# Unpacking Data-driven Technologies and Resource Mobilization Landscape Shifts in New Ventures

Sung ho Park

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
University of Washington
2024

Reading Committee:
Benjamin Hallen, Chair
Warren Boeker
Emily Cox Pahnke
Alicia DeSantola

Program Authorized to Offer Degree:
Foster School of Business

University of Washington

## Abstract

# Unpacking Data-driven Technologies and Resource Mobilization Landscape Shifts in New Ventures

Sung Ho Park

Chair of the Supervisory Committee:

Benjamin Hallen

Department of Management and Organization

Foster School of Business

Over the last two decades, the entrepreneurial landscape has undergone significant change. New ventures now are in the presence of low-cost data-driven technologies, novel entrepreneurial support systems, and evolving fundraising landscapes. While these developments have made entrepreneurship more accessible, they have simultaneously increased the complexity of the venture creation and growth process. As a result, these changes raise important questions about how they affect a venture's ability to gain traction, mobilize resources, and pursue growth. My dissertation examines several facets of these changes. In the first chapter, I examine whether and when data analytics technologies enhance the visibility of a venture. I find what while these technologies are beneficial on average, there is significant variation among ventures in terms of the informational value gained by data analytics and the founding team's capability to effectively leverage the technology. In the second chapter, with my co-authors, I examine the factors contributing to venture valuation, focusing on the relative explanatory power of environmental conditions, founding team characteristics, and venture progress. We find that a significant amount of variance in valuation is linked to current and previous investment rounds, whereas the contributions of the founding team and venture progress are relatively minor. Lastly, in the third chapter, with my co-authors, I assess the role of seed accelerators in helping ventures match with high-status investors. Our results indicate that while some seed accelerators are effective in facilitating these connections, others are not. Further analysis provides a nuanced view of the benefits of seed accelerators in the investment matchmaking process.

# TABLE OF CONTENTS

# ACKNOWLEDGEMENTS

As the saying goes, *'No one achieves anything alone.'*—Leslie Knope. This dissertation is a testament to that phrase, made possible by the amazing support and contributions of many people around me.

First, I would like to express my deepest gratitude to my advisor/chair. Dr. Benjamin Hallen has provided truly invaluable learning opportunities, from teaching me how to formulate important research questions to guiding me in writing code efficiently. Additionally, through the many office hours and phone calls, I have learned a great deal about how to navigate challenges and move forwards in the wonderfully unique world of academic research. Furthermore, I am particularly grateful for the intellectual freedom he gave me in exploring novel ideas, while guiding me to maintain rigor and balance.

I am also immensely thankful for the steadfast support from my committee members. Dr. Warren Boeker has always been welcoming of impromptu discussions in his office, providing feedback that helped bring parsimony to my convoluted ideas. Dr. Emily Cox Pahnke has been instrumental during my internal presentations, offering constructive feedback that enhanced my papers with greater nuance, not to mention ensuring I had sufficient support for networking and conference attendances. Dr. Alicia DeSantola's guidance in navigating reviews and the job market has been equally invaluable, as have the practical insights I have gained from her brown bag presentations. Finally, Dr. Jerald Herting has enriched my skillsets with accessible technical knowledge, which I am extremely grateful for.

Further appreciation goes to faculty outside of my committee who have been mentors and co-authors. Dr. Suresh Kotha generously shared his deep insights into entrepreneurship during our office meetings, and he also instilled in me the drive to advance research. I am deeply grateful to Dr. Susan Cohen at the University of Georgia for her assistance in the paper revision process and

for helping to expand my network of scholars. My gratitude also goes out to Alex Murray at the University of Oregon for his encouragement during my first ever PhD presentation that help me to push forward. Many others including Drs. David Sirmon, Kevin Steensma, Abhinav Gupta, Ming Zhu Wang, and Ryan Allen have also greatly enriched my academic journey.

Additionally, my fellow doctoral students (some of whom are faculty members)—Yu Tse Heng, Michelle Lee, Jen Rhymer, Anna Fung, Joseph Shin, Majid Majzoubi, Niki West, Wei Xia, Sanjana Prabhu, Scott Taylor, Brit McClellan, Keehyuk Ra, Rohan Radke, Vishal Rituraj, Evy Smith, Anusuya Banerjee, Ekonkar Kaur, Jenson Lau, Shawn Quan, and SM Musa—have made this journey both enjoyable and worthwhile. A special thank you to Sorah Seong, who has profoundly influenced my teaching pedagogy and curriculum design.

Finally, I extend an extra special thanks to my family. My parents, YT Park and Myung Suk Lee, were instrumental in my decision to enter academia, which was one of the best decisions of my life. Their unwavering moral support and love have been the absolute cornerstone of my achievements. I also thank my sister, Sumi Park, for making the life journey enjoyable and memorable through our ventures across the US.

I would like to thank all people who are not mentioned here but have provided richness and joy throughout my doctoral journey. Your help has not gone unnoticed and I deeply appreciate you.

# LIST OF CO-AUTHORS

Chapter 2. Benjamin Hallen and Suresh Kotha

Chapter 3. Benjamin Hallen and Susan Cohen[1]

---

[1] A similar version of this chapter was published as: Hallen, B. L., Cohen, S. L., & Park, S. H. (2023). Are seed accelerators status springboards for startups? Or sand traps? *Strategic Management Journal*, 44(8), 2060–2096. The reuse was done with permission from **Wiley/Strategic Management Journal**.

# INTRODUCTION

As a means to mobilize critical resources and pursue growth, entrepreneurs often strategize by engaging in activities that combine ‘doing’ and ‘thinking’ (Ott, Eisenhardt, & Bingham, 2017). ‘Doing’ involves strategizing based on direct experience that yields learning opportunities (e.g., Baker & Nelson, 2005; Politis, 2005; Kerr, Nanda, & Rhodes-Kropf, 2014; Peterson & Wu, 2021; Bingham & Davis, 2012; Contigiani & Young-Hyman, 2021; Hallen, Cohen, & Bingham, 2020; McDonald & Eisenhardt, 2020). For instance, founders pursue bootstrapping via bricolage to find and pursue novel opportunities, which involves learning to combine various resources that are currently available at hand (Baker & Nelson, 2005) or repurpose existing resources for further added value creation (Welter, Mauer, & Wuebker, 2016). Additionally, in the midst of high uncertainty surrounding the business model, customer base, and appropriate technologies, founders often engage in trial-and-error (Bingham & Davis, 2012) or experimentation (Andries, Debackere, & Van Looy, 2013; Koning, Hasan, & Chatterji, 2022; McDonald & Eisenhardt, 2020) to uncover new and relevant information useful for strategizing.

On the other side of the strategy formulation spectrum is ‘thinking’, which involves cognition and is affected by organizational and individual history (Delmar & Shane, 2003; Gary & Wood, 2011; Powell & Baker, 2014). Facing high complexity, entrepreneurs engage in thinking through the development of mental models that help create a simplified version of the highly convoluted entrepreneurial situations they face (Gary & Wood, 2011; Kiss & Barr, 2015). Similarly, entrepreneurs also utilize heuristics (Bingham & Eisenhardt, 2011) and invoke analogies (Gavetti & Rivkin, 2007) as a means to strategizing.

While these entrepreneurial approaches of strategizing by ‘doing’ and ‘thinking’ explain much about why some entrepreneurs achieve superior competitive positions or greater success in resource mobilization and growth (Ott et al., 2017), how entrepreneurs engage in strategizing behaviors have drastically evolved over the past decade. For instance, strategizing by experimenting (a form of ‘doing’) has become more scientifically rigorous and structured with the help of technology tools. Entrepreneurs can now conduct low-cost experimentation through randomized control trials via online A/B testing technologies, facilitating the acquisition of causal knowledge among low-resource ventures (Koning et al., 2022). Moreover, entrepreneurs have access to data analytics technologies that help gather large amounts of data on their potential customers, allowing them to strategically target individuals that are more likely to convert into actual customers (Berman & Israeli, 2022).

Concurrently, advancements in ‘thinking’ have occurred. For instance, rich data gathered through data analytics technologies shapes how entrepreneurs think and what issues are given greater attention, likely altering the depth/breadth of their cognitive maps of the entrepreneurial landscape and the audience (Murray, Rhymer, & Sirmon, 2020; Adner & Levinthal, 2008; Gary & Wood, 2011). Additionally, entrepreneurs have opportunities to enroll in sponsor organizations that shape their strategic thinking processes (Hallen et al., 2020; Cohen, Hallen, & Bingham, 2019; Assenova, 2020). For instance, seed accelerators have increasingly become a popular approach for early-stage venture founders to engage in inter/intra-organizational learning. Through consultation and peer interactions in accelerators, founders enhance their business strategies by incorporating the experiences gleaned from accelerators into their own business ideas (Hallen et al., 2020; Cohen et al., 2019).

Despite the increasingly evolving ways entrepreneurs think and act to develop strategies and the growing body of research exploring these shifts, many questions still remain about the effectiveness of these new approaches to entrepreneurial strategizing in regard to resource mobilization and growth. For instance, it is uncertain whether using data analytics provides a unique advantage for new ventures over relying on intuition or anecdotal evidence, especially since smaller organizations often lack the capability or the right circumstances to fully exploit these benefits. Moreover, there is the lack of clarity about the role of accelerators in enhancing a venture's ability to secure higher-status investors post-graduation, particularly given the potential negative or weak signaling effects associated with accelerator enrollment.

Furthermore, a related question remains around *how much* the outcomes of strategies from ‘doing’ and ‘thinking’—in the form of achievements and progress—influence entrepreneurial resource mobilization and growth. Numerous studies underscore the strong impact of broader environmental factors and stakeholder dynamics on venture outcomes, which are beyond an individual entrepreneur's control. For example, the geographic location of the venture (Hochberg, Ljungqvist, & Lu, 2010; Wasserman, 2017), the competitive industry environment (Gompers & Lerner, 2000; Hsu, 2007), and temporality (Gompers & Lerner, 2000; Hsu, 2007; Moghaddam, Bosse, & Provance, 2016) all are found to have substantial influence of which entrepreneurs do not have the complete agency.

To better understand several important facets of the newly evolving ways that entrepreneurs attempt to augment strategizing by ‘doing’ and/or ‘thinking’, I explore two prominent means: *audience data analytics* and *seed accelerators.*

Chapter 1 examines the outcomes of new ventures adopting audience data analytics technologies, which imply a fundamental shift in how entrepreneurs learn about their audience and, consequently, how they ‘think’ to strategize. Moving beyond reliance on intuition, simple rules, small-scale anecdotal evidence, and gut feelings, data analytics offers an opportunity to enhance decision-making with information that is more objective, precise, voluminous, and accurate (Brynjolfsson, Hitt, & Kim, 2011; Brynjolfsson & McElheran, 2016). Additionally, from the perspective of organizational interpretation systems, this represents a fundamental shift in how data is interpreted and assigned meaning (Daft & Weick, 1984), as data analytics either supplements or replaces the traditional roles of managers and entrepreneurs in gathering, analyzing, and interpreting audience information.

Whether data analytics is superior in helping ventures mobilize resources is still largely unknown among new ventures. On one hand, these technologies can enhance prominence by addressing information gaps and reducing the cognitive load of managing audience engagement—benefits that are particularly valuable for new ventures without historical data and facing informational constraints (Huber, 1990; Brynjolfsson & McElheran, 2019; Berman & Israeli, 2022; Simon, 1955; Gavetti, Levinthal, & Ocasio, 2007). On the other hand, new organizations are more susceptible to struggling with novel cognitive biases when employing data analytics (March, 2006; Muller, 2008; Ridway, 1956; Mazmanian & Beckman, 2018). Furthermore, these tools may offer limited value to new ventures because they primarily focus on reflecting external data. This approach may not be as beneficial in entrepreneurship, where generative thinking—creating new ideas and solutions rather than merely retrieving existing information—is often more critical (Felin, Gambardella, Novelli, & Zenger, 2023; Felin & Zenger, 2009; Felin & Zenger, 2017).

I investigate the use of web-based audience data analytics among 6,903 new U.S. ventures to assess their impact on enhancing prominence/traction on the web. Despite the theorized limitations, I find ventures that use data analytics yield net benefits. Employing a novel time-series cross-sectional matching approach, I find a 28-52% growth in web visibility and a 1-10% reduction in its volatility among ventures that use audience data analytics.

These results imply that data analytics as a means to augmenting ‘thinking’ (and even possibly ‘doing’) is more effective than traditional modes of information acquisition and analysis of audience. Additional findings suggest that data analytics are especially beneficial in well-established market segments, result in increased prominence/traction when paired with A/B testing, and are influenced by the entrepreneurial prominence of the founding team's previous employers. These contingencies indicate that entrepreneurs should selectively choose the environments in which they implement data analytics-augmented forms of ‘thinking’ and the actions that follow accordingly.

In Chapter 3, alongside my co-authors, we examine the role of seed accelerators in assisting ventures to attract high-status investors post-graduation. Accelerators offer distinctive educational experiences through intensive and structured consultations that shape entrepreneurial mindsets and actions (Hallen et al., 2020; Cohen et al., 2019). As a result, it can alter and enhance how entrepreneurs think and act to formulate strategies.

While much of recent research highlights how such programs influence venture performance, growth, raising capital, gaining customer traction, time until shutdown, Mitigating bounded rationality, learning, and alters the entrepreneurial ecosystem (e.g., abundance of funds) (Hallen et al., 2020; Gonzalez-Uribe & Leatherbee, 2018; Cohen, Hallen, & Bingham, 2019; Assenova, 2020; Yu, 2020; Fehder & Hochberg, 2022), less is known about whether such program influences their chances of obtaining prestigious partners.

This chapter extends the investigation to explore whether participation in accelerators enables early-stage startups to secure high-status investors post-graduation. From one perspective, participating in accelerators can enhance a startup's quality and/or effectively signal this improvement to potential partners, attracting desirable collaborations (Podolny, 1994; Azoulay, Stuart, & Wang, 2014; Sauder, 2008; Connelly, Certo, Ireland, & Reutzel, 2011; Spence, 1973; Gonzalez-Uribe & Leatherbee, 2018; Hallen et al., 2020; Lall, Chen, & Roberts, 2020). However, studies also indicate that these signaling effects might be minimal or even negative since accelerators could be perceived as a "market for lemons," where the presence of lower-quality startups diminishes the overall perceived value (Simcoe & Waguespack, 2011; Azoulay et al., 2014; Akerlof, 1970; Feltovich, Harbaugh, & To, 2002).

By analyzing two unique and proprietary U.S. startup datasets, we discovered that while certain accelerators significantly help startups attract high-status investors, others have limited or slightly negative effects. In summary, the use of accelerators as a tool to enhance venture strategy formulation through thinking and/or doing brings considerable benefits in securing high-status investors, as evidenced by our findings.

Finally, Chapter 2 offers a distinct yet related angle on the topic of entrepreneurial strategizing. Here, with my co-authors, we explore how much of venture achievements and progress—which results from strategizing by 'doing' and/or 'thinking'—matters in achieving greater valuation outcomes.

A venture's valuation is a key component of venture capital equity investment. It holds theoretical significance for researchers focused on strategy and entrepreneurship because it encapsulates the consensus between entrepreneurs and major resource providers (investors) on the distribution of future value created by the venture (Campbell, Coff, & Kryscynski, 2012; Gans &

Ryall, 2017; Mindruta, Moeen, & Agarwal, 2016; Wasserman, 2017). Furthermore, valuation acts as an informative indicator, revealing the cost of capital for new ventures and mirroring investor perceptions of a venture's value or quality (Hsu, 2007). The financial assessment of a firm is influenced by various elements such as its network, business concept, and appropriability regime (Baum & Silverman, 2004; Hall & Hofer, 1993; Hsu, 2007). Consequently, similar to fundraising efforts, research has pinpointed multiple determinants of venture valuation including attributes of the entrepreneur, the venture, the industry, the region, and the investors.

However, what is unclear is how much the achievements and progress made by the entrepreneur matters in determining valuation. Research identifies three main categories that affect early venture valuations: venture progress, founding team attributes, and environmental factors. Environmental factors include the industry, region, and fluctuations in investment capital supply (Eisenhardt & Schoonhoven, 1990; Gompers & Lerner, 2000; Kotha, Shin, & Fisher, 2022). Attributes of the founding team reflect the team's human and social capital (Agarwal, Echambadi, Franco, & Sarkar, 2004; Batjargal & Liu, 2004; Chatterji, 2009; Eesley & Roberts, 2012; Hsu, 2007; Wasserman, 2017). Venture progress indicates reduced uncertainty through metrics like revenue, profitability, and customer engagement (Claes & Vissa, 2020; Hallen & Eisenhardt, 2012; Hsu, 2007; Hsu & Ziedonis, 2013). Despite a wealth of studies on each factor independently, how important the outcomes stemming from entrepreneurial strategizing *relative to* other factors such as team attributes or environmental factors remains unclear. Extant research has not simultaneously examined all three and has focused more on estimating the absolute impact of specific attributes rather than their comparative importance.

Our research analyzes 1,737 U.S. technology-based ventures from 2007 to 2022 across three stages of funding, using comprehensive data on founding team attributes and venture progress. We apply the Shapley Value Approach (SVA), a sophisticated method from game theory that assesses the marginal contributions of different factors without needing strict assumptions about their interactions (Sharapov, Kattuman, Rodriguez, & Velazquez, 2021; Shapley, 1953; Ghorbani & Zou, 2019).

Our study reveals that environmental factors, particularly regional disparities, significantly influence valuations, accounting for 12-16% of the variance across various funding stages. In contrast, tangible measures of venture progress explain only 3-4% of valuation differences in the early stages, though this percentage increases to 11.5% in later rounds. In summary, our findings show that the outcomes of strategizing through doing and/or thinking, manifested in milestones and progress markers like joining accelerators, generating revenue, achieving profitability, and gaining audience traction, contribute relatively little to the overall variance in venture valuation in the early stages. However, these predictors grow considerably more influential as the venture matures.

Throughout its three chapters, this dissertation reveals that new technologies such as audience data analytics and innovative support systems like seed accelerators, which change the way entrepreneurs strategize through ‘doing’ and ‘thinking’, are generally linked to positive resource mobilization outcomes. However, it also demonstrates that the results of entrepreneurial strategizing, as evidenced by achievements and progress markers, have limited explanatory power in comparison to other critical determinants of resource mobilization and venture growth.

# Chapter 1. The Effect of Audience Data Analytics Technologies on Startup Prominence

## ABSTRACT

Over the past decade, audience-facing data-driven technologies have seen widespread adoption within organizations. However, the precise efficacy of audience data analytics in enhancing the prominence of new organizations is still unclear. On one hand, these technologies can significantly improve prominence by mitigating bounded rationality—addressing information gaps and reducing the cognitive load associated with managing audience engagement, which is particularly valuable for new ventures lacking historical data and facing informational constraints pertaining to audiences. Conversely, new organizations typically lack the necessary capabilities to circumvent the unique cognitive biases introduced by data analytics, thereby reducing their ability to extract its full potential utility. Moreover, data analytics may offer limited informational value to new organizations as they primarily focus on mirroring external data rather than fostering the generative thinking required for effectively gaining audience in entrepreneurship. This study examines 6,903 U.S. new ventures interacting through web interfaces to explore the effectiveness of web-based audience data analytics in enhancing venture prominence on the web. Using a novel time-series cross-sectional matching method, I show that these technologies are associated with a 28-52% increase in web prominence and a 1-10% decrease in its volatility, demonstrating that data analytics not only boost but also stabilize prominence. Further analysis reveals that data analytics are particularly effective in more typical market spaces, achieves greater prominence when accompanied by A/B testing, and influenced by the founding team's prior employer entrepreneurial prominence.

# INTRODUCTION

Building organizational prominence—a process of augmenting the degree of collective organizational attention and awareness garnered from external audiences (Burton, Sørensen, & Beckman, 2002; Ravasi, Rindova, Etter, & Cornelissen, 2018; Rindova, Williamson, Petkova, & Sever, 2005)—brings a wealth of benefits to new organizations. The process starts with attracting initial attention and recognition from the audience, which opens up the pathway for organizations to be evaluated (Fiske & Taylor, 1991; Roberts & Lattin, 1997), thereby helping establish perceptions of quality (Petkova, 2016; Petkova, Rindova, & Gupta, 2013), and mobilizing critical resources necessary for organizational growth (Clough, Fang, Vissa, & Wu, 2019; Rindova, Petkova, & Kotha, 2007; Pollock & Gulati, 2007).

Despite its advantages, acquiring and growing awareness to achieve a critical mass of prominence from the ground up is a challenging task. Many new organizations encounter difficulty in securing the initial awareness and interest from key resource providers due to the lack of resources and track records of performance (Aldrich, 1999; Aldrich & Fiol, 1994; Blank, 2020; Stinchcombe, 1965). Furthermore, the challenge does not conclude with the successful attainment of initial/early attention and engagement. Newly established organizations confront the challenge of growing their prominence (Petkova, Rindova, & Gupta, 2008; Rindova et al., 2007). At different stages of the venture, entrepreneurs must navigate distinct audiences to effectively mobilize resources necessary for growth (Fisher, Kotha, & Lahiri, 2016; Fisher, Kuratko, Bloodgood, & Hornsby, 2017).

As the organization evolves, managing the numerous audiences becomes increasingly burdensome due to the cognitive constraints of entrepreneurs, which limit their capacity to effectively address the multitude of audience expectations/demands while balancing other critical tasks necessary for growth (Covin & Slevin, 1997; Zuzul & Edmondson, 2017). In response to the acute challenges around information processing constraints, organizations have traditionally mitigated these issues by creating organizational structures like task hierarchies to improve information processing efficiency (March & Simon, 1958; Sah & Stiglitz, 1986). Additionally, standardizing practices and routines help further reduce the decision-makers' cognitive burden (Cyert & March 1963; March & Simon, 1958; Nelson & Winter, 1982).

Despite these adaptations, the limitations persist for organizations like new ventures that do not have the experience or resources required for intricate information architectures (Mintzberg, 1989; Russell Merz & Sauber, 1995). Simultaneously, advancements in digital and web communication technologies have significantly broadened the reach/accessibility of organizations, exposing them to a larger and more diverse audience than ever before (Alaimo & Kallinikos, 2022). As a result, modern new organizations face a unique situation of managing a wider-ranging audience of greater numbers than ever before. However, the past decade has also seen an increased emphasis on leveraging data-driven intelligence as a means to addressing these challenges (Brynjolfsson, Hitt, & Kim, 2011; Brynjolfsson & McElheran, 2016).

Specific to managing external audience attention and engagement, the role of data driven decision-making technologies has become increasingly prevalent for organizations (Brynjolfsson et al., 2011; Brynjolfsson & McElheran, 2016). In particular, *audience data analytics technologies* help collect and analyze digital footprints or ‘eyeballs’, providing a systematic and partially automated means to gathering information on audiences (Bailey, Faraj, Hinds, Leonardi, & von Krogh, 2022; Berman & Israeli, 2022). For example, Google Analytics, a widely adopted web-based audience analytics tool provides insights into visitor profiles for owners of web domains (e.g., entrepreneurs). By collecting and aggregating data on metrics like 'days since user last visited' and 'unique product purchases by user' and presenting them in user-friendly interfaces, the software offers an organized summary and breakdown of the visitors to any given web domain (Mokalis & Davis, 2018). In essence, these technologies serve as information receptors (see Daft & Weick, 1984), aiding in the collection and interpretation of data on audience attention and engagement to comprehend the audience's identity, preferences they reveal, and behaviors they display.

In the recent decade, these technologies have become particularly relevant for new organizations with a robust online presence due to the large amounts of digital footprints and exhausts that allow to capture audience attention in greater scale and abundance (Bailey et al., 2022; Ewens, Nanda, & Rhodes-Kropf, 2018). At the same time, advancements in algorithms embedded in these technologies enable the identification of increasingly complex patterns in data, yielding more nuanced insights than before (Choudhury, Allen, & Endres, 2021). However, despite its prevalence, the precise impact of audience data analytics on assisting new organizations manage their prominence among audience remains unclear, reflecting two contrasting perspectives.

On the one hand, audience data analytics can be *highly effective* in growing prominence due to informational and cognitive benefits. From the perspective of *bounded rationality* (e.g., March & Simon, 1958; Simon, 1955), these technologies can help mitigate the severity of information incompleteness and alleviate the cognitive/computational burden that is traditionally placed on humans in managing (acquiring, retaining, and growing) audiences. Given the acute constraints of bounded rationality that entrepreneurs face—such as limitations in information gathering, susceptibility to cognitive biases, and often a lack of access to comprehensive data (Cohen, Bingham, & Hallen, 2019; Eggers & Song, 2015; Hallen & Pahnke, 2016)—data analytics emerges as a particularly valuable tool for new ventures. This technology can have a particularly profound effect on new organizations as they often do not have a substantial history to leverage experiential insights as alternative or complementary sources of information (Blank, 2013; Blank & Dorf, 2005; Murray & Tripsas, 2004; Stinchcombe, 1965).

However, new organizations face distinct challenges and operate under unique circumstances that complicate their use of data analytics to manage audience attention. These factors can diminish their ability to extract valuable insights and/or result in obtaining information of marginal utility. First, while using data analytics helps mitigate certain cognitive biases in general, new ventures often lack the capacity to counteract novel biases specific to data analytics. One notable bias is the *measurability bias.* This bias occurs when the organizational focus revolves around what is easily 'measurable', which does not necessarily equate to what is 'important' (Muller, 2008). Data analytics constructs a digitized version of reality using what can be captured through technological means (e.g., cookies, pixels, and tags). This shift can lead users to place undue emphasis on quantifiable (yet not always important) metrics, driven by the constraints of data analytics technologies in capturing the full spectrum of important data. (March, 2006; Mazmanian & Beckman, 2018; Muller, 2008; Ridway, 1956). The decoupling between what is measurable versus what is important can be particularly prevalent in new organizations, which do not prioritize or have the resources to build/maintain custom analytics software that can offset this issue (First Round Capital, n.d.; Poppo & Zenger, 1998).

Simultaneously, what is measured is not fully reflective of the complex external environment (Albin & Foley, 1998; March, 2006), leading to potential reductionistic insights of inaccurate/imprecise nature regarding the audience behavior and preferences. Given that ventures lack experiential knowledge to assess and triangulate the validity or relevance of information generated via these technologies, the potential of accepting subpar insights without proper scrutiny is greater (Allen & Choudhury, 2022; Choudhury, Starr, & Agarwal, 2020; Dane, Rockmann, & Pratt, 2012).

Second, the insight gained via data analytics may be of limited value for new organizations. The *theory-based view* (Felin, Gambardella, & Zenger, 2023; Felin & Zenger, 2009; Felin & Zenger, 2017) suggests that efforts to mitigate limited rationality by accounting for information on audience behavior does not necessarily yield valuable learning opportunities for new organizations. This perspective suggests that in the context of entrepreneurship, solutions often demand generative thinking or—*creating* an answer—rather than simply *retrieving* an existing answer from the external environment (Felin, Koenderink, & Krueger, 2017). Thus, data analytics technologies, which are primarily designed to mirror and retrieve data of the external environment, can have limited utility in helping grow visibility and engagement in new ventures. Put short, simply *knowing* the audience well may not yield enough of a value.

Thus, the answer to the question of '*Do audience data analytics technologies help achieve greater prominence in new ventures, on average?'* remains unclear. The opposing perspectives suggest a need to understand the precise magnitude of data analytics' impact, since drawbacks may offset the advantages, thereby altering the overall net effect. Additionally, this prompts further inquiry into whether there are conditions where the positive effects are significantly greater (or negative effects are smaller). To answer these questions, I examine 6,903 U.S. new ventures whose audience engage with the organization through web interfaces (i.e., interaction through venture's website). In this setting, I assess whether using web-based audience data analytics technologies that detect, collect, and analyze the behavior, preferences, and actions of the audience are helpful for new ventures in gaining greater prominence on the web. In evaluating the technology's efficacy, I employ a novel time-series cross-sectional matching method developed by Imai, Kim and Wang (2023).

I find that utilizing data analytics technologies is associated with an average of 28-52% increase in web prominence, measured via the number of web crawls conducted by the Internet Archive. Furthermore, I find that the utilization also correlates with a 1-10% decrease in the standard deviation of web prominence in the 3 quarters following utilization. Together, these results show that audience data analytics technologies help not only achieve greater prominence but also stabilize the degree of fluctuation in the process.

Additionally, I explore the conditions under which audience data analytics usage is more effective/is of greater value. At the *environmental level*, I find that ventures founded in well-defined, conventional market spaces benefit more compared to those situated in atypical spaces. Simultaneously, however, the difference in outcome between usage and non-usage is greatest in least conventional spaces. Observing the venture *founding team level* contingency, I find that teams that have worked for employers with greater entrepreneurial prominence are associated with greater data analytics effectiveness. Finally, at the *strategic action* level, find that ventures that implement experiments (i.e., A/B testing) alongside data analytics experience greater prominence compared to when using data analytics alone.

My research contributes to the organizational reputation literature. Extensive work has examined how organizations gain prominence through outbound activities and communication like securing favorable media rankings, affiliating with high-status entities, obtaining third-party certifications, and using distinct linguistic framing (e.g., Deephouse, 2000; Fombrun & Shanley, 1990; Pollock & Gulati, 2007; Rindova et al., 2005; Rindova et al., 2007; Stuart, 2000). My study, in contrast, focuses on the significance of managing *inbound* information—behavioral data of the audiences—via data analytics as an additional strategy for gaining and enhancing prominence. I also make contributions to the data-driven decision making (DDD) literature, which has primarily focused on established organizations. I highlight that new organizations, despite facing significant challenges in effectively leveraging data analytics, still achieve positive outcomes. Lastly, I contribute to the literature on resource mobilization by exploring how non-social (and rather, technical) routes are another valuable means to which ventures acquire information needed for organizational growth.

# THEORY

## New Venture Prominence

As one of the two dimensions of organizational reputation, *prominence* reflects the degree of attention and awareness the organization garners in the collective minds of the audience (Ravasi et al., 2018; Rindova, Williamson, & Petkova, 2010; Rindova et al., 2005). To gain prominence, organizations have to initially seek out catalysts that help it become visible and recognized (Fiske & Taylor 1991, Pollock, Rindova, & Maggitti, 2008; Tversky & Kahneman, 1973). In exploring the catalysts that promote greater organizational prominence, prior research identifies influential third parties, including *institutional intermediaries* and *high-status actors*, as key sources.

For instance, Rindova, Williamson, Petkova, and Sever (2005) highlight how mainstream media outlets can affect an organization's prominence through evaluations, such as rankings and ratings. Additionally, the authors underscore the significant impact of expert intermediaries' certifications on an organization's visibility. Specifically, they illustrate that business schools gain greater prominence among recruiters when their faculty publish in leading academic journals, with these publications acting as certifications. Similarly, Pollock and Gulati (2007) demonstrate the significant influence of specialized market intermediaries. Analyzing 404 IPOs between 1995-

2000, they find that the number of analysts covering a public company correlates with the amount of visibility it receives from potential strategic partners. Extending these insights, Pollock, Rindova, and Maggitti (2008) explore the dynamics between different intermediaries, emphasizing how one significantly influences another in determining how much attention is allocated to which organizations. For instance, they note that the media's focus is heavily swayed by the presence and recency of actions taken by investors.

Additional research underscores the significant role high-status actors play in enhancing organizational prominence. For example, Pollock and Gulati (2007) find that a firm's affiliation with prestigious investors (VCs) and stock underwriters during its IPO marginally increases the collective attention it receives from potential future alliance partners. Additionally, Rindova, Williamson, Petkova, & Sever (2005) find that universities that hire faculty with degrees from prestigious institutions are associated with greater prominence among peer schools.

Furthermore, studies show that organizations engage in *agentic actions*—deliberate external facing efforts that are observable to the audience—to gain visibility and attention. For example, ventures that perform a greater degree of market actions, including symbolic (i.e., actions invoking cultural/institutional resources) and innovative actions (i.e., new service development), gain greater visibility (Rindova et al., 2007). In addition, ventures engaging in sensegiving activities intended for strategic legitimation purposes (e.g., updating websites and adding informational links) attract media attention, impacting the overall visibility of the firm across a broader set of audiences including investors (Petkova et al., 2013).

Despite numerous studies focusing on whether outbound communication and strategic actions help capture audience attention and engagement, research on the value of *inbound* flow of audience information in gaining prominence is less explored. This void is surprising given that prominence building is interactive. While the process of building prominence is initiated by organizations sending out signals to gain attention and visibility (Fombrun & Shanley, 1990), the audience then responds to the stimuli (e.g., in the form of direct feedback provision, endorsement, actions that reveal preferences, inaction etc.). In turn, this information can be used by organizations to interpret, evaluate, and adjust future actions and decisions to gain a greater audience. This cycle is reflected in the interpretative systems view (Daft & Weick, 1984), which describes organizations as social systems engaged in collecting, interpreting, and utilizing external information for reducing uncertainty and fostering learning of the external environment (Cyert & March, 1963; Daft & Weick, 1984; Fiol & Lyles, 1985; Huber, 1991; Levitt & March, 1988). From this perspective, the collection and interpretation of information flowing into the organization is central to an organization's survival and success as much as outbound communication and action.

In utilizing audience information, traditional interpretation systems that rely on humans to collect and interpret data are increasingly met with greater challenges due to the overabundance of data (Bailey et al., 2022), making it more difficult for managers to effectively process and utilize this information in its entirety (March, 1978; Simon, 1947; Simon 1955). However, simultaneously, there have been new transitions – from the traditional human-centric interpretative systems approach to a more data-driven machine intelligence-based approach where data analytics technologies assume critical roles in collecting and interpreting external information in lieu of or complementing the human counterparts. These technologies, acting as sophisticated receptors of external data, have become an important component of managing the nuances of organizational prominence—by capturing and analyzing data on attention and engagement of audiences (Berman & Israeli, 2022).

## Effect of Audience Data Analytics on New Venture Prominence

Data analytics are broadly categorized as data-driven technologies that help inform decision-making and action (Brynjolfsson et al., 2011; Brynjolfsson & McElheran, 2016; Brynjolfsson & McElheran, 2019). These technologies help capture, process, and organize data regarding individuals/groups/organizations/systems to equip decision-makers with relevant information (Berman & Israeli, 2022; Conti, de Matos, & Valentini, 2023; Mokalis & Davis, 2018; Wu, Hitt, & Lou, 2020). Due to the increased digitization of everyday objects and organizations (e.g., smart phones, digital warehouses) that emit large amounts of digital exhausts, the application of analytics technologies have become more commonplace and central to organizations (Bailey et al., 2022; Ewens et al., 2018; McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012).

Among the wide variety of data analytics, audience data analytics technologies are particularly relevant to managing organizational prominence. Audience data analytics presents the user with a wide variety of longitudinal and cross-sectional observations regarding the audience (and more broadly, the platform in which audiences engages with the organization), arranged in a systematic and organized manner through visuals, tables, and statistical language (Mokalis & Davis, 2018).

### *The Benefits of Audience Data Analytics*

The concept of bounded rationality acknowledges the limitations of individuals and organizations in being able to make perfectly rational decisions (March & Simon, 1958; Simon, 1955). The limitations largely stem from imperfect and incomplete information gathered by decision makers who are simultaneously limited in their cognitive capabilities in avoiding cognitive biases (Fiske & Taylor, 1991; Gavetti, Levinthal, & Ocasio, 2007; March, 1978; Simon, 1947, Simon, 1955; Tversky & Kahneman, 1974). Furthermore, these limitations persist due to environmental-level risk and uncertainty, where the consequence of each alternative is often unknown/uncertain for any given individual/organization (March & Simon, 1958; Simon, 1964; Tversky & Kahneman, 1980).

Boundedness of rationality is especially profound in new ventures, marked by several distinct challenges. New ventures often start with limited history and accumulated knowledge (Blank & Dorf, 2005; Murray & Tripsas, 2004), providing little datapoints for reference when evaluating strategies and decision-making. Even with founders having extensive prior founding experience, the value of such experience can diminish as it is less applicable to current venture conditions (Parker, 2013). Additionally, ventures often have sparse social networks which limits their ability to processing information with great accuracy (Hallen, 2008; Hallen & Pahnke, 2016). Internally, new ventures face the tricky task of managing numerous tasks simultaneously, such as gaining legitimacy and fostering organizational learning, leading to frequent cognitive overload (Zuzul & Edmondson, 2017). Thus, ventures rely on external support and advice to mitigate issues regarding bounded rationality (Cohen et al., 2019; Hallen, Cohen, & Bingham, 2020; Miller, O'Mahony, & Cohen, 2024).

*Informational Value of Data Analytics* In light of these limitations, audience data analytics presents avenues to alleviate the constraints of bounded rationality in new ventures by providing both informational and computational benefits. First, audience data analytics mitigates the severity of audience information incompleteness/imperfections. Computer assisted information technologies like audience data analytics help to acquire accurate and time-relevant information in greater quantities and precision (Berman & Israeli, 2022; Brynjolfsson & McElheran, 2019; Huber, 1990). Thus, while data analytics, serving as an information collecting system, does not directly enhance the cognitive capacity of entrepreneurs/new ventures to allow greater amounts of information to be processed, they enable greater *potential* amounts of information that can be retrieved because they serve as memory systems with superior cataloguing/indexing functions (e.g., friendly user interface and logs that allow easy retrieval of information) (Mokalis & Davis, 2018; Simon, 2013). As a result, the severity of 'satisficing' behavior in decision-makers (e.g., entrepreneurs), which involves reduced efforts in collecting information and thus utilizing a smaller/fewer choice sets for decision-making (Gavetti et al., 2007; Simon, 1955), can be mitigated. Additionally, the cataloguing/indexing functions in analytics can help mitigate common cognitive biases like the availability biases where greater attention is given to information that is more accessible/available (Kahneman, 2013; Tversky & Kahneman, 1973). These technologies can allow for a broader range of information to be browsed and presented with lower cognitive effort, reducing the tendency to prioritize readily available data that is easier to retrieve.

*Computational Value of Data Analytics* Second, data analytics can alleviate the computational burden that is traditionally placed on entrepreneurs in managing (acquiring, retaining, and growing) audience. These technologies are not only adept in collecting information but also in processing it in a systematic manner through engineering and statistical protocols (Murray, Rhymer, & Sirmon, 2021; Simon, 2013). Complex calculations that require extreme brainpower can be efficiently and accurately calculated via data analytics (Mokalis & Davis, 2018). Furthermore, data analytics presents the processed data in organized formats such as tables, statistics, and visuals to make information easily digestible (Berman & Israeli, 2022). In theory, such presentation allows users to better synthesize complex relationships and detect emergent patterns/connect the dots to derive generalized theories and conjectures regarding the audience with greater ease, ultimately contributing to easier comprehension of the processed information (Arthur, 1994; Wu, Lou, & Hitt, 2019). Consequently, using audience data analytics can reduce the cognitive burden of ventures in terms of both processing and comprehending critical information surrounding its audience such as why the firm is not gaining enough attention/engagement or how the organization can grow its prominence.

These benefits around memory and processing can be particularly profound in understanding and strategizing how to manage (acquire and grow) attention/engagement for new ventures given that the counterfactual case is relying solely on intuition or past experience as a reference. The latter case can be particularly detrimental since new ventures often have little or no historical performance data to make informed interpretations or assessments of their audience, especially in their early stages (Blank & Dorf, 2005; Murray & Tripsas, 2004). Even with founders having extensive prior founding experience, the value of such experience can diminish when applying the learning to a different (current) venture (Parker, 2013). Simultaneously, past founding experience can create biases that lead to misattribution of causal linkages (Eggers & Song, 2015), making intuition or experience-based knowledge less useful information derived via data-driven technologies.

### *The Limitations of Audience Data Analytics*

Nonetheless, new organizations are subject to unique challenges in extracting value from data analytics to effectively manage audience attention and engagement. Simultaneously, they face unique circumstances that may diminish the value of information acquired from data analytics technologies.

*The Limited Ability of New Ventures in Utilizing Data Analytics* In efforts to mitigate bounded rationality through data analytics, new cognitive biases can emerge. A notable example is the measurability bias, which redirects organizational focus to aspects that are easily quantifiable (Muller, 2008). Audience data analytics constructs a reality heavily influenced by data that can be captured through technological traces, such as cookies, pixels, and tags (Mokalis & Davis, 2018).

As a result, organizations are able to capture certain types of data—like the duration of a website visit or page interactions through tools such as Google Analytics—but are less effective at discerning others, such as the motivations behind visitor behaviors. This prompts users to place undue emphasis on quantifiable metrics, which do not always align with critical or relevant organizational goals (Halberstam, 1972; March, 2006; Mazmanian & Beckman, 2018; Muller, 2008; Ridway, 1956; Wildavsky, 1979). The discrepancy between what is measured and what is important can be particularly pronounced for new ventures because they often buy off-the-shelf applications over developing customized solutions due to resource constraints. Ventures lack the skillsets necessary to both build and maintain their own technology tools, increasing the likelihood of buying over making (First Round Capital, n.d.; Poppo & Zenger, 1998). As a result, new organizations have restrictions in customizing measurements to better align with metrics that are contextually important.

Moreover, even when there is alignment between what is important and what is measurable, the data captured through technology still fails to fully reflect the complexity of the external environment that surrounds the venture (Albin & Foley, 1998; March, 2006). This limitation often results in overly simplistic insights into audience behavior and preferences. For example, intricate relationships such as mediations and interactions are difficult to recognize due to the technical constraints. Additionally, correlations might be erroneously interpreted as causations, leading to further confusion. Given that new ventures often lack the experiential or practical knowledge needed to assess and validate the relevance of insights generated by these technologies, there is an increased risk of accepting inaccurate or imprecise insights without adequate filtering of information (Allen & Choudhury, 2022; Choudhury et al., 2020; Dane et al., 2012).

*The Limited Value of Insights from Data Analytics* The theory-based view, often referred to as "theory-as-a-strategy" (Felin & Zenger, 2009; Felin & Zenger, 2017), suggests that efforts to mitigate boundedness by gathering information on audience behavior (e.g., how visitors engage on company's website) does not always yield valuable learning opportunities (Felin, et al., 2023).

This perspective argues that technologies like audience data analytics, which are designed to reflect the external environment *as is*, operate well under the premise that an objective answer or truth exists and can be discovered (Felin & Koenderink, 2022). Under this premise, solutions are found via retrieval of information. However, in the context of entrepreneurship, where solutions often demand generative thinking or—*creating* an answer—rather than simply retrieving an existing answer through search, the utility of mere environmental reflections (gathered through data analytics) can be severely limited (Felin et al., 2023). A similar sentiment is shared in March (2006)'s work on rationality-based technologies that derive abstractions of a system using data and protocols (e.g., data analytics). March argues that in environments of high complexity and dynamism like in entrepreneurship (McDonald & Eisenhardt, 2020; McDonald & Gao, 2019; McGrath, 2023; Rindova & Kotha, 2001; Townsend, Hunt, McMullen, & Sarasvathy, 2018), these technologies are severely less efficacious or even detrimental for exploratory activities that do not have a set answer.

## Contingencies in the Efficacy of Audience Data Analytics in New Ventures

In light of the contrasting perspectives on the efficacy of data analytics, I explore whether certain conditions could reduce the impact of its disadvantages, thereby enhancing the overall suggested benefits of data analytics. I explore contingencies that look at the conditions under which 1) the *value of information* extracted from data analytics is greater and 2) the venture's *capability* of utilizing data analytics is greater. I assess three different contingencies at the environmental, founding team, and strategic action-level.

*Market Space Atypicality* At the environmental level, I explore whether the venture's market space atypicality influences the informational value of audience data analytics. New ventures operate in a wide variety of markets ranging from conventional ones that are well defined to distinct spaces that aim to create new demands but are less understood by the audience (Durand & Khaire 2017, McDonald & Eisenhardt 2020; Santos & Eisenhardt, 2009).

The value of data analytics lies in reducing the extremity of bounded rationality of decision-makers by providing greater amounts of more accurate and precise information (Brynjolfsson & McElheran, 2019; Huber, 1990; Simon, 2013). However, in atypical or unconventional market spaces, the value of data analytics may be diminished because increasing audience engagement is a matter of creating interest and demand through persuasion rather than simply listening and reacting to the behavior of the audience (DeSantola, Gulati, & Zhelyazkov, 2023; Felin et al., 2023; Felin & Zenger, 2009). Thus, in market spaces that are defined as more atypical or unconventional, ventures may build greater prominence by generating solutions based on their own internal theory building rather than referencing audience information retrieved through data analytics technologies.

*Experimentation (A/B testing)* At the strategic action level, I explore whether the informational value of data analytics is influenced by accompanying it with experimentation (A/B testing). Data analytics enriches the knowledge base of entrepreneurs because it nudges the user to embrace inductive or abductive reasoning logics to construct theories (Arthur, 1994; Farjoun, 2008; Mantere & Ketokivi, 2013). To elaborate, analytics technologies expose the user to a multitude of observations in neatly organized formats including visuals, tables, and statistics. In turn, this enables the user to easily connect the dots and pattern match across time and/or observations to derive generalized/specific theories on what may be happening or has happened (Nisbett, Krantz, Jepson, & Kunda, 1983).

In contrast, experimentation represents a deductive reasoning logic, which involves using theory and generalized knowledge to establish and test hypotheses. Data analytics can boost the experimentation process by helping formulate more robust theory and generalized understanding of the underlying landscape (Adner & Levinthal, 2008; Eisenhardt & Graebner, 2007), which then directly or indirectly feeds into hypothesis testing and interpretation of outcomes. Entrepreneurs can benefit *directly* by feeding the analytics insights or variables into experimentation via technology integration between analytics and A/B testing tools. Even when direct integration is not implemented, entrepreneurs can *indirectly* benefit by using data analytics to establish a more robust or well-rounded understanding of the entrepreneurial landscape they are situated in, which in turn enriches the theory and generalized knowledge used to conduct experiments (Csaszar & Levinthal, 2016; Gary & Wood, 2011). The benefit of embracing abductive/inductive reasoning via analytics tools alongside experimentation is particularly relevant and useful in the context of entrepreneurship where uncertainty, novelty, and high ambiguity makes using deductive reasoning alone very tricky (Arthur, 1994; Gavetti, 2012; Gavetti, Levinthal, & Rivkin, 2005; Gilboa & Schmeidler 2000; March 2006; Neustadt & May 1986; Santos & Eisenhardt, 2009).

*Prior Employer Prominence* At the founding team level, I explore whether the venture's ability to extract value out of data analytics is influenced by the founding team's background. I specifically explore whether the prominence of the founding team members' past employer(s) impacts the efficacy of data analytics. Burton, Sørensen, & Beckman (2002) suggest that venture teams gain informational benefits in having had worked in organizations where entrepreneurship is more prevalent. As a result of assuming more central network positions, these organizations gain favorable access to diverse information that not only helps identify unexplored opportunities but also know-how on how to pursue such opportunities. This includes information on novel technologies and the practices around it. Additionally, individuals affiliated with organizations of greater prominence gain access to a richer entrepreneurial network consisting of ex-employees that are further along the entrepreneurship route. As a result, these individuals have the opportunity to vicariously learn from others about entrepreneurial best practices and tactics, as well as gaining access to human capital.

Additionally, venture founding teams gain reputational benefits by being affiliated with organizations of greater prominence (Burton et al., 2002). Teams associated with prominent organizations can effectively attract superior talent, both within and beyond their networks, since prominence signals high perceived quality (Spence, 1974). Simultaneously, the probability of a successful match between the founding team and superior talent is likely greater due to the comparable perceived quality between them (Hallen, 2008; Gulati & Gargiulo, 1999; Podolny, 1994).

Thus, new ventures with founding team members who have previously worked in organizations of greater entrepreneurial prominence are likely to extract greater efficacy of data analytics technologies given their 1) direct knowledge on best practices of technology usage in entrepreneurship, 2) access to rich networks to leverage indirect experience on technology usage in entrepreneurial contexts, and 3) the ability to mobilize high quality human capital through network access and matching dynamics. As a result, venture teams with greater prior employer prominence may mitigate issues around the confounding of what is important versus what is measurable with greater competence and/or be able to assess the quality or relevance of information outputs with better judgement.

## METHODS

### Empirical Context & Data

In this study, I track the longitudinal usage of audience data analytics technologies used in the websites of new ventures. Simultaneously, I capture the number of web crawls for each venture's website, which reflects the general prominence of the venture on the web. This is achieved using two databases: the *Internet Archive* and *BuiltWith*.

First, BuiltWith, an online sales lead generator and web technology tracker, offers precise data on web technology usage patterns, including the exact date, time, name, and type of technology used by a web domain (e.g., www.facebook.com). By monitoring each venture web domain over time, I collect data on the venture's use of data analytics technologies from inception to 10 years from founding. BuiltWith sources its data through extensive web crawls, either weekly or quarterly based on the website's traffic. As of March 2024, it covers nearly 100% of currently active domains that end in .com, .net, or .org[2]. Recent studies in strategy and entrepreneurship have utilized BuiltWith to analyze web technology usage in new ventures (e.g., Hallen, Cohen, & Park, 2023; Koning, Hasan, & Chatterji, 2022).

Considering that the data analytics technologies detected by BuiltWith are designed with the primary goal of increasing its presence and success on the web, the ideal outcome measure (i.e., prominence) would also be web-based. Thus, I use the Internet Archive, which stores over 835 billion historical web pages, to capture the web prominence of new ventures as of May 2024[3].

