suppress founding rates, such as when competition is mild (Dobbin and Dowd 1997). For industries, the institutional environment also shapes industry and regional practices. Strang and Bradburn (2001) show how the diffusion and modification of health maintenance industry policy was shaped after the government eliminated restrictive regulations. Rajagopalan and Finkelstein (1992) find that industries tend to offer greater compensation and performance pay when the regulatory environment allows for high degrees of discretion. Eesley et al. (2018) find that national and regional differences in the types of entrepreneurial activities and levels of business growth can be explained by misaligned formal and informal institutions.

While the literature explores the critical impact of institutional environments on entrepreneurial activities, the extant research rarely examines how start-ups can actively meet the challenges posed by their institutional environments. This is particularly salient in sectors where regulations usually lag technological development, such as in the high-tech industry. Understanding how start-ups can actively change institutional environments in their favor is critical because existing mechanisms to mitigate institutional risks are often restricted to established firms that have the resources to influence regulation (Feinberg and Gupta 2009). Start-ups, lacking these resources, need to develop different strategies to mitigate institutional risk especially when it involves new sectors and new business models (Li and Zhang 2007, Armanios et al. 2017).

2.2 Corporate Political Connections and Institutions

Nonmarket-based corporate strategies have examined how firms use political action to manage their institutional environments (Lux et al. 2011, Mellahi et al. 2016). These actions include lobbying efforts, donating to campaigns, and building political connections (Sun et al., 2012, Hadani and Schuler 2013). For start-ups, political connections, such as having government ties, are even more important than for established firms because they do not have the financial resources and social influence to participate in campaign donations and lobbying (Khaire 2010, Armanios et al. 2017).

Prior literature examines how companies use political connections to enact public policy that favors their business environments (Hillman 2005, Sun et al. 2012, Dorobantu et al. 2017). Economic literature in a resource-based view suggests that firms can deploy political connections to capture economic rents (Holburn and Zelner 2010) or "excess returns" that are inconsistent with competitive market outcomes (Tollison 1982). Opportunities to pursue economic rents can emerge when there are "government restrictions upon economic activity," such as barriers to market entry and the discretionary authority to award or punish firms (Krueger 1974). When such opportunities arise, politically connected firms can often capture those rents using special contracts, licenses, bailouts, rules, or exceptions. These rents can be substantial even when corruption levels are low (Faccio et al. 2006, Amore and Bennedsen 2013, Goldman et al. 2013). According to the corporate political strategy literature, political decision makers behave as self-interested actors and enact public policies favorable to certain businesses in return for constituency supports or financial contributions such as taxes and employment (Buchanan 1987, Hillman and Hitt 1999, Schuler et al. 2002). Thus, companies that can demonstrate public legitimacy and social values are more likely to participate in such exchanges with the government. Traditionally, political connections are the primary vehicle for facilitating the exchange (Hillman 2005, Frynas et al. 2006). They not only provide firms with access to policy makers to influence policies, but also provide political legitimacy and information advantage (Dieleman and Boddewyn 2012, Cui et al. 2018, Meyer et al. 2018).

Despite the many advantages of corporate political connections, forming these relationships is difficult because politicians already have existing social networks and interests (Siegel 2007, Cohen et al. 2011, Haselman et al. 2018). Thus, unless the founding teams have already established political connections derived from their personal relationships, it is difficult for a start-up to quickly establish relationships with policy makers (Armanios et al. 2017, Bertrand et al. 2018). If an alternative that democratizes information transmission and promotes the legitimacy of a start-up to the public exists, it would help start-ups that lack these connections to access political bodies and to improve their standing in the institutional environment.

2.3 Social Media as an Information Tool

It has been well established in the literature that by reducing the cost of information dissemination and improving efficiency, social media are effective at brand building, consumer engagement, and long-term company performance (Angst et al. 2010, Aral and Walker 2011, Wu 2013). For entrepreneurship, social media can reduce information asymmetry and thereby help start-ups acquire greater venture financing. For example, Aggarwal et al. (2012) find that online reviews and judgments can serve as quality signals about start-ups and influence their subsequent chance of garnering such capital. Wang and Wu (2019) illustrate that social media content relating to business operations can alleviate information asymmetries in the venture capital market and, further, that the benefit of using social media is especially high for start-ups founded by women and new entrepreneurs. These studies collectively show that social media can facilitate entrepreneurial success by enabling the founders to signal quality, reduce information asymmetry between the company and external markets, and strategically access key resources.

A second stream of the literature examines the mobilizing role of social media in directing public opinions and mobilizing social support, from the Arab Spring and the Egyptian protests (Ghannam 2011, Tufekci and Wilson 2012) to the MeToo and Black Lives Matter movements (Tarafdar and Kajal Ray 2021). Research has found that social media can facilitate social movements by broadcasting information, gathering support, and building social networks (Alterman 2011, Choudhary et al. 2012, Gladwell and Shirky 2011). Accordingly, social media are not just passively transmitting information in the public space, but are also fostering, forming, and directing public opinion. Start-ups can thus use social media to mobilize support and change both formal and informal institutional environments in their favor. Managing both types of institutional environment is important because the two systems can interact with each other: government policies can reorient public opinion, especially in China since most media are under the control and censorship of the government (Tong 2010, Xu and Albert 2014). Conversely, public opinion can also affect government policy (North 1990, Scott 1995). For example, in nascent industrial sectors, the Chinese government often uses public opinion to decide whether to give a company exemption to an existing regulation (Luo 2014, Tang et al. 2018). Social media are extraordinarily efficient at broadcasting information about a start-up to the entire public space that includes both the formal administrative institutions and informal public opinion, so companies are more likely to establish public legitimacy in public sectors and receive favorable rulings on policies from policy makers by using social media. Thus, we hypothesize that social media can provide institutional benefits for start-ups:

H1: Using social media has a positive effect on a start-up's standing in its institutional environment.

As stated above, firms with fewer political connections lack the channel to signal to administrative agencies about their quality and establish legitimacy. Social media could alleviate this information disadvantage by increasing a government's knowledge about a start-up. Start-ups can demonstrate quality and social values through social media to the public, and the success of such efforts (i.e., public approval) on social media can be observed by administrative agencies (Xiao 2011, Brouwer 2020). Collectively, social media can help firms gain institutional support. Although social media can help all firms, the relative benefit can be larger for firms that lack political connections. This is because the government usually grants licenses once the company has attained certain legitimacy, but past a certain point, legitimacy would not necessarily bring additional institutional benefits to the business. Thus, for politically connected companies that already attained legitimacy with the government, the added benefit from social media would be minimal. However, for firms lacking these relationships, additional signals on social media can increase their visibility and help initiate relationships with government agencies which can potentially improve how they are perceived by the institutional environment they face.

Therefore, we hypothesize that firms with fewer political connections should gain more institutional benefits from using social media than do firms with more political connections:

H2: Start-ups with fewer political connections receive more institutional benefits from using social media than do start-ups with more political connections.

2.4 Corporate Social Responsibility (CSR) News as a Signal

Although social media can help transfer all kinds of information, certain types of information might be more helpful than others in improving a start-up's legitimacy. Literature has shown that CSR-related news is more likely to attract attention from the public sector than news on other topics and can accordingly help firms improve their reputation and obtain political resources (Deng et al. 2013, Cahan et al. 2015, Lins et al. 2017, Hwang et al. 2021). For instance, Werner (2015) shows that policy makers usually evaluate firms based on their sociopolitical reputations, and firms with higher CSR reputations have greater access to the government. Hong and Liskovich (2015) find that firms with high CSR ratings pay lower fines if they are prosecuted for bribery. CSR is also important in China as the central government has taken the social governance level of local businesses as a critical indicator to evaluate and promote local government officials.[17] By propagating one's social values on social media, such as showing commitment to local people, especially to minorities (Malpani and Ghosh 2020, Soundararajan et al. 2017), start-ups are more likely to establish a social reputation, thereby receiving more institutional benefits. Therefore, we

[17] For example, in 2012 Beijing's municipal government issued an official document to include the environmental protection index and other civil welfare indicators in the performance evaluation of officials.

hypothesize that firms should receive more institutional benefits from sending social media posts that are CSR-related than from posts on other topics:

H3: Start-ups receive more institutional benefits from sending social media posts that are CSR-related than from posts on other topics.

As stated above, the institutional environment is comprised of formal and informal institutions, both of which can affect a firm's public legitimacy. Thus, the institutional benefits that start-ups obtain from using social media should be reflected in both formal and informal institutions simultaneously.

3. Chinese Entrepreneurship and Policy Change

To answer our research questions, we empirically study whether entrepreneurs in China can use social media to reduce the institutional barriers that they face, and if so, whether these effects alleviate or exacerbate existing disparities due to political connections. There are a few benefits to setting the research context in China. The most important reason is that China has one of the fastest-changing set of economic institutions in the past half century (Wang 2016, Gilley 2008, Li et al. 2000, Tsai 2006). Since the 1970s, after many large-scale reforms and countless smaller reforms, the Chinese government has gradually explored a set of socialist market-based economic systems. On the one hand, there is an internal drive to change existing planning-based systems to improve efficiency; on the other hand, the rapid development of the economy and technology is also forcing institutional changes (Krug and Hendrischke 2008). Under the combined effects of internal and external factors, the Chinese government has iterated several versions of economic and institutional systems over the past few decades. These reforms provide an ideal setting for examining the effect of the institutional environment on entrepreneurship and IT strategy (Eesley 2016, Armanios et al. 2017, Eesley et al. 2016).

One significant reform is the Mass Entrepreneurship and Innovation Policy (MEIP), which was implemented in September 2014. It has drastically lowered institutional barriers on entrepreneurial activities and thus provided a natural experiment to understand how institutional changes affect these activities. Before the reform, institutional barriers primarily involved administrative examinations and pre-registration approvals. Prior to MEIP, to start operations in any business area, start-ups needed to submit applications to the government and go through a stringent but opaque vetting process. This process can be especially difficult when it involves an area that is previously restricted by law or by monopolies from state-owned enterprises. For example, for ride-haling companies like Uber, before entering any city, one must first obtain a business license from the local government. It could take thousand pages of materials and several weeks or longer to go through an approval process; in some cases, it could take years.[18] Despite the lengthy process, the approval rate was extremely low (Zhang and Rosenbloom 2018).

After the MEIP was implemented, the need for administrative examinations and pre-registration approvals has been rapidly eliminated. This allowed start-ups to enter business areas that had been previously prohibited, such as telecommunications, finance, and public transportations (Ahlstrom et al. 2018, Lo et al. 2019). Led by the premier of China, this policy was implemented at all levels of administrative branches across the country (Chen and Wang 2021, Gao and Mu 2021). The scale and intensity of this policy reform was so unprecedented that many scholars and entrepreneurs see it as an exogenous shock to alleviate institutional barriers for entrepreneurial activities (Wu et al. 2020, He et al. 2019). This policy has also effectively eliminated the need for firms to find ways to improve their approval chances. Prior to MEIP, firms often use social media to make public disclosure and civic engagement to improve their chances of

[18] For example, http://www.npc.gov.cn/npc/c541/200204/2b818b78b91841648e7b2caf9b65568a.shtml from the National People's Congress of the People's Republic of China has demonstrated a real case: to obtain the administrative approval to build a wholesale market in 2004, one needed to acquire approvals from 112 different government departments. In an optimistic scenario, if each department takes three days to process, the total approval time would take at least a year.

receiving public support and favorable policies and exemptions (Li et al. 2016, Ye et al. 2017). With the implementation of MEIP, firms have less need to mobilize public support and to obtain preferential policies, and firms accordingly have less need to use social media to attain these goals.

As a mechanism check, we expect the social media benefits enjoyed by Chinese entrepreneurs should be more salient when the institutional barriers are high, and lowering the barriers would subsequently reduce the benefits. We assess the impact of social media posts on firms' abilities to work with their institutional environments. However, MEIP reduced the need to accomplish what the social media posts provided: mobilize public support and obtain preferential policies. We would then expect that the impact of using social media was vastly diminished after MEIP was implemented.

4. Data and Measures

4.1 Company Information Data

We acquired firmographics and financial information of start-ups from Crunchbase which is one of the most widely used databases for global entrepreneurial research (Cumming et al. 2016, Dalle et al. 2017). In Crunchbase, there are a total of about 1,800 Chinese start-ups[19] that were founded after the year 2011. We then match these start-ups with the Chinese Enterprise Database, which curates comprehensive information about all the registered companies in China, either from public data sources or from proprietary governmental databases. From this data, we obtained firm location, industry category, subsidiary and branch information, yearly personnel information (the number of social security payers), social media accounts, founder and board member information (gender, education level and resume of past experiences), the detailed administrative approvals and punishments from all governmental departments, the detailed

[19] Data acquired in 2020.

certificate, trademark, patent and copyright information, and also all the online news reports that are related to the focal company (including full report and posting time) from over 100 major news agencies (including both private and state-owned agencies) in China, and coded the emotions of each news report (which were rated as positive, negative or no emotion). These data are collected by a government-funded smart city company and are validated through other similar data sources such as Tianyancha. The detailed collection of these companies also shows how the Chinese regulatory agencies monitor business environments and public opinions about firms.

4.2 Social Media Data

We use WeChat Public Accounts to measure a company's social media activities as WeChat is the most widely used social media in China and nearly all firms have a presence on the platform. Firms use WeChat heavily to communicate with investors and local administrative agencies.[20] On WeChat, each company can register only one official public account (registration and verifications are needed), and they can publish multi-media articles through that account. The published content is public and appears in the feed of subscribers to the account. Several metrics are used to measure engagement. A "reading" is recorded if the published article is visited by a unique IP address. Readers can also comment on or "like" an article. We identify a start-up's official WeChat public account through the Chinese Enterprise Database and manually check those that do not match. A firm decides whether to create a WeChat public account, and if it does, what content to provide. For each start-up's official WeChat account, we obtain the time the firm started to use this account and all the articles published in this account, including the title, text, publication time, and the total number of readings, comments and likes for each article. On average, companies post about 80

[20] For example, http://www.baohe.gov.cn/zmhd/zwwxq/index.html shows the QR code for the WeChat discussion groups established by local administrations in Hefei, China. In China, many local governments have created such WeChat groups to better communicate with residents and local businesses.

articles and receive 500,000 readings, 200 comments and 4,000 likes annually.

4.3 Variables

Dependent Variables

To measure one's formal and informal institutional environments, we use the **Number of Administrative Approvals** and the **Number of Positive Media Stories** that a company receives in each year as dependent variables. They represent the level of legitimacy attained by both the formal and informal institutions. The number of administrative approvals that the company receives in each year is a proxy of a company's legitimacy level as recognized by the formal institutions because it directly determines whether a company can operate in a specific industry. The process of issuing administrative approvals is not only very stringent, but also the Administration Law of China does not allow licenses to be transferred or purchased. The number of positive media stories that the company receives each year is a proxy of a company's level of legitimacy in informal institutions because major media outlets can serve as a proxy for a start-up's standing in public opinion. By using the number of administrative approvals and the number of positive media stories, we can comprehensively measure the effect of social media on both the formal and informal institutional environment.

Independent Variables

To measure one's engagement level on social media, we construct the dummy variable **Has WeChat** to indicate whether the company has created a WeChat public account that year, and **WeChat Engagement**, which is calculated as the first principal component of the total number of articles posted, number of readings, comments and likes received for each company in each year. After demeaning the variable, a one standard deviation of WeChat Engagement is associated with about 200 articles, 5 million readings, 2,000 comments and 30,000 likes per year.

To measure different social media topics, we select several sets of keywords to capture the common content topics based on a firms' WeChat posts. These topics include business announcements, marketing promotions and customer relations and they have been shown to have a positive impact on start-ups' future funding amounts (Wang and Wu 2019). To measure a company's social values, we create a new keyword dictionary to capture the content related to CSR (Appendix B1). We translate these keywords into Chinese words and their synonyms and identify the number of articles in each of the four topics that each company posts in each year by counting the number of articles that contain the specific keywords. Thus, one article can contain several topics. The number of articles in each of the four topics are used to measure different social media topics posted by each company in each year. We then demean the measure and divide it by its standard deviation.

Other Variables

For each start-up, starting with the year that it was founded, we construct firm level control variables such as Age, Number of Subsidiaries, Number of Branches and Number of Employees, company quality information such as Number of Trademarks, Number of Patents, Number of Copyrights and Number of Certificates, company funding information such as Total Funding Rounds, Historical Funding Amounts. We also use Funded by Tencent to indicate whether the focal company has received investment from the technology company that developed and owns WeChat.

There are also time-invariant variables including the **Location (City)**, **Industry Category** and resumes of founding team members based on the Crunchbase dataset and Chinese Enterprise Data. Using the resumes of founding team and board members, we identify politically connected people as those who have prior working experiences at governments or state-owned enterprises (Li et al. 2008, Okhmatovskiy 2010). We calculate the **Company Political Distance** of each start-up by taking the reciprocal of the number of politically connected people plus 1 (so that politically distant companies have fewer connections). By construction, this variable is fixed over time. To enhance identification, we also create moderators to measure whether a company faces extraordinarily constrained institutional environments. Based on the metric suggested by Campa and Kedia (2002), we calculate the level of **Administrative Competitiveness** of each industry category as the fraction of companies within an industry sector that received administrative approvals in each year (competitiveness is one minus this value suggesting administrative approvals are harder to obtain in a more competitive industry). We also construct the **Number of Administrative Punishments** for each company in each year, which are the penalties imposed by regulators on the company's illegal activities, usually in the form of fines. Such punishments rarely occur, and we use this variable to capture the circumstances when the focal company's legitimacy level is low as recognized by formal institutions. We also construct the **Number of Negative Media Stories** for each company in each year to capture the circumstances that correspond to low legitimacy levels and low public opinion. Appendix B1 shows detailed variable descriptions.

The summary statistics and correlation table for the data are reported in Table 9 and Table 10. Our unbalanced panel dataset includes 10,641 observations for about 1,800 companies from 2011 to 2020.

5. Empirical Methodology

5.1 Main Model

To examine whether social media can alleviate institutional barriers and mitigate existing disparities, we explore the effect of WeChat usage as well as the interaction effect between WeChat usage and company political distance after controlling for firm characteristics. We first fit the following OLS model with company fixed effect (γ_i) and time dummies (τ_t) for all companies focusing the effect of social media on formal institutions, as measured by administrative approvals and then we estimate the effect on informal institutions, as measured by positive media stories:

$$
\begin{aligned}
\{&Administrative\ Approvals_{i,t+1}, Positive\ Media\ Coverages_{i,t+1}\} \\
&= \beta_0 + \boldsymbol{\beta_1 has\_wechat_{i,t}} + \boldsymbol{\beta}_2 \boldsymbol{wechat\_engagement_{i,t}} + \boldsymbol{\beta}_3 \boldsymbol{has\_wechat_{i,t}} \\
&\times \boldsymbol{political\_distance_i} + \boldsymbol{\beta_4 wechat\_engagement_{i,t}} \times \boldsymbol{political\_distance_i} \\
&+ controls + \tau_t + \gamma_i + \varepsilon_{i,t}
\end{aligned}
$$

A significant benefit of a firm fixed-effect model is that all the time-invariant characteristics of the start-ups, such as industry categories and the backgrounds of founders, are controlled in the regression. Note, we lead the dependent variable by one year to avoid potential reverse causality and simultaneity issues. As most companies do not obtain institutional benefits every year and some may never seek such benefits, there are many zeros in the dependent variables. This problem is partially addressed by our company fixed effect model if such firm characteristics are fixed over time, and we also use a Heckman Selection model and calculate the inverse Mills ratios to account for a firm's decision for institutional benefits.

Furthermore, beyond two-way fixed effect models, we also conduct additional robustness checks using alternative econometric specifications including coarsened exact matching and negative binomial regressions. To avoid potential biases caused by different social media adoption timing in our panel data (Goodman-Bacon 2021), we verify our fixed effect estimators using a staggered difference-in-difference doubly-robust estimator introduced by Callaway and Sant'Anna (2021). To reduce the impact of endogenous factors in the difference-in-difference setting caused by unbalanced parallel trends between WeChat adopters and non-adopters, we test our results using the synthetic difference-in-difference method proposed by Arkhangelsky et al. (2021). All these results are listed in Appendix B2. They are consistent with our main results using two-way fixed effects models.

5.2 Identification Strategy

We use three ways to identify for both the choice to adopt social media as well as the level of social media engagement in each firm. First, all our models are firm-fixed effect models. We then use two sets of instrumental variables to address the endogenous firm choice to adopt and use social media in the first place. For adoption choices, we also use Heckman selection models. Lastly, we use the implementation of MEIP to show that the effect of social media should diminish as it becomes less important to obtain exemptions in a looser regulatory environment.

It is possible that both the social media activities and the institutional outcomes are spuriously correlated due to unobserved characteristics such as start-up quality. For example, while any firm can create a WeChat public account, bigger firms have the motivation and resources to build social media engagement. These issues can be partially addressed using a firm fixed effect model, separating social media adoption and engagement to avoid the potential "slack resources" hypothesis, and using institutional outcomes in future years as dependent variables. Moreover, one of the advantages of our data is that we have detailed company quality information in each year, including the number of employees, certificates, trademarks, patents and copyrights, etc., which can help us control for the influence of some factors related to the quality of the company.

In addition, we use two sets of instrumental variables to address other forms of endogeneity or omitted variable biases associated with company-specific, time-varying characteristics (such as unobserved start-up "quality") that may affect both social media activity and institutional success simultaneously. The first instrumental variable is the yearly social media adoption level of the local municipal government in the city where the focal company is located. First, the adoption of WeChat by a local government is highly relevant to the adoption and use by local firms because local businesses tend to follow actions adopted by a local government such as the use of communication technologies. Second, the specific timing of adoption by a local government varies from place to place because a higher administrative unit (e.g., the provincial government) determines when its local governments adopt social media. As a result, the adoption decision by a municipal government is exogenous to the municipal institutional environment or local business activities because the adoption is made at the provincial level. By being highly relevant to local business adoptions as well as being quasi-random, we can use a local government's adoption of social media as an instrumental variable for a firm's adoption and use. We collect the total number of WeChat public accounts for all departments of each municipal government directly from official government websites. We manually identify the time when each account was registered and use the yearly total number of accounts as the instrumental variable.

Second, we use the average level of social media activities of firms in the same provincial region across different industries to instrument for firms' own social media activities. Start-ups in the same region are likely to be influenced by similar media factors, such as having similar types of social media users. Previous literature has shown that the diffusion of internet technology and online interactions are similar within a geographical region (Forman et al. 2002, Forman et al. 2005, Kulshrestha et al. 2012, Kim et al. 2013). Accordingly, WeChat usage for firms located in the same region is likely to be influenced by the number of WeChat users in the region and users' propensity to interact with start-ups online. However, other firms' social media activities should not directly influence a start-up's own institutional outcomes, especially when they are in different industries. Figure 7 shows the provincial distribution of the start-ups: 25% in Beijing, 15% in Shanghai, 15% in Guangdong, and the rest are almost evenly distributed across the regions. The distribution has sufficient variations to provide necessary first stage power for two-stage least squares (2SLS) regressions.

For every 2SLS estimation in the following sections, we use these two sets of instrumental variables to instrument for the WeChat activities and use the product of the above-mentioned two sets of instrumental variables and political distances of one's geographical neighbors to instrument for the interaction terms in the regressions. We also run two specification tests for every 2SLS estimation in the following sections, including the weak instrumental variable test and Hansen's over-identification test. For each 2SLS estimation, the weak instrument F-statistic is greater than 20, indicating that the two sets of instruments are relevant; Hansen's J-statistic is not statistically significant, suggesting that the two sets of instruments do not move in opposite directions. Since the two instruments are collected independently from separate sources, such results provide further credibility for our overall results.

Lastly, we use the implementation of MEIP as an exogenous shock on the need to use social media to receive public support. Given MEIP is largely an exogenous change that affects all of China, we can examine the effect of social media on the changing institutional environment before and after the MEIP. We further explored the mechanisms by examining the content of social media posts.

6. Results and Interpretation

6.1 Main effects

We first analyze the effect of social media on formal institutions in Table 11. Columns (1) and (2) are fixed-effect regression results, and columns (3) and (4) use 2SLS. The dependent variable is the number of administrative approvals in the next year. All four models use a Heckman selection model with an inverse Mills ratio to correct for possible selection biases, with control variables including start-up age, number of employees, number of patents, number of certificates, total funding rounds, etc. to control for possible unobserved start-up qualities. Column (1) shows that WeChat Engagement is positively associated with more administrative approvals in the next year: about 0.7 more administrative approvals are generated for a one-standard-deviation increase in WeChat engagement; that deviation is about 200 articles, 5 million readings, 2000 comments and 30,000 likes per year, and the estimated coefficient is statistically significant. Such an effect is substantial as administrative approvals are usually hard to obtain and even one license can have a huge impact on the company's future business. Column (3) shows that the 2SLS results are directionally consistent to fixed effects results. Notably, the coefficient of social media adoption is not statistically significant when controlling for engagement, suggesting that having a WeChat public account is not enough; companies need to engage actively on the platform to receive the benefits. Such results are also inconsistent with reverse causality that better or better resourced firms adopt social media due to slack resources, in which we should observe a strong positive effect on WeChat adoption. All these results support H1 that social media activities help start-ups receive more institutional benefits.

We then explore whether social media can alleviate the current institutional disparities. Column (2) shows that social media use is correlated with more future administrative approvals for companies that start out politically distant: for companies without any political connections (so that the political distance is 1), a one-standard-deviation increase in WeChat engagement is associated with about 0.2 additional administrative approvals. Similarly, after controlling WeChat engagement, the coefficient of WeChat adoption is not statistically significant, suggesting that mere adoption is not useful: politically distant companies need to actively engage on social media to generate the institutional benefits. Column (4) shows that the 2SLS results are directionally consistent to fixed effects results. All these results support H2 that social media help start-ups with fewer political connections to obtain more institutional benefits than connected ones.

Next, we analyze the effect of social media on informal institutions in Table 12. Columns (1) and (2) are fixed-effect regression results, and columns (3) and (4) use 2SLS. The dependent variable is the number of positive media stories in the next year. All models use a Heckman selection model with control variables that are the same as in Table 11. Column (1) shows that social media activities are positively associated with more positive media stories in the next year: about 18 more positive media reports outcome for a one-standard-deviation increase in WeChat engagement. Such results suggest that social media engagements help improve start-ups' legitimacy level in public opinions. Column (3) shows that the 2SLS results are directionally consistent with fixed effects results.[21] All these results support H1 that social media activities help start-ups receive more institutional benefits.

Table 12 also shows whether politically connected individuals receive greater or less benefit from using social media when compared to less connected individuals. Column (2) shows that social media use is correlated with more future positive media stories for companies that are politically distant: for companies without any political connections, a one-standard-deviation increase in WeChat engagement is associated with about six more pieces of positive media reporting. Similarly, after controlling for WeChat engagement, the coefficient of WeChat adoption is not statistically significant, suggesting that mere adoption is not driving the result. 2SLS results are directionally consistent with fixed effects results. All these results support H2 that social media help start-ups with fewer political connections obtain more institutional benefits from both formal and informal institutions compared to those with more political connections.

6.2 Content Analysis

We then explore whether our results are driven by a certain type of social media content. In Table 13, we show the effects of different topics on the company's future institutional environments. We explore four broad categories of social media topics: (1) Business Announcements, (2) Marketing Promotions, (3) Customer Relations, and (4) CSR.

Columns (1) to (5) show the results on future administrative approvals and columns (6) to (10) show the results on future positive media stories. Among different topics, social media content related to CSR are the most beneficial for improving the start-up's institutional environments, both statistically and

[21] Due to the significant increases in the coefficients in the 2SLS regressions, we restrict our discussion to the sign of these coefficients, although these could suggest the magnitude of the effects we observe are in reality larger than what is suggested by the OLS results.

economically. All other topics have much smaller effects both in their effect size as well as statistical significance. These results suggest that to obtain institutional benefits, start-ups need to signal their social values rather than other topics (such as business values), supporting H3.

7. Mechanism Check

In this section we examine the impact of social media in different institutional settings to verify the mechanism, as the effect of social media should be different when start-ups face different institutional environments. Specifically, we study the effect of social media before and after MEIP was implemented. We also examine the effect of social media in industries where institutional benefits are harder to obtain, and during the time when the start-up's public legitimacy level is low.

In Figure 8, we show the yearly regression results of WeChat engagement on one's institutional environments. Specifically, the left figure shows the yearly point estimator with 95% confidence intervals of WeChat engagement regressing on the lead number of administrative approvals. After 2014, the estimated coefficient has significantly decreased toward 0, and is no longer statistically significant after 2016, suggesting that the effect of social media in obtaining formal institutional benefits is significantly weakened after MEIP was implemented. Similarly, using the lead number of positive media reports as the dependent variable, the right figure shows that the effect of social media in obtaining informal institutional benefits is also weakened after MEIP's implementation, while the effect vanishes more slowly in comparison to the left figure as informal institutions usually take longer to change compared with formal institutions (Scott 1995). Both results support the mechanism that social media posts are more effective in obtaining institutional benefits when institutional barriers are high.

