Analysis of Founder Background as a Predictor for Start-up Success in Achieving Successive Fundraising Rounds

by

Dolan Dworak

Thesis Advisor: Jerry Davis

A thesis proposal submitted in fulfillment of the requirements of the Michigan Ross Senior Thesis Seminar (BA 480), April 22nd, 2022.

Abstract

The culture of Silicon Valley has created some of the most valuable companies in the world. Successful start-ups build on these companies' innovations, becoming large tech firms themselves. This paper first explores start-up history and contextual reasons for why this might be the case. We then attempt to measure the effect of working in what we define as “Big Tech” before forming a start-up. To do so, we use a series of logistic regression and multivariate logistic regression models based on firm, founder, and funding round data from the CrunchBase database. We show that working in Big Tech leads to more successful outcomes and fewer negative outcomes in the likelihood of raising venture capital but has limited to no effects beyond the funding pipeline.

Introduction: Defining Start-ups

In academic literature, the concept of a start-up is not uniformly described. The term arose in the 1980s, largely in tandem with the increased usage of words like "semiconductor, minicomputer, and venture capital" (Cockayne 2019).

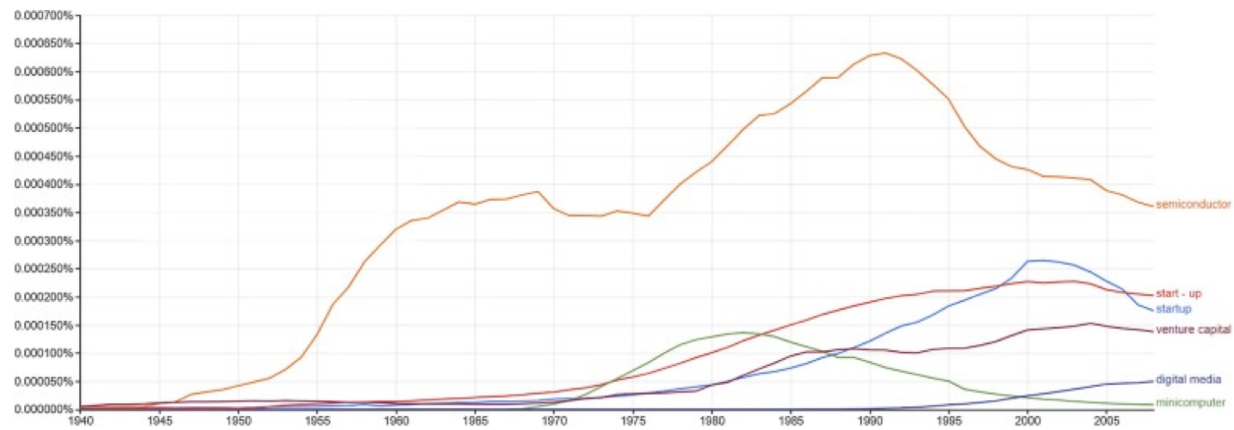

Figure 1: Frequency of Relevant Word Usage Over Time (Google NGram & Cockayne 2019)

Here, through the use of Google's Ngram Corpus, we see the usage of these terms over time (Cockayne 2019). The rise begins near 1980, partially because the Bayh-Dole Act allowed researchers and universities the right to claim ownership of intellectual property whereas it had previously gone to the government (Engel 2015). Not surprisingly, this term was formalized in California's "Silicon Valley," which was the birthplace for early internet pioneers such as Intel (founded 1968) and Cisco (founded 1984). In her seminal book, "Regional Advantage," Annalee Saxenian (1996) contended that it's the unspoken norms that allow for fluid job mobility, the sharing of ideas, and the California legislature's non-enforcement of non-compete agreements that made Silicon Valley a "start-up" hotspot. To this point, access to capital is another touchpoint for the Valley's success (Engel 2015).