---

[2] https://builtwith.com/data-coverage

[3] https://archive.org/about/

Similar to BuiltWith, the Internet Archive, as an influential 3rd party of the web, scrapes a website's front-end code through repeated crawls. The frequency of these crawls in a given timeframe reflects the web centrality of a website[4], where better-connected websites rank higher on search engines queries, attracting more visits (Arriola, 2021). The centrality of a website is often determined by how many other websites create linkages to or makes citation of the focal website, indicating how cognitively salient a given website is to others in the relevant web space[5]. To assess whether this measure captures the degree of prominence, I test its correlation to web traffic, a commonly used measure for the collective attention or prominence an entity garners (e.g., Burford, Shipilov, & Furr, 2022; Hallen & Pahnke, 2016; Koning et al., 2022). I find moderately high correlation (See *Dependent variable* section for more details).

### Sample

I construct the initial sample using Crunchbase, a database that captures a large fraction of high-potential U.S-based technology ventures that are funded by angel/VC investments. It is often utilized to sample technology-based new ventures in both strategy and entrepreneurship literature (e.g., Conti & Roche, 2021; DeSantola et al., 2023; Hallen, Cohen & Bingham, 2020; Koning et al., 2022). I select ventures that receive any equity-based capital between 2005-2018 since Crunchbase's quality of coverage drastically increases after mid-2000s (Koning et al., 2022).

To make sure that the web-based audience data analytics are relevant and applicable to the ventures chosen for this study, my sample is confined to technology-based ventures in industry categories characterized by a high likelihood of firms maintaining an active web presence, wherein

---

[4] According to Jeff Kaplan, a content management representative at the Internet Archive, the number of web crawls is indicative of how well linked (degree of centrality) the focal website is to other websites (https://archive.org/post/407082/how-often-does-archive-crawl-a-specific-website)
[5] https://www.cloudflare.com/learning/bots/what-is-a-web-crawler/

the website serves as an integral component of the product/service offering[6]. I note that the presence of an active website is a prerequisite for BuiltWith and the Internet Archive to effectively and reliably detect and extract data from it. Thus, I include the following industries: data & analytics, internet services, gaming, software, apps, artificial intelligence, mobile, commerce & shopping, and platform businesses. I exclude industries such as utilities and raw materials, where web-based technologies have minimal relevance or substance. This leads to the initial sample of 9,253 ventures.

Using the web domain of the 9,253 ventures, I collect venture-level longitudinal data on the adoption and abandonment of web-based analytics and A/B testing technologies via BuiltWith. I identify 8,146 out of 9,253 ventures that have at least partial records of adoption/abandonment in the time frame between the date of founding and 10 years from founding. Together, these ventures utilize 1,260 unique analytics technology brands in this period. Simultaneously, data on performance is collected using the Internet Archive (via the Wayback Machine), where I find information on 9,197 out of 9,253 ventures.

Combining the three sources of data (Crunchbase, BuiltWith, and Internet Archive), the final sample consists of 6,903 ventures founded between 2005-2018 that have complete information on data analytics adoption/abandonment dates, number of web crawls, and other covariates. Each venture is observed on a year-quarter basis, resulting in 81,566 unique venture year-quarter observations. On average, this equates to 11.8 quarters (approximately 3 years) of observation for each venture.

---

[6] For comparison, a website of a physical mom-and-pop style store that sells books and only uses its website to announce the opening hours and directions is less likely to be a critical component compared to that of a SaaS business that heavily relies on the website to interact with customers and deploy its product.

## Variables

### *Dependent variable*

*Web Crawl Count (Logged)*: For each venture, I measure the logged frequency of web crawls conducted by the internet archive at the venture year-quarter level. Web crawls serve as a proxy for how prominent a website is on the web. This is because greater web crawls are a natural function of how well cited and/or linked a given website is to other websites – reflecting how aware and recognized a website is to other relevant audience (websites) on the web[7]. This measure was used in a prior study by Hallen, Cohen, and Park (2023).

While other measures of prominence on the web exists such as web traffic (e.g., Burford et al., 2022; Koning et al., 2022), I use web crawls instead of web traffic since most web traffic databases (e.g., Similarweb, SemRush, & SE Rankings) typically offer data only for the most recent 3 years from the date of extraction. Given that my observations begin in 2005, obtaining comprehensive web traffic data is not realistically feasible[8].

### *Independent variables*

*Data Analytics Utilization (Binary)*[9]: Data analytics utilization is a binary measure of whether the focal venture uses any data analytics technology brands for the given year-quarter.

---

[7] While the measure is conceptually similar to extant measures of social network-based prominence that utilizes network centrality (e.g., Ozmel, Reuer, & Gulati, 2013; Ozmel, Reuer, & Wu, 2017), this measure is more expansive in that prominence does not *require* a direct/indirect social tie – simply being cognitive salient would contribute to obtaining prominence. Furthermore, while this measure is similar to extant studies in that it stems from the decision of an influential 3rd party (in my case, the Internet Archive) (e.g., Rindova et al., 2005; Vanacker & Forbes, 2016), the decision-making process of determining prominence is decentralized – the Internet Archive's decision on who is prominent or not (more crawls or less) is largely a function of what a group of other websites consider prominent.

[8] I use SE Rankings, a web traffic data provider, to test the correlation between web traffic and web crawls using a pseudo-random sample of 300 ventures (of which I find data on 282) in my sample in the past three years: 2020, 2021, and 2022. I find that a fraction of ventures has data covering more than 3 years of web traffic, and thus include these observations in testing the correlation. This resulted in 7,122 venture year-quarter observations. I compare the quarterly web crawl count (logged) to the corresponding web traffic size (logged) to find a relatively strong positive correlation (r=0.65). Conducting a simple linear regression, I find that a 1% increase in web crawls is associated with a 1.8% increase in web traffic.

[9] Some data analytics brands go beyond descriptive and diagnostic data analytics to provide experimentation services such as A/B testing without extensive programming or added financial cost. 32 out of 1260 data analytics brands had a/b testing functionality as an additional or peripheral feature. Given that these features may be influencing the efficacy

Here, I use the categorization made by BuiltWith to verify which technologies qualify as data analytics. These data analytics are largely descriptive and/or diagnostic in nature.

***Moderator/Control variables***

*Cumulative web crawl (Count Logged)*: I measure the cumulative sum of the logged web crawl count for each venture year-quarter, representing the cumulative quantity (volume) of prominence achieved by the venture since the date of founding, as well as the maturity of the venture. Serving as a source of data, greater cumulation of web crawls can influence the quality of insights generated by data-driven technologies (Kohavi, Deng, & Vermeer, 2022).

*Raised to date (USD Logged)*: I measure the cumulative logged amount of funds that a venture has raised from investors at the venture year-quarter level. The cumulative amount raised represents the abundance of financial resources in a given venture year-quarter, which can impact the venture's future resource mobilization opportunities and performance outcomes arising from direct/indirect investments in IT or other complementary assets (Brynjolfsson, Jin, & McElheran, 2021; Clough et al., 2019).

*Investor status (Score between 0 and 1)*: A venture's investor status can influence the terms of future resource mobilization and the level of attention garnered from audiences (Hallen, 2008; Hsu, 2004; Pollock & Gulati, 2007; Simcoe & Waguespack, 2011). Thus, for each venture year-quarter, I identify the highest status investor associated with the venture to date. Status is measured using eigenvector centrality, derived from the investor's co-investment syndication network over the previous five years leading up to the focal round (Hallen, 2008; Hochberg, Ljungqvist, & Lu, 2007).

---

of descriptive and diagnostic data analytics technologies, I run the same analysis excluding these hybridized tools. I find nearly identical outcomes (see robustness check section).

*Round type (Dummies)*: Fundraising announcements can attract more attention to a website since entrepreneurs utilize popular media outlets such as TechCrunch or GeekWire and their social media accounts to spread the news (Petkova et al., 2013). Thus, I include round type dummies for each venture year-quarter. Each round type dummy is given a value of 1 if the round takes place in the focal year-quarter. I include the following dummies: *Angel round, Seed round, all Series A to Series I rounds* (respectively)*, Convertible note round, Corporate round, Private equity round, Series unknown, Series undisclosed* (reference category is *No round event at given quarter*).

*Other technology stack (Count):* New ventures utilize a range of web technologies, not just analytics. These include various widgets, frameworks, online experiments (A/B/n testing), and CDNs (Content Delivery Network). These technological choices can impact the efficacy of data analytics (Berman & Israeli, 2022). Following Koning et al. (2022), I count the number of unique web technologies used by a venture in the focal year-quarter excluding data analytics.

*Cumulative quarters of data analytics utilization (Count)*: Data analytics efficacy in the current period can be affected by how long it has been utilized in the past. There can be potential learning effects that stem from familiarity and experience using the same technology over time (Koning et al., 2022; Krakowski, Luger, & Raisch, 2022). Furthermore, some insights generated from prior usage can spill over to influence outcomes in the current period. Thus, I measure how many quarters in the past the venture has utilized data analytics.

In the time-series cross-sectional (TSCS) matching and random effects models, additional controls are incorporated that are absent in the fixed-effects models. First, I include venture *Region (State)* dummies. The five most represented states *CA, MA, NY, TX,* and *WA* are given a separate dummy while other states are grouped under the *other stat*e category. Second, I add the *Broad*

*Industry Category* dummies: *apps, artificial intelligence, commerce & shopping, data & analytics, gaming, information technology, internet services, mobile, platforms,* and *software*.

*Market Space Atypicality:* Ventures differ in the degree to how well defined the market space they occupy is at the time of birth. Some ventures are situated in completely new spaces that are a novel combination of possible market categories whereas others are located in much more conventional and well-defined spaces. To operationalize ventures in terms of their market space unconventionality or atypicality at the time of their founding, I adopt the measure used by DeSantola, Gulati, and Zhelyazkov (2023) and McDonald and Allen (2022), where I compute a proximity index for each pair of categories based on how often they appeared together within a venture's categorical descriptions. The approach starts with the baseline assumption that each firm could belong to multiple categories. I then calculated a proximity score, $P_{ij}$, for each distinct pair of categories, i and j. This involves first, identifying the frequency of links between categories i and j across firms, followed by calculating $P_{ij}$ using the formula $P_{ij} = ½*(C_{ij}/C_i + C_{ij}/C_j)$, where $C_{ij}$ denotes the number of times categories i and j co-occur up until the firm's founding, Ci represents the total occurrences of category i, and $C_j$ signifies the total occurrences of category j. Finally, I computed Typicality (T) as the sum of all $C_{ij}$ combinations divided by half the product of the total number of categories claimed by the venture (N) and N-1. Atypicality was defined inversely as T*-1. I note that ventures with only one category designation are dropped, following DeSantola et al. (2023).

*Additional A/B Testing (Binary)*: I create a binary variable where the value is 1 if the venture utilizes A/B testing alongside data analytics and 0 if only data analytics is used in the given year-quarter. I assess A/B testing utilization by assessing whether the focal venture adopts any A/B testing technology brands for the given year-quarter. I benchmark the list used by Koning et al.

(2022) to determine whether a web technology qualifies as an A/B testing technology or not. However, given that my study has an extended time frame compared to Koning et al. (2022)'s study and thus potentially covering a non-identical list of A/B testing technology brands, I additionally utilize BuiltWith's own classification. BuiltWith analyzes the type of computing language, framework, and location (within the code script) to determine what category each technology falls into. Simultaneously, it also triangulates via user feedback and inquiries for this purpose[10].

*Prior Employer Entrepreneurial Prominence (Logged):* Based on Burton, Sørensen, & Beckman (2002), I measure the prior employer entrepreneurial prominence of a venture using data on the founding team's prior employment experience accumulated at the time of the venture's founding. For each venture founding team member, I calculate the number of other venture founders in my sample who worked for the same company up until the time of the focal venture's founding year-quarter. Then, I take the logged maximum count of this measure among the founding team as a representative value of that venture, benchmarking Hallen, Cohen, and Bingham (2020) and Hallen, Cohen, and Park (2023).

## Statistical model

My study does not claim any strict causal identification. I do not establish an experimental design that randomizes the assignment of treatment. Furthermore, I do not find a viable exogenous shock or set of instrumental variables that uniformly applies to all of the different brands of analytics technologies I assess in my study. While Koning et al. (2022) utilizes the latter two approaches to assess the effectiveness of A/B testing technologies, the same approach is more difficult to apply to data analytics technologies because of the sheer number (1,260 unique analytics brands) of

---

[10] https://builtwith.com/suggest

brands associated with my sample of ventures. I instead triangulate between several different statistical models and specifications as an attempt to mitigate potential endogeneity issues.

My main models utilize time-series cross-sectional (TSCS) matching that derives a DiD (Difference-in-Difference) estimator. This is a novel matching method developed by Imai, Kim and Wang (2023) found in the R package '*PanelMatch*'. This method is a design-based approach (an approach that emulates a randomized experiment using archival data) to measuring treatment effects using panel data (see Hernan & Robins, 2016). Unlike existing matching methods that are largely geared towards cross-sectional data, the TSCS method takes into account the temporal structure of time series data by matching each treated observation (i.e., focal venture) with control observations (i.e., other ventures) in the same time period (i.e., same year-quarter) that have identical history of treatment. Here, history refers to the number of time periods ($L$) preceding the current period ($L$ is determined via user discretion). This matching design aims to prevent pairing ventures with varied histories of treatments that can influence both the likelihood of receiving treatment in the current period and the corresponding outcome (Imai, Kim, & Wang, 2023).

When selecting the value of $L$, Imai and co-authors recommend a careful balance between too long versus too short of a historical comparison in treatment. This is because while longer periods provide greater coverage of delayed or longer-lasting carryover effects, they simultaneously diminish the efficiency of the estimates (i.e., number of matches possible; precision of estimates). In my analyses I set $L = 4$ (4 quarters; 1 year). In other words, I assume any treatment differences in periods exceeding 1 year prior to the current period is unlikely to influence the probability of treatment in the current period and the corresponding outcome. I also run the same analyses with a more conservative assumption where $L = 8$ (8 quarters; 2 years). The results are comparable to when $L = 4$ and can be found in the robustness check section.

While matching based on historical patterns of treatment is helpful for more accurate inferences, other factors that influence the outcome can still exist. Thus, after the matched sets are selected (a matched set consists of one treatment observation and *N* other control observations that have *L* prior quarters of identical history of treatment; *N* varies depending on the treated observation), the TSCS method further refines each matched set so that control observations which are relatively more similar to the treatment observation are given greater emphasis.

Of the several alternative refinement approaches offered by the *PanelMatch* package, I first use similarity-based matching. Similarity between a treated and control observation can be measured using the Mahalanobis distance or propensity scores. For the Mahalanobis distance-based matching (matching based on distance between two points in a multivariate space), I calculate the mean of the standardized Mahalanobis distance between each treated observation and a corresponding control observation across the past *L* time periods ($L = 4$). Distance is calculated using the following covariates: *Web Crawls* (Indicating pre-treatment outcome), *Other technology stack, Cumulative maximum of investor status, Cumulative web crawl, $ Raised to date, Cumulative quarters of data analytics utilization, Round type, Industry category, and Region*. By doing so, each control observation in a matched set receives a distinct distance measure. Next, for each matched set, I select the 10 most similar controls (closest Mahalanobis distance) to the treated observation. These 10 controls are then given a non-zero weight based on the calculated Mahalanobis distance (the closer it is to the treated observation, the greater the weight allocated to the control). All other controls within the same matched set receive no weight (weight = 0), thus effectively being excluded from estimation.

In parallel, I also use the propensity score-based matching refinement approach. This approach utilizes the exact same process and covariates as the Mahalanobis distance-based approach, where I select 10 matched controls most similar to each treatment observation but in terms of propensity scores. The propensity score represents the conditional probability of assignment to the treatment condition, contingent upon observed pre-treatment covariate values (Rosenbaum & Rubin, 1983). In this refinement approach, the propensity score is determined by looking at the historical values of covariates. Specifically, I use the values of covariates from each of the preceding $L$ periods to calculate the score. In essence, the propensity score-based matching approach aims to ensure that the treated and matched control units have similar propensity scores. This helps reduce potential biases in the estimation of treatment effects.

I also simultaneously evaluate a third refinement approach which utilizes weighting – specifically, propensity score weighting. Unlike propensity score matching, the weighting approach allocates weights to *every* control observation in a given matched set[11]. Using the exact same approach to calculate propensity scores as the matching approach, weights are assigned based on inverse probability of treatment weighting (IPTW) for the treatment and control observations in a matched set. In essence, IPTW gives more importance to "unexpected" choices, helping to balance out the differences between the treatment and control observations.

Each of these three refinement processes helps ensure that any differences in outcomes between the treatment and control groups are more likely due to the treatment itself, rather than other factors. Without such refinements, disparities in past outcomes or covariates between the groups could compromise the parallel trends assumption, affecting the credibility of the DiD estimator.

After going through the process of creating and refining the matched sets, the matched sets are then used to calculate the average treatment effect on the treated (ATT) by applying a DiD

---

[11] I note that while all control observations in a matched set are included, they are technically excluded from estimation if weight = 0.

design to assess the contemporaneous (t) treatment effect, as well as the effect at t+1 and t+2 (Imai et al., 2023)[12]. The DiD estimator, a type of weighted two-way fixed effects (TWFE) regression, measures the treatment effect by comparing outcome changes between treated and untreated groups over time. Standard error of estimates is calculated using bootstrapping (1000 iterations)[13].

The TSCS matching method's main benefit is that the estimates derived are more robust to model misspecification compared to two-way fixed effects regression estimators and are applicable to both balanced and unbalanced panel datasets (Imai et al., 2023). Furthermore, compared to synthetic controls which have limitations to how many units receive treatment or the time period in which the treatment is received (Abadie, Diamond, & Hainmueller, 2010; Ben-Michael, Feller, & Rothstein, 2019; Doudchenko & Imbens, 2016), it allows many units to receive treatment regardless of timing and also enables observations to repeatedly pivot back and forth between treatment and non-treatment status over time (Imai et al., 2023).

## RESULTS

### Descriptive statistics

Table 1 provides descriptive statistics of each variable. Data analytics utilization is very common (utilized in 79% of all venture-quarters observed). Figure 1 reveals a more longitudinal view of venture-level utilization trends by showing the % of total ventures in my sample (6,903 ventures) that utilize data analytics in the given quarter from founding. 39% of the ventures utilized data analytics at the quarter of founding. For the ventures that hit 10 years (40 quarters) from founding, utilization becomes approximately 95%.

---

[12] The interpretation of potential causal effects is easiest when the observed period after treatment is shorter (contemporaneous being the shortest). Thus, I focus on three periods t, t+1, and t+2.

[13] Block bootstrapping is used to preserve the temporal dependencies among observations.

Looking at the pairwise Pearson-correlation matrix in table 2, moderate correlation exists between using data analytics in the current year-quarter and the cumulative number of quarters the venture has utilized it ($r$=0.49). Interpreting the moderately strong association, it is possible that data analytics technologies are sticky; the longer the ventures utilize these technologies, the more likely they are to continue utilizing them. While very uncommon in my sample (3%), technology utilization at the venture level is not necessarily permanent post-initial utilization – there are short periods where data analytics technologies are abandoned and then picked up again at a later date at different points in life.

## Time-Series Cross-Sectional (TSCS) Matching Models

Given that the TSCS matching models uses a DiD design[14], I explore whether there is suggestive evidence for meeting various assumptions. First is the parallel trends assumption. The parallel trends assumption stipulates that without any treatment, both the treated and control groups would have exhibited consistent trends in outcomes over time (Goodman-Bacon, 2021). Put simply, any variation in trends between the two groups can be exclusively credited to the treatment, rather than any other influences. If the treated and untreated groups have differing pre-treatment characteristics that are suspected to affect the outcome's progression, this assumption could be challenged (Abadie, 2005). I explore whether there is suggestive evidence of meeting this assumption using the covariate balance function that comes with the '*PanelMatch'* package in R.

Covariate balance is measured for each covariate included in refining the matched sets. It indicates the average difference in the covariate value between the treated observation and that of the weighted average of control observations across all matched sets[15]. For each covariate, the

---

[14] While it adopts the DiD design, the split between treated and control observations are not a result of an exogenous shock or other randomization.

[15] This expresses the difference between a treated unit and its matched controls in relation to how much that specific covariate typically varies among all treated units. If the value is 0.1, the treated unit's covariate value is slightly

balance is then standardized across the treated observations using the standard deviation of the focal covariate (Imai et al., 2023). A greater absolute covariate balance value indicates a greater average difference between the treated versus control observations.

Additionally, this function can be used to compare the dependent variable in the periods preceding treatment in regard to the parallel trends assumption. In my study, I compare web crawl counts for each of the 4 quarters preceding the focal quarter. If the covariate balance of the dependent variable remains relatively constant/stable over time between t – 4 and t – 1 without any sharp breakaways, it suggests that the treated and control groups were on parallel trends before the treatment. Consistency is observed in the models using propensity score-based matching and weighting approaches to refinement (see Figures A2 & A3 '*Web Crawl'* graph), providing some suggestive evidence for meeting the assumption[16]. I note that in the Web Crawl graph for the Mahalanobis distance-based refinement approach found in figure A1, there is a marginal increase in the difference over the 4 quarters preceding treatment, where the increase is a s.d. of 0.15 for data analytics. This means that as time passes from t – 4 to t – 1, the treatment observations have increasingly greater web crawl counts relative to the control observations in each matched set on average, albeit a very minor increase.

To strengthen the inferences drawn from the matching estimators, I additionally verify whether the covariates of the treated versus control observations in a matched set are comparable at t – 4, t – 3, t – 2, and t – 1, respectively. I do this for each of the three models using different refinement approaches (Figures A1, A2, & A3) to find that the majority of covariates show consistent differences across time, reflected in the covariate balance trendline over the 4 quarters

---

different from its matched controls. This difference is small, equivalent to 10% of the typical variation we see among all treated units for that covariate. If the value was 1, it would mean the difference is as large as the full typical variation (or standard deviation) of that covariate among all treated units.

preceding treatment remaining largely flat. Furthermore, for each covariate, the magnitude of the difference is also marginal per observed time period given that the line remains close to the 0 standard deviation point of the y-axis (0 = no difference). Such findings bolster the argument that differences in outcomes between the treated versus controls is more likely due to the treatment itself and not the differences in other observed confounders. However, the issue of omitted variable biases that cannot be measured for this study still remains.

Next, I explore the potential of spillover effects between treatment and control observations where an observation's treatment can influence the outcome value of other observations. I assume that a spillover is less likely, meaning that the usage of data analytics in one venture will not impact the growth outcome of others. This is because there is little value or motivation for a venture to share their extracted insights with others because it is neither applicable to other ventures' situations nor do the internal bylaws allow easy sharing of private information (e.g., customer data) generated using these technologies with outsiders.

However, carryover effects, where past treatment influences future outcomes within a venture, are likely. My main models assume that the effect of utilizing data analytics technologies will influence outcomes concerning the next 4 quarters but not beyond. Due to the realistic potential of treatment effects being delayed by or prolonged for more than 1 year, I run additional robustness checks by applying a more conservative threshold of 2 years (e.g., utilization of data analytics at the start of year 1 influences the venture's outcome up until the startup of year 3). I find no meaningful differences in the estimates (See Robustness Checks Section). I assume 2 years to be reasonable given the average venture age of my sample is 3.2 years.

Table 3, 4, and 5 show the results for the TSCS matching estimators using three different (alternative) refinement methods, *Mahalanobis distance-based matching, propensity score-based* *matching, and propensity score-based weighting*. For each refinement method, the matching estimators reveal the average treatment effect on the treated (ATT) for data analytics usage. Regarding the number of matched sets, I find 3,561 matched sets for data analytics with a mean set size of 169 for each treatment observation. I note that the matched set size is irrelevant for the matching-based refinement approaches (Mahalanobis distance and Propensity score-based matching) since only the 10 best controls are considered (given any non-zero weight) in the estimation stage.

Table 3 reveals the estimate using Mahalanobis distance-based matching refinement. The results show that the contemporaneous effect of data analytics usage is 0.357 (95% CI: [0.318, 0.396]). This means that on average, ventures utilizing data analytics are associated with a 42.9% greater web crawls relative to comparable ventures not utilizing it at time t. At t+1, the effect was slightly smaller at 0.302 (95% CI: [0.251, 0.350]), equating to a 35.3% greater web crawl count. At t+2, the usage is associated with a 27.7% greater web crawl count, reflected in the effect size of 0.245 (95% CI: [0.189, 0.295]).

Next, table 4 shows the estimate using propensity score distance-based matching refinement. Data analytics shows a contemporaneous effect estimate of 0.400 (95% CI: [0.360, 0.440]), which indicates that ventures using data analytics experience a 49.2% greater count of web crawls compared to non-users on average. At t+1, the effect remains largely comparable at 45% – reflected in the effect size of 0.370 (95% CI: [0.323, 0.418]). At t+2, the effect size is noticeably smaller but still substantial at 36.6% increase (0.312; 95% CI: [0.258, 0.364]).

Finally, table 5 shows the estimates for the model concerning propensity score-based weighting refinement. In this model, data analytics yield a contemporaneous effect of 0.418 (95% CI: [0.378, 0.454]). This means that ventures using data analytics are correlated with 51.9% greater web crawl counts compared to comparable ventures that do not use it at the time of analytics usage. For t+1, the effect size remains similar at 48%, reflected in the effect size of 0.392 (95% CI: [0.349, 0.437]). Finally, the effect size at t+2 is 0.326 (95% CI: [0.281, 0.377]), which can be interpreted as venture having 38.5% greater web crawl counts when utilizing data analytics.

For each model and time period, the positive effect of data analytics is not only significant but large in effect size. At the same time, comparing the three models, there is notable range in effect sizes. Depending on the model, At time t, the effect ranges between 42.9-51.9%. At t+1, the range is between 35.3-48% and at t+2, 27.7-38.5%. Given that TSCS matching is a between-firm comparison approach, I also use fixed-effects regression to assess within-firm comparison. The results are consistent and are shown in the Robustness Checks section.

### ***Dependent variable as the Standard Deviation in Web Crawl Counts***

I additionally observe the *standard deviation* in web crawl counts over the next three quarters (including the current quarter) as a measure of outcome. The results across three different versions of the TSCS matching (Mahalanobis distance-based matching, Propensity score-based matching, & Propensity score-based weighting) show a consistent pattern of reduction in the variance.

First, the Mahalanobis distance-based matching shows an estimate of -0.109 (95% CI: [-0.126, -0.090]) at t, -0.110 (95% CI: [-0.133, -0.086]) at t+1, and -0.108 (95% CI: [-0.134, -0.086]) at t+2. These results show that the utilization of data analytics is associated with a decrease in the standard deviation of web crawl counts by approximately 10% consistently in the year following treatment. While the directionality of the other two TSCS models are the same, the magnitude of the decrease is noticeably smaller. In the propensity score-based matching models, the result shows a 3.5-6.2% decrease in standard deviation across the observed time periods. The propensity score-based weighting models show an even smaller result (ranging from 1.4-5% decrease in standard deviation), where the estimate is not significant at t+1 at the 95% confidence level (95% CI: [-0.041, 0.010]). Together, the results show evidence that the utilization of data analytics is associated with a significant reduction in the standard deviation of web crawl counts, albeit the decrease being small at 10% at the maximum.

## Additional Analyses

To explore the contingencies around the utilization of additional experimentation (A/B testing), I utilize a fixed effects model. For the contingencies around market space atypicality and prior employer prominence, I use random effects models due to the nature of the moderators being time-invariant at the venture-level.

*Atypical Market Spaces* Running a random effects model that interacts data analytics utilization and market space atypicality, I find significant positive interaction (B=0.261; p=0.041[17]) (see table 6). Observing the interaction effect in more detail (see Figure 2), I find a unique pattern among ventures that utilize data analytics. In more typical market spaces (prominence being closer to -1), ventures achieve greater web crawl counts on average compared to their counterparts in less typical spaces. However, within these typical spaces, there is only a small difference in web crawl counts between using and not using data analytics. In contrast, the difference is substantial in less typical market spaces – ventures that utilize data analytics perform noticeably better than those that do not.

*Experimentation in Ventures* For nearly all of venture year-quarters in my sample (99.9%), the usage of A/B testing is accompanied by the usage of data analytics but not the other way around. Hence, a moderation effect would potentially yield inaccurate results due to the inherent dependency between the two technologies. Thus, I compare the effect of using both technologies

---

[17] Using cluster-robust standard error. In an alternative specification that does implement cluster-robust standard errors, the p-value becomes p=0.000.

versus just data analytics as the baseline observation – this leads to the assessment of 6,255 ventures. Using a fixed-effects regression, I find additionally utilizing A/B testing alongside data analytics is associated with 8.8% greater number of web crawls (B=0.084; p=0.000) compared when using data analytics without any A/B testing tools (see table 6).

*Prior Employer Entrepreneurial Prominence* Utilizing a random effects model, I interact data analytics usage and prior employer entrepreneurial prominence measured at the time of the venture's founding (see table 6). I find that ventures with founding teams that have work experience in companies of greater entrepreneurial prominence (ex: Microsoft, Google, IBM) are able to utilize data analytics with greater efficacy (see figure 3). The interaction term (B=0.014; p=0.079[18]) indicates that a 1% increase in prominence score is associated with approximately a 0.014% increase in web crawl counts. Applying these numbers to my sample, a founding team with a very high prior employer prominence score (e.g., 5) would be associated with 5.6% greater web crawl counts compared to those with a very low prominence score (e.g., 1). Given that founder data is available for only 4,129 out of the initial sample of 9,253 ventures (45%), I add survey weights to the regression to see whether issues around data representativeness changes the outcome. In creating the survey weights, I use a logit model that predicts the existence of web crawls, data analytics technology utilization, *and* founding team prior employer prominence data based on industry category, state, founded year-quarter, focal year-quarter, dollars raised to date, and venture investor status. I find an even greater positive effect (B=0.023; p=0.000).

## Robustness Checks

First, to supplement the estimates derived via TSCS matching models, I conduct fixed-effects regression analysis to derive within-firm estimates concerning the treatment effect of data analytics

---

[18] Using cluster-robust standard errors. Using an alternative specification that does not implement cluster-robust standard errors, the p-value becomes p=0.025.

usage. I include venture and year-quarter fixed effects and cluster the standard errors at the venture-level. Table B1 in Appendix B reveals the within-firm estimation of data analytics usage. Similar to the results of the matching models, the regression shows that analytics are associated with an effect size of 0.283 (p=0.000), indicating that data analytics utilization is associated with a 32.7% increase in web crawl counts at t+1 – comparable to the lower end of the range observed in the inference made via TSCS matching.

Second, I run the same fixed-effects regression model with survey weights. Survey weights help account for the biases in sample representation relative to the population. They adjust the influence of individual observations in a dataset. Observations that are underrepresented (or less likely to be included) are given greater weight, while overrepresented observations (or those more likely to be included) receive smaller weights. This ensures a balanced representation of all observations in the analysis (Bollen, Biemer, Karr, Tueller, & Berzofsky, 2016). In my sample, 2,350 out of 9,253 ventures are not included in the final sample due to missing web crawls and data analytics technology utilization data. Thus, the 6,903 ventures in my final sample may not accurately depict the nature of the initial sample of 9,523 ventures. Thus, I create survey weights to adjust for the potential misrepresentation. In creating the survey weights, I run a logit model that predicts the existence of both web crawls *and* data analytics technology utilization data using the venture's broad industry category, state (region), founded year-quarter, current year-quarter (accounting for temporal trends), maximum cumulative investor status for venture, and raised to date USD (logged) as the predictors. I trim the weights at the 1 and 99% percentile via winsorization (limiting extreme values without omitting observations at 1 and 99% percentile). The results (see table B2 in Appendix B) show no meaningful change in the coefficient of data analytics usage (B=0.278; p=0.000).

Next, I conduct a series of analyses to explore the sensitivity of the main models. First, I run models excluding 32 data analytics technologies that have A/B testing as an add-on feature. This is to assess whether the inclusion of these technologies confounds the effect of descriptive and diagnostic data analytics with the effects of A/B testing given the hybrid nature of the 32 brands. The exclusion leads to only marginal changes to the estimates. Results are shown in tables B3, B4, and B5 in Appendix B. Lastly, I change the length of treatment history (*L*) used to create matched sets from 4 quarters (1 year) to 8 quarters (2 years). I find that the results are very similar (See tables B6, B7, & B8 in Appendix B).

## DISCUSSION

In this paper, I sought out to understand the effect of data analytics on organizational prominence in new ventures. Across different models that observe differences within and between-ventures, I find consistent results: Utilizing web-based audience data analytics is positively associated 35-48% greater organizational prominence on the web. Simultaneously, I find that the variance in organizational prominence is reduced by 1-10% in the 3 quarters following data analytics utilization, thus reducing the probability of experiencing extreme outcomes (i.e., drastic fall or peaks in prominence).

Additionally, I find that data analytics efficacy is highly contingent in the context of deployment. First, it is contingent on the degree of atypicality of the market space the venture is situated in. While data analytics is more effective in more typical market spaces, ventures are able to achieve comparable prominence without – potentially due to other sources of information being sufficient in strategizing how to increase salience and engagement among audience. These results align with the theory-based view that obtaining precise mirroring of the external environment is less useful for ventures in markets that aim to provide novel or radical value (Felin et al., 2023).

On the contrary, while the average efficacy of data analytics is notably lower in more atypical market spaces, utilizing these tools gives substantially greater boosts in prominence versus not using them. These findings indicate that while audience data may be less valuable in atypical market spaces compared to typical ones, the disadvantages of lacking such information are considerably greater too.

Next, the accompanying of A/B testing, representing experimentation, alongside data analytics is correlated with 9% greater venture web prominence. This indicates that A/B testing provides distinctive value from data analytics by providing additional insights or validation that cannot be captured through data analytics alone. This suggests that data analytics can be great for uncovering trends and patterns within existing data, but A/B testing provides experimental evidence that can confirm or refute hypotheses generated by analytics – offering an opportunity to validate the trends and patterns with greater rigor in precision and accuracy. Such evidence aligns with work that highlights the complementary relationship between theory building and testing as a means to developing entrepreneurial strategies and decision-making (e.g., Agarwal et al., 2023; Miller et al., 2024).

Lastly, I find that prior employer prominence of the venture founding team influences audience data analytics efficacy. My result indicates that teams with greater prior employer prominence are associated with greater effectiveness in data analytics usage. While not tested in this study, there may be several different mechanisms at play in tandem. First, the founding team may be directly demonstrating and disseminating better practices of data analytics adoption/usage within the organization (Burton et al., 2002). Second, founding teams with greater prior employer prominence may have preferential access to and greater likelihood of success in matching with human capital (e.g., employees) that can deploy these technologies more effectively due to reputational benefits (Burton et al., 2002; Gulati & Gargiulo, 1999; Hallen, 2008; Podolny, 1994; Spence, 1974).

## Contributions

First, my study offers contributions to the notion of prominence in the *organizational reputation* literature. While much of extant literature explores how prominence is initial gained and grown through *outbound activities* such as gaining favorable rankings in the media, acquiring certifications from 3rd party intermediaries, being affiliated with high-status actors, and linguistical framing to grab attention (e.g., Deephouse, 2000; Fombrun & Shanley, 1990; Pollock & Gulati, 2007; Rao, 1994; Rindova et al., 2007; Rindova et al., 2005; Stuart, 2000; Taeuscher & Rothe, 2023), my study explores the efficacy of audience data analytics—a tool specialized in managing the *inbound* information (data on audience behavior/preference/feedback) relevant to gaining and growing prominence. My study, in tandem with prior studies, helps draw a more complete cycle of prominence building in the early stages of organizing by revealing the importance of utilizing the behavioral data of the audience as a form of feedback to grow in prominence.

To the *data-driven decision making* (DDD) literature that has predominantly focused on established organizations and facilities[19], I add that new organizations, despite having greater limitations in resources to effectively harnessing the benefits of data analytics compared to established/larger counterparts (Brynjolfsson et al., 2021; LaValle et al., 2011), reap net positive effects when used on average.

At the same time, I also find that ventures are subject to factors that are outside of their control in determining the efficacy of data-driven technologies. I find that effectiveness of data analytics

---

[19] This is with the exception of Koning et al. (2022) that observes A/B testing in new ventures. While Conti et al. (2023) explores small firms, the study does not specify whether they are young firms.

is in part predetermined by the *initial conditions* at an organization's inception. This is reflected in the fact that accumulated social and human capital of the founding team by the time of founding is significantly associated with data analytics efficacy. Such finding complements the DDD literature which has predominantly focused on the factors determined *after* the establishment of the organization such as its structure/design, cumulated financial capital, and managerial capabilities (Allen & Choudhury, 2022; Anthony, 2021; Brynjolfsson et al., 2021; Choudhury et al., 2020; Ghosh, Thomke, & Pourkhalkhali, 2020; Lebovitz, Lifshitz-Assaf, & Levina, 2022; Lebovitz, Levina, & Lifshitz-Assaf, 2021). My findings help provide a more balanced view of how much agency an organization has in influencing the efficacy of data-driven decision-making and the technologies that underlie it.

Finally, I contribute to the literature on *entrepreneurial resource mobilization* by exploring the value of utilizing non-social conduits to acquire informational resources. While a substantial amount of literature has focused on *social* means (i.e., social networks) as conduits for valuable informational/knowledge resources (e.g., Sorenson & Stuart, 2001; Hallen, Davis, & Murray, 2020; Yli-Renko, Autio, & Sapienza, 2001), I find there is also significant value in the information acquired through technological means such as web-based data analytics technologies that does not require any social ties for inflow to occur. Additionally, I also find that the value of informational resources is dependent on the nature of the market itself. Market spaces of greater typicality yield greater value from data analytics than more conventional counterparts. Despite this, using data analytics is more crucial in atypical spaces given that non-usage is likely to yield significantly worse outcomes than in typical spaces.

Finally, this study employs a novel matching technique developed by Imai et al. (2023) specifically designed for use with panel datasets. Considering the substantial informational benefits of panel datasets and the existing constraints of matching methods (i.e., predominantly suited for cross-sectional data), this new approach to matching offers a highly effective solution for a common frustration amongst strategy and entrepreneurship scholars.

However, there are strict limitations to this study. It does not directly observe the mechanisms through which data analytics enhance decision-making, due to the archival nature of the dataset. Future research designs might consider examining changes in the cognitive processes of decision-makers to develop a unique cognitive map of the entrepreneurial landscape (Adner & Levinthal, 2008). Alternatively, exploring how task structures facilitate the effective digestion and utilization of data could prove valuable (Lee, 2022; Salancik & Pfeffer, 1978). Further investigation into these mechanisms is recommended to deepen the understanding of data analytics' effectiveness in entrepreneurial settings.

## CONCLUSION

This study explores the efficacy of data analytics technologies used by new ventures in achieving greater prominence. I find that new ventures, despite having limitations to resources and susceptibility to cognitive biases, are able to reap positive benefits when using data analytics technologies, on average. At the same time, I show that environmental, founding team, and strategic action-level contingencies affect how effective data analytics tools can be by influencing either how skillfully the tool is utilized or how useful the derived information from the tool is. My findings imply that utilizing data-driven technologies to manage organizational prominence requires careful scrutiny of the boundary conditions in which it is deployed for extracting greater efficacy. Given that entrepreneurial ventures must make trade-offs in which activities/tasks to prioritize, understanding when data analytics are more effective can be highly beneficial in deciding how to incorporate data analytics into other critical tasks necessary for organizational growth.

# Chapter 2. A Variance-Decomposition Analysis of Startup Valuation Drivers: The Surprisingly Limited Role of Founder Attributes and Venture Progress

## ABSTRACT

Numerous studies suggest differences in venture valuations are greatly influenced by factors related to the environment, attributes of the founding team, and the venture's progress. Yet it remains unclear which factors are more important, by how much, or how their influence may shift as ventures evolve. To address these questions, we conduct a variance decomposition analysis using the modern Shapley-Value Approach (SVA) on 1,737 U.S. ventures at three different funding stages – initial, early follow-on, and later rounds. We find that founder attributes and venture progress explain a surprisingly small amount of valuation differences at each stage. In addition, we find that the initial resource mobilization outcomes are the dominant predictor of valuations in subsequent rounds. Our findings raise important questions about the extent to which more experienced founders have ongoing resource mobilization advantages and their ability to capture value and preserve control relative to the providers of capital to their ventures.

# INTRODUCTION

Resource mobilization is vital for venture survival and growth (Dencker & Gruber, 2015; Eesley & Roberts, 2012; Eisenhardt & Schoonhoven, 1990; Hashai & Zahra, 2022; Shane, & Stuart, 2002). Resources enable ventures to seek attractive opportunities, build sustainable competitive advantages, and beneficially shape their environments (Aldrich & Fiol, 1994; DeSantola & Gulati, 2017; Gao & McDonald, 2022; Katila, Chen, & Piezunka, 2012; Stam, Arzlanian, & Elfring, 2014). Beyond its practical importance, resources provide a sound theoretical lens for understanding the origins of competitive advantages and the mechanisms nascent firms employ to manage critical resource dependencies (Clough, Fang, Vissa, & Wu, 2019; Hallen, Davis, & Murray, 2020).

However, resources are not free but come at a cost. More specifically, entrepreneurs *trade* equity in a venture to obtain resources from others, especially those seeking substantial capital to pursue high-potential opportunities (Hall & Woodward, 2010). Angel investors and venture capitalists (VCs) provide financial capital expecting their acquired equity stake will produce substantial returns should the venture exit via IPO or acquisition (Mallaby, 2022). In the VC fundraising context, the extent to which entrepreneurs must give up more or less equity in return for a given dollar of capital depends on the venture's pre-money *valuation* (hereafter referred to simply as valuation for brevity[20]), or the agreed-upon price that is assigned to the existing equity in the venture (Gompers & Lerner, 2000; Hochberg, Ljungqvist, & Lu, 2010; Hsu, 2007).

---

[20] In entrepreneurial contexts, it is typically important for there to be clarity whether valuation refers to the valuation right before the investment (the pre-money valuation) or valuation of the venture right after the investment (the post-money valuation). The reason this distinction occurs and is important is that the invested capital typically, especially in earlier stages, becomes the property of the venture and is exchanged for *newly issued shares*. In other words, the investment both dilutes all existing owners and makes a venture more valuable because of the additional capital that it now owns. The pre- and post-money valuation are mechanically related, however, with the post-money valuation simply being the pre-money valuation plus the capital invested. To better isolate how much the existing equity is worth from the amount raised, we focus in this paper on pre-money valuation.

A venture's valuation is determined through a negotiated process between investors and venture founders (Kotha, Shin, & Fisher, 2022). In principle, a venture's risk-adjusted perceived future value and an entrepreneur's relative negotiating power affect a venture's valuation (Hellmann, 2002; Sahlman & Willis, 2009). Ventures perceived as more likely to succeed are in a better negotiating position than capital providers looking to invest. Such ventures can obtain more capital by giving up the same amount of equity or less of their venture ownership for the same amount of capital than others. The converse is true for entrepreneurs with riskier or lower prospects that limits their bargaining power, and as result their ventures compete from a less-resourced position for capital.

A venture's valuation has important implications for various strategic and organizational questions. These include the extent to which ventures can mobilize additional resources and their resulting strategic positioning vis-à-vis competitors, the future value capture of entrepreneurs (and existing shareholders) relative to new capital providers, and the relative control of different shareholders over the firm (Berle & Means, 1932; Hall & Woodward, 2010; Wasserman, 2012; Magnuson, 2022). The following example illustrates our point. Uber and Lyft, founded within two years of each other, compete in the ride-hailing industry with similar business models, providing an illustrative case. Both secured Series A funding in 2011, but the equity they traded for capital differed significantly. Uber received more than twice Lyft's valuation ($49M vs. $23M pre-money valuation) while giving up almost identical equity (approximately 18%). Accordingly, Uber acquired substantially more capital than Lyft for the same percentage of equity ($11M vs. $6M). This, in turn, provided Uber with substantial strategic advantages in investing more in product development, customer acquisition, and expansion to new regions. Moreover, such stark differences in valuations for similar ventures are common (Chatterji, 2009; Hsu, 2004; Hsu, 2007).

Research suggests three broad categories of *observable* factors influencing early venture valuations—venture progress, founding team attributes, and the environment. *Environmental factors* represent the supply and demand dynamics of the broader market that influence venture valuations and include the venture's *industry*, *region*, and *time-varying* factors that impact the supply of investment capital, which in turn impact valuations ventures receive (Eisenhardt & Schoonhoven, 1990; Gompers & Lerner, 2000; Kotha et al., 2022). *Founding team attributes* are observable cues of the entrepreneurial team's human and social capital that influences valuation (Agarwal, Echambadi, Franco, & Sarkar, 2004; Batjargal & Liu, 2004; Baum, Locke & Smith, 2001; Chatterji, 2009; Eesley & Roberts, 2012; Hsu, 2007; Wasserman, 2017). Finally, *venture progress* represents tangible evidence about resolving uncertainty around the venture's prospects, such as having revenues or attaining profitability and gaining sufficient customer traction (Claes & Vissa, 2020; Hallen & Eisenhardt, 2012; Hsu, 2007; Hsu & Ziedonis, 2013).

Despite this extensive research on the observable factors linked to venture valuation, significant gaps remain. First, past research is unclear about their *relative* importance because research has yet to examine all three factors simultaneously. Research examining factors independently has focused on estimating and highlighting absolute effect sizes (i.e., how much changes in certain attributes are associated with a change in valuation) and not their relative importance. This distinction between absolute and relative importance lies in whether an explanation (predictor) is measured by its strength versus the completeness of prediction. The latter measures how much variation in the equity price or valuation outcomes can be attributed to the focal predictor/factor as a percentage (e.g., coefficient of determination), enabling direct comparison with other factors in the model. Knowing both is critical since the large absolute importance of a predictor only sometimes equates to large relative importance and vice versa.

Second, with a few exceptions, research is yet to examine to what extent (magnitude) the relative importance of a factor *changes* as a venture grows. Understanding such temporal shifts in valuation drivers is crucial because they reveal which factors become more/less important when and by how much. And while Fitza, Matusik, and Mosakowski (2009) examine how different factors influence changes in valuation *across* different funding rounds, their study focuses only on changes in valuation as an outcome and assumes that the relative role of each factor remains constant as ventures mature.

To address the gaps and the limitations of past studies highlighted, we focus on the following questions: *What is the relative importance of environmental factors, founding team attributes, and venture progress in the valuations a venture receives? And how (much) does their relative importance change as a venture evolves?* To answer our research questions, we study 1,737 U.S. technology-based ventures with deal rounds between 2007–2022 at three stages—Initial, early follow-on, and later round stages—using unusually rich data on founding team attributes and venture progress for a sample of this size. We measure equity prices using *pre-money valuations*, which reflect the price at which investors purchase equity in a venture for each dollar invested (Gompers & Lerner, 2000; Hochberg et al., 2010; Hsu, 2007).

We conduct a variance decomposition-style analysis using the cutting-edge Shapley Value Approach (SVA) (Sharapov et al., 2021), which incorporates the concept of weighted average marginal contributions from game theory (Shapley, 1953). This approach is widely used in computational sciences, such as machine learning to improve interpretability and promote fairly distributed prediction (Ghorbani & Zou, 2019; Lundberg & Lee, 2017), as well as more recently in management studies (e.g., Kumar, Liu, & Zaheer, 2022). Unlike prior variance decomposition approaches, this modern approach does not require strong (and possibly arbitrary) assumptions about how effects are nested and is not restricted to treating attributes as fixed or random effects repeatedly observed over time (Athey & Imbens, 2019; Choudhury, Allen, & Endres, 2021).

We find that environmental factors have a sizable influence on predicting differences in valuations, collectively explaining 12-16% of the total variance in valuation depending on the financing round. Notable, however, is that in the initial round, most of the effect is due to regional differences. Regarding venture progress, our observable measures explain only 3-4% of valuation differences on the initial and early follow-on rounds. However, we observe a large spike in the later rounds (11.5%). Next, in contrast to the substantial literature (e.g., Dencker & Gruber, 2015; Eesley & Roberts, 2012; Eisenhardt & Schoonhoven, 1990; Hsu, 2007; Shane & Stuart, 2002) examining the impact of founder education and industry and entrepreneurial experience, we find that observable founder attributes explain only 4.7% of the variance in initial rounds. This magnitude remains static during the subsequent rounds, early follow-on, and later. Also surprisingly, we find that the characteristics of past investment rounds (past valuation, past investor status) predict a substantial portion of valuation differences in later rounds (38.62% of early follow-on rounds and 45.95% in later-stage rounds).

These findings offer numerous contributions. To the venture finance and growth literatures. First, our findings highlight that observable founder attributes and venture progress—two factors touted as predictors of various entrepreneurial outcomes—explain a relatively small portion of differences in venture valuations, especially in the early rounds of venture financing. While our study is focused on establishing the relative predictive power of each factor for valuation differences, we believe this raises several substantiative theoretical and empirical questions for future research and possible explanations that need exploring. Second, we highlight the substantial path dependency in valuations across rounds, suggesting that the initial valuation is a unique opportunity for entrepreneurs and ventures. We also highlight the implications of our findings for entrepreneurial strategy and the evolution of ventures' competitive resource positions. Our findings surprisingly indicate that stronger entrepreneur backgrounds might offer less of an advantage than prior literature would suggest while raising questions about the efficacy of early-stage venture progress as an entrepreneurial strategy.