Next, we examine the effect of social media in industry categories where administrative approvals are harder to obtain. The institutional requirements vary substantially from one industry category to the next. In industries once monopolized by state-owned enterprises or involved with national security, the administrative requirements for licensing could be very stringent. Prior literature has shown that additional information can be crucial for decision-making when there are many alternatives or when facing high uncertainty (Forman et al 2008, Lu et al 2019, Snir and Hitt 2003). In such industries where institutional barriers are extraordinarily high, having an additional information channel to reach out to administrative agencies, signal social values, and establish public legitimacy could be critical for entrepreneurs. We show the results in Table 14.

In Table 14 we show the effect of social media activities in different levels of institutional competitiveness. Columns (1) to (4) use two-way fixed effects regression model and column (5) uses 2SLS. The results show that, despite the negative association between institutional competitiveness and future administrative approvals, WeChat Engagement is more beneficial in industry categories where institutional competition is stronger. The 2SLS results are also directionally consistent.

Finally, we examine the effect of social media under institutional shocks. Here we identify 2 types of institutional shocks that have a negative impact on the institutional environments of a company: administrative punishments and negative news reports, both of which can hurt the public legitimacy of a start-up. We define the two years since such institutional shocks happen to the focal company as an institutional shock period and create a dummy variable indicating whether the company is in an institutional shock period at the focal year. During such periods when one's public legitimacy is extraordinarily low, activities that can signal one's social value and mobilize public support would be especially helpful in establishing one's public legitimacy. We show the results in Table 15.

In Table 15 we show the effect of social media activities during institutional shocks when the public legitimacy of a company is challenged. Columns (1) to (4) use a two-way fixed effects regression model and column (5) uses 2SLS. The results show that WeChat Engagement is more beneficial when companies are facing legitimacy challenges. The 2SLS results are directionally consistent, supporting our mechanism that social media are more effective in periods when the company's institutional barriers are higher.

8. Discussion

Institutional environments play an increasingly important role in entrepreneurial activities. Our results provide new evidence that social media can help entrepreneurs overcome institutional barriers, and that such an effect is more beneficial for start-ups with fewer political connections over better connected firms. These results suggest that firms can mitigate the political disparities they face by effectively using social media to promote their business interests. By using social media to signal one's social values, start-ups can influence public opinion to obtain preferential policies from administrative agencies, thereby establishing public legitimacy. Consistent with this mechanism, we find that the effect of social media is stronger in situations when and where institutional barriers are higher (such as before MEIP, in more competitive industries, or during negative institutional shocks), and only social media topics related to CSR (rather than business activities) are effective in managing institutional environments. We also find that our results are robust under various model specifications and matched samples. These results are promising as they reveal some concrete steps that can improve reputation in start-ups, such as encouraging social media engagement and direct communications with administrative agencies. Such results can be generalized across various national contexts due to the prevalence of institutional challenges for entrepreneurship, the frequent institutional changes in the recent years, and the digital divide caused by network capital around the world.

There are a few limitations to our study. First, entrepreneurs may spread inaccurate signals about their social values on social media platforms. For example, Li and Wu (2020) show that some companies never make effective efforts at CSR even after publicly committing to implementing certain CSR programs. Although government agencies usually conduct rigorous scrutiny before granting administrative licenses, which can partly mitigate the risk of distorted signals, we recognize that we cannot distinguish between the "real social value" and "inflated value" from social media articles in our context. Regardless of the degree of fraud or inflation in company claims about their CSR programs, the fact remains that only posts about CSR have a significant impact on a company's standing in its institutional environment.

Second, there are other forms of institutional barriers and resources faced by start-ups and entrepreneurs, such as laws, regulations, government bailouts and procurements, etc., while in this article we examine only administrative approvals and online news reports as proxies for formal and informal institutional environments. Future research could further examine the institutional nuances.

This paper offers important managerial and policy implications. For start-ups and entrepreneurs, social media posts focused on CSR offer a low-cost channel to signal social value and establish public legitimacy, which can in turn help companies gain preferential treatment from the government, especially for disadvantaged companies that have fewer political connections. This would benefit not only start-ups and entrepreneurs in their business operations, but also the broader innovation and entrepreneurship environment. For administrators in public sectors, social media offer an important information channel to better evaluate the quality of a company, enhancing their ability to invest public resources focused on companies with higher social values. Government should pay more attention on social media platforms to better monitor the market development, and direct communications on social media between companies and governmental agencies should be further encouraged.

This paper contributes to the institutional theory of entrepreneurship. Most of the existing studies in institutional theories treat institutional environments as exogenous factors affecting entrepreneurial activities, while in this study we examine how start-ups can utilize social media posts on CSR efforts to improve their standing in institutional environments and obtain access to institutional legitimacy. Our results suggest that early-stage firms should not only establish a presence on social media platforms, but also effectively engage by signaling the social value of the company and mobilizing public support, especially for those firms with fewer political resources.

This paper also provides insights into the “digital divide” and promoting diversity. Over the past few decades there has been a great amount of research on how information technologies affect diversity. While many studies have found that information technology increases disparities (e.g., the digital divide), we show how social media can also be used to equalize information transmission to some extent. This has reduced barriers to access resources and improved equity and diversity in entrepreneurship. We recognize that political connections and other factors continue to play a significant role in creating disparities, and it is difficult for any technology to fully erase these disadvantages. However, by showing that technology can help even a small part of the institutional disadvantages that actors face, it is a promising direction because it is much harder to change culture, norms, and institutions than it is to adopt and use social media technology. We also show that simply using social media is not enough, a specific social media strategy needs to be tailored for each situation. To improve a firm’s standing in its institutional environment, we show that CSR related posts are most effective. By exploring the mechanism of how social media can alleviate disparities in political connections and overcome barriers in institutional environments, we show the potential of technology to help disadvantaged entrepreneurs establish public legitimacy.

Figure 7: Start-up Regional Concentration at Provincial Level

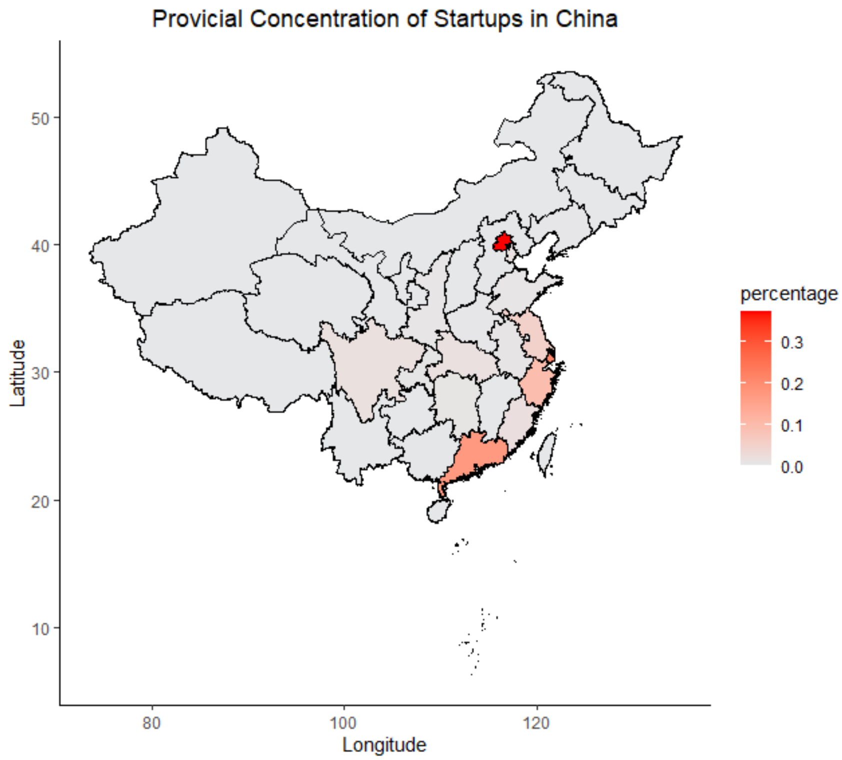
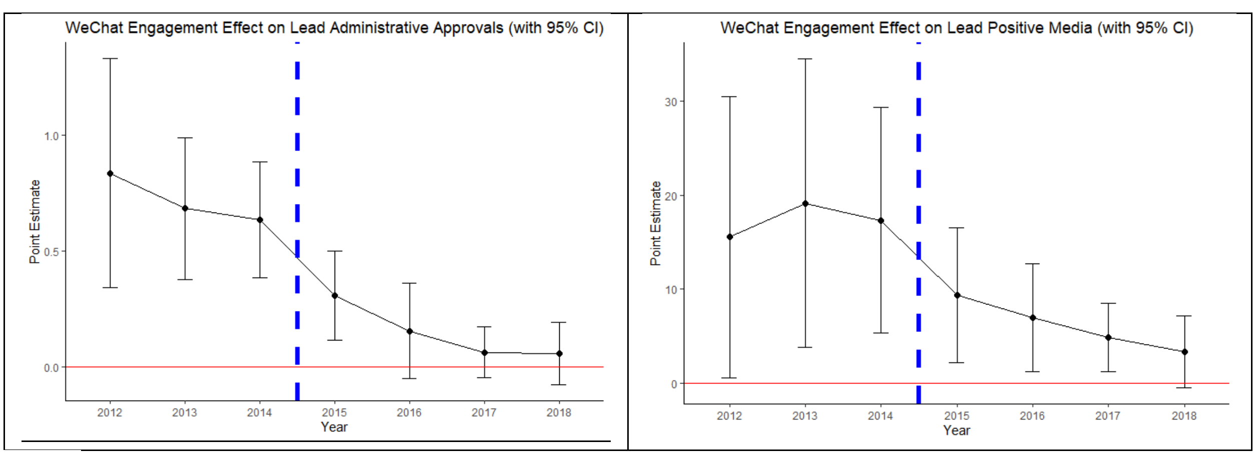


Figure 8: Event Study: Social Media Effect before and after MEIP



Notes: The left figure shows the point estimator (with 95% confidence interval) of regressing WeChat Engagement on lead number of administrative approvals using yearly data, with control variables, and the right figure shows the point estimator (with 95% confidence interval) of regressing WeChat Engagement on lead number of positive media stories using yearly data, with control variables.

Table 9: Summary Statistics

Variable Category	Variable Name	No. Obs.	Mean	Std. Dev.	Min	Max
Company Basic Information	Age	10,641	3.84	2.10	1	9
	Registration Capital	10,641	2.4e7	1.5e8	2.5e3	3.7e9
	Number of Branches	10,641	3.00	17.91	0	367
	Number of Subsidiaries	10,641	5.13	8.16	0	82
	Number of Employees	10,641	134.65	714.94	0	35935
Company Quality Information	Number of Trademarks	10,641	54.30	194.47	0	6931
	Number of Patents	10,641	23.40	128.37	0	4134
	Number of Certificates	10,641	1.44	9.56	0	553
	Number of Work Copyrights	10,641	11.10	557.84	0	47611
	Number of Software Copyrights	10,641	0.66	10.36	0	575
Company Institutional Information	Political Distance	10,641	0.95	0.17	0.07	1
	Number of Administrative Approvals	10,641	0.52	1.63	0	92
	Number of Administrative Punishments	10,641	0.02	0.59	0	48
	Number of Positive Media Stories	10,641	20.53	90.94	0	2,273
	Number of Negative Media Stories	10,641	4.32	47.02	0	2,586
Company Funding Information	Total Funding Rounds	10,641	1.32	1.39	0	17
	Total Funding Amount	10,641	4.0e7	2.0e8	0	5.0e9
	Funded by Tencent	10,641	0.04	0.29	0	1
Social Media	Has WeChat	10,641	0.56	0.50	0	1
	WeChat Engagement	10,641	0	1	-2.72	4.03
	Business Announcement	10,641	0	1	-0.28	30.5
	Marketing Promotion	10,641	0	1	-0.31	23.2
	Customer Relations	10,641	0	1	-0.24	33.3
	CSR	10,641	0	1	-0.31	29.9
Other Metrics	Industry Administrative Competitiveness	10,641	0.73	0.10	0	1

Table 10: Correlation between Main Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) Age	1.00															
(2) Registration Capital	0.02	1.00														
(3) Number of Branches	0.01	0.07	1.00													
(4) Number of Subsidiaries	0.05	0.21	0.20	1.00												
(5) Number of Employees	0.12	0.12	0.09	0.15	1.00											
(6) Number of Trademarks	0.20	0.13	0.07	0.26	0.50	1.00										
(7) Number of Patents	0.14	0.14	0.00	0.15	0.16	0.27	1.00									
(8) Number of Certificates	0.13	0.01	0.01	0.05	0.03	0.12	0.09	1.00								
(9) Number of Work Copyrights	0.06	0.01	0.00	0.09	0.05	0.16	0.07	0.02	1.00							
(10) Number of Software Copyrights	0.06	0.01	0.00	0.08	0.06	0.29	0.09	0.00	0.03	1.00						
(11) Political Distance	-0.44	-0.05	-0.04	-0.10	-0.11	-0.19	-0.08	-0.06	-0.04	-0.05	1.00					
(12) Number of Administrative Approvals	0.00	0.02	0.06	0.05	0.08	0.08	0.07	0.02	0.01	0.06	-0.04	1.00				
(13) Number of Administrative Punishments	0.02	0.03	0.19	0.00	0.04	0.01	0.02	0.00	0.00	0.00	-0.03	0.02	1.00			
(14) Number of Positive Media Stories	0.22	0.06	0.05	0.14	0.13	0.24	0.24	0.05	0.35	0.06	-0.20	0.09	0.08	1.00		
(15) Number of Negative Media Stories	0.21	0.06	0.07	0.12	0.13	0.22	0.24	0.04	0.28	0.06	-0.19	0.08	0.16	0.62	1.00	
(16) Total Funding Rounds	0.46	0.08	0.09	0.15	0.16	0.25	0.15	0.06	0.07	0.08	-0.68	0.04	0.05	0.28	0.28	1.00
(17) Total Funding	0.10	0.47	0.03	0.18	0.09	0.17	0.13	0.01	0.02	0.03	-0.13	0.03	0.02	0.13	0.13	0.20
(18) Funded by Tencent	0.09	0.08	0.00	0.13	0.08	0.20	0.17	0.01	0.00	0.05	-0.16	0.01	0.00	0.11	0.15	0.28
(19) Has WeChat	0.16	0.02	0.05	0.05	0.06	0.08	0.02	0.03	0.00	0.03	-0.13	0.07	0.02	0.09	0.09	0.16
(20) WeChat Engagement	0.21	0.02	0.03	0.03	0.05	0.06	0.04	0.04	0.01	0.03	-0.17	0.08	0.02	0.12	0.14	0.19
(21) Business Announcement	0.04	0.02	0.01	0.05	0.03	0.03	-0.01	0.00	0.02	0.03	-0.03	0.01	0.00	0.06	0.04	0.05
(22) Marketing Promotion	0.04	0.03	0.13	0.08	0.07	0.07	0.00	0.00	0.03	0.04	-0.04	0.06	0.02	0.07	0.09	0.08
(23) Customer Relations	0.02	0.02	0.02	0.06	0.04	0.05	-0.00	-0.00	0.03	0.04	-0.05	0.02	-0.01	0.05	0.05	0.06
(24) CSR	0.05	0.02	0.08	0.11	0.05	0.05	0.00	0.00	0.01	0.02	-0.05	0.05	0.00	0.07	0.06	0.07
(25) Industry Administrative Competitiveness	-0.30	0.00	-0.01	0.02	-0.07	-0.09	-0.07	-0.07	-0.02	-0.04	0.19	-0.20	-0.01	-0.15	-0.15	-0.18

Continue from above	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)
(17) Total Funding	1.00								
(18) Funded by Tencent	0.13	1.00							
(19) Has WeChat	0.06	0.04	1.00						
(20) WeChat Engagement (Standardized)	0.08	0.04	0.70	1.00					
(21) Business Announcement	0.05	0.01	0.25	0.49	1.00				
(22) Marketing Promotion	0.10	0.04	0.27	0.66	0.19	1.00			
(23) Customer Relations	0.07	0.02	0.23	0.51	0.39	0.18	1.00		
(24) CSR	0.08	0.00	0.25	0.39	0.56	0.39	0.31	1.00	
(25) Industry Administrative Competitiveness	-0.04	-0.02	-0.13	-0.17	0.00	0.00	-0.03	-0.01	1.00

Table 11: Social Media Effect on Formal Institutions (Administrative Approvals)

FE: Company ID	FE		FE + 2SLS	
Dependent Variable: Lead Administrative Approvals	(1)	(2)	(3)	(4)
Inverse Mills Ratio (Admin)	-1.343	-0.754	-3.486	1.168
	(4.713)	(4.715)	(3.407)	(3.802)
Has WeChat	0.085	-0.083	0.123	-0.660*
	(0.064)	(0.386)	(0.058)	(0.390)
WeChat Engagement	0.710***	0.941***	**0.570*****	2.506***
	(0.086)	(0.105)	**(0.081)**	(0.346)
Has WeChat * Political Distance		0.153		0.636
		(0.385)		(0.393)
WeChat Engagement * Political Distance		0.209***		**1.918*****
		(0.053)		**(0.345)**
Other Controls	Age, Number of Employee, Number of Trademark, Number of Patent, Number of Certificate, Number of Work Copyright, Number of Software Copyright, Total Funding Rounds, Total Funding Amounts, Funded by Tencent, Year Index			
Observations	7,127	7,127	7,127	7,127
R-squared	0.417	0.419		

Notes:

1. Dependent variable is the number of administrative approvals in the next year. WeChat engagement is standardized, where one-standard-deviation change is associated with about 200 articles, 5 million readings, 2000 comments and 30,000 likes per year.

2. Columns (1) and (2) report the OLS results with company fixed effect. Columns (3) and (4) report the 2SLS results with company fixed effect. The instrumental variables are average WeChat adoptions and WeChat Engagement of the start-ups in the same city/region while in different industry categories.

3. Robust standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 12: Social Media Effect on Informal Institutions (Positive Media Stories)

FE: Company ID	FE		FE + 2SLS	
Dependent Variable: Lead Positive Media Stories	(1)	(2)	(3)	(4)
Inverse Mills Ratio (Media)	580.38***	579.29***	580.40***	572.40***
	(29.25)	(29.25)	(29.26)	(29.71)
Has WeChat	3.852	22.46	3.857	14.04
	(2.648)	(15.91)	(2.653)	(16.58)
WeChat Engagement	18.37***	22.35***	**18.32*** **	47.22***
	(3.563)	(4.315)	**(3.979)**	(12.07)
Has WeChat * Political Distance		-19.89		-13.31
		(16.26)		(16.71)
WeChat Engagement * Political Distance		5.676***		**26.45*** *
		(2.205)		**(10.46)**
Other Controls	Age, Number of Employee, Number of Trademark, Number of Patent, Number of Certificate, Number of Work Copyright, Number of Software Copyright, Total Funding Rounds, Total Funding Amounts, Funded by Tencent, Year Index			
Observations	7,127	7,127	7,127	7,127
R-squared	0.460	0.461		

Notes:

1. Dependent variable is the number of positive media stories in the next year. WeChat engagement is standardized, where one-standard-deviation change is associated with about 200 articles, 5 million readings, 2000 comments and 30,000 likes per year.

2. Columns (1) and (2) report the OLS results with company fixed effect. Columns (3) and (4) report the 2SLS results with company fixed effect. The instrumental variables are average WeChat adoptions and WeChat Engagement of the start-ups in the same city/region while in different industry categories.

3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 13: Topic Posts Effect on Institutions

Dependent Variable:	Administrative Approvals					Positive Media				
Model	FE				FE+2SLS	FE				FE+2SLS
FE: Company ID	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Inverse Mills Ratio (Admin)	-4.612	-4.671	-4.613	-5.152	-5.897*					
	(3.425)	(3.423)	(3.424)	(3.410)	(3.438)					
Inverse Mills Ratio (Media)						466.69***	465.32***	465.32***	462.19***	455.36***
						(26.33)	(26.43)	(26.40)	(26.32)	(26.49)
Has WeChat	0.170***	0.170***	0.170***	0.124**	0.061	5.148**	5.132**	5.125**	3.802	2.060
	(0.058)	(0.058)	(0.058)	(0.058)	(0.060)	(2.588)	(2.588)	(2.588)	(2.602)	(2.669)
Business Announcement	-0.005					3.514*				
	(0.041)					(1.834)				
Marketing Promotion		-0.038					1.325			
		(0.032)					(1.428)			
Customer Relations			-0.013					5.021		
			(0.095)					(4.242)		
Corporate Social Responsibility				0.207***	**0.487*** **				5.877***	**13.53*** *
				(0.030)	**(0.063)**				(1.342)	**(2.810)**
Other Controls	Age, Number of Employee, Number of Trademark, Number of Patent, Number of Certificate, Number of Work Copyright, Number of Software Copyright, Total Funding Rounds, Total Funding Amounts, Year Index									
Observations	7,127	7,127	7,127	7,127	7,127	7,127	7,127	7,127	7,127	7.127
R-squared	0.389	0.389	0.389	0.394		0.427	0.426	0.427	0.428	

Notes: 1. Dependent variable for Columns (1) through (5) is the number of administrative approvals in the next year. Dependent variable for Columns (6) through (10) is the number of positive media stories in the next year. Business Announcement, Marketing Promotion, Customer Relations and Corporate Social Responsibility are standardized.

2. All columns add company fixed effects and year fixed effects. Column (5) and Column (10) are estimated using instrumental variable 2SLS.

3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 14: Social Media and Industry Competitiveness

FE: Company ID	FE				FE + 2SLS
Dependent Variable: Lead Administrative Approvals	(1)	(2)	(3)	(4)	(5)
Inverse Mills Ratio (Admin)	-3.838	-3.694	-3.665	-3.642	-3.453
	(3.408)	(3.406)	(3.393)	(3.394)	(3.552)
Admin Competitiveness	-6.358***	-17.15***	-12.08***	-14.06***	-2.962
	(2.302)	(4.027)	(2.419)	(4.039)	(4.814)
Has WeChat	0.122**	0.069	0.133**	0.122**	0.310***
	(0.058)	(0.060)	(0.058)	(0.060)	(0.074)
WeChat Engagement	0.635***	0.634***	0.535***	0.538***	0.194*
	(0.075)	(0.075)	(0.076)	(0.076)	(0.107)
Has WeChat * Admin Competitiveness		15.75***		3.167	-42.02***
		(4.825)		(5.167)	(10.83)
WeChat Engagement * Admin Competitiveness			18.68***	18.07***	**82.94*** **
			(2.529)	(2.718)	**(13.76)**
Other Controls	Age, Number of Employees, Number of Trademarks, Number of Patents, Number of Certificates, Number of Work Copyrights, Number of Software Copyrights, Total Funding Rounds, Total Funding Amounts, Year Index				
Observations	7,127	7,127	7,127	7,127	7,127
R-squared	0.397	0.398	0.402	0.402	

Notes:

1. Dependent variable is the number of administrative approvals in the next year. WeChat engagement is standardized, where one-standard-deviation change is associated with about 200 articles, 5 million readings, 2000 comments and 30,000 likes per year.

2. Columns (1) to (4) report the OLS results with company fixed effect. Column (5) reports the 2SLS results with company fixed effect. The instrumental variables are average WeChat adoptions and WeChat Engagement of the start-ups in the same city/region while in different industry categories.

3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 15: Social Media and Institutional Shock

FE: Company ID	FE				FE + 2SLS
Dependent Variable: Lead Administrative Approvals	(1)	(2)	(3)	(4)	(5)
Inverse Mills Ratio (Admin)	-3.545	-3.449	-3.242	-3.216	0.369
	(3.407)	(3.407)	(3.406)	(3.406)	(4.166)
Institutional Shock	-0.248*	-0.496**	-0.169	-0.284	2.966**
	(0.136)	(0.225)	(0.138)	(0.238)	(1.405)
Has WeChat	0.119**	0.110*	0.120**	0.116**	0.199***
	(0.058)	(0.058)	(0.058)	(0.058)	(0.075)
WeChat Engagement	0.631***	0.632***	0.615***	0.616***	0.371***
	(0.075)	(0.075)	(0.075)	(0.057)	(0.134)
Has WeChat * Institutional Shock		0.386		0.173	-3.098**
		(0.281)		(0.291)	(1.427)
WeChat Engagement * Institutional Shock			1.186***	1.121***	**18.33*** **
			(0.394)	(0.409)	**(7.313)**
Other Controls	Age, Number of Employee, Number of Trademark, Number of Patent, Number of Certificate, Number of Work Copyright, Number of Software Copyright, Total Funding Rounds, Total Funding Amounts, Year Index				
Observations	7,127	7,127	7,127	7,127	7,127
R-squared	0.396	0.397	0.397	0.397	

Notes:

1. Dependent variable is the number of administrative approvals in the next year. WeChat engagement is standardized, where one-standard-deviation change is associated with about 200 articles, 5 million readings, 2000 comments and 30,000 likes per year. Institutional shock refers to the 2 years since the company has received an administrative punishment or a negative media report.

2. Columns (1) to (4) report the OLS results with company fixed effect. Column (5) reports the 2SLS results with company fixed effect. The instrumental variables are average WeChat adoptions and WeChat Engagement of the start-ups in the same city/region while in different industry categories.

3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

CHAPTER 4: ARTIFICIAL INTELLIGENCE, LEAN START-UP METHOD AND PRODUCT INNOVATIONS

1. Introduction and Motivation

New product development is defined as a set of activities that start with responding to a market opportunity and end with a differentiated product or service to meet the market need (Krishnan and Loch 2005). These activities are crucial for start-ups to subsequently capture the market gains that a successful product may generate. But innovation is also risky because a failed product can be costly; and particularly so for a cash-strapped start-up, because failure could lead to bankruptcy (Bhaskaran et al. 2021). To aid with product development, many start-ups have employed artificial intelligence (AI), hoping to create more innovative products that could prove to be popular with consumers (Berente et al. 2021, Verganti et al. 2020). However, research shows that most companies have not benefited from AI despite the surge in its investment. Simply adopting AI is likely to fail because AI is multifaceted and multi-purposed; as such, successful use of AI requires organizational support that is tailored to the particulars of the AI being used (Varian 2018, Brynjolfsson et al. 2021, Dixon et al. 2021). Some firms use AI to analyze external markets to explore new business opportunities, while others use AI to improve internal production processes. Furthermore, the organizational processes required to support different aspects of AI capabilities in product development may also be different.

To understand how different AI capabilities affect product development, we first distinguish specific aspects of AI capabilities and then we examine the organizational practices that complement them. First, we distinguish between internally and externally focused AI based on whether the company uses AI to process internal or external information. The different types of abilities needed to leverage internal vs. external information are well-documented in information systems and marketing literature. (Tambe et al. 2012, Kohli et al. 1993, Mendelson and Pillai 1999). We similarly distinguish internally and externally focused AI by the types of information that they process, and the context knowledge required for that processing. Specifically, internally focused AI refer to the ability to make predictions by processing internal process information. These types of capabilities are important for fine-tuning production and engineering activities and automating aspects of the product design process. For example, Airbnb has been using AI to classify room pictures (e.g., bedroom, kitchen, etc.) to provide a better search experience for guests. By removing the repetitive work in labeling the pictures manually, Airbnb was able to launch a new premium version, Airbnb Plus that significantly improved the consumer search experience.[22] Externally focused AI, on the other hand, is characterized by using AI to process external market information for forecasting customer demand and market trends. For example, the Chinese beverage start-up Genki Forest originally made tea beverages. The company used AI to analyze consumer data on major social platforms and found that young people in China were highly concerned about sugar, diabetes, and obesity. This led to Genki Forest's later creation of zero-sugar sparkling water. Because they used AI to process market data, quickly identify a new market niche, and launch a product for the niche, the product was widely successful.[23]

We then examine the organizational practices that best support the use of internally and externally focused AI. Specifically, we focus on the choice of using the lean start-up method (LSM)[24], a commonly adopted organizational practice in both early-stage and mature firms (Koning et al. 2022, Yoo et al. 2021) to develop new products. Inspired by the concept of "lean manufacturing" from the Toyota Production System, LSM advocates the use of external market information to validate a new product and to continuously collect

[22] See https://medium.com/airbnb-engineering/widetext-a-multimodal-deep-learning-framework-31ce2565880c.

[23] See https://techcrunch.com/2021/07/25/data-driven-iteration-helped-chinas-genki-forest-become-a-6b-beverage-giant-in-5-years/.

[24] Despite having the word "start-up" in the middle of its name, our use of the term—and our focus on studying start-ups—does not imply our findings only apply to start-ups. The application of our findings to companies past the start-up phase is a matter for future research.

process data to further improve the product (Ries 2011, Blank 2013). It favors market feedback over a rigorous business plan or entrepreneurial insights. For example, start-ups can build a prototype to learn about consumer preferences and modify and iteratively improve upon the product design according to consumer feedback. If the product fails, it can pivot to something else. Although LSM has the word "start-up" in the middle of its name, these principles of LSM are widely embraced by both start-ups and mature firms.