To attain start-up classification, a firm must be referred to "as a *kind of firm*, defined sectorally or financially, and/or [conduct] a *kind of work*, defined through adjectives like horizontal and informal. [Other] qualities like firm size and stage, growth, and absence of a finished product … are notably absent from academic discussion" (Cockayne 2019). More specifically a start-up is a firm performing "creative destruction [in] a resilient innovation habitat," within which there is "a complex ecosystem of relationships among entrepreneurs, researchers, venture capitalists, [and other stakeholders]" (Henton and Held 2013). While start-ups are not relegated to software (nor the technology sector broadly), many start-ups rely heavily on technological innovation.

Characteristics of Founded Start-ups

In determining how start-ups are founded, it is critical to understand the lifecycle of such technological innovation. This is an idea first identified by Joseph Schumpeter in the early 1930's. In *The Theory of Economic Development*, John Elliot says "entrepreneurs … make innovative investments embodying new technologies … [and] if these innovative investments are successful, imitators follow … and the economy embarks on a dramatic upward surge [into] prosperity" (Schumpeter 1934). For clarity, this is important as it formalizes an otherwise amorphous relationship between large industry incumbents and entrants relying on their platforms and technology.

Although technology S-Curves are familiar, software companies are uniquely positioned within them. Apple's App Store generated over $643 BN in value for developers, and it was the birthplace of the newest crop of unicorns: the likes of DoorDash, AirBnb, and Uber (Apple Newsroom 2021). Case studies like the failed grocery delivery service Webvan affirm the difficulty of achieving success without the reliance on larger incumbents (Deighton 1999). It is impossible to imagine what a firm like AirBnB would look like without the existence of Apple's App Store, Google's mapping service, Amazon's cloud architecture, or Facebook's React Native development framework: the four of which it used throughout its infancy (Peal 2018). Similarly, Shopify, which emerged as a competitor to Amazon, built its cloud infrastructure on Amazon Web Service. In 2018, it began to transition to Google's cloud services, after already surpassing $10BN in value (Neufeld 2018). To further emphasize this idea, Shopify itself is the home to upstart companies like Warby Parker and AllBirds, which have respectively IPO'd to $4 and $5BN market capitalizations (Buckland 2021).

Determinants for Start-up Success

There are many factors that determine the success of start-ups and the entrepreneurs behind them. Legislative factors like the Bayh-Dole Act, the JOBS Act, and Supreme Court rulings directly affect aspects of start-up formation. However, these apply somewhat uniformly to all American start-ups. The ability to commercialize research or the deregulation of accredited investors does not necessarily explain how *specific* start-ups sustain a competitive advantage or invent new technologies. Rather, the literature identifies two main touchpoints behind entrepreneurs' success: a "consistency with innovation and continuous flow of funds" (Okrah et al. 2018)[1]. In later sections, we identify specifically how these factors must be achieved simultaneously and how that leads to incumbent firms losing market share to entrants. Since the literature identifies access to capital as a critical determinant, we explore the criteria venture capitalists use to allocate funding.

Determinants for Attaining Venture Funding

Venture capital is a term that describes how a certain subset of firms receive funding for the purpose of commercializing new technologies. While it is a small part of the greater research and development sector, it is for the specific purpose of investing "in a company … until it reaches a sufficient size and credibility" (Zider 1998). Due to regulatory restrictions on other institutional financiers, venture capitalists characteristically have higher risk profiles and expect greater returns over the venture firm's lifetime.

1

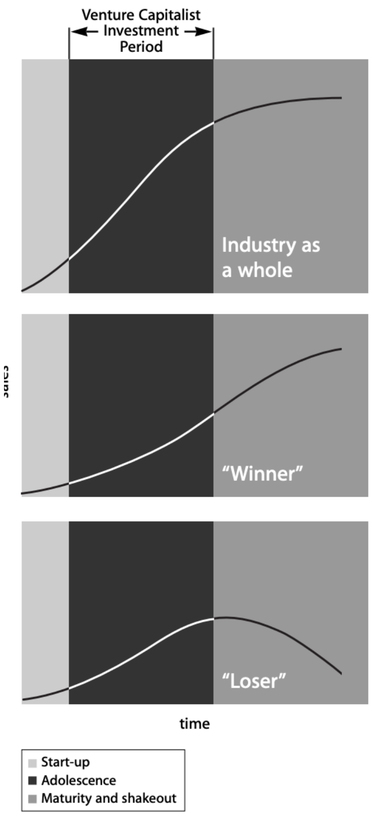

Figure 2: Allocation of Venture Funding at the Adolescent Stage (Zider 1998)

In the previous figure, we see a simplification of where "80% of the money" invested goes and how a start-ups adolescence is critical to its potential success (Zider 1998). It is therefore an obligation of the venture capitalists, when determining who gets funding, to assess the start-up firm, the founder, and market potential.