## THEORETICAL BACKGROUND

Venture valuation refers to the inferred value of the collective equity in a venture, as calculated from the price at which investors most recently purchased a fraction of equity in exchange for capital (Gompers & Lerner, 2000; Hochberg et al., 2010; Hsu, 2007). The price paid for equity reflects the fundamental or expected value of the venture given its current/future resources, capabilities, and risk/opportunity sets (Downes & Hienkel, 1982; Hsu, 2007). The valuation process entails a negotiation between the current venture owners (including founders and extant investors) and new (prospective) investors (Davila, Foster, & Gupta, 2003; Hsu, 2007; Kotha et al., 2022). Valuation also has important implications for employees compensated in equity and options, though they rarely have an influence on new valuations.

Although there is no 'one-size-fits-all' valuation process (Gompers, Gornall, & Kaplan, 2020; Guler, 2007; Matusik, George, & Heeley, 2008), research suggests that the value a venture receives is contingent on three observable sets of factors: 1) the environment (Gompers & Lerner, 2000; Hochberg et al., 2010; Hsu, 2007), 2) founder attributes including their pre-entry experience and other attributes (Chatterji, 2009; Claes & Vissa, 2020; Hsu, 2007; Wasserman, 2017), and 3) the startup's accomplishments or venture progress since founding (Claes & Vissa, 2020; Hsu & Ziedonis, 2013).

### *Environmental Factors*

Environmental factors represent the supply and demand dynamics of the broader market that influence venture valuations. These include temporal, regional, and industry differences, each of which may contribute to variance in the availability of capital and competition for investments. Apart from differences in general inflation (which are generally accounted for prior to estimating valuation differences), temporal factors are changes in the availability of and competition for venture investment capital as an asset class. This includes but also goes beyond yearly fluctuations in VC inflow (Gompers & Lerner, 2000). For instance, Hsu (2007) finds that ventures receiving Series A investments in 1999 and 2000 obtained approximately 93% and 97.2% higher valuations, respectively, compared with other years in the study, reflecting the extreme inflow of venture capital during the dot-com bubble era.

Regional differences may also play a key role in valuations. Venture investors tend to primarily focus on local ventures who they can better evaluate and monitor (Shane & Cable, 2002; Sorenson & Stuart, 2001). To the extent that some regions have more or fewer investors with more or less capital to invest, this is likely to drive differences in valuations (Chen, Gompers, Kovner, & Lerner, 2010). Valuations within a region may in turn be influenced by regional networking patterns that may deter or facilitate the entry of new investors in a region. Hochberg et al. (2010) show that ventures in densely networked markets (e.g., Biotechnology industries in Massachusetts) experience lower valuation than those in sparsely networked markets (e.g., Low-tech industries in New York) and that a reduction of one standard deviation in network density is correlated with a 10% plus increase in venture valuation.

Lastly, studies indicate that industry-level factors can significantly impact ventures' prospects, which can influence the size of opportunity being pursued, its competitiveness, and the profit potential (Eisenhardt & Schoonhoven, 1990; Porter, 1980).[21] While industry effects are often found to play an important role in explaining differences in business-unit profits in established corporations (McGahan & Porter, 1997; Rumelt, 1991), support for persistent industry-level differences in valuations is mixed. Gompers and Lerner (2000) found that valuations varied systematically across six industries. In contrast and using early-stage financing associated with MIT E-Lab, Hsu (2007) found that industry dummies for Internet, software, biotech, communications, and other industries were statistically insignificant at traditional levels. Likewise, Fitza et al. (2009) failed to find systematic differences across industries in the change in ventures' valuations from one round to the next. More recently, Kotha et al. (2022) show that the industry effects are not associated with the speed with which firms receive extremely high valuations.

### *Founding team attributes*

Founding venture team's attributes are signals of the human and social capital that founders bring to their ventures. Founders with more work experience and education are likely to be better equipped to pursue opportunities and thus their ventures may have increased chances of success (Eisenhardt & Schoonhoven, 1990; Hallen, 2008; Hsu, 2007; Shane & Stuart, 2002). This in turn may allow more experienced founders to secure higher valuations, allowing them either to mobilize more investment capital for the same equity or to preserve more ownership by giving up less equity for the same amount of capital.

Prior literature highlights various founder attributes that are likely to be associated with venture success and valuations. And to be clear up front and as further elaborated in our methods section, our emphasis is on entrepreneurial human and social capital attributes that may be

---

[21] Among public companies, scholars found that industry can explain 9–19% of the differences in firm profitability (McGahan & Porter, 1997; Rumelt, 1991).

observed across all funded founders versus focusing on founder fixed-effects that can only be estimated for the small subset of founders that go onto become serial entrepreneurs.

One such key attribute is whether founders’ industry experience influences valuation because such founders may be better positioned to identify and pursue opportunities. For instance, Chatterji (2009) finds medical device industry founders who had worked for prominent companies obtained approximately 44% higher equity valuations.

Education may also predict future success through a combination of the skills learned, networks developed, and the credentials obtained. For instance, Hsu (2007) shows that founding teams with doctorates in internet-based industries tend to receive higher valuations due to signaling effects. More broadly, founding teams with degrees for elite universities often have advantages in attracting investments from more desirable investors (Hallen, 2008).

Also important is prior founding experience. Experience helps build competencies and networks specifically around entrepreneurship and helps navigate uncertainty (Mullins, 1996). Experienced founders make better decisions than relatively untested ones (Gilbert, McDougall, & Audretsch, 2006) and are more likely to receive funding at more favorable terms (Eesley & Roberts, 2012; Hsu, 2007). Kotha et al. (2022) show that founding experience is associated with venture valuation in highly valued startups such as unicorns. Furthermore, Hsu (2007) finds that serial founders whose previous venture achieved high financial returns (liquidated prior venture at >99% internal rate of return) experienced a 48% higher valuation in their current venture. Similarly, Claes and Vissa (2020) find that each additional prior startup founded by members of the founding team is associated with a 22% higher valuation.

Team size may also affect venture valuation, although different studies have inconsistent results. Consistent with the argument that larger teams have more human and social capital to leverage, Wasserman (2017) finds that a 1% increase in founding team size correlates to a 16% increase in pre-money valuation in a large sample study. On the other hand, in a smaller sample of 149 early-stage technology-based ventures, Hsu (2007) finds that the number of founders does not significantly alter valuation size – perhaps reflecting that entrepreneurial teams can often hire for needed tasks and that larger teams may have reduced incentives as they further split their equity ownership.

Beyond these founder-related signals of venture potential, common biases around founder attributes may also influence valuations. For instance, paralleling research on gender-related bias in organizations and entrepreneurship, female founders have been found to often receive lower valuations (Lee & Huang, 2018; Marlow & McAdam, 2013; Veer & Bringmann, 2021).

### *Venture progress markers*

Post-funding venture progress is the third broad category influencing a venture's valuation. As ventures make more progress toward demonstrating revenue, sales growth, and profitability, the uncertainty about a venture's future prospects is reduced (Camuffo, Cordova, Gambardella, & Spina, 2020; Eckhardt, Shane, & Delmar, 2006; Hallen & Eisenhardt, 2012; Levinthal & Contigiani, 2019; McDonald & Eisenhardt, 2020). Whereas founder attributes are fixed before founding, progress markers provide a more *dynamic* or *evolutionary* view of venture valuation where the equity they trade for resources depends on the successful entrepreneurial actions taken.

Supporting these arguments, a few studies have found venture progress is associated with differences in venture valuations. Gompers and Lerner (2000) show that ventures which reach profitability are valued 39-68% higher on average. Firm age, which reflects a venture's survival to date, is also associated with higher valuations (Gompers & Lerner, 2000; Hsu, 2007). In the semi-conductor domain, Hsu and Ziedonis (2013) find that ventures with a greater stock of patents receive higher valuations. For instance, they found that doubling the patent filings from 3.5 (the mean) to 7 is associated with an increase in pre-money valuation of $5.7 million.

While distinct from progress in validating the business viability and model, a second form of venture progress that may influence valuations is progress resulting from *past resource mobilization*. Past resources raised provide social proof of prior investors supporting a venture and also reflect the resource advantages of the venture to date (Claes & Vissa, 2020; Moghaddam, Bosse, & Provance, 2016). Hochberg et al. (2010) find that acquiring more funding rounds, a proxy for the venture's progress (more rounds meaning more milestones achieved), positively correlates with valuation. Prior valuation size and amount raised can also serve as a reference point for determining current valuation since it yields baseline expectations around exit potential and attractive market positioning (Gompers et al., 2020; Moghaddam et al., 2016). Given matching dynamics in the market for venture investments, having obtained higher-status past investors also signals venture has been perceived as having stronger prospects (Hallen, 2008). However, while Claes and Vissa (2020) have studied the relationship between VC status of the focal deal round and the corresponding valuation outcome, we are not aware of studies exploring whether past investor status influences future deal round valuation. (Note that while some studies have examined the role of current investor attributes – e.g., Hsu (2004) – we break these out in a separate exploratory analysis as described later, since which investors invest is in part driven by the other factors reviewed above.)

In summary, past assessments of important factors are generally limited to measuring their absolute rather than relative importance. The focus is on the absolute importance of a predictor or factor based on effect sizes (e.g., coefficient beta). Effect size reveals whether the focal predictor/factor has any non-zero influence on equity prices and thus provides insights about the *magnitude* of the effect controlling for other covariates (Batjargal & Liu, 2004; Bengtsson & Hsu, 2015; Claes & Vissa, 2020; Hsu, 2007). In contrast, relative importance measures the *completeness* of the explanation via the coefficient of determination (e.g., adjusted $R^2$). It measures how much variation in the equity price or valuation outcomes is attributed to the focal predictor/factor in the form of a percentage, which allows direct comparison with other factors included in the model. Further, a large absolute importance of a predictor only sometimes equates to large relative importance and vice versa.

Fitza et al. (2009)'s study is the only study we found that explicitly measures relative importance in venture valuations. The study finds that time-invariant venture effects account for significant variance in equity prices. However, since Fitza et al. (2009) fail to distinguish between founder and venture progress factors, it is difficult to discern their individual effects on venture valuations. Also, the ANOVA approach they use, highly appropriate for their context, is no longer in vogue, having been superseded by better approaches for variance decomposition, such as the Shapley Values approach, as explained below. Thus, the question of which factor is more important in explaining venture equity price differentials remains unclear. Answering the question is also important for entrepreneurs and those who support them, offering insight into their ability to create strategic advantages in resource mobilization (more capital for less equity) and relative value capture compared to investors that provide capital for building their firms.

Simultaneously, there is limited knowledge of how the relative importance of factors change over time in predicting differences in venture valuations. This is important because investors are seen applying different standards for evaluating a startup contingent on the venture's stage of development (Bernstein, Korteweg, & Laws, 2017; Gompers et al., 2020; Hallen, 2008).

Despite this evidence, no study has yet measured which factor becomes more/less important over time in predicting valuation differences and by precisely how much.

In essence, our understanding of the complex and potentially evolving nature of venture valuations is incomplete – neither do we know whether/which factors have greater influence on valuation outcomes, nor do we understand whether/which factors gain or lose relative importance over different stages of venture maturity by how much.

## METHODOLOGY

We use a variance decomposition design to explore to what extent differences in venture valuations at different stages are predicted by the factors identified above. While traditional regression analyses focus on the slope (coefficient) of the relationship between a particular attribute and a given outcome, variance decomposition analyses explore what fraction of an outcome's variance is explained by a factor or set of attributes. This approach has a long history in strategy research in examining the classic debate around the importance of industry, business-unit, and corporate-parent effects on firm performance (Kotha & Nair, 1995; McGahan & Porter, 1997; Misangyi, Elms, Greckhamer, & Lepine, 2006; Rumelt, 1991; Schmalansee, 1985). Studies have also utilized the method to examine the influence of CEOs on firm performance (Fitza, 2014; Quigley & Graffin, 2017; Quigley & Hambrick, 2015) and acquisition behavior and performance (Meyer-Doyle, Lee, & Helfat, 2019).

Unique features of our context required us to adapt our approach somewhat from more traditional variance decomposition analyses. Since we are interested in how the examined factors' impact may dynamically shift across different venture stages or levels of maturity, we ran separate variance decomposition analyses on three different deal stages in our dataset: (a) the initial round raised from VCs; (b) early follow-on rounds (where a venture was still at an early stage but had participated in a prior VC round); and (c) later rounds. Our approach mirrors Karniouchina, Carson, Short, and Ketchen (2013), who examined the impact of industry maturity on the drivers of firm performance via separate variance decomposition analyses across three different stages. Also, and as elaborated shortly and inline with this focus on distinct variance decomposition analyses for each stage, we focus on estimating the variance explained by observable attributes and not simply time-invariant fixed effects for founder attributes and venture progress.

We conduct our variance decomposition analyses using the Shapley Value approach. This approach originates from Shapley (1953)'s framework of cooperative game theory, which measures how much each player in a coalition contributes to the total value created in a game's outcome. We apply Shapley's framework to OLS regression analyses where each factor is analogous to a unique player, and the regression outcome of interest ($R^2$) is analogous to the game's outcome value. We mirror Sharapov et al. (2021)'s approach to calculating the weighted average value of each factor's incremental contribution over all possible sequential orderings (permutations) of the factors. This approach overcomes the limitations of prior variance decomposition analyses such as ANOVA, which is sensitive to the sequential orderings of factors that are included into the final model, and variance components analysis (VCA), which rely on stringent assumptions of joint-normal distributions regarding random effects and independence in the effects when the factors are non-orthogonal (Garson, 2012; Guo, 2017).

A key advantage of the Shapley Value approach is it provides an increasingly accurate assessment of the relative importance of regressors as the sample size decreases compared to ANOVA and MLM (Sharapov et al., 2021). While several recent variance decomposition studies of established firm performance have utilized multilevel modeling (MLM) (Meyer-Doyle et al., 2019; Misangyi et al., 2006; Quigley & Graffin, 2017), we chose the Shapley value approach as there was no clear hierarchical order for nesting the examined factors in our context.

Another advantage of the Shapley Value approach is its versatility on estimating the overall predictive power in a set of measures. In line with other modern machine learning methods, these measures do not need to be restricted to fixed-effect attributes repeatedly observed and assumed to have a consistent influence over time (Athey & Imbens, 2019; Choudhury, Allen, & Endres, 2021). This is key in examining how factors' impact might change across different investment stages and in helping us estimate the portion of valuation variance predicted by observable founder and venture progress attributes.

We use the *Shapley* package in R due to its computational efficiency (speed) in calculating models with many regressors or factors. We elaborate on our approach in greater detail below.

## Data Sample

Our sample is a subset of a larger sample constructed to address a research agenda around high-potential technology-based new ventures and the heterogeneity in their pathways to growth and high performance. The larger sample is sourced from Crunchbase, a venture investment and deal database often utilized as the source of venture deal data in studies of new ventures (e.g., DeSantola et al., 2023; Hallen et al., 2020; Koning et al., 2022). The database sources its data from prominent investment firms around the world in addition to voluntary user entries that go through a process of both manual and algorithmic (e.g., machine learning) validation internally.

We collect supplementary data using two other sources. First, to collect deal specification data such as amount raised and valuation per deal round, we use Pitchbook, another prominent subscription service detailing VC, private equity, and private M&A transactions. Pitchbook is the official data provider for the National Venture Capital Association. Pitchbook compiles its information from surveys of VC firms and entrepreneurs, public data, press releases, and SEC filings[22]. Second, for the founding team data, we use LinkedIn to gather the historical work experience and educational backgrounds of individual founders. LinkedIn is an online career networking and job search platform that provides public access to data on historical work experience and education with great accuracy (Ge, Huang, & Png, 2016).

The starting Crunchbase sample consists of 9,547 U.S-based ventures with any equity-based round between 2005-2017 that raise at least $250,000 and are 10 years or younger at the time of the round. We observed years after 2005 since the data quality in Crunchbase drastically increased post-mid-2000s (see Koning, Hasan, & Chatterji (2022) Online Appendix). This sample excludes several industry sectors, including biotechnology, energy, and natural resources. We chose to focus on the narrower selection of industries as they represent an oft-recognized divide between ventures grounded in business model innovations and internet-related technological and social disruptions versus biotechnology, energy, and material science startups that are often science-driven (Kerr, Nanda, & Rhodes-Kropf, 2014).

As part of the larger project, we gathered data from LinkedIn for those ventures where Crunchbase listed founders' LinkedIn profiles and profiles were publicly accessible (4,058 ventures or 42.5% of that starting sample). We identified information on the members of the founding team using Crunchbase. For each founder, Crunchbase provides a link to each founder's LinkedIn URL that can be used to navigate and code the individual's entrepreneurial, professional, and educational experiences. Appendix C provides further analyses examining where founder information was available versus absent. We observed generally small differences in mean values across these two groups. Further, as elaborated below, we use survey-based weights in this study

---

[22] In our own interactions with VCs, many have told us they use Pitchbook as their primary source to examine valuations in the market for both potential investments and competing ventures; A limitation of Pitchbook, however, is that subscriptions typically do not allow the export of large samples by industry or over time – making it less suited for identifying the original sample of ventures that may have raised funds.

to account for observable differences in founder profiles being more likely available for some types of ventures. Overall, this yielded what we believe is an unusually large sample of venture deals with rich founder information.

Given our research question, we further narrowed our sample for this study to focus on those ventures that raised early-stage rounds where a valuation was agreed upon between the venture and the investors. These are called “priced” rounds. Most rounds involving a few million or more in investment tend to be priced. In contrast, smaller rounds that ventures raise very early in their life tend to often use convertible notes or similar structures (e.g., SAFEs). In such non-priced rounds, the investment is made with an agreement that a valuation will be assigned when the venture later raises a priced rounds using some agreed-upon structure (e.g., a valuation at a 20% discount to the first priced round, with the resulting valuation capped at $20M). Many ventures, though, never gain enough traction to raise priced rounds.

We used Pitchbook to identify valuations of priced rounds, a dataset widely utilized by venture capital firms and regarded as the industry standard for valuation data. Of the 4,058 ventures in our starting sample for whom we had identified LinkedIn profiles, we identified 4,032 in Pitchbook. However, only 1,737 of these had valuation data for initial priced rounds that met the following conditions: and 1) The round is an early-stage deal that is both priced and involves VC investment and 2) the round is raised by a venture that is less than 8 years at the time of the deal (since raising an early round very late in age may indicate a substantatively different private equity deal).

Using the same set of 1,737 ventures in the initial round sample, we then constructed the other two samples, early follow-on and later rounds. The “early follow-on rounds” sample is comprised of rounds labeled in Pitchbook as “Early Stage VC” rounds but which came after the initial VC round; these were often but not always Series B rounds. This sample includes 1,170 rounds from 837 ventures, averaging 1.39 rounds per venture. Our “later rounds sample” sample are rounds in these ventures that are listed as “Later Stage VC”; these are typically Series C and D rounds. There are 602 rounds in this sample from 374 ventures, averaging 1.6 rounds per venture. The gradual reduction in the number of ventures over the two samples compared to the initial sample reflects the fact that not all ventures raise subsequent rounds from VCs after their initial deal for various reasons (e.g., going out of business, relying on sales to fund growth, etc.). Compared to the initial round, 48% percent of ventures raised early follow-on rounds and only 21% raised later-stage rounds. We also report supplementary analyses restricting the samples to one venture per round in the results section (see Appendix D for results).

In exploring our sample, an examination of the descriptive statistics revealed that a small number of initial rounds appeared to have very high valuations, suggesting that they were unlikely traditional VC investments but rather some form of private equity round (e.g., rollups). Accordingly, we limit the corresponding values for the initial round via winsorization. We winsorize the bottom and top 2.5% of rounds by valuation and the bottom and top 5% of rounds by the amount of previously raised funds (i.e., changing valuations above or below these thresholds to be at most/least the threshold values). Given that outliers were less pronounced in the early follow-on and later round valuations, we did not winsorize those samples.

To account for potential biases in representation arising from availability of founder and valuation data, we conduct all analyses with survey weights allocated to each venture. See Hallen, Cohen, and Park (2023) for a similar approach. Weights were created based on the probability in the larger sample of founder data being available and the venture raising a priced round. These probabilities were calculated using a logit model based on observable attributes from Crunchbase of the initial round regardless of whether it was unpriced or priced. These include the (logged) amount raised in the round, the number of investors, initial industry category, state (region), and deal year of initial round. See Table C2 in Appendix C for further details. Survey-based weights were then calculated as one divided by the estimated probability. In effect these weights give greater emphasis to observations where similar ventures were more likely to have not been included in the smaller sample due to either founder information not being available or their not raising a priced round. The weights are trimmed at the 99% percentile. We also ran analyses omitting these survey-based weights and obtained highly similar results. Comparing the model with and without weights, the largest difference in adjusted R-squared value was regarding past resource mobilization in the early follow-on rounds, with a 4% difference. All other factors exhibited differences of less than 2% across all three rounds.

## Measures

*Dependent Variable*: Our dependent variable is the *premoney valuation* provided by Pitchbook[23]. This variable represents the agreed upon valuation between the venture and the investors' syndicate participating in a given round. Pre-money valuation reflects the price at which investors are purchasing equity in a venture for each dollar invested. Following past studies, it is measured in millions of dollars, adjusted for inflation to 2003 dollars, and logged to reduce skew (Gompers & Lerner, 2000; Hochberg et al., 2010; Hsu, 2007).

*Region*: We measured region at the MSA level. MSAs with less than three appearances in the data were aggregated into a single "Other MSA" category. The initial round sample contained 45 region

---

[23] Pre-money valuation represents the valuation of a venture in the moment just before an investment occurs. Distinguishing between pre-money and post-money valuations is necessary in the venture investment context because ventures are both issuing new shares to investors and because the invested capital becomes the property of the venture (not the prior owners). Thus, the moment after the investment occurs, the post-money valuation is the pre-money valuation plus the invested capital (which now sits in the venture's bank account).

categories.

*Industry*: Industry was determined using the the initial category designation by Pitchbook. This yielded a set of 31 industry categories in the initial round sample.

*Year*: We measured year effects consisted of dummies for each deal year. The initial round sample contained 11 categories for the years of investment (2007–2017).

*Founding Team Attributes*: Given the limited repetition of founders across our datasets, we measured the founding team's human and social capital utilizing attributional measures from the entrepreneurship literature on venture fundraising and performance (Eesley & Roberts, 2012; Eisenhardt & Schoonhoven, 1990; Hallen, Cohen, & Bingham, 2020; Hsu, 2007). We include the number of founders by creating dummies *single founder, team of 2-3 founders (reference being the team of 4 or more founders)* to account for larger teams with more human and social capital, and whether the *founding team has female,* as research suggests that investors may discriminate against women (Kanze, Huang, Conley, & Higgins, 2018; Lee & Huang, 2018). Founder gender is inferred using Genderize.io. We utilize an 80% confidence interval following Marx (2022). In the case that Genderize.io was not able to infer, we used gender information entered into the Linkedin profile by the user. In the case we were not able to assign a gender, we gave the *gender unknown* a value of 1 while we assigning has female value to 0. We captured a founding team's educational experience by including dummy variables indicating if any founders possessed an *MBA*, *JD, technical doctorate, other doctorate, technical masters,* or *other Masters*. Following Rider (2012) and Hallen, Cohen, and Bingham (2020), we controlled for the maximum educational prestige of founders' undergraduate and graduate degrees using U.S. News & World Report's 2012 worldwide ranking of the top 400 global universities, with scores ranging from 31.3 to 100. Founders whose university was not included in the list were assigned the score of 28.

We also included numerous measures of founder's professional experience by capturing the maximum value of the *years of work experience (logged)* between founders' undergraduate graduation and the deal year among the founding team members. We measured *prior employer prominence (logged)* in accordance with Hallen, Cohen, and Bingham (2020) by measuring how many other ventures have founders who share the same employer as the founders of the focal venture (within our sample). Finally, we measured entrepreneurial experience using dummy variables to indicate if founders had *previously founded another venture, number of previously founded ventures,* and *previously founded a VC-backed venture*. Data for each founder were coded from their LinkedIn profiles, with these identified via Crunchbase.

We used the identified founders where some, but not all, of a founding team's profiles could not be identified on LinkedIn. In cases where only founder education or job data was listed on LinkedIn, we set the other measures to 0. We included a dummy variable indicating this: founder *education linkedin missing* and *founder job linkedin missing*.

*Venture Progress*: We measured a venture's progress using Pitchbook's data on *business status* at the time of a round. This categorical variable indicates if a venture is in one of the phases: startup, stealth, beta test, product development, revenue, or profitable. We used startup as the omitted category and aggregated stealth and beta test into this phase, as only 22 instances (1.3%) of stealth and beta rounds were found in the initial round sample. We included a dummy if a venture had previously *participated in an accelerator program* as certain accelerators have been found to improve venture quality (Cohen et al., 2019; Gonzalez-Uribe & Leatherbee, 2018; Hallen, Cohen, & Bingham, 2020; Yu, 2020). We also controlled for a *venture's age (years)* from founding and logged this measure to reduce skewness.[24] Lastly, we include *internet archive crawls*, a measure that counts the number of times the Internet Archive crawled the focal website in the past 90 days prior to the fundraising deal date (Hallen et al., 2023). The number of crawls indicate how central the domain is on the internet (the more crawls, the higher the centrality) as a proxy for web traffic, the latter increasingly being adopted as a measure of venture growth and firm performance (Koning et al., 2022; Dushnitsky, Piva, & Rossi-Lamastra, 2022). We use web crawls instead of web traffic since web traffic historical data providers only release data up to 3 years from the time of data collection. In a separate working paper, we find strong positive correlation ($r = 0.65$) between web traffic (logged) and web crawls (logged) using a pseudo-random sample of 300 ventures of our initial sample over the past three years (2020-2023) at the quarter-level. We assign a value of 0 in cases where there are no web crawls observed in the time frame. Additionally, we log the measure to reduce skew.

*Past resource mobilization:* For the analyses of early follow-on and later rounds, we include past resource mobilization outcomes as a separate factor to capture any path-dependent influence of this early outcome. First, we include the *maximum pre-money valuation (logged)* obtained in previous rounds. Second, we accounted for the *total size of funds raised in prior rounds*, measured in millions (adjusted for inflation) and logged to reduce skewness. Third, to capture the impact of a venture's prior inevstors, we incluced *previous maximum investor status* among investors in all prior rounds, measured using eigenvector centrality.

*A note on current investor characteristics:* We chose not include characteristics of the current-round investors (such as investor status and type) in our main analyses. These characteristics are

---

[24] We considered gathering patent data as research has indicated that patents generally provide less protection in the industries we were studying (Cohen, Nelson, & Walsh, 2000); instead, we chose to focus our data gathering resources on founding team attributes.

typically determined by quality-based matching, where high-quality ventures are more likely to match with similar quality investors (the same applies to lower quality venture and investor pairs) (Hallen, 2008; Mindruta, Moeen, & Agarwal, 2016; Sørensen, 2007). Since these matching dynamics are likely driven by other factors already included in the analysis such as the attributes of the founding team, the progress of the venture, or the environment, omitting these characteristics helps better capture the net (direct and indirect) effects of our focal factors. However, we do conduct exploratory analyses including them; these results are discussed in the 'Exploratory Analyses' of the results section below and are detailed in Appendix E.

## Estimation Approach

The Shapley Value approach to variance decomposition utilizes the exhaustive set of permutations regarding a set of factors/regressors (all possible orders in which factors/regressors can be included into a model; each permutation is a unique sequence of factor or regressor orderings) to calculate the average incremental contribution of each factor/regressor (Elbers, 2021; Sharapov et al., 2021; Shorrocks, 2013). The average incremental contribution is measured as the model's increment in total variance explained ($R^2$). Using the general form of a linear regression model where there are $\boldsymbol{i}$ number of factors ($\boldsymbol{X_1}, \boldsymbol{X_2}, \cdots, \boldsymbol{X_i}$),

$$\boldsymbol{Y} = \boldsymbol{\beta_0} + \sum_{\boldsymbol{i=1}}^{\boldsymbol{i}} \boldsymbol{\beta_i X_i} + \boldsymbol{u}$$

each factor ($\boldsymbol{X_i}$) is assumed to be able to take any position in the sequential order in which it is included into the model. For instance, if there are two factors to be measured ($\boldsymbol{S_2}: \{\boldsymbol{X_1}, \boldsymbol{X_2}\}$), each factor can take on the position of being included in the model either first or second, respectively ($\boldsymbol{S_2}: \{(\boldsymbol{X_1}\ \boldsymbol{X_2}), (\boldsymbol{X_2}\ \boldsymbol{X_1})\}$). The total number of possible permutations in this case would be two ($\boldsymbol{i}!$). Then, each factor's unique contribution to the model's $R^2$ is measured as the average value of its incremental contribution to $R^2$ of the two permutations. The average incremental contribution of a given factor $\boldsymbol{X_i}$ in regard to the full model's $R^2$ can be expressed as follows:

$$\frac{\mathbf{1}}{\boldsymbol{i}!}\sum_{(\boldsymbol{all\ permutations\ of\ a})}\overline{\boldsymbol{R}}^{\mathbf{2}}\big(\boldsymbol{X_i}\cup\boldsymbol{S_i}(\boldsymbol{a})\big)-\ \overline{\boldsymbol{R}}^{\mathbf{2}}(\boldsymbol{S_i}(\boldsymbol{a}))$$

Here, there are $\boldsymbol{i}$ number of factors and each unique permutation is expressed as $\boldsymbol{a} = (\boldsymbol{a_1}, \boldsymbol{a_2}, \cdots, \boldsymbol{a_i})$. A unique permutation where factors are included before $\boldsymbol{X_i}$ in the specific order of $\boldsymbol{a}$ is expressed as $\boldsymbol{S_i}(\boldsymbol{a})$ and total variance explained by $\boldsymbol{S_i}(\boldsymbol{a})$ is denoted as $\overline{\boldsymbol{R}}^{\mathbf{2}}(\boldsymbol{S_i}(\boldsymbol{a}))$. Finally, the incremental contribution to total variance explained by factor $\boldsymbol{X_i}$ is denoted as $\overline{\boldsymbol{R}}^{\mathbf{2}}\big(\boldsymbol{X_i}\cup\boldsymbol{S_i}(\boldsymbol{a})\big)-\ \overline{\boldsymbol{R}}^{\mathbf{2}}(\boldsymbol{S_i}(\boldsymbol{a}))$. We use the formula to measure both $R^2$ and adjusted $R^2$.

Following Sharapov et al. (2021)'s suggestion that inferencing from Shapley value-based variance decomposition results are better executed via non-parametric boostrapping (with confidence intervals) due to the lack of finite sample sampling distribution, we run 100 iterations of sample size equivalent to the original sample size of each stage (initial, early follow-on, and later rounds) with replacement using a 95% confidence interval (two-tailed) (Johnson, 2001). For robustness purposes, we compare the outputs derived via bootstrapping to the non-boostrapped version (See Appendix F) that utilizes the full sample of each deal stage.

## RESULTS

Table 7 presents descriptive statistics for the continuous measures within each of our three data samples: *Initial rounds, early follow-on rounds, and later follow-on rounds*. The median inflation-adjusted pre-money valuations for each sample are \$4.9 million for initial rounds, \$16.5 million for early follow-on, and \$60.4 million for later rounds. Interestingly, though only 2% of teams had previously received VC backing, serial entrepreneurship was relatively common in the sample and was seen in 40% of teams (suggesting much of their earlier attempts was far less successful in raising venture capital). As expected, more ventures were earning revenue in the early follow-on and later follow-on samples. Only a fraction of ventures (6%) was profitable by later rounds, though, likely reflecting the tendency of VC-backed ventures to accelerate growth at the expense of profitability. To clarify the empirical distinctiveness of founding team attributes to venture progress, we checked the correlation between the variables for these two factors and found no correlation larger than 0.3 or smaller than -0.3.

## Initial Rounds

The first two columns in Table 8 present the variance explained for each factor in the initial rounds sample. We focus on adjusted $R^2$ (instead of $R^2$) to prevent artificially inflating the explanatory power arising from the sheer number of independent variables we include in the regression model. We find that region explains sizeable mean variance in terms of adjusted $R^2$ (10.21%). Industry has a relatively marginal effect (1.84%) on valuation differentials. Similarly, year also explained a marginal amount of the variance (3.24%). These results indicate that region explains a notable amount of total variance in valuation while industry and year not so much in the initial rounds. Combined, environmental factors explain a sizeable amount of total variance (15.29%).

Founding team attributes also uniquely explained a relatively small amount, 4.67%, of valuation variance. This limited influence of founding team attributes is particularly striking, as literature have highlighted the advantages that accrue to more experienced, connected, and skilled entrepreneurs and senior executives (Eisenhardt & Schoonhoven, 1990; Gompers, Kovner, Lerner, & Scharfstein, 2010; Hallen, 2008; Klepper & Sleeper, 2005; Quigley & Graffin, 2017; Shane & Stuart, 2002). We return to this unexpected finding later in the discussion. Finally, venture progress uniquely explained 3.39% of the variance. These results indicate that region (location) explains the greatest variance, followed by founding team attributes, venture progress, year effects, and industry effects in that order. Notable is that most of the variation remained unexplained even after accounting for all these factors; the adjusted $R^2$ of the full model is 23%, meaning that 77% of the variance remains unexplained.

## Early Follow-on Rounds

Estimates for early follow-on rounds are also shown in Table 8. These are rounds where a venture had previously raised another priced round, but the current round was still classified as "Early-Stage VC." Region uniquely explained only 4.62% of the total variance, a much smaller figure compared to the initial priced round. Industry fixed effects again had no discernible relationship to valuation (3.41%). Year-fixed effects also had a minimal effect on valuation (3.46%). Together, the three environmental factors explain 11.49%.

Founding team attributes continued to explain a limited amount of variance in valuation (3.78%). Venture progress in early follow-on rounds also explained little variance, like the initial rounds, explaining 4.26% of the variance. Past resource mobilization, though, accounted for a substantial portion of the variance by explaining 38.62% of the total variance in valuation outcomes, far outweighing the influence of the other measured factors – something we expand on in the Discussion.

## Later Rounds

Table 8 presents the estimates for later rounds in the last two columns. Compared to the initial or early follow-on rounds, region (6.4%), industry (5.53%), and year (4.47%) each explained slightly larger amounts of variance compared to early follow-on rounds but still largely non-substantial portions. Environmental factors thus continued to explain a sizeable portion of the total variance (16.4%). Founding-team attributes accounted for 5.8%, while venture progress explained 11.55%

of the variance. Venture progress explains a noticeably larger amount of variance compared to both initial and early follow-on rounds. Our analyses reveal that the increase can be attributed to the web crawls that proxy for traction. Past resource mobilization was again the dominant predictor of valuation variance in later rounds. Compared to early follow-on rounds, past resource mobilization accounted for an even larger amount of variance (45.95%). The adjusted $R^2$ of the model also gets much larger as ventures mature (Initial round: 23%; Early-rounds: 58%; Later-rounds: 80%), suggesting valuations become much more comprehensively predictable from observable attributes as ventures mature.

## Exploratory Analyses

We ran a variety of analyses to further explore the sensitivity of our results to alternative empirical choices. These are detailed in the Appendix and summarized here.

Including Current Round Characteristics: We omitted the characteristics of the current round investors from our main analyses to better highlight the net effect of each focal factor and as these factors would likely function as mediators given assortative matching in funding dynamics (Hallen, 2008; Mindruta et al., 2016; Sørensen, 2007). However, we ran further analyses that included the characteristics of the current investment round to examine both the importance of this factor and its effect on the other focal factors. Here we include investor-level fixed effects for the highest-status investor, as well as measures for number of investors, types of involved investors (individual angels, angel groups, CVCs, etc.), percentage of equity acquired by inevstors, and measures of founder control (whether the CEO position is still held by a founder). See Appendix E for further details.

Overall, we find that the investor characteristics/deal attributes uniquely account for a sizeable amount of total variance explained (Adj. $R^2$ for Initial round: 21.38%; early follow-on rounds: 23.12%; later rounds: 27.09%). This finding aligns with Fitza et al. (2009)'s study that shows the substantial effect of VC ownership on changes in valuation between rounds, though we find effects larger than they do. Strikingly, though, we find that the results for environmental factors, founding team attributes, and venture progress, remain highly similar to those shown in Table 8 as estimated without these current-round characteristics (<2.7% difference in Adj. $R^2$ for each factor between model with and without investor-investment attributes). The one place we do see more of a shift is the influence of past resource mobilization by some margin (e.g., a change of 5.92% in Adj. $R^2$ in early follow-on rounds and a change of 12.16% in later rounds). We elaborate on these findings in Appendix E and the Discussion.

Exploring Sample and Measurement Choices: We first restrict the follow-on early-stage and later-stage samples to the earliest round per venture in each sample. Results were largely consistent with those reported in Tables 8 (see Appendix D). We also explored whether region effects were more pronounced if we restricted region fixed effects to just the top 6 MSA; this, however, yielded very similar results aside from 3-8% smaller region effects (potentially due to limited regional heterogeneity in environmental munificence in top MSAs) in the three rounds and approximately 5% larger industry effects and past resource mobilization, respectively in the later rounds. Additionally, we ran the same Shapley value variance decomposition analysis without using bootstrapping and instead, utilizing the full original sample for each of the three stages. We find similar results to the bootstrapped analysis. The results are reported in Appendix F.

## DISCUSSION

We sought to understand the *relative* importance of key observable factors that influence a venture's valuation at *various* stages of maturity, utilizing a state-of-the-art approach to variance decomposition, the Shapley Value Approach (SVA).

We find that environmental factors (i.e., region, industry, and year) collectively explain a notable portion of valuation differences (15% of total variance in initial rounds, 12% in follow-on early rounds, and 16% in later-stage rounds). Initial round valuations are predominantly driven by regional differences, though the effects are relatively more even across environmental factors in subsequent rounds. Unexpectedly, observable founding team attributes explain only a modest portion of valuation differences across all rounds (4.67% in initial rounds, 3.78% in early follow-on rounds, and 5.8% in later-stage rounds). Also, while our measures of venture progress explain a small amount of variance in the initial (3.39%) and early follow-on rounds (4.26%), the amount increases substantially in the later rounds (11.55%).

Although not part of our main theorizing, other factors exhibit prominent effects. The characteristics of the current rounds were highly predictive of future venture valuations; they explain 21.4% to 27.1% of the total variance in our data. Likewise, past investment round attributes explained 32.7% to 33.8% when accounting for current-round characteristics (and a greater percentage of the variance when not accounting for the current round). Also striking is that despite accounting for each of these factors, current or prior round investment still had limited impact on the variance explained by environmental, founder attribute, and venture progress attributes, suggesting that these factors are largely not mediated by these characteristics.

## Contributions

These findings offer important contributions to entrepreneurial finance and venture growth, resource mobilization, and value capture literatures.

*Contributions to Entrepreneurial Finance and Venture Growth literatures.* Our first contributions lie at the intersection of entrepreneurial finance and venture growth literatures. As noted above, our findings regarding the collective impact of environmental factors are broadly consistent with the past literature on both venture valuations and firm performance (measured as growth) literature more broadly (e.g., Brush & Bromiley, 1997; Fitza, 2014; Fitza et al., 2009; McGahan & Porter, 1997; Misangyi et al., 2006; Rumelt, 1991; Wang, 2023). So here, our contributions largely reinforce the role of region, industry, and year variables as determinants of venture valuation differences.

However, more striking are the relatively limited effects of founding team attributes on venture valuations. We found that these attributes explained only a modest portion of valuation differences in early rounds (4.67% in initial rounds, 3.78% in early follow-on rounds) and slightly more in later-stage rounds (5.8%). This finding is striking as observable founding team attributes have been a substantial focus of past research on numerous entrepreneurial outcomes, including scaling, success in finding partners and investors, and the amount of capital raised in investments (Eisenhardt & Schoonhoven, 1990; Klepper & Sleeper, 2005; Hsu, 2007; Agarwal, Campbell, Franco, & Ganco, 2016; DeSantola & Gulati, 2017; Hallen, 2008; Hashai & Zahra, 2022). The broad logic posited is that founding attributes often reflect the human and social capital that founders bring to their ventures. Prior research suggests that combining these resources and the signals they convey often creates path-dependent advantages amplified over time. Despite including a rich set of externally observable founder attributes drawn from past literature, we found their effects were rather muted.

So how do we interpret observable founding team attributes having a relatively modest effect in predicting valuation differences, especially in earlier-stage rounds? A few possible explanations could offer intriguing avenues for follow-on research. First, even though more experienced and capable entrepreneurs may be better able to attract capital and grow their ventures, they may have relatively limited bargaining power when compared to investors – especially those from whom they most wish to raise capital investments (cf., Hsu, 2004). Second, investors may be primarily attuned to hard-to-quantify founder differences not quantifiable via external data sources such as personality traits (cf., Kaplan, Klebanov, and Sorensen, 2012). Third, founders may primarily influence valuations through astute negotiation when raising investments (Clough et al., 2019). This may be an especially likely explanation as past literature suggests that entrepreneurs often lack multiple competing investment offers due to "high frictions" in fundraising (Hallen & Eisenhardt, 2012). As a result, they often accept the valuation that is offered. While exploring each of these potential explanations is beyond our current research, we see our findings raising important questions about the extent to which founder attributes might have a distinctive relationship to venture valuations than other entrepreneurial outcomes.

Next, venture progress in the initial and early follow-on rounds also had a comparably minor influence in predicting valuation differences (they accounted for about 3.39% and 4.26% of the total variance explained). This finding is surprising since venture progress has been shown to predict other entrepreneurial outcomes, including raising capital from higher-status investors. However, this finding is limited to fewer studies when compared to founding team attributes (Claes & Vissa, 2020; Hallen, 2008; Hallen & Eisenhardt, 2012; Hsu, 2007; Hsu & Ziedonis, 2013).

Paralleling our arguments for founding team attributes provided above, one possible explanation for this limited effect during the early phases fund raising is that investors might be focusing on less externally observable measures of venture progress, such as prospective customers feedback regarding their offering. Such feedback that entrepreneurs receive is hard to capture or quantify via external data sources. Another possibility is that even though greater venture progress may help attract investors in the early stages, friction in the fundraising process may still make it difficult for entrepreneurs to obtain competing offers simultaneously, thus limiting their power to negotiate more advantageous offers. Again, further work is needed.

Our findings suggest that a venture's progress becomes substantially more important in the later rounds. The increased explanatory power of venture progress in the later stages (i.e., 11.5%) may be due to progress markers becoming more useful and salient quality signals as ventures grow. This is evidenced in Table 7, where the standard deviation of internet archive crawls or the number of profitable ventures increases as the stage progresses, providing a stronger and possibly more salient contrast over time.

Also striking was the predictive power of the current and past investment rounds. This finding encompasses who the investors are, how much capital is being raised, and any prior valuations. As such, they represent proxies for otherwise hard-to-observe venture-level characteristics such as those considered above. These may play a key role in the matching process by which entrepreneurs receive and accept offers from particular investors (Hallen, 2008; Mindura et al., 2016). Still, another possibility is that these factors primarily capture investors' desirability and ability to negotiate more attractive deal terms. While further research is needed, a key takeaway from this study is that whom entrepreneurs raise money from (both currently and previously) have a far bigger effect on the valuations they receive than the entrepreneurs' backgrounds or the venture progress.

## Implications for Theories of Entrepreneurial Strategy

Our findings also hold important implications for entrepreneur research focused entrepreneurial strategy. First, while we focus on valuation, ventures that receive higher valuations can secure more capital while giving up the same amount of equity. This suggests that valuations may be a key lever in the extent to which ventures have early resource advantages, being better able to invest in their products or accelerate customer acquisition while outpacing competitors. Our findings unexpectedly suggest that many classically emphasized advantages of new ventures, including their founding team attributes and early-stage venture progress, may offer limited benefits regarding these resource mobilization advantages conferred primarily through better valuations.

On the one hand, this suggests that the founding of a new venture may often act as a “fresh start” that helps create a level playing field. For strategy research, it raises important questions about the transferability of advantage from founder and top management team characteristics – key sources of competitive advantage often emphasized in the literature (Eisenhardt & Schoonhoven, 1990; Shane, & Stuart, 2002; Fitza, 2014; Quigley & Hambrick, 2015; Agarwal et al., 2016; Quigley & Graffin, 2017).

On the other hand, our findings also raise important questions for the classic question of the relative value capture and control of the different parties involved in building and managing firms (Berle & Means, 1932; Hall & Woodward, 2010; Wasserman, 2012; Magnuson, 2022). Here we find that both founders’ prior experiences and early-stage accomplishments post founding have relatively limited bearing on their ability to secure greater valuations (while preserving greater ownership equity), at least in the early stages of the venture. In contrast, current and past investor, and deal characteristics – as well as region (in the initial round) – significantly influence valuations garnered. This suggests that apart from their choice of whom to raise capital from or locate to different region, entrepreneurs may have less direct influence over retaining larger ownership of their ventures while raising the same amount of capital, especially in the early stages. The important caveat, though, is that future research may be able to identify other – likely hard to observe from external data sources – attributes that may exert a greater influence on valuations venture receive.

## CONCLUSION

Our study provides a uniquely comprehensive illustration of venture valuations' multifaceted and (not so) dynamic nature. Our findings unexpectedly indicate that within our focal sample, factors that have shown to exhibit substantial advantages in the entrepreneurial process – namely observable founder attributes and venture progress – exhibit a much smaller role in predicting differences in venture valuations in the early stages of the venture. In contrast, the attributes of the investors and the current and past deal structure exhibit substantial roles. While ours is only one study, we believe the surprising nature of the results invites important avenues for future research. This may include quasi-replications seeing if results hold in other types of ventures, eras, or samples of ventures. Our findings also raise questions as to whether there are additional levers to be identified by which entrepreneurs may exert a more direct and substantial influence on their valuations. As a whole, we see our study not as making a definitive statement about venture valuations but rather as a springboard suggesting several avenues for promising future research.

# Chapter 3. Are Seed Accelerators Status Springboards For Startups? Or Sand Traps?[25]

## ABSTRACT

Recent research finds participating in seed accelerators can improve startups' access to growth and funding, but effects are highly heterogeneous. Less understood is whether accelerator participation can also help startups obtain higher-status early partners. On one hand, participation may improve startups' quality and signal quality to desirable partners. Yet research suggests reasons any signaling effects might be quite small or even negative, since accelerators might be perceived as a "market for lemons". Using two complementary and proprietary datasets of U.S. startups, we find many seed accelerators are indeed springboards – associated with startups raising from higher-status investors – but that others have very limited and some slight negative effects. We examine contingencies and mechanisms, and conclude with contributions to literatures on accelerators and organizational status.

## INTRODUCTION

Accelerators are short, fixed-length entrepreneurship support programs providing batched-cohorts of early-stage startups with intensive mentoring and education (Cohen, Bingham, & Hallen, 2019a). Thousands of accelerators have sprouted across the globe. These include seed accelerators focused on high-potential startups (e.g., Y Combinator, Techstars), impact-oriented accelerators focused on supporting disadvantaged entrepreneurs, regions, and societally-important sectors (e.g., Village Capital and Unreasonable Institute) and corporate accelerators connecting

---

[25] A similar version of this chapter was published as: Hallen, B. L., Cohen, S. L., & Park, S. H. (2023). Are seed accelerators status springboards for startups? Or sand traps? *Strategic Management Journal*, 44(8), 2060–2096. The reuse was done with permission from **Wiley/Strategic Management Journal**.

startups to sponsoring companies (e.g., programs by Disney and Target) (Cohen, Fehder, Hochberg, & Murray, 2019b; Dushnitsky & Sharkar, 2022; Roberts & Lall, 2019; Shankar & Shepherd, 2019; Younger & Fisher, 2020). Research has found that some (but not all) accelerators improve the performance of participating startups in terms of growth, customer traction, and capital raising (Gonzalez-Uribe & Leatherbee, 2018; Hallen, Cohen, & Bingham, 2020). Participation is also associated with startups closing faster (Yu, 2020). Further, accelerator, region, and entrepreneur factors can influence heterogeneity in accelerator effects (Assenova, 2020; Lall, Chen, & Roberts, 2020). Overall, research indicates *some* accelerators substantially help startups gain traction, grow, and raise funds, but *many do not*.

Less clear, and our research question, is: *whether attending a seed accelerator influences obtaining high-status partners, and if so, does it help or hurt?* Partners such as investors, alliance partners, or key supply and distribution partners are often vital to startups' success because they provide access to resources, knowledge, capabilities, and legitimacy that startups often otherwise lack (Clough, Fang, Vissa, & Wu, 2019; Eisenhardt & Schoonhoven, 1996; Mindruta, Moeen, & Agarwal, 2016; Ozcan, 2018; Ozmel, Reuer, & Gulati, 2013; Ozmel & Guler, 2015; Rothaermel & Deeds, 2004). But not all partners are equally valuable. High-status partners are particularly desirable since they are likely to offer superior resources, more valuable advice, and confer high-status (Hallen, Katila, & Rosenberger, 2014; Hsu, 2004; Stuart, Hoang, & Hybels, 1999). Status is a socially-constructed hierarchy of an organization's perceived quality relative to its peers (Gould, 2002; Jensen, Kim, & Kim, 2011; Pollock et al., 2015; Podolny, 1993; Shipilov, 2005). Startups often seek high-status partners since status is a proxy for a potential partner's otherwise unobservable quality. But, whether short interventions, like accelerators, can influence startups obtaining higher-status partners is theoretically unclear (Azoulay, Stuart, & Wang, 2014; Simcoe

& Waguespack, 2011). Even more, the directionality of any effect, even by higher-status accelerators, is also unclear.