AI capabilities are similar in that they can also help with processing external market information and improving internal processes. Thus, it is unclear whether AI can substitute or complement LSM in product innovations. Exploring their interactions can help explain why some start-ups successfully use AI to innovate while others fail.

Therefore, our goal is threefold. First, to understand differences in AI, we examine internal and externally focused AI and how they differ in helping firms develop new products. Second, to understand AI's impact on inventing, we explore how the two types of AI substitute or complement different aspects of LSM in product innovations. Finally, by looking at these collectively, we seek to explain why some firms reap substantial returns on their AI investments while others do not.

Using a longitudinal dataset with comprehensive information for about 2,000 start-ups in China from 2011 to 2020, we find that start-ups with higher AI capabilities can develop more products every year, including both novel products (defined as new products for the company) and incremental products (defined as new versions of existing products). We then show that start-ups' AI capabilities and the general use of LSM are complementary in developing products.

To understand the complementarities better, we separate general AI capabilities into internally and externally focused AI. We also refine general LSM into validated learning (e.g., building prototypes) and iterative experimentation (e.g., simulations and A/B testing), the two important aspects of LSM. We find that validated learning complements externally focused AI in developing novel products, but not in developing incremental ones. This is because finding a market niche with real-time market demand is key for developing novel products. AI can aid in this task by expanding the search scope of business opportunities, and validated learning can provide market feedback through prototyping based on these opportunities. Together, AI and LSM are complements in expanding business scope and thereby in alleviating market uncertainty.

On the other hand, we find iterative experimentation complements internally focused AI in developing incremental products, but not in developing novel ones. To help start-ups update existing products, AI can improve the efficiency of existing development procedures through automation and optimization, and iterative experimentation in LSM can help to improve the development process by providing quality feedback through continuous experimentation. This creates another complementary system for improving and optimizing processes. We find that the joint adoption of internally focused AI and iterative experimentation is associated with faster product iterations and higher product quality (as measured by third-party certifications described later), and the complementarity becomes stronger as the start-up faces stronger market competition.

To ensure our results do not suffer from obvious reverse causality and selection biases, we create two instrumental variables that are exogenous shocks to AI adoption. The first is the number of local Chinese government policies with tax benefits and cash subsidies that encourage firms to adopt AI. The other variable is the number of news reports announcing that a similar firm has adopted AI. We also create two instruments for the adoption of LSM, which are the average adoption of validated learning and iterative experimentation for start-ups in a different industry and geographic location but share a board member with the focal start-up. Our results are robust after applying these identification methods. To summarize, our results provide new evidence for the effect of different types of AI capabilities on start-up product innovations, and secondly that different AI capabilities complement different aspects of LSM. The results also provide an explanation for why investment in AI can be a losing proposition when these nuances are not taken into account.

2. Literature and Theory

In this section we first review the literature about start-up product development and describe how start-ups use LSM for product innovations. Then we theorize on the complementarities between AI capabilities and general LSM. Lastly, we show how different AI foci complement two different aspects of LSM in product innovation.

2.1 Start-up Product Innovation and LSM

Recent decades witnessed the remarkable impact of start-ups on product innovation that disrupted many industries (Tambe 2014, Liu et al. 2016, Ransbotham et al. 2019, Hellmann and Thiele 2015, Burtch et al. 2016, Mei et al. 2022 Gompers and Lerner 2001, Brynjolfsson and McAfee 2014). Even the disrupters are being disrupted as fierce competition rages among start-ups in part due to rapid changes in the technology landscape. To adapt to the fast-changing technology and market dynamics, many start-ups and others adopt LSM for product innovation (Blank 2005, Ries 2011, Blank 2013, Rigby et al. 2018). In contrast to strictly adhering to a business plan or an entrepreneur's insights, LSM advocates that entrepreneurs should create products by first validating an idea in the marketplace and then iteratively experimenting to continuously improve the product design. LSM has gained significant acceptance across a broad range of entrepreneurs, and it has also been shown to improve product innovations in various settings, including both software and hardware companies (Sarasvathy 2001, Read et al. 2011).

Past literature has documented two major applications of LSM: validated learning (creating prototypes and minimum viable products) and iterative experimentation (rigorous simulations or A/B testing) (Yoo et al. 2021, Harms and Schwery 2020, Calantone et al. 2002, Koning et al. 2022, Cui and Wu 2017). In validated learning, start-ups build and launch prototypes to learn about consumer preferences and then modify the product based on the feedback received from the market. This can substantially alleviate market uncertainties when creating a novel product for a new market (Yoo et al. 2021, Camuffo et al. 2020). In iterative experimentation, start-ups continuously experiment with features (Kohavi et al. 2020) to improve existing products (Koning et al. 2022, Thomke 2020). Together validated learning and iterative experimentation in LSM help firms create new products and to improve upon existing ones.

However, despite its many benefits, LSM is also associated with an increase in marketing and operating costs which can deter start-ups from adopting it. To effectively use LSM, companies need to continuously monitor market changes and adjust internal operations (Rigby et al. 2018, Sharma et al. 2018). Iterative experimentations such as A/B testing often require a large sample of users, which can be too costly for early-stage companies to obtain or simply unavailable. Furthermore, using LSM alone tends to favor creating incremental innovations over novel ones (Felin et al. 2020, Thiel and Masters 2014, Mollick 2019) because frequent responses to market can improve products to capture short-term demands, but such focus is often pursued at the expense of long-term radical innovations (Bhamu and Sangwan 2014, Felin and Foss 2011, Anand et al. 2016).[25] To address the drawbacks of LSM, start-ups are actively looking for new technologies such as AI to boost product innovation of both types, novel and incremental.

2.2 AI and Product Innovation

In this study we follow the recent literature that defines AI as an advanced computing power, mainly driven by machine-learning algorithms (Agrawal et al. 2018, Taddy 2018, Berente et al. 2021, Brynjolfsson et al.

---

[25] Many entrepreneurs have criticized the lean start-up method because of its over-emphasis on immediate validation and lack of market insights. For example, PayPal co-founder Peter Thiel quipped, “Why should you expect your own business to succeed without a plan? Darwinism may be a fine theory in other contexts, but in start-ups, intelligent design works the best.”

2021). As with prior generations of information technology, AI is multifaceted and multi-purposed (Bresnahan and Trajtenberg 1995, Varian 2018, Berente et al. 2021, Brynjolfsson et al. 2021). To understand how different aspects of AI affect product innovation in start-ups, we divide the general AI capability into two categories based on the context knowledge they require and the types of information they process. Specifically, the externally focused AI refers to the ability to analyze external market information, while the internally focused AI refers to the ability to optimize internal development processes. The distinction between processing external and internal information in organizations is well-documented (Tambe et al. 2012, Kohli et al. 1993, Mendelson and Pillai 1999). Start-ups have similarly used different types of AI foci to address challenges in different aspects of product innovation. Developing new products and refining existing ones require different types of capabilities. For developing novel products in a new market niche, the major challenge lies in the uncertain customer demand—will customers want it (Lynn et al. 1996, Hoeffler 2003, Kwark et al. 2018)—while for updating an existing product in a relatively known market, improving internal development processes is key to creating high quality products and to capturing market share (Wilson and Norton 1989, Netessine and Taylor 2007, Hashai and Markovich 2017, Gans et al. 2002).

Using AI to analyze a vast amount of information quickly is key to capturing market demand (Berente et al. 2021, Lou and Wu 2021, Wu et al. 2019). Digitalization of market information has enabled AI to predict market dynamics and identify business opportunities (Wu and Brynjolfsson 2014, Aral et al. 2013, Bigsby et al. 2017, Qiu and Kumar 2017, Cui et al. 2018). Having the AI capability to process external market information is key to capturing these opportunities. For example, Centricity Inc., a start-up that creates customized service contracts in retail and manufacturing, uses AI to predict consumer demand for specific products in particular communities. By analyzing billions of data points in real time that include the information about individual location, the websites they are visiting and their online search results, Centricity Inc. can predict, even months in advance, certain product characteristics that will be popular in certain demographic and geographic communities. By providing critical insights to product development,[26] externally focused AI can alleviate uncertainties about market demand.

On the other hand, internally focused AI is also key to automating and optimizing internal product development processes. They can remove some of the scaling limitations that involve repetitive and labor-intensive activities in product development (Verganti et al. 2020, Iansiti and Lakhani 2020, Dixon et al. 2021). For example, the start-up Turing Labs uses AI to test and filter new formulations for existing products. By simulating different ingredient combinations for new products, Turing Labs can identify a smaller subset of the ingredients for prototyping. Using AI to select the appropriate features greatly enhances Turing Labs' ability to explore a wide range of possible input features and accordingly increases its efficiency in the production process. According to the company, "the process was 90% manual but now is 90% digital".[27] This is especially helpful for optimizing the existing development procedures and improving existing products. Using AI to optimize product development processes and remove the traditional scaling limitations can help start-ups expedite product iterations, and thereby creating higher-quality versions of existing products faster.

Thus, AI can accelerate the ideation and the creation of a novel product that can capture new market opportunities. It can also facilitate internal development processes for creating incremental products. Together, we hypothesize that using AI is beneficial for start-ups to develop new products:

H1: Using AI can raise the quantity of a start-up's product innovations.

[26] See https://www.wsj.com/articles/ai-start-up-sees-opportunity-forecasting-pandemic-era-consumer-demand-11610496808.

[27] See https://techcrunch.com/2022/01/19/turing-labs-which-uses-ai-to-help-formulate-cpg-products-raises-16-5m/

2.3 Complementarity

Despite the potential benefits of AI on innovation and productivity, many organizations that invested in AI have not benefited from it (Lou and Wu 2021, Berente et al 2021). An important mediating factor is having the organizational practices and conditions that are complementary to AI and thereby enable firms to take advantage of what AI can offer. When these conditions are not met, firms can even experience negative productivity caused by the technology investments (Aral et al 2012).

Furthermore, the literature has shown that AI is less helpful when available data are sparse and unreliable (Lou and Wu 2021, Wu et al 2019). Paradoxically, the greater the novelty of an innovation, the less likely it is that reliable data will be available (Thomke 2020). Using AI to predict market demand and develop a radically new product may not be effective because there is no consumer data about these products and consumers are poor at conceptualizing how a new product can be used (Cooper et al. 2002). Even in developing incremental products where start-ups face less market uncertainty, the characteristics that constitute a higher quality product remain ambiguous, as production environments and customer needs vary widely and change over time (Allon et al. 2021). The complexity makes it hard to meet the needs of long-term product improvement, and companies not only need to calibrate and improve their AI algorithms frequently, but also consider implementing complementary organizational processes to meet the needs.

LSM is a popular tool that start-ups and others use to facilitate product innovation. It differs from AI in navigating the market landscape because AI uses existing data to find optimal solutions while LSM generates new data through prototyping and experimentations. The data from LSM are key to improving judgment about market conditions and product quality (Kerr et al. 2014, Ewens et al. 2018, Camuffo et al. 2020), and AI is in good position to leverage the data from LSM. This is because LSM provides the key parameters AI needs to optimize decision making: information about the objective functions as well as the key input features needed to train and use AI algorithms. By gathering market intelligence, LSM can be helpful to decision makers to determine the right objective functions. A mis-specified objective for AI can even lead to detrimental consequences. For instance, the objective to maximize short-term revenue can come at the expense of long-term profitability. By gathering many internal process parameters, LSM can also inform on the input features needed to use AI algorithms (Agrawal et al. 2018). For example, Wu and Brynjolfsson (2014) show that simply adding Google Trend as an input into a prediction model for housing price can significantly improve prediction accuracy.

Different aspects of LSM also support different aspects of product innovation. For developing novel products, validated learning can complement externally focused AI by providing feedback data on prototypes and validate the market niche (or the objective function) that the firm chooses to enter. Prototyping can help determine whether there is real market demand or help to narrow down the design spectrum. For example, the Chinese beverage start-up Genki Forest, mentioned earlier, launched several low-sugar test products after using AI to analyze the market. The company initially targeted zero-sugar tea beverages at consumers, but the market feedback data created through testing products rejected this approach. Through trial and error, the company created the best-selling zero-sugar sparkling water. In an uncertain and fast-moving market, launching test products into the market and using AI to analyze the results can provide companies with first-hand feedback data about demand distributions. Based on the information, firms can either validate their objectives or pivot to a different product. This will help reduce the market uncertainty faced by firms when developing novel products.

For developing incremental products where there is already a clear objective function for product innovation, iterative experimentation can complement AI by experimenting with the universe of input features in product development and providing feedback data to further improve the AI algorithms. For instance, in order to provide the best guest experience through AI-enabled room picture classification, Airbnb also conducted A/B testing to improve their AI algorithm. Since it is inefficient to use AI to process every pixel of the high-fidelity room pictures, the company used A/B testing to specify the best subset of features in the pictures so that they can simplify the inputs for the AI algorithm. The company iteratively experimented with the effect of different image feature choices for the room picture classification without changing other features, and then selected the optimized archetype for each image feature to train the AI algorithm. Consequently, the AI algorithm could achieve the optimal customer satisfaction level by predicting consumer preferences fast and accurately. These experimentation approaches, such as A/B testing, can help companies better understand which features of their product design are best for customer retention and profitability, thereby improving the efficiency and effectiveness of AI algorithms. Thus, start-ups can continuously iterate and improve their existing products to outperform in the market competition.

Moreover, AI can also improve the efficiency of LSM. AI insights derived from market data can help broaden the scope of market exploration and provide guidance on where to launch test products. This effect could be especially helpful in product innovations when the search space is too large for LSM to navigate alone. AI can also improve the efficiency of LSM by optimizing the process of experimentation. For example, AI algorithms such as those that are based on the multi-armed bandit theory have been widely adopted in experimentation contexts, where companies can exploit the experimental results during exploration. The combination of AI and LSM can create a closed-loop system in market search, validation and iteration, which can turbocharge the scaling process of a start-up in product innovations (DeSantola and Gulati 2017).

Thus, LSM and AI should be complements in exploring external market opportunities and creating novel products, as well as facilitating internal product iterations. Together, we hypothesize that LSM complements AI capability in start-up product innovation:

H2: There are positive interaction effects between AI capability and LSM in a start-up's product innovations.

Specifically, validated learning should complement externally focused AI in creating novel products, and iterative experimentation should complement internally focused AI in creating incremental products. We show the theoretical model of this paper in Figure 9.

H3a: Externally focused AI and validated learning are complements for creating novel products in start-ups.

H3b: Internally focused AI and iterative experimentation are complements for creating incremental products in start-ups.

3. Data and Measures

3.1 Company Information Data

We acquired start-up information from Crunchbase, one of the most widely used databases for global entrepreneurial research (Cumming et al. 2016, Dalle et al. 2017). We examine start-ups that are founded after 2011. In total, there are about 1,800 Chinese start-ups, including about 600 software companies and 1,200 companies that develop physical products.[28] For each of these start-ups, we obtain the founding year, founder demographics, and specific funding outcomes—the number of successful rounds, the amount of financing received, and the financing sources in all funding series (Seed, Series A, Series B, etc.). The average number of funding rounds is about 1.5. We match these start-ups with the Chinese Enterprise Database. The database curates comprehensive information about all the registered companies in China, from public data sources and proprietary governmental databases, including location, industry category, subsidiary and branch information, job postings on major hiring websites, yearly personnel information, social media accounts, founder and board member information (gender, education level and resume of past

[28] Data acquired in 2020.

experiences), the detailed administrative information from all governmental departments, the detailed certificate, trademark, patent and copyright information, and also all the online news reports that relate to the focal company (including report title and posting time) from over 100 major news agencies (including both private and state-owned agencies) in China, and emotional orientation of each news report (positive, negative or no emotion). This data is validated through other similar data sources such as *Tianyancha*.

Using the company's WeChat public accounts shown in the Chinese Enterprise Database, we acquired all the historical articles published in each account, including the title, text, and time of the publication for each article, and whether the article is an original article published by the company or a reposted article. WeChat is the most widely used social media platform in China and nearly all firms have a presence on it. Firms use WeChat heavily to make public announcements and broadcast product information, which can help us infer the company's adoption level of AI capability and LSM.

3.2 Variables

Dependent Variables

For software companies, we measure the number of new products using the **Number of Software Copyrights** that a company applies for each year. For physical product companies, we measure the number of new products using the **Number of Product Trademarks** that a company applies for each year. These numbers are useful because they are common practices for companies in China to protect intellectual properties before launching new products. An example of software copyrights and product trademarks is shown in Figure 10.

To identify whether software is novel, we collect the registered software version from the copyright certificate. To register a software copyright, the assignee needs to declare the version number of its software, such as "V1.0" or "V4.3". We treat all "V1.0" software as **novel** of the firm and subsequent numbers as **incremental** versions of an existing product. To identify whether a physical product is novel, we collect the registered product sub-classes from the trademark certificates. To register a trademark, the assignee needs to declare the product sub-classes, and most trademarks have over 10 different sub-classes. We treat a trademark as **novel** if it contains at least one new sub-class in comparison to the company's existing trademarks as it suggests the company is expanding to a new market niche, and the rest are treated as **incremental** products.

We also calculate the **Number of High-Quality Products** that each company creates in each year based on the number of product certificates, except the China Compulsory Certifications (3C). Except for 3C, product certificates are issued as certifications of quality by third-party agencies such as the International Organization for Standardization (ISO).

One concern about the dependent variables (software copyright and product trademark) is that they are not representative of a start-up's overall productivity. To measure the overall performance of start-ups, traditional metrics such as revenue and profit are not publicly available and not verified by public accountants. To alleviate this concern, we also use the **log Amount of Funding** as the dependent variable to further verify the robustness of the results as funding is one of the most important financial indicators of start-up performance (Bardhan et al. 2013). We also use the **Search Index** of each start-up in each year on the major search engine in China, as prior literature has shown that online traffic data can be used to predict and verify firm revenue (Froot et al. 2017, Huang 2018, Zhu 2019). The results are listed in the Appendix C1 and are consistent with the main results.

AI Capabilities

We measure a start-up's **AI Capability** by combining AI-related job posts, original social media posts and news reports about a firm into a single measure[29] and taking the first principal component. AI job posts have been used in previous studies to measure a company's AI capability (Lou and Wu 2021, Babina et al. 2020). We complement this measure by using social media posts and news articles to measure a firm's overall AI capability. We identify a job post, social media post, and news report as AI-related if it contains any of the AI keywords (Table 16) in the summary texts. In Figure 11, we plot the yearly average cumulative number of AI-related items under different measures since a start-up is founded. Although different measures start and grow in various patterns, the overall trends are similar, which suggests that using principal components is a valid choice. We then separately estimate each of the three AI metrics into our main regression and find that we cannot reject the hypothesis that all three measures have the same coefficients.

We also differentiate the start-ups' internally versus externally focused AI based on the company's usage of AI in processing internal information about production and engineering activities, or external information about market and customer demand. We have manually labeled 1,000 randomly selected AI-related job postings, news articles and social media articles as internally- or externally focused, and then use a bag-of-word method to identify the internal and external keywords with the greatest coefficients in predicting whether a text is internally or externally AI-oriented (Table 1). We identify a job post, social media post or news report as externally AI-related if it contains both AI and external keywords within the same post, or as internally AI-related if it contains both AI and internal keywords.[30] We take the first principal component to measure a company's **Externally focused AI** and **Internally focused AI**. We also use neural networks

[29] Cumulative quantity with a yearly depreciation rate of 15% following literature convention (Lou and Wu 2021). The same is true for the LSM Level below.

[30] An example of an externally focused AI job posting is: "…Proficient in using machine learning to analyze market demand…". An example of an internally focused AI job posting is: "…The algorithm engineer is responsible for using machine learning to optimize internal R&D processes, for example, image de-noising…".

to predict whether a text is internally AI-oriented or externally AI-oriented. Our best prediction model has an accuracy of 97% in testing data, and the results are consistent with the method of using keywords. For simplicity and interpretability, we use the keyword method.

Since the dependent variable is product development, there are concerns that start-ups may invest in AI capabilities in other areas, such as human resource management (e.g., AI-powered recruiting), and there are thereby errors in our AI measurements. However, more than 95% of AI-related job postings in our sample directly mention keywords such as product development and research and development (R&D), suggesting that start-ups' AI investments are mostly devoted to product innovations in the sample we observe.

Lean Start-up Method (LSM)

To operationalize the extent to which start-ups use LSM, we count the number of job posts, social media articles and news reports that contain keywords related to the lean start-up method (Table 1). We follow the previous literature that operationalizes the two major applications of LSM (Ries 2011, Bortolini et al. 2018, Harms and Schwery 2020, Koning et al. 2022) by looking for keywords that represent (1) Validated Learning (e.g., trial product) and (2) Iterative Experimentation. Start-ups that hire employees with expertise in experimentation or announce trial product launches are more likely to have adopted LSM than other start-ups. We then take the first principal component of the factors derived from the three data sources as the firm's **LSM Level** in each year. Similar to the AI capability, we found that we cannot reject the hypothesis that all three measures have the same coefficients. For simplicity, we combine them into a single measure. We also separately measure the extent to which companies adopt each of the two major applications of LSM, **Validated Learning** and **Iterative Experimentation,** using the keywords. To further ensure that we are measuring the underlying construct, we capture whether the start-up has received funding from government or state-owned funds, which can be a proxy for LSM in our context. This is because start-ups receiving support from public institutions have a greater propensity to produce and follow business plans (Honig and Karlsson 2004), and hence are less likely to adopt the principles of LSM. We also examine the concentration level of investor voting rights as dispersed voting rights are linked to a being slower at changing business directions, or lower LSM adoption (Brinckmann et al. 2010). Higher values on these measures represent that the company is more likely to have adopted LSM. The results are listed in Appendix C2 and are consistent with the LSM measurement.

However, firms that do not have any LSM-related job posts, social media articles or news reports can still be using LSM (Type II error). On the other hand, it is less likely that a firm that rates high on our LSM metric will know little about LSM methods (Type I error). Regardless, on the extent that our construct mismeasures the true underlying level of LSM of some firms, we are more likely to underestimate the effect than overestimate it, hence the measurement error is likely to bias the estimates on our LSM measures in the panel data models downward (Griliches and Hausman 1986, Tambe et al. 2012). The same reasoning applies to the AI measures as well. For example, some start-ups may outsource part of their AI capabilities, creating Type II measurement errors. We can further alleviate some of the measurement error concerns using instrumental variables.

Other Variables

For each start-up in each year, we construct firm-level control variables such as **Age**, **Number of Subsidiaries**, **Number of Branches**, **Total Number of Job Posts** and **Total Number of Employees**, company quality information such as **Total Number of Administrative Approvals**, **Total Number of Patents**, **Total Number of Copyrights** and **Total Number of Certificates**, company funding information such as **Total Funding Rounds**, and **Historical Funding Amounts**. By taking the average number of days between two versions of the same product, we calculate **Average Product Iteration Days** for each company in each year as a proxy for product iteration speed. We also calculate the **Market Uncertainty** for each company in each year using the failure rate of start-ups within its industry (Hyytinen et al. 2015). We calculate **Market Competition** for each company in each year as the total number of similar products in the market, where we define similar products as the trademarks owned by all companies in our sample with the same subclasses as the trademarks owned by the focal company. The greater the number of similar products in the market, the higher the competition is that the company faces. Our unbalanced panel dataset includes 10,641 observations for about 1,800 companies from 2011 to 2020. Table 17 shows summary statistics and Table 18 shows correlations between the variables.

4. Empirical Methodology

4.1 Base Model

We explore the effect of AI capability on the number of product innovations produced in a start-up in a typical year. We first fit the following OLS model with company fixed effect (γ_i) and time dummies (τ_t) for all companies after applying various control variables:

$$\{Number\_of\_Products_{i,t+1}, Number\_of\_Novel\_Products_{i,t+1}, Number\_of\_Incremenal\_Products_{i,t+1}\}$$
$$= \beta_0 + \boldsymbol{\beta_1 AI\_capability_{i,t}} + controls + \tau_t + \gamma_i + \varepsilon_{i,t}$$

A significant benefit of a firm fixed-effect model is that all the time-invariant characteristics of the start-ups, such as industry categories and the backgrounds of founders, are controlled in the regression. We lead the dependent variable by one year to avoid potential reverse causality and simultaneity issues. Control variables include firm age, number of branches, number of subsidiaries, number of employees, total trademarks, total patents, total certificates, total copyrights, total funding rounds and amount, and total job posts. This comprehensive set of information about start-ups can help us control for the unobserved effects of “firm quality” that can potentially cause omitted variable biases.

4.2 Complementarity Test

In order to test the complementarities between company AI capability and lean practices, we follow the previous literature and conduct two statistical tests: the demand equation that examines the correlations in inputs, and the performance test that examines the performance differences when complementary inputs are used in combination and independently (Arora and Gambardella 1990, Arora 1996, Athey and Stern 1998, Aral and Weill 2007, Brynjolfsson and Milgrom 2013, Aral et al. 2012, Tambe 2014, Wu et al. 2020). Specifically, we fit the following OLS model with company fixed effect (γ_i) and time dummies (τ_t) to test whether companies that have adopted LSM are more likely to invest in AI capability:

$$AI\_capability_{i,t} = \beta_0 + \boldsymbol{\beta}_1 \boldsymbol{LSM_{i,t}} + controls + \tau_t + \gamma_i + \varepsilon_{i,t}$$

Then we fit the following OLS model with company fixed effect (γ_i) and time dummies (τ_t) to test whether AI capability and LSM form a system of complements that provides additional performance improvements in product development when used together:

$$\{Number\_of\_Products_{i,t+1}, Number\_of\_Novel\_Products_{i,t+1}, Number\_of\_Incremental\_Products_{i,t+1}\}$$
$$= \beta_0 + \boldsymbol{\beta_1 AI\_capability_{i,t}} + \boldsymbol{\beta}_2 \boldsymbol{LSM_{i,t}} + \boldsymbol{\beta_3 AI\_capability_{i,t}} \times \boldsymbol{LSM_{i,t}} + controls + \tau_t + \gamma_i + \varepsilon_{i,t}$$

The complementarities are supported by a positive $\boldsymbol{\beta_3}$ coefficient. One concern about our measurement is that the company's AI capability might be spuriously correlated with its LSM level because they are computed using overlapping sources of information (job posts, social media articles and news reports). However, the performance test can help rule out such concerns if the estimated coefficient of interaction term is significantly positive, because there would hardly be any performance differentiations if the correlations are spuriously high and there are no true complementarities (Aral et al. 2012).

4.3 Identification Strategy

Complementarity approaches are naturally robust to some types of endogeneity and reverse causality problems because they are about matching two or more investment decisions rather than about the effectiveness of the decisions themselves. Any biases that affect the complementarity term must be present only at the confluence of both factors, and not when factors are present individually (Tambe et al. 2012). In addition, the two tests for complementarities are subject to different sources of bias, so consistent results provide greater confidence that the results are not primarily due to endogenous issues (Wu et al. 2020).

Nonetheless, the decision to adopt AI and LSM can be endogenous due to unobservable characteristics such as start-up quality. Such issues are partially addressed using a firm fixed effect model and the detailed company quality information in each year as control variables. To avoid potential biases caused by different timings of AI adoption in our panel data (Goodman-Bacon 2021), we also verify our fixed effect estimators using an improved doubly-robust difference-in-difference estimator (Callaway and Sant'Anna 2021, Sant'Anna and Zhao 2020). The results are listed in Appendix C4 and are consistent with our panel data of regression results.

In addition, we use instrumental variables to address other forms of endogeneity. We use two sets of instruments for AI capabilities. The first set of instrumental variables includes all the AI policies issued by provincial-level governments in China from 2011 to 2020 collected from the official government websites. Over the past decade, many local governments in China have issued incentives (e.g., cash subsidies and tax abatements[31]) to encourage local businesses adopt AI technologies. Local Chinese government incentives are often unexpected and can thus serve as an exogenous source of variation to study company performance.

[31] An example of these policies: "…R&D expenses in core areas such as artificial intelligence can be deducted at a rate of 175%…".

Such preferential policies are critical incentives for start-ups to adopt new technologies because the tax benefits they enjoy after adopting AI compensate, at least partially for the investment required in AI. This could translate to an exogenous shock that reduces the cost of AI adoption. Thus, we create an instrumental variable for a firm's AI investment using the number of annual AI policies in the province where the company is located. We find a total of 3,231 policies that contain AI-related keywords from official government websites in China. Figure 12 shows the provincial distribution of the number of AI incentives in China from 2011 to 2020.

The geographic distribution of the sample start-ups is as follows: 25% are in Beijing, 15% in Shanghai, 15% in Guangdong, and the rest are distributed almost evenly across China's remaining regions. The distribution has sufficient variations to provide the necessary power for the first-stage analysis of 2SLS.

The second set of instrumental variables is the number of news reports that a company in the same industry as the focal start-up has adopted AI. The news reports of other companies are exogenous shocks to the focal start-up; news of a competitor adopting a new technology can spur the company to imitate (Gatignon and Robertson 1989, Biemans 2018). Where possible, we use both instrumental variables to generate exogenous variations for internally and externally focused AI capabilities.