The criteria that venture capitalists use can be broken into subjective and objective decision factors. Subjective factors include characteristics about the entrepreneur, specifically, whereas objective factors detail the market offerings and investor requirements. Of the entrepreneur, the literature identifies 20 criteria, which include the "referral by a trustworthy source, ability to sustain intense effort, [and] familiarity with [the] technology" (Bachher & Guild 1996). These can be classified as subjective, because the venture capitalist must critically assess the proposal based on their 'subjective' opinions about the founder. Conversely, objective factors (24 total) like "there is a strong dominant competitor with a large market share" or the amount of required funding exceeds portfolio requirements, are more easily verifiable. Largely, venture capitalists are "primarily interested in understanding the characteristics of the entrepreneur(s)" (Baccher et. al 1999). In such cases, it is important to consider founder background as a predictor for the ability to achieve success in raising money.

What is Big Tech?

Big Tech has been colloquially defined as FAANG, including Facebook (now Meta), Apple, Amazon, Netflix, and Google. This is indeed a shorthand, as it conspicuously omits companies of similar size like Microsoft. Big Tech firms are characteristically software companies with multi-billion dollar market capitalizations. Other core characteristics include relying heavily on user data to improve their products and prioritizing growth over profitability (Beard 2021). When defining "Big Tech" it is important to understand how the introduction of new technology leads to dominant firms losing market share to entrants. With an example from the FAANG framework, we regard how Netflix 'disrupted' Blockbuster through the advent of CD's and streaming; now just as Netflix was once the entrant, start-ups are disrupting Netflix (whether it be the defunct Quibi or ByteDance's thriving TikTok).

Since the broadly defined technology sector is so large, it can uniquely accomodate large firms and a growing amount of new firms. This means even as the dominant platform changes (like in the case of mobile usage surpassing desktop), firms like Intel, Dell, and HP are doing well, especially compared to firms in other mature industries. Intel, founded in the late 1960's, is down from its peak in early 2000's, but it is still humming along to the tune of $200BN market capitalization; Intel's respective 'disruptor,' NVIDIA, is worth over $700BN. PayPal ($134.88BN), eBay ($33.88BN), and Zynga ($10.44BN) are other examples of this concept, with their competitors Square ($76.08BN), Shopify ($86.24BN), and Roblox ($26.81BN). Consider these against the hemorrhaging of value in the automotive, retail, and manufacturing industries.

Motivating Big Tech Case Study: Neeva

In 2003, Sridhar Ramaswamy was recruited to Google to work on its nascent ads platform, AdWords. As money began to flow through the system, Ramaswamy and his team began to race against the speed of Google's other products, keeping ads as fast and relevant as the search queries that inspired them. Eventually, the AdWords team "built the world's largest real-time auction system, [even] bigger than [the] NASDAQ," with many more innovations to boot (Efrati 2015). At the end of his 15 year tenure, Ramaswamy was reporting to Google's CEO and was VP of the $115BN AdWords division, thereby responsible for over 80% of Google's total revenue (Google 2020).

Yet, Sridhar became disillusioned with the impact of the team he helped create, noting that "the size of the Google Ads team … dwarfs that of the search team" (Kantrowitz 2021). To him, it was a signal that Google came to serve those with the deepest pockets at the expense of its users. This manifested itself when it was discovered Google ran ads on YouTube content that featured alleged child abuse. Sridhar's "last straw" was seeing advertisements of blue-chip companies served by his team's software on these videos (Wakabayasi 2020). After leaving Google, he admitted "an ad-supported model had limitations," and saw a future in subscription based search.