On the one hand, accelerators may improve startups' quality through learning (Gonzalez-Uribe & Leatherbee, 2018; Hallen et al., 2020; Lall et al., 2020; Miller, O'Mahony, & Cohen, 2022; Yu, 2020), making participating startups more attractive to higher-status partners. Additionally, accelerators that have previously graduated high-quality startups may have a signaling effect that further helps participating startups attract higher-status partners (Connelly, Certo, Ireland, & Reutzel, 2011; Merton, 1968; Podolny, 1993; Spence, 1973). But on the other hand, recent studies have found status-related signaling effects are often smaller in magnitude than previously believed (Azoulay, Stuart, & Wang, 2014; Malter, 2014; Simcoe & Waguespack, 2011). Accelerators may also run the risk of becoming perceived as "markets for lemons" that primarily attract and graduate lower-quality startups (Akerlof, 1970). In such cases, participation might be a negative signal that might even overwhelm any quality-improving effects. Thus, even amongst accelerators found to aid startup development, it might be that they help startups form partnerships but with lower-status partners. For these collective reasons, better understanding the impact (if any) of accelerators on startups obtaining high-status partners may address important gaps in the literatures on accelerators and organizational status, in addition to offering important implications for practice.

We conduct our study in the context of U.S. seed accelerators and focus on their impact on high-potential startups' obtaining higher-status equity investors, a critical and prevalent early partner (Hallen, 2008; Ozmel et al., 2013; Pahnke, Katila, & Eisenhardt, 2015). We use two proprietary datasets with complementary strengths and tradeoffs around causal identification and generalizability. Our first dataset captures highly confidential data on 235 startups that either participated in or were “almost accepted” to four accelerator cohorts in 2011-2012 across three seed accelerators that were relatively high-status with respect to other accelerators. Of these, 85 startups received investments from disclosed investors. While focusing on just a few accelerators and unable to show heterogeneity across programs, this first sample is particularly well-suited for rigorous causal inference in identifying a close comparison set of startups that did not participate in the focal accelerators but were deemed to be of highly-similar potential through the expert evaluations of accelerator directors and their mentors. We find these accelerators have an average treatment effect that is positive and non-trivial: for a startup that would otherwise be expected to raise from an investor ranked 200, participation in one of these accelerators is associated with raising from an investor ranked 106 instead.

The second, larger dataset consists of 3,702 early-stage U.S. technology startups that raised investments from 2005 through 2017, of which 1,501 were participants in 111 different seed accelerators. The wider breadth of this sample allows examination of heterogeneity in effects as well as contingencies at the accelerator, region and entrepreneur-levels. We find participation in 21 of the 25 largest accelerators (by number of funded startups) is associated with positive effects in helping startups obtain higher-status investors, though the magnitude varies substantially and many have only small effects. We also find patterns consistent with effects being driven in part by quality-relaying signaling, though our tests of quality-improving learning are less conclusive. Unexpectedly, the status of investors in accelerated startups does not appear to be influenced by regional entrepreneurial activity or recursively by the status of past graduates’ investors.

Our study offers several contributions. First, we contribute to the growing literature on accelerators as key programs that help develop startups, entrepreneurs, and regions. While prior work has shown that some (but not all) accelerators help startups grow and raise funds, we add by showing that many also help startups obtain higher-status partners post-acceleration – an important outcome for many startups. Second, while literature on organizational status and networks has emphasized that status often changes gradually, we find relatively short interventions may act as status springboards (and not status sand traps): impactful and rapid paths to higher-status partners. Finally, for entrepreneurs, accelerator directors, and policy makers, our results indicate a broader number of accelerators may offer a broader set of benefits than suggested by prior research.

## THEORY

### Accelerators' Effects and Mechanisms

An emerging body of work has examined the effects of accelerators on participating startups' growth, survival, and funding. Collectively prior work suggests that *some* accelerators help startups gain customer traction and raise funds, but that *many do not.* Two stylized mechanisms have been theorized to causally drive accelerators' effects: *learning* and *signaling* (Cohen et al. 2019a; Gonzalez-Uribe & Leatherbee, 2018; Hallen et al., 2020; Lall et al., 2020; Yu, 2020). To be clear, these mechanisms may be complementary and co-exist (i.e., we do not wish to setup a "horse race"). Additionally, research suggests that a correlational dynamic, *sorting*, may also arise and result in further systematic differences in outcomes.

Learning refers to accelerators improving the quality of startups through multiple modalities, including intensive consultation with many mentors, training workshops with experienced experts, and transparent interactions with other startups in their cohort (Assenova 2020; Cohen et al., 2019b; Miller et al., 2022). Drawing on fieldwork at eight U.S. accelerators, Cohen and co-authors found that these activities in accelerators help startups learn "how" to do particular tasks (customer development, set the pace of product development, fundraise, etc.), and learn "what" to do in terms of their business model and how to prioritize their next steps (which customer segment to target, go-to-market strategy, etc.) (Cohen et al., 2019a; Hallen et al., 2020). Quantitatively, learning as a key driver of accelerators' effects is supported by Gonzalez-Uribe and Leatherbee's (2018) finding that Start-Up Chile's benefits are contingent on participation in its internal entrepreneurship school and Yu's (2020) finding that participants in top seed accelerators are likely to shutdown sooner.

Unlike learning, the second theorized mechanism – signaling – does not alter the underlying quality of startups. Signaling refers to the logic that accelerators may reliably broadcast the quality of startups to third parties[26]. Credible signals of quality are observable indicators that are: (a) at least partially manipulable by the signaler and (b) have a cost inversely correlated with the underlying quality (Connelly et al., 2011; Spence, 1973). Subsequent work has often defined these criteria somewhat broadly; see for instance Podolny’s (1993) arguments for why the status arising from partner affiliations is a form of signal. Accelerator participants meet these criteria in that: (a) they have to apply to the accelerator and have previously achieved a quality level sufficient to gain acceptance (both of which are partially under their control) and (b) it would be costly / difficult for lower-quality startups to gain admittance and the corresponding signal without spending the time and effort to find an attractive opportunity, make progress, or otherwise improve their prospects. This suggests that accelerator participation may come to be highly correlated with startup quality and be a credible signal of this otherwise hard to observe attribute. To date, however, the empirical evidence for accelerators’ signaling effects is less clear than for learning. For instance, Hallen et al. (2020) failed to find that the effects of examined accelerators were dampened for serial entrepreneurs or those with less startup traction, contingencies that would

---

[26] Following prior accelerator literature, we define the mechanism of signaling broadly to include both any signals arising from publicly-observable participation in an accelerator as well as the effects of accelerators privately endorsing startups to their network of acquaintances (Hallen et al., 2020).

have been expected if signaling were a key driver of accelerators' growth and fundraising benefits.

Researchers have also theorized matching between startups and accelerators may exhibit *sorting*, or the two-sided process by which matches emerge from both parties' preferences (Fox, Yang, & Hsu, 2018; Mindurta et al., 2016). Sorting dynamics may result in a correlational, but not directly causal, relationship between accelerators and various startup outcomes. This is because certain accelerators may be more likely to attract, select, and graduate higher-quality startups. As we develop later, patterns of sorting may give rise to signaling effects over time.

Overall, research to date offers compelling support for seed accelerators' effects on startup growth and capital raising. What is missing is whether accelerator participation helps, hurts, or has no discernible effect on startups obtaining higher-status early partners.

## How SEED Accelerators May Impact Startups' Obtaining Higher or Lower-Status Partners

Status, which may be inferred from and reflected in the status of a firm's affiliates, is often a reliable and insightful proxy for a potential partner's otherwise unobservable quality (Ozmel et al., 2013; Podolny, 1993, 1994). Other organizations' decisions to work with a firm reflect and summarize the detailed, often unobservable information available to the firm's past partners. Status is theoretically distinct from, though often partially correlated with, a firm's reputation – or perceptions of a firm's quality based on its observable past performance and accomplishments (Pollock et al., 2015; Pollock, Lashley, Rindova & Han, 2019).

Startups often seek out high-status partners because they are likely to be especially beneficial (Stuart et al., 1999; Hallen, 2008; Hsu, 2004). Yet inter-organizational partnerships often tend toward status-based (assortative) matching (Gulati & Gargiulo, 1999; Podolny, 1994). High-status firms seek out partnerships with one another, assuming greater mutual value creation, leaving moderate-status firms to partner with other moderate-status firms, and lowest-status firms to partner with one another.

So how do firms, especially startups, form partnerships with firms whose status is higher than their own? The literature suggests two ways, which largely map to the discussed accelerator mechanisms of learning and signaling. First, firms may engage higher-status partners by displaying non-status signals of quality. Such alternative signals are more likely to be impactful on obtaining high-status partners when they are especially reliable (Podolny, 1994) or when a startup is younger and has limited status-conferring affiliations (Hallen, 2008). Thus, by improving startups' observable quality through learning, accelerators may help startups attract higher-status partners.

Second, status may also be improved via the signaling effect of endorsements from arbiters, including evaluators whose job is to discern and communicate quality that might otherwise be opaque. This may occur through rankings, such as those for business and law schools, and prize committees common in scientific and artistic fields (Azoulay et al., 2014; Sauder, 2008). To the extent accelerators are highly-selective, they may be viewed as status-conferring arbiters and this may create a signal of quality that also helps participating startups obtain higher-status partners.

Yet there are also arguments for why signaling efforts made by an accelerator may have no discernible effect. Recent empirical work suggests the signaling effects of status may often be more circumscribed than earlier work had suggested. In a study of life scientists, Azoulay et al. (2014) found that after being named as prestigious Howard Hughes Medical Institute Investigators there was a subsequent increase in citations to recipients' prior publications relative to a closely matched sample, but the average effect was surprisingly small (less than 1 additional citation per a year) and fleeting. Similarly, Simcoe and Waguespack (2011) study leveraged a natural experiment where the Internet Engineering Task Force randomly replaced some individual author names with "et al." for proposals under consideration. While this replacement had a substantial impact on publication rates by high-status authors, the effect disappeared completely for a prescreened set of proposals whose quality was clearer. Overall, this research suggests that when potential partners can reliably evaluate a startup's quality through other means, accelerators may not have a discernible signaling effect for participants.

It is also possible that any signaling effect of accelerators might be negative. Even if an accelerator selects only the highest-quality applicants and further improves startup quality through learning, at graduation accelerated startups might still be of lower-quality than startups seeking high-status partners directly (i.e., without participating in an accelerator). This may arise if there are sorting dynamics whereby the highest-quality startups do not apply to an accelerator, either due to associated costs (time, equity) or out of concerns for what participation might signal about their startup. Over time, this might lead to a recursive "market for lemons" dynamic, whereby accelerators are perceived as having lower-quality graduates, which may further deter higher-quality applicants (Akerlof, 1970; Feltovich, Harbaugh, & To, 2002). This may in turn lead high-status partners to interpret accelerator participation as a negative signal of quality and so participants may ultimately obtain lower-status partners than they might have otherwise.

## Accelerator, Region, and Entrepreneur Contingencies

Our above arguments suggest the directionality of accelerators' aggregate effects on helping startups obtain high-status partners is somewhat unclear theoretically and likely to be heterogeneous. We thus consider potential contingencies at the accelerator, region, and entrepreneur-levels that may contribute to this heterogeneity. Understanding contingencies may also offer insight into the presence of theorized mechanisms of learning and signaling (cf. Elfenbein, Hamilton, & Zenger, 2010).

At the accelerator-level, there are reasons to suspect time-invariant (fixed effect) differences. Accelerators may have relatively persistent design differences (Cohen et al., 2019a), imprinted at establishment (Boeker, 1989; Stinchcombe, 1965). The status of an accelerators' graduates might also lead to a recursive rich-get-richer and poor-get-poorer dynamic referred to in sociology as a "Matthew Effect" (Azoulay et al., 2014; Merton, 1968; Zuckerman, 1977). Accelerators whose graduates previously obtained high-status partners may broadcast a stronger signal of quality, attracting even higher-quality applicants, further amplifying positive effects over time. In contrast, accelerators whose graduates obtain lower-status partners may enter a downward spiral where they are perceived as a market-for-lemons, which may increasingly deter high-quality applicants and high-status partners.

Contingencies may also exist at the region-level. Startups in less active entrepreneurial regions may have fewer alternative sources for learning and signaling, which might amplify any accelerator effect. Then again, accelerators in less active entrepreneurial regions may have more limited and less knowledgeable mentor pools, as well as fewer high-status partners with whom to connect graduates, possibly reducing any effect.

Finally, effects may depend on entrepreneur-level contingencies. We would expect effects driven by signaling to be dampened for startups whose founders have other strong signals of human capital (Podolny, 1994; Stuart et al., 1999). Additionally, entrepreneurial and scientific fields often exhibit gender-based biases (Ding, Ohyama, & Agarwal, 2021; Guzman & Kacperczyk, 2019). Thus, following Lall and coauthors' (2020) finding amongst impact accelerators, all male teams may benefit more from seed accelerators.

### Summary

Overall, these logics suggest accelerators' effect on the status of startups' early partners might range from positive to negative, as well as heterogeneity in effects and contingencies at the accelerator, region, and entrepreneur-levels. We next examine these outcomes using two complementary analyses, with our first drawing on a smaller dataset offering more rigorous causal inference but less ability to examine generalizability across accelerators and our second drawing on a larger dataset that allows greater examination of accelerator heterogeneity, contingencies, and mechanisms but with less rigorous causal identification.

## METHODS & RESULTS

### Analysis I: Almost Accepted Sample

#### Almost Accepted Sample and Methods

Our first set of analyses draw on a sample of confidential quantitative data on 235 accepted or almost accepted applicants to four accelerator cohorts across three different accelerators in 2011 and 2012. A key advantage of this sample is it allows a rigorous (though not perfect) causal inference approach consistent with that used in other recent research on how accelerators impact survival, growth and funding (Gonzalez-Uribe & Leatherbee, 2018; Hallen et al., 2020; Lall et al., 2020). These accelerators, including their selection processes are described in further detailed in Hallen et al. (2020). All sampled accelerators were amongst the earliest accelerators founded in the U.S., relatively high-status at the time and have been found to have some combination of effects on startup performance (see also Appendix G). Each cohort operated in a different region, all with at least moderately active entrepreneurial ecosystems and none in Silicon Valley. To the extent our theory suggests any accelerators might have positive influences on helping startups obtain high-status partners, we would expect to find it in accelerators such as these.

Selecting startups is an immensely important part of acceleration, playing a key role in any signaling effects of accelerators; additionally, it underlies our causal inference strategy. Out of more than 3,100 original applicants to these four cohorts, our sample began with the 263 startups that made it to the "round before last" in a multi-round selection process. Cohorts were selected in a holistic manner to ensure a balance of industries and reduce competition between startups and there were often heated disagreements amongst directors as to which startups to ultimately accept – attesting to the comparability of the accepted and almost accepted startups' potential.

We believe this "almost accepted" sample allows a second-best approach to disentangling a causal treatment effect from any underlying, non-causal, sorting dynamics – especially as we are not aware of a valid instrument for accelerator participation. This allows sampling a comparison set of startups based on rich information and expert evaluations generally unobservable to researchers. While particularly well-suited to estimating *if any accelerators have a causal effect on obtaining higher-status investors*, there are tradeoffs. The samples' small size reduces the ability to examine generalizability and heterogeneity across accelerators and makes estimates more sensitive to random variance in a small number of outcomes. These limitations are a motivation for our use of a second, larger complementary Sample II that makes a different set of trade-offs and may be viewed as "quasi-replication" within our paper (Bettis, Helfat, & Shaver, 2016).

Access to these confidential, inside data was generously provided by the accelerators. To preserve confidentiality, we refer to applicant pools by pseudonyms and years: B-2012, E-2012, X-2012, and E-2011 (with E representing the accelerator for which we have both 2012 and 2011 cohort data). We agreed not to contact applicants directly, but were able to gather founder biographies from LinkedIn and public sources for 44 of the accepted and 191 of the almost accepted startups (235 startups total). Five of the almost accepted startups applied to multiple sampled cohorts and three were ultimately accepted to one of them; we thus include these startups multiple times (though some covariates change as they applied to cohorts at different times).

We first collected investment data from Crunchbase, a prominent source of startup investment in recent academic studies (Lyons & Zhang, 2018; Piezunka & Dahlander, 2015; Wang, Pahnke, & McDonald, 2022; Yu, 2020). Crunchbase collects its data from 3,500 global investment firms as well as an active community of contributors, and then verifies it through a human data team and AI/machine learning algorithms. See Retterath and Braun (2020) for validation of Crunchbase in European venture capital investments. For startups not listed as having investments in Crunchbase, we searched Pitchbook, which uses similar data gathering methods and is the official data provider for the National Venture Capital Association, as well as searching the Web. This brought the total number of startups with identified equity investors to 85 out of 235[27]. We also identified 5 startups with undisclosed investors and 2 receiving grants; we omit these from the main analyses but report robustness analyses including them.

To account for remaining observable differences between startups, we hand-collected a rich set of founder and startup characteristics. We used these data in two of our main estimation approaches — inverse probability of treatment weights (IPTW) and propensity score matching (PSM) — to further improve the observable similarity of the comparison sets and reduce sample selection bias. We detail both approaches in the Results section.

## Measures

Dependent Variable: Our focal dependent variable is the maximum status of any identified investor in a startup in the five years following the start of accelerator participation. As cohort start dates are not all public, we assumed the start date to be 90 days before each accelerator's demo

---

[27] We identified investments from disclosed investors for 58 startups in Crunchbase, an additional 25 startups in Pitchbook, and then a further 2 via web searches. 15 of the Pitchbook additions were in accelerator startups, with one-third in a focal accelerator and two-thirds in others. The additional startups tended to have raised less capital (with a median of 0.4M vs. 1.38M) and from lower status investors (with the highest status investor having a median status of 0 vs. 1.32). Robustness tests utilizing just the investments identified in Crunchbase yielded broadly similar results. Important for future research, this indicates Pitchbook may have more complete coverage of investments into startups that have raised less than $1M and from lower-status investors.

day. Early investments came from a combination of venture capitalists, angels, and other forms of equity investors and therefore we did not impose restrictions around investor type. We explicitly exclude investments from accelerators or dedicated follow-on funds that guarantee investment.

We constructed our status measure for each investor based on their syndication-based affiliations as measured in Crunchbase. Our status measure is a transformation of ranked investor eigenvector centrality in the five-year co-investment network. Eigenvector centrality in the venture finance co-investment network has been found predictive of both investor outcomes and the attraction of higher-potential entrepreneurs (Hallen, 2008; Hochberg, Ljungqvist, & Lu, 2007). While the ordinal ranking we created by eigenvector centrality aligned well with perceptions in the industry, the *raw eigenvector centralities* exaggerated differences between very top-ranked investors relative to industry perceptions (e.g., showing very large differences between status ranks 2 vs 3 that did not align with our experience in the field). At the same time, the *raw ordinal rankings* exaggerated differences amongst lower-status investors, where perceived status differences are very slight.

Accordingly, we use a transformed status measure that preserves the ordinal ranking of eigenvector centrality but improves cardinality by diminishing differences between consecutive status positions at lower levels of status. Drawing on Phillips and Zuckerman's (2001) argument that there are generally three tiers of status in a field but that audiences still draw distinctions within tiers, we adjusted the status differences between consecutive status positions (e.g., position 75 vs. 76) using two inflection points. We selected inflection points around the 50th and 51st ranked investors in a given year and between the 250th and 251st based on field interactions and recent academic research (McDonald & Eisenhardt, 2019). We assigned each investor a score depending on their annual tier (top, middle, bottom) and ranking within tiers. We began by rank ordering all investors in a given year by their eigenvector centrality $ec_{i,y}$, such that $ec_{i,y} = 1$ for the investor with the highest eigenvector centrality in a year y, the second ranked investor in that year had a score of $ec_{i,y} = 2$, and so forth. We then calculated status $s_{i,y}$ for investor *i* in year *y*:

$$\boldsymbol{s_{i,y}} = \begin{cases} \boldsymbol{3 - ec_{i,y} / 50} & if\ 1 \le ec_{i,y} \le 50 \\ \boldsymbol{2 - (ec_{i,y} - 50) / 200} & if\ 51 \le ec_{i,y} \le 250 \\ \boldsymbol{1 - (ec_{i,y} - 250)} / 750 & if\ 251 \le ec_{i,y} \le 1000 \\ 0 & if\ 1001 \le ec_{i,y} \end{cases}$$

Higher scores indicate higher status. See Appendix H for further analysis of “who are high-status investors”. Robustness tests using alternative cutoffs, as well as raw eigenvector centrality, yielded broadly similar results as described later.

Accelerator Participation: Given that our empirical analyses focus on the subset of startups that ultimately raise equity investments, we examine the collective effect of participating in any of these accelerator cohorts as measured by the dummy *focal accelerator participant*. Additionally, 25 of the initially “almost accepted” startups later participated in other accelerator programs, some of equal prestige to the focal cohorts and some lower. We capture this with the dummy variable *other accelerator participant*.

Control Variables: We follow prior literature and include a number of founding team and startup measures to control for remaining differences between the accepted and almost accepted startups (Eisenhardt & Schoonhoven, 1996; Hsu, 2007). Many of these control variables are further elaborated in Hallen et al. (2020). We include *2 or 3 founders* as our fieldwork suggested such teams were often desirable to accelerators as they struck a balance between having greater human and social capital than solo founders while requiring fewer early resources to support than larger teams. We controlled for *has female founders* following literature suggesting that investors may exhibit biases toward female entrepreneurs. We include *fulltime at application* as the percentage of founders working full time on the startup at their time of application to the accelerator. We also include dummy variables indicating if any of a startup's founders had completed the following degrees at the time of application: MBA, JD (law degree), technical masters, or technical PhD. Following Rider (2012), we include *university prominence* as the maximum prestige of the universities according to *U.S. News & World Report's* 2012 ranking of the top 400 global universities, with scores ranging from 100 to 29.2 and graduates of universities outside this list being assigned a score of 28. We control for *prior employer prominence* using the same scale as Hallen et al. (2020), measured as the logged number of startups in their sample that share the same prior employer as one of the startups' founders.

We controlled for prior entrepreneurial experience by measuring if any startups' founders had previously started another startup (*serial entrepreneur)*, and whether any of these startups had raised venture capital *(previously raised VC)*. We also accounted for the extent of founders' professional experience by measuring *years work experience* as the average number of years between their undergraduate degree and a startups' founding; this measure was logged to reduce skew. We set this measure to 0 in cases where founders did not list their graduation year and include a dummy variable indicating such startups.

Finally, we controlled for startups' *prior web traffic* as the average daily page views (in thousands) of startups' website over the thirty days preceding the start of the accelerator.

### Results for Almost Accepted Analyses

Table 9 presents the descriptive statistics and pairwise correlations for the set of 85 startups that were either accepted or almost accepted to the focal cohorts and which went onto raise investments from disclosed investors within five years. See Appendix G for the same statistics for the full set of 235 startups that were either accepted or almost accepted, as well as for a summary of the impact of these accelerators on survival / being acquired, employee growth, web traffic, and raising investments (with the first three outcomes also examined in Hallen et al. 2020).

Table 9 examines whether sampled accelerators helped or hurt startups in obtaining higher-status investors. Model 1 begins with an OLS estimate amongst the 85 startups that raise investments. This estimate indicates participation in the focal accelerators is associated with a 0.472 increase in investor status (p=0.064).

Models 2 and 3 address two forms of potential bias. In Model 2, we utilize inverse probability of treatment weights (IPTW) to place the greatest emphasis on startups that appeared closest to admission cutoff at the time of application (i.e., barely accepted or barely rejected) based on observable factors (Hirano & Imbens, 2001). Weights were calculated using logit estimates of the likelihood of a startup participating in the focal accelerators based on the control variables. Model 2 estimates an average accelerator treatment effect of 0.546 (p=0.013).

Model 3 uses propensity score matching (PSM) to help account for sample selection bias (Certo, Busenbark, Woo, & Semadeni, 2016; Heckman, 1977), following Wolfolds and Siegel's (2019) recommendation to do so in contexts where a valid exclusion restriction (instrument) is not available and as weak instruments might introduce greater bias (Bettis, Gambardella, Helfat, & Mitchell, 2014; Certo et al., 2016). Since participating in an accelerator may also increase the likelihood of a startup raising funds, accelerators may have some graduates that raise funds from lower-status investors but would have been unlikely to have raised at all had they not participated in an accelerator. Estimations not correcting for such sample selection bias might thus underestimate any accelerator effect. Bias could also arise if our multisource data gathering systematically missed any investments in startups. PSM helps correct for these biases by focusing comparisons on accelerator versus non-accelerator startups that go onto raise investments and that appear highly similar on observables. Given we are focused on the treatment effect of the focal accelerators, we omit startups participating in other accelerators prior to the matching. PSM yields an estimated average focal accelerator treatment effect of 0.446 (p=0.025).

Overall, the effects of participating in the focal accelerators are relatively similar across these estimates, with the strongest evidence (lowest p values) coming from the matching models that account for remaining observable differences between treated and untreated startups and potential sample selection biases. The propensity score estimates indicate that participating in the focal accelerators provides a status increase of just under a half-tier (e.g., increasing from an investor ranked 200 to one ranked 106), an effect we interpret as relatively sizable. The consistency of the focal accelerator effect across the alternative weightings also reduces concerns our results may be driven by randomness in the outcomes of a small number of startups.

The participants in other accelerators reflect a broader range of accelerators, but not a representative sample. We observe they on average also had a discernible, though slightly lower, effect in both the OLS estimate (β=0.370; p=0.042) and in the IPTW estimate (β=0.443; p=0.041); by design, these startups are excluded from the PSM estimate.

Figure 4 presents a coefficient plot based on the propensity score matching estimates from Model 3. The point estimate for focal accelerator participation is one of the larger effects relative to other measures. We say this with some caution, though, as the sample was constructed based on a set of startups that the focal accelerators perceived as quite similar in potential. We also note that some attributes are quite infrequent (previously raised VC or JD) and their estimates shift more across Models 1 through 3, suggesting a few observations may drive these results.

We ran additional analyses to explore the impact of key empirical choices that are further detailed in the Appendix I. For our status measure, we explored alternative inflection points and using eigenvector centrality; in all cases finding broadly similar results. Results were also broadly consistent when including the five startups with undisclosed investors (and assuming these investors were of status 0) or including grants. Exploratory analyses omitting the top 10% of outcomes from each group also yielded broadly similar results, though there were indications that the effect of *other accelerator participant* might be driven by the highest status outcomes. This suggests there may be heterogeneity in the extent to which accelerators primarily aid their average versus top participants. The estimated effect was also higher ($\beta=0.810$; $p=0.013$) when using the *teffects psmatch* command in Stata that accounts for the estimated nature of the propensity scores in calculating standard errors, suggesting our estimates in Table 10 may be conservative.

Taken as a whole, our almost accepted analyses repeatedly indicate the focal accelerators likely had a causal and non-trivial impact on the status of startups' early investors that extended above any underlying sorting dynamics. Thus, it appears at least *some* accelerators positively help startups gain higher-status early partners.

## Analysis II: U.S. Startup DataSet

### U.S. Startup Sample and Methods

A limitation of Sample I is that its smaller size restricts the extent to which it may be used to explore generalizability and heterogeneity across accelerators, as well as accelerator, region, and entrepreneur-level contingencies. Yet it is these weaknesses of Sample I that are the strengths of our Sample II. It is a larger dataset that captures a high-fraction of early-stage venture capital investments in U.S. technology startups from 2005 through 2017. A tradeoff, though, is it allows weaker causal inference, relying on observable control variables to disentangle accelerators' treatment effects from the confounding effect of any sorting. It is for these reasons that we view our two samples as both important, together having complementary strengths and tradeoffs while offering a degree of "quasi-replication" of core findings (Bettis et al., 2016).

We draw investment data from Crunchbase because it allows access to all investment activity for a given time window. Given the geographic representation of Sample I, we similarly restricted Sample II to startups headquartered in the United States. We also excluded biotechnology startups from the sample, as their participation in accelerators was limited and as research has not yet validated the efficacy of accelerators on aiding startups in this sector.

While Crunchbase sometimes lists accelerator participation, it was often incomplete. Accordingly, we triangulated accelerator participation using three additional datasets: the Seed Accelerator Rankings Project (SARP), Seed-DB, and yclist.com[28]. We manually verified that all accelerators were batched, fixed-length programs focusing on mentoring and education and were seed accelerators (i.e., not corporate or impact accelerators). We retained accelerators that limited participation to particular founder groups (e.g., StartX focusing on founders with Stanford affiliations) so long as they met our other criterion. We focus on accelerator participants raising initial post-accelerator funding between 2005 (the year of the first Y Combinator cohort) and 2017 (we obtained our Crunchbase data in the summer of 2018, thus providing a six-month window for completed deals to enter our dataset).

We restricted our sample to first rounds from at least one identifiable investor and involving investments of $250,000 or more as entrepreneurs reported this as the minimum for meaningfully impacting startup success and longevity. We included rounds occurring within the first five years of a startup’s life and excluded initial rounds exceeding $25 million (the largest initial round observed of an accelerator participant). For accelerator startups, we also excluded

---

[28] SARP is an annual survey of accelerator program directors who provide information on their alumni; this project grew out of Northwestern University in 2010 and is now overseen by a team of academics at Rice University, the University of Georgia, and the University of Southern California. Seed-DB is a list of accelerator graduates maintained by Jed Christensen who began the project as a graduate student at Oxford University and then was employed by Techstars. yclist.com is a crowdsourced listing of companies that have participated in Y Combinator.

initial rounds occurring more than six-months before the accelerator graduation; we included rounds after this point, as they might reflect signaling effects of being accepted into an accelerator. We also restricted our analysis to rounds that were either equity investments or convertible into equity investments (i.e., convertible notes or SAFEs). We excluded 14 startups missing state information. This yielded 9,966 early rounds from the same number of startups, of which 1,501 participated in 111 different accelerators (some had only a few graduates raising investments).

To account for differences between founding teams, we gathered data from LinkedIn for the subset of startups where Crunchbase listed founders' LinkedIn profiles and profiles were publicly accessible (3,702 startups or 37% of the sample). As discussed below, we used survey-based weights to account for observable differences in this information being available for only some startups. See Appendix H for further analyses of where founder information was available versus absent. Finally, and as elaborated below, we also gathered additional founder and startup data from Genderize.io, BuiltWith, and the Internet Archive.

## Measures

*Dependent Variable:* Our dependent variable, *investor status*, is constructed in the same manner and using the same Crunchbase data as in Sample I. As this analysis is at the round-level, we took the highest status of all investors in a focal round and omitted investments from the focal accelerator or any dedicated follow-on funds.

*Independent Variables:* We use *accelerator-level fixed effects* to explore accelerator heterogeneity and explore design and other time-invariant differences.

*Accelerator-level Matthew Effects:* We measured *status of investors in accelerator alumni* to capture potential recursive dynamics around a Matthew Effect. We measured this as the 90th percentile of investor status (for the highest status investor) for graduates from the same accelerator over the three years ending the day before the focal round. We chose the 90th percentile to balance audiences being more tuned to the actions of higher-status actors (Simcoe & Waguespack, 2011) but to also avoid sensitivity to outliers. This measure was set to zero for rounds where an accelerator did not have any graduates raise VC rounds in the last 3 years.

*Region-level Contingencies*: We measured *state 3-year VC rounds* as the number of venture capital deals (rounds) in the state over the last 3 years, ending the year before the focal deal and logged to reduce skew. We also report robustness to alternative measures.

*Founder-level Contingencies and Control Variables:* We controlled for observable differences in founding teams, startups, and deals. Our founding team measures largely mirror those from the almost accepted sample, except we measured *prior employer prominence* using the set of prior employers in this sample. Additionally, some startups' founders were missing either educational or job data on their LinkedIn profiles. Where available for some founders, we determined the control variables using the available data. Where one of these was missing for all of a startups' founders, we set the relevant controls to zero and noted this with the dummy variables *founders' jobs missing* and *founders' education missing*. We inferred *has female founders* using Genderize.io and an 80% confidence threshold (Marx, 2021). Where inconclusive, we used the gender data specified by the user on LinkedIn or inferred by LinkedIn from names and pronouns.

*Further Control Variables:* While we used web traffic to control for startup quality in the almost accepted sample, these data are generally only available for the most recent three years. Accordingly, we measured the startup's pre-accelerator quality 436 days before the round[29] using

---

[29] We estimate this is when a startup would likely have been applying to an accelerator. We estimated this time by examining the distribution of the days between graduating from an accelerator (demo day) and the focal rounds in the subsample of accelerator startups. We then took the 75th percentile of 316 days from graduation to raising and added 120 days to reflect that most accelerators last 3-4 months – so a total of 436 days before the round. Using the 50th percentile (151 days + 120 days) and the 90th percentile (587 days + 120 days) yielded highly similar results.

*Internet Archive crawls* and *BuiltWith technology stack*. Internet Archive crawls is the number of times the Internet Archive "crawled" the startup's website over the 3 months preceding this date, logged to reduce skew. This is a function of the number of other sites that are linking to a website and reflects a websites' centrality on the overall internet.

Our second measure draws on data from BuiltWith, which captures many aspects of the technical stacks used to build startups' websites and analytics (Koning, Hasan, & Chatterji, 2020). A total of 1,091 different technologies were used by at least 50 startups in the sample. Use of each technology was measured longitudinally as a dummy variable. Given the high dimensionality of the original data and a lack of theoretical priors, we used a LASSO to select technologies to include in our main analyses while avoiding overfitting (Tibshirani, 1996). To ensure our feature selection models were trained on different data than our main analyses, we ran the LASSO regression on the 6,264 startups lacking founder education and work history and using the status of investors as the dependent variable; in doing so, we used the adaptive approach to select the lambda (Choudhury, Allen, & Endres, 2021). This yielded 33 technologies most predictive of differences in the status of a startups' investors. Following Belloni, Chernozhukov, and Hansen (2014), we ran a second LASSO model on this training set predicting accelerator participation (our treatment) as otherwise the LASSO might create omitted variable bias by excluding variables more correlated with the treatment; this yielded 25 measures, 1 of which overlapped with the prior selection.

At the deal level, we measured *startup age* in years at the time of the round. We measured *current round amount* as the size of the round in millions of dollars; logged to reduce skew (after adding a value of 1). We measured *previously raised capital* in millions of dollars and logged to reduce skew (after adding a value of 1). We included a dummy variable indicating if Crunchbase listed the round as a convertible note. We also included the *percentage of VC* investors in a round to account for baseline differences in the status of different types of investors.

To account for regional differences, we included state level fixed effects, combining states with fewer than 10 entries in our sample to reduce overfitting. Crunchbase specified startups' industries using a nonexclusive tagging system, such that a startup may be tagged as being part of zero, one, or any combination of 47 categories. Accordingly, we captured startups' industry using a 47-dimensional vector, with the measure for each category listed for a startup being assigned a value of: *1 / # Categories Listed for the Startup;* startups not listed as belonging to that category were assigned a value of 0. We captured temporal differences using round-year fixed effects.

## Results for Analysis II using Accelerator Participation within U.S. Startup Sample

### *Exploring Effects of and Heterogeneity Across Accelerators*

Table 11 reports descriptive statistics and correlations on the full sample. Startups were on average 2.28 years old. The average syndicate's highest status investor had a status of 0.995, or ranked about 250th in a given year. Table 12 presents models estimating how accelerators differ in the extent to which they help or hurt startups in gaining higher-status investors. We include fixed effects for all accelerators, with non-participation being the omitted category and report the coefficients for the 12 largest accelerators by volume of startups in the sample for brevity. Model 4 is estimated on the full sample, but omits founder and startup performance-level measures because they are only available for a subset of startups. Model 4 is thus a *naïve analysis*, in that accelerator coefficients are likely capturing a combination of accelerator treatment effects *and* sorting. As expected, the point estimates of the coefficients vary substantially across accelerators (e.g., MassChallenge having a coefficient of 0.175 (p=0.000) while Y Combinator has a coefficient of 1.091 (p=0.000)).

Model 5 restricts the analysis to startups where founder data was available, but does not yet add founder controls, to understand if narrowing to this subset introduces bias in estimates.

This and subsequent models utilize sample survey weights to give greater emphasis to startups more representative of others dropped for lack of founder data. Weights were calculated as the inverse of the estimated probability of founder data being available for a given startup. See Appendix J for details. In most cases accelerator coefficients remain quite similar across the full sample estimate in Model 4 and the founder subsample estimate in Model 5. For a small number of accelerators, however, changes in estimates are more pronounced – e.g., Techstars Seattle shrinking from $\beta=0.920$ to $\beta=0.598$ and Alchemist Accelerator from $\beta=0.564$ to $\beta=0.112$. For these accelerators, this indicates point estimates may be more influenced by potential bias arising from founder information only being available for some startups.

Model 6 adds measures for founder controls and startup progress. Coefficients for founder control variables are broadly in line with expectations. Relative to Model 4 that omits founder measures, Model 6 improves the predictive power (Adjusted $R^2$) from 0.266 to 0.304, likely accounting for some sorting dynamics. We thus utilize Model 6 to answer our research question: *do accelerators help, hurt, or have no discernible effect on startups in obtaining higher-status partners?* All shown top 12 accelerators by volume of funded startups have positive effects (with the highest p-value being 0.01) relative to otherwise similar startups that did not participate in an accelerator. Yet, these accelerators are highly heterogeneous in their effects.

Model 7 reports the robustness of these estimates accounting for potential sample-selection bias in terms of which startups are raising investment rounds using propensity-score matching. This approach also helps account for other potential forms of selection bias, including founder or funder information only being available for some startups, at least to the extent this bias is correlated with observables in our founding team, round, or other controls (Wolfolds & Siegel, 2019). While matching drops the sample size by about half, the coefficients across Model 6 (using founder survey weights) and Model 7 (using propensity score matching) are similar with some exceptions (e.g., DreamIt's coefficient drops from 0.442 to 0.282; Techstars Seattle's coefficient increases from 0.452 to 0.787, bringing it closer to the estimate in Model 4).

Coefficient estimates for the top 25 accelerators by number of funded startups are displayed in Figure 5. The effects range from slightly negative (Blueprint Health's β=-0.155, p=0.000) to nearly a full jump in status "level" (Y Combinator's β=0.890, p=0.000). To put Y Combinator's estimated coefficient in context, the effect is roughly equivalent to jumping from an investor ranked #250 in the status hierarchy to one ranked #72. Striking is how many accelerators have a positive effect, even if only slight. Of note, accelerators with positive effects are located in a variety of regions, not only top entrepreneurial hubs. Offering further evidence for heterogeneity in effects across accelerators, the confidence intervals for the effects of accelerators often do not overlap (Figure 5). Additionally, in unreported tests, we obtained highly similar results using the same alternative status measures that we used in robustness tests for Sample I[30]. Overall, our results consistently indicate that many of the largest accelerators likely have a positive effect, though some may have a slightly negative effect and heterogeneity is substantial.

### ***Exploring Accelerator, Region, and Entrepreneur-Level Contingencies***

Table 13 presents analyses exploring potential drivers of this heterogeneity at the accelerator, region, and entrepreneur-levels. Models in Table 13 are estimated using the subsample of startups for which founder data were available and using survey-based weights to account for the availability of founder information. We continue to include accelerator fixed effects with non-

---

[30] Coefficients were generally smaller when we defined the inflection points as occurring at rankings 26 and 100, while coefficients were generally larger when we defined the inflection points as occurring at rankings 101 and 500. For eigenvector centrality, the rank orderings of the coefficients across accelerators were generally consistent, though some had high p-values (e.g., Alchemist Accelerator, Boost VC, DreamIt, and StartX all had p values greater than 0.50 when status was measured using eigenvector centrality).

participation as the omitted category to account for time-invariant differences across accelerators.

In Model 8 we explore potential status-based Matthew Effects at the accelerator-level, which would be evident if there is a recursive influence of the status of the investors in an accelerator's alumni over the preceding three years. Counter to the expected positive effect predicted by a Matthew Effect (Azoulay et al., 2014; Merton, 1968; Zuckerman, 1977), the coefficient for *status of investors in accelerator alumni* is slightly negative and has a high p-value ($\beta$=-0.037; p=0.424). Results were similar when we ran variants of this measure using the average of alumni status (instead of the 90th percentile) and 2-year and 1-year windows for alumni investments[31]. Moreover, these results did not appear to be driven by a measurement issue, as unreported models omitting accelerator fixed effects showed a clearer positive relationship between the status of investors in an accelerator's alumni and the status of investors in its newest graduates ($\beta$=0.175; p=0.003). We interpret these findings as indicating that the status of investors in an accelerator's alumni is a good proxy for an accelerator's effect, but that we do not observe evidence of a rich-get-richer and poor-get-poorer Matthew Effect in accelerators over time. We return to this unexpected finding in the discussion.

In Model 9 we explore whether accelerators' effects vary with entrepreneurial activity in a region, measured by logged VC rounds in the state over the preceding three years. We do not observe any clear pattern from this interaction ($\beta$=-0.035; p=0.221), nor in an unreported model using regional entrepreneurial activity in place of accelerator fixed effects ($\beta$=0.0480; p=0.198). Results were highly similar when we instead used total rounds in the state (not just VC rounds), total number of investors, and total number of VC investors. These results suggest accelerator

---

[31] One exception to note is that when measuring alumni investor status using a 2-year window and the 90th percentile, the coefficient had a negative and more consistent effect ($\beta$=-0.076; p=0.12).

status effects do not systematically vary with regional entrepreneurial activity.

Next, we focus on entrepreneur-level contingencies. We first explore signaling-related contingencies in Models 10 and 11. Following arguments that any signaling effects of status should be dampened when other signals of quality are present (or vice-versa) (Azoulay et al., 2014; Podolny, 1994; Simcoe & Waguespack, 2011), we interact accelerator participation with the entrepreneurs' prior employer prominence (Model 10) and university prominence (Model 11). We see negative effects in both cases. For accelerator participants, the effect of prior employer prominence falls by 59% (the interaction β=-0.038, p=0.071, with prior employer prominence otherwise having a β=0.064, p=0.000). Similarly, we see accelerator participation lowers the effect of university prominence by 50% (the interaction β=-0.002, p=0.053, with university prominence otherwise having a β=0.004, p=0.000). Overall, this pattern is consistent with signaling being at least a partial driver of the observed accelerator effects. This also suggests accelerators may have a reduced effect for startups with other strong signals.

We next explore potential gender-related contingencies in Model 12. Here we do not observe any clear pattern around teams with female founders exhibiting a dampened accelerator effect (β=-0.085; p=0.333). In an unreported robustness test, we also ran the analysis using *all female founders*. This interaction also did not exhibit a clear pattern (β=-0.101; p=0.406), though the main effect of *all female founders* was negative (β=-0.098; p=0.000). These results indicate that accelerators in this sample do not appear to amplify gender-related inequalities, but they also do not appear to dampen these inequalities.

### *Exploring Improvements in Startup Quality via Mediation Analysis*

While our entrepreneur-level contingencies provide some support for the presence of signaling mechanisms in driving observed accelerator effects, they did not directly test the theorized learning and quality-improvement mechanism of accelerators. Accordingly, we conducted a mediation analysis and examined whether accelerator participation was associated with changes in observable startup quality. We again measured startup quality using *Internet Archive crawls* and the *BuiltWith technical stack*, though measuring each in the month before the focal investment round (after any accelerator participation). Our expectation was that these measures might only capture a fraction of any quality changes and we did not necessarily expect a large mediation effect. Following current best practices, we used bootstrapping procedures on a structural equation model (Shrout & Bolger, 2002; Vissa, 2012). As with our contingency analyses, we focus on the mediated effect of *any* accelerator participation.

We estimated our structural equation models using the Lavaan package in R, which restricts the number of measures included in structural equation models. We reduced the dimensionality of our models using a LASSO regression, calculated in Stata using investor status as the dependent variable, including all controls from Model 6 plus the additional quality measures, and forcing inclusion of accelerator participation. This identified 45 measures, including eight BuiltWith measures (Apache, Facebook CDN, Global Sign, GoDaddy SSL, Google, Script Aculo US, Twitter CDN, and the Twitter Follow Button). We then added time-lagged versions of these BuiltWith measures and *Internet Archive crawls* measured at 436 days before the round to ensure mediation was capturing changes in observable quality, bringing the number of measures to 54.

Results of the bootstrap mediation analysis are presented in Table 14. We find that the estimated net indirect effect of participating in any accelerator is 0.016 (with a 95% confidence interval from 0.003 to 0.030), which is about 3.5% of the estimated total effect of 0.455 (with a 95% confidence interval from 0.379 to 0.531). Most of this appears due to changes in Internet Archive crawls, and not technical stacks. While the mediated effect is small, it indicates accelerated startups have greater traction and visibility – though it is less clear to what extent this is due to learning versus signaling (had the BuiltWith technical stack been a stronger mediator, this would have offered more compelling evidence of a quality-improvement effect).

### *Summary of Sample II Analyses*

Overall, findings from Sample II complement our findings from Sample I: many accelerators help startups obtain higher-status early investors though heterogeneity is high. Observed effects are largely independent of accelerator-level Matthew Effects, the extent of entrepreneurial activity in the region, and the gender of founding teams. We also observe patterns consistent with accelerator effects being driven in part by signaling, though our results around learning-based quality improvements are less conclusive.

## DISCUSSION

In this paper we examine how seed accelerators impact startups obtaining high-status early partners. Across two complementary samples, we find some accelerators are associated with startups raising from higher-status investors than they would likely have obtained otherwise. Yet the heterogeneity of effects appears high across accelerators, with many accelerators having only slight positive effects or even slight negative effects. We also observe indications that these effects are driven in part by quality-relaying signaling, though our tests of improvements in startup quality are less conclusive. Unexpectedly, we *do not* observe indications of an accelerator-level Matthew Effect, a consistent relationship between regional entrepreneurial activity and accelerator effects, or that effects vary with the gender composition of founding teams.

### Implications for Research on Accelerators and Entrepreneurial Development Programs

Our primary contributions are to the literature on accelerators as a rising and critical component of the entrepreneurial ecosystem (Cohen et al., 2019a; Gonzalez-Uribe & Leatherbee, 2018; Grimes, 2018; Hallen et al., 2020; Lall et al., 2020; Yu, 2020). While much extant work has explored if accelerators impact startup development, including customer traction and raising capital, we show that many accelerators also impact startups' social positions in terms of the status of early post-accelerator partners. This is a particularly important outcome for many entrepreneurs, since high-status partners generally offer greater benefits to startups in terms of their resources, knowledge, connections, and the status they confer (Hallen et al., 2014; Hsu, 2004; Ozmel et al., 2013; Podolny, 1993; Stuart et al., 1999). Many accelerators thus not only aid in startup survival and growth, but act as "status springboards" quickly vaulting startups' social positions.

We see evidence for this across both samples, with Sample I providing particularly compelling evidence that at least some well-regarded accelerators have a causal treatment effect and Sample II providing greater insight into generalizability and heterogeneity. We were particularly struck by finding positive effects for 21 of the top 25 accelerators by volume of funded startups, though effect size varied substantially (Figure 5). Stepping back, our findings indicate that not only do some accelerators have a broader range of beneficial effects than previously identified, but that a broader set of accelerators may offer positive effects than previously suggested – even if some effects are mild and other accelerators have slight negative effects.

The observed contingencies also offer novel insights. The dampened accelerator effect for entrepreneurs with university and employment-related signals of quality is consistent with the previously theorized signaling effect of accelerators. However, to our knowledge, ours is the first quantitative evidence supporting this mechanism. On the other hand, we did not observe differences in effects by gender, suggesting many seed accelerators may have managed to avoid biases found in other parts of the entrepreneurial ecosystem (Ding, Ohyama, & Agarwal, 2021; Guzman & Kacperczyk, 2019; Lall et al., 2020). This is consistent with our own experience with seed accelerators, but further research is needed to better understand when and how gender differences may alter the effects of accelerators. Surprisingly accelerator effects appear independent from entrepreneurial activity in a region. Possibly underlying our observation is that many seed accelerators are founded by entrepreneurs who were successful in major entrepreneurial hubs prior to moving home to start accelerators in less-entrepreneurial regions (Fehder & Hochberg, 2022); thus, accelerators may import knowledge and networks through their founders - another opportunity for future research.

## Contributions to the Literature on Networks and Status

The literature has long highlighted the importance of partnerships and networks for firms and startups, and that high-status partners are especially valuable (Podolny, 1993; Stam, Arzlanian & Elfring, 2014; Stuart et al., 1999). Yet this literature also suggests that obtaining high-status partners is both difficult, due to high-status partners tending to partner with one another, and frequently slow, due to the recursive influence of other status-conferring affiliations (Ozmel et al., 2013; Podolny & Phillips, 1996; Pollock et al., 2015).