One concern about the instrumental variable is that the AI policies published by local governments may be correlated with regional socio-economic levels and thereby correlated with unobserved firm qualities. The number of AI-related news items from companies in the same industry may also be correlated with unobserved market dynamics and firm qualities. If that is the case, then we should also observe that instrumental variables are associated with start-up innovation activities even when the focal start-up has not adopted AI capability. We follow the procedures in Martin and Yurukoglu (2017) to conduct the reduced-form instrumental variable test with results reported in Appendix C3. Specifically, we create a binary variable of AI Adoption by identifying the year the company first lists AI-related job postings, social media articles or is the subject of a news report. We then separate our sample into AI adopters versus non-adopters and use the instrumental variables as the independent variable. The results show that our instrumental variables are only significantly associated with start-up innovation activities when start-ups have adopted AI, which provides further evidence for the validity of the instrumental variables.

To instrument the level of LSM, we use the average level of validated learning and iterative experimentation of other companies[32] in each year that (1) share a board member with the focal start-up and (2) are in a different industry and province from the focal start-up. Companies with a shared board member are likely to receive similar advice and consequently likely to formulate similar managerial and operational strategies (Kroll et al. 2008, Dechow and Tan 2021). At the same time, they are less likely to experience the same systematic shocks in performance faced by firms in the same industry or region. Thus, their product innovation performances are less likely to be correlated. Where possible, we use both instrumental variables to generate exogenous variations for validated learning and iterative experimentations. The first-stage F-statistics for all 2SLS results are over 20, suggesting that the results do not suffer from weak instrument issues.

5. Results and Interpretation

We first examine the effect of AI adoption on start-ups' product innovations in Table 19. Columns 1 to 6 show the effect of AI on software start-ups. We show that AI capability is positively associated with increased future software copyrights, including both novel and incremental ones after controlling for various firm characteristics, and 2SLS results are also similar. Economically, the results show that a one-standard-deviation increase in AI capability, which includes about six AI-related job postings, 30 AI-related social media articles and eight AI-related news reports, is associated with creating about one novel software

[32] Including companies that are outside our sample.

product and one incremental software product in the following year. To show that this effect is not only applicable to digital products, Columns 7 to 12 show the positive effect of AI capability for firms that produce physical products. Specifically, a one-standard-deviation increase of AI capability is associated with creating about one novel product and three incremental products in the following year. These results support H1.

We then examine the complementarity between AI capability and LSM level. Results are reported in Table 20. We first show the correlation test results in Column 1 for software companies and Column 6 for physical product companies. The estimated coefficient is positive. Next, we test the interaction effects between AI capability and LSM level in the performance regression. For software companies, Columns 2 and 3 show that AI capability and LSM are separately positively associated with increases in future software products, and Column 4 shows that the adoption of both as a system is significantly associated with producing more future software products. Both the correlation and performance tests suggest a complementary relationship between AI capability and LSM. The 2SLS results are also consistent. Note that after adding the interaction term between AI capability and LSM, the coefficient of AI capability is no longer statistically significant, suggesting that adopting AI capability without LSM has a limited effect on product innovation.

Similarly, for other physical product companies, Columns 7 and 8 separately show that AI capability and LSM are also positively associated with an increase in future products. Column 9 shows the complementarity between AI capability and LSM, and Column 10 shows the 2SLS results for physical products. These results largely mirror the results for digital products. Collectively, they indicate that a complementarity exists between AI capability and LSM for developing both software and physical products. All these results support H2.

Next, we separate AI capabilities into externally focused and internally focused AI capabilities and separate LSM into its two components of validated learning and iterative experimentation. In Table 21, we explore the complementarities between externally focused AI and validated learning and between internally focused AI and iterative experimentation. We explore their differences by examining how they increase novel and incremental products. Columns 1 and 2 show that validated learning and externally focused AI are complements in developing novel software that is new to the company, while the interaction between iterative experimentation and internally focused AI capability is not statistically significant. Columns 3 and 4, on the contrary, show that iterative experimentation and internally focused AI are complements in developing incremental software (a new version of existing products), while the interaction between validated learning and externally focused AI shows no complementarity. Columns 5 to 8 show similar results in companies that develop physical products. Our 2SLS results are directionally consistent. Collectively, such results support the mechanism that externally focused AI and validated learning can create a complementary system in market search and validation, which is especially beneficial for developing novel products; this supports H3a. On the other hand, internally focused AI and iterative experimentation can create another complementary system in product iteration and improvement, which is especially beneficial for developing incremental products; this supports H3b.

To further verify the mechanism, we also empirically test the complementarity between validated learning and internally focused AI, as well the complementarity between iterative experimentation and externally focused AI in Appendix C5. The interaction terms are not statistically significant, which is likely due to the different context knowledge for the externally and internally focused AI, as well as the different R&D objectives for developing novel and incremental products. These results suggest that different AI capabilities are complements to different types of product innovations. The correct matching between the type of AI capability and the type of LSM method is essential for creating the best match between AI and developing novel and incremental products that will result in the greatest desired impact.

6. Mechanism Check

To further verify the above mechanisms, in Table 22 we show that the joint adoption of externally focused AI and validated learning is helpful for start-ups to expand their business scopes. We measure the business scopes of each company in each year using the number of business scope entries on the company's business license. By law in China, companies can only operate in business domains that are shown in one's business scope entries. Companies that want to expand their business scope are required to apply for permission. Thus, business scopes and their expansion are a matter of record. The results in Table 22 support the mechanism that externally focused AI and validated learning are complements in expanding business scopes.

Also, if externally focused AI and validated learning create a complementary system in market search and validation, one should observe that companies adopting both to perform the best when firms face high market uncertainties. In Table 23 we show the three-way complementarity of validated learning, externally focused AI and market uncertainty of the start-ups. The results suggest that the complementarity of externally focused AI and validated learning in developing novel products is stronger when the start-up faces a higher level of market uncertainty, supporting the mechanism.

In Table 24 we show that the joint adoption of internally focused AI and iterative experimentation is associated with a higher speed of product iteration. We measure iteration speed using the average number of days between two versions of the same product. We also measure product quality using the number of product certificates that are issued by a neutral third party such as standard setting organizations. Using product speed and quality in the following year, we show that iterative experimentation and internally focused AI are complementary in product iterations and improvements.

We also analyze whether complementarities between internally focused AI and iterative experimentation would strengthen when facing greater market competition. In Table 25 we show the three-way complementarity of iterative experimentation, internally focused AI capabilities and market competition level. The results suggest that the complementarity of internally focused AI and iterative experimentation in developing incremental products is stronger when the start-up faces a higher level of market competition, supporting the mechanism.

7. Discussion

Our findings show the importance of treating AI as a heterogeneous construct because different AI capabilities require different organizational processes for support. Specifically, our results provide new evidence for the effect of AI capability on start-up product development and offer explanations for how such an effect is complemented by using LSM, a commonly adopted organizational practice. We show that adopting AI and LSM as a system in product development is associated with an increase in product innovation, including both novel products that are new to the company as well as incremental products that are new versions of existing products. We also show that adopting AI without also adopting LSM has a limited effect on product innovation. Our further analyses show that the validated learning aspect of LSM complements externally focused AI in developing novel products and expanding company business scopes, and the effect is stronger when the start-up faces a higher degree of market uncertainty; on the other hand, the iterative experimentation aspect of LSM complements internally focused AI in developing incremental products, facilitating product iterations and creating higher-quality products, and the effect is stronger when the start-up faces a greater degree of market competition.

As many recent studies have documented, although there has been a surge in AI capability adoptions in business over the past decade, most companies have not benefited from such investments. This paper provides one possible explanation for this phenomenon by empirically testing the complementarities between AI capabilities and LSM, and more importantly the matching between different types of AI focus to different aspects of LSM. Although the research context is in China, which has been a leading power in adopting AI capabilities in the past decade, we postulate that the results are generalizable to other business contexts because the types of AI tools and the novel and incremental types of innovation are universal. Also, the results in this study can be generalized into established firms as many mature companies are also using AI and LSM for intrapreneurship activities.

This paper has important managerial implications for multiple stakeholders in the entrepreneurship world. For start-ups and entrepreneurs, current AI tools with high prediction accuracy can improve productivity while leading to unintended problems due to their low interpretability and reliance on high-quality data. Thus, when deploying such tools, it is even more important for companies to iteratively obtain first-hand market feedback data to continuously improve their AI tools. Coupling LSM with AI could help mitigate market risks because together they help firms adjust the objective function AI is maximizing and create high fidelity features that AI uses to improve prediction accuracy. For venture investors and policy makers, it is critical to think about the organizational complementarities (what tool fits what objective, and how should the organizational capacity be built out) when encouraging new technology adoptions. Our research indicates that simply encouraging "AI adoption" without respect to its varying tools and uses might lead to limited business benefits, or worse, costs that do not get recouped.

This paper contributes to the emerging empirical literature on the business value of AI and AI complementarities. As an increasing number of companies adopt AI in business practices, it is important to understand how AI can benefit companies and what organizational complements are required to maximize the benefits. Our study examines the nuance effect of different AI capabilities in various usage scenarios, provides new evidence to explain the uneven returns of AI among early adopters and suggests that a better match between tools and objectives may improve those results.

This paper also contributes to the literature on innovation and entrepreneurship and offers managerial insights for start-ups. Although general purpose technologies like AI can benefit companies in product development, start-ups should also invest in complementary organizational practices such as LSM for product innovation and apply different types of AI capabilities to different aspects of LSM according to the different development objectives. Appropriate investment in AI and LSM can help start-ups to better scale up product innovation.

Future studies could consider other types of organizational complements to the AI capabilities, such as employee skillsets and organizational structure. Understanding how nascent technological tools should be better deployed inside organizations would be a fruitful future direction. Also, as AI technologies have been continuously evolving in recent years, future studies could further examine the business value of new AI tools such as large generative models and interpretable machine learning. As our findings suggest, we believe there is a rich vein of opportunity for future research that aims to pinpoint with ever-greater precision what it takes to achieve the maximum impact of AI by examining its varied forms and their organizational requirements.

Figure 9: Theoretical Model

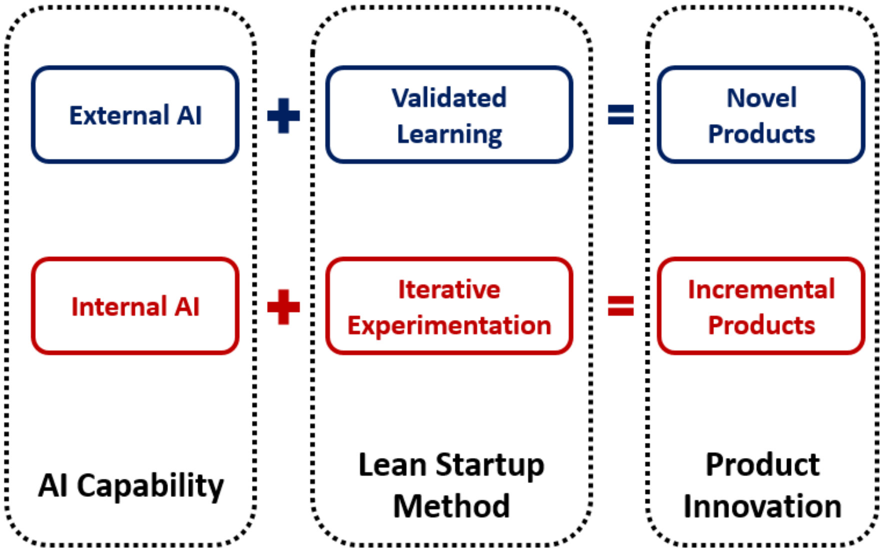


Figure 10: An Example of Software Copyright and Trademark (Translated in English)

Software Copyright (Example)

Trademark (Example)

Figure 11: Cumulative AI-related Capability using Different Measures

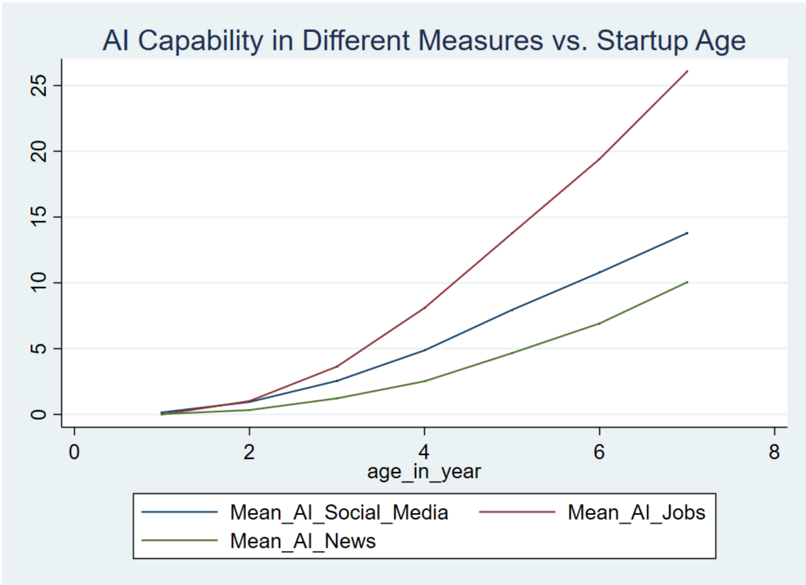


Figure 12: Provincial Distributions of AI Incentives in China (2011-2020)

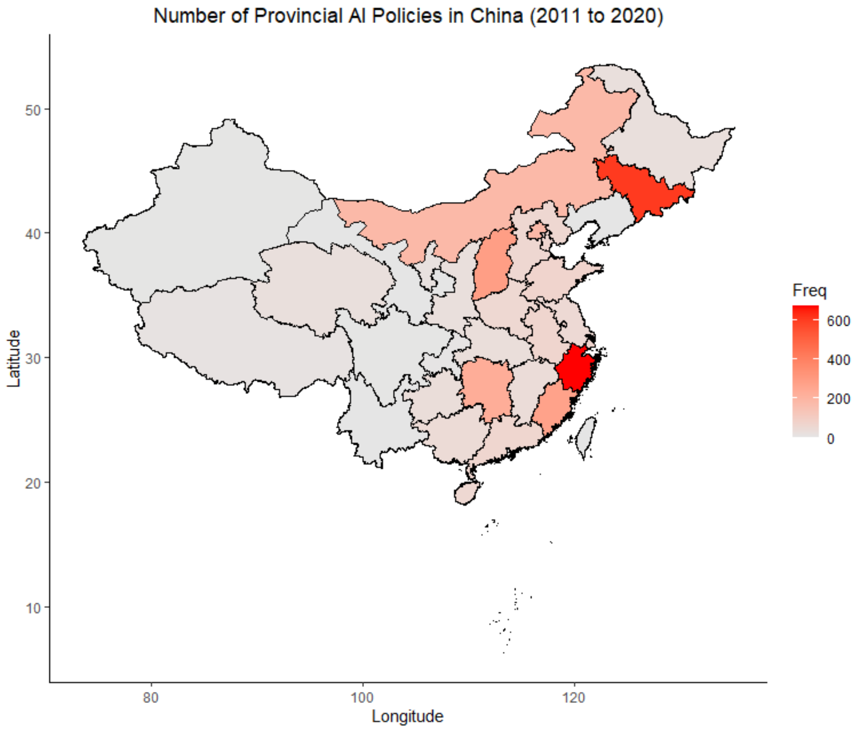


Table 16: Key Words Dictionary

Topic	Keywords	Translation in Chinese
AI	Artificial Intelligence, AI, Machine Learning, Deep Leaning, Neural Network	人工智能，AI，机器学习，深度学习，神经网络
Internal	Engineering, (Software) Architect, Algorithm, Hardware, System, Simulation	工程，架构，算法，硬件，系统/体系，模拟
External	Market, Demand, Marketing, Customer, User	市场，需求，营销，顾客，用户
Lean Start-up Method (LSM)	Test Product, Experimentation, Lab (Technician), (A/B) Testing, Scrum, Kanban, Validation, Verification, Trial and Error, Customer/Market Demand, Customer/Market Learning/Response	试产品，试验，实验室/员，(A/B)测试，验证，检验，顾客/市场需求，顾客/市场学习/响应
Validated Learning	Test Product, Validation, Verification, Trial and Error, Customer/Market Demand, Customer/Market Learning/Response	试产品，验证，检验，试错，顾客/市场需求，顾客/市场学习/响应
Iterative Experimentation	Experimentation, Lab (Technician), (A/B) Testing	试验，实验室/员，(A/B)测试

Table 17: Summary Statistics

Variable Category	Variable Name	No. Obs.	Mean	Std. Dev.	Min	Max
Dependent Variable	Lead Number of Software Copyrights	8,954	2.79	5.92	0	117
	Lead Number of Trademarks	8,954	17.15	66.93	0	2305
	Lead Number of Novel Software Copyrights	8,954	2.18	5.06	0	108
	Lead Number of Novel Trademarks	8,954	5.29	24.76	0	958
	Lead Number of Incremental Software Copyrights	8,954	6.54	13.19	0	321
	Lead Number of Incremental Trademarks	8,954	11.86	46.65	0	1347
	Lead Number of High-Quality Products	8,954	0.70	3.56	0	167
Independent Variable	AI Capability	10,641	0	1	-1.37	36.40
	Externally focused AI	10,641	0	1	-1.45	25.25
	Internally focused AI	10,641	0	1	-2.01	37.12
	LSM Level	10,641	0	1	-0.26	37.04
	Validated Learning	10,641	0	1	-0.34	33.15
	Iterative Experimentation	10,641	0	1	-0.22	37.37
Company Basic Information	Age	10,641	3.84	2.10	1	9
	Number of Branches	10,641	3.00	17.91	0	367
	Number of Subsidiaries	10,641	5.13	8.16	0	82
	Total Number of Job Posts	10,641	161.26	805.61	0	44938
	Total Number of Employees	10,641	134.65	714.94	0	35935
	Market Uncertainty	10,641	0.06	0.14	0.01	1
	Market Competition	10,641	15.03	7.21	1	97
Company Quality Information	Total Number of Administrative Approvals	10,641	1.77	3.20	0	96
	Total Number of Patents	10,641	23.40	128.37	0	4134
	Total Number of Certificates	10,641	1.44	9.56	0	553
	Total Number of Work Copyrights	10,641	11.10	557.84	0	47611
	Average Product Iteration Days	10,641	57.12	125.11	0.02	360
Company Funding Information	Total Funding Rounds	10,641	1.32	1.39	0	17
	Total Funding Amount	10,641	4.0e7	2.0e8	0	5.0e9

Table 18: Correlations between Main Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Lead Number of Software Copyrights	1.00								
(2) Lead Number of Trademarks	0.16	1.00							
(3) Lead Number of Novel Software Copyrights	0.64	0.14	1.00						
(4) Lead Number of Novel Trademarks	0.12	0.68	0.11	1.00					
(5) Lead Number of Incremental Software Copyrights	0.48	0.20	0.38	0.16	1.00				
(6) Lead Number of Incremental Trademarks	0.17	0.67	0.14	0.63	0.21	1.00			
(7) Lead Number of Certificates	0.12	0.13	0.10	0.09	0.12	0.13	1.00		
(8) AI Capability	0.22	0.17	0.19	0.17	0.32	0.16	0.08	1.00	
(9) Externally Focused AI	0.41	0.20	0.37	0.16	0.39	0.20	0.12	0.31	1.00
(10) Internally Focused AI	0.06	0.09	0.05	0.06	0.14	0.10	0.37	0.37	0.45
(11) LSM Level	0.20	0.41	0.18	0.34	0.29	0.41	0.08	0.39	0.18
(12) Validated Learning	0.08	0.15	0.06	0.10	0.13	0.17	0.01	0.12	0.06
(13) Iterative Experimentation	0.17	0.14	0.14	0.10	0.30	0.15	0.08	0.19	0.14
(14) Age	0.12	0.06	0.09	0.02	0.35	0.07	0.11	0.14	0.09
(15) Number of Branches	0.17	0.22	0.14	0.16	0.24	0.23	0.07	0.08	0.14
(16) Number of Subsidiaries	0.03	0.07	0.02	0.04	0.03	0.08	0.01	0.01	0.02
(17) Total Number of Job Posts	0.16	0.23	0.14	0.18	0.27	0.24	0.06	0.41	0.14
(18) Total Number of Employees	0.20	0.33	0.17	0.28	0.28	0.32	0.08	0.16	0.17
(19) Market Uncertainty	-0.16	-0.20	-0.14	-0.14	-0.27	-0.21	-0.07	-0.09	-0.14
(20) Market Competition	0.13	0.08	0.12	0.06	0.24	0.09	0.09	0.12	0.12
(21) Total Number of Administrative Approvals	0.13	0.58	0.10	0.49	0.34	0.58	0.11	0.30	0.10
(22) Total Number of Patents	0.13	0.14	0.11	0.12	0.28	0.13	0.11	0.28	0.11
(23) Total Number of Certificates	0.06	0.09	0.05	0.06	0.14	0.10	0.47	0.07	0.05
(24) Total Number of Work Copyrights	0.01	0.04	0.00	0.02	0.06	0.05	0.01	0.02	0.00
(25) Total Funding Rounds	0.08	0.15	0.06	0.10	0.13	0.17	0.01	0.12	0.06
(26) Total Funding Amount	0.17	0.14	0.14	0.10	0.30	0.15	0.08	0.19	0.14
Continue from above	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(10) Internally Focused AI	1.00								
(11) LSM Level	0.39	1.00							
(12) Validated Learning	0.18	0.45	1.00						
(13) Iterative Experimentation	0.09	0.50	0.43	1.00					
(14) Age	0.26	0.22	0.13	0.09	1.00				
(15) Number of Branches	0.38	0.15	0.23	0.14	0.35	1.00			
(16) Number of Subsidiaries	0.20	0.24	0.11	0.06	0.16	0.19	1.00		
(17) Total Number of Job Posts	0.15	0.24	0.09	0.14	0.16	0.15	0.17	1.00	
(18) Total Number of Employees	0.20	0.15	0.16	0.01	0.18	0.29	0.09	0.18	1.00
(19) Market Uncertainty	-0.27	-0.20	-0.26	-0.07	-0.51	-0.36	0.25	-0.15	-0.02
(20) Market Competition	0.08	0.15	0.06	0.01	0.07	0.08	0.05	0.11	0.15
(21) Total Number of Administrative Approvals	0.04	0.04	0.07	0.00	0.01	0.04	0.02	0.01	0.04
(22) Total Number of Patents	0.28	0.35	0.24	0.03	0.26	0.27	0.27	0.24	0.35
(23) Total Number of Certificates	0.21	0.12	0.19	0.08	0.18	0.22	0.08	0.09	0.12
(24) Total Number of Work Copyrights	0.27	0.47	0.16	0.10	0.25	0.19	0.17	0.18	0.47
(25) Total Funding Rounds	0.08	0.15	0.06	0.10	0.13	0.17	0.01	0.12	0.15
(26) Total Funding Amount									
Continue from above	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	
(19) Market Uncertainty	1.00								
(20) Market Competition	-0.47	1.00							
(21) Total Number of Administrative Approvals	-0.15	0.14	1.00						
(22) Total Number of Patents	-0.10	0.02	0.01	1.00					
(23) Total Number of Certificates	-0.34	0.28	0.13	0.07	1.00				
(24) Total Number of Work Copyrights	-0.22	0.20	0.02	0.00	0.12	1.00			
(25) Total Funding Rounds	-0.25	0.17	0.08	0.06	0.29	0.20	1.00		
(26) Total Funding Amount	-0.01	0.04	0.00	0.02	0.06	0.05	0.26	1.00	

Table 19: AI Capability Effect on Start-up Product Development

Company Industry	Software Companies						Physical Product Companies					
Dependent Variables	(1) Lead # Software Copyright	(2)	(3) Lead # Novel Software Copyright	(4)	(5) Lead # Incremental Software Copyright	(6)	(7) Lead # Product Trademark	(8)	(9) Lead # Novel Product Trademark	(10)	(11) Lead # Incremental Product Trademark	(12)
Model	FE	FE+2SLS	FE	FE+2SLS	FE	FE+2SLS	FE	FE+2SLS	FE	FE+2SLS	FE	FE+2SLS
AI Capability	1.78***	**2.71*****	0.97***	**1.23*****	0.81***	**1.08*****	3.84***	**6.13*****	1.20**	**1.73*****	2.65***	**4.43*****
	(0.30)	**(0.44)**	(0.27)	**(0.39)**	(0.11)	**(0.31)**	(1.32)	**(1.22)**	(0.52)	**(0.55)**	(0.94)	**(0.12)**
Controls	Company Fixed Effect, Year Fixed Effect, Company Age, Total Number of Employees, Total Number of Job Postings, Total Number of Patents, Total Number of Certificates, Total Number of Work Copyrights, Total Number of Administrative Approvals, Total Amount of Funding											
Observations	2,490	2,454	2,490	2,454	2,490	2,454	6,464	6,396	6,464	6,396	6,464	6,396
R-squared	0.477		0.439		0.548		0.585		0.528		0.666	

Notes:

1. Dependent variable is the number of new products (total number, novel product, incremental product) in the next year. AI Capability is standardized, where one-standard-deviation change is associated with about 1 AI patent, 6 AI job postings, 30 AI social media articles and 8 AI news reports.

2. Odd number columns report the OLS results with company fixed effect. Even number columns report the 2SLS results with company fixed effect. 2 sets of instruments: (1) Number of AI policies issued by local governments in each year; (2) Number of news reports that a company in the same industry has adopted AI.

3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 20: Test for Complementarity between AI Capability and Lean Start-up Method (LSM)

Company Industry	Software Companies					Physical Product Companies				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent Variables	AI Capability	Lead # Software Copyright				AI Capability	Lead # Product Trademark			
Model	FE	FE	FE	FE	FE+2SLS	FE	FE	FE	FE	FE+2SLS
AI Capability		1.777***		0.308	1.221		3.844***		-0.669	0.856
		(0.301)		(0.404)	(1.419)		(1.318)		(1.490)	(1.069)
LSM	0.078***		0.995***	0.341	1.056	0.047***		11.51***	11.19***	20.17***
	(0.017)		(0.234)	(0.251)	(0.983)	(0.016)		(1.427)	(1.438)	(2.085)
AI * LSM				3.183***	**5.718***				8.104***	**15.33***
				(0.661)	**(0.825)**				(1.804)	**(2.174)**
Controls	Company Fixed Effect, Year Fixed Effect, Company Age, Total Number of Employees, Total Number of Job Postings, Total Number of Patents, Total Number of Certificates, Total Number of Work Copyrights, Total Number of Administrative Approvals, Total Amount of Funding									
Observations	2,490	2,490	2,490	2,490	2,454	6,464	6,464	6,464	6,464	6,396
R-squared	0.646	0.477	0.472	0.485		0.612	0.585	0.590	0.591	

Notes:

1. For Column (1) and (6), dependent variable is AI Capability in the current year. For Column (2)-(5) and (7)-(10), dependent variable is the number of new products (total number) in the next year.

2. AI Capability is standardized, where one-standard-deviation change is associated with about 1 AI patent, 6 AI job postings, 30 AI social media articles and 8 AI news reports.

3. Column (5) and (10) report the 2SLS results with both AI Capability and LSM instrumented.

4. Robust standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 21: Complementarity Tests for Novel Products vs. Incremental Products

Company Industry	Software Companies				Physical Product Companies			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variables	Lead # Novel Software Copyright		Lead # Incremental Software Copyright		Lead # Novel Product Trademark		Lead # Incremental Product Trademark	
Model	FE	FE+2SLS	FE	FE+2SLS	FE	FE+2SLS	FE	FE+2SLS
Validated Learning	0.48	-0.14	0.07	0.04	0.48***	0.07	0.14	0.15
	(0.48)	(0.49)	(0.09)	(0.12)	(0.04)	(0.13)	(0.23)	(0.23)
External AI	0.44	0.43	0.42	0.16	0.22**	1.49	1.27	5.15*
	(0.34)	(0.32)	(0.31)	(0.32)	(0.09)	(1.59)	(1.32)	(2.81)
Validated Learning * External AI	0.43***	**1.29*** **	-0.09	**0.12***	1.38***	**5.60** **	1.57*	**1.24**
	(0.07)	**(0.40)**	(0.54)	**(0.07)**	(0.19)	**(2.75)**	(0.91)	**(0.92)**
Iterative Experimentation	-0.04	0.01	0.55***	0.46***	0.33	0.23	0.51*	0.53**
	(0.11)	(0.10)	(0.11)	(0.10)	(0.56)	(0.19)	(0.26)	(0.26)
Internal AI	0.21	0.65	0.34*	0.46	-1.07	0.92	1.39*	4.76*
	(0.26)	(0.58)	(0.19)	(0.43)	(2.01)	(1.56)	(0.78)	(2.75)
Iterative Experimentation * Internal AI	0.20	**-0.14**	0.23**	**0.98*** **	0.33	**0.59**	1.72***	**5.52** **
	(0.14)	**(0.37)**	(0.09)	**(0.28)**	(1.21)	**(1.70)**	(0.37)	**(2.77)**
Controls	Company Fixed Effect, Year Fixed Effect, Company Age, Total Number of Employees, Total Number of Job Postings, Total Number of Patents, Total Number of Certificates, Total Number of Work Copyrights, Total Number of Administrative Approvals, Total Amount of Funding							
Observations	2,490	2,454	2,490	2,454	6,464	6,396	6,464	6,396
R-squared	0.488		0.789		0.460		0.550	

Notes: 1. For Column (1)-(2) and (5)-(6), dependent variable is the number of novel products in the next year. For Column (3)-(4) and (7)-(8), dependent variable is the number of incremental products in the next year.
2. Column (2), (4), (6) and (8) report the 2SLS results with both AI Capability and LSM instrumented.
3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 22: External AI and Validated Learning in Expanding Business Scope

Company Industries	Software Companies			Physical Product Companies		
Dependent Variable	Lead # Business Scope Entries			Lead # Business Scope Entries		
	(1)	(2)	(3)	(10)	(11)	(12)
Validated Learning	-0.035		0.031	0.122***		0.051
	(0.100)		(0.133)	(0.036)		(0.034)
External AI		0.302**	0.202*		0.163***	0.164***
		(0.117)	(0.108)		(0.028)	(0.036)
Validated Learning * External AI			1.175***			0.203***
			(0.227)			(0.036)
Controls	Company Fixed Effect, Year Fixed Effect, Company Age, Total Number of Employees, Total Number of Job Postings, Total Number of Patents, Total Number of Certificates, Total Number of Work Copyrights, Total Number of Administrative Approvals, Total Amount of Funding					
Observations	2,490	2,490	2,490	6,464	6,464	6,464
R-squared	0.583	0.584	0.589	0.699	0.699	0.702

Notes:

1. Dependent variable is the number of business scope entries in the next year. Column (1)-(3) are software companies. Column (4)-(6) are physical product companies.