After leaving Google, two years later, he was the founder of an ads-free, subscription based search engine, Neeva, a product seemingly in direct competition with Google (Wakabayashi 2020). This is to say throughout the 20th century, there has been a matter of fact relationship between intellectual property ownership and the relationship between employers and their firms. Yet in the case of Neeva, an executive had the chance to see the company's inner workings (and shortcomings) and started a competing firm. Even in the San Francisco Bay Area, which has long been the birthplace of huge, innovative software companies, this is an eyebrow raising outcome.

As it stands, the Supreme Court ruled against a software company alleging Hawaiian Airlines infringed upon its patent for an algorithm to convert customers through a loyalty rewards program (*Loyalty v. American* 2014). The dominant reasoning was that patents need to be more than the "abstract idea" of an algorithm (and software by logical extension). If Sridhar had a similar insight about electric-vehicles' battery technology, having been an engineer at Ford, the firm would assuredly assume the right to an equivalent invention via the Shop-Right Doctrine:

> "Shop rights arise when an employee, 'during his hours of employment, [and] working with his [employer's] materials and appliances, conceives and perfects an invention for which he obtains a patent.' The courts have held that in such circumstances, the employee by force of law must give his employer a nonexclusive, royalty-free right to practice the invention" (Simmons 2012).

For Neeva, there is no device or code-base to which Google can claim ownership. This marks a shift between the obligations of employees and how innovation can arise from participation with incumbent firms.

The Neeva case study is an illustrative example within a far larger ecosystem. There are dozens of moments within SIlicon Valley lore of back-of-the-napkin moments between two engineers, hypothesizing on their future endeavors. Indeed, even Steve Jobs, the archetype entrepreneur, worked at Hewlett-Packard as an intern and later at Atari, before starting Apple—two companies of significant size and influence at the time. If software companies can no longer protect their *ideas* with the same vigor as traditional intellectual property, there must be an advancement of how working at a Big Tech firm qualifies as an "advantage" and how it applies to entrant firms.

Rationale for Method: What makes you a Big Tech alumnus?

For the purposes of this research, we will specify what makes someone an alumnus of Big Tech. Similar to how someone becomes an alumnus of a university, a Big Tech alumnus is someone who has worked at a firm before going on to start their own venture. Conversely, not working in Big Tech will classify founders as having created a "non-alumni" start-up. While we hoped to further qualify "non-alumni" between law, consulting, the military, or other prestigious industries, there were too few entrepreneurs within our sampling criteria. Within the methodology section, we flesh out this definition.

This alumni concept is not new to the software industry. Google employees have particularly leaned into the idea, referring to themselves as Xooglers (ex-Googlers). There is even a syndicate of former employees organized under the Xoogler name who host demodays, pitch competitions, and networking events with top venture capital firms (like Sequoia Capital, Greylock Partners, and Kleiner Perkins). These VCs affirm our previous venture criteria, noting “two things [they’re] looking for are quality of the founder and the market” (Nicas 2016). To Xooglers, the transition between working at Google to starting a company is not nuanced: founders “must have something great if [they’re] willing to leave (Google)” and they already “have a potential hiring pool to draw from—[their] former colleagues” (Nicas 2016). For reference, Google alumni started the most companies in our sampling frame, but it is only one company in our set of Big Tech firms.

Foundation of our Hypothesis

Our hypothesis relies on understanding Complex Network Theory (CNT). The following section will explain the presumed advantages working in Big Tech gives founders. To begin, CNT draws a distinction between innovation clusters and industrial clusters. Hollywood and Wall Street are two examples of industrial clusters, focusing on “incremental innovations that reinforce … excellence and competitiveness in a specific … domain” (Ferrary and Granovetter 2009). On the other hand, innovation clusters create value by reimagining the state of the world. This is most commonly achieved by the formation of start-ups, as evidenced by the aforementioned disruption theory of technological innovation.

Regard our Shopify example mentioned earlier. We note that Shopify was built on the cloud infrastructure of Amazon, its competitor. While this may seem surprising, it is within the understanding of CNT’s innovation clusters:

> “There is a virtuous self-reinforcing dynamic of creation of high-tech start-ups. Several large firms that currently contribute to the complex network of Silicon Valley have previously been high-tech start-ups … and have been developed with the support of other agents of