We contribute to this literature by showing that relatively short programs, like accelerators, can function as springboards that help organizations quickly obtain higher-status partners than have been observed in other settings (Azoulay et al., 2014; Simcoe & Waguespack, 2011). What drives this difference? One possibility is accelerators' coupling of quality-improving activities with quality-relaying signaling, which may produce a more unique and informative signal. Future research might explore if status springboards effects generalize to other interventions and contexts.

The unexpected absence of rich-get-richer Matthew Effects raises interesting questions for the literature on organizational status (Merton, 1977; Sauder, Lynn, & Podolny, 2012; Zuckerman, 1977). One possibility is that the initial managing directors (founders) of an accelerator may exert a strong imprinting effect, influencing the network of the accelerator via their personal networks (Hallen, 2008; Milanov & Shepherd, 2013). It may also be that Matthew Effect dynamics exist but so do counteracting effects. For instance, it may be higher-status accelerators foster their own ecosystems of newer (and lower status) investors. Overall, we believe this research highlights the value of revisiting classic theories in novel settings (O'Mahony & Cohen, 2022).

### Implications for Entrepreneurs, Accelerator Directors, and Policy Makers

For entrepreneurs, our findings consistently suggest that participating in the right accelerators can help them obtain higher-status investors and partners than they would likely obtain on their own. This is an important insight, since the market for lemons concern is not simply an academic theory, but something we repeatedly heard from entrepreneurs and other stakeholders in entrepreneurial ecosystems. While the theorized "market for lemons" dynamic (Akerlof, 1970) does not appear to be dominant amongst the studied accelerators, future research might examine if market for lemons dynamics are more prevalent in other types of accelerators or regions.

Also, important for entrepreneurs are our findings that a wide set of accelerators may have positive benefits and these benefits do not appear dependent on local entrepreneurial activity. For creators of new accelerators or the policy makers supporting them, this finding also suggests accelerators may be beneficial across many areas in a high-income country like the United States, including outside of major entrepreneurial hubs.

## Future Directions

While our first sample is highly suggestive of causality, future research could scrutinize causality further, for example using instruments or randomized control trials (e.g., perhaps following the approach of Camuffo, Cordova, Gambardella, & Spina, 2020). Likewise, while we utilize two samples with different inference approaches, there remains an opportunity for further quasi-replication especially across other geographies or types of accelerators (Bettis et al., 2016). And while our analyses indicate that past status and regional entrepreneurial activity are not key drivers in the observed heterogeneity across accelerators, this leaves open what does drive differences across accelerators' effects? Empirically, we also note that our structure of Sample II may offer a template for further exploring heterogeneity across accelerators' effects, especially given the challenges in assembling almost accepted data across many accelerators. Here it may especially be interesting to use this approach to explore how accelerator participation impacts experimentation (Camuffo et al., 2020; Leatherbee & Kaitla, 2020). It might also be interesting to examine how accelerators impact other status dynamics, including who startups hire or their ability to win grants or awards in fields where that is more prevalent. In conclusion, we contribute to theory and practice by showing a broader set of seed accelerators provide a broader set of benefits than previously suggested.

# CONCLUSION

The strategies entrepreneurs use, characterized by 'doing' and 'thinking,' are continually evolving, influenced by the introduction of new support systems and the development of innovative technology tools. This dissertation has explored the dynamic nature of entrepreneurial strategy, highlighting that modern approaches such as employing data analytics or participating in seed accelerators can significantly enhance resource mobilization outcomes. However, it also demonstrates that the impact of these strategizing outcomes has limitations within the broader context.

Looking ahead, my research will continue to delve into these topics, focusing particularly on a deeper theoretical and empirical analysis of technology as a tool for entrepreneurial learning and resource mobilization. This exploration will extend beyond merely examining the adoption or abandonment of tools to explore how and why technology can either enhance or constrain the cognitive abilities of entrepreneurs, as well as uncovering optimal practices and patterns around the adoption and utilization of these tools as a means to facilitate learning, resource mobilization, and ultimately, growth.

To expound on the gap in literature, over the past two decades, technology tools have increasingly played a critical role in shaping organizational outcomes (Brynjolfsson & McElheran, 2016; Ewens, Nanda, & Rhodes-Kropf, 2018; Raisch & Krakowski, 2021; Argote & Hora, 2017). These tools influence not only minor day-to-day decisions but also major strategic pivots across a wide array of organizations, including new ventures (Brynjolfsson & McElheran, 2016; Koning, Hasan, & Chatterji, 2022).

Despite their increased prevalence and theoretical significance, the optimal practices and patterns of technology tool usage in new ventures remain largely unexplored. Insights from tool adoption in mature organizations (e.g., Kane & Alavi, 2007; Berman & Israeli, 2022) offer a starting point, but these insights do not always translate directly to new ventures, which operate under different conditions characterized by uncertainty, resource constraints, and distinct managerial cognition (Stinchcombe, 1965; Huang & Pearce, 2015; Packard, Clark, & Klein, 2017).

These fundamental differences likely influence how, why, and by whom a tool is adopted, in turn impacting its overall effectiveness. For instance, the effectiveness of a tool can vary greatly depending on its fit with the user's capabilities (Garfield, Taylor, Dennis, & Satzinger, 2001) and the specific mix of technology tools employed, which can have synergistic or antagonistic effects depending on their configuration (Katila & Ahuja 2002, Hansen, Nohria, & Tierney, 1999). Additionally, the timing of tool adoption is crucial, potentially increasing or decreasing its effectiveness (Carlson & Zmud, 1999; Hansen & Haas, 2001). Understanding the conditions under which tool efficacy is maximized is thus an important yet underexplored area of research, particularly given the diverse skillsets and experiences of entrepreneurs who utilize these tools (Koning et al., 2022).

In summary, while the importance and complexities of technology tools are well recognized, there is a significant gap in our understanding of how specific technology tools augment learning or provide an advantage in resource mobilization or growth outcomes in new ventures. Existing studies on larger established organizations provide a foundation for predictions, but these may not always be applicable due to the highly contingent nature of tool efficacy on specific members, tasks, other tools, and the surrounding environment. Given that new ventures often operate in contexts markedly different from those of established organizations, the factors affecting tool efficacy are likely to differ as well. Moreover, it is crucial to consider that tool efficacy may be influenced not only by the treatment effects but also by selection effects. Therefore, my future work aims to focus more on theorizing and empirically testing how entrepreneurs and new ventures effectively use technology tools to their advantage.

# REFERENCES

## Introduction & Conclusion

Adner, R., & Levinthal, D. (2008). Doing versus seeing: Acts of exploitation and perceptions of exploration. *Strategic Entrepreneurship Journal*, *2*(1), 43–52.

Agarwal, R., Echambadi, R., Franco, A. M., & Sarkar, M. B. (2004). Knowledge transfer through inheritance: Spin-out generation, development, and survival. *Academy of Management Journal*, *47*(4), 501–522.

Akerlof, G. A. (1970). The Market for "Lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics, 84*(3), 488-500. https://doi.org/10.2307/1879431

Andries, P., Debackere, K., & Van Looy, B. (2013). Simultaneous experimentation as a learning strategy: Business model development under uncertainty. *Strategic Entrepreneurship Journal*, *7*(4), 288–310.

Argote, L., & Hora, M. (2017). Organizational learning and management of technology. *Production and Operations Management*, *26*(4), 579–590.

Assenova, V. A. (2020). Early-stage venture incubation and mentoring promote learning, scaling, and profitability among disadvantaged entrepreneurs. *Organization Science*, *31*(6), 1560–1578.

Azoulay, P., Stuart, T., & Wang, Y. (2014). Matthew: Effect or fable? *Management Science, 60*(1), 92-109. https://doi.org/10.1287/mnsc.2013.1755

Baker, T., & Nelson, R. E. (2005). Creating something from nothing: Resource construction through entrepreneurial bricolage. *Administrative Science Quarterly*, *50*(3), 329–366.

Batjargal, B., & Liu, M. (2004). Entrepreneurs' access to private equity in China: The role of social capital. *Organization Science*, *15*(2), 159–172.

Baum, J. A., & Silverman, B. S. (2004). Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology startups. *Journal of Business Venturing*, *19*(3), 411–436.

Berman, R., & Israeli, A. (2022). The value of descriptive analytics: Evidence from online retailers. *Marketing Science*, *41*(6), 1074–1096.

Berman, R., & Israeli, A. (2022). The value of descriptive analytics: Evidence from online retailers. *Marketing Science*, *41*(6), 1074–1096.

Bingham, C. B., & Davis, J. P. (2012). Learning sequences: Their existence, effect, and evolution. *Academy of Management Journal*, *55*(3), 611–641.

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.

Brynjolfsson, E., & McElheran, K. (2016). The rapid adoption of data-driven decision-making. *American Economic Review*, *106*(5), 133–139.

Brynjolfsson, E., & McElheran, K. (2016). The rapid adoption of data-driven decision-making. *American Economic Review*, *106*(5), 133–139.

Brynjolfsson, E., & McElheran, K. (2019). Data in action: Data-driven decision making and predictive analytics in US manufacturing. *Rotman School of Management Working Paper, 3422397*.

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in numbers: How does data-driven decision making affect firm performance? *Available at SSRN 1819486*.

Campbell, B. A., Coff, R., & Kryscynski, D. (2012). Rethinking sustained competitive advantage from human capital. *Academy of Management Review*, *37*(3), 376–395.

Carlson, J. R., & Zmud, R. W. (1999). Channel expansion theory and the experiential nature of media richness perceptions. *Academy of Management Journal*, *42*(2), 153–170.

Chatterji, A. K. (2009). Spawned with a silver spoon? Entrepreneurial performance and innovation in the medical device industry. *Strategic Management Journal*, *30*(2), 185–206.

Claes, K., & Vissa, B. (2020). Does social similarity pay off? Homophily and venture capitalists' deal valuation, downside risk protection, and financial returns in India. *Organization Science*, *31*(3), 576–603.

Cohen, S. L., Bingham, C. B., & Hallen, B. L. (2019). The role of accelerator designs in mitigating bounded rationality in new ventures. *Administrative Science Quarterly*, *64*(4), 810–854.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, *37*(1), 39-67. https://doi.org/10.1177/0149206310388419

Contigiani, A., & Young-Hyman, T. (2022). Experimentation, planning, and structure in early-stage ventures: Evidence from pitch decks. *Strategic Entrepreneurship Journal*, *16*(3), 425–459.

Daft, R. L., & Weick, K. E. (1984). Toward a model of organizations as interpretation systems. *Academy of Management Review*, *9*(2), 284–295.

Delmar, F., & Shane, S. (2003). Does business planning facilitate the development of new ventures? *Strategic Management Journal*, *24*(12), 1165–1185.

Eesley, C. E., & Roberts, E. B. (2012). Are you experienced or are you talented?: When does innate talent versus experience explain entrepreneurial performance? *Strategic Entrepreneurship Journal*, *6*(3), 207–219.

Eisenhardt, K. M., & Schoonhoven, C. B. (1990). Organizational growth: Linking founding team, strategy, environment, and growth among US semiconductor ventures, 1978-1988. *Administrative Science Quarterly*, 504–529.

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.

Felin, T., & Zenger, T. R. (2009). Entrepreneurs as theorists: On the origins of collective beliefs and novel strategies. *Strategic Entrepreneurship Journal*, *3*(2), 127–146.

Felin, T., & Zenger, T. R. (2017). The theory-based view: Economic actors as theorists. *Strategy Science*, *2*(4), 258–271.

Felin, T., Gambardella, A., Novelli, E., & Zenger, T. (2023). A scientific method for startups. *Journal of Management*.

Feltovich, N., Harbaugh, R., & To, T. (2002). Too cool for school? Signalling and counter signalling. *RAND Journal of Economics*, 630-649. https://doi.org/10.2307/3087478

Gans, J., & Ryall, M. D. (2017). Value capture theory: A strategic management review. *Strategic Management Journal*, *38*(1), 17–41.

Garfield, M. J., Taylor, N. J., Dennis, A. R., & Satzinger, J. W. (2001). Modifying paradigms—Individual differences, creativity techniques, and exposure to ideas in group idea generation. *Information Systems Research*, *12*(3), 322–333.

Gary, M. S., & Wood, R. E. (2011). Mental models, decision rules, and performance heterogeneity. *Strategic Management Journal*, *32*(6), 569–594.

Gavetti, G., & Rivkin, J. W. (2007). On the origin of strategy: Action and cognition over time. *Organization Science*, *18*(3), 420–439.

Gavetti, G., Levinthal, D., & Ocasio, W. (2007). Perspective—Neo-Carnegie: The Carnegie school's past, present, and reconstructing for the future. *Organization Science*, *18*(3), 523–536.

Ghorbani, A., & Zou, J. (2019). *Data shapley: Equitable valuation of data for machine learning*. 2242–2251.

Gompers, P., & Lerner, J. (2000). Money chasing deals? The impact of fund inflows on private equity valuations. *Journal of Financial Economics*, *55*(2), 281–325.

Gonzalez-Uribe, J., & Leatherbee, M. (2018). The effects of business accelerators on venture performance: Evidence from start-up Chile. *The Review of Financial Studies*, *31*(4), 1566–1603.

Gonzalez-Uribe, J., & Leatherbee, M. (2018). The effects of business accelerators on venture performance: Evidence from start-up chile. *Review of Financial Studies*, *31*(4), 1566-1603. https://doi.org/10.1093/rfs/hhx103

Hall, J., & Hofer, C. W. (1993). Venture capitalists' decision criteria in new venture evaluation. *Journal of Business Venturing*, *8*(1), 25–42.

Hallen, B. L., & Eisenhardt, K. M. (2012). Catalyzing strategies and efficient tie formation: How entrepreneurial firms obtain investment ties. *Academy of Management Journal*, *55*(1), 35–70.

Hallen, B. L., Cohen, S. L., & Bingham, C. B. (2020). Do accelerators work? If so, how? *Organization Science*, *31*(2), 378–414.

Hallen, B. L., Cohen, S. L., & Bingham, C. B. (2020). Do accelerators work? If so, how? *Organization Science*, *31*(2), 378-414. https://doi.org/10.1287/orsc.2019.1304

Hansen, M. T., & Haas, M. R. (2001). Competing for attention in knowledge markets: Electronic document dissemination in a management consulting company. *Administrative Science Quarterly*, *46*(1), 1–28.

Hansen, M. T., Nohria, N., & Tierney, T. (1999). What's your strategy for knowledge management. *Harvard Business Review*, *77*(2), 106–116.

Hochberg, Y. V., Ljungqvist, A., & Lu, Y. (2010). Networking as a barrier to entry and the competitive supply of venture capital. *The Journal of Finance*, *65*(3), 829–859.

Hsu, D. H. (2007). Experienced entrepreneurial founders, organizational capital, and venture capital funding. *Research Policy*, *36*(5), 722–741.

Hsu, D. H., & Ziedonis, R. H. (2013). Resources as dual sources of advantage: Implications for valuing entrepreneurial-firm patents. *Strategic Management Journal*, *34*(7), 761–781.

Huang, L., & Pearce, J. L. (2015). Managing the unknowable: The effectiveness of early-stage investor gut feel in entrepreneurial investment decisions. *Administrative Science Quarterly*, *60*(4), 634–670.

Huber, G. P. (1990). A Theory of the Effects of Advanced Information Technologies on Organizational Design, Intelligence, and Decision Making. *Academy of Management Review*, *15*(1), 47–71.

Kane, G. C., & Alavi, M. (2007). Information technology and organizational learning: An investigation of exploration and exploitation processes. *Organization Science*, *18*(5), 796–812.

Katila, R., & Ahuja, G. (2002). Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of Management Journal*, *45*(6), 1183–1194.

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

Kiss, A. N., & Barr, P. S. (2015). New venture strategic adaptation: The interplay of belief structures and industry context. *Strategic Management Journal*, *36*(8), 1245–1263.

Koning, R., Hasan, S., & Chatterji, A. (2022). Experimentation and start-up performance: Evidence from A/B testing. *Management Science*, *68*(9), 6434–6453.

Koning, R., Hasan, S., & Chatterji, A. (2022). Experimentation and start-up performance: Evidence from A/B testing. *Management Science*, *68*(9), 6434–6453.

Kotha, S., Shin, S. J., & Fisher, G. (2022). Time to unicorn status: An exploratory examination of new ventures with extreme valuations. *Strategic Entrepreneurship Journal*, *16*(3), 460–490.

Lall, S. A., Chen, L.W., & Roberts, P. W. (2020). Are we accelerating equity investment into impact-oriented ventures? *World Development*, *131*, 104952. https://doi.org/10.1016/j.worlddev.2020.104952

March, J. G. (2006). Rationality, foolishness, and adaptive intelligence. *Strategic Management Journal*, *27*(3), 201–214.

Mazmanian, M., & Beckman, C. M. (2018). “Making” Your Numbers: Engendering Organizational Control Through a Ritual of Quantification. *Organization Science*, *29*(3), 357–379. https://doi.org/10.1287/orsc.2017.1185

McDonald, R. M., & Eisenhardt, K. M. (2020). Parallel play: Startups, nascent markets, and effective business-model design. *Administrative Science Quarterly*, *65*(2), 483–523.

Mindruta, D., Moeen, M., & Agarwal, R. (2016). A two-sided matching approach for partner selection and assessing complementarities in partners’ attributes in inter-firm alliances. *Strategic Management Journal*, *37*(1), 206–231.

Moghaddam, K., Bosse, D. A., & Provance, M. (2016). Strategic alliances of entrepreneurial firms: Value enhancing then value destroying. *Strategic Entrepreneurship Journal*, *10*(2), 153–168.

Muller, J. (2018). *The tyranny of metrics*. Princeton University Press.

Murray, A., Rhymer, J., & Sirmon, D. G. (2021). Humans and technology: Forms of conjoined agency in organizations. *Academy of Management Review*, *46*(3), 552–571.

Ott, T. E., Eisenhardt, K. M., & Bingham, C. B. (2017). Strategy formation in entrepreneurial settings: Past insights and future directions. *Strategic Entrepreneurship Journal*, *11*(3), 306–325.

Packard, M. D., Clark, B. B., & Klein, P. G. (2017). Uncertainty types and transitions in the entrepreneurial process. *Organization Science*, *28*(5), 840–856.

Peterson, A., & Wu, A. (2021). Entrepreneurial learning and strategic foresight. *Strategic Management Journal*, *42*(13), 2357–2388.

Podolny, J. M. (1994). Market uncertainty and the social character of economic exchange. *Administrative Science Quarterly*, 458–483.

Politis, D. (2005). The process of entrepreneurial learning: A conceptual framework. *Entrepreneurship Theory and Practice*, *29*(4), 399–424.

Powell, E. E., & Baker, T. (2014). It’s what you make of it: Founder identity and enacting strategic responses to adversity. *Academy of Management Journal*, *57*(5), 1406–1433.

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, *46*(1), 192–210.

Ridway, V. F. (1956). Dysfunctional consequences of performance measurement. *Administrative Science Quarterly*, 240–247.

Sauder, M. (2008). Interlopers and field change: The entry of US news into the field of legal education. *Administrative Science Quarterly, 53*(2), 209-234. https://doi.org/10.2189/asqu.53.2.209

Shapley, L. S. (1953). *A value for n-person games*.

Sharapov, D., Kattuman, P., Rodriguez, D., & Velazquez, F. J. (2021). Using the SHAPLEY value approach to variance decomposition in strategy research: Diversification, internationalization, and corporate group effects on affiliate profitability. *Strategic Management Journal*, *42*(3), 608–623.

Simcoe, T. S., & Waguespack, D. M. (2011). Status, quality, and attention: What's in a (missing) name? *Management Science*, *57*(2), 274-290. https://doi.org/10.1287/mnsc.1100.1270

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 99–118.

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, *87*(3), 355-374. https://doi.org/10.1016/B978-0-12-214850-7.50025-5

Stinchcombe, A. L. (1965). Social structure and organizations. In *Handbook of Organizations* (pp. 142–193). Rand McNally.

Wasserman, N. (2017). The throne vs. The kingdom: Founder control and value creation in startups. *Strategic Management Journal*, *38*(2), 255–277.

Welter, C., Mauer, R., & Wuebker, R. J. (2016). Bridging behavioral models and theoretical concepts: Effectuation and bricolage in the opportunity creation framework. *Strategic Entrepreneurship Journal*, *10*(1), 5–20.

## Chapter 1.

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, *72*(1), 1–19.

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.

Adner, R., & Levinthal, D. (2008). Doing versus seeing: Acts of exploitation and perceptions of exploration. *Strategic Entrepreneurship Journal*, *2*(1), 43–52.

Agarwal, R., Bacco, F., Camuffo, A., Coali, A., Gambardella, A., Msangi, H., Sonka, S. T., Temu, A., Waized, B., & Wormald, A. (2023). Does a theory-of-value add value? Evidence from a randomized control trial with Tanzanian entrepreneurs. *Evidence from a Randomized Control Trial with Tanzanian Entrepreneurs (April 6, 2023)*.

Alaimo, C., & Kallinikos, J. (2022). Organizations Decentered: Data Objects, Technology and Knowledge. *Organization Science*, 33(1), 19–37. https://doi.org/10.1287/orsc.2021.1552

Albin, P., & Foley, D. K. (1998). *Barriers and bounds to rationality: Essays on economic complexity and dynamics in interactive systems*.

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., & Ruef, M. (2006). *Organizations Evolving*.

Allen, R., & Choudhury, P. (2022). Algorithm-augmented work and domain experience: The countervailing forces of ability and aversion. *Organization Science*, *33*(1), 149–169.

Anthony, C. (2021). When knowledge work and analytical technologies collide: The practices and consequences of black boxing algorithmic technologies. *Administrative Science Quarterly*, *66*(4), 1173–1212.

Arriola, B. (2021, March 13). Why Links Are Important for SEO. Search Engine Journal. https://www.searchenginejournal.com/seo/why-links-important-seo/

Arthur, W. B. (1994). Inductive reasoning and bounded rationality. *The American Economic Review*, *84*(2), 406–411.

Bailey, D. E., Faraj, S., Hinds, P. J., Leonardi, P. M., & von Krogh, G. (2022). We are all theorists of technology now: A relational perspective on emerging technology and organizing. *Organization Science*, *33*(1), 1–18.

Ben-Michael, E., Feller, A., & Rothstein, J. (2019). Synthetic controls and weighted event studies with staggered adoption. *arXiv Preprint arXiv:1912.03290*, *2*.

Berman, R., & Israeli, A. (2022). The value of descriptive analytics: Evidence from online retailers. *Marketing Science*, *41*(6), 1074–1096.

Blank, S. (2020). *The four steps to the epiphany: Successful strategies for products that win*. John Wiley & Sons.

Blank, S., & Dorf, B. (2005). The path to epiphany: The customer development model. *The Four Steps to the Epiphany*, 17–28.

Bollen, K. A., Biemer, P. P., Karr, A. F., Tueller, S., & Berzofsky, M. E. (2016). Are survey weights needed? A review of diagnostic tests in regression analysis. *Annual Review of Statistics and Its Application*, *3*, 375–392.

Brynjolfsson, E., & McElheran, K. (2016). The rapid adoption of data-driven decision-making. *American Economic Review*, *106*(5), 133–139.

Brynjolfsson, E., & McElheran, K. (2019). Data in action: Data-driven decision making and predictive analytics in US manufacturing. *Rotman School of Management Working Paper*, *3422397*.

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in numbers: How does data-driven decision making affect firm performance? *Available at SSRN 1819486*.

Brynjolfsson, E., Jin, W., & McElheran, K. (2021). The power of prediction: Predictive analytics, workplace complements, and business performance. *Business Economics*, *56*, 217–239.

Burford, N., Shipilov, A. V., & Furr, N. R. (2022). How ecosystem structure affects firm performance in response to a negative shock to interdependencies. *Strategic Management Journal*, *43*(1), 30–57.

Burton, M. D., Sørensen, J. B., & Beckman, C. M. (2002). Coming from good stock: Career histories and new venture formation. In *Social structure and organizations revisited* (pp. 229–262). Emerald Group Publishing Limited.

Choudhury, P., Allen, R. T., & Endres, M. G. (2021). Machine learning for pattern discovery in management research. *Strategic Management Journal*, *42*(1), 30–57.

Choudhury, P., Starr, E., & Agarwal, R. (2020). Machine learning and human capital complementarities: Experimental evidence on bias mitigation. *Strategic Management Journal*, *41*(8), 1381–1411.

Clough, D. R., Fang, T. P., Vissa, B., & Wu, A. (2019). Turning lead into gold: How do entrepreneurs mobilize resources to exploit opportunities? *Academy of Management Annals*, *13*(1), 240–271.

Cohen, S. L., Bingham, C. B., & Hallen, B. L. (2019). The role of accelerator designs in mitigating bounded rationality in new ventures. *Administrative Science Quarterly*, *64*(4), 810–854.

Conti, A., & Roche, M. P. (2021). Lowering the bar? External conditions, opportunity costs, and high-tech start-up outcomes. *Organization Science*, *32*(4), 965–986.

Conti, R., de Matos, M. G., & Valentini, G. (2023). Big Data Analytics, Firm Size, and Performance. *Strategy Science*.

Covin, J. G., & Slevin, D. P. (1997). High growth transitions: Theoretical perspectives and suggested directions. *Entrepreneurship*, *2000*, 99–126.

Csaszar, F. A., & Levinthal, D. A. (2016). Mental representation and the discovery of new strategies. *Strategic Management Journal*, *37*(10), 2031–2049.

Cyert, R., & March, J. (1963). *A behavioral theory of the firm*.

Daft, R. L., & Weick, K. E. (1984). Toward a model of organizations as interpretation systems. *Academy of Management Review*, *9*(2), 284–295.

Dane, E., Rockmann, K. W., & Pratt, M. G. (2012). When should I trust my gut? Linking domain expertise to intuitive decision-making effectiveness. *Organizational Behavior and Human Decision Processes*, *119*(2), 187–194.

Deephouse, D. L. (2000). Media reputation as a strategic resource: An integration of mass communication and resource-based theories. *Journal of Management*, *26*(6), 1091–1112.

DeSantola, A., Gulati, R., & Zhelyazkov, P. I. (2023). External interfaces or internal processes? Market positioning and divergent professionalization paths in young ventures. *Organization Science*, *34*(1), 1–23.

Doudchenko, N., & Imbens, G. W. (2016). *Balancing, regression, difference-in-differences and synthetic control methods: A synthesis*. National Bureau of Economic Research.

Durand, R., & Khaire, M. (2017). Where do market categories come from and how? Distinguishing category creation from category emergence. *Journal of Management*, *43*(1), 87–110.

Eggers, J., & Song, L. (2015). Dealing with failure: Serial entrepreneurs and the costs of changing industries between ventures. *Academy of Management Journal*, *58*(6), 1785–1803.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of Management Journal*, *50*(1), 25–32.

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.

Farjoun, M. (2008). Strategy making, novelty and analogical reasoning—Commentary on Gavetti, Levinthal, and Rivkin (2005). *Strategic Management Journal*, *29*(9), 1001–1016.

Felin, T., & Koenderink, J. (2022). A generative view of rationality and growing awareness. *Frontiers in Psychology*, *13*, 807261.

Felin, T., & Zenger, T. R. (2009). Entrepreneurs as theorists: On the origins of collective beliefs and novel strategies. *Strategic Entrepreneurship Journal*, *3*(2), 127–146.

Felin, T., & Zenger, T. R. (2017). The theory-based view: Economic actors as theorists. *Strategy Science*, *2*(4), 258–271.

Felin, T., Gambardella, A., Novelli, E., & Zenger, T. (2023). A scientific method for startups. *Journal of Management*.

Felin, T., Koenderink, J., & Krueger, J. I. (2017). Rationality, perception, and the all-seeing eye. *Psychonomic Bulletin & Review*, *24*, 1040–1059.

Fiol, C. M., & Lyles, M. A. (1985). Organizational learning. *Academy of Management Review*, *10*(4), 803–813.

First Round Capital. (n.d.). *The tools early-stage startups actually need to understand their customers*. First Round Review. https://review.firstround.com/the-tools-early-stage-startups-actually-need-to-understand-their-customers/

Fisher, G., Kotha, S., & Lahiri, A. (2016). Changing with the times: An integrated view of identity, legitimacy, and new venture life cycles. *Academy of Management Review*, *41*(3), 383–409.

Fisher, G., Kuratko, D. F., Bloodgood, J. M., & Hornsby, J. S. (2017). Legitimate to whom? The challenge of audience diversity and new venture legitimacy. *Journal of Business Venturing*, *32*(1), 52–71.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. Mcgraw-Hill Book Company.

Fombrun, C., & Shanley, M. (1990). What's in a name? Reputation building and corporate strategy. *Academy of Management Journal*, *33*(2), 233–258.

Gary, M. S., & Wood, R. E. (2011). Mental models, decision rules, and performance heterogeneity. *Strategic Management Journal*, *32*(6), 569–594.

Gavetti, G. (2012). PERSPECTIVE—Toward a behavioral theory of strategy. *Organization Science*, *23*(1), 267–285.

Gavetti, G., Levinthal, D. A., & Rivkin, J. W. (2005). Strategy making in novel and complex worlds: The power of analogy. *Strategic Management Journal*, *26*(8), 691–712.

Gavetti, G., Levinthal, D., & Ocasio, W. (2007). Perspective—Neo-Carnegie: The Carnegie school's past, present, and reconstructing for the future. *Organization Science*, *18*(3), 523–536.

Ghosh, S., Thomke, S., & Pourkhalkhali, H. (2020). *The effects of hierarchy on learning and performance in business experimentation*. *2020*(1), 20500.

Gilboa, I., & Schmeidler, D. (2000). Case-based knowledge and induction. *IEEE Transactions on Systems, Man, and Cybernetics-PART A: Systems and Humans*, *30*(2), 85–95.

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

Gulati, R., & Gargiulo, M. (1999). Where do interorganizational networks come from? *American Journal of Sociology*, *104*(5), 1439–1493.

Halberstam, D. (1972). *The best and the brightest*.

Hallen, B. L. (2008). The causes and consequences of the initial network positions of new organizations: From whom do entrepreneurs receive investments? *Administrative Science Quarterly*, *53*(4), 685–718.

Hallen, B. L., & Pahnke, E. C. (2016). When do entrepreneurs accurately evaluate venture capital firms' track records? A bounded rationality perspective. *Academy of Management Journal*, *59*(5), 1535–1560.

Hallen, B. L., Cohen, S. L., & Bingham, C. B. (2020). Do accelerators work? If so, how? *Organization Science*, *31*(2), 378–414.

Hallen, B. L., Cohen, S. L., & Park, S. H. (2023). Are seed accelerators status springboards for startups? Or sand traps? *Strategic Management Journal*, *44*(8), 2060–2096.

Hallen, B. L., Davis, J. P., & Murray, A. (2020). Entrepreneurial network evolution: Explicating the structural localism and agentic network change distinction. *Academy of Management Annals*, *14*(2), 1067–1102.

Hernán, M. A., & Robins, J. M. (2016). Using big data to emulate a target trial when a randomized trial is not available. *American Journal of Epidemiology*, *183*(8), 758–764.

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

Hsu, D. H. (2004). What do entrepreneurs pay for venture capital affiliation? *The Journal of Finance*, *59*(4), 1805–1844.

Huber, G. P. (1990). A Theory of the Effects of Advanced Information Technologies on Organizational Design, Intelligence, and Decision Making. *Academy of Management Review*, *15*(1), 47–71.

Huber, G. P. (1991). Organizational learning: The contributing processes and the literatures. *Organization Science*, *2*(1), 88–115.

Imai, K., Kim, I. S., & Wang, E. H. (2023). Matching methods for causal inference with time-series cross-sectional data. *American Journal of Political Science*, *67*(3), 587–605.

Kahneman, D. (2013). A perspective on judgment and choice: Mapping bounded rationality. *Progress in Psychological Science around the World. Volume 1 Neural, Cognitive and Developmental Issues.*, 1–47.

Kohavi, R., Deng, A., & Vermeer, L. (2022). *A/B testing intuition busters: Common misunderstandings in online controlled experiments*. 3168–3177.

Koning, R., Hasan, S., & Chatterji, A. (2022). Experimentation and start-up performance: Evidence from A/B testing. *Management Science*, *68*(9), 6434–6453.

Krakowski, S., Luger, J., & Raisch, S. (2023). Artificial intelligence and the changing sources of competitive advantage. *Strategic Management Journal*, *44*(6), 1425–1452.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2010). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*.

Lebovitz, S., Levina, N., & Lifshitz-Assaf, H. (2021). IS AI GROUND TRUTH REALLY TRUE? THE DANGERS OF TRAINING AND EVALUATING AI TOOLS BASED ON EXPERTS'KNOW-WHAT. *MIS Quarterly*, *45*(3).

Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science*, *33*(1), 126–148.

Lee, S. (2022). The myth of the flat start-up: Reconsidering the organizational structure of start-ups. *Strategic Management Journal*, *43*(1), 58–92.

Levitt, B., & March, J. G. (1988). Organizational learning. *Annual Review of Sociology*, *14*(1), 319–338.

Mantere, S., & Ketokivi, M. (2013). Reasoning in organization science. *Academy of Management Review*, *38*(1), 70–89.

March, J. G. (1978). Bounded rationality, ambiguity, and the engineering of choice. *The Bell Journal of Economics*, 587–608.

March, J. G. (2006). Rationality, foolishness, and adaptive intelligence. *Strategic Management Journal*, *27*(3), 201–214.

March, J. G., & Simon, H. A. (1958). *Organizations*.

Mazmanian, M., & Beckman, C. M. (2018). "Making" Your Numbers: Engendering Organizational Control Through a Ritual of Quantification. *Organization Science*, *29*(3), 357–379. https://doi.org/10.1287/orsc.2017.1185

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., & Barton, D. (2012). Big data: The management revolution. *Harvard Business Review*, *90*(10), 60–68.

McDonald, R. M., & Allen, R. T. (2022). A spanner in the works: Category-spanning entrants and audience valuation of incumbents. *Strategy Science*, *7*(3), 190–209.

McDonald, R. M., & Eisenhardt, K. M. (2020). Parallel play: Startups, nascent markets, and effective business-model design. *Administrative Science Quarterly*, *65*(2), 483–523.

McDonald, R., & Gao, C. (2019). Pivoting isn't enough? Managing strategic reorientation in new ventures. *Organization Science, 30*(6), 1289–1318.

McGrath, R. G. (2023). Who Learns Fastest, Wins: Lean Startup and Discovery Driven Growth. *Journal of Management*, 01492063231204870.

Miller, A., O'Mahony, S., & Cohen, S. L. (2024). Opening the aperture: Explaining the complementary roles of advice and testing when forming entrepreneurial strategy. *Organization Science*, *35*(1), 1–26.

Mintzberg, H. (1989). *The structuring of organizations*. Springer.

Mokalis, A. L., & Davis, J. J. (2018). *Google Analytics Demystified (4th Edition)*.

Muller, J. (2018). *The tyranny of metrics*. Princeton University Press.

Murray, A., Rhymer, J., & Sirmon, D. G. (2021). Humans and technology: Forms of conjoined agency in organizations. *Academy of Management Review*, *46*(3), 552–571.

Murray, F., & Tripsas, M. (2004). The exploratory processes of entrepreneurial firms: The role of purposeful experimentation. In *Business strategy over the industry lifecycle* (pp. 45–75). Emerald Group Publishing Limited.

Nelson, R. R., & Winter G, S. (1982). *An evolutionary theory of economic change*.

Neustadt, R. E., & May, E. R. (1986). *Thinking in time: The uses of history for decision makers*.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review*, *90*(4), 339.

Ozmel, U., Reuer, J. J., & Gulati, R. (2013). Signals across Multiple Networks: How Venture Capital and Alliance Networks Affect Interorganizational Collaboration. *Academy of Management Journal*, *56*(3), 852–866. https://doi.org/10.5465/amj.2009.0549

Ozmel, U., Reuer, J. J., & Wu, C. (2017). Interorganizational Imitation and Acquisitions of High-tech Ventures. *Strategic Management Journal*, *38*(13), 2647–2665. https://doi.org/10.1002/smj.2666

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

Petkova, A. P. (2016). Standing out or blending in? The formation of new firms' legitimacy and reputation under different levels of market uncertainty. *Corporate Reputation Review, 19*, 22–34.

Petkova, A. P., Rindova, V. P., & Gupta, A. K. (2008). How can new ventures build reputation? An exploratory study. *Corporate Reputation Review*, *11*, 320–334.

Petkova, A. P., Rindova, V. P., & Gupta, A. K. (2013). No news is bad news: Sensegiving activities, media attention, and venture capital funding of new technology organizations. *Organization Science*, *24*(3), 865–888.

Podolny, J. M. (1994). Market uncertainty and the social character of economic exchange. *Administrative Science Quarterly*, 458–483.

Pollock, T. G., & Gulati, R. (2007). Standing out from the crowd: The visibility-enhancing effects of IPO-related signals on alliance formation by entrepreneurial firms. *Strategic Organization*, *5*(4), 339–372.

Pollock, T. G., Rindova, V. P., & Maggitti, P. G. (2008). Market watch: Information and availability cascades among the media and investors in the US IPO market. *Academy of Management Journal*, *51*(2), 335–358.

Poppo, L., & Zenger, T. (1998). Testing alternative theories of the firm: Transaction cost, knowledge-based, and measurement explanations for make-or-buy decisions in information services. *Strategic Management Journal*, *19*(9), 853–877.

Rao, H. (1994). The social construction of reputation: Certification contests, legitimation, and the survival of organizations in the American automobile industry: 1895–1912. *Strategic Management Journal*, *15*(S1), 29–44.

Ravasi, D., Rindova, V., Etter, M., & Cornelissen, J. (2018). The formation of organizational reputation. *Academy of Management Annals*, *12*(2), 574–599.

Ridway, V. F. (1956). Dysfunctional consequences of performance measurement. *Administrative Science Quarterly*, 240–247.

Rindova, V. P., & Kotha, S. (2001). Continuous "morphing": Competing through dynamic capabilities, form, and function. *Academy of Management Journal*, *44*(6), 1263–1280.

Rindova, V. P., Petkova, A. P., & Kotha, S. (2007). Standing out: How new firms in emerging markets build reputation. *Strategic Organization*, *5*(1), 31–70.

Rindova, V. P., Williamson, I. O., & Petkova, A. P. (2010). Reputation as an intangible asset: Reflections on theory and methods in two empirical studies of business school reputations. *Journal of Management*, *36*(3), 610–619.

Rindova, V. P., Williamson, I. O., Petkova, A. P., & Sever, J. M. (2005). Being good or being known: An empirical examination of the dimensions, antecedents, and consequences of organizational reputation. *Academy of Management Journal*, *48*(6), 1033–1049.

Roberts, J. H., & Lattin, J. M. (1997). Consideration: Review of research and prospects for future insights. *Journal of Marketing Research*, *34*(3), 406–410.

Rosenbaum, P. R., & Rubin, D. B. (1983). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society: Series B (Methodological)*, *45*(2), 212–218.

Russell Merz, G., & Sauber, M. H. (1995). Profiles of managerial activities in small firms. *Strategic Management Journal*, 16(7), 551–564. https://doi.org/10.1002/smj.4250160705

Sah, R. K., & Stiglitz, J. E. (1986). The architecture of economic systems: Hierarchies and polyarchies. *American Economic Review*, *76*(4).

Salancik, G. R., & Pfeffer, J. (1978). A social information processing approach to job attitudes and task design. *Administrative Science Quarterly*, 224–253.

Santos, F. M., & Eisenhardt, K. M. (2009). Constructing markets and shaping boundaries: Entrepreneurial power in nascent fields. *Academy of Management Journal*, *52*(4), 643–671.

Simcoe, T. S., & Waguespack, D. M. (2011). Status, quality, and attention: What's in a (missing) name? *Management Science*, *57*(2), 274–290.

Simon, H. A. (1947). *Administrative behavior*.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 99–118.

Simon, H. A. (1964). Theories of Bounded Rationality. *CIP Working Paper #66.*

Simon, H. A. (1979). Rational decision making in business organizations. *The American Economic Review*, *69*(4), 493–513.

Simon, H. A. (2013). *Administrative behavior (4th Edition)*.

Sorenson, O., & Stuart, T. E. (2001). Syndication networks and the spatial distribution of venture capital investments. *American Journal of Sociology*, *106*(6), 1546–1588.

Spence, A. M. (1974). *Market Signaling: Informational Transfer in Hiring and Related Screening Processes*.

Stinchcombe, A. L. (1965). Social structure and organizations. In *Handbook of Organizations* (pp. 142–193). Rand McNally.

Stuart, T. E. (2000). Interorganizational alliances and the performance of firms: A study of growth and innovation rates in a high-technology industry. *Strategic Management Journal*, *21*(8), 791–811.

Taeuscher, K., & Rothe, H. (2024). Entrepreneurial framing: How category dynamics shape the effectiveness of linguistic frames. *Strategic Management Journal*, *45*(2), 362–395.

Townsend, D. M., Hunt, R. A., McMullen, J. S., & Sarasvathy, S. D. (2018). Uncertainty, knowledge problems, and entrepreneurial action. *Academy of Management Annals*, *12*(2), 659–687.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science, 185*(4157), 1124–1131.

Tversky, A., & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In *Progress in social psychology* (pp. 49–72). Psychology Press.

Vanacker, T., & Forbes, D. P. (2016). Disentangling the Multiple Effects of Affiliate Reputation on Resource Attraction in New Firms. *Organization Science*, *27*(6), 1525–1547. https://doi.org/10.1287/orsc.2016.1090

Wildavsky, A. (1979). *The art and craft of policy analysis*.

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.

Yli-Renko, H., Autio, E., & Sapienza, H. J. (2001). Social capital, knowledge acquisition, and knowledge exploitation in young technology-based firms. *Strategic Management Journal*, *22*(6-7), 587–613.

Zuzul, T., & Edmondson, A. C. (2017). The advocacy trap: When legitimacy building inhibits organizational learning. *Academy of Management Discoveries*, *3*(3), 302–321.

## Chapter 2.

Agarwal, R., Campbell, B. A., Franco, A. M., & Ganco, M. (2016). What do I take with me? The mediating effect of spin-out team size and tenure on the founder–firm performance relationship. *Academy of Management Journal*, *59*(3), 1060–1087.

Agarwal, R., Echambadi, R., Franco, A. M., & Sarkar, M. B. (2004). Knowledge transfer through inheritance: Spin-out generation, development, and survival. *Academy of Management Journal*, *47*(4), 501–522.

Aldrich, H. E., & Fiol, C. M. (1994). Fools rush in? The institutional context of industry creation. *Academy of Management Review*, *19*(4), 645–670.

Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, *11*, 685–725.

Batjargal, B., & Liu, M. (2004). Entrepreneurs' access to private equity in China: The role of social capital. *Organization Science*, *15*(2), 159–172.

Baum, J. R., Locke, E. A., & Smith, K. G. (2001). A multidimensional model of venture growth. *Academy of Management Journal*, *44*(2), 292–303.

Bengtsson, O., & Hsu, D. H. (2015). Ethnic matching in the US venture capital market. *Journal of Business Venturing*, *30*(2), 338–354.

Berle, A. A., & Means, G. G. C. (1932). *The modern corporation and private property*. Transaction publishers.

Bernstein, S., Korteweg, A., & Laws, K. (2017). Attracting early-stage investors: Evidence from a randomized field experiment. *The Journal of Finance*, *72*(2), 509–538.

Brush, T. H., & Bromiley, P. (1997). What does a small corporate effect mean? A variance components simulation of corporate and business effects. *Strategic Management Journal*, *18*(10), 825–835.

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.

Chatterji, A. K. (2009). Spawned with a silver spoon? Entrepreneurial performance and innovation in the medical device industry. *Strategic Management Journal*, *30*(2), 185–206.

Chen, H., Gompers, P., Kovner, A., & Lerner, J. (2010). Buy local? The geography of venture capital. *Journal of Urban Economics*, *67*(1), 90–102.

Choudhury, P., Allen, R. T., & Endres, M. G. (2021). Machine learning for pattern discovery in management research. *Strategic Management Journal*, *42*(1), 30–57.

Claes, K., & Vissa, B. (2020). Does social similarity pay off? Homophily and venture capitalists' deal valuation, downside risk protection, and financial returns in India. *Organization Science*, *31*(3), 576–603.

Clough, D. R., Fang, T. P., Vissa, B., & Wu, A. (2019). Turning lead into gold: How do entrepreneurs mobilize resources to exploit opportunities? *Academy of Management Annals*, *13*(1), 240–271.

Cohen, S. L., Bingham, C. B., & Hallen, B. L. (2019). The role of accelerator designs in mitigating bounded rationality in new ventures. *Administrative Science Quarterly*, *64*(4), 810–854.

Cohen, W. M., Nelson, R., & Walsh, J. P. (2000). *Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not)*. National Bureau of Economic Research Cambridge, Mass., USA.

Cole, S. R., & Hernán, M. A. (2008). Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology*, *168*(6), 656–664.

Contigiani, A., & Levinthal, D. A. (2019). Situating the construct of lean start-up: Adjacent conversations and possible future directions. *Industrial and Corporate Change*, *28*(3), 551–564.

Davila, A., Foster, G., & Gupta, M. (2003). Venture capital financing and the growth of startup firms. *Journal of Business Venturing*, *18*(6), 689–708.

Dencker, J. C., & Gruber, M. (2015). The effects of opportunities and founder experience on new firm performance. *Strategic Management Journal*, *36*(7), 1035–1052.

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

DeSantola, A., Gulati, R., & Zhelyazkov, P. I. (2022). External Interfaces or Internal Processes? Market Positioning and Divergent Professionalization Paths in Young Ventures. *Organization Science*.

Downes, D. H., & Heinkel, R. (1982). Signaling and the valuation of unseasoned new issues. *The Journal of Finance*, *37*(1), 1–10.

Dushnitsky, G., Piva, E., & Rossi-Lamastra, C. (2022). Investigating the mix of strategic choices and performance of transaction platforms: Evidence from the crowdfunding setting. *Strategic Management Journal*, *43*(3), 563–598.

Eckhardt, J. T., Shane, S., & Delmar, F. (2006). Multistage selection and the financing of new ventures. *Management Science*, *52*(2), 220–232.

Eesley, C. E., & Roberts, E. B. (2012). Are you experienced or are you talented?: When does innate talent versus experience explain entrepreneurial performance? *Strategic Entrepreneurship Journal*, *6*(3), 207–219.

Eisenhardt, K. M., & Schoonhoven, C. B. (1990). Organizational growth: Linking founding team, strategy, environment, and growth among US semiconductor ventures, 1978-1988. *Administrative Science Quarterly*, 504–529.

Elbers, B. (2021). A method for studying differences in segregation across time and space. *Sociological Methods & Research*, 0049124121986204.

Fitza, M. A. (2014). The use of variance decomposition in the investigation of CEO effects: How large must the CEO effect be to rule out chance? *Strategic Management Journal*, *35*(12), 1839–1852.

Fitza, M., Matusik, S. F., & Mosakowski, E. (2009). Do VCs matter? The importance of owners on performance variance in start-up firms. *Strategic Management Journal*, *30*(4), 387–404.

Fox, J. T., Yang, C., & Hsu, D. H. (2018). Unobserved heterogeneity in matching games. *Journal of Political Economy*, *126*(4), 1339–1373.

Gao, C., & McDonald, R. (2022). Shaping nascent industries: Innovation strategy and regulatory uncertainty in personal genomics. *Administrative Science Quarterly*, *67*(4), 915–967.

Garson, G. D. (2012). Testing statistical assumptions. *Asheboro, NC: Statistical Associates Publishing*.

Ge, C., Huang, K.-W., & Png, I. P. (2016). Engineer/scientist careers: Patents, online profiles, and misclassification bias. *Strategic Management Journal*, *37*(1), 232–253.

Ghorbani, A., & Zou, J. (2019). Data shapley: Equitable valuation of data for machine learning. *International Conference on Machine Learning*, 2242–2251.

Gilbert, B. A., McDougall, P. P., & Audretsch, D. B. (2006). New venture growth: A review and extension. *Journal of Management*, *32*(6), 926–950.

Gompers, P. A., Gornall, W., Kaplan, S. N., & Strebulaev, I. A. (2020). How do venture capitalists make decisions? *Journal of Financial Economics*, *135*(1), 169–190.

Gompers, P., Kovner, A., Lerner, J., & Scharfstein, D. (2010). Performance persistence in entrepreneurship. *Journal of Financial Economics*, *96*(1), 18–32.

Gompers, P., & Lerner, J. (2000). Money chasing deals? The impact of fund inflows on private equity valuations. *Journal of Financial Economics*, *55*(2), 281–325.

Gonzalez-Uribe, J., & Leatherbee, M. (2018). The effects of business accelerators on venture performance: Evidence from start-up Chile. *The Review of Financial Studies*, *31*(4), 1566–1603.

Guler, I. (2007). Throwing good money after bad? Political and institutional influences on sequential decision making in the venture capital industry. *Administrative Science Quarterly*, *52*(2), 248–285.

Guo, G. (2017). Demystifying variance in performance: A longitudinal multilevel perspective. *Strategic Management Journal*, *38*(6), 1327–1342.