2. Externally Focused AI and Validated Learning are standardized.

3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 23: External AI and Validated Learning in Uncertain Markets

Company Industries	Software Companies	Physical Product Companies
Dependent Variable	Lead # Novel Software Copyright	Lead # Novel Product Trademark
	(1)	(2)
Externally Focused AI	0.702*	1.784**
	(0.358)	(0.839)
Validated Learning	0.304	2.715***
	(0.222)	(0.572)
External AI * Validated Learning	0.840	-1.792
	(0.605)	(1.825)
Market Uncertainty	-0.084*	-0.467***
	(0.05)	(0.115)
External AI * Validated Learning * Market Uncertainty	52.90***	5.246***
	(5.57)	(0.906)
Controls	Company Fixed Effect, Year Fixed Effect, Company Age, Total Number of Employees, Total Number of Job Postings, Total Number of Patents, Total Number of Certificates, Total Number of Work Copyrights, Total Number of Administrative Approvals, Total Amount of Funding	
Observations	2,490	6,464
R-squared	0.470	0.425

Notes:

1. Dependent variable is the number of novel products in the next year.
2. Externally Focused AI and Validated Learning are standardized.
3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 24: Internal AI and Iterative Experimentation in Iteration Speed and Product Quality

Industry	Software Companies						Physical Product Companies					
DV	Lead Average Product Iteration Days			Lead # Certificated Products			Lead Average Product Iteration Days			Lead # Certificated Products		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Iterative Experimentation	-1.43		-11.78**	0.20*		0.36	-2.90		-11.07***	0.12**		0.21**
	(4.96)		(5.60)	(0.12)		(0.26)	(2.16)		(2.66)	(0.06)		(0.09)
Internal AI		-24.40***	-2.95		0.30***	0.13		-29.72***	-32.38***		0.22***	0.14*
		(4.18)	(4.53)		(0.09)	(0.10)		(2.79)	(2.83)		(0.06)	(0.07)
Iterative Experimentation * Internal AI			-46.62***			0.33**			-16.02***			0.21***
			(9.57)			(0.15)			(2.82)			(0.05)
Controls	Company Fixed Effect, Year Fixed Effect, Company Age, Total Number of Employees, Total Number of Job Postings, Total Number of Patents, Total Number of Certificates, Total Number of Work Copyrights, Total Number of Administrative Approvals, Total Amount of Funding											
Observations	2,490	2,490	2,490	2,490	2,490	2,490	6,464	6,464	6,464	6,464	6,464	6,464
R-squared	0.461	0.468	0.474	0.547	0.549	0.550	0.448	0.459	0.460	0.666	0.667	0.667

Notes:

1. For Column (1)-(3) and (7)-(9), dependent variable is the average number of days between two versions of the same products in the next year. For Column (4)-(6) and (10)-(12), dependent variable is the number of certificated products in the next year.

2. Internally focused AI and Iterative Experimentation are standardized.

3. Robust standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 25: Internal AI and Iterative Experimentation in Competitive Markets

Company Industries	Software Companies	Physical Product Companies
Dependent Variable	Lead # Incremental Software Copyright	Lead # Incremental Product Trademark
	(1)	(2)
Internally Focused AI	0.556	3.139**
	(0.644)	(1.495)
Iterative Experimentation	0.124	8.794***
	(0.089)	(1.019)
Internal AI * Iterative Experimentation	0.864***	3.213***
	(0.242)	(1.470)
Market Competition	-0.009	1.917
	(0.020)	(1.405)
Internal AI * Iterative Experimentation * Market Competition	0.403*	0.334**
	(0.223)	(0.161)
Controls	Company Fixed Effect, Year Fixed Effect, Company Age, Total Number of Employees, Total Number of Job Postings, Total Number of Patents, Total Number of Certificates, Total Number of Work Copyrights, Total Number of Administrative Approvals, Total Amount of Funding	
Observations	2,490	6,464
R-squared	0.789	0.580

Notes:

1. Dependent variable is the number of incremental products in the next year.
2. Internally focused AI and Iterative Experimentation are standardized.
3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

APPENDIX A: APPENDIX FOR CHAPTER 2

A1. Instrumental Variable Validity Check

To provide support for the validity of our instrument, we check whether the instrumental variables predict the future levels of funding in companies that have not adopted Twitter following the procedures in Martin and Yurukoglu (2017). The logic is as follows: if the instrument is valid, then it should be correlated *only* to the focal company's funding through its effect on the focal company's Twitter activities. For companies without Twitter activities, we should not observe a significant correlation between the instrument and the outcome variable. To test this, we regress the dependent variable on the instrumental variables directly, using only data from observations that did or did not adopt Twitter.

The first two columns of Table A1 show the result of isolated companies. We find that there is statistically significant relationship between the instruments and the dependent variable when the focal observations have adopted Twitter, and that there is no statistically significant relationship between the instrumental variable and the dependent variable when the focal observations have not adopted Twitter. In Column (3) to (4) we show the results of connected companies. Similarly, we find that there is statistically significant relationship between the instruments and the dependent variable when the focal observations have adopted Twitter, and that there is no statistically significant relationship between the instruments and the dependent variable when the focal observations have not adopted Twitter. Collectively, these results provide support for the validity of our instrumental variables.

Table A1: IV Validity Check

Sample of Observations	Isolated Companies		Connected Companies	
DV: Log lead funding	(1)	(2)	(3)	(4)
Whether Adopted Twitter?	No	Yes	No	Yes
Location Gender Diversity * Location Twitter Engagement	0.062	0.133***		
	(0.063)	(0.035)		
Location Network Constraint * Location Twitter Engagement			0.346*	0.022***
			(0.192)	(0.007)
Investor Network Constraint * Network Twitter Engagement			-0.056	0.222***
			(0.104)	(0.069)
Other Controls	Year Indicators, Market-Year Fixed Effect, Firm Age, Total Funding Rounds, Website Rank, Google Trends		Year Indicators, Market-Year Fixed Effect, Firm Age, Total Funding Rounds, Website Rank, Google Trends, Network Degree, Network Transitivity, Network Constraint	
Observations	37,883	26,193	19,974	14,096
R-squared	0.518	0.635	0.512	0.614

Notes: 1. Dependent variable is the log funding in the next year. Location Gender Diversity is the average gender diversity of geographical neighbors of the focal company. Location Network Constraint is the average network constraint of network neighbors of the focal company. Investor Network Constraint is the average network constraint of network neighbors of the focal company.

2. Location Twitter Engagement is the average Twitter Engagement of geographical neighbors of the focal company. Network Twitter Engagement is the average Twitter Engagement of investor network neighbors of the focal company.

3. Columns (1) and (2) report the company fixed effect regression results for isolated companies. Columns (3) and (4) report the company fixed effect regression results for connected companies.

4. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

A2. Coarsened Exact Matching Results

To control the *ex-ante* differences between the treatment and control samples, we use coarsened exact matching (Blackwell et al. 2009) to match explicitly on observable characteristics. Table A2 shows the summary statistics for matched isolated observations using gender diversity as treatment, and Table A3 shows the fixed effect regression result using the matched groups. The investor network controls are not matched here because isolated companies are not part of the network. The results are consistent with our above results using full sample.

Similarly, Table A4 shows the summary statistics for matched observations for connected firms using network positions as treatment, where high diversity group includes firms with above-median network positions, and low diversity group includes firms with below-median network positions. The investor network controls are not matched here because we already use the investor network diversity as the treatment to divide the sample. Table A5 shows the fixed effect regression results using the matched groups, and the results are consistent with our above results using full sample.

Table A2: Summary Statistics for Gender Treatment using Coarsened Exact Matching

	Female	Male
Age	3.49 (4.88)	4.53 (5.59)
Total Funding Rounds	0.39 (0.81)	0.48 (0.80)
Website Rank	2,033,403 (1,383,965)	2,003,188 (1,404,067)
Google Trends	12.67 (19.31)	11.92 (18.23)
Founder-Alumni Degree Maximum	30.79 (98.66)	17.13 (65.95)
Observations	6,552	59,392

Table A3: Twitter Effect on Gender Diversity after Coarsened Exact Matching

DV: Log lead funding	Fixed Effect: Company ID	
VARIABLES	(1)	(2)
Has twitter	0.622***	0.631***
	(0.077)	(0.077)
Twitter engagement	0.397***	0.385***
	(0.046)	(0.045)
Gender diversity * Has twitter	0.392***	0.245***
	(0.056)	(0.070)
Gender diversity * Twitter Engagement		0.105**
		(0.051)
Other Controls	Year Indicators, Market-Year Fixed Effect, Firm Age, Total Funding Rounds, Website Rank, Google Trends, Alumni Network Degree	
Observations	57,329	57,329
R-squared	0.489	0.489

Notes: 1. Dependent variable is the log funding in the next year. Twitter engagement is standardized, where one-standard-deviation change is associated with about 300 Tweets and 30,000 reposts, comments and favorites posted each year.

2. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Summary Statistics for Network Diversity using Coarsened Exact Matching

	High diversity	Low diversity
Age	7.04 (4.61)	7.27 (5.31)
Total Funding Rounds	2.37 (1.61)	1.76 (1.44)
Website Rank	1,884,523 (1,423,300)	1,811,665 (1,452,000)
Google Trends	14.09 (18.46)	12.07 (17.73)
Founder-Alumni Degree Maximum	31.31 (86.66)	18.12 (64.23)
Observations	19,073	17,818

Table A5: Twitter Effect on Network Diversity after Coarsened Exact Matching

DV: Log lead funding	Fixed Effect: Company ID	
VARIABLES	(1)	(2)
Has twitter	-0.109	-0.125
	(0.118)	(0.118)
Twitter engagement	0.165***	0.181***
	(0.058)	(0.057)
Investor network constraint * Has twitter	0.084	-0.049
	(0.057)	(0.064)
Investor network constraint * Twitter engagement		0.189***
		(0.049)
Other Controls	Year Indicators, Market-Year Fixed Effect, Firm Age, Total Funding Rounds, Website Rank, Google Trends, Alumni Network Degree	
Observations	31,048	31,048
R-squared	0.503	0.503

Notes: 1. Dependent variable is the log funding in the next year. Twitter engagement is standardized, where one-standard-deviation change is associated with about 300 Tweets and 30,000 reposts, comments and favorites posted each year.

2. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

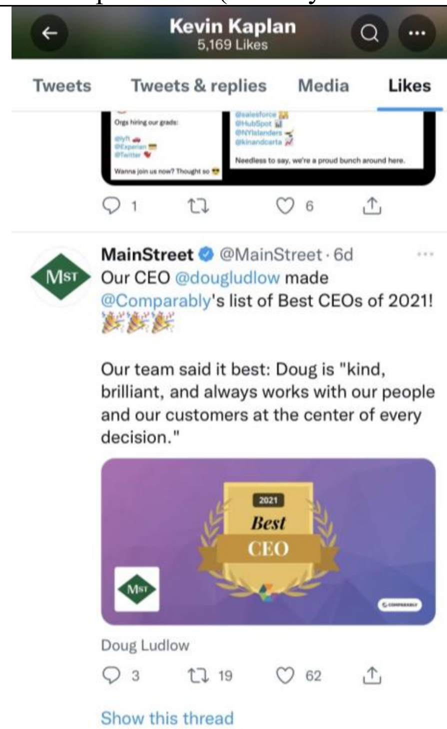
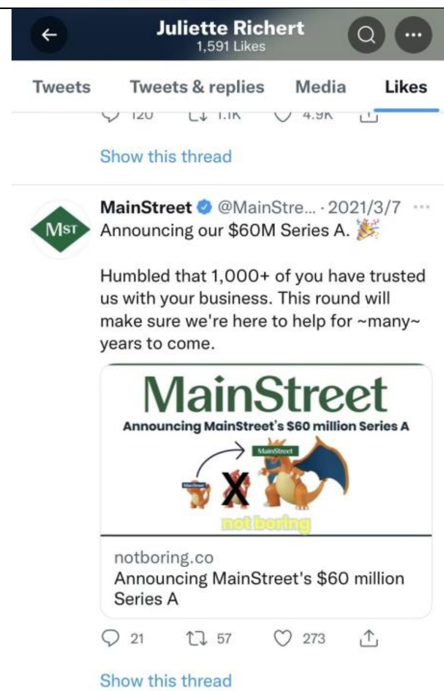
A3. Twitter Content and Natural Language Processing

We use natural language processing techniques to measure the quality of signals provided by start-ups. Under our hypotheses, the effect of Twitter should be stronger for information salient to (potential) start-up performance. We believe there are (at least) 3 main topics of quality signals on Twitter: (a) The quality of the start-up itself, including the product quality and team quality; (b) The quality of the start-up's marketing mix, i.e., how they promote the sales; (c) The quality of their customer relationships, i.e., after-sale services. These signals cover the entire process of a product from production to sales and to repurchase.

To better illustrate the quality signals or information provided by start-ups, we show in the following paragraphs the quantitative analyses together with qualitative examples. As discussed in the paper, we find the top 3 topics discussed by these start-ups on Twitter: Business Announcement, Product Promotion and Customer Relations. We now show Tweets of such topics attract the attention of investors and are helpful for raising future funding, especially for start-ups that face more serious information asymmetry problems. We also add Human Resource Management into our analysis because it measures team quality which is associated with funding outcomes. For each of the above 4 topics, we build a dictionary of word stems that help us identify specific Tweets using Latent Dirichlet Allocation. We show the dictionary and example Tweet in the following table.

The first column in the above table shows the different topics of Tweets that we identify. The second column shows the key word stem dictionary for each topic, and the third column shows an example Tweet of such topic posted by a start-up and liked by an investor who identifies themselves as an investor on their Twitter accounts and has no investments in the focal company. Note, these Tweets are liked by investors who have not invested or made announcements about investing in that firm. This can help us directly show how investors get to know about a firm that they have not invested before on Twitter. These qualitative examples show that there is indeed a quality signal transmission from start-ups to investors through social media.

Table A6: Twitter Topics, Key Word Dictionary and Example Tweets

Topics	Key Word (Stem) Dictionary	Example Tweet (liked by an investor)
Human Resource Management	cXo, chief, develop, coach, expert, team, join, crew, executive, founder, entrepre, head, office	
Business Announcement	Annouc, app, connect, creat, award, improv, deal, busi, design, cto, develop, coach, expert, crew, executive, founder, entrepre, head	

Product Promotion	Festiv, discount, free, fan, custom, buy, campus, deal, offer, special	
Customer Relations	Custom, connect, account, call, face, blog, claim, appolog, thank, friendship	

A4. Funding Probability and Dynamic Panel Model

In Table A7, we show the results if we use the probability of being funded as opposed to using the amount of funding as the dependent variable. The results in Table A7 are directionally consistent with prior results - Twitter activities have positive associations with the probability of getting funded in the next year with the effect being larger for less connected firms in the connected firm sample and for gender diverse firms in the isolated firm sample.

In Table A8, we show our results are robust after using system GMM estimators (Arellano and Bond 1991, Blundell and Bond 1998), suggesting that the effects are unlikely to be driven by time-series endogeneity.

Table A7: Twitter Effect for Increasing Funding Probability

DV: Pr(Lead Fund)	Connected Companies		Isolated Companies	
	LPM + FE	Logit + FE	LPM + FE	Logit + FE
VARIABLES	(1)	(2)	(3)	(4)
Has Twitter	0.015*	-0.006	0.056***	0.394***
	(0.008)	(0.090)	(0.005)	(0.073)
Twitter engagement	0.006	0.234***	0.008***	0.226***
	(0.004)	(0.049)	(0.003)	(0.044)
Investor network constraint * Has twitter	0.012**	0.094		
	(0.006)	(0.084)		
Investor network constraint * Twitter engagement	0.005	0.085**		
	(0.004)	(0.041)		
Gender diversity * Has twitter	0.000	-0.035	0.013***	0.113*
	(0.007)	(0.083)	(0.004)	(0.062)
Gender diversity * Twitter Engagement	-0.001	0.007	0.001	0.009
	(0.003)	(0.043)	(0.003)	(0.039)
Other Controls	Year Indicators, Firm Age, Total Funding Rounds, Website Rank, Google Trends, Network Degree, Network Transitivity, Network Constraint		Year Indicators, Firm Age, Total Funding Rounds, Website Rank, Google Trends	
Observations	40,959	35,488	75,025	71,017

Notes: 1. Dependent variable is binary (1 if get funded the next year, 0 otherwise). Twitter engagement is standardized, where one-standard-deviation change is associated with about 300 Tweets and 30,000 reposts, comments and favorites posted each year. Column (1) and (2) report results for connected companies. Column (3) and (4) report results for isolated companies.
2. Column (1) and (3) use linear probability model. Column (2) and (4) use logit model. Company fixed effects are added.
3. Standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A8: Twitter Effect on Connected Companies with Dynamic Panel Model

DV: Log Lead Funding	Dynamic Panel Model	
VARIABLES	(1)	(2)
Inverse mills ratio (connected)	0.166	0.164
	(0.151)	(0.152)
Log funding this year	-0.164***	-0.164***
	(0.014)	(0.014)
Has twitter	-0.492***	-0.491***
	(0.132)	(0.132)
Twitter engagement	0.060	0.059
	(0.054)	(0.055)
Investor network constraint * Has twitter	0.094	0.093
	(0.091)	(0.091)
Investor network constraint * Twitter engagement	0.109**	0.109**
	(0.050)	(0.050)
Gender Diversity * Has twitter		-0.027
		(0.119)
Gender Diversity * Twitter engagement		0.009
		(0.041)
Other Controls	Year Indicators, Market-Time Fixed Effect, Firm Age, Total Funding Rounds, Website Rank, Google Trend, Alumni Network Degree, Investor Network Degree, Investor Network Transitivity, Investor Network Constraint	
Observations	34,070	34,070

Notes: 1. Dependent variable is the log funding in the next year. Twitter engagement is standardized, where one-standard-deviation change is associated with about 300 Tweets and 30,000 reposts, comments and favorites posted each year. Inverse mills ratio for connected companies is added.
2. Columns (1) to (2) report the system GMM results with company fixed effects and market-time fixed effects using the STATA package *xtdpdsys*.
3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

A5. Falsification Test for Founder-Level Networks

One endogeneity concern about our results is that the investor network constraint could also measure the social capital of a company in general, not only in the perspective of investor relationships. In order to differentiate a company's social capital in the investor relationship networks versus a company's social capital in general, in Table A9 we do a falsification test. Specifically, we replace Investor Network Constraint with Alumni Network Degree, which is calculated based on the founder alumni networks and measures the social capital of founders. The results show that: (1) Alumni Network Degree is neither statistically nor economically significant for future funding outcome, as the estimated coefficient is 0.000. This result shows that founders' personal social capital is less helpful on start-up funding performances once the company is already connected. (2) Social media activities are not helpful for founders with less social capital in general, as the estimated coefficients for interaction terms are not statistically significant. These results show that social media engagement helps firms lacking social capital in investor networks instead of firms lacking social capital in general. Such results highlight the importance of social capital in investor networks rather than personal networks after the company has secured prior funding. Moreover, such results also indicate that the use of social media and funding outcomes are not correlated with the founder's social connections in general, alleviating potential endogeneity concerns about our results that start-ups with "higher quality" are more likely to engage on social media.

Table A9: Twitter Effect on Connected Companies (with Alumni Network Degree)

DV: Lead Log Funding	Company Fixed Effects		
VARIABLES	(1)	(2)	(3)
Inverse mills ratio (connected)	-0.117	-0.121	-0.119
	(0.123)	(0.123)	(0.123)
Has twitter	-0.457***	-0.453***	-0.449***
	(0.123)	(0.123)	(0.123)
Twitter engagement	0.179***	0.172***	0.169***
	(0.053)	(0.054)	(0.054)
Alumni Network Degree	0.000	0.000	0.000
	(0.001)	(0.000)	(0.001)
Alumni Network Degree * Has twitter	-0.017		-0.047
	(0.080)		(0.086)
Alumni Network Degree * Twitter engagement		0.020	0.030
		(0.029)	(0.029)
Other Controls	Year Indicators, Market-Time Fixed Effects, Firm Age, Total Funding Rounds, Website Rank, Google Trends, Investor Network Degree, Investor Network Transitivity, Investor Network Constraint, Linear Time Trend		
Observations	34,070	34,070	34,070
R-squared	0.492	0.492	0.492

Notes: 1. Dependent variable is the log funding in the next year.

2. All columns report the OLS results with company fixed effect. The sample only include the connected companies.

3. Twitter engagement is standardized, where one-standard-deviation change is associated with about 300 Tweets and 30,000 reposts, comments and favorites posted each year. Inverse mills ratio for connected companies is added.

APPENDIX B: APPENDIX FOR CHAPTER 3

B1. Variable Descriptions

In this section we provide the detailed descriptions for the constructs of variables used in the manuscript. Table A1 shows the keywords that are used to identify different social media topics, and Table A2 shows the detailed explanations for each variable.

Table A1: Social Media Contents and Key Words Dictionary

Topic	**Key Word Dictionary**
Business Announcement	Announce, announcement, app, application, product, connect, create, award, improve, deal, business, design, cXo, develop, coach, expert, crew, executive, founder, entrepreneur, head.
Marketing Promotion	Festival, discount, free, fan, custom, buy, campus, deal, offer, special.
Customer Relations	Customer, connect, account, call, face, blog, claim, apology, thank, reply, friendship.
Corporate Social Responsibility	Society, support, help, responsibility, environment, sustainability, together, quality, safety, ethnicity, fair, green, legal.

Table A2: List of Variables

Variable Category	Variable Name	Time Attribute	Variable Detail
Company Basic Information	Age	Time-Variant	The age of the focal company at the focal year.
	Location (City)	Time-Fixed	The name of the city that the focal company is registered in.
	Registration Capital	Time-Fixed	The registration capital of the focal company (in US dollars) at the focal year. Very few companies have ever changed their registration capital, so we treat this variable as time-fixed.
	Industry Category	Time-Fixed	The registered industry category of the focal company based on National Economic Category (GB/T4754-2017), including 15 industry sectors (A, C, E, F, G, H, I, J, L, M, N, O, P, Q, R).
	Number of Branches	Time-Variant	The number of branches of the focal company at the focal year.
	Number of Subsidiaries	Time-Variant	The number of subsidiaries of the focal company at the focal year.
	Number of Employees	Time-Variant	The number of social security payers of the focal company at the focal year.
Company Quality Information	Number of Trademarks	Time-Variant	The cumulative number of granted trademarks of the focal company at the end of the focal year.
	Number of Patents	Time-Variant	The cumulative number of granted patents of the focal company at the end of the focal year.
	Number of Certificates	Time-Variant	The cumulative number of granted certificates of the focal company at the end of the focal year.
	Number of Work Copyrights	Time-Variant	The cumulative number of granted work copyrights of the focal company at the end of the focal year.
	Number of Software Copyrights	Time-Variant	The cumulative number of granted software copyrights of the focal company at the end of the focal year.
Company Institutional Information	Political Distance	Time-Fixed	1/(The number of political connections + 1). The number of political connections is calculated as the number of founders and board members who have prior working experiences at governments or state-owned enterprises.
	Number of Administrative Approvals	Time-Variant	The number of newly added administrative approvals received by the focal company at the focal year.
	Number of Administrative Punishments	Time-Variant	The number of newly added administrative punishments received by the focal company at the focal year.
	Number of Positive Media Stories	Time-Variant	The number of newly added positive media reports from the major news agencies in China

			(including private and state-owned agencies) received by the focal company at the focal year.
	Number of Negative Media Stories	Time-Variant	The number of newly added negative media reports from the major news agencies in China (including private and state-owned agencies) received by the focal company at the focal year.
Company Funding Information	Total Funding Rounds	Time-Variant	The total number of funding rounds that the focal company has already had at the beginning of the focal year.
	Total Funding	Time-Variant	The total amount of funding (in US dollars) that the focal company has already had at the beginning of the focal year.
	Funded by Tencent	Time-Variant	The dummy variable that indicates whether the focal company has been invested by Tencent-related venture funds at the beginning of the focal year.
Social Media	Has WeChat	Time-Variant	The dummy variable that indicates whether the focal company has started to use WeChat Public Account at the focal year.
	WeChat Engagement (Standardized)	Time-Variant	The first principal component of the total number of articles posted, the total number of readings, comments and likes received by the focal company's WeChat Public Account at the focal year.
	Business Announcement	Time-Variant	The standardized value of the number of business announcement-related articles posted by the focal company's WeChat Public Account at the focal year.
	Marketing Promotion	Time-Variant	The standardized value of the number of marketing promotion related articles posted by the focal company's WeChat Public Account at the focal year.
	Customer Relations	Time-Variant	The standardized value of the number of customer relations related articles posted by the focal company's WeChat Public Account at the focal year.
	CSR	Time-Variant	The standardized value of the number of CSR related articles posted by the focal company's WeChat Public Account at the focal year.
Other Metrics	Industry Administrative Competitiveness	Time-Variant	1 minus the proportion of companies in the same industry category of the focal company that have received at least one administrative approval at the focal year.

B2. Robustness Check

In this section we check whether our results are robust under various econometric models. We first conduct negative binomial regressions to substitute the OLS regressions before, as the numbers of administrative approvals and positive media stories are positive integers. The results are shown in Table B3. Columns (1) and (2) show the effect of social media activities on future administrative approvals, and columns (3) and (4) show the effect of social media activities on future positive media reports. All four models add company fixed effects and year fixed effects. The results are similar to the results in Table 3 and Table 4: social media engagement is positively associated with more institutional benefits in the future, and the effect is stronger for politically distant companies. Also, columns (2) and (4) show that engagement on social media platform matters while pure adoption is helpless. All these results show that our findings are robust using count models.

Furthermore, we also conduct coarsened exact matching to verify that our results are robust to possible sample selection issues (e.g., better companies choose to use social media). We divide the sample into treatment group (companies with WeChat account) and control group (companies without WeChat account) and match the observations according to the firm characteristics in the data. The matching results are shown in Table B4.

Next, using the matched observations, we run OLS regressions with fixed effects to test the effect of social media activities on formal and informal institutions. The results are shown in Table B5. As shown here, social media engagement is still positively associated with more future administrative approvals and positive media reports, especially for politically distant companies, and the economic scales of the results are also close to the results before in Table 3 and Table 4, suggesting that the results are robust under sample selections.

Moreover, as two-way fixed effect regression estimators can be biased due to inappropriate weights among time-variant treatment effects (Goodman-Bacon 2021), we conduct additional event study tests based on the staggered difference-in-difference doubly robust estimator introduced by Callaway and Sant'Anna (2021). We take the company's social media adoption as treatment and take companies that never adopt WeChat as controls. The aggregated average treatment effects (ATT) with bootstrapped standard errors (SE) on institutional environments in different group-time are shown in Table B6, and Figure B1 shows the relative-time event-study plot. Overall, the results are consistent with our two-way fixed effect regression results that social media adoption is associated with better institutional environments including both formal institutions and informal institutions. Specifically, the group-time average treatment effects are more statistically significant in years before 2015 for administrative approvals, and before 2017 for positive media stories, which also verifies the mechanism that the institutional benefits of social media have decreased after MEIP was implemented.

To verify H2, we decompose our sample into two subgroups of companies: the companies that are politically disconnected (without any political connections in the company) versus the companies that are politically connected (with at least one political connection in the company). In Figure B2 and B3, we show the event studies for the two groups using Callaway and Sant'Anna (2021). The results demonstrate that social media adoption is more beneficial for politically disconnected companies than politically connected companies in obtaining institutional benefits.