Hall, R. E., & Woodward, S. E. (2010). The burden of the nondiversifiable risk of entrepreneurship. *American Economic Review*, *100*(3), 1163–1194.

Hallen, B. L. (2008). The causes and consequences of the initial network positions of new organizations: From whom do entrepreneurs receive investments? *Administrative Science Quarterly, 53*(4), 685–718.

Hallen, B. L., Cohen, S. L., & Bingham, C. B. (2020). Do accelerators work? If so, how? *Organization Science, 31*(2), 378–414.

Hallen, B. L., Cohen, S. L., & Park, S. H. (2023). Are seed accelerators status springboards for startups? Or sand traps? *Strategic Management Journal*.

Hallen, B. L., Davis, J. P., & Murray, A. (2020). Entrepreneurial network evolution: Explicating the structural localism and agentic network change distinction. *Academy of Management Annals*, *14*(2), 1067–1102.

Hallen, B. L., & Eisenhardt, K. M. (2012). Catalyzing strategies and efficient tie formation: How entrepreneurial firms obtain investment ties. *Academy of Management Journal*, *55*(1), 35–70.

Hashai, N., & Zahra, S. (2022). Founder team prior work experience: An asset or a liability for startup growth? *Strategic Entrepreneurship Journal, 16*(1), 155–184.

Hellmann, T. (2002). A theory of strategic venture investing. *Journal of Financial Economics*, *64*(2), 285–314.

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

Hochberg, Y. V., Ljungqvist, A., & Lu, Y. (2010). Networking as a barrier to entry and the competitive supply of venture capital. *The Journal of Finance*, *65*(3), 829–859.

Hsu, D. H. (2004). What do entrepreneurs pay for venture capital affiliation? *The Journal of Finance, 59*(4), 1805–1844.

Hsu, D. H. (2007). Experienced entrepreneurial founders, organizational capital, and venture capital funding. *Research Policy*, *36*(5), 722–741.

Hsu, D. H., & Ziedonis, R. H. (2013). Resources as dual sources of advantage: Implications for valuing entrepreneurial-firm patents. *Strategic Management Journal, 34*(7), 761–781.

Johnson, R. W. (2001). An introduction to the bootstrap. *Teaching Statistics, 23*(2), 49–54.

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

Kaplan, S. N., Klebanov, M. M., & Sorensen, M. (2012). Which CEO characteristics and abilities matter? *The Journal of Finance*, *67*(3), 973–1007.

Katila, R., Chen, E. L., & Piezunka, H. (2012). All the right moves: How entrepreneurial firms compete effectively. *Strategic Entrepreneurship Journal*, *6*(2), 116–132.

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

Klepper, S., & Sleeper, S. (2005). Entry by spinoffs. *Management Science, 51*(8), 1291–1306.

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

Kotha, S., & Nair, A. (1995). Strategy and environment as determinants of performance: Evidence from the Japanese machine tool industry. *Strategic Management Journal*, *16*(7), 497–518.

Kotha, S., Shin, S. J., & Fisher, G. (2022). Time to unicorn status: An exploratory examination of new ventures with extreme valuations. *Strategic Entrepreneurship Journal*, *16*(3), 460–490.

Kumar, P., Liu, X., & Zaheer, A. (2022). How much does the firm’s alliance network matter? *Strategic Management Journal, 43*(8), 1433–1468.

Lee, M., & Huang, L. (2018). Gender bias, social impact framing, and evaluation of entrepreneurial ventures. *Organization Science*, *29*(1), 1–16.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*.

Magnuson, W. (2022). *For Profit: A History of Corporations*. Hachette UK.

Mallaby, S. (2022). *The power law: Venture capital and the making of the new future*. Penguin.

Marlow, S., & McAdam, M. (2013). Gender and entrepreneurship: Advancing debate and challenging myths; exploring the mystery of the under-performing female entrepreneur. *International Journal of Entrepreneurial Behavior & Research*, *19*(1), 114–124.

Marx, M. (2022). Employee non-compete agreements, gender, and entrepreneurship. *Organization Science*, *33*(5), 1756–1772.

Matusik, S. F., George, J. M., & Heeley, M. B. (2008). Values and judgment under uncertainty: Evidence from venture capitalist assessments of founders. *Strategic Entrepreneurship Journal*, *2*(2), 95–115.

McDonald, R. M., & Eisenhardt, K. M. (2020). Parallel play: Startups, nascent markets, and effective business-model design. *Administrative Science Quarterly*, *65*(2), 483–523.

McGahan, A. M., & Porter, M. E. (1997). How much does industry matter, really? *Strategic Management Journal*, *18*(S1), 15–30.

Meyer-Doyle, P., Lee, S., & Helfat, C. E. (2019). Disentangling the microfoundations of acquisition behavior and performance. *Strategic Management Journal*, *40*(11), 1733–1756.

Mindruta, D., Moeen, M., & Agarwal, R. (2016). A two-sided matching approach for partner selection and assessing complementarities in partners' attributes in inter-firm alliances. *Strategic Management Journal*, *37*(1), 206–231. https://doi.org/10.1002/smj.2448

Misangyi, V. F., Elms, H., Greckhamer, T., & Lepine, J. A. (2006). A new perspective on a fundamental debate: A multilevel approach to industry, corporate, and business unit effects. *Strategic Management Journal*, *27*(6), 571–590.

Moghaddam, K., Bosse, D. A., & Provance, M. (2016). Strategic alliances of entrepreneurial firms: Value enhancing then value destroying. *Strategic Entrepreneurship Journal*, *10*(2), 153–168.

Mullins, J. W. (1996). Early growth decisions of entrepreneurs: The influence of competency and prior performance under changing market conditions. *Journal of Business Venturing*, *11*(2), 89–105.

Podolny, J. M. (1994). Market uncertainty and the social character of economic exchange. *Administrative Science Quarterly*, 458–483.

Porter, M. E. (1980). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. New York: Free Press.

Quigley, T. J., & Graffin, S. D. (2017). Reaffirming the CEO effect is significant and much larger than chance: A comment on Fitza (2014). *Strategic Management Journal*, *38*(3), 793–801.

Quigley, T. J., & Hambrick, D. C. (2015). Has the "CEO effect" increased in recent decades? A new explanation for the great rise in America's attention to corporate leaders. *Strategic Management Journal*, *36*(6), 821–830.

Rider, C. I. (2012). How employees' prior affiliations constrain organizational network change: A study of US venture capital and private equity. *Administrative Science Quarterly*, *57*(3), 453–483.

Rumelt, R. P. (1991). How much does industry matter? *Strategic Management Journal*, *12*(3), 167–185.

Sahlman, W. A., & Willis, M. (2009). *The Basic Venture Capital Formula*. Harvard Business School.

Schmalensee, R. (1985). Do markets differ much? *The American Economic Review*, *75*(3), 341–351.

Shane, S., & Cable, D. (2002). Network ties, reputation, and the financing of new ventures. *Management Science*, *48*(3), 364–381.

Shane, S., & Stuart, T. (2002). Organizational endowments and the performance of university start-ups. *Management Science*, *48*(1), 154–170.

Shapley, L. (1953). Quota solutions op n-person games1. *Edited by Emil Artin and Marston Morse*, 343.

Sharapov, D., Kattuman, P., Rodriguez, D., & Velazquez, F. J. (2021). Using the SHAPLEY value approach to variance decomposition in strategy research: Diversification, internationalization, and corporate group effects on affiliate profitability. *Strategic Management Journal*, *42*(3), 608–623.

Shorrocks, A. F. (2013). Decomposition procedures for distributional analysis: A unified framework based on the Shapley value. *Journal of Economic Inequality*, *11*(1), 99–126.

Sørensen, M. (2007). How smart is smart money? A two-sided matching model of venture capital. *The Journal of Finance*, *62*(6), 2725–2762.

Sorenson, O., & Stuart, T. E. (2001). Syndication networks and the spatial distribution of venture capital investments. *American Journal of Sociology*, *106*(6), 1546–1588.

Stam, W., Arzlanian, S., & Elfring, T. (2014). Social capital of entrepreneurs and small firm performance: A meta-analysis of contextual and methodological moderators. *Journal of Business Venturing*, *29*(1), 152–173.

Thoemmes, F., & Ong, A. D. (2016). A primer on inverse probability of treatment weighting and marginal structural models. *Emerging Adulthood*, *4*(1), 40–59.

Wang, M. zhu. (2023). Changes in industry and corporate effects in the United States, 1978–2019. *Strategic Management Journal*, *44*(2), 477–490.

Wasserman, N. (2012). *The founder's dilemmas: Anticipating and avoiding the pitfalls that can sink a startup*. Princeton University Press.

Wasserman, N. (2017). The throne vs. the kingdom: Founder control and value creation in startups. *Strategic Management Journal*, *38*(2), 255–277.

Yu, S. (2020). How do accelerators impact the performance of high-technology ventures? *Management Science*, *66*(2), 530–552.

## Chapter 3.

Abadie, A., & Imbens, G. W. (2016). Matching on the estimated propensity score. *Econometrica*, *84*(2), 781–807. https://doi.org/10.3982/ECTA11293

Akerlof, G. A. (1970). The Market for "Lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, *84*(3), 488-500. https://doi.org/10.2307/1879431

Assenova, V. A. (2020). Early-stage venture incubation and mentoring promote learning, scaling, and profitability among disadvantaged entrepreneurs. *Organization Science, 31*(6), 1560-1578. https://doi.org/10.1287/orsc.2020.1367

Azoulay, P., Stuart, T., & Wang, Y. (2014). Matthew: Effect or fable? *Management Science, 60*(1), 92-109. https://doi.org/10.1287/mnsc.2013.1755

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, *28*(2), 29-50.

http://doi.org/10.1257/jep.28.2.29

Bettis, R., Gambardella, A., Helfat, C., & Mitchell, W. (2014). Quantitative empirical analysis in strategic management. *Strategic Management Journal*, *35*(7), 949-953. https://doi.org/10.1002/smj.2278

Bettis, R.A., Helfat, C.E. & Shaver, J.M. (2016). The necessity, logic, and forms of replication. *Strategic Management Journal*, *37*(11), 2193-2203. https://doi.org/10.1002/smj.2580

Boeker, W. (1989). Strategic change: The effects of founding and history. *Academy of Management Journal*, *32*(3), 489-515. https://doi.org/10.5465/256432

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. https://doi.org/10.1287/mnsc.2018.3249

Certo, S. T., Busenbark, J. R., Woo, H., & Semadeni, M. (2016). Sample selection bias and Heckman models in strategic management research. *Strategic Management Journal*, *37*(13), 2639-2657. https://doi.org/10.1002/smj.2475

Choudhury, P., Allen, R. T., & Endres, M. G. (2021). Machine learning for pattern discovery in management research. *Strategic Management Journal*, *42*(1), 30-57. https://doi.org/10.1002/smj.3215

Clough, D. R., Fang, T. P., Vissa, B., & Wu, A. (2019). Turning lead into gold: How do entrepreneurs mobilize resources to exploit opportunities? *Academy of Management Annals*, *13*(1), 240-271. https://doi.org/10.5465/annals.2016.0132

Cohen, S. L., Bingham, C. B., & Hallen, B. L. (2019a). The role of accelerator designs in mitigating bounded rationality in new ventures. *Administrative Science Quarterly*, *64*(4), 810-854. https://doi.org/10.1177/0001839218782131

Cohen, S., Fehder, D. C., Hochberg, Y. V., & Murray, F. (2019b). The design of startup accelerators. *Research Policy*, *48*(7), 1781-1797. https://doi.org/10.1016/j.respol.2019.04.003

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, *37*(1), 39-67. https://doi.org/10.1177/0149206310388419

Ding, W. W., Ohyama, A., & Agarwal, R. (2021). Trends in gender pay gaps of scientists and engineers in academia and industry. *Nature Biotechnology*, *39*(8), 1019-1024. https://doi.org/10.1038/s41587-021-01008-0

Dushnitsky G & Sarkar S. (2022). Here comes the sun: The impact of incidental contextual factors on entrepreneurial resource acquisition. *Academy of Management Journal*, *65*(1): 66-92. https://doi.org/10.5465/amj.2019.0128

Dushnitsky, G. & Matusik, S.F., (2019). A fresh look at patterns and assumptions in the field of entrepreneurship: What can we learn?. *Strategic Entrepreneurship Journal*, *13*(4), pp.437-447. https://doi.org/10.1002/sej.1324

Eisenhardt, K. M., & Schoonhoven, C. B. (1996). Resource-based view of strategic alliance formation: Strategic and social effects in entrepreneurial firms. *Organization Science*, *7*(2), 136-150. https://doi.org/10.1287/orsc.7.2.136

Eisenmann, T. R., & Kind, L. (2014) Andreessen Horowitz. *Harvard Business School Case Study*, 814060-PDF-ENG.

Elfenbein, D. W., Hamilton, B. H., & Zenger, T. R. (2010). The small firm effect and the entrepreneurial spawning of scientists and engineers. *Management Science*, *56*(4), 659-681. http://doi.org/10.1287/mnsc.1090.1130

Fehder, D. C., & Hochberg, Y. V. (2022). Accelerators and the regional supply of venture capital investment. *Available at SSRN 2518668*. https://doi.org/10.2139/ssrn.2518668

Feltovich, N., Harbaugh, R., & To, T. (2002). Too cool for school? Signalling and counter signalling. *RAND Journal of Economics*, 630-649. https://doi.org/10.2307/3087478

Fox, J. T., Yang, C., & Hsu, D. H. (2018). Unobserved heterogeneity in matching games. *Journal of Political Economy*, *126*(4), 1339-1373. https://doi.org/10.1086/697740

Gonzalez-Uribe, J., & Leatherbee, M. (2018). The effects of business accelerators on venture performance: Evidence from start-up chile. *Review of Financial Studies*, *31*(4), 1566-1603. https://doi.org/10.1093/rfs/hhx103

Gould, R. V. (2002). The origins of status hierarchies: A formal theory and empirical test. *American Journal of Sociology*, *107*(5), 1143-1178. https://psycnet.apa.org/doi/10.1086/341744

Grimes, M. G. (2018). The pivot: How founders respond to feedback through idea and identity work. *Academy of Management Journal*, *61*(5), 1692-1717. https://doi.org/10.5465/amj.2015.0823

Gulati, R., & Gargiulo, M. (1999). Where do interorganizational networks come from? *American Journal of Sociology*, *104*(5), 1439-1493. https://doi.org/10.1086/210179

Guzman, J., & Kacperczyk, A. O. (2019). Gender gap in entrepreneurship. *Research Policy*, *48*(7), 1666-1680.

Hallen, B. L. (2008). The causes and consequences of the initial network positions of new organizations: From whom do entrepreneurs receive investments? *Administrative Science Quarterly*, *53*(4), 685-718. https://doi.org/10.2189/asqu.53.4.685

Hallen, B. L., Cohen, S. L., & Bingham, C. B. (2020). Do accelerators work? If so, how? *Organization Science*, *31*(2), 378-414. https://doi.org/10.1287/orsc.2019.1304

Hallen, B. L., Katila, R., & Rosenberger, J. D. (2014). How do social defenses work? A resource-dependence lens on technology ventures, venture capital investors, and corporate relationships. *Academy of Management Journal*, *57*(4), 1078-1101. https://doi.org/10.5465/amj.2012.0003

Heckman, J. J. (1977). *Sample selection bias as a specification error (with an application to the estimation of labor supply functions)*. National Bureau of Economic Research. https://doi.org/10.2307/1912352

Hirano, K., & Imbens, G. W. (2001). Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. *Health Services and Outcomes Research Methodology*, *2*(3), 259-278. https://doi.org/10.1023/A:1020371312283

Hochberg, Y. V., Ljungqvist, A., & Lu, Y. (2007). Whom you know matters: Venture capital networks and investment performance. *Journal of Finance*, *62*(1), 251-301. https://doi.org/10.1111/j.1540-6261.2007.01207.x

Hsu, D. H. (2004). What do entrepreneurs pay for venture capital affiliation? *Journal of Finance*, *59*(4), 1805-1844. https://doi.org/10.1111/j.1540-6261.2004.00680.x

Hsu, D. H. (2007). Experienced entrepreneurial founders, organizational capital, and venture capital funding. *Research Policy*, *36*(5), 722-741. https://doi.org/10.1016/j.respol.2007.02.022

Jensen, M., Kim, B. K., & Kim, H. (2011). The importance of status in markets: A market identity perspective. *Status in Management and Organizations*, *48*(2), 87-113. https://doi.org/10.1017/CBO9780511760525.008

Koning, R., Hasan, S., & Chatterji, A. (2022). Experimentation and start-up performance:

Evidence from A/B testing. *Management Science*, *68*(9), 6434-6453. https://doi.org/10.1287/mnsc.2021.4209

Lall, S. A., Chen, L.W., & Roberts, P. W. (2020). Are we accelerating equity investment into impact-oriented ventures? *World Development*, *131,* 104952. https://doi.org/10.1016/j.worlddev.2020.104952

Leatherbee, M., & Gonzalez-Uribe, J. (2018). Selection issues. In Accelerators. Edward Elgar Publishing. https://doi.org/10.4337/9781786434098.00009

Leatherbee, M., & Katila, R. (2020). The lean startup method: Early-stage teams and hypothesis-based probing of business ideas. *Strategic Entrepreneurship Journal*, *14*(4), 570-593. https://doi.org/10.1002/sej.1373

Lyons, E., & Zhang, L. (2018). Who does (not) benefit from entrepreneurship programs? *Strategic Management Journal*, *39*(1), 85-112. https://doi.org/10.1002/smj.2704

Mallaby, S. (2022). *The power law: Venture capital and the making of the new future*. Penguin.

Malter, D. (2014). On the causality and cause of returns to organizational status: Evidence from the Grands Crus Classés of the Médoc. *Administrative Science Quarterly*, *59*(2), 271-300.

Marx, M. (2021). Employee non-compete agreements, gender, and entrepreneurship. *Organization Science*, *33*(5), 1756-1772. https://doi.org/10.1177/0001839214532428

McDonald, R. M., & Eisenhardt, K. M. (2019). Parallel play: Startups, nascent markets, and effective business-model design. *Administrative Science Quarterly*, *65*(2), 483-523. https://doi.org/10.1177/0001839219852349

McDonald, R. M., Burke, A., Franking, E., & Tempest, N. (2017). Floodgate: On the Hunt for Thunder Lizards. *Harvard Business School Case Study,* 617044-PDF-ENG.

Merton, R. K. (1968). The Matthew Effect in Science: The reward and communication systems of science are considered. *Science*, *159*(3810), 56-63. https://doi.org/10.1126/science.159.3810.56

Milanov, H., & Shepherd, D. A. (2013). The importance of the first relationship: The ongoing influence of initial network on future status. *Strategic Management Journal*, *34*(6), 727-750. https://doi.org/10.1002/smj.2109

Miller, A., O’Mahony, S., & Cohen, S. L. (2022). Opening the aperture: Explaining the complementary roles of advice and testing when forming entrepreneurial strategy. *forthcoming, Organization Science.*

Mindruta, D., Moeen, M., & Agarwal, R. (2016). A two-sided matching approach for partner selection and assessing complementarities in partners’ attributes in inter-firm alliances. *Strategic Management Journal*, *37*(1), 206-231. https://doi.org/10.1002/smj.2448

O’Mahony, S., & Cohen, S. L. (2022). Navigating the promises and perils of researching emerging phenomena in strategy and organizations. *Strategic Organization, 20*(4), 872-885. https://doi.org/10.1177/14761270221122394

Ozcan, P. (2018). Growing with the market: How changing conditions during market growth affect formation and evolution of interfirm ties. *Strategic Management Journal*, *39*(2), 295-328. https://doi.org/10.1002/smj.2740

Ozmel, U. & Guler, I. (2015). Small fish, big fish: The performance effects of the relative standing in partners' affiliate portfolios. *Strategic Management Journal*, *36*(13), 2039-2057. https://doi.org/10.1002/smj.2332

Ozmel, U., Reuer, J. J., & Gulati, R. (2013). Signals across multiple networks: How venture capital and alliance networks affect interorganizational collaboration. *Academy of Management Journal*, *56*(3), 852-866. https://doi.org/10.5465/amj.2009.0549

Pahnke, E. C., Katila, R., & Eisenhardt, K. M. (2015). Who takes you to the dance? How partners' institutional logics influence innovation in young firms. *Administrative Science Quarterly*, *60*(4), 596-633. https://doi.org/10.1177/0001839215592913

Park, S. H., Hallen, B., & McDonald, R. M. (2022) Disruption of the elite: What drove shifts in the venture capital status hierarchy from 2005 to 2019? *Working Paper*.

Phillips, D. J., & Zuckerman, E. W. (2001). Middle-status conformity: Theoretical restatement and empirical demonstration in two markets. *American Journal of Sociology*, *107*(2), 379-429. https://doi.org/10.1086/324072

Piezunka, H., & Dahlander, L. (2015). Distant search, narrow attention: How crowding alters organizations' filtering of suggestions in crowdsourcing. *Academy of Management Journal*, *58*(3), 856-880. https://doi.org/10.5465/amj.2012.0458

Podolny, J. M. (1993). A status-based model of market competition. *American Journal of Sociology, 98*(4), 829-872. https://doi.org/10.1086/230091

Podolny, J. M. (1994). Market uncertainty and the social character of economic exchange. *Administrative Science Quarterly, 39*(3), 458-483. https://doi.org/10.2307/2393299

Podolny, J. M. (2001). Networks as the pipes and prisms of the market. *American Journal of Sociology, 107*(1), 33–60. http://doi.org/10.1086/323038

Podolny, J. M., & Phillips, D. J. (1996). The dynamics of organizational status. *Industrial and Corporate Change*, *5*(2), 453-471. https://doi.org/10.1093/icc/5.2.453

Pollock, T. G., Lee, P. M., Jin, K., & Lashley, K. (2015). (Un) tangled: Exploring the asymmetric coevolution of new venture capital firms' reputation and status. *Administrative Science Quarterly*, *60*(3), 482-517. https://doi.org/10.1177/0001839215585602

Pollock, T.G., Lashley, K., Rindova, V.P., & Han, J.H. (2019). Which of these things are not like the others? Comparing the rational, emotional, and moral aspects of reputation, status, celebrity, and stigma. *Academy of Management Annals*, *13*(2), 444-478. https://doi.org/10.5465/annals.2017.0086

Retterath, A., & Braun, R. (2020). Benchmarking venture capital databases. *Available at SSRN 3706108*. https://doi.org/10.2139/ssrn.3706108

Rider, C. I. (2012). How employees' prior affiliations constrain organizational network change: A study of US venture capital and private equity. *Administrative Science Quarterly*, *57*(3), 453-483. https://doi.org/10.1177/0001839212461633

Roberts, P. W., & Lall, S. A. (2019). *Observing acceleration: Uncovering the effects of accelerators on impact-oriented entrepreneurs*. Springer. https://doi.org/10.1007/978-3-030-00042-4

Rothaermel, F. T., & Deeds, D. L. (2004). Exploration and exploitation alliances in biotechnology: A system of new product development. *Strategic Management Journal*, *25*(3), 201-221. https://doi.org/10.1002/smj.376

Sauder, M. (2008). Interlopers and field change: The entry of US news into the field of legal education. *Administrative Science Quarterly*, *53*(2), 209-234. https://doi.org/10.2189/asqu.53.2.209

Sauder, M., Lynn, F., & Podolny, J. M. (2012). Status: Insights from organizational sociology. *Annual Review of Sociology*, *38*, 267-283. https://doi.org/10.1146/annurev-soc-071811-145503

Shankar, R. K., & Shepherd, D. A. (2019). Accelerating strategic fit or venture emergence: Different paths adopted by corporate accelerators. *Journal of Business Venturing*, *34*(5), 105886. https://doi.org/10.1016/j.jbusvent.2018.06.004

Shipilov, A. V. (2005). Should you bank on your network? Relational and positional embeddedness in the making of financial capital. *Strategic Organization*, *3*(3), 279-309. https://doi.org/10.1177/1476127005055793

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*(4), 422-445. https://doi.org/10.1037/1082-989X.7.4.422

Simcoe, T. S., & Waguespack, D. M. (2011). Status, quality, and attention: What's in a (missing) name? *Management Science*, *57*(2), 274-290. https://doi.org/10.1287/mnsc.1100.1270

Sorenson, O., and Stuart, T. E. (2001). Syndication networks and the spatial distribution of venture capital investments. *American Journal of Sociology*, *106*(6), 1546-1588. https://doi.org/10.1086/321301

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, *87*(3), 355-374. https://doi.org/10.1016/B978-0-12-214850-7.50025-5

Stam, W., Arzlanian, S., & Elfring, T. (2014). Social capital of entrepreneurs and small firm performance: A meta-analysis of contextual and methodological moderators. *Journal of Business Venturing*, *29*(1), 152-173. https://doi.org/10.1016/j.jbusvent.2013.01.002

Stinchcombe, A. L. (1965). Organizations and social structure. *Handbook of Organizations*, *44*(2), 142-193.

Stuart, T. E., Hoang, H., & Hybels, R. C. (1999). Interorganizational endorsements and the performance of entrepreneurial ventures. *Administrative Science Quarterly*, *44*(2), 315-349. https://doi.org/10.2307/2666998

Ter Wal, A. L., Alexy, O., Block, J., & Sandner, P. G. (2016). The best of both worlds: The benefits of open-specialized and closed-diverse syndication networks for new ventures' success. *Administrative Science Quarterly*, *61*(3), 393–432. https://doi.org/10.1177/0001839216637849

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Vissa, B. (2012). Agency in action: Entrepreneurs' networking style and initiation of economic exchange. *Organization Science*, *23*(2), 492-510. https://doi.org/10.1287/orsc.1100.0567

Wang, D., Pahnke, E. C., & McDonald, R. (2022). The past is prologue? Venture-capital syndicates' collaborative experience and start-up exits. *Academy of Management Journal*, *65*(2), 371-402. https://doi.org/10.5465/amj.2019.1312

Wolfolds, S. E., & Siegel, J. (2019). Misaccounting for endogeneity: The peril of relying on the Heckman two-step method without a valid instrument. *Strategic Management Journal*, *40*(3), 432-462. https://doi.org/10.1002/smj.2995

Younger, S., & Fisher, G. (2020). The exemplar enigma: New venture image formation in an emergent organizational category. *Journal of Business Venturing*, *35*(1), 105897. https://doi.org/10.1016/j.jbusvent.2018.09.002

Yu, S. (2020). How do accelerators impact the performance of high-technology ventures? *Management Science*, *66*(2), 530-552. https://doi.org/10.1287/mnsc.2018.3256

Zuckerman, H. (1977). *Scientific elite: Nobel laureates in the United States*. Transaction Publishers.

# TABLES & FIGURES

## TABLE 1. VARIABLE DESCRIPTIVE SUMMARY

| Variable | Mean | S.D. | Min | Max |
|---|---|---|---|---|
| Web Crawl (Count Logged) | 1.72 | 1.19 | 0.00 | 9.30 |
| Raised To Date (USD Logged) | 7.66 | 7.76 | 0.00 | 23.77 |
| Other Technology Stack (Count) | 24.28 | 23.80 | 0.00 | 165.00 |
| Cumulative Max Investor Status (Eigenvector Centrality) | 0.09 | 0.20 | 0.00 | 1.00 |
| Series A Round (Binary) | 0.03 | 0.17 | 0.00 | 1.00 |
| Series B Round (Binary) | 0.02 | 0.13 | 0.00 | 1.00 |
| Series C Round (Binary) | 0.01 | 0.09 | 0.00 | 1.00 |
| Series D Round (Binary) | 0.00 | 0.06 | 0.00 | 1.00 |
| Series E Round (Binary) | 0.00 | 0.04 | 0.00 | 1.00 |
| Series F Round (Binary) | 0.00 | 0.02 | 0.00 | 1.00 |
| Series G Round (Binary) | 0.00 | 0.01 | 0.00 | 1.00 |
| Series H Round (Binary) | 0.00 | 0.00 | 0.00 | 1.00 |
| Series I Round (Binary) | 0.00 | 0.00 | 0.00 | 1.00 |
| Series Unknown (Binary) | 0.02 | 0.14 | 0.00 | 1.00 |
| Angel Round (Binary) | 0.01 | 0.08 | 0.00 | 1.00 |
| Convertible Note Round (Binary) | 0.01 | 0.08 | 0.00 | 1.00 |
| Corporate Round (Binary) | 0.00 | 0.02 | 0.00 | 1.00 |

| | | | | |
|---|---|---|---|---|
| Private Equity Round (Binary) | 0.00 | 0.04 | 0.00 | 1.00 |
| Seed Round (Binary) | 0.05 | 0.22 | 0.00 | 1.00 |
| Series Undisclosed (Binary) | 0.00 | 0.03 | 0.00 | 1.00 |
| State CA (Binary) | 0.45 | 0.50 | 0.00 | 1.00 |
| State MA (Binary) | 0.06 | 0.24 | 0.00 | 1.00 |
| State NY (Binary) | 0.13 | 0.34 | 0.00 | 1.00 |
| State TX (Binary) | 0.05 | 0.21 | 0.00 | 1.00 |
| State WA (Binary) | 0.04 | 0.18 | 0.00 | 1.00 |
| Other State (Binary) | 0.27 | 0.44 | 0.00 | 1.00 |
| Apps (Binary) | 0.14 | 0.35 | 0.00 | 1.00 |
| Artificial intelligence (Binary) | 0.12 | 0.33 | 0.00 | 1.00 |
| Commerce and shopping (Binary) | 0.22 | 0.41 | 0.00 | 1.00 |
| Data and analytics (Binary) | 0.19 | 0.39 | 0.00 | 1.00 |
| Gaming (Binary) | 0.02 | 0.14 | 0.00 | 1.00 |
| Information technology (Binary) | 0.14 | 0.35 | 0.00 | 1.00 |
| Internet services (Binary) | 0.09 | 0.28 | 0.00 | 1.00 |
| Mobile (Binary) | 0.02 | 0.15 | 0.00 | 1.00 |
| Platforms (Binary) | 0.00 | 0.01 | 0.00 | 1.00 |
| Software (Binary) | 0.06 | 0.23 | 0.00 | 1.00 |
| Data Analytics Utilization (Binary) | 0.79 | 0.41 | 0.00 | 1.00 |
| Cumulative Quarters of Data Analytics Utilization (Count) | 6.87 | 7.16 | 0.00 | 43.00 |
| Cumulative Web Crawl (Count Logged) | 14.78 | 17.87 | 0.00 | 202.11 |
| Market Space Atypicality | -0.17 | 0.10 | -0.75 | -0.01 |
| Additionally uses A/B testing (Binary) | 0.14 | 0.35 | 0.00 | 1.00 |
| Prior Employer Prominence (Logged) | 0.93 | 1.68 | 0.00 | 5.54 |

# TABLE 2. CORRELATION MATRIX

| | Web Crawl | Raised To Date USD | Other Technology Stack | Cumulative Max Investor Status | Data Analytics Utilization | Cumulative Quarters of Data Analytics Utilization | Cumulative Web Crawl |
|---|---|---|---|---|---|---|---|
| Web Crawl | 1.00 | | | | | | |
| Raised To Date USD | 0.35 | 1.00 | | | | | |
| Other Technology Stack | 0.40 | 0.39 | 1.00 | | | | |
| Cumulative Max Investor Status | 0.30 | 0.46 | 0.18 | 1.00 | | | |
| Data Analytics Utilization | 0.32 | 0.25 | 0.43 | 0.13 | 1.00 | | |
| Cumulative Quarters of Data Analytics Utilization | 0.39 | 0.36 | 0.61 | 0.19 | 0.49 | 1.00 | |
| Cumulative Web Crawl | 0.58 | 0.40 | 0.57 | 0.30 | 0.32 | 0.80 | 1.00 |

*Note.* excludes moderators

## TABLE 3. TIME-SERIES CROSS SECTIONAL MATCHING MODEL USING MAHALANOBIS DISTANCE-BASED MATCHING REFINEMENT

Estimate of Average Treatment Effect on the Treated (ATT)
*(1000 bootstrapped iterations at 95% CI)*

| | estimates | std.error | estimate at 2.5% | estimate at 97.5% |
|---|---|---|---|---|
| *t* | 0.357 | 0.021 | 0.318 | 0.396 |
| *t+1* | 0.302 | 0.025 | 0.251 | 0.350 |
| *t+2* | 0.245 | 0.029 | 0.189 | 0.295 |

## TABLE 4. TIME-SERIES CROSS SECTIONAL MATCHING MODEL USING PROPENSITY SCORE-BASED MATCHING REFINEMENT

Estimate of Average Treatment Effect on the Treated (ATT)
*(1000 bootstrapped iterations at 95% CI)*

| | estimates | std.error | estimate at 2.5% | estimate at 97.5% |
|---|---|---|---|---|
| *t* | 0.400 | 0.020 | 0.360 | 0.440 |
| *t+1* | 0.370 | 0.025 | 0.323 | 0.418 |
| *t+2* | 0.312 | 0.027 | 0.258 | 0.364 |

## TABLE 5. TIME-SERIES CROSS SECTIONAL MATCHING MODEL USING PROPENSITY SCORE-BASED WEIGHTING REFINEMENT

Estimate of Average Treatment Effect on the Treated (ATT)
*(1000 bootstrapped iterations at 95% CI)*

| | estimates | std.error | estimate at 2.5% | estimate at 97.5% |
|---|---|---|---|---|
| *t* | 0.418 | 0.019 | 0.378 | 0.454 |
| *t+1* | 0.392 | 0.022 | 0.349 | 0.437 |
| *t+2* | 0.326 | 0.025 | 0.281 | 0.377 |

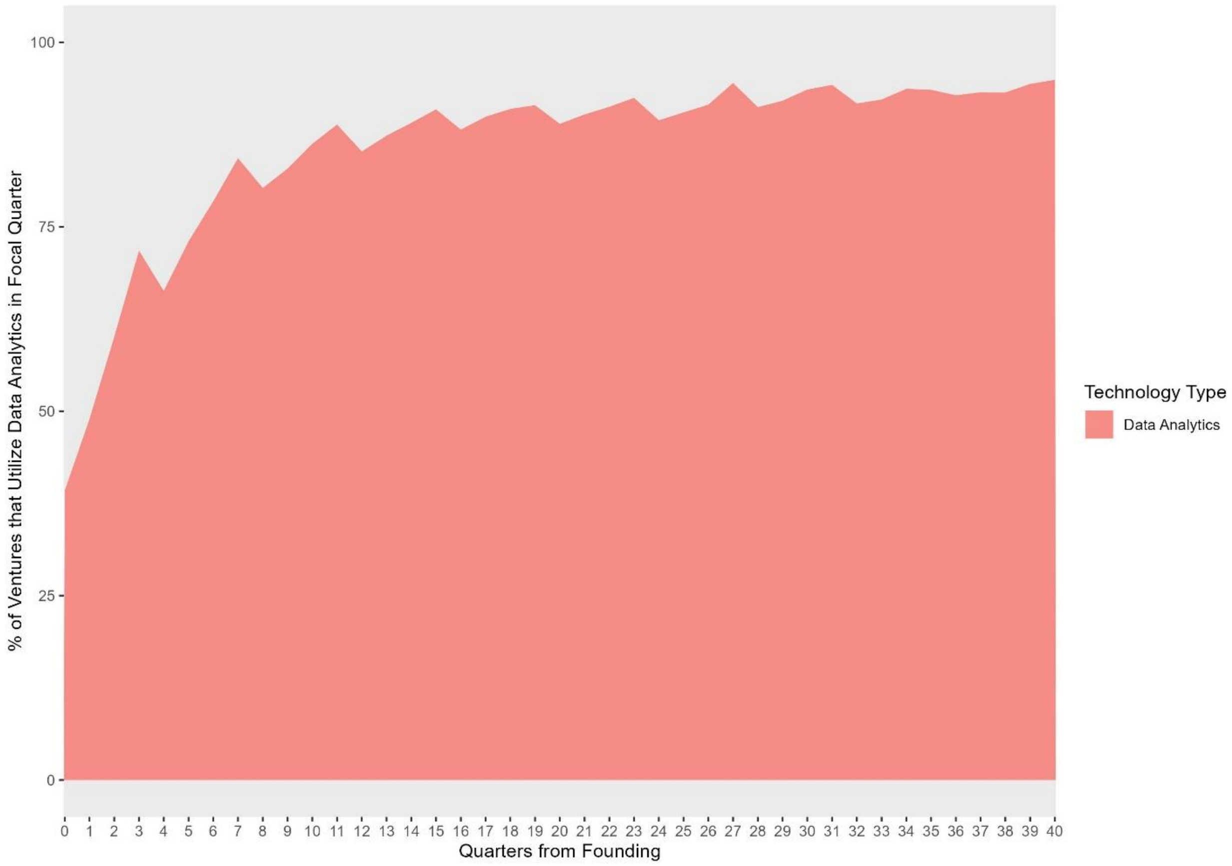
# TABLE 6. REGRESSION WITH MODERATORS

| | (1) RE - Market Space Atypicality | | (2) FE - A/B Testing | | (3) RE - Prior Employer Prominence | |
|---|---|---|---|---|---|---|
| (Intercept) | 0.581 | [0.000] | | | 0.601 | [0.000] |
| | (0.04) | | | | (0.038) | |
| Raised To Date (USD Logged) | 0.013 | [0.000] | 0.009 | [0.000] | 0.015 | [0.000] |
| | (0.001) | | (0.001) | | (0.001) | |
| Other Technology Stack (Count) | 0.006 | [0.000] | | | 0.006 | [0.000] |
| | (0.000) | | | | (0.001) | |
| Investor Status (Cumulative Max Centrality) | 0.510 | [0.000] | 0.484 | [0.000] | 0.488 | [0.000] |
| | (0.048) | | (0.082) | | (0.056) | |
| Cumulative Web Crawl (Count Logged) | 0.024 | [0.000] | 0.000 | [0.947] | 0.025 | [0.000] |
| | (0.001) | | (0.002) | | (0.001) | |
| Cumulative Quarters of Data Analytics Utilization (Count) | -0.022 | [0.000] | -0.017 | [0.363] | -0.021 | [0.000] |
| | (0.002) | | (0.018) | | (0.003) | |
| Data Analytics Utilization (Binary) | 0.480 | [0.000] | | | 0.418 | [0.000] |
| | (0.025) | | | | (0.021) | |
| Market Space Atypicality | -0.433 | [0.000] | | | | |
| | (0.115) | | | | | |
| Data Analytics Utilization (Binary) x Market Space Atypicality | 0.261 | [0.041] | | | | |
| | (0.128) | | | | | |
| Other Technology Stack (excl. A/B testing) (Count) | | | 0.008 | [0.000] | | |
| | | | (0.001) | | | |
| Additionally uses A/B testing (Binary) | | | 0.084 | [0.000] | | |
| | | | (0.024) | | | |
| Prior Employer Prominence (Logged) | | | | | 0.003 | [0.713] |
| | | | | | (0.008) | |
| Data Analytics Utilization (Binary) x Prior Employer Prominence (Logged) | | | | | 0.014 | [0.079] |
| | | | | | (0.008) | |
| S.E.: Clustered | by: Firm | | by: Firm | | by: Firm | |
| Firm fixed effect | No | | Yes | | No | |
| Year-Quarter fixed effect | No | | Yes | | No | |
| Round Controls | Yes | | Yes | | Yes | |
| Industry Controls | Yes | | No | | Yes | |
| Region Controls | Yes | | No | | Yes | |
| Observations | 77,535 | | 64,389 | | 48,469 | |

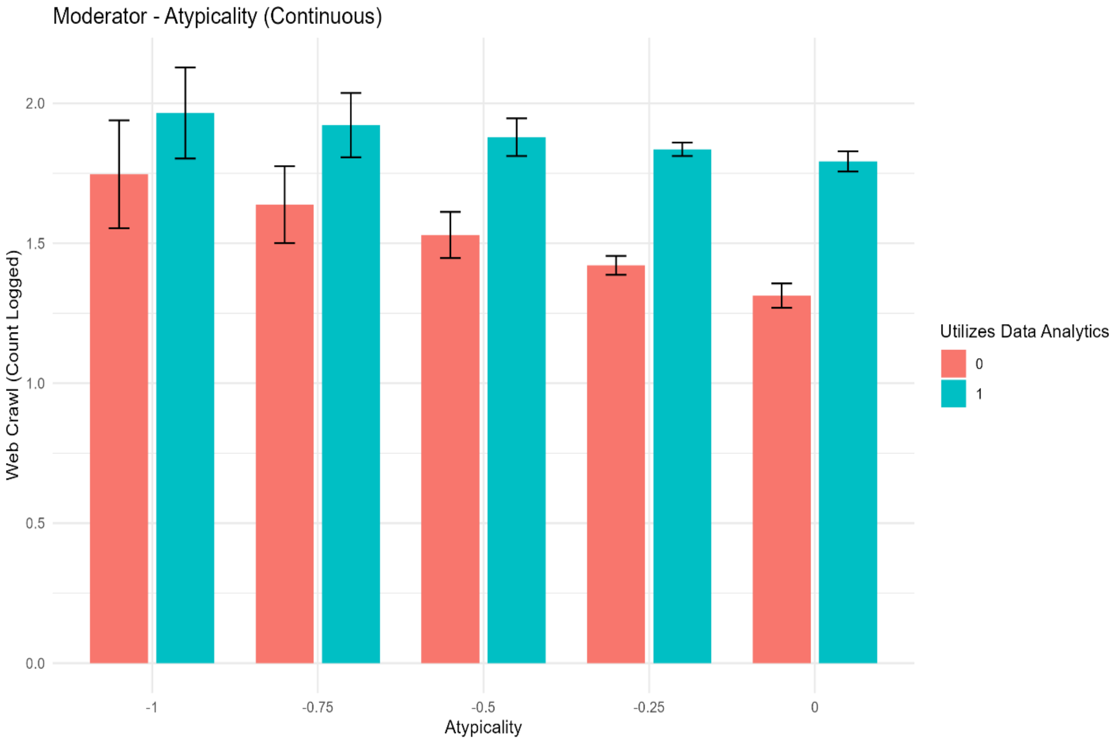
| | (1) RE - Market Space Atypicality | (2) FE - A/B Testing | (3) RE - Prior Employer Prominence |
|---|---|---|---|
| No. of firms | 6,435 | 6,255 | 4,129 |
| R2 | 0.2552 | 0.6838 | 0.2662 |
| adjusted R2 | 0.2548 | 0.6493 | 0.2656 |

cluster-robust standard errors in parentheses; p-value in square brackets

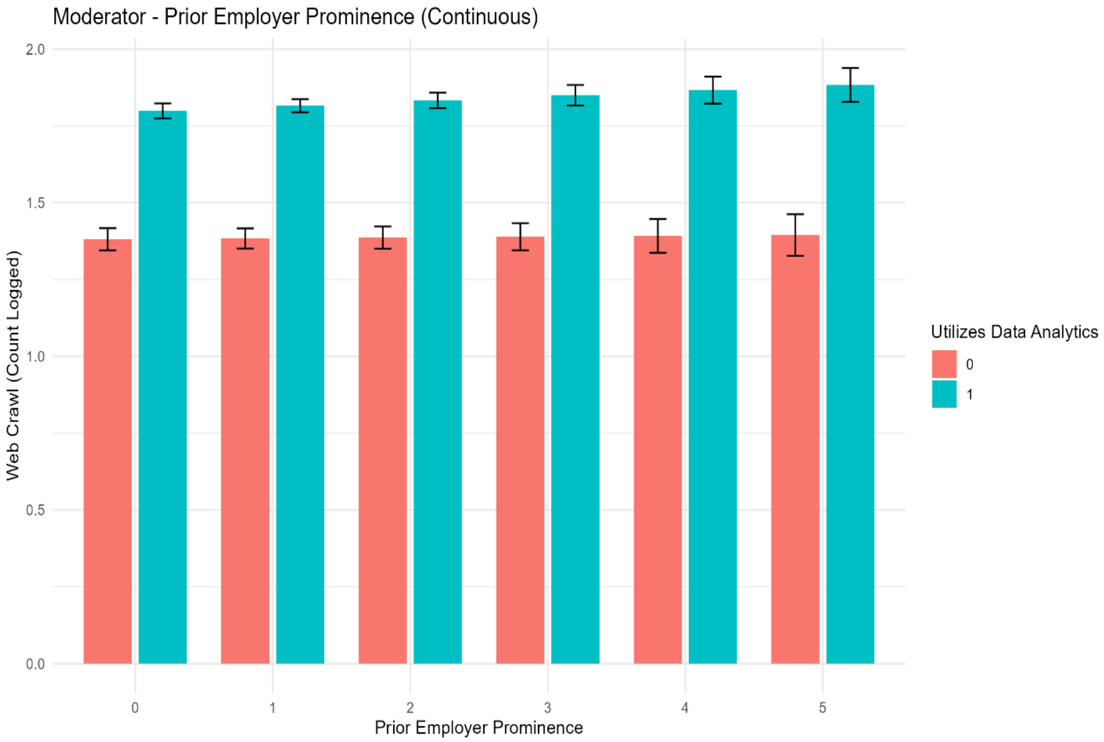
## FIGURE 1. DATA ANALYTICS UTILIZATION TRENDS



## FIGURE 2. INTERACTION EFFECT OF DATA ANALYTICS & MARKET SPACE ATYPICALITY ON WEB CRAWL COUNT



## FIGURE 3. INTERACTION EFFECT OF DATA ANALYTICS & PRIOR EMPLOYER PROMINENCE ON WEB CRAWL COUNT



# TABLE 7: DESCRIPTIVE STATISTICS OF THREE SAMPLES

| | Initial Stage (N = 1,737) | | Follow-on Early Rounds (N = 1,170) | | Later Stage (N = 602) | |
|---|---|---|---|---|---|---|
| | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| Dependent Variable | | | | | | |
| Pre-money valuation (log) | 1.83 | 0.63 | 2.93 | 0.88 | 4.14 | 1.41 |
| Region | | | | | | |
| MSA dummies | 45 Categories | | 32 Categories | | 27 Categories | |
| Industry | | | | | | |
| Industry dummies | 31 Categories | | 27 Categories | | 24 Categories | |
| Year | | | | | | |
| Year dummies | 11 Categories | | 13 Categories | | 14 Categories | |
| Founder Attributes | | | | | | |
| Single founder | 0.36 | 0.48 | 0.35 | 0.48 | 0.44 | 0.50 |
| Founding team of 2 or 3 | 0.59 | 0.49 | 0.59 | 0.49 | 0.52 | 0.50 |
| Founding team of 4 or more | 0.05 | 0.21 | 0.06 | 0.25 | 0.05 | 0.21 |
| Founding team has Female | 0.16 | 0.36 | 0.15 | 0.36 | 0.12 | 0.32 |
| Founding team gender is unknown | 0.02 | 0.13 | 0.01 | 0.12 | 0.01 | 0.11 |
| Founder education Linkedin missing | 0.10 | 0.30 | 0.11 | 0.31 | 0.14 | 0.35 |
| Founder job experience Linkedin missing | 0.02 | 0.14 | 0.01 | 0.11 | 0.03 | 0.16 |
| Founders work experience (log Years) | 2.34 | 0.74 | 2.33 | 0.73 | 2.25 | 0.79 |
| U.S. News 2012 Score | 64.14 | 29.28 | 65.24 | 29.32 | 62.06 | 29.51 |
| MBA | 0.19 | 0.39 | 0.21 | 0.41 | 0.17 | 0.38 |
| JD | 0.03 | 0.17 | 0.02 | 0.14 | 0.03 | 0.17 |
| Technical Masters | 0.20 | 0.40 | 0.21 | 0.40 | 0.19 | 0.39 |
| Other Masters | 0.05 | 0.21 | 0.04 | 0.19 | 0.02 | 0.16 |
| Technical Doctorate | 0.09 | 0.29 | 0.08 | 0.28 | 0.09 | 0.28 |
| Other Doctorate | 0.04 | 0.20 | 0.04 | 0.20 | 0.03 | 0.17 |
| Prior employer prominence (log) | 2.50 | 1.78 | 2.59 | 1.76 | 2.22 | 1.72 |
| Previously founded another venture (Binary) | 0.40 | 0.49 | 0.40 | 0.49 | 0.36 | 0.48 |
| Previously founded another venture (Count) | 0.66 | 1.03 | 0.65 | 1.02 | 0.59 | 0.98 |
| Previously founded VC-backed venture | 0.02 | 0.15 | 0.02 | 0.14 | 0.02 | 0.13 |
| Venture Progress | | | | | | |
| Venture age | 2.53 | 1.22 | 3.75 | 1.22 | 6.23 | 1.85 |
| Business status - Product Development Phase | 0.03 | 0.16 | 0.03 | 0.16 | 0.00 | 0.07 |
| Business status - Revenue Phase | 0.41 | 0.49 | 0.57 | 0.50 | 0.83 | 0.38 |
| Business status - Profitable Phase | 0.01 | 0.09 | 0.02 | 0.12 | 0.06 | 0.24 |
| Previously participated in accelerator | 0.01 | 0.12 | 0.02 | 0.15 | 0.03 | 0.17 |
| Internet Archive Crawls (Count logged) | 0.96 | 1.11 | 1.73 | 1.30 | 2.54 | 1.54 |

|  | Initial Stage (N = 1,737) |  | Follow-on Early Rounds (N = 1,170) |  | Later Stage (N = 602) |  |
|---|---|---|---|---|---|---|
|  | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| Prior Resource Mobilization |  |  |  |  |  |  |
| Raised previously cumulative (log $M) | 0.13 | 0.28 | 1.50 | 0.73 | 2.87 | 1.14 |
| Previous maximum eigenvector centrality | 0.03 | 0.10 | 0.23 | 0.26 | 0.30 | 0.27 |
| Previous maximum Pre-money valuation (log) | 1.28 | 0.74 | 2.12 | 0.74 | 3.59 | 1.23 |

# TABLE 8. SUMMARY OF VARIANCE EXPLAINED BY ROUND

| | Initial Round | | Early Follow-On Rounds | | Later Rounds | |
|---|---|---|---|---|---|---|
| factor | R squared (%)[a] | Adj. R squared (%)[a] | R squared (%)[a] | Adj. R squared (%)[a] | R squared (%)[a] | Adj. R squared (%)[a] |
| *Region* | 12.1 | 10.21 | 6.23 | 4.62 | 8.22 | 6.4 |
| | [11.83,12.36] | [9.93,10.49] | [6.06,6.4] | [4.44,4.79] | [7.95,8.49] | [6.11,6.69] |
| *Industry* | 3.22 | 1.84 | 4.75 | 3.41 | 7.07 | 5.53 |
| | [3.09,3.35] | [1.7,1.98] | [4.5,4.99] | [3.16,3.66] | [6.86,7.29] | [5.3,5.76] |
| *Year* | 3.65 | 3.24 | 4.02 | 3.46 | 5.3 | 4.47 |
| | [3.48,3.83] | [3.06,3.42] | [3.86,4.18] | [3.29,3.63] | [5.08,5.52] | [4.24,4.7] |
| *Founding team attributes* | 5.45 | 4.67 | 4.7 | 3.78 | 7.03 | 5.8 |
| | [5.27,5.63] | [4.48,4.85] | [4.55,4.85] | [3.62,3.93] | [6.85,7.22] | [5.6,6] |
| *Venture progress* | 3.6 | 3.39 | 4.47 | 4.26 | 11.58 | 11.55 |
| | [3.43,3.76] | [3.22,3.57] | [4.28,4.66] | [4.06,4.46] | [11.21,11.94] | [11.17,11.94] |
| *Past resource mobilization* | | | 37.36 | 38.62 | 43.45 | 45.95 |
| | | | [36.87,37.86] | [38.11,39.14] | [42.95,43.96] | [45.41,46.5] |
| *Total variance explained* | 28.01 | 23.35 | 61.53 | 58.15 | 82.66 | 79.7 |
| | [27.62,28.41] | [22.93,23.77] | [61.08,61.98] | [57.66,58.64] | [82.12,83.2] | [79.07,80.34] |
| | *N = 1,737* | | *N = 1,170* | | *N = 602* | |

[a] Mean variance explained; Bootstrapped 95% Confidence Interval in square brackets (100 iterations)

## TABLE 9. SUMMARY AND PAIRWISE CORRELATION STATISTICS FOR SAMPLE I STARTUPS THAT RAISED INVESTMENTS

| | | Mean | S.D. | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) | Investor Status | 0.952 | 1.026 | | | | | | | | | | | |
| (2) | Focal Accelerator Participant | 0.376 | 0.487 | 0.14 | | | | | | | | | | |
| (3) | Other Accelerator Participant | 0.294 | 0.458 | 0.10 | -0.50 | | | | | | | | | |
| (4) | Cohort B-2012 Applicant | 0.329 | 0.473 | -0.06 | 0.02 | -0.07 | | | | | | | | |
| (5) | Cohort X-2012 Applicant | 0.329 | 0.473 | 0.19 | 0.02 | -0.01 | -0.49 | | | | | | | |
| (6) | Cohort E-2012 Applicant | 0.176 | 0.383 | -0.10 | 0.02 | -0.03 | -0.32 | -0.32 | | | | | | |
| (7) | Prior Web Traffic | 1.806 | 2.058 | 0.14 | 0.05 | 0.02 | -0.11 | 0.02 | -0.08 | | | | | |
| (8) | 2 or 3 Founders | 0.694 | 0.464 | 0.24 | 0.15 | -0.08 | 0.03 | -0.02 | 0.04 | 0.14 | | | | |
| (9) | Has Female Founders | 0.247 | 0.434 | -0.03 | -0.11 | -0.07 | 0.00 | -0.05 | 0.09 | -0.11 | 0.03 | | | |
| (10) | Fulltime at Application | 0.652 | 0.380 | 0.09 | 0.04 | -0.08 | 0.13 | 0.05 | -0.12 | 0.06 | 0.15 | 0.11 | | |
| (11) | Years Work Experience (L) | 2.118 | 0.692 | -0.07 | 0.25 | -0.03 | -0.03 | -0.18 | 0.10 | -0.13 | 0.12 | 0.03 | 0.12 | |
| (12) | Years Work Exp. Missing | 0.012 | 0.108 | -0.09 | -0.08 | -0.07 | 0.16 | -0.08 | -0.05 | 0.14 | -0.16 | -0.06 | -0.19 | -0.34 |
| (13) | Prior Employer Prominence | 0.204 | 0.673 | -0.09 | 0.14 | -0.09 | 0.08 | -0.16 | 0.05 | -0.10 | 0.20 | -0.07 | 0.05 | 0.02 |
| (14) | University Prominence | 51.500 | 23.444 | 0.03 | -0.06 | 0.02 | 0.01 | -0.10 | 0.02 | 0.07 | 0.13 | 0.14 | -0.02 | -0.14 |
| (15) | MBA | 0.376 | 0.487 | -0.20 | -0.10 | -0.02 | 0.08 | -0.23 | 0.09 | -0.30 | -0.17 | 0.12 | -0.03 | 0.25 |
| (16) | JD | 0.047 | 0.213 | -0.15 | 0.06 | -0.14 | -0.04 | -0.04 | 0.04 | -0.13 | 0.15 | -0.13 | -0.11 | 0.07 |
| (17) | Technical Masters | 0.224 | 0.419 | 0.03 | 0.05 | 0.03 | 0.04 | -0.02 | -0.10 | 0.01 | 0.17 | -0.18 | -0.18 | 0.03 |
| (18) | Technical PhD | 0.071 | 0.258 | 0.16 | 0.17 | -0.08 | -0.10 | 0.20 | -0.01 | -0.05 | 0.08 | 0.16 | 0.01 | 0.03 |
| (19) | Serial Entrepreneurs | 0.706 | 0.458 | 0.28 | 0.13 | 0.08 | -0.10 | 0.18 | -0.04 | 0.25 | 0.24 | 0.07 | 0.16 | 0.02 |
| (20) | Previously Raised VC | 0.082 | 0.277 | 0.00 | 0.30 | -0.19 | 0.06 | 0.06 | -0.03 | -0.06 | -0.08 | 0.03 | 0.04 | 0.07 |

| | | (12) | (13) | (14) | (15) | (16) | (17) | (18) | (19) |
|---|---|---|---|---|---|---|---|---|---|
| (13) | Prior Employer Prominence | -0.03 | | | | | | | |
| (14) | University Prominence | 0.09 | 0.22 | | | | | | |
| (15) | MBA | -0.08 | 0.15 | 0.01 | | | | | |
| (16) | JD | -0.02 | 0.16 | 0.04 | -0.06 | | | | |
| (17) | Technical Masters | -0.06 | 0.36 | 0.08 | 0.05 | 0.01 | | | |
| (18) | Technical PhD | -0.03 | 0.20 | -0.13 | -0.02 | -0.06 | 0.18 | | |
| (18) | Serial Entrepreneurs | -0.17 | 0.00 | 0.16 | -0.14 | 0.02 | 0.04 | 0.18 | |
| (19) | Previously Raised VC | -0.03 | 0.09 | 0.01 | -0.06 | -0.07 | 0.15 | 0.25 | 0.19 |

N = 85 for all variables. Restricted to startups in Sample I that raised equity investments from disclosed investors. See Appendix G for descriptive statistics for original sample of 235 startups accepted and almost accepted to these accelerators.

## TABLE 10. ACCELERATOR IMPACT ON STATUS OF EARLY INVESTORS IN SAMPLE I

| | (1) | | (2) | | (3) | |
|---|---|---|---|---|---|---|
| Accelerator Participation | | | | | | |
| Focal Accelerator Participant | 0.472 | [0.064] | 0.546 | [0.013] | 0.446 | [0.025] |
| | (0.214) | | (0.165) | | (0.151) | |
| Other Accelerator Participant | 0.370 | [0.042] | 0.443 | [0.041] | | |
| | (0.149) | | (0.178) | | | |
| Control Variables | | | | | | |
| Cohort B-2012 Applicant | -0.010 | [0.968] | 0.311 | [0.340] | 0.114 | [0.811] |
| | (0.238) | | (0.303) | | (0.456) | |
| Cohort X-2012 Applicant | 0.130 | [0.569] | 0.209 | [0.533] | 0.103 | [0.819] |
| | (0.217) | | (0.319) | | (0.430) | |
| Cohort E-2012 Applicant | -0.131 | [0.445] | -0.045 | [0.863] | 0.068 | [0.871] |
| | (0.161) | | (0.252) | | (0.404) | |
| Prior Web Traffic | -0.009 | [0.889] | -0.003 | [0.970] | -0.085 | [0.130] |
| | (0.065) | | (0.077) | | (0.048) | |
| 2 or 3 Founders | 0.432 | [0.203] | 0.160 | [0.617] | -0.086 | [0.158] |
| | (0.308) | | (0.306) | | (0.053) | |
| Has Female Founders | -0.119 | [0.585] | 0.063 | [0.798] | 0.077 | [0.859] |
| | (0.207) | | (0.239) | | (0.415) | |
| Fulltime at Application | 0.102 | [0.844] | -0.060 | [0.926] | 0.632 | [0.462] |
| | (0.498) | | (0.626) | | (0.804) | |
| Years Work Experience (L) | -0.166 | [0.515] | -0.300 | [0.392] | -0.480 | [0.021] |
| | (0.242) | | (0.330) | | (0.154) | |
| Years Work Exp. Missing | -0.452 | [0.431] | -0.918 | [0.339] | 0.000 | [.] |
| | (0.541) | | (0.894) | | (.) | |
| Prior Employer Prominence | -0.209 | [0.174] | -0.072 | [0.731] | -0.132 | [0.574] |
| | (0.138) | | (0.201) | | (0.223) | |
| University Prominence | 0.002 | [0.681] | 0.003 | [0.573] | 0.002 | [0.822] |
| | (0.005) | | (0.006) | | (0.008) | |
| MBA | -0.146 | [0.557] | -0.312 | [0.177] | -0.037 | [0.819] |
| | (0.237) | | (0.207) | | (0.156) | |
| JD | -0.697 | [0.015] | -0.631 | [0.148] | -0.877 | [0.096] |
| | (0.216) | | (0.389) | | (0.444) | |
| Technical Masters | 0.034 | [0.933] | -0.238 | [0.604] | -0.042 | [0.945] |
| | (0.391) | | (0.439) | | (0.584) | |
| Technical PhD | 0.522 | [0.290] | 0.325 | [0.424] | -0.024 | [0.942] |
| | (0.456) | | (0.382) | | (0.313) | |
| Serial Entrepreneurs | 0.355 | [0.396] | 0.900 | [0.127] | 1.133 | [0.013] |
| | (0.392) | | (0.521) | | (0.324) | |
| Previously Raised VC | -0.337 | [0.473] | -0.466 | [0.245] | -0.753 | [0.005] |
| | (0.445) | | (0.367) | | (0.172) | |
| Constant | 0.437 | [0.558] | 0.478 | [0.588] | 0.809 | [0.201] |
| | (0.711) | | (0.842) | | (0.564) | |
| Estimation model | OLS | | IPTW | | PSM | |
| N | 85 | | 85 | | 64 | |
| R squared | 0.242 | | 0.396 | | 0.604 | |
| Adj. R squared | 0.020 | | 0.219 | | 0.458 | |

P-values in brackets. Robust standard errors in parentheses, clustered at level of accelerator cohort / application pool. IPTW = Inverse Probability of Treatment Weights; PSM = Propensity Score Matching, restricted to startups participating in the focal accelerators or with no accelerator participation.

(Note that the reported N for propensity score matching is 64, reflecting some non-accelerator startups being matched multiple times.)

## TABLE 11. SUMMARY AND PAIRWISE CORRELATION STATISTICS FOR SAMPLE II

| | | Mean | S.D. | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) | Investor Status | 0.995 | 1.093 | | | | | | | | | | |
| (2) | Accelerator Participation | 0.151 | 0.358 | 0.14 | | | | | | | | | |
| (3) | 2 or 3 Founders | 0.569 | 0.495 | 0.10 | 0.11 | | | | | | | | |
| (4) | Has Female Founders | 0.144 | 0.351 | 0.02 | 0.04 | 0.03 | | | | | | | |
| (5) | Prior Employer Prominence | 2.324 | 1.759 | 0.18 | 0.05 | 0.19 | 0.00 | | | | | | |
| (6) | Years Work Exper. (L) | 2.327 | 0.747 | -0.08 | -0.13 | 0.04 | 0.00 | 0.33 | | | | | |
| (7) | Serial Entrepreneurs | 0.395 | 0.489 | 0.05 | 0.05 | 0.03 | -0.03 | 0.09 | 0.23 | | | | |
| (8) | Previously Raised VC | 0.021 | 0.143 | 0.03 | -0.01 | -0.04 | -0.04 | 0.01 | 0.06 | 0.14 | | | |
| (9) | Founders Jobs Missing | 0.022 | 0.145 | 0.00 | 0.00 | -0.05 | 0.01 | -0.20 | -0.46 | -0.12 | -0.02 | | |
| (10) | MBA | 0.172 | 0.377 | -0.04 | 0.00 | 0.03 | 0.04 | 0.13 | 0.09 | 0.00 | 0.02 | 0.07 | |
| (11) | Technical Masters | 0.183 | 0.387 | 0.07 | 0.02 | 0.09 | -0.02 | 0.19 | 0.08 | 0.01 | -0.02 | -0.01 | 0.00 |
| (12) | Other Masters | 0.058 | 0.234 | -0.03 | 0.00 | 0.00 | 0.08 | -0.01 | 0.02 | 0.01 | 0.00 | -0.03 | -0.02 |
| (13) | Technical PhD | 0.076 | 0.266 | 0.02 | -0.06 | 0.05 | -0.02 | 0.12 | 0.07 | -0.04 | 0.00 | 0.00 | -0.06 |
| (14) | Other PhD | 0.033 | 0.179 | 0.01 | -0.02 | 0.03 | 0.01 | 0.07 | 0.07 | 0.03 | 0.00 | 0.00 | -0.02 |
| (15) | JD | 0.028 | 0.164 | -0.01 | 0.01 | 0.00 | 0.00 | -0.01 | 0.00 | 0.01 | 0.00 | 0.01 | -0.02 |
| (16) | University Prominence | 58.7 | 33.1 | 0.14 | 0.08 | 0.13 | 0.08 | 0.31 | 0.05 | 0.09 | 0.02 | 0.04 | 0.27 |
| (17) | Founders Education Missing | 0.098 | 0.297 | 0.00 | -0.05 | -0.10 | -0.04 | -0.15 | -0.11 | -0.07 | -0.02 | -0.05 | -0.15 |
| (18) | Internet Archive Crawls (-436 days; L) | 0.476 | 0.825 | 0.00 | 0.00 | 0.00 | 0.03 | -0.04 | 0.00 | -0.01 | -0.02 | 0.02 | -0.01 |
| (19) | Startup Age | 2.281 | 1.182 | -0.10 | -0.02 | -0.04 | -0.01 | -0.11 | 0.07 | -0.04 | -0.04 | 0.00 | -0.01 |
| (20) | Current Round Amount (M$; L) | 1.187 | 0.713 | 0.28 | -0.20 | -0.01 | -0.05 | 0.12 | 0.13 | 0.01 | 0.07 | -0.01 | 0.00 |
| (21) | Previously Raised Capital (M$; L) | 0.009 | 0.033 | 0.00 | 0.30 | 0.04 | 0.01 | -0.03 | -0.06 | 0.01 | -0.03 | 0.03 | 0.00 |
| (22) | Convertible Note | 0.020 | 0.141 | -0.01 | 0.12 | 0.04 | 0.05 | 0.00 | 0.01 | 0.03 | 0.01 | -0.01 | 0.01 |
| (23) | Percentage VCs in Round | 0.644 | 0.391 | 0.20 | -0.07 | 0.04 | -0.03 | 0.05 | 0.04 | -0.01 | 0.02 | -0.01 | -0.01 |
| (24) | State 3-Year VC Rounds (L) | 5.652 | 1.706 | 0.28 | 0.10 | 0.09 | 0.06 | 0.16 | -0.03 | 0.07 | 0.03 | -0.01 | -0.01 |

| | | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) | (19) | (20) | (21) | (22) | (23) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (12) | Other Masters | -0.02 | | | | | | | | | | | | |
| (13) | Technical PhD | 0.22 | -0.02 | | | | | | | | | | | |
| (14) | Other PhD | 0.09 | 0.06 | 0.06 | | | | | | | | | | |
| (15) | JD | -0.04 | -0.01 | -0.04 | -0.01 | | | | | | | | | |
| (16) | University Prominence | 0.21 | 0.06 | 0.19 | 0.11 | 0.08 | | | | | | | | |
| (17) | Founders Education Missing | -0.16 | -0.08 | -0.09 | -0.06 | -0.06 | -0.58 | | | | | | | |
| (18) | Internet Archive Crawls (-436 days; L) | 0.00 | -0.02 | -0.04 | 0.00 | -0.01 | -0.02 | 0.01 | | | | | | |
| (19) | Startup Age | 0.02 | 0.03 | 0.04 | 0.01 | 0.01 | -0.05 | 0.00 | 0.33 | | | | | |
| (20) | Current Round Amount (M$; L) | 0.09 | -0.02 | 0.14 | 0.06 | 0.00 | 0.07 | 0.03 | 0.11 | 0.19 | | | | |
| (21) | Previously Raised Capital (M$; L) | -0.01 | 0.02 | -0.04 | 0.00 | -0.03 | 0.00 | -0.02 | 0.03 | 0.05 | -0.14 | | | |
| (22) | Convertible Note | 0.02 | 0.03 | -0.03 | 0.00 | 0.01 | 0.01 | -0.01 | -0.01 | -0.02 | -0.12 | 0.05 | | |
| (23) | Percentage VCs in Round | 0.05 | -0.03 | 0.04 | 0.03 | -0.01 | 0.01 | 0.02 | 0.03 | 0.04 | 0.26 | -0.06 | -0.02 | |
| (24) | State 3-Year VC Rounds (L) | 0.08 | 0.02 | 0.04 | 0.02 | -0.04 | 0.12 | -0.02 | -0.01 | -0.1 | 0.09 | 0.01 | 0.03 | 0.04 |

N = 9,966 for round and startup-level measures, 3,702 for founder-level measures.

# TABLE 12. ACCELERATOR IMPACT ON STATUS OF EARLY INVESTORS IN SAMPLE II

| | (4) | | (5) | | (6) | | (7) | |
|---|---|---|---|---|---|---|---|---|
| 500 Startups | 0.569 | [0.000] | 0.401 | [0.000] | 0.386 | [0.000] | 0.373 | [0.000] |
| | (0.009) | | (0.006) | | (0.018) | | (0.027) | |
| Alchemist Accelerator | 0.564 | [0.000] | 0.112 | [0.000] | 0.145 | [0.000] | 0.139 | [0.009] |
| | (0.011) | | (0.023) | | (0.021) | | (0.052) | |
| AngelPad | 0.520 | [0.000] | 0.394 | [0.000] | 0.323 | [0.000] | 0.307 | [0.000] |
| | (0.009) | | (0.009) | | (0.011) | | (0.025) | |
| Boost VC | 0.545 | [0.000] | 0.554 | [0.000] | 0.548 | [0.000] | 0.536 | [0.000] |
| | (0.012) | | (0.016) | | (0.022) | | (0.063) | |
| DreamIt | 0.402 | [0.000] | 0.554 | [0.000] | 0.442 | [0.000] | 0.282 | [0.000] |
| | (0.009) | | (0.010) | | (0.021) | | (0.064) | |
| MassChallenge | 0.175 | [0.000] | 0.195 | [0.000] | 0.073 | [0.010] | 0.067 | [0.305] |
| | (0.008) | | (0.028) | | (0.028) | | (0.065) | |
| Mucker Capital | 0.763 | [0.000] | 0.523 | [0.000] | 0.546 | [0.000] | 0.628 | [0.000] |
| | (0.010) | | (0.022) | | (0.026) | | (0.078) | |
| Plug and Play | 0.242 | [0.000] | 0.179 | [0.000] | 0.140 | [0.000] | 0.167 | [0.005] |
| | (0.009) | | (0.023) | | (0.034) | | (0.058) | |
| StartX | 0.318 | [0.000] | 0.219 | [0.000] | 0.129 | [0.006] | -0.025 | [0.710] |
| | (0.006) | | (0.022) | | (0.046) | | (0.067) | |
| Techstars New York | 0.945 | [0.000] | 0.770 | [0.000] | 0.710 | [0.000] | 0.753 | [0.000] |
| | (0.015) | | (0.025) | | (0.040) | | (0.050) | |
| Techstars Boston | 0.888 | [0.000] | 0.754 | [0.000] | 0.751 | [0.000] | 0.759 | [0.000] |
| | (0.020) | | (0.025) | | (0.026) | | (0.071) | |
| Techstars Boulder | 0.419 | [0.000] | 0.439 | [0.000] | 0.314 | [0.000] | 0.271 | [0.000] |
| | (0.012) | | (0.036) | | (0.043) | | (0.070) | |
| Techstars Seattle | 0.920 | [0.000] | 0.598 | [0.000] | 0.452 | [0.000] | 0.787 | [0.000] |
| | (0.019) | | (0.025) | | (0.056) | | (0.068) | |
| Techstars (Unclassified) | 0.367 | [0.000] | 0.316 | [0.000] | 0.302 | [0.000] | 0.347 | [0.000] |
| | (0.022) | | (0.031) | | (0.057) | | (0.063) | |
| Y Combinator | 1.091 | [0.000] | 1.003 | [0.000] | 0.890 | [0.000] | 0.880 | [0.000] |
| | (0.016) | | (0.015) | | (0.014) | | (0.016) | |
| 2 or 3 Founders | | | | | 0.025 | [0.067] | 0.012 | [0.859] |
| | | | | | (0.013) | | (0.068) | |
| Has Female Founders | | | | | -0.026 | [0.145] | -0.075 | [0.144] |
| | | | | | (0.018) | | (0.051) | |
| Prior Employer Prominence | | | | | 0.058 | [0.000] | 0.032 | [0.008] |
| | | | | | (0.005) | | (0.012) | |
| Years Work Exper. (L) | | | | | -0.082 | [0.000] | -0.037 | [0.369] |
| | | | | | (0.010) | | (0.041) | |
| Serial Entrepreneurs | | | | | 0.011 | [0.327] | 0.130 | [0.006] |
| | | | | | (0.011) | | (0.046) | |
| Previously Raised VC | | | | | 0.186 | [0.000] | 0.461 | [0.000] |
| | | | | | (0.025) | | (0.103) | |
| Founders Jobs Missing | | | | | -0.007 | [0.920] | 0.195 | [0.299] |
| | | | | | (0.071) | | (0.186) | |
| MBA | | | | | -0.127 | [0.000] | -0.117 | [0.141] |
| | | | | | (0.026) | | (0.079) | |
| Technical Masters | | | | | 0.047 | [0.155] | 0.093 | [0.109] |
| | | | | | (0.033) | | (0.058) | |
| Other Masters | | | | | -0.104 | [0.000] | -0.000 | [0.997] |
| | | | | | (0.018) | | (0.114) | |
| Technical PhD | | | | | -0.122 | [0.001] | -0.086 | [0.411] |
| | | | | | (0.037) | | (0.104) | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Other PhD | | | | | -0.085 | [0.000] | 0.069 | [0.693] |
| | | | | | (0.023) | | (0.174) | |
| JD | | | | | 0.017 | [0.670] | -0.077 | [0.383] |
| | | | | | (0.040) | | (0.088) | |
| University Prominence | | | | | 0.003 | [0.000] | 0.003 | [0.009] |
| | | | | | (0.000) | | (0.001) | |
| Founders Education Missing | | | | | 0.155 | [0.001] | 0.136 | [0.358] |
| | | | | | (0.045) | | (0.147) | |
| Internet Archive Crawls (-436 days; L) | | | | | -0.004 | [0.651] | 0.051 | [0.159] |
| | | | | | (0.010) | | (0.036) | |
| Startup Age | -0.094 | [0.000] | -0.133 | [0.000] | -0.113 | [0.000] | -0.115 | [0.012] |
| | (0.006) | | (0.015) | | (0.016) | | (0.045) | |
| Current Round Amount (M$; L) | 0.433 | [0.000] | 0.381 | [0.000] | 0.362 | [0.000] | 0.640 | [0.000] |
| | (0.016) | | (0.024) | | (0.027) | | (0.066) | |
| Previously Raised Capital (M$; L) | -0.008 | [0.985] | -0.207 | [0.719] | -0.179 | [0.771] | -0.857 | [0.347] |
| | (0.406) | | (0.574) | | (0.613) | | (0.907) | |
| Convertible Note | 0.134 | [0.188] | 0.176 | [0.019] | 0.156 | [0.057] | 0.338 | [0.004] |
| | (0.101) | | (0.073) | | (0.081) | | (0.114) | |
| Percentage VCs in Round | 0.400 | [0.000] | 0.389 | [0.000] | 0.387 | [0.000] | 0.265 | [0.000] |
| | (0.012) | | (0.015) | | (0.015) | | (0.040) | |
| State 3-Year VC Rounds (L) | 0.207 | [0.000] | 0.394 | [0.000] | 0.402 | [0.000] | 0.257 | [0.232] |
| | (0.019) | | (0.036) | | (0.045) | | (0.214) | |
| BuiltWith FE | No | | No | | Yes | | Yes | |
| Round Year FE | Yes | | Yes | | Yes | | Yes | |
| State FE | Yes | | Yes | | Yes | | Yes | |
| Category Vector | Yes | | Yes | | Yes | | Yes | |
| Sample | Full Sample | | Founder Subsample | | Founder Subsample | | Founder Subsample, After Matching | |
| Weights / Matching | No | | Survey-based Weights | | Survey-based Weights | | Propensity Score Matching | |
| N | 9,966 | | 3,702 | | 3,702 | | 1,816 | |
| R squared | 0.282 | | 0.330 | | 0.355 | | 0.422 | |
| Adj. R squared | 0.266 | | 0.293 | | 0.304 | | 0.331 | |

P-values in brackets. Robust standard errors in parentheses, clustered at the level of accelerators / no accelerator. Fixed effects are included for all accelerators, but only shown for the top 12 accelerators by volume of startups (ordered alphabetically). The omitted category is startups that did not participate in an accelerator.

# TABLE 13. CONTINGENCIES IN ACCELERATOR EFFECTS IN SAMPLE II

| | Accelerator-Level | Region-Level | Entrepreneur-Level | | |
|---|---|---|---|---|---|
| | (8) | (9) | (10) | (11) | (12) |
| Status of Investors in Accelerator Alumni | -0.037 [0.424] | | | | |
| | (0.046) | | | | |
| Accelerator Participant X State 3-Year VC Rounds (L) | | -0.035 [0.221] | | | |
| | | (0.029) | | | |
| Accelerator Participant X Prior Employer Prominence | | | -0.038 [0.071] | | |
| | | | (0.021) | | |
| Accelerator Participant X University Prominence | | | | -0.002 [0.053] | |
| | | | | (0.001) | |
| Accelerator Participant X Has Female Founders | | | | | -0.085 [0.333] |
| | | | | | (0.087) |
| Accelerator Fixed Effects (Not all shown) | | | | | |
| 500 Startups | 0.491 [0.000] | 0.616 [0.002] | 0.469 [0.000] | 0.495 [0.000] | 0.403 [0.000] |
| | (0.133) | (0.192) | (0.049) | (0.062) | (0.025) |
| Alchemist Accelerator | 0.241 [0.056] | 0.402 [0.070] | 0.239 [0.000] | 0.273 [0.000] | 0.164 [0.000] |
| | (0.125) | (0.220) | (0.052) | (0.074) | (0.028) |
| AngelPad | 0.430 [0.002] | 0.555 [0.004] | 0.417 [0.000] | 0.445 [0.000] | 0.332 [0.000] |
| | (0.134) | (0.188) | (0.047) | (0.064) | (0.013) |
| Boost VC | 0.635 [0.000] | 0.795 [0.000] | 0.641 [0.000] | 0.664 [0.000] | 0.554 [0.000] |
| | (0.112) | (0.199) | (0.059) | (0.059) | (0.023) |
| DreamIt | 0.536 [0.000] | 0.703 [0.001] | 0.557 [0.000] | 0.574 [0.000] | 0.459 [0.000] |
| | (0.120) | (0.212) | (0.060) | (0.079) | (0.029) |
| MassChallenge | 0.151 [0.148] | 0.269 [0.120] | 0.164 [0.000] | 0.210 [0.015] | 0.098 [0.024] |
| | (0.104) | (0.171) | (0.045) | (0.084) | (0.043) |
| Mucker Capital | 0.652 [0.000] | 0.789 [0.000] | 0.626 [0.000] | 0.647 [0.000] | 0.571 [0.000] |
| | (0.132) | (0.191) | (0.052) | (0.060) | (0.041) |
| Plug and Play | 0.203 [0.022] | 0.372 [0.048] | 0.213 [0.001] | 0.233 [0.000] | 0.143 [0.000] |
| | (0.087) | (0.185) | (0.062) | (0.045) | (0.034) |
| StartX | 0.215 [0.078] | 0.291 [0.040] | 0.206 [0.000] | 0.255 [0.002] | 0.140 [0.004] |
| | (0.120) | (0.140) | (0.048) | (0.078) | (0.048) |
| Techstars New York | 0.819 [0.000] | 0.915 [0.000] | 0.793 [0.000] | 0.826 [0.000] | 0.733 [0.000] |
| | (0.140) | (0.187) | (0.049) | (0.085) | (0.051) |
| Techstars Boston | 0.851 [0.000] | 0.944 [0.000] | 0.832 [0.000] | 0.854 [0.000] | 0.759 [0.000] |
| | (0.128) | (0.166) | (0.037) | (0.068) | (0.029) |
| Techstars Boulder | 0.396 [0.000] | 0.476 [0.001] | 0.404 [0.000] | 0.438 [0.000] | 0.324 [0.000] |
| | (0.105) | (0.133) | (0.056) | (0.086) | (0.046) |
| Techstars Seattle | 0.551 [0.000] | 0.627 [0.000] | 0.587 [0.000] | 0.584 [0.000] | 0.464 [0.000] |
| | (0.136) | (0.158) | (0.092) | (0.094) | (0.057) |
| Techstars (Unclassified) | 0.369 [0.000] | 0.482 [0.005] | 0.393 [0.000] | 0.414 [0.000] | 0.319 [0.000] |
| | (0.102) | (0.167) | (0.071) | (0.093) | (0.061) |
| Y Combinator | 1.000 [0.000] | 1.127 [0.000] | 0.983 [0.000] | 1.012 [0.000] | 0.900 [0.000] |
| | (0.139) | (0.199) | (0.046) | (0.069) | (0.017) |
| Interacted Controls | | | | | |
| State 3-Year VC Rounds (L) | 0.401 [0.000] | 0.405 [0.000] | 0.402 [0.000] | 0.399 [0.000] | 0.403 [0.000] |
| | (0.046) | (0.043) | (0.045) | (0.048) | (0.045) |
| Prior Employer Prominence | 0.058 [0.000] | 0.058 [0.000] | 0.064 [0.000] | 0.058 [0.000] | 0.058 [0.000] |
| | (0.005) | (0.005) | (0.003) | (0.005) | (0.005) |
| University Prominence | 0.003 [0.000] | 0.003 [0.000] | 0.003 [0.000] | 0.004 [0.000] | 0.003 [0.000] |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Has Female Founders | -0.026 [0.142] | -0.026 [0.151] | -0.027 [0.135] | -0.025 [0.152] | -0.011 [0.160] |
| | (0.018) | (0.018) | (0.018) | (0.017) | (0.008) |
| Other Controls | Yes | Yes | Yes | Yes | Yes |
| Survey-based Weights | Yes | Yes | Yes | Yes | Yes |
| N | 3,702 | 3,702 | 3,702 | 3,702 | 3,702 |
| R squared | 0.355 | 0.355 | 0.355 | 0.355 | 0.355 |
| Adj. R squared | 0.304 | 0.304 | 0.304 | 0.304 | 0.304 |

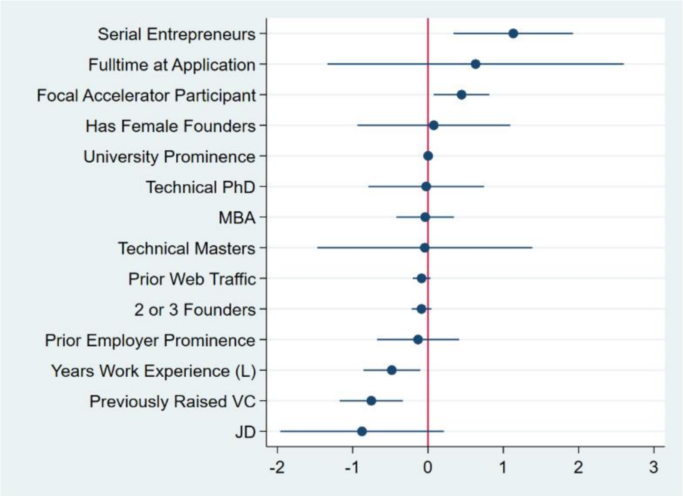
P-values in brackets. Robust standard errors in parentheses, clustered at the level of accelerators / no accelerator. Fixed effects are included for all accelerators, but only shown for the top 12 accelerators by volume of startups (ordered alphabetically). The omitted category is startups that did not participate in an accelerator. Sample is the set of startups with founder educational and/or job history data. All estimates utilize survey-based weights to account for the availability of founder information.

# TABLE 14. BOOTSTRAPPED MEDIATION RESULTS IN SAMPLE II

|  | Direct Effect | Indirect via Internet Archive Crawls | Indirect via BuiltWith Tech. Stack Measures | Net Indirect Effect | Total Effect |
|---|---|---|---|---|---|
| Accelerator Participation | 0.438<br>(CI: 0.362 to 0.515) | 0.015<br>(CI: 0.005 to 0.026) | 0.002<br>(CI: -0.008 to 0.012) | 0.016<br>(CI: 0.003 to 0.030) | 0.455<br>(CI: 0.379 to 0.531) |

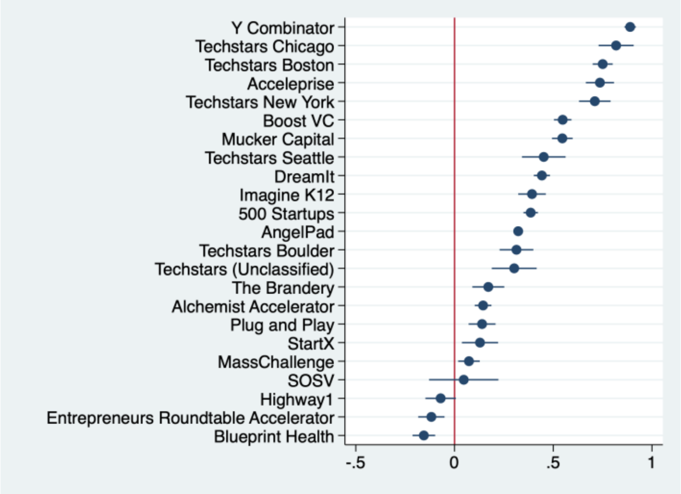
N = 3,702. CI: 95% confidence interval estimated using bootstrapping, 5,000 replications with replacement, and bias correction. Coefficients are the effects on investor status. Shading indicates aggregate effects.

# FIGURE 4. COEFFICIENT PLOT FOR SAMPLE I ESTIMATES



Coefficients are from the PSM estimates in Model 3 of Table 10, and ordered by size of point estimates. Dots represent a coefficient's point estimate and lines the 95% confidence interval.

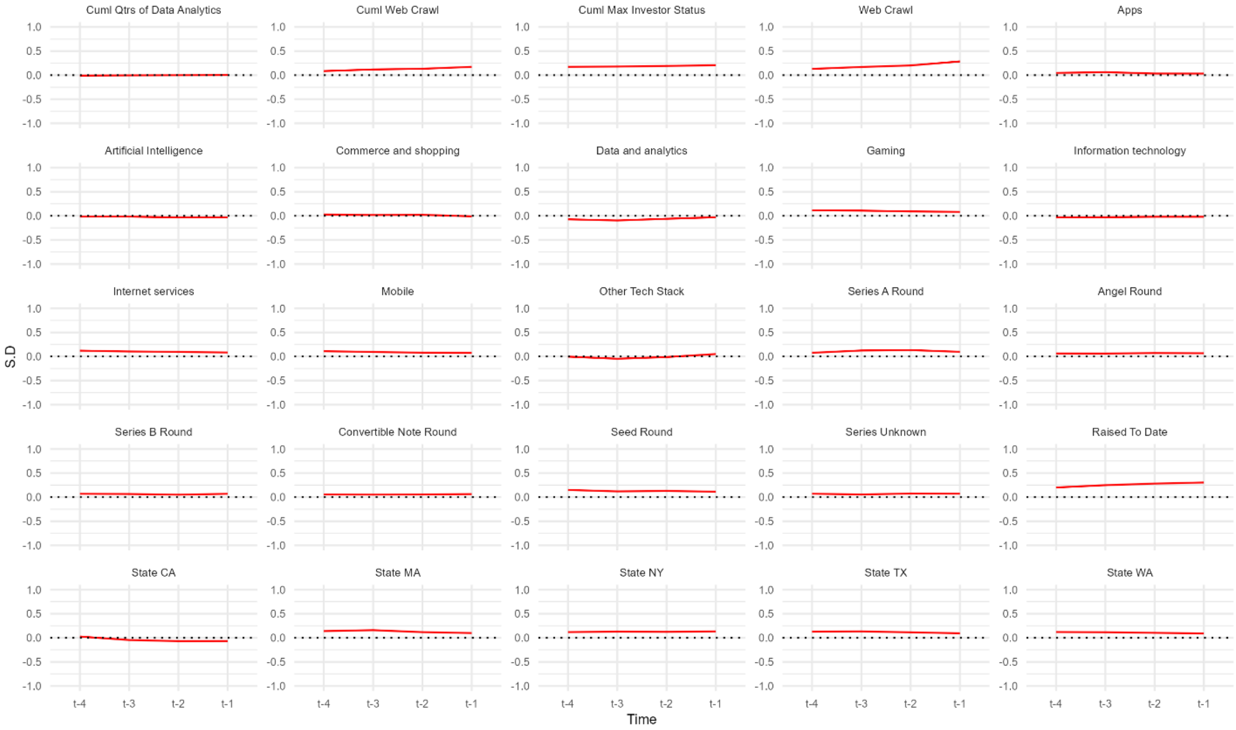
# FIGURE 5. COEFFICIENT PLOT OF MAIN ACCELERATOR EFFECTS FOR SAMPLE II



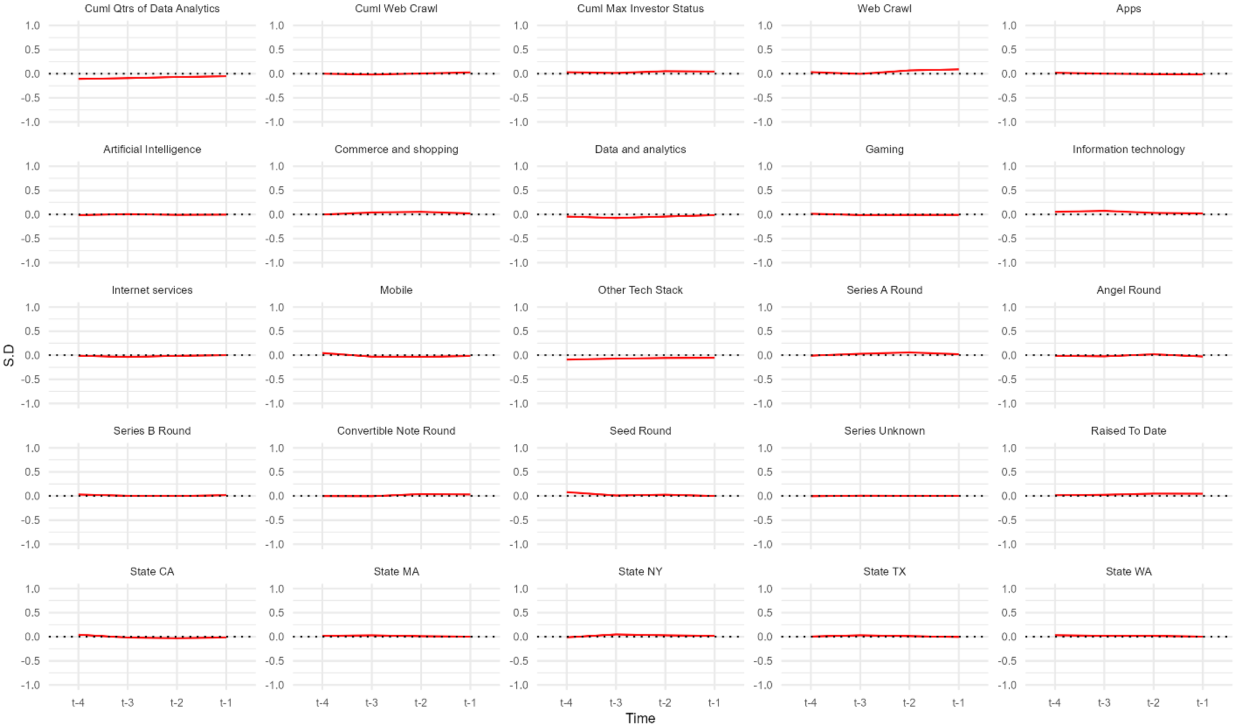
Coefficients are from Model 6 in Table 12, and are relative to startups similar on observables that did not participate in an accelerator. Coefficients are only shown for the top 25 accelerators by volume of startups in the sample, and are ordered by magnitude of point estimates. Dots represent a coefficient's point estimate and lines the 95% confidence interval.

# APPENDIX A – COVARIATE BALANCE FOR TSCS MODELS

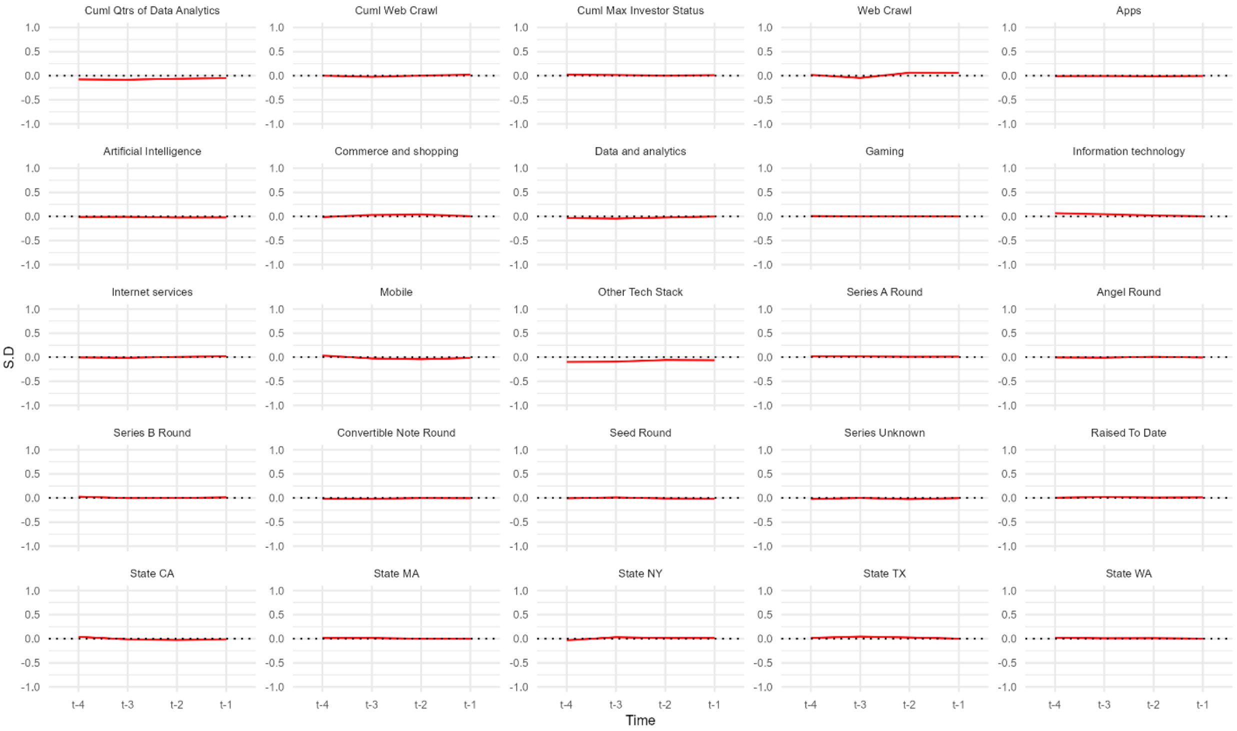
## FIGURE A1. COVARIATE BALANCE FOR MAHALANOBIS DISTANCE-BASED MATCHING MODEL



# FIGURE A2. COVARIATE BALANCE FOR PROPENSITY SCORE-BASED MATCHING MODEL



# FIGURE A3. COVARIATE BALANCE FOR PROPENSITY SCORE-BASED WEIGHTING MODEL



# APPENDIX B – ROBUSTNESS CHECKS

## TABLE B1. REGRESSION ON DATA ANALYTICS UTILIZATION

| | (1) FE - Controls | | (2) FE - Main Effect | |
|---|---|---|---|---|
| Raised To Date (USD Logged) | 0.011 | [0.000] | 0.01 | [0.000] |
| | (0.001) | | (0.001) | |
| Other Technology Stack (Count) | 0.011 | [0.000] | 0.009 | [0.000] |
| | (0.000) | | (0.001) | |
| Investor Status (Cumulative Max Centrality) | 0.514 | [0.000] | 0.478 | [0.000] |
| | (0.063) | | (0.062) | |
| Cumulative Web Crawl (Count Logged) | 0.003 | [0.069] | 0.005 | [0.001] |
| | (0.001) | | (0.001) | |
| Cumulative Quarters of Data Analytics Utilization (Count) | -0.045 | [0.000] | -0.033 | [0.000] |
| | (0.005) | | (0.005) | |
| Data Analytics Utilization (Binary) | | | 0.283 | [0.000] |
| | | | (0.015) | |
| S.E.: Clustered | by: Firm | | by: Firm | |
| Firm fixed effect | Yes | | Yes | |
| Year-Quarter fixed effect | Yes | | Yes | |
| Round Controls | Yes | | Yes | |
| Observations | 81,566 | | 81,566 | |
| No. of firms | 6,903 | | 6,903 | |
| R2 | 0.6692 | | 0.6729 | |
| adjusted R2 | 0.6382 | | 0.6423 | |

Cluster-robust standard error in parentheses; p-value in square brackets

# TABLE B2. REGRESSION ON DATA ANALYTICS UTILIZATION W/ SURVEY WEIGHTS

| | (1) FE - Controls | | (2) FE - Main Effect | |
|---|---|---|---|---|
| DV | Web Crawl Count (Logged) | | Web Crawl Count (Logged) | |
| Raised To Date (USD Logged) | 0.014 | [0.000] | 0.012 | [0.000] |
| | (0.001) | | (0.001) | |
| Other Technology Stack (Count) | 0.012 | [0.000] | 0.009 | [0.000] |
| | (0.001) | | (0.001) | |
| Investor Status (Cumulative Max Centrality) | 0.467 | [0.000] | 0.436 | [0.000] |
| | (0.073) | | (0.072) | |
| Cumulative Web Crawl (Count Logged) | 0.005 | [0.001] | 0.007 | [0.000] |
| | (0.002) | | (0.001) | |
| Cumulative Quarters of Data Analytics Utilization (Count) | -0.038 | [0.000] | -0.03 | [0.000] |
| | (0.006) | | (0.005) | |
| Data Analytics Utilization (Binary) | | | 0.278 | [0.000] |
| | | | (0.018) | |
| S.E.: Clustered | by: Firm | | by: Firm | |
| Firm fixed effect | Yes | | Yes | |
| Year-Quarter fixed effect | Yes | | Yes | |
| Round Controls | Yes | | Yes | |
| Observations | 81,566 | | 81,566 | |
| No. of firms | 6,903 | | 6,903 | |
| R2 | 0.6453 | | 0.6497 | |
| adjusted R2 | 0.6121 | | 0.6169 | |

Cluster-robust standard error in parentheses; p-value in square brackets

## TABLE B3. TIME-SERIES CROSS SECTIONAL MATCHING MODEL USING MAHALANOBIS DISTANCE-BASED MATCHING REFINEMENT WITH NARROWER SET OF DATA ANALYTICS

Estimate of Average Treatment Effect on the Treated (ATT)
*(1000 bootstrapped iterations at 95% CI)*

| | estimates | std.error | estimate at 2.5% | estimate at 97.5% |
|---|---|---|---|---|
| *t* | 0.353 | 0.020 | 0.312 | 0.391 |
| *t+1* | 0.304 | 0.024 | 0.255 | 0.352 |
| *t+2* | 0.245 | 0.028 | 0.189 | 0.295 |

## TABLE B4. TIME-SERIES CROSS SECTIONAL MATCHING MODEL USING PROPENSITY SCORE-BASED MATCHING REFINEMENT WITH NARROWER SET OF DATA ANALYTICS

Estimate of Average Treatment Effect on the Treated (ATT)
*(1000 bootstrapped iterations at 95% CI)*

| | estimates | std.error | estimate at 2.5% | estimate at 97.5% |
|---|---|---|---|---|
| *t* | 0.400 | 0.020 | 0.361 | 0.439 |
| *t+1* | 0.376 | 0.025 | 0.331 | 0.422 |
| *t+2* | 0.314 | 0.027 | 0.260 | 0.366 |

## TABLE B5. TIME-SERIES CROSS SECTIONAL MATCHING MODEL USING PROPENSITY SCORE-BASED WEIGHTING REFINEMENT WITH NARROWER SET OF DATA ANALYTICS

Estimate of Average Treatment Effect on the Treated (ATT)
*(1000 bootstrapped iterations at 95% CI)*

| | estimates | std.error | estimate at 2.5% | estimate at 97.5% |
|---|---|---|---|---|
| *t* | 0.411 | 0.019 | 0.372 | 0.447 |
| *t+1* | 0.393 | 0.022 | 0.352 | 0.436 |
| *t+2* | 0.326 | 0.024 | 0.281 | 0.375 |

## TABLE B6. TIME-SERIES CROSS SECTIONAL MATCHING MODEL USING MAHALANOBIS DISTANCE-BASED MATCHING REFINEMENT WITH 8 QUARTERS OF TREATMENT HISTORY COMPARISON

Estimate of Average Treatment Effect on the Treated (ATT)
*(1000 bootstrapped iterations at 95% CI)*

| | estimates | std.error | estimate at 2.5% | estimate at 97.5% |
|---|---|---|---|---|
| *t* | 0.362 | 0.021 | 0.321 | 0.402 |
| *t+1* | 0.307 | 0.025 | 0.255 | 0.353 |
| *t+2* | 0.249 | 0.029 | 0.191 | 0.302 |

## TABLE B7. TIME-SERIES CROSS SECTIONAL MATCHING MODEL USING PROPENSITY SCORE-BASED MATCHING REFINEMENT WITH 8 QUARTERS OF TREATMENT HISTORY COMPARISON

Estimate of Average Treatment Effect on the Treated (ATT)
*(1000 bootstrapped iterations at 95% CI)*

| | estimates | std.error | estimate at 2.5% | estimate at 97.5% |
|---|---|---|---|---|
| *t* | 0.406 | 0.020 | 0.365 | 0.445 |
| *t+1* | 0.363 | 0.025 | 0.316 | 0.411 |
| *t+2* | 0.303 | 0.027 | 0.248 | 0.356 |

## TABLE B8. TIME-SERIES CROSS SECTIONAL MATCHING MODEL USING PROPENSITY SCORE-BASED WEIGHTING REFINEMENT WITH 8 QUARTERS OF TREATMENT HISTORY COMPARISON

Estimate of Average Treatment Effect on the Treated (ATT)
*(1000 bootstrapped iterations at 95% CI)*

| | estimates | std.error | estimate at 2.5% | estimate at 97.5% |
|---|---|---|---|---|
| *t* | 0.422 | 0.019 | 0.384 | 0.458 |
| *t+1* | 0.393 | 0.022 | 0.352 | 0.438 |
| *t+2* | 0.328 | 0.025 | 0.282 | 0.378 |

# APPENDIX C. SAMPLE COMPARISON BASED ON FOUNDER DATA AVAILABILITY

We compare whether there are significant and substantive differences between the sample with founder data available versus the sample without. Crunchbase lists founder profile links (LinkedIn URLs) for 4,058 out of 9,547 ventures (42.5%) in our initial sample. Using the founder data availability as the grouping (founder profile link available = 1; founder profile link not available = 0), we conduct a t-test to assess the differences in means of variables that are readily observable in both samples. We include the year of initial priced round deal, amount raised in initial priced round, number of investors in initial priced round, region (state), and initial industry category. Results of the t-test show that many state and industry category dummies are significantly different, yet most are substantively not too dissimilar (Table C1). For instance, while some industry categories and geographical location show a p-value below 0.05, the numerical difference is not meaningfully distinct. On the contrary, there is a noticeably larger difference in the amount raised between the two samples, likely influenced by the skew in distribution. The mean in the sample with missing founder information was $2.6 Million whereas that of the sample with founder information was $1.7 Million. In addition, there is a greater abundance of deals across the years 2012-2016 for the sample with available founder information.