Lastly, we conduct event study tests based on synthetic difference-in-difference to address the potential parallel trend differences between social media adopters and non-adopters (Arkhangelsky et al. 2021), as there could be concerns in the difference-in-difference setting about the validity of parallel trends between WeChat adopters and non-adopters. As a generalization of the difference-in-difference approach, synthetic control methods offer a formalized and more rigorous method for identifying comparison cases. By constructing a "synthetic" control unit that represents a weighted combination of many untreated cases, we reduce the biases from endogenous factors as a result of being in unbalanced parallel trends (Abadie and

Gardeazabal 2003, Abadie et al. 2010, Arkhangelsky et al. 2021). We take the company's social media adoption as treatment and take companies that later adopt WeChat as components of synthetic controls.

In Figure B4, we show the yearly average number of administrative approvals/positive media reports received by the treatment group, including the companies that have adopted WeChat public accounts in 2014, versus the synthetic control group.[33] The latter is a weighted combination of companies that have not adopted WeChat public accounts and are matched to the treatment group using a synthetic difference-in-difference method (Arkhangelsky et al. 2021).

Figure B4 shows that, after adopting social media, the treatment group starts to receive more administrative approvals/positive media reports than the synthetic control group, and the differences have grown in the subsequent years. This suggests that social media help firms improve their standing in institutional environments, and such benefits become stronger over time. Meanwhile, the differences in administrative approvals and positive media stories gradually shrink in 2016 and 2017, suggesting that the effect of social media on institutional barriers faced by firms has diminished following the wide implementation of the MEIP.

As a summary of the robustness tests, we find our results are robust under various model specifications and matched samples. The coarsened exact matching and synthetic difference-in-difference results provide further evidence that our findings are unlikely due to selection biases such as "better firms choose to use social media".

[33] The *synthdid* package requires (1) block adoption/treatment and (2) at least two time periods before adoption/treatment, so we choose to plot companies who adopt WeChat in 2014, which is also before MEIP was implemented.

Table A3: Social Media Effects using Fixed Effects Negative Binomial Regression

Dependent Variable	Lead Administrative Approvals		Lead Positive Media Reports	
FE: Company ID	(1)	(2)	(3)	(4)
Inverse Mills Ratio (Admin)	-2.197	-2.036		
	(1.715)	(1.721)		
Inverse Mills Ratio (Media)			-11.71***	-11.55***
			(3.089)	(3.090)
Has WeChat	0.331***	0.450	0.653***	0.413
	(0.037)	(0.279)	(0.048)	(0.286)
WeChat Engagement	0.269***	0.375***	0.387***	0.526***
	(0.037)	(0.044)	(0.060)	(0.072)
Has WeChat * Political Distance		-0.133		0.238
		(0.284)		(0.293)
WeChat Engagement * Political Distance		0.094***		0.130***
		(0.022)		(0.036)
Other Controls	Age, Number of Employee, Number of Trademark, Number of Patent, Number of Certificate, Number of Work Copyright, Number of Software Copyright, Total Funding Rounds, Total Funding Amounts, Funded by Tencent, Year Index			
Observations	7,127	7,127	7,127	7,127

Notes:

1. Negative binomial regression model with fixed effects are used here.

2. Column (1) to (2) report the results using the number of administrative approvals in the next year as the dependent variable. Column (3) to (4) report the results using the number of positive media reports in the next year as the dependent variable. WeChat engagement is standardized, where one-standard-deviation change is associated with about 200 articles, 5 million readings, 2000 comments and 30,000 likes per year.

3. Standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A4: Summary Statistics for Coarsened Exact Matched Samples

	Treated (Has WeChat = 1)	Control (Has WeChat = 0)
Matched Variables	**Mean (Std. Dev.)**	**Mean (Std. Dev.)**
Age	3.89 (1.88)	3.22 (2.04)
Registration Capital	7.0e6 (2.4e7)	6.9e6 (2.5e7)
Number of Branches	1.07 (2.89)	0.87 (2.82)
Number of Subsidiaries	3.18 (3.66)	2.69 (3.69)
Number of Employees	73.93 (132.76)	43.87 (96.91)
Number of Trademarks	29.61 (49.09)	17.77 (40.27)
Number of Patents	11.06 (28.91)	5.74 (21.04)
Number of Certificates	1.15 (2.83)	0.56 (2.16)
Number of Work Copyrights	1.13 (17.27)	1.31 (23.04)
Number of Software Copyrights	0.24 (1.77)	0.12 (0.91)
Total Funding Rounds	1.10 (0.93)	0.77 (0.87)
Total Funding Amounts	1.5e7 (4.0e7)	8.2e6 (3.5e7)
Funded by Tencent	0.001 (0.042)	0.001 (0.042)
Industry Admin Competitiveness	0.73 (0.09)	0.73 (0.09)
Observations	3,902	3,956

Table A5: Social Media Effect on Institutions after Coarsened Exact Matching

Dependent Variable	Lead Administrative Approvals		Lead Positive Media Reports	
FE: Company ID	(1)	(2)	(3)	(4)
Has WeChat	0.136**	-0.100	2.137	15.98
	(0.055)	(0.330)	(1.975)	(11.88)
WeChat Engagement	0.504***	0.708***	16.40***	21.88***
	(0.080)	(0.096)	(2.868)	(3.471)
Has WeChat * Political Distance		0.227		-14.84
		(0.337)		(12.13)
WeChat Engagement * Political Distance		0.215***		6.041***
		(0.058)		(2.082)
Other Controls	Age, Number of Employee, Number of Trademark, Number of Patent, Number of Certificate, Number of Work Copyright, Number of Software Copyright, Total Funding Rounds, Total Funding Amounts, Funded by Tencent, Year Index			
Observations	6,613	6,613	6,613	6,613
R-squared	0.416	0.418	0.396	0.397

Notes: 1. Column (1) to (2) report the results using the number of administrative approvals in the next year as the dependent variable. Column (3) to (4) report the results using the number of positive media reports in the next year as the dependent variable. WeChat engagement is standardized, where one-standard-deviation change is associated with about 200 articles, 5 million readings, 2000 comments and 30,000 likes per year.

2. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Aggregated Group-Time Average Treatment Effect

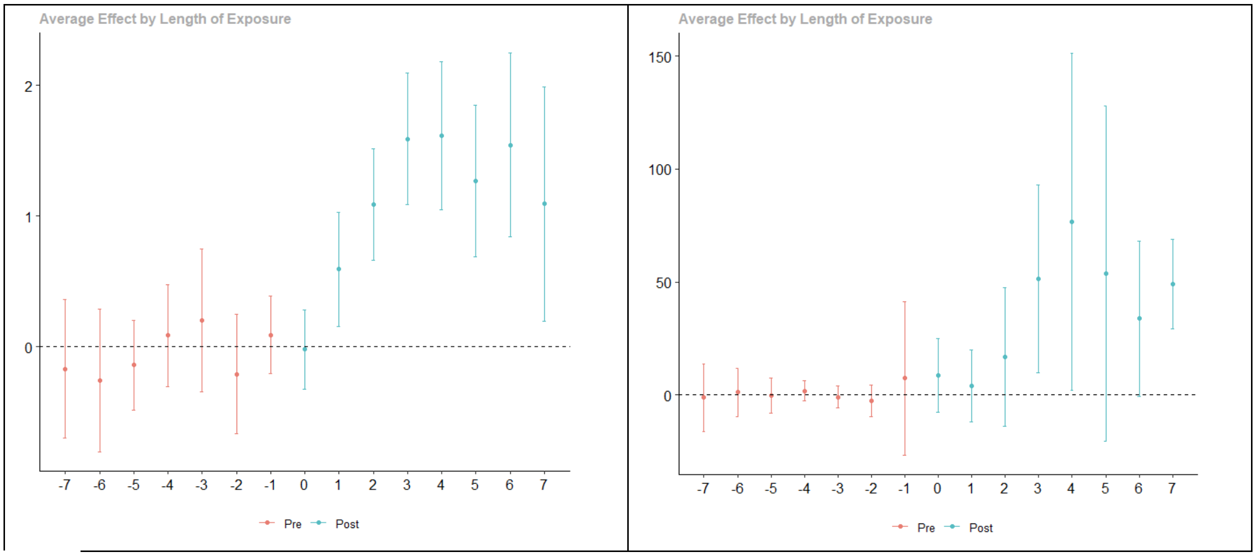
DV: Lead Administrative Approvals			DV: Lead Positive Media		
Group	**ATT**	**SE**	**Group**	**ATT**	**SE**
All Groups	0.9455*	0.1227	All Groups	28.5321*	10.0654
2012	0.9249*	0.2141	2012	42.2201*	16.5064
2013	1.4073*	0.2392	2013	37.8278*	12.9817
2014	1.0216*	0.2611	2014	34.1359*	15.3193
2015	0.8864	0.4367	2015	30.5010*	14.3229
2016	1.0454	0.5697	2016	38.6301*	18.4827
2017	0.4667	0.4069	2017	20.0943	22.1224
2018	0.6789	0.4614	2018	22.1568	19.7617
2019	0.3780	0.4842	2019	26.5854	21.9718

Notes:

1. Left: Average treatment effect of Has WeChat on Lead Administrative Approvals for all groups and subgroups in years. Right: Average treatment effect of Has WeChat on Lead Positive Media for all groups and subgroups in years.

2. Table created following the procedures in Callaway and Sant'Anna (2021) using *did* package in R. The * indicates the estimator is significant at 95% level.

Figure B1: Event Study based on Callaway and Sant'Anna (2021)

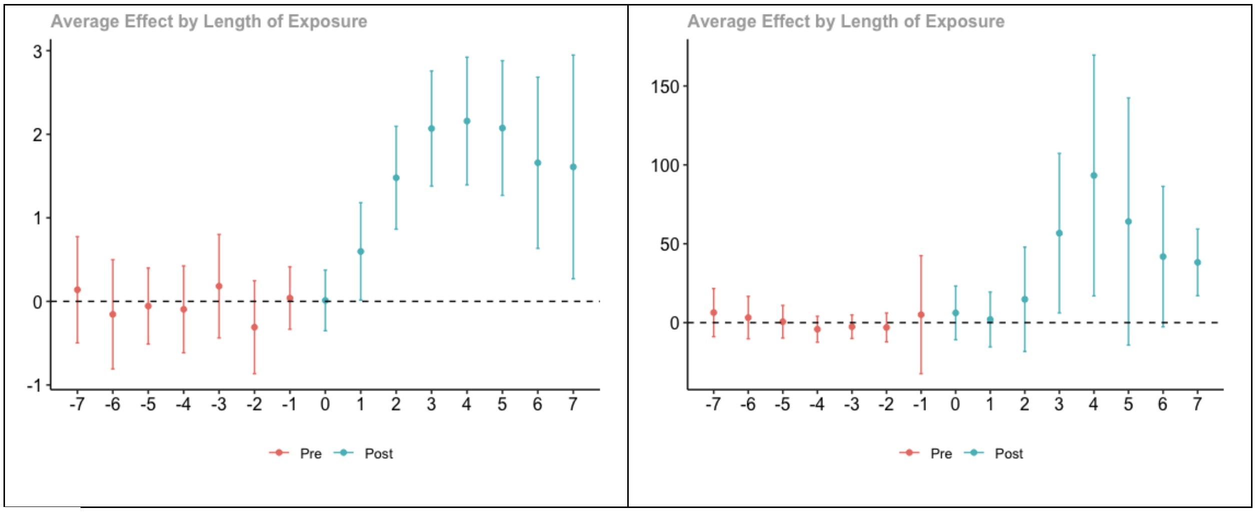


Notes:

1. Left: Coefficients of Has WeChat on the number of Administrative Approvals in years relative to adopting WeChat, with 95% confidence intervals. Right: Coefficients of Has WeChat on the number of Positive Media in years relative to adopting WeChat, with 95% confidence intervals.

2. Figure created following the procedures in Callaway and Sant'Anna (2021) using *did* package in R.

Figure B2: Event Study for Politically Disconnected Companies



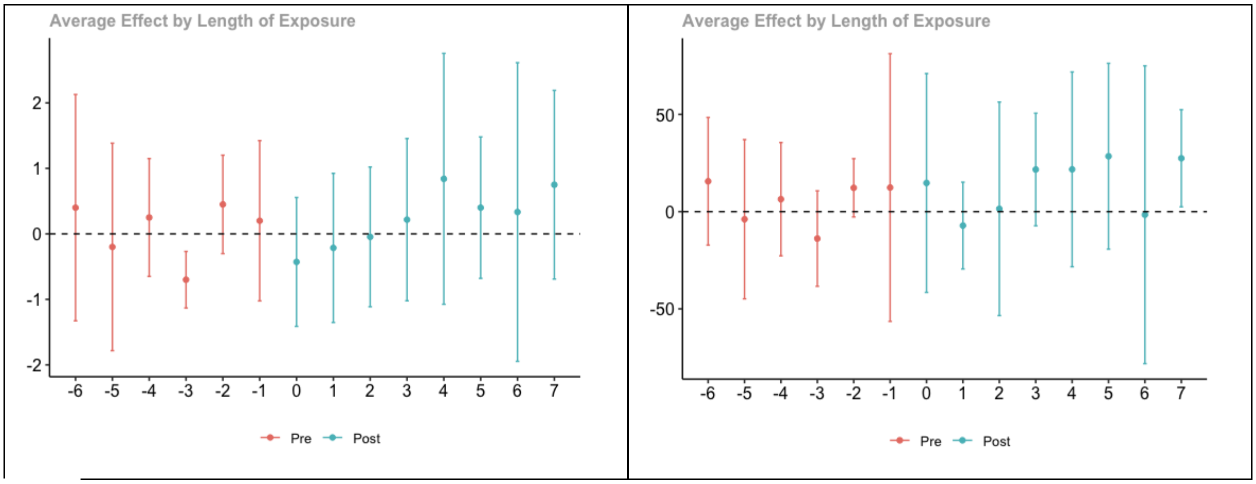
Notes:

1. Left: Coefficients of Has WeChat on the number of Administrative Approvals in years relative to adopting WeChat, with 95% confidence intervals. Right: Coefficients of Has WeChat on the number of Positive Media in years relative to adopting WeChat, with 95% confidence intervals.

2. The sample only includes companies that have 0 political connections.

3. Figure created following the procedures in Callaway and Sant'Anna (2021) using *did* package in R.

Figure B3: Event Study for Politically Connected Companies



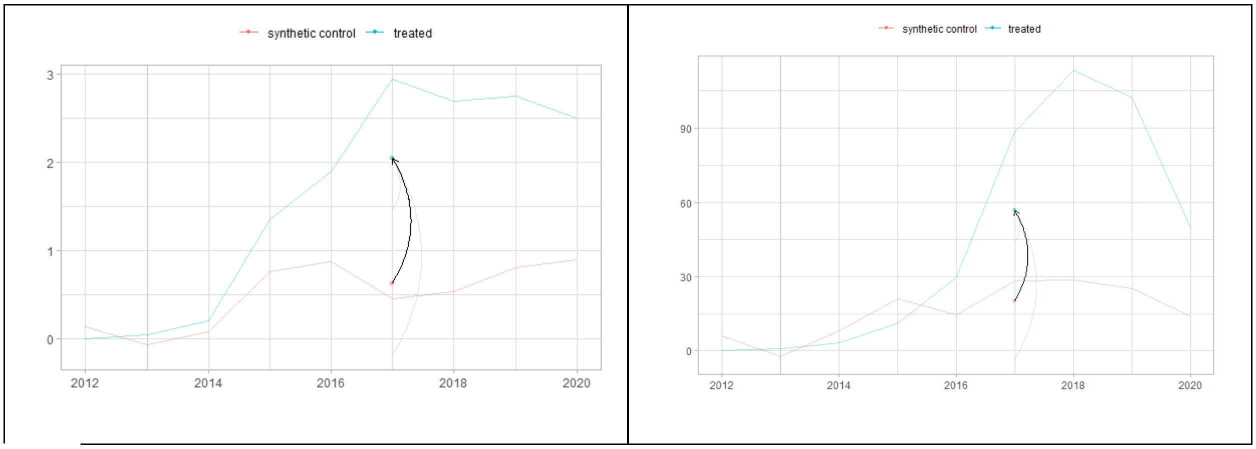
Notes:

1. Left: Coefficients of Has WeChat on the number of Administrative Approvals in years relative to adopting WeChat, with 95% confidence intervals. Right: Coefficients of Has WeChat on the number of Positive Media in years relative to adopting WeChat, with 95% confidence intervals.

2. The sample only includes companies that have at least 1 political connections.

3. Figure created following the procedures in Callaway and Sant'Anna (2021) using *did* package in R.

Figure B4: Event Study based on Synthetic Difference-in-Difference (Arkhangelsky et al. 2021)



Notes:

1. Left: Dependent variable is the number of administrative approvals in each year. Event study for companies who adopted WeChat in 2014 versus synthetic controlled companies who never adopted WeChat.

2. Right: Dependent variable is the number of positive media stories in each year. Event study for companies who adopted WeChat in 2014 versus synthetic controlled companies who never adopted WeChat.

3. Figure created following the procedures in Arkhangelsky (2021) using *synthdid* package in R.

APPENDIX C: APPENDIX FOR CHAPTER 4

C1. Alternative Measurement of Start-up Performance

One concern about the dependent variables (software copyright and product trademark) is that they are not representative of start-up's overall performances, because companies may create products that are not well received in the market. The overall performance of start-ups would be an ideal dependent variable but traditional metrics such as revenue and profit are not publicly available and verified by public accountants. To alleviate this concern, we replace the dependent variable using the log amount of funding in the next year in Table C1, as funding is one of the most important financial indicators for start-ups (Bardhan et al. 2013). We also use the Baidu search index as the dependent variable in Table C2. Baidu is the major search engine in China, and the Baidu Search Index reflects the absolute Baidu search volume of certain keywords weighted over total search volume in the same period of time. We treat such search volume as a proxy for online traffic to the companies, and literature has shown that online traffic data can be used to predict and verify company revenue (Froot et al. 2017, Huang 2018, Zhu 2019). The results are consistent in economic and statistical significance with our earlier results that use products.

Table C1: Alternative Performance Measure using Funding Amount

Company Industry	Software Companies					Physical Product Companies				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent Variables	AI Capability	Lead Log Funding Amount				AI Capability	Lead Log Funding Amount			
Model	FE	FE	FE	FE	FE+2SLS	FE	FE	FE	FE	FE+2SLS
AI Capability		0.405**		1.164**	1.823*		0.571***		0.414***	0.605
		(0.200)		(0.458)	(0.963)		(0.133)		(0.127)	(1.150)
LSM	0.078***		0.081	0.198	0.458	0.047***		0.135	0.167**	0.116
	(0.017)		(0.235)	(0.268)	(0.390)	(0.016)		(0.102)	(0.082)	(0.097)
AI * LSM				0.626***	**2.555*** **				0.471***	**0.497*** **
				(0.217)	**(0.841)**				(0.135)	**(0.135)**
Controls	Company Fixed Effect, Year Fixed Effect, Company Age, Total Number of Employees, Total Number of Job Postings, Total Number of Patents, Total Number of Certificates, Total Number of Work Copyrights, Total Number of Administrative Approvals									
Observations	2,490	2,490	2,490	2,490	2,454	6,464	6,464	6,464	6,464	6,396
R-squared	0.646	0.312	0.313	0.315		0.612	0.316	0.318	0.591	

Notes: 1. For Column (1) and (6), dependent variable is AI Capability in the current year. For Column (2)-(5) and (7)-(10), dependent variable is the log funding amount in the next year.
2. AI Capability is standardized, where one-standard-deviation change is associated with about 1 AI patent, 6 AI job postings, 30 AI social media articles and 8 AI news reports.
3. Column (5) and (10) report the 2SLS results with company fixed effect. 2 sets of instruments: (1) Number of AI policies issued by local governments in each year; (2) Number of news reports that a company in the same industry has adopted AI.
4. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table C2: Alternative Performance Measure using Search Index from Baidu

Company Industry	Software Companies					Physical Product Companies				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent Variables	AI Capability	Lead Search Index				AI Capability	Lead Search Index			
Model	FE	FE	FE	FE	FE+2SLS	FE	FE	FE	FE	FE+2SLS
AI Capability		2.423***		0.406	1.134**		0.897***		0.035*	0.068
		(0.339)		(0.546)	(0.498)		(0.016)		(0.019)	(0.110)
LSM	0.078***		1.585***	0.641	0.234	0.047***		0.023	-0.089	-0.153
	(0.017)		(0.552)	(0.546)	(0.404)	(0.016)		(0.020)	(0.098)	(0.112)
AI * LSM				1.673***	**1.752***				0.734***	**0.915***
				(0.196)	**(0.194)**				(0.031)	**(0.015)**
Controls	Company Fixed Effect, Year Fixed Effect, Company Age, Total Number of Employees, Total Number of Job Postings, Total Number of Patents, Total Number of Certificates, Total Number of Work Copyrights, Total Number of Administrative Approvals, Total Amount of Funding									
Observations	2,490	2,490	2,490	2,490	2,454	6,464	6,464	6,464	6,464	6,396
R-squared	0.646	0.481	0.470	0.501		0.612	0.316	0.624	0.697	

Notes: 1. For Column (1) and (6), dependent variable is AI Capability in the current year. For Column (2)-(5) and (7)-(10), dependent variable is the Baidu Search Index in the next year.
2. AI Capability is standardized, where one-standard-deviation change is associated with about 1 AI patent, 6 AI job postings, 30 AI social media articles and 8 AI news reports.
3. Column (5) and (10) report the 2SLS results with company fixed effect. 2 sets of instruments: (1) Number of AI policies issued by local governments in each year; (2) Number of news reports that a company in the same industry has adopted AI.
4. Robust standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

C2. Validation of LSM Level

We adopt two ways to verify our empirical measurement of LSM. First, we measure whether the start-up has received funding from government or state-owned funds. Prior literature has documented that start-ups receiving support from public institutions are more likely to produce and follow business plans (Honig and Karlsson 2004), because government funding usually requires a rigorous business plan and subsequent execution from the invested start-up. By contrast, start-ups funded by venture capital firms are more flexible to pivot and use the LSM. In Figure C1, we show the distribution of the LSM level of start-ups that have or have not received funding from state-owned funds. Consistent in theory, companies without state-owned funding have a significantly higher level of LSM in comparison to companies with state-owned funding.

We also verify our LSM measure using the concentration level of investor voting rights. Prior literature has documented that it is usually harder for start-ups to pivot when voting rights of the investors are dispersed among shareholders because it is harder to reach consensus at the top management team (Brinckmann et al. 2010). We calculate the concentration voting rights in each year for a firm using the Herfindahl-Hirschman Index (HHI) based on the proportion of investment each investor has made on the firm. A higher HHI suggests that a few people hold the majority of voting rights in the firm. In Figure C2, we show the distribution of the LSM level of start-ups that have an above-median HHI versus the start-ups that have a below-median HHI. Companies with more concentrated voting rights have a significantly higher level of LSM than firms with more dispersed voting rights. These results help validate our empirical measurement of LSM level.

Figure C1: LSM Level for Start-ups with or without Government Funding

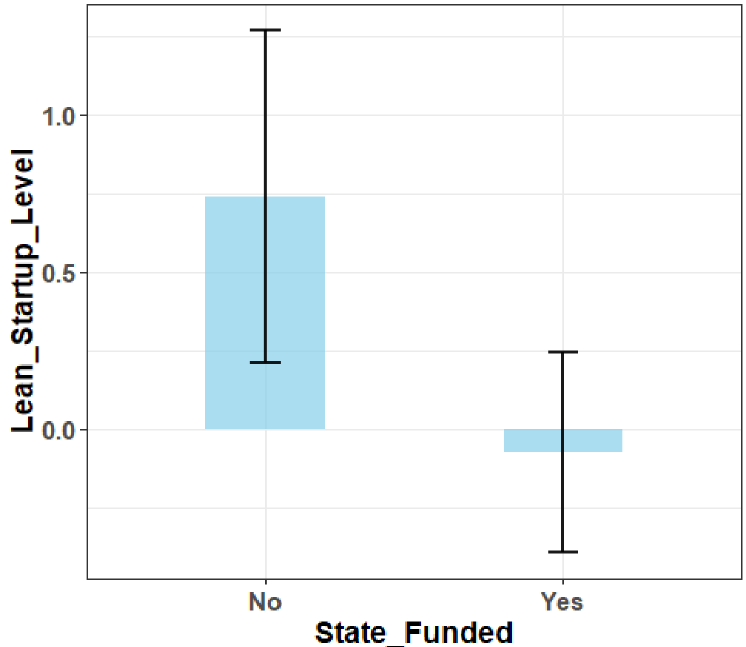
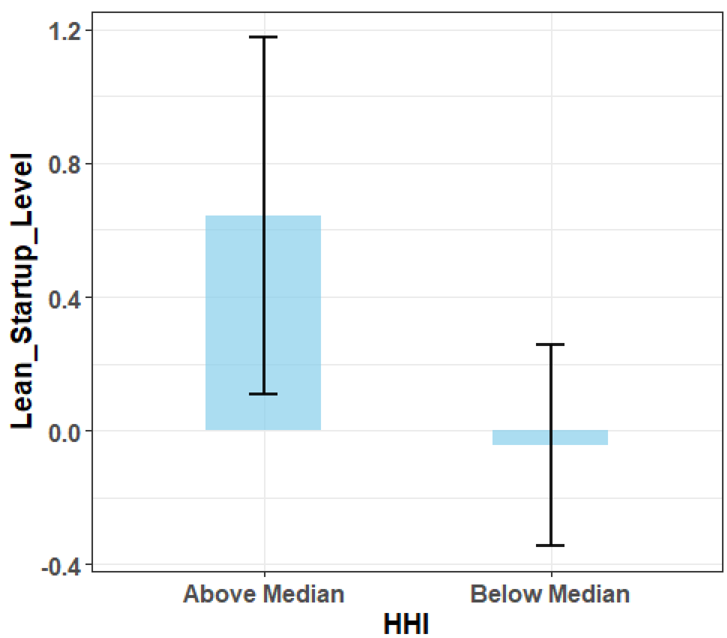


Figure C2: LSM Level for Start-ups above or below the Median of Share HHI



C3. Reduced-Form Instrumental Variable Test

A reasonable concern about the instrumental variable is that the AI policies published by local governments may be correlated with regional socio-economic levels and thereby correlated with unobserved firm qualities. The number of AI-related news about companies in the same industry may also be correlated with unobserved market-level dynamics and firm qualities. If that is the case, then we should also observe that instrumental variables are associated with start-up innovation activities even when the focal start-up has not adopted AI capability. We show the reduced-form instrumental variable test in Table C3 following the procedures in Martin and Yurukoglu (2017). Specifically, we create a binary variable of AI Adoption by identifying the earliest year when the company has any AI-related patent, job posting, social media article or news report. We then separate our sample into AI adopters versus non-adopters use the instrumental variables as the independent variable directly in the regression. The results show that the instruments are only significantly associated with start-up innovation activities when start-ups have adopted AI, suggesting that our instrumental variable only correlates the dependent variable through AI adoption.