# TABLE C1. T-TEST BETWEEN SAMPLE WITH & WITHOUT FOUNDER DATA

| Variable | Mean (Founder Information Missing) | Mean (Founder Information Available) | Difference | T-Statistic | P-value |
|---|---|---|---|---|---|
| Amount Raised in Initial Priced Round (Million USD Logged) | 1.307 | 1.008 | 0.299 | 21.932 | 0.000 |
| Number of Investors in Initial Priced Round (Logged) | 0.702 | 1.031 | -0.329 | -20.146 | 0.000 |
| Year 2012 | 0.079 | 0.115 | -0.036 | -5.814 | 0.000 |
| Year 2013 | 0.077 | 0.151 | -0.074 | -11.021 | 0.000 |
| Year 2014 | 0.079 | 0.170 | -0.091 | -13.243 | 0.000 |
| Year 2015 | 0.090 | 0.160 | -0.070 | -10.061 | 0.000 |
| Year 2016 | 0.075 | 0.128 | -0.053 | -8.376 | 0.000 |
| WA State | 0.038 | 0.035 | 0.003 | 0.592 | 0.554 |
| MA State | 0.071 | 0.060 | 0.011 | 2.229 | 0.026 |
| TX State | 0.043 | 0.034 | 0.009 | 2.255 | 0.024 |
| NY State | 0.125 | 0.166 | -0.041 | -5.649 | 0.000 |
| CA State | 0.468 | 0.470 | -0.002 | -0.202 | 0.840 |
| Information Technology | 0.065 | 0.059 | 0.006 | 1.199 | 0.231 |
| Advertising | 0.075 | 0.065 | 0.010 | 1.854 | 0.064 |
| Apps | 0.064 | 0.084 | -0.020 | -3.639 | 0.000 |
| Data & Analytics | 0.070 | 0.090 | -0.020 | -3.532 | 0.000 |
| Commerce & Shopping | 0.081 | 0.098 | -0.017 | -2.748 | 0.006 |
| *Sample size* | *N = 5,489* | *N = 4,058* | | | |

Note: Only Top 5 states, industry categories, and year of initial priced round in frequency are included

Table C2 shows the logistic regression model used to calculate the survey weights. We start by constructing the dependent variable as a binary measure of whether both the focal venture's founder data *and* initial priced round pre-money valuation data are available (1950 out of 9547 ventures had both data available). The probability of both data being available is then calculated using a logit model based on the following variables: *the raised amount (logged) at initial round (regardless of whether unpriced or priced), investor count at initial round, initial industry category, state (region),* and *deal year of initial round*. The probability is then inversed (=1/probability of both data being available) to create survey-based weights for each venture. Adding survey-based weights to the analyses helps reduce the effect of overrepresented observations while increasing the influence of underrepresented observations, both of which occurs from differences in the likelihood of founder and valuation information availability. We trim the weights at the 99% percentile to reduce variability among the weights, which helps to improve precision (Cole & Hernán, 2008), and prevent extreme weights that can the skew analysis (Thoemmes & Ong, 2016).

# TABLE C2. LOGISTICS REGRESSION - PREDICTING LIKELIHOOD OF FOUNDER INFORMATION & VALUATION DATA BEING AVAILABLE

| | *Dependent variable:* |
|---|---|
| | Has Founder AND Pre-money Valuation Data (Binary) |
| Amount Raised in Initial Priced Round (Million USD Logged) | 0.0005 |
| | (0.042) |
| Number of Investors in Initial Priced Round (Logged) | $0.413^{***}$ |
| | (0.034) |
| Initial Deal Year 2012 | $3.281^{***}$ |
| | (0.459) |
| Initial Deal Year 2013 | $3.170^{***}$ |
| | (0.458) |
| Initial Deal Year 2014 | $3.380^{***}$ |
| | (0.457) |
| Initial Deal Year 2015 | $3.383^{***}$ |
| | (0.457) |
| Initial Deal Year 2016 | $3.283^{***}$ |
| | (0.459) |
| CA State | -0.254 |
| | (0.339) |
| MA State | 0.049 |
| | (0.352) |
| NY State | -0.183 |
| | (0.343) |
| TX State | -0.256 |
| | (0.364) |
| WA State | -0.240 |
| | (0.365) |
| Industry: Advertising | -0.082 |
| | (0.276) |
| Industry: Apps | -0.109 |
| | (0.271) |
| Industry: Commerce/Shopping | 0.076 |
| | (0.267) |
| Industry: Data/Analytics | 0.305 |
| | (0.268) |
| Information Technology | 0.107 |
| | (0.276) |
| Constant | $-4.484^{***}$ |
| | (0.619) |
| Observations | 9,547 |
| Log Likelihood | -4,268.019 |
| Akaike Inf. Crit. | 8,716.037 |

*Note:*
*Standard Error shown in parentheses; Only top Only Top 5 states, industry categories, and year of initial priced round shown in focal table.*

$^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

# APPENDIX D. VARIANCE DECOMPOSITION WITH ONE ROUND PER VENTURE

## TABLE D1. SUMMARY OF VARIANCE EXPLAINED BY ROUND INCLUDING ONLY ONE ROUND PER VENTURE

| | Initial Round | | Early Follow-On Rounds | | Later Rounds | |
|---|---|---|---|---|---|---|
| factor | R squared (%)[a] | Adj. R squared (%)[a] | R squared (%)[a] | Adj. R squared (%)[a] | R squared (%)[a] | Adj. R squared (%)[a] |
| *Region* | 12.1 | 10.21 | 7.07 | 4.89 | 8.87 | 5.92 |
| | [11.83,12.36] | [9.93,10.49] | [6.82,7.32] | [4.63,5.16] | [8.55,9.18] | [5.56,6.28] |
| *Industry* | 3.22 | 1.84 | 5.05 | 3.17 | 8.89 | 6.56 |
| | [3.09,3.35] | [1.7,1.98] | [4.83,5.27] | [2.94,3.41] | [8.47,9.32] | [6.09,7.04] |
| *Year* | 3.65 | 3.24 | 4.7 | 3.96 | 6.99 | 5.88 |
| | [3.48,3.83] | [3.06,3.42] | [4.46,4.93] | [3.71,4.2] | [6.71,7.27] | [5.58,6.19] |
| *Founding team attributes* | 5.45 | 4.67 | 5.78 | 4.52 | 9.15 | 7.22 |
| | [5.27,5.63] | [4.48,4.85] | [5.52,6.04] | [4.25,4.79] | [8.79,9.5] | [6.82,7.62] |
| *Venture progress* | 3.6 | 3.39 | 3.81 | 3.49 | 11.46 | 11.47 |
| | [3.43,3.76] | [3.22,3.57] | [3.6,4.02] | [3.27,3.71] | [11.08,11.85] | [11.05,11.89] |
| *Past resource mobilization* | | | 35.52 | 37.14 | 37.39 | 40.69 |
| | | | [34.92,36.13] | [36.5,37.78] | [36.77,38.01] | [40,41.38] |
| *Total variance explained* | 28.01 | 23.35 | 61.93 | 57.18 | 82.75 | 77.75 |
| | [27.62,28.41] | [22.93,23.77] | [61.33,62.53] | [56.51,57.85] | [82.21,83.29] | [77.06,78.44] |
| | *N = 1,737* | | *N = 837* | | *N = 374* | |

[a] Mean variance explained; Bootstrapped 95% Confidence Interval in square brackets (100 iterations)

# APPENDIX E. VARIANCE DECOMPOSITION INCLUDING CURRENT ROUND ATTRIBUTES

All analyses in Appendix E include an additional factor – current-round attributes. To measure the *investor-investment attributes* of the current round, we utilized investor fixed effects to both capture the tendency of investors to command higher or lower price premiums (Hsu, 2004) and sorting dynamics whereby more desirable investors are likely to attract and invest in higher quality ventures that garner greater competition (Fox, Yang, & Hsu, 2018; Mindruta et al., 2016). To ensure we included only one investor fixed effect per round, we used the *highest status investor in the round* as literature suggests these may serve as proxies for quality of advice, network, and signaling benefits that a venture receives from investors, as well as the desirability of investors and their ability to command a price discount (Hallen, 2008; Hsu, 2004; Sørensen, 2007). We measured investor status using eigenvector centrality in the five-year co-investment syndication network ending in the year prior to the focal round (Hallen, 2008; Hochberg, Ljungqvist, & Lu, 2007).[32] Like our other fixed effects, we included dummies for investors that appeared at least three times in the data as the most central investor, aggregating other investors into an “other investor” category. This yielded 160 investor fixed effects for our initial round sample.

As a greater number of investors can indicate more attractive rounds, we included *number of investors*, logged to reduce skewness. Since certain types of investors may be more or less desirable, or willing to invest at higher valuations, we included dummy variables to indicate if a round’s syndicate contained any *individual angel investors*, *angel investor groups*, or *CVCs*.

Finally, we accounted for the relative control obtained by investors. We controlled for the *percentage of equity acquired* by an investor syndicate in a given round, calculated as the capital invested divided by the post-money valuation. Following Wasserman (2017), we included a

---

[32] We thank Pitchbook for generously providing us with this industry-level network.

dummy for *CEO position still held by founder* to indicate whether any founders were CEO at the time of the round. As the CEO was not listed for some rounds in Pitchbook, we also included a dummy *CEO is unknown* and set *CEO position still held by founder* to zero in such circumstances.[33]

Table E1 shows the result for the initial round. A large proportion of the total variance explained in valuation in the initial round is accounted for by investor-investment attributes (Adj. $R^2$=21.38%) – much larger than any other factor in the model. Region accounts for 8.11%, followed by other factors that explain a relatively small amount of variance: founding team attributes (3.96%), venture progress (3.48%), year (3.33%), and industry (1.86%). Comparing the factors' adjusted $R^2$ to our main analysis, there is no substantial difference. The largest difference is in region, with a 2.1% difference in adjusted $R^2$ (all others have a difference of <1%).

Table E1 shows the results for the early follow-on rounds. Comparing the model with and without investor-investment attributes for the early follow-on rounds, there is no meaningful difference in the adjusted $R^2$ values for founding team attributes, region, year, industry, and venture progress factors (all have a difference of <1%). Moreover, similar to the initial round, the investor-investment attributes account for a substantial amount of variance in valuation (Adj. $R^2$=23.12%). Simultaneously, the inclusion of investor-investment attributes leads to a small but noticeable reduction in adjusted $R^2$ of past resource mobilization compared to our main analysis (Adj. $R^2$ shifts from 38.62% → 32.7%).

Table E1 shows the results for the later rounds. The same pattern seen in the early follow-on rounds is repeated in the later rounds. All factors aside from past resource mobilization and investor-investment attributes show little difference in the variance explained compared to the main analysis (< 3% difference in Adj. $R^2$), respectively. Moreover, investor-investment attributes

---

[33] Our data do not describe qualitative control rights utilized by investors at the deal level (antidilution protections, liquidation preferences, etc.). Thus, like past studies (Hochberg et al. 2010), we left this as an opportunity for future research.

explain a substantial portion of variance (Adj. $R^2$=27.09%). Again, the inclusion of investor-investment attributes leads to a sizeable reduction in the variance explained by past resource mobilization (Adj. $R^2$ shifts from 45.95% → 33.79%)

These results are rather surprising given the existing literature's elaboration on mediating effects of observable venture and founder-level factors that may come to play, as briefly discussed above. Our results instead imply that investor-investment attributes may be much more strongly driven by difficult to observe/hard to quantify factors such as venture strategy, opportunity size, founder personality, or other uncaptured venture/investor-level idiosyncrasies than observable factors measured in our main analysis – a future research agenda that requires more attention.

On the other hand, the inclusion of investor-investment attributes noticeably reduces the influence of past resource mobilization by some margin, especially so in the later rounds (difference in Adj. $R^2$ between model with and without investor-investment attributes in early follow-on rounds: 5.92%; later rounds: 12.16%). We may be seeing such results potentially because past investor status is included as part of the past resource mobilization factor, which is likely correlated to the current deal investor status. This is due to matching dynamics around status where an affiliation with high-status VCs in the prior rounds can help attract comparable status VCs in the next (Hallen, 2008; Podolny, 1994). Another reason may be that prior valuation and amount raised can be closed associated with current investor status or degree of control conferred (Gompers et al., 2020; Hsu, 2007), leading to a larger degree of shared explanatory power between past resource mobilization and investor-investment attributes.

# TABLE E1. SUMMARY OF VARIANCE EXPLAINED BY ROUND INCLUDING INVESTOR/INVESTMENT ATTRIBUTES

| | Initial Round | | Early Follow-On Rounds | | Later Rounds | |
|---|---|---|---|---|---|---|
| factor | R squared (%)[a] | Adj. R squared (%)[a] | R squared (%)[a] | Adj. R squared (%)[a] | R squared (%)[a] | Adj. R squared (%)[a] |
| *Region* | 9.54 | 8.11 | 5.24 | 4 | 6.45 | 5.04 |
| | [9.33,9.74] | [7.89,8.33] | [5.1,5.37] | [3.85,4.14] | [6.27,6.63] | [4.84,5.24] |
| *Industry* | 3 | 1.86 | 3.8 | 2.72 | 5.79 | 4.65 |
| | [2.88,3.11] | [1.74,1.98] | [3.63,3.96] | [2.54,2.89] | [5.62,5.96] | [4.46,4.83] |
| *Year* | 3.54 | 3.33 | 3.33 | 2.93 | 4.15 | 3.53 |
| | [3.4,3.69] | [3.17,3.49] | [3.22,3.45] | [2.81,3.05] | [3.99,4.31] | [3.36,3.71] |
| *Founding team attributes* | 4.53 | 3.96 | 3.79 | 3.06 | 5.79 | 4.9 |
| | [4.38,4.67] | [3.8,4.12] | [3.67,3.91] | [2.93,3.19] | [5.64,5.94] | [4.73,5.06] |
| *Venture progress* | 3.46 | 3.48 | 3.54 | 3.44 | 8.91 | 8.91 |
| | [3.33,3.6] | [3.33,3.62] | [3.4,3.68] | [3.29,3.59] | [8.66,9.17] | [8.64,9.19] |
| *Investor-investment attributes* | 27.03 | 21.38 | 26.49 | 23.12 | 28.33 | 27.09 |
| | [26.63,27.42] | [20.94,21.82] | [26.06,26.92] | [22.64,23.61] | [27.94,28.72] | [26.65,27.53] |
| *Past resource mobilization* | | | 30.59 | 32.7 | 31.48 | 33.79 |
| | | | [30.19,31] | [32.26,33.15] | [31.11,31.85] | [33.38,34.2] |
| *Total variance explained* | 51.09 | 42.12 | 76.79 | 71.96 | 90.91 | 87.91 |
| | [50.69,51.48] | [41.65,42.59] | [76.42,77.16] | [71.52,72.41] | [90.64,91.17] | [87.56,88.26] |
| | *N = 1,737* | | *N = 1,170* | | *N = 602* | |

[a] Mean variance explained; Bootstrapped 95% Confidence Interval in square brackets (100 iterations)

# APPENDIX F. NON-BOOTSTRAPPED SHAPLEY VALUE VARIANCE DECOMPOSITION

TABLE F1. SUMMARY OF VARIANCE EXPLAINED BY ROUND NON-BOOTSTRAPPED OUTPUT

| | Initial Round | | Early Follow-On Rounds | | Later Rounds | |
|---|---|---|---|---|---|---|
| factor | R squared (%) | Adj. R squared (%) | R squared (%) | Adj. R squared (%) | R squared (%) | Adj. R squared (%) |
| *Region* | 10 | 8 | 4.8 | 3 | 6.6 | 4.5 |
| *Industry* | 2 | 0.4 | 3.3 | 1.8 | 5.9 | 4 |
| *Year* | 3 | 2.6 | 3.2 | 2.5 | 4.4 | 3.4 |
| *Founding team attributes* | 4.5 | 3.6 | 3.7 | 2.6 | 6.1 | 4.7 |
| *Venture progress* | 3.4 | 3.2 | 4.1 | 3.9 | 11.4 | 11.3 |
| *Past resource mobilization* | | | 38.5 | 39.8 | 45.5 | 48.3 |
| *Total variance explained* | 22.9 | 17.8 | 57.6 | 53.6 | 79.9 | 76.2 |
| | *N = 1,737* | | *N = 1,170* | | *N = 602* | |

The non-bootstrapped results show largely similar results to the bootstrapped version in the main analysis. Comparing factor by factor per round, the largest difference in adjusted $R^2$ was past resource mobilization in the later rounds (2.35% difference in adjusted $R^2$). All other factors in each round of the non-bootstrapped model showed a difference of less than 2.3% compared to the bootstrapped model.

# APPENDIX G

To better isolate our theoretical question of accelerators' impact on startups obtaining higher or lower-status early partners, we focus on a set of accelerators in Sample I whose impact on startup development was examined in Hallen et al. (2020). In this earlier work, we assessed startup development by estimating the impact of the four focal accelerator cohorts on three startup performance outcomes gathered in the summer of 2014 project: whether a startup was still ongoing or had been acquired (versus has shutdown), whether it had 11 or more employees, and web traffic levels (note that these outcomes are generally not available longitudinally). Additionally, and consistent with our focus in this paper on equity investors and using updated data, we also examined whether a startup was listed as having any investors in the five years following what would have been the start of focal accelerator participation.

Table G1 presents descriptive statistics and correlations for the 235 accepted and almost accepted startups. Table G2 presents the estimates for these startup outcomes. Models utilize inverse probability of treatment weights, placing the most emphasis on barely-accepted or barely-rejected startups. Estimates reveal accelerator participation has substantial effects on these outcomes. Focal accelerator participants are additively 30.9% more likely to still be currently ongoing in the summer of 2014 (p=0.000), 37.2% more likely additively to have eleven or more employees (p=0.018), and to have 233% more web traffic one year after graduation (e^1.204-1; p=0.017). They also enjoy a greater likelihood of raising investments, with an average additive increase of 41.3% (p=0.001). The evidence on participating in another (non-focal) accelerator is more limited, with some marginal evidence for an impact on currently ongoing and an effect on having future investors quite similar to that of the focal accelerators. Overall, estimates in Table A2 indicate the focal accelerators had a positive and substantial average impact on participating startups’ performance.

## TABLE G1. SUMMARY AND PAIRWISE CORRELATION STATISTICS FOR FULL ALMOST ACCEPTED SAMPLE

| | | Mean | S.D. | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) | Currently Ongoing | 0.562 | 0.497 | | | | | | | | | | | | |
| (2) | 11 or more employees | 0.237 | 0.427 | *NA* | | | | | | | | | | | |
| (3) | Web traffic in one year | 1.527 | 2.190 | 0.34 | *0.18* | | | | | | | | | | |
| (4) | Has Future Investors | 0.362 | 0.482 | 0.45 | 0.23 | 0.36 | | | | | | | | | |
| (5) | Focal Accelerator Participant | 0.187 | 0.391 | 0.25 | 0.35 | 0.24 | 0.37 | | | | | | | | |
| (6) | Other Accelerator Participant | 0.166 | 0.373 | 0.05 | -0.11 | 0.04 | 0.26 | -0.21 | | | | | | | |
| (7) | Cohort B-2012 Applicant | 0.268 | 0.444 | 0.09 | 0.11 | 0.06 | 0.10 | 0.01 | 0.01 | | | | | | |
| (8) | Cohort X-2012 Applicant | 0.268 | 0.444 | 0.13 | 0.13 | 0.09 | 0.10 | 0.01 | 0.01 | -0.37 | | | | | |
| (9) | Cohort E-2012 Applicant | 0.226 | 0.419 | -0.10 | -0.19 | -0.12 | -0.09 | 0.00 | -0.10 | -0.33 | -0.33 | | | | |
| (10) | Prior Web Traffic | 1.325 | 1.914 | 0.15 | 0.04 | 0.71 | 0.19 | 0.12 | 0.07 | -0.04 | 0.13 | -0.14 | | | |
| (11) | 2 or 3 Founders | 0.609 | 0.489 | 0.10 | -0.07 | 0.10 | 0.13 | 0.16 | 0.03 | 0.09 | 0.09 | -0.03 | 0.13 | | |
| (12) | Has Female Founders | 0.191 | 0.394 | 0.10 | -0.15 | 0.01 | 0.11 | -0.01 | 0.02 | 0.02 | 0.10 | -0.06 | 0.01 | 0.01 | |
| (13) | Fulltime at Application | 0.529 | 0.431 | 0.20 | 0.04 | 0.13 | 0.21 | 0.17 | 0.05 | 0.19 | 0.08 | -0.06 | 0.16 | 0.07 | 0.04 |
| (14) | Years Work Experience (L) | 1.853 | 0.900 | 0.19 | -0.02 | 0.06 | 0.22 | 0.17 | 0.01 | 0.05 | -0.01 | 0.06 | -0.01 | 0.08 | 0.16 |
| (15) | Years Work Exp. Missing | 0.089 | 0.286 | -0.05 | -0.07 | -0.04 | -0.20 | -0.11 | -0.06 | -0.05 | -0.12 | -0.03 | 0.02 | -0.18 | -0.11 |
| (16) | Prior Employer Prominence | 0.145 | 0.573 | -0.03 | 0.04 | 0.03 | 0.08 | 0.18 | 0.02 | 0.12 | -0.05 | -0.01 | -0.03 | 0.14 | 0.09 |
| (17) | University Prominence | 50.402 | 22.833 | -0.03 | -0.03 | 0.04 | 0.04 | -0.06 | 0.04 | -0.04 | 0.05 | -0.04 | 0.05 | 0.13 | 0.05 |
| (18) | MBA | 0.345 | 0.476 | 0.05 | -0.20 | -0.14 | 0.05 | -0.05 | -0.01 | -0.05 | -0.05 | 0.04 | -0.13 | -0.02 | 0.12 |
| (19) | JD | 0.026 | 0.158 | 0.09 | 0.08 | -0.08 | 0.10 | 0.06 | -0.07 | -0.04 | -0.04 | 0.11 | -0.05 | 0.02 | -0.08 |
| (20) | Technical Masters | 0.243 | 0.430 | 0.06 | 0.13 | -0.02 | -0.03 | 0.03 | 0.01 | 0.06 | -0.01 | -0.04 | -0.03 | 0.17 | -0.07 |
| (20) | Technical PhD | 0.068 | 0.252 | 0.14 | 0.01 | 0.00 | 0.01 | 0.09 | -0.03 | -0.09 | 0.10 | -0.02 | -0.03 | 0.04 | 0.04 |
| (21) | Serial Entrepreneurs | 0.502 | 0.501 | 0.25 | 0.20 | 0.23 | 0.31 | 0.13 | 0.06 | 0.14 | 0.16 | -0.09 | 0.24 | 0.26 | 0.07 |
| (22) | Previously Raised VC | 0.034 | 0.182 | 0.17 | 0.17 | 0.07 | 0.20 | 0.27 | -0.08 | 0.05 | 0.10 | -0.05 | -0.01 | -0.04 | 0.09 |

| | | (13) | (14) | (15) | (16) | (17) | (18) | (19) | (20) | (21) | (22) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| (14) | Years Work Experience (L) | 0.19 | | | | | | | | | |
| (15) | Years Work Exp. Missing | -0.18 | -0.65 | | | | | | | | |
| (16) | Prior Employer Prominence | 0.10 | 0.06 | -0.08 | | | | | | | |
| (17) | University Prominence | -0.06 | -0.02 | -0.11 | 0.11 | | | | | | |
| (18) | MBA | -0.01 | 0.20 | -0.04 | 0.06 | 0.04 | | | | | |
| (19) | JD | -0.02 | 0.09 | -0.05 | 0.09 | 0.05 | -0.06 | | | | |
| (20) | Technical Masters | -0.08 | 0.06 | -0.04 | 0.16 | 0.01 | 0.07 | -0.03 | | | |
| (20) | Technical PhD | 0.04 | 0.00 | 0.09 | 0.14 | -0.04 | 0.02 | -0.04 | 0.24 | | |
| (21) | Serial Entrepreneurs | 0.20 | 0.20 | -0.17 | -0.06 | 0.05 | -0.07 | 0.05 | 0.03 | 0.07 | |
| (22) | Previously Raised VC | 0.07 | 0.10 | -0.06 | 0.07 | 0.00 | 0.01 | -0.03 | 0.06 | 0.14 | 0.19 |

N = 235 for all variables except web traffic in one year (N = 227) and 11 or more employees (N = 118); web traffic is missing some values due to data availability, while 11 or more employees only takes on values for startups that had not been acquired or shutdown at the time of data collection. Has future investors restricted to rounds with disclosed future investors (5 startups had future rounds where listed investors were undisclosed); this is the sample used in the main analyses.

# TABLE G2. ACCELERATOR IMPACT ON STARTUP PERFORMANCE (SAMPLE I)

| | (1) | | (2) | | (3) | | (4) | |
|---|---|---|---|---|---|---|---|---|
| | Currently Ongoing | | 11 or more employees | | Web traffic in one year | | Has Future Investors | |
| Focal Accelerator Participant | 0.309 | [0.000] | 0.372 | [0.018] | 1.204 | [0.017] | 0.413 | [0.001] |
| | (0.040) | | (0.121) | | (0.388) | | (0.081) | |
| Other Accelerator Participant | 0.124 | [0.108] | -0.086 | [0.225] | 0.340 | [0.355] | 0.419 | [0.000] |
| | (0.067) | | (0.064) | | (0.342) | | (0.053) | |
| Cohort B-2012 Applicant | 0.171 | [0.004] | 0.335 | [0.085] | 1.488 | [0.059] | 0.324 | [0.039] |
| | (0.041) | | (0.167) | | (0.661) | | (0.128) | |
| Cohort X-2012 Applicant | 0.243 | [0.001] | 0.041 | [0.708] | 0.574 | [0.309] | 0.167 | [0.179] |
| | (0.043) | | (0.104) | | (0.523) | | (0.112) | |
| Cohort E-2012 Applicant | 0.071 | [0.259] | -0.022 | [0.640] | 0.994 | [0.198] | 0.149 | [0.239] |
| | (0.058) | | (0.046) | | (0.700) | | (0.116) | |
| Prior Web Traffic | 0.020 | [0.158] | 0.006 | [0.901] | 0.827 | [0.000] | 0.038 | [0.045] |
| | (0.013) | | (0.045) | | (0.087) | | (0.016) | |
| 2 or 3 Founders | -0.027 | [0.745] | -0.189 | [0.117] | -0.117 | [0.685] | 0.005 | [0.953] |
| | (0.079) | | (0.106) | | (0.277) | | (0.084) | |
| Has Female Founders | -0.020 | [0.838] | 0.148 | [0.414] | -0.067 | [0.767] | 0.087 | [0.114] |
| | (0.095) | | (0.170) | | (0.217) | | (0.048) | |
| Fulltime at Application | 0.045 | [0.636] | -0.109 | [0.208] | -0.680 | [0.274] | -0.091 | [0.348] |
| | (0.091) | | (0.078) | | (0.573) | | (0.090) | |
| Years Work Experience (L) | 0.080 | [0.058] | -0.048 | [0.431] | 0.221 | [0.532] | 0.105 | [0.143] |
| | (0.036) | | (0.058) | | (0.336) | | (0.064) | |
| Years Work Exp. Missing | 0.297 | [0.015] | -0.144 | [0.433] | -0.206 | [0.804] | -0.133 | [0.291] |
| | (0.093) | | (0.173) | | (0.797) | | (0.116) | |
| Prior Employer Prominence | -0.092 | [0.067] | -0.032 | [0.846] | -0.042 | [0.901] | 0.004 | [0.946] |
| | (0.042) | | (0.158) | | (0.327) | | (0.064) | |
| University Prominence | 0.002 | [0.135] | -0.004 | [0.031] | 0.008 | [0.314] | -0.001 | [0.560] |
| | (0.001) | | (0.001) | | (0.007) | | (0.002) | |
| MBA | 0.130 | [0.144] | -0.111 | [0.480] | 0.180 | [0.687] | 0.091 | [0.133] |
| | (0.079) | | (0.148) | | (0.429) | | (0.054) | |
| JD | 0.329 | [0.039] | 0.280 | [0.399] | -0.930 | [0.253] | 0.353 | [0.090] |
| | (0.130) | | (0.313) | | (0.747) | | (0.180) | |
| Technical Masters | 0.081 | [0.316] | 0.118 | [0.449] | -0.120 | [0.725] | -0.115 | [0.083] |
| | (0.075) | | (0.147) | | (0.329) | | (0.057) | |
| Technical PhD | 0.124 | [0.329] | -0.142 | [0.520] | 0.290 | [0.494] | -0.040 | [0.773] |
| | (0.119) | | (0.210) | | (0.402) | | (0.135) | |
| Serial Entrepreneurs | 0.183 | [0.044] | 0.176 | [0.088] | 0.179 | [0.340] | 0.203 | [0.059] |
| | (0.075) | | (0.089) | | (0.175) | | (0.090) | |
| Previously Raised VC | 0.130 | [0.304] | -0.019 | [0.862] | 0.122 | [0.898] | 0.156 | [0.057] |
| | (0.117) | | (0.105) | | (0.918) | | (0.068) | |
| Constant | -0.127 | [0.315] | 0.411 | [0.146] | -1.046 | [0.433] | -0.222 | [0.262] |
| | (0.118) | | (0.251) | | (1.258) | | (0.182) | |
| Nature of Outcome | Binary | | Binary | | Ln(thousands of daily page views) | | Binary | |
| Estimation model | Linear Probability Model | | Linear Probability Model | | Least Squares | | Linear Probability Model | |
| Weighting | IPTW | | IPTW | | IPTW | | IPTW | |
| N | 235 | | 118 | | 227 | | 235 | |
| R squared | 0.292 | | 0.389 | | 0.551 | | 0.455 | |
| Adj. R squared | 0.230 | | 0.270 | | 0.510 | | 0.406 | |

P-values in brackets. Robust standard errors in parentheses, clustered at level of accelerator cohort / application pool. Has future investors is restricted to disclosed future investors. For ease of interpretation, we use linear probability models for discrete outcomes and ordinary least squares for continuous outcomes.

# APPENDIX H - FURTHER DETAILS ON HIGH-STATUS INVESTORS

Venture capital has been a popular context for studying status and inter-organizational networks for over twenty-five years (Stuart et al., 1999; Podolny, 2001; Sorenson & Stuart, 2001; Ozmel et al., 2013; Ter Wal et al., 2016; Wang et al., 2021). Yet the competitive landscape of venture finance has recently evolved, including the introduction of new organizational forms like the "Micro VC" (McDonald, Burke, Franking, & Tempest, 2017) and new mega venture capital firms like Andreessen Horowitz, with some adopting novel approaches to competing and deal structures (Eisenmann & Kind, 2014).

Consequently, classic patterns and assumptions in the field of entrepreneurship may need a "fresh look" (Dushnitsky and Matusik, 2019). While we examine this more fully in separate research (Park, Hallen, and McDonald, 2022), here we present descriptive statistics showing the composition of investors across status tiers and how this changed over the duration of our study (e.g., VC firms, individual angels, angel networks, etc.). This table was constructed utilizing the Crunchbase investor data and status measure described in the main text. We also note that while our status measure shifts the scalar differences between status positions, it preserves the ordinality of investors' eigenvector centrality. Thus, any shifts either represent changes in the venture investment landscape or differences in the availability of investment data.

Table H1 displays the distribution of investor types at different status levels in 2005 and 2017 (the first and final years of Sample II). We restricted this sample to the top 2,000 investors in each year for simplicity and to best focus on investors of particular interest to entrepreneurs.

Striking is that the relative portion of VC firms in the top status levels noticeably declines between 2005 and 2017. In 2005, VC firms comprised 82% of the highest status tier (the top 50 investors), with the rest of the tier comprised of investment banks (10%), corporate venture capital (4%), and private equity firms (4%). In contrast in 2017, traditional VCs only comprised 32%, though micro VCs comprised an additional 24%. Also noteworthy is the rise of accelerators as an investment type.  By 2017 we see that 30% of the top tier investors were accelerators. Overall, we see this rise of accelerators as high-status investors as the result of the status-enhancing effects we observe in our results across Samples I and II. We also see some shift in who are top tier investors. In 2005, some of the highest status investors include Bessemer Venture Partners, Kleiner Perkins Caufield & Byers, and Sequoia. While many of these remain in the top status tier by 2017, we also see newer entrants like Andreessen Horowitz, First Round Capital, and Founders Fund.

While beyond the scope of this paper, we believe exploring the dynamics and causes of these shifts more granularly is a promising avenue for future research. We believe these results speak to the value in using an affiliation-based status measure that captures the longitudinal shifts across different investor types and allows for greater comparison across types of investors.

# TABLE H1. INVESTOR TYPE BREAKDOWN BY STATUS LEVELS OVER TIME

| | 2005 Top 2,000 Investors | | | | 2017 Top 2,000 Investors | | | |
|---|---|---|---|---|---|---|---|---|
| Investor Type | Status = 0 | 0<Status<1 | 1≤Status<2 | 2≤Status<3 | Status = 0 | 0 < Status<1 | 1≤Status<2 | 2≤Status<3 |
| *Ranked orderings* | *1000 and higher* | *999 to 251* | *250 to 51* | *50 to 1* | *1000 and higher* | *999 to 251* | *250 to 51* | *50 to 1* |
| Accelerator | 0.1% | 0.0% | 0.0% | 0.0% | 4.1% | 5.9% | 9.0% | 30.0% |
| Angel (Individual Investor) | 4.3% | 0.3% | 0.0% | 0.0% | 30.5% | 16.8% | 9.5% | 4.0% |
| Angel Group | 1.7% | 0.8% | 0.0% | 0.0% | 3.5% | 2.9% | 2.5% | 2.0% |
| Co-Working Space | 0.0% | 0.1% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Corporate Venture Capital | 1.6% | 4.7% | 7.5% | 4.0% | 3.1% | 2.8% | 5.0% | 2.0% |
| Entrepreneurship Program | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% | 0.4% | 0.5% | 0.0% |
| Family Office | 0.8% | 0.5% | 0.0% | 0.0% | 0.9% | 1.7% | 1.0% | 0.0% |
| Fund of Funds | 0.1% | 0.1% | 2.0% | 0.0% | 0.3% | 0.3% | 0.0% | 0.0% |
| Government Office | 1.2% | 1.1% | 2.0% | 0.0% | 0.9% | 0.7% | 0.0% | 0.0% |
| Hedge Fund | 0.7% | 0.8% | 0.5% | 0.0% | 0.9% | 0.4% | 0.5% | 0.0% |
| Incubator | 0.2% | 0.4% | 1.0% | 0.0% | 0.6% | 0.7% | 0.5% | 0.0% |
| Investment Bank | 2.0% | 4.8% | 4.5% | 10.0% | 0.3% | 0.9% | 1.5% | 0.0% |
| Investment Partner | 1.3% | 0.7% | 0.0% | 0.0% | 7.7% | 10.4% | 5.5% | 6.0% |
| Micro VC | 2.9% | 2.4% | 0.5% | 0.0% | 7.3% | 10.7% | 20.0% | 24.0% |
| Private Equity Firm | 12.2% | 8.9% | 14.0% | 4.0% | 2.2% | 2.9% | 1.0% | 0.0% |
| Secondary Purchaser | 0.1% | 0.0% | 0.0% | 0.0% | 0.2% | 0.0% | 0.0% | 0.0% |
| Syndicate | 0.0% | 0.0% | 0.0% | 0.0% | 0.1% | 0.0% | 0.0% | 0.0% |
| University Program | 0.3% | 0.5% | 0.5% | 0.0% | 0.2% | 0.1% | 0.0% | 0.0% |
| Venture Capital | 70.3% | 73.3% | 67.5% | 82.0% | 37.1% | 41.9% | 43.5% | 32.0% |
| Venture Debt | 0.2% | 0.5% | 0.0% | 0.0% | 0.1% | 0.4% | 0.0% | 0.0% |

## APPENDIX I – EXAMINING KEY EMPIRICAL CHOICES IN SAMPLE I

We ran a number of additional analyses to explore the impact of key empirical choices in Sample I. For our status measure, we tested sensitivity by using alternative inflection points and measures. Defining the inflection points as occurring at rankings 26 and 100 in a given year yielded estimates ranging from 12% lower to 4% higher relative to results shown in Table 10, while defining the inflection points as rankings 101 and 500 yielded coefficients 2-6% higher than those in Table 10 ($p<0.07$ in all cases). Estimates using raw eigenvector centrality were more mixed ($p=0.282$ for OLS; $p=0.15$ for IPTW; $p=0.078$ for PSM, all with positive coefficients but lower $R^2$ values). In aggregate, these results support both the improved predictive power of our status transformation and also that our results are broadly robust to alternative measurement specifications.

We also probed various sampling decisions. Coefficients were higher and had lower p values when we included the five additional startups for whom individual investors were not disclosed. When including these startups, we assumed the undisclosed investors were of low status (status = 0) given the incentives for startups to disclose higher-status investors. We also explored including the two startups that were listed as primarily raising future funds via grants. This also yielded broadly similar results with higher coefficients and lower p values. These estimates suggest our estimates might be conservative if funding is defined more broadly.

Additionally, we explored the extent to which our effects might be primarily driven by accelerators helping their top graduates raise from very high-status investors – something they may be incentivized to do given power laws in venture finance (McDonald et al., 2017; Malby, 2022). We first examined the distributions of investor status for both the accelerator and non-accelerator startups. We found that the accelerator-backed startups consistently had investors with higher status at equivalent percentiles in their distributions. We also ran models omitting the top 10% of startups amongst each of the treatment groups in terms of the status of their investors (top 10% by outcome of focal accelerator participants, top 10% of other accelerator participants, and top 10% of non-accelerator startups). Estimates using the PSM model which helps account for sample selection bias were relatively consistent ($\beta=0.473$, $p=0.052$), but those using the OLS and IPTW models were less conclusive ($\beta=0.329$, $p=0.212$ for OLS; $\beta=0.398$, $p=0.155$ for IPTW). Interestingly, we did not observe effects for *other accelerator participant* in these models. Together, we interpret these results as indicating there may be heterogeneity in the extent to which accelerators primarily aid average versus top participants. Further research, though, is needed, perhaps using other accelerators or larger samples.

Finally, we calculated propensity score estimates using the *teffects psmatch* command in Stata, following the approach suggested by Abadie and Imbens (2016) which accounts for the estimated nature of the propensity scores in calculating standard errors. This yielded an average accelerator treatment effect of 0.810 ($p=0.013$), suggesting our estimates in Table 10 may be conservative.

# APPENDIX J – EXAMINING FOUNDER DATA AVAILABILITY IN SAMPLE II AND CALCULATING FOUNDER SURVEY WEIGHTS

We also further examined the availability of founder data for Sample II. Crunchbase listed founders' LinkedIn profiles for 3,702 (37%) of the 9,966 startups in Sample II. We examined the extent to which there might be systematic differences across startups where founder information is or is not available.

Table J1 reports the means for the subsets of the sample where founder information is missing and is known, as well as t tests comparing these means. We show these differences for the outcome of investor status, all round controls and internet archive crawls, all of which are observable across the full sample. Many of the t-tests have low p-values, though the magnitude of many of differences is often relatively slight. Perhaps the most noticeable difference is that the subsample where founder information is available has a higher percentage of accelerator participation (24.6% vs. 9.5%).

Table J2 reports the logit model used to calculate the survey-based weights utilized in our main regressions in Tables 12 and 13 to adjust for observable differences around which startups are likely to have founder information available. For interpretability, we also report a linear probability model (though this was not used to calculate the weights). We used the logit model to estimate the probability that founder information is available for a given startup based on observables. We then calculated the survey-based weights as the inverse of this probability. Using survey-based weights gives additional weight to startups where we were less likely to have founder information, thus assuming that these startups are particularly representative of other startups dropped in the narrowing of the analysis to where founder information is available.

## TABLE J1. COMPARING STARTUPS WHERE FOUNDER DATA IS / IS NOT AVAILABLE (SAMPLE II)

| Measure | Mean (Founder Information Missing) | Mean (Founder Information Available) | Difference | T Statistic | P Value |
|---|---|---|---|---|---|
| Investor Status | 0.969 | 1.037 | 0.068 | -2.988 | 0.003 |
| Accelerator Participation | 0.095 | 0.246 | 0.151 | -18.921 | 0.000 |
| Venture Age | 2.335 | 2.191 | -0.144 | 6.105 | 0.000 |
| Current Round Amount (M$; L) | 1.299 | 0.998 | -0.301 | 22.432 | 0.000 |
| Percentage VCs in Round | 0.664 | 0.611 | -0.053 | 6.648 | 0.000 |
| Convertible Note | 0.020 | 0.020 | 0.000 | 0.153 | 0.879 |
| Previously Raised Capital (M$; L) | 0.008 | 0.010 | 0.003 | -3.647 | 0.000 |
| State 3-Year VC Rounds | 5.506 | 5.901 | 0.395 | -11.321 | 0.000 |
| Internet Archive Crawls (-436 days; L) | 0.502 | 0.433 | -0.070 | 4.249 | 0.000 |
| *Sample Size* | *N = 6,264* | *N = 3,702* | | | |

## TABLE J2. PREDICTING LIKELIHOOD OF FOUNDER INFORMATION BEING AVAILABLE (SAMPLE II)

| | Logit Estimate of Founder Information Available (Used to Calculate Survey-Based Weights) | | Linear Probability Model of Founder Information Available (Presented for Clarity around Effect Sizes) | |
|---|---|---|---|---|
| 500 Startups | 0.488 | [0.001] | 0.124 | [0.000] |
| | (0.148) | | (0.005) | |
| Alchemist Accelerator | 1.093 | [0.005] | 0.249 | [0.000] |
| | (0.393) | | (0.008) | |
| AngelPad | 0.839 | [0.017] | 0.203 | [0.000] |
| | (0.350) | | (0.006) | |
| Boost VC | 0.888 | [0.055] | 0.209 | [0.000] |
| | (0.463) | | (0.005) | |
| DreamIt | 0.949 | [0.015] | 0.208 | [0.000] |
| | (0.392) | | (0.009) | |
| MassChallenge | 0.815 | [0.018] | 0.192 | [0.000] |
| | (0.344) | | (0.013) | |
| Mucker Capital | -0.010 | [0.982] | -0.005 | [0.821] |
| | (0.422) | | (0.020) | |
| Plug and Play | 0.590 | [0.032] | 0.142 | [0.000] |
| | (0.275) | | (0.004) | |
| StartX | 0.607 | [0.098] | 0.140 | [0.000] |
| | (0.367) | | (0.010) | |
| Techstars New York | 1.025 | [0.001] | 0.242 | [0.000] |
| | (0.319) | | (0.018) | |
| Techstars Boston | 1.283 | [0.002] | 0.284 | [0.000] |
| | (0.419) | | (0.023) | |
| Techstars Boulder | 0.444 | [0.195] | 0.090 | [0.000] |
| | (0.343) | | (0.010) | |
| Techstars Seattle | 0.893 | [0.055] | 0.213 | [0.000] |
| | (0.465) | | (0.010) | |
| Techstars (Unclassified) | 0.714 | [0.037] | 0.174 | [0.000] |
| | (0.343) | | (0.023) | |
| Y Combinator | 0.830 | [0.000] | 0.186 | [0.000] |
| | (0.130) | | (0.013) | |
| Internet Archive Crawls (-436 days; L) | -0.024 | [0.515] | -0.001 | [0.645] |
| | (0.037) | | (0.003) | |
| Venture Age | -0.063 | [0.006] | -0.011 | [0.000] |
| | (0.023) | | (0.003) | |
| Current Round Amount (M$; L) | -0.517 | [0.000] | -0.088 | [0.000] |
| | (0.040) | | (0.005) | |
| Previously Raised Capital (M$; L) | -4.483 | [0.000] | -0.910 | [0.001] |
| | (0.761) | | (0.265) | |
| Convertible Note | -1.039 | [0.000] | -0.213 | [0.000] |
| | (0.169) | | (0.022) | |
| Percentage VCs in Round | 0.094 | [0.137] | 0.015 | [0.028] |
| | (0.063) | | (0.007) | |
| State 3-Year VC Rounds (L) | -0.147 | [0.333] | -0.003 | [0.744] |
| | (0.151) | | (0.008) | |
| Tech stack FE | Yes | | Yes | |
| Round Year FE | Yes | | Yes | |
| State FE | Yes | | Yes | |
| Category Vector | Yes | | Yes | |