Table C3: Instrumental Variable Test

AI Adopters or Not	AI Adopters		Non-Adopters	
Company Industries	Software Companies	Physical Product Companies	Software Companies	Physical Product Companies
Dependent Variable	Lead # Software Copyright	Lead # Product Trademark	Lead # Software Copyright	Lead # Product Trademark
	(1)	(2)	(3)	(4)
Provincial AI Policies	0.025***	0.076***	-0.003	0.015
	(0.005)	(0.026)	(0.007)	(0.011)
AI News from Companies in the Same Industry	0.011**	0.031*	0.015	0.006
	(0.006)	(0.017)	(0.010)	(0.006)
Controls	Company Fixed Effect, Year Fixed Effect, Company Age, Total Number of Employees, Total Number of Job Postings, Total Number of Patents, Total Number of Certificates, Total Number of Work Copyrights, Total Number of Administrative Approvals, Total Amount of Funding			
Observations	1,262	2,014	1,228	4,450
R-squared	0.189	0.254	0.178	0.254

Notes:

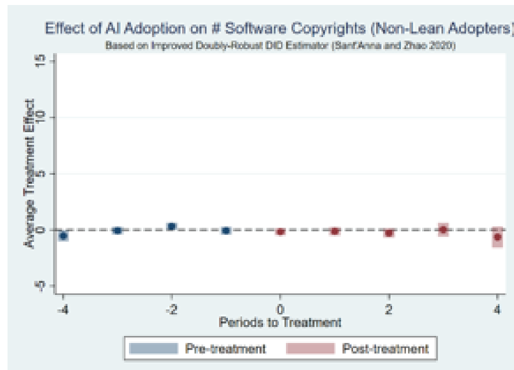
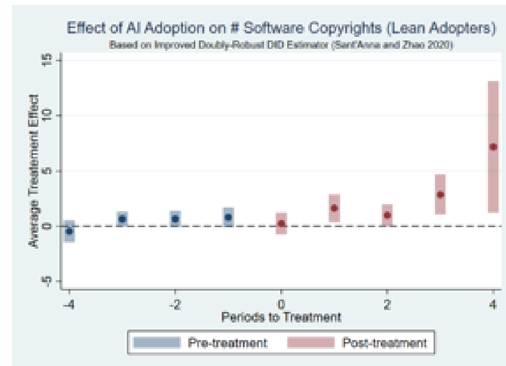
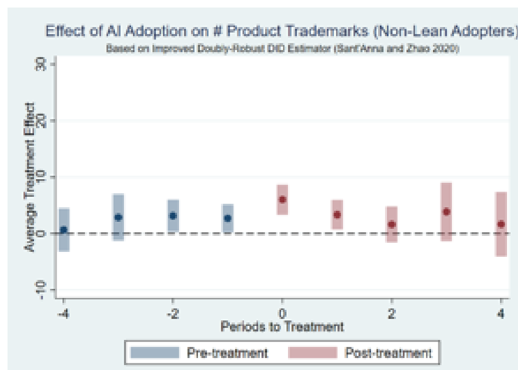
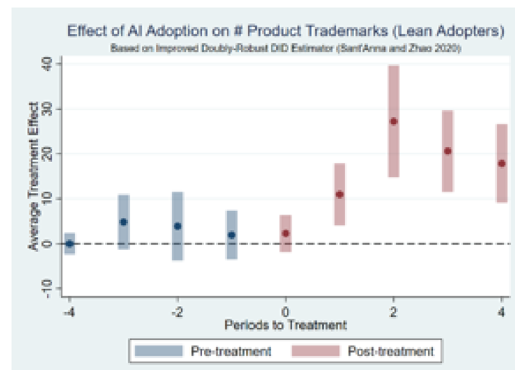
1. Dependent variable is the number of new products in the next year.
2. Independent variable is the number of AI-related policies published by local Chinese governments in each year.
3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

C4. Doubly Robust DiD

Due to recent development in difference-in-difference estimators (Goodman-Bacon 2021), we apply the doubly robust difference-in-difference estimator (Callaway and Sant'Anna 2021, Sant'Anna and Zhao 2020) to verify the regression results. As the estimator can only be applied on binary treatment effects, we use the binary variable of AI adoption as independent variable. Also, to verify the complementarity results, we divide the sample start-ups into LSM adopter and non-adopters based on whether the company has ever published an LSM-related job posting, social media article or new report. Each group has at least 35% of total observations so that the statistical power is enough for identification. The results are shown in Figure C3.

The two plots at the top of Figure C3 show the effect of AI adoption on software start-ups; the left is for firms that did not adopt LSM, and the left is for LSM adopters. They show that AI adoption has limited effect on non-adopters but has significant effect on LSM adopters. There is a one-year lag between adoption and performance, and it is likely due to that it takes time to develop a new product. These results not only verify the positive impact of AI adoption on start-up product innovations, but also support the complementarities between AI capability and LSM. The bottom two plots are for physical product start-ups, and the results are similar to that for software products.

Figure C3: Event Study: Effect of AI Adoption on Product Innovation based on Callaway and Sant'Anna (2021) and Sant'Anna and Zhao (2020)

	Non LSM Adopters	LSM Adopters
Software Start-ups (DV: Number of Software Copyrights)		
Physical Product Start-ups (DV: Number of Product Trademarks)		

Notes: Event study plots created using *csdid* package in STATA. Top 2 plots show AI adoption effect on software start-ups. Bottom 2 plots show AI adoption effect on physical product start-ups. Left 2 plots are companies that never adopted LSM. Right 2 plots are companies that have at least one job posting, social media article or news report that is lean-related.

C5. Cross Complementarity Tests

To further verify the mechanism, in this section we empirically test the complementarity between validated learning and internally focused AI, as well the complementarity between iterative experimentation and externally focused AI. We would expect that there should not be complementarities if our hypothesis is correct that complementarities should happen between extern AI and validated learning and between internal AI and iterative experimentation. The results are shown in Table C4.

The interaction terms are not statistically significant, showing that different AI capabilities are complements to different types of lean start-up method. This is likely due to that context knowledge needed for external AI is different from that for internal AI, and that the differences in R&D objectives for developing novel is different from that for incremental products. The context knowledge required for externally focused AI is marketing oriented, such as product positioning and customer segmentations, while internally focused AI requires specialized context knowledge in internal product development processes. These results suggest that start-ups should deploy AI capabilities according to the objectives of product innovations.

They can also help to rule out possible concerns about omitted variable biases such as "tech-savvy start-ups are better at innovations and are also more likely to adopt AI and LSM", because if that is the case, we should also observe an effect for the cross complementarities, not just in the specific complementarities we show.

Table C4: Complementarity Tests for Novel Products vs. Incremental Products

Company Industry	Software Companies				Physical Product Companies			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variables	Lead # Novel Software Copyright		Lead # Incremental Software Copyright		Lead # Novel Product Trademark		Lead # Incremental Product Trademark	
Model	FE	FE+2SLS	FE	FE+2SLS	FE	FE+2SLS	FE	FE+2SLS
Validated Learning	0.39	0.48**	0.84	0.57	1.45*	3.35***	2.02	0.94
	(0.24)	(0.23)	(0.71)	(0.98)	(0.79)	(0.69)	(1.42)	(2.54)
Internal AI	0.26	0.37	0.31***	0.78	0.19	0.22	1.42	3.86**
	(0.36)	(0.44)	(0.10)	(0.81)	(1.32)	(0.14)	(1.93)	(1.59)
Validated Learning * Internal AI	-0.07	**0.08**	0.96	**-0.23**	-4.53	**-1.51**	3.33	**2.69**
	(0.16)	**(0.36)**	(0.74)	**(0.15)**	(3.99)	**(1.16)**	(2.51)	**(2.09)**
Iterative Experimentation	-0.04	0.32	0.77***	0.68***	0.19	0.24	7.06***	6.78***
	(0.16)	(0.32)	(0.23)	(0.10)	(0.31)	(0.88)	(2.14)	(1.24)
External AI	2.03**	0.41***	0.08	-0.65	2.35*	2.76**	0.88	0.80
	(0.80)	(0.16)	(0.05)	(0.68)	(1.41)	(1.31)	(2.45)	(2.37)
Iterative Experimentation * External AI	0.26	**0.25**	0.46	**0.88**	0.71	**0.35**	2.45	**0.69**
	(0.48)	**(0.66)**	(0.80)	**(0.81)**	(1.37)	**(0.65)**	(2.36)	**(2.29)**
Controls	Company Fixed Effect, Year Fixed Effect, Company Age, Total Number of Employees, Total Number of Job Postings, Total Number of Patents, Total Number of Certificates, Total Number of Work Copyrights, Total Number of Administrative Approvals, Total Amount of Funding							
Observations	2,490	2,454	2,490	2,454	6,464	6,396	6,464	6,396
R-squared	0.436		0.899		0.534		0.578	

Notes: 1. For Column (1)-(2) and (5)-(6), dependent variable is the number of novel products in the next year. For Column (3)-(4) and (7)-(8), dependent variable is the number of incremental products in the next year.

2. Column (2), (4), (6) and (8) report the 2SLS results with both AI Capability and LSM instrumented.

3. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

BIBLIOGRAPHY

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. American economic review, 93(1), 113-132. Agrawal, A., Gans, J., & Goldfarb, A. (2018). Prediction machines: the simple economics of artificial intelligence. Harvard Business Press.

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. Journal of the American statistical Association, 105(490), 493-505.

Aggarwal, R., Gopal, R., Gupta, A., & Singh, H. (2012). Putting money where the mouths are: The relation between venture financing and electronic word-of-mouth. Information Systems Research, 23(3-part-2), 976-992.

Agrawal, A.K., Catalini, C. and Goldfarb, A., (2011). The geography of crowdfunding (No. w16820). National bureau of economic research.

Ahl, H., (2006). Why research on women entrepreneurs needs new directions. Entrepreneurship theory and practice, 30(5), pp.595-621.

Ahlstrom, D., Yang, X., Wang, L., & Wu, C. (2018). A global perspective of entrepreneurship and innovation in China. Multinational Business Review.

Aldrich, H. E., & Baker, T. (2001). Learning and legitimacy: Entrepreneurial responses to constraints on the emergence of new populations and organizations. The entrepreneurship dynamic: Origins of entrepreneurship and the evolution of industries, 1, 207-35.

Aldrich, H. E., & Fiol, C. M. (1994). Fools rush in? The institutional context of industry creation. Academy of management review, 19(4), 645-670.

Aldrich, H. E., & Yang, T. (2012). Lost in translation: Cultural codes are not blueprints. Strategic Entrepreneurship Journal, 6(1), 1-17.

Aldrich, H.E. and Fiol, C.M., (1994). Fools rush in? The institutional context of industry creation. Academy of management review, 19(4), pp.645-670.

Allon, G., Askalidis, G., Berry, R., Immorlica, N., Moon, K., & Singh, A. (2021). When to be agile: Ratings and version updates in mobile apps. Management Science.

Alterman, J. B. (2011). The revolution will not be tweeted. The Washington Quarterly, 34(4), 103-116.

Amit, R., Glosten, L. and Muller, E., (1990). Entrepreneurial ability, venture investments, and risk sharing. Management science, 36(10), pp.1233-1246.

Amore, M. D., & Bennedsen, M. (2013). The value of local political connections in a low-corruption environment. Journal of Financial Economics, 110(2), 387-402.

Anand, J., Mulotte, L., & Ren, C. R. (2016). Does experience imply learning?. Strategic Management Journal, 37(7), 1395-1412.

Angst, C.M., Agarwal, R., Sambamurthy, V. and Kelley, K., (2010). Social contagion and information technology diffusion: the adoption of electronic medical records in US hospitals. Management Science, 56(8), pp.1219-1241.

Angulo-Ruiz, F., Pergelova, A., & Wei, W. X. (2019). How does home government influence the internationalization of emerging market firms? The mediating role of strategic intents to internationalize. International Journal of Emerging Markets.

Aral, S. and Van Alstyne, M. (2011). The diversity-bandwidth trade-off. American journal of sociology, 117(1), 90-171.

Aral, S. and Walker, D., (2011). Creating social contagion through viral product design: A randomized trial of peer influence in networks. Management Science, 57(9), pp.1623-1639.

Aral, S., & Weill, P. (2007). IT assets, organizational capabilities, and firm performance: How resource allocations and organizational differences explain performance variation. Organization science, 18(5), 763-780.

Aral, S., Brynjolfsson, E., & Wu, L. (2012). Three-way complementarities: Performance pay, human resource analytics, and information technology. Management Science, 58(5), 913-931.

Aral, S., Dellarocas, C., & Godes, D. (2013). Introduction to the special issue—social media and business transformation: a framework for research. Information Systems Research, 24(1), 3-13.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The review of economic studies, 58(2), 277-297.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. American Economic Review, 111(12), 4088-4118.

Armanios, D. E., Eesley, C. E., Li, J., & Eisenhardt, K. M. (2017). How entrepreneurs leverage institutional intermediaries in emerging economies to acquire public resources. Strategic Management Journal, 38(7), 1373-1390.

Arora, A. (1996). Testing for complementarities in reduced-form regressions: A note. Economics letters, 50(1), 51-55.

Arora, A., & Gambardella, A. (1990). Complementarity and external linkages: the strategies of the large firms in biotechnology. The journal of industrial economics, 361-379.

Athey, S., & Stern, S. (1998). An empirical framework for testing theories about complimentarity in organizational design.

Babar, Y., & Burtch, G. (2020). Examining the Heterogeneous Impact of Ride-Hailing Services on Public Transit Use. Information Systems Research.

Babina, T., Fedyk, A., He, A. X., & Hodson, J. (2020). Artificial intelligence, firm growth, and industry concentration.

Baker, T., & Nelson, R. E. (2005). Creating something from nothing: Resource construction through entrepreneurial bricolage. Administrative science quarterly, 50(3), 329-366.

Balachandra, L., Briggs, T., Eddleston, K. and Brush, C., (2019). Don't Pitch Like a Girl!: How Gender Stereotypes Influence Investor Decisions. Entrepreneurship Theory and Practice, 43(1), pp.116-137.

Baltagi, B.H. and Rich, D.P., (2005). Skill-biased technical change in US manufacturing: a general index approach. Journal of Econometrics, 126(2), pp.549-570.

Bardhan, I., Krishnan, V., & Lin, S. (2013). Research note—business value of information technology: testing the interaction effect of IT and R&D on Tobin's Q. Information Systems Research, 24(4), 1147-1161.

Baron, J.N., Burton, M.D. and Hannan, M.T., (1996). The road taken: Origins and evolution of employment systems in emerging companies. Industrial and Corporate Change, 5(2), pp.239-275.

Behaghel, L. and Greenan, N., (2010). Training and age-biased technical change. Annals of Economics and Statistics/Annales d'Économie et de Statistique, pp.317-342.

Behaghel, L., Caroli, E. and Roger, M., (2014). Age-biased technical and organizational change, training and employment prospects of older workers. Economica, 81(322), pp.368-389.

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. MIS Q, 45(3), 1433-1450.

Berman, E., Bound, J., & Machin, S. (1998). Implications of skill-biased technological change: international evidence. The Quarterly Journal of Economics, 113(4), 1245-1279.

Bertrand, M., Kramarz, F., Schoar, A., & Thesmar, D. (2018). The cost of political connections. Review of Finance, 22(3), 849-876.

Bhamu, J., & Sangwan, K. S. (2014). Lean manufacturing: literature review and research issues. International Journal of Operations & Production Management.

Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. (2013). Digital business strategy: toward a next generation of insights. MIS quarterly, 471-482.

Bhaskaran, S., Erzurumlu, S. S., & Ramachandran, K. (2021). Sequential product development and introduction by cash-constrained start-ups. Manufacturing & Service Operations Management, 23(6), 1505-1523.

Bhaskaran, V., Mirchandani, R., & Shah, M. (2013). Creating a Measurable Social Media Marketing Strategy: Increasing the Value and ROI of Intangibles and Tangibles for Hokey Pokey. Marketing Science, 32(2), 194-212.

Biemans, W. G. (2018). Managing innovation within networks. Routledge.

Bigsby, K. G., Ohlmann, J. W., & Zhao, K. (2017). Online and off the field: Predicting school choice in college football recruiting from social media data. Decision analysis, 14(4), 261-273.

Bingham, C. B., & Eisenhardt, K. M. (2011). Rational heuristics: the ‘simple rules’ that strategists learn from process experience. Strategic management journal, 32(13), 1437-1464.

Bird, B. and Brush, C., (2002). A gendered perspective on organizational creation. Entrepreneurship theory and practice, 26(3), pp.41-65.

Blackwell, M., Iacus, S., King, G. and Porro, G., (2009). cem: Coarsened exact matching in Stata. The Stata Journal, 9(4), pp.524-546.

Blank, S. (2013). Why the lean start-up changes everything. Harvard business review, 91(5), 63-72.

Blank, S. A. (2005). The four steps to the epiphany: Successful Strategies for Products that Win CafePress. com.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of econometrics, 87(1), 115-143.

Bond, S. D., He, S. X., & Wen, W. (2019). Speaking for "free": Word of mouth in free-and paid-product settings. Journal of Marketing Research, 56(2), 276-290.

Bortolini, R. F., Cortimiglia, M. N., Danilevicz, A. D. M. F., & Ghezzi, A. (2018). Lean Start-up: a comprehensive historical review. Management Decision.

Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies 'Engines of growth'?. Journal of econometrics, 65(1), 83-108.

Bresnahan, T. F., Brynjolfsson, E., & Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. The quarterly journal of economics, 117(1), 339-376.

Breznitz, D., Forman, C., & Wen, W. (2018). The role of venture capital in the formation of a new technological ecosystem: Evidence from the cloud. MIS Quarterly, 42(4).

Brinckmann, J., Grichnik, D., & Kapsa, D. (2010). Should entrepreneurs plan or just storm the castle? A meta-analysis on contextual factors impacting the business planning–performance relationship in small firms. Journal of business Venturing, 25(1), 24-40.

Brooks, A.W., Huang, L., Kearney, S.W. and Murray, F.E., (2014). Investors prefer entrepreneurial ventures pitched by attractive men. Proceedings of the National Academy of Sciences, 111(12), pp.4427-4431.

Brouwer, J. (2020). Government Partners with Private Corporations to Monitor China's Internet. China Digital Times. Retrieved December 22, 2020, https://chinadigitaltimes.net/2020/12/government-partners-with-private-corporations-to-monitor-chinas-internet/

Brown, S. L., & Eisenhardt, K. M. (1997). The art of continuous change: Linking complexity theory and time-paced evolution in relentlessly shifting organizations. Administrative science quarterly, 1-34.

Bruni, A., Gherardi, S. and Poggio, B., (2004). Doing gender, doing entrepreneurship: An ethnographic account of intertwined practices. Gender, Work & Organization, 11(4), pp.406-429.

Brush, C., Carter, N.M., Gatewood, E.J., Greene, P.G. and Hart, M., (2001). The Diana Project: Women business owners and equity capital: The myths dispelled. Babson College Center for Entrepreneurship Research Paper, (2009-11).

Brush, C., Greene, P., Balachandra, L., Davis, A. and Blank, A.M., (2014). Women entrepreneurs 2014: bridging the gender gap in venture capital. Arthur M. Blank Center for Entrepreneurship Babson College.

Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies. WW Norton & Company.

Brynjolfsson, E., & Milgrom, P. (2013). Complementarity in organizations. The handbook of organizational economics, 11-55.

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. Science, 358(6370), 1530-1534.

Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The productivity J-curve: How intangibles complement general purpose technologies. American Economic Journal: Macroeconomics, 13(1), 333-72.

Buchanan, J. M. (1987). The constitution of economic policy. Science, 236(4807), 1433-1436.

Burt, R. S., Kilduff, M., & Tasselli, S. (2013). Social network analysis: Foundations and frontiers on advantage. Annual review of psychology, 64, 527-547.

Burt, R.S., (2009). Structural holes: The social structure of competition. Harvard university press.

Burtch, G., Ghose, A., & Wattal, S. (2016). Secret admirers: An empirical examination of information hiding and contribution dynamics in online crowdfunding. Information Systems Research, 27(3), 478-496.

Bygrave, W.D., (1988). The structure of the investment networks of venture capital firms. Journal of Business Venturing, 3(2), pp.137-157.

Cahan, S. F., Chen, C., Chen, L., & Nguyen, N. H. (2015). Corporate social responsibility and media coverage. Journal of Banking & Finance, 59, 409-422.

Calantone, R. J., Cavusgil, S. T., & Zhao, Y. (2002). Learning orientation, firm innovation capability, and firm performance. Industrial marketing management, 31(6), 515-524.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. Journal of Econometrics, 225(2), 200-230.

Campa, J.M. and Kedia, S., (2002). Explaining the diversification discount. The journal of finance, 57(4), pp.1731-1762.

Camuffo, A., Cordova, A., Gambardella, A., & Spina, C. (2020). A scientific approach to entrepreneurial decision making: Evidence from a randomized control trial. Management Science, 66(2), 564-586.

Castelvecchi, D. (2016). Can we open the black box of AI?. Nature News, 538(7623), 20.

Cejka, M.A. and Eagly, A.H., (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. Personality and social psychology bulletin, 25(4), pp.413-423.

Chandler, A. D., Jr. (1990). Scale and scope: The dynamics of industrial capitalism. Cambridge: Harvard University Press/Belknap.

Chen, J., & Wang, L. (2021). Entrepreneurship and Innovation of Small and Medium-Sized Enterprises in China. The Oxford Handbook of China Innovation, 156.

Chen, Y., Li, H., & Zhou, L. A. (2005). Relative performance evaluation and the turnover of provincial leaders in China. Economics Letters, 88(3), 421-425.

Choudhary, A., Hendrix, W., Lee, K., Palsetia, D., & Liao, W. K. (2012). Social media evolution of the Egyptian revolution. Communications of the ACM, 55(5), 74-80.

Christensen, C. M. (2013). The innovator's dilemma: when new technologies cause great firms to fail. Harvard Business Review Press.

Chung, S., Animesh, A., Han, K. and Pinsonneault, A., (2015). Firm's Social Media Efforts, Consumer Behavior, and Firm Performance: Evidence from Facebook.

Cohen, L., Coval, J., & Malloy, C. (2011). Do powerful politicians cause corporate downsizing?. Journal of Political Economy, 119(6), 1015-1060.

Cooper, R. G., & Edgett, S. J. Kleinschimidt (2002) Optimizing the Stage-Gate Process. What Best Practice Companies are doing. Research-Technology Management, 45(5), 21-27.

Cui, A. S., & Wu, F. (2017). The impact of customer involvement on new product development: Contingent and substitutive effects. Journal of Product Innovation Management, 34(1), 60-80.

Cui, L., Hu, H. W., Li, S., & Meyer, K. E. (2018). Corporate political connections in global strategy. Global Strategy Journal, 8(3), 379-398.

Cui, R., Ding, H., & Zhu, F. (2022). Gender inequality in research productivity during the COVID-19 pandemic. Manufacturing & Service Operations Management, 24(2), 707-726.

Cui, R., Gallino, S., Moreno, A., & Zhang, D. J. (2018). The operational value of social media information. Production and Operations Management, 27(10), 1749-1769.

Cui, R., Gallino, S., Moreno, A., & Zhang, D. J. (2018). The operational value of social media information. Production and Operations Management, 27(10), 1749-1769.

Cumming, D., (2006). Adverse selection and capital structure: Evidence from venture capital. Entrepreneurship Theory and Practice, 30(2), pp.155-183.

Cumming, D., Walz, U., & Werth, J. C. (2016). Entrepreneurial spawning: experience, education, and exit. Financial Review, 51(4), 507-525.

Dalle, J. M., Den Besten, M., & Menon, C. (2017). Using Crunchbase for economic and managerial research.

Davila, A., Foster, G. and Gupta, M., (2003). Venture capital financing and the growth of start-up firms. Journal of business venturing, 18(6), pp.689-708.

Davis, J. P., & Eisenhardt, K. M. (2011). Rotating leadership and collaborative innovation: Recombination processes in symbiotic relationships. Administrative Science Quarterly, 56(2), 159-201.

Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2009). Optimal structure, market dynamism, and the strategy of simple rules. Administrative Science Quarterly, 54(3), 413-452.

Dechow, P. M., & Tan, S. T. (2021). How do accounting practices spread? An examination of law firm networks and stock option backdating. The Accounting Review, 96(1), 431-464.

Deng, X., Kang, J. K., & Low, B. S. (2013). Corporate social responsibility and stakeholder value maximization: Evidence from mergers. Journal of financial Economics, 110(1), 87-109.

DeSantola, A., & Gulati, R. (2017). Scaling: Organizing and growth in entrepreneurial ventures. Academy of Management Annals, 11(2), 640-668.

Dessein, W., (2005). Information and control in ventures and alliances. The Journal of Finance, 60(5), pp.2513-2549.

Dieleman, M., & Boddewyn, J. J. (2012). Using organization structure to buffer political ties in emerging markets: A case study. Organization Studies, 33(1), 71-95.

DiStaso, M. W., McCorkindale, T., & Wright, D. K. (2011). How public relations executives perceive and measure the impact of social media in their organizations. Public relations review, 37(3), 325-328.

Dixon, J., Hong, B., & Wu, L. (2021). The robot revolution: Managerial and employment consequences for firms. Management Science, 67(9), 5586-5605.

Dobbin, F., & Dowd, T. J. (1997). How policy shapes competition: Early railroad foundings in Massachusetts. Administrative science quarterly, 501-529.

Dorobantu, S., Kaul, A., & Zelner, B. (2017). Nonmarket strategy research through the lens of new institutional economics: An integrative review and future directions. Strategic Management Journal, 38(1), 114-140.

Eberhart, R. N., Eesley, C. E., & Eisenhardt, K. M. (2017). Failure is an option: Institutional change, entrepreneurial risk, and new firm growth. Organization Science, 28(1), 93-112.

Eddleston, K.A., Ladge, J.J., Mitteness, C. and Balachandra, L., (2016). Do you see what I see? Signaling effects of gender and firm characteristics on financing entrepreneurial ventures. Entrepreneurship Theory and Practice, 40(3), pp.489-514.

Eddleston, K.A., Ladge, J.J., Mitteness, C. and Balachandra, L., (2016). Do you see what I see? Signaling effects of gender and firm characteristics on financing entrepreneurial ventures. Entrepreneurship Theory and Practice, 40(3), pp.489-514.

Eesley, C. (2016). Institutional barriers to growth: Entrepreneurship, human capital and institutional change. Organization Science, 27(5), 1290-1306.

Eesley, C., & Wu, L., (2020). For start-ups, adaptability and mentor network diversity can be pivotal: evidence from a randomized experiment on a mooc platform. MIS Quarterly, 44(2), 661-697.

Eesley, C. E., & Wu, L., (2019). For Start-ups, Adaptability and Mentor Network Diversity can be Pivotal: Evidence from a Randomized Experiment on a MOOC Platform. MIS Quarterly, Forthcoming.

Eesley, C. E., Eberhart, R. N., Skousen, B. R., & Cheng, J. L. (2018). Institutions and entrepreneurial activity: The interactive influence of misaligned formal and informal institutions. Strategy Science, 3(2), 393-407.

Eesley, C., Li, J. B., & Yang, D. (2016). Does institutional change in universities influence high-tech entrepreneurship? Evidence from China's Project 985. Organization Science, 27(2), 446-461.

Eisenhardt, K. M., Furr, N. R., & Bingham, C. B. (2010). CROSSROADS—Microfoundations of performance: Balancing efficiency and flexibility in dynamic environments. Organization science, 21(6), 1263-1273.

Ewens, M., Nanda, R., & Rhodes-Kropf, M. (2018). Cost of experimentation and the evolution of venture capital. Journal of Financial Economics, 128(3), 422-442.

Faccio, M., Masulis, R. W., & McConnell, J. J. (2006). Political connections and corporate bailouts. The journal of Finance, 61(6), 2597-2635.

Feinberg, S. E., & Gupta, A. K. (2009). MNC subsidiaries and country risk: Internalization as a safeguard against weak external institutions. Academy of Management journal, 52(2), 381-399.

Felin, T., & Foss, N. J. (2011). The endogenous origins of experience, routines, and organizational capabilities: the poverty of stimulus. Journal of Institutional Economics, 7(2), 231-256.

Felin, T., Gambardella, A., Stern, S., & Zenger, T. (2020). Lean start-up and the business model: Experimentation revisited. Long Range Planning 53(2020), 101889.

Feng, L., Hu, Y., Li, B., Stanley, H. E., Havlin, S., & Braunstein, L. A. (2015). Competing for attention in social media under information overload conditions. PloS one, 10(7), e0126090.

Flamholtz, E. G., & Randle, Y. (2015). Growing pains: Building sustainably successful organizations (5th ed.). San Francisco, CA: Jossey-Bass.

Fleming, N. (2018). How artificial intelligence is changing drug discovery. Nature, 557(7706), S55-S55.

Forman, C. (2005). The corporate digital divide: Determinants of Internet adoption. Management Science, 51(4), 641-654.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. Information systems research, 19(3), 291-313.

Forman, C., Goldfarb, A. and Greenstein, S., (2002). Digital dispersion: An industrial and geographic census of commercial internet use (No. w9287). National Bureau of Economic Research.

Forman, C., Goldfarb, A. and Greenstein, S., (2005). Geographic location and the diffusion of Internet technology. Electronic Commerce Research and Applications, 4(1), pp.1-13.

Forret, M.L. and Dougherty, T.W., (2004). Networking behaviors and career outcomes: differences for men and women?. Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior, 25(3), pp.419-437.

Froot, K., Kang, N., Ozik, G., & Sadka, R. (2017). What do measures of real-time corporate sales say about earnings surprises and post-announcement returns?. Journal of Financial Economics, 125(1), 143-162.

Frynas, J. G., Mellahi, K., & Pigman, G. A. (2006). First mover advantages in international business and firm‐specific political resources. Strategic Management Journal, 27(4), 321-345.

Furr, N. R., Cavarretta, F., & Garg, S. (2012). Who changes course? The role of domain knowledge and novel framing in making technology changes. Strategic Entrepreneurship Journal, 6(3), 236-256.

Gans, J., Hsu, D.H., & Stern, S. (2002) When does start-up innovation spur the gale of creative destruction? Rand J. Econom. 33(4):571–586.

Gao, J., & Mu, R. (2021). Mass Entrepreneurship and Mass Innovation in China. The Oxford Handbook of China Innovation, 254.

Gatignon, H., & Robertson, T. S. (1989). Technology diffusion: an empirical test of competitive effects. Journal of marketing, 53(1), 35-49.

Ghannam, J. (2011). Social Media in the Arab World: Leading up to the Uprisings of 2011. Center for international media assistance, 3(1), 1-44.

Ghose, A., Ipeirotis, P. G., & Li, B. (2019). Modeling consumer footprints on search engines: An interplay with social media. Management Science, 65(3), 1363-1385.

Gibson, D.E. and Lawrence, B.S., (2010). Women's and men's career referents: How gender composition and comparison level shape career expectations. Organization Science, 21(6), pp.1159-1175.

Gilley, B. (2008). Legitimacy and institutional change: the case of China. Comparative Political Studies, 41(3), 259-284.

Gladwell M. and Shirky C. (2011). From innovation to revolution: Do social media make protests possible? ProQuest. Foreign Affairs 90(2):153–154

Goldman, E., Rocholl, J., & So, J. (2013). Politically connected boards of directors and the allocation of procurement contracts. Review of Finance, 17(5), 1617-1648.

Gompers, P., & Lerner, J. (2001). The venture capital revolution. Journal of economic perspectives, 15(2), 145-168.

Gompers, P.A., (1995). Optimal investment, monitoring, and the staging of venture capital. The journal of finance, 50(5), pp.1461-1489.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics, 225(2), 254-277.

Granovetter, M. (1993). The Nature of Economic. Explorations in economic sociology, 3.

Green, P. E., & Srinivasan, V. (1978). Conjoint analysis in consumer research: issues and outlook. Journal of consumer research, 5(2), 103-123.

Green, P. E., & Srinivasan, V. (1990). Conjoint analysis in marketing: new developments with implications for research and practice. Journal of marketing, 54(4), 3-19.

Greenwood, B. N., & Gopal, A. (2015). Research note—Tigerblood: Newspapers, blogs, and the founding of information technology firms. Information Systems Research, 26(4), 812-828.

Greve, A. and Salaff, J.W., (2003). Social networks and entrepreneurship. Entrepreneurship theory and practice, 28(1), pp.1-22.

Griliches, Z., & Hausman, J. A. (1986). Errors in variables in panel data. Journal of econometrics, 31(1), 93-118.

Guo, R. j., Lev, B., and Zhou, N. (2004). "Competitive Costs of Disclosure by Biotech Ipos," Journal of Accounting Research (42:2), pp. 319-355.

Hadani, M., & Schuler, D. A. (2013). In search of El Dorado: The elusive financial returns on corporate political investments. Strategic Management Journal, 34(2), 165-181.

Haltiwanger, J., Jarmin, R.S. and Miranda, J., (2013). Who creates jobs? Small versus large versus young. Review of Economics and Statistics, 95(2), pp.347-361.

Harms, R., & Schwery, M. (2020). Lean start-up: operationalizing lean start-up capability and testing its performance implications. Journal of small business management, 58(1), 200-223.

Harnish, V. (2014). Scaling up. Ashburn, VA: Gazelles, Inc.

Harrison, R.T. and Mason, C.M., (2007). Does gender matter? Women business angels and the supply of entrepreneurial finance. Entrepreneurship Theory and Practice, 31(3), pp.445-472.

Hashai, N., & Markovich, S. (2017). Market entry by high technology start-ups: The effect of competition level and start-up innovativeness. Strategy Science, 2(3), 141-160.

Hausman, J.A., (1996). Valuation of new goods under perfect and imperfect competition. In The economics of new goods(pp. 207-248). University of Chicago Press.

He, C., Lu, J., & Qian, H. (2019). Entrepreneurship in China. Small Business Economics, 52(3), 563-572.

Hellmann, T. and Puri, M., (2000). The interaction between product market and financing strategy: The role of venture capital. The review of financial studies, 13(4), pp.959-984.

Hellmann, T., & Thiele, V. (2015). Friends or foes? The interrelationship between angel and venture capital markets. Journal of Financial Economics, 115(3), 639-653.

Hillman, A. J. (2005). Politicians on the board of directors: Do connections affect the bottom line?. Journal of management, 31(3), 464-481.

Hillman, A. J., & Hitt, M. A. (1999). Corporate political strategy formulation: A model of approach, participation, and strategy decisions. Academy of management review, 24(4), 825-842.

Hillman, A. J., & Wan, W. P. (2005). The determinants of MNE subsidiaries' political strategies: Evidence of institutional duality. Journal of International Business Studies, 36(3), 322-340.

Hitt, L. M., Jin, F., Wu, L. (2016). Data analytics skills and the corporate value of social media. Kelley School of Business Research Paper, (16-61).

Hoang, H. and Antoncic, B., (2003). Network-based research in entrepreneurship: A critical review. Journal of business venturing, 18(2), pp.165-187.

Hochberg, Y.V., Ljungqvist, A. and Lu, Y., (2007). Whom you know matters: Venture capital networks and investment performance. The Journal of Finance, 62(1), pp.251-301.

Hoeffler, S. (2003). Measuring preferences for really new products. Journal of marketing research, 40(4), 406-420.

Holburn, G. L., & Zelner, B. A. (2010). Political capabilities, policy risk, and international investment strategy: Evidence from the global electric power generation industry. Strategic Management Journal, 31(12), 1290-1315.

Hong, H., & Liskovich, I. (2015). Crime, punishment and the halo effect of corporate social responsibility (No. w21215). National Bureau of Economic Research.

Hong, S., (2013). Who benefits from Twitter? Social media and political competition in the US House of Representatives. Government Information Quarterly, 30(4), pp.464-472.

Honig, B., & Karlsson, T. (2004). Institutional forces and the written business plan. Journal of management, 30(1), 29-48.

Horowitz, B. (2010). Taking the Mystery out of Scaling a Company. Ben’s Blog. http://www.bhorowitz.com/taking_the_mystery_out_of_scaling_a_company.

Hsu, D. H., Roberts, E. B., & Eesley, C. E. (2007). Entrepreneurs from technology-based universities: Evidence from MIT. Research Policy, 36(5), 768-788.

Huang, J. (2018). The customer knows best: The investment value of consumer opinions. Journal of Financial Economics, 128(1), 164-182.

Huang, S., Aral, S., Hu, Y. J., & Brynjolfsson, E. (2020). Social Advertising Effectiveness Across Products: A Large-Scale Field Experiment. Marketing Science, forthcoming.

Hultin, M. and Szulkin, R., (1999). Wages and unequal access to organizational power: An empirical test of gender discrimination. Administrative Science Quarterly, 44(3), pp.453-472.

Hwang, C. Y., Titman, S., & Wang, Y. (2021). Investor Tastes, Corporate Behavior, and Stock Returns: An Analysis of Corporate Social Responsibility. Management Science.

Hyytinen, A., Pajarinen, M., & Rouvinen, P. (2015). Does innovativeness reduce start-up survival rates?. Journal of business venturing, 30(4), 564-581.

Iansiti, M., & Lakhani, K. R. (2020). Competing in the age of AI: strategy and leadership when algorithms and networks run the world. Harvard Business Press.

Ingram, P. and Roberts, P.W., (2000). Friendships among competitors in the Sydney hotel industry. American journal of sociology, 106(2), pp.387-423.

Ingram, P. and Roberts, P.W., (2000). Friendships among competitors in the Sydney hotel industry. American journal of sociology, 106(2), pp.387-423.

Jiang, J. (2018). Making Bureaucracy Work: Patronage Networks, Performance Incentives, and Economic Development in China. American Journal of Political Science, 62 (4), 982-999

Jin, F., Wu, A. and Hitt, L., (2017), January. Social Is the New Financial: How Start-up Social Media Activity Influences Funding Outcomes. In Academy of Management Proceedings (Vol. 1, p. 13329). Briarcliff Manor, NY 10510: Academy of Management.

Johnson, G.E., (1997). Changes in earnings inequality: the role of demand shifts. Journal of economic perspectives, 11(2), pp.41-54.

Jones, Q., Ravid, G., & Rafaeli, S., (2004). Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. Information systems research, 15(2), 194-210.

Joseph, D., Ang, S. and Slaughter, S.A., (2015). Turnover or turnaway? Competing risks analysis of male and female IT professionals' job mobility and relative pay gap. Information Systems Research, 26(1), pp.145-164.

Jung, M.J., Naughton, J.P., Tahoun, A. and Wang, C., (2017). Do firms strategically disseminate? Evidence from corporate use of social media. The Accounting Review, 93(4), pp.225-252.

Kanze, D., Huang, L., Conley, M.A. and Higgins, E.T., (2018). We ask men to win and women not to lose: Closing the gender gap in start-up funding. Academy of Management Journal, 61(2), pp.586-614.

Katila, R., Rosenberger, J. D., & Eisenhardt, K. M. (2008). Swimming with sharks: Technology ventures, defense mechanisms and corporate relationships. Administrative science quarterly, 53(2), 295-332.

Kerr, W. R., Nanda, R., & Rhodes-Kropf, M. (2014). Entrepreneurship as experimentation. Journal of Economic Perspectives, 28(3), 25-48.

Khaire, M. (2010). Young and no money? Never mind: The material impact of social resources on new venture growth. Organization Science, 21(1), 168-185.

Kim, H.G., Lee, S. and Kyeong, S., (2013), August. Discovering hot topics using Twitter streaming data social topic detection and geographic clustering. In 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013) (pp. 1215-1220). IEEE.

Kim, J. B., & Zhang, L. (2016). Accounting conservatism and stock price crash risk: Firm‐level evidence. Contemporary accounting research, 33(1), 412-441.

King, M. R. (2015). Political bargaining and multinational bank bailouts. Journal of International Business Studies, 46(2), 206-222.

Kleis, L., Chwelos, P., Ramirez, R.V. and Cockburn, I., (2012). Information technology and intangible output: The impact of IT investment on innovation productivity. Information Systems Research, 23(1), pp.42-59.

Koellinger, P., Minniti, M. and Schade, C., (2013). Gender differences in entrepreneurial propensity. Oxford bulletin of economics and statistics, 75(2), pp.213-234.

Kohavi, R., Tang, D., & Xu, Y. (2020). Trustworthy online controlled experiments: A practical guide to a/b testing. Cambridge University Press.

Kohli, A. K., Jaworski, B. J., & Kumar, A. (1993). MARKOR: a measure of market orientation. Journal of Marketing research, 30(4), 467-477.

Koning, R., Hasan, S., & Chatterji, A. (2022). Experimentation and Start-up Performance: Evidence from A/B Testing. Management Science.

Krishnan, V., & Loch, C. H. (2005). A retrospective look at production and operations management articles on new product development. Production and Operations Management, 14(4), 433-441.

Krishnan, V., & Ulrich, K. T. (2001). Product development decisions: A review of the literature. Management science, 47(1), 1-21.

Kroll, M., Walters, B. A., & Wright, P. (2008). Board vigilance, director experience, and corporate outcomes. Strategic Management Journal, 29(4), 363-382.

Krueger, A. O. (1974). The political economy of the rent-seeking society. The American economic review, 64(3), 291-303.

Krug, B., & Hendrischke, H. (2008). Framing China: Transformation and institutional change through co-evolution. Management and Organization Review, 4(1), 81-108.

Kulshrestha, J., Kooti, F., Nikravesh, A. and Gummadi, K.P., (2012), May. Geographic dissection of the twitter network. In Sixth International AAAI Conference on Weblogs and Social Media.

Kumar, V., Bhaskaran, V., Mirchandani, R., & Shah, M. (2013). Practice prize winner—creating a measurable social media marketing strategy: increasing the value and ROI of intangibles and tangibles for hokey pokey. Marketing Science, 32(2), 194-212.

Kwark, Y., Chen, J., & Raghunathan, S. (2018). User-generated content and competing firms' product design. Management Science, 64(10), 4608-4628.

Lambrecht, A., Tucker, C., & Wiertz, C. (2018). Advertising to early trend propagators: Evidence from Twitter. Marketing Science, 37(2), 177-199.

Lee, D., Hosanagar, K., Nair, H. S. (2018). Advertising content and consumer engagement on social media: Evidence from Facebook. Management Science, 64(11), 5105-5131.

Lee, H. S. (2017). Peer networks in venture capital. Journal of Empirical Finance, 41, 19-30.

Lee, J. K., Choi, J., Kim, C., & Kim, Y., (2014). Social media, network heterogeneity, and opinion polarization. Journal of communication, 64(4), 702-722.

Lerner, J. and Malmendier, U., (2013). With a little help from my (random) friends: Success and failure in post-business school entrepreneurship. The Review of Financial Studies, 26(10), pp.2411-2452.

Li, H., & Zhang, Y. (2007). The role of managers' political networking and functional experience in new venture performance: Evidence from China's transition economy. Strategic management journal, 28(8), 791-804.

Li, H., Meng, L., Wang, Q., & Zhou, L. A. (2008). Political connections, financing and firm performance: Evidence from Chinese private firms. Journal of development economics, 87(2), 283-299.

Li, J., & Wu, D. (2020). Do corporate social responsibility engagements lead to real environmental, social, and governance impact?. Management Science, 66(6), 2564-2588.

Li, J., Li, M., Wang, X., & Thatcher, J. B. (2021). Strategic Directions for AI: The Role of CIOs and Boards of Directors. MIS Quarterly, 45(3).

Li, S., Li, S., & Zhang, W. (2000). The road to capitalism: Competition and institutional change in China. Journal of Comparative Economics, 28(2), 269-292.

Li, S., Schulze, W. and Li, Z., (2009). Plunging into the sea, again? A study of serial entrepreneurship in China. Asia Pacific Journal of Management, 26(4), pp.667-680.

Li, X. , & Wu, L., (2018). Herding and social media word-of-mouth. MIS Quarterly, 42(4), 1331-1351.

Lin, M., Prabhala, N.R. and Viswanathan, S., (2013). Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. Management Science, 59(1), pp.17-35.

Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. the Journal of Finance, 72(4), 1785-1824.

Liu, X., Singh, P. V., & Srinivasan, K. (2016). A structured analysis of unstructured big data by leveraging cloud computing. Marketing Science, 35(3), 363-388.

Lo, Y. L., Li, Y., & Chan, K. C. (2019). Contemporary Innovation in China. The Chinese Economy, 52(5), 387-399.

Lou, B., & Wu, L. (2021). AI on Drugs: Can Artificial Intelligence Accelerate Drug Development? Evidence from a Large-scale Examination of Bio-pharma Firms. Evidence from a Large-scale Examination of Bio-pharma Firms. MIS Quarterly, 45(3).

Lu, T., Zhang, Y. and Li, B., (2019). The Value of Alternative Data in Credit Risk Prediction: Evidence from a Large Field Experiment.

Luo, X., Zhang, Y., Zeng, F., Qu, Z. (2020). Complimentarity and Cannibalization of Offline-Toonline Targeting: A Field Experiment on Omnichannel Commerce. MIS Quarterly, 44(2)

Luo, Y. (2014). The Internet and agenda setting in China: The influence of online public opinion on media coverage and government policy. International Journal of Communication, 8, 24.

Lux, S., Crook, T. R., & Woehr, D. J. (2011). Mixing business with politics: A meta-analysis of the antecedents and outcomes of corporate political activity. Journal of management, 37(1), 223-247.

Lynn, G. S., Morone, J. G., & Paulson, A. S. (1996). Marketing and discontinuous innovation: the probe and learn process. California management review, 38(3), 8-37.

Malpani, R., & Ghosh, A. (2020). Can a start-up be sustainable and socially responsible? A glimpse of activities for sustainability. A Glimpse of Activities for Sustainability (July 22, 2020).

Martin, G. J., & Yurukoglu, A. (2017). Bias in cable news: Persuasion and polarization. American Economic Review, 107(9), 2565-99.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. AI magazine, 27(4), 12-12.

McIntyre, K. E. (2014). The evolution of social media from 1969 to 2013: A change in competition and a trend toward complementary, niche sites. The Journal of Social Media in Society, 3(2).

Meek, W. R., Pacheco, D. F., & York, J. G. (2010). The impact of social norms on entrepreneurial action: Evidence from the environmental entrepreneurship context. Journal of Business Venturing, 25(5), 493-509.

Mei, X., Cheng, H. K., Bandyopadhyay, S., Qiu, L., & Wei, L. (2022). Sponsored Data: Smarter Data Pricing with Incomplete Information. Information Systems Research, 33(1), 362-382.

Meier, R.L., (1963). Information input overload: features of growth in communications-oriented institutions. Libri, 13(1), pp.1-44.

Mellahi, K., Frynas, J. G., Sun, P., & Siegel, D. (2016). A review of the nonmarket strategy literature: Toward a multi-theoretical integration. Journal of management, 42(1), 143-173.

Mendelson, H., & Pillai, R. R. (1999). Information age organizations, dynamics and performance. Journal of Economic Behavior & Organization, 38(3), 253-281.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. American journal of sociology, 83(2), 340-363.

Meyer, K. E., Ding, Y., Li, J., & Zhang, H. (2018). Overcoming distrust: How state-owned enterprises adapt their foreign entries to institutional pressures abroad. In State-Owned Multinationals (pp. 211-251). Palgrave Macmillan, Cham.

Milgrom, P., & Roberts, J. (1994). Comparing equilibria. The American Economic Review, 441-459.

Mollick, E. (2019). What the lean start-up method gets right and wrong. Harvard Business Review, 10, 1-4.

Naik, P. A., Raman, K. (2003). Understanding the impact of synergy in multimedia communications. Journal of marketing research, 40(4), 375-388.

Nanda, R. and Sørensen, J.B., (2010). Workplace peers and entrepreneurship. Management Science, 56(7), pp.1116-1126.

Nanda, R. and Sørensen, J.B., (2010). Workplace peers and entrepreneurship. Management Science, 56(7), pp.1116-1126.

Netessine, S., & Taylor, T. A. (2007). Product line design and production technology. Marketing Science, 26(1), 101-117.

Nicolaou, N. and Birley, S., (2003). Social networks in organizational emergence: The university spinout phenomenon. Management science, 49(12), pp.1702-1725.

North, D. C. (1990). Economic performance (Vol. 3). Cambridge: Cambridge University Press.

Ofek, E., & Turut, O. (2008). To innovate or imitate? Entry strategy and the role of market research. Journal of Marketing Research, 45(5), 575-592.

Okhmatovskiy, I. (2010). Performance implications of ties to the government and SOEs: A political embeddedness perspective. Journal of management studies, 47(6), 1020-1047.

Oliver, C., & Holzinger, I. (2008). The effectiveness of strategic political management: A dynamic capabilities framework. Academy of Management Review, 33(2), 496-520.

Orser, B.J., Riding, A.L. and Manley, K., (2006). Women entrepreneurs and financial capital. Entrepreneurship Theory and Practice, 30(5), pp.643-665.

Oster, E., (2019). Unobservable selection and coefficient stability: Theory and evidence. Journal of Business & Economic Statistics, 37(2), pp.187-204.

Parker, S.C., (2013). Do serial entrepreneurs run successively better-performing businesses?. Journal of Business Venturing, 28(5), pp.652-666.

Peng, M. W., Wang, D. Y., & Jiang, Y. (2008). An institution-based view of international business strategy: A focus on emerging economies. Journal of international business studies, 39(5), 920-936.

Pérez Santín, E., Rodríguez Solana, R., González García, M., García Suárez, M. D. M., Blanco Díaz, G. D., Cima Cabal, M. D., ... & López Sánchez, J. I. (2021). Toxicity prediction based on artificial intelligence: A multidisciplinary overview. Wiley Interdisciplinary Reviews: Computational Molecular Science, 11(5), e1516.

Petrova, M., Sen, A., & Yildirim, P. (2020). Social Media and Political Contributions: The Impact of New Technology on Political Competition. Management Science, forthcoming.

Piva, M., Santarelli, E. and Vivarelli, M., (2005). The skill bias effect of technological and organisational change: Evidence and policy implications. Research Policy, 34(2), pp.141-157.

Qiu, L., & Kumar, S. (2017). Understanding voluntary knowledge provision and content contribution through a social-media-based prediction market: A field experiment. Information Systems Research, 28(3), 529-546.

Que, W., Zhang, Y., & Schulze, G. (2019). Is public spending behavior important for Chinese official promotion? Evidence from city-level. China Economic Review, 54, 403-417.

Rajagopalan, N., & Finkelstein, S. (1992). Effects of strategic orientation and environmental change on senior management reward systems. Strategic Management Journal, 13(S1), 127-141.

Ransbotham, S., Lurie, N. H., & Liu, H. (2019). Creation and consumption of mobile word of mouth: how are mobile reviews different?. Marketing Science, 38(5), 773-792.

Rathore, A. K., Kar, A. K., & Ilavarasan, P. V. (2017). Social media analytics: Literature review and directions for future research. Decision Analysis, 14(4), 229-249.

Read, S., Sarasvathy, S., Dew, N., Wiltbank, R., & Ohlsson, A. V. (2011). Effectual Entrepreneurship: 240.

Ridgeway, C.L., (2001). Gender, status, and leadership. Journal of Social issues, 57(4), pp.637-655.

Ries, E. (2011). The lean start-up: How today's entrepreneurs use continuous innovation to create radically successful businesses. Currency.

Rigby, D. K., Sutherland, J., & Noble, A. (2018). Agile at scale. Harvard Business Review, 96(3), 88-96.

Rishika, R., J. Ramaprasad. (2019). The Effect of Asymmetric Social Ties on Online Contribution: The Role of Tie Strength and Homophily. Management Science. 65(7), 2947-3448

Rishika, R., Kumar, A., Janakiraman, R., & Bezawada, R. (2013). The effect of customers' social media participation on customer visit frequency and profitability: an empirical investigation. Information systems research, 24(1), 108-127.

Rudman, L.A. and Glick, P., (2001). Prescriptive gender stereotypes and backlash toward agentic women. Journal of social issues, 57(4), pp.743-762.

Sant’Anna, P. H., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. Journal of Econometrics, 219(1), 101-122.

Sarasvathy, S. D. (2001). Causation and effectuation: Toward a theoretical shift from economic inevitability to entrepreneurial contingency. Academy of management Review, 26(2), 243-263.

Saxenian, A., (1991). The origins and dynamics of production networks in Silicon Valley. Research policy, 20(5), pp.423-437.

Schuler, D. A., Rehbein, K., & Cramer, R. D. (2002). Pursuing strategic advantage through political means: A multivariate approach. Academy of Management Journal, 45(4), 659-672.

Schumpeter, J. A. (1966). Capitalism, Socialism and Democracy. New York: Harper and Brothers.

Scott, W. R. (1995). Institutions and organizations (Vol. 2). Thousand Oaks, CA: Sage.

Shane, S. and Cable, D., (2002). Network ties, reputation, and the financing of new ventures. Management science, 48(3), pp.364-381.

Shane, S., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. Academy of management review, 25(1), 217-226.

Sharma, A., Saboo, A. R., & Kumar, V. (2018). Investigating the influence of characteristics of the new product introduction process on firm value: The case of the pharmaceutical industry. Journal of Marketing, 82(5), 66-85.

Siegel, J. (2007). Contingent political capital and international alliances: Evidence from South Korea. Administrative Science Quarterly, 52(4), 621-666.

Sine, W. D., Haveman, H. A., & Tolbert, P. S. (2005). Risky business? Entrepreneurship in the new independent-power sector. Administrative Science Quarterly, 50(2), 200-232.

Singh, J., Hansen, M.T. and Podolny, J.M., (2010). The world is not small for everyone: Inequity in searching for knowledge in organizations. Management Science, 56(9), pp.1415-1438.

Snir, E. M., & Hitt, L. M. (2003). Costly bidding in online markets for IT services. Management Science, 49(11), 1504-1520.

Sorenson, O. and Stuart, T.E., (2001). Syndication networks and the spatial distribution of venture capital investments. American journal of sociology, 106(6), pp.1546-1588.

Soundararajan, V., Jamali, D., & Spence, L. J. (2018). Small business social responsibility: A critical multilevel review, synthesis and research agenda. International Journal of Management Reviews, 20(4), 934-956.

Sridhar, S., Sriram, S. (2015). Is online newspaper advertising cannibalizing print advertising?. Quantitative Marketing and Economics, 13(4), 283-318.

Stephen, A. T., & Galak, J. (2012). The effects of traditional and social earned media on sales: A study of a microlending marketplace. Journal of marketing research, 49(5), 624-639.

Strang D., Bradburn E.M. (2001). Theorizing legitimacy or legitimating theory? Neoliberal discourse and HMO policy. Campbell JL, Pedersen OK, eds. The Rise of Neoliberalism and Institutional Analysis: Neoliberalism in Perspective (Princeton University Press, Princeton, NJ), 129–158.

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. Academy of management review, 20(3), 571-610.

Sun, P., Mellahi, K., & Wright, M. (2012). The contingent value of corporate political ties. Academy of Management Perspectives, 26(3), 68-82.

Sutton, R. I., & Rao, H. (2014). Scaling up excellence. New York: Crown Business.

Taddy, M. (2018). The technological elements of artificial intelligence. In The economics of artificial intelligence: An agenda (pp. 61-87). University of Chicago Press.

Tambe, P. (2014). Big data investment, skills, and firm value. Management science, 60(6), 1452-1469.

Tambe, P., Hitt, L. M., & Brynjolfsson, E. (2012). The extroverted firm: How external information practices affect innovation and productivity. Management Science, 58(5), 843-859.

Tang, X., Chen, W., & Wu, T. (2018). Do authoritarian governments respond to public opinion on the environment? Evidence from China. International journal of environmental research and public health, 15(2), 266.

Tarafdar, M., & Kajal Ray, D. (2021). Role of Social Media in Social Protest Cycles: A Sociomaterial Examination. Information Systems Research.

Tarasconi, G. and Menon, C., (2017). Matching Crunchbase with patent data. OECD Science, Technology and Industry Working Papers, No. 2017/07, OECD Publishing, Paris, https://doi.org/10.1787/15f967fa-en.

Thiel, P. A., & Masters, B. (2014). Zero to one: Notes on start-ups, or how to build the future. Currency.

Thomke, S. H. (2020). Experimentation works: The surprising power of business experiments. Harvard Business Press.

Tollison, R. D. (1982). Rent Seeking: A Survey. Kylos, 35 (November): 575-602.

Tong, J. (2010). The crisis of the centralized media control theory: how local power controls media in China. Media, Culture & Society, 32(6), 925-942.

Tsai, K. S. (2006). Adaptive informal institutions and endogenous institutional change in China. World Politics, 59(1), 116-141.

Tufekci, Z., & Wilson, C. (2012). Social media and the decision to participate in political protest: Observations from Tahrir Square. Journal of communication, 62(2), 363-379.

Van Alstyne, M., & Brynjolfsson, E., (2005). Global village or cyber-balkans? Modeling and measuring the integration of electronic communities. Management Science, 51(6), 851-868.

Van Gelderen, M., Thurik, R. and Bosma, N., (2005). Success and risk factors in the pre-start-up phase. Small business economics, 24(4), pp.365-380.

Varian, H. (2018). Artificial intelligence, economics, and industrial organization. In The economics of artificial intelligence: an agenda (pp. 399-419). University of Chicago Press.

Verganti, R., Vendraminelli, L., & Iansiti, M. (2020). Innovation and design in the age of artificial intelligence. Journal of Product Innovation Management, 37(3), 212-227.

Wang, C., Hong, J., Kafouros, M., & Wright, M. (2012). Exploring the role of government involvement in outward FDI from emerging economies. Journal of International Business Studies, 43(7), 655-676.

Wang, F. L. (2016). Institutions and institutional change in China: Premodernity and modernization. Springer.

Wang, X., & Wu, L. (2019). Can Social Media Alleviate Inequalities? Evidence from Venture Capital Financing. Proceedings of International Conference on Information Systems.

Werner, T. (2015). Gaining access by doing good: The effect of sociopolitical reputation on firm participation in public policy making. Management Science, 61(8), 1989-2011.

Wilson, L. O., & Norton, J. A. (1989). Optimal entry timing for a product line extension. Marketing Science, 8(1), 1-17.

Wright, M., Robbie, K. and Ennew, C., (1997). Venture capitalists and serial entrepreneurs. Journal of business venturing, 12(3), pp.227-249.

Wu, L. (2013). Social network effects on productivity and job security: Evidence from the adoption of a social networking tool. Information systems research, 24(1), 30-51.

Wu, L., & Brynjolfsson, E. (2015). The future of prediction: How Google searches foreshadow housing prices and sales. In Economic analysis of the digital economy (pp. 89-118). University of Chicago Press.

Wu, L., & Kane, G. C. (2021). Network-Biased Technical Change: How Modern Digital Collaboration Tools Overcome Some Biases but Exacerbate Others. Organization Science, 32(2), 273-292.

Wu, L., (2013). Social network effects on productivity and job security: Evidence from the adoption of a social networking tool. Information systems research, 24(1), pp.30-51.

Wu, L., Hitt, L., & Lou, B. (2020). Data analytics, innovation, and firm productivity. Management Science, 66(5), 2017-2039.

Wu, L., Lou, B., & Hitt, L. (2019). Data analytics supports decentralized innovation. Management Science, 65(10), 4863-4877.

Wu, Y., Eesley, C., Eisenhardt, K. (2020). Entrepreneurship in Dynamic Environments: A Comparison Between the U.S. and China. Quarterly Journal of Management.

Xiao, Q. (2011). The rise of online public opinion and its political impact. Changing media, changing China, 202-24.

Xu, B., & Albert, E. (2014). Media censorship in China. Council on Foreign Relations, 25, 243.

Ye, Y., Xu, P., & Zhang, M. 2017. Social media, public discourse and civic engagement in modern China. Telematics and Informatics, 34(3), 705-714.

Yoo, O. S., Huang, T., & Arifoğlu, K. (2021). A theoretical analysis of the lean start-up method. Marketing Science, 40(3), 395-412.

Zhu, C. (2019). Big data as a governance mechanism. The Review of Financial Studies, 32(5), 2021-2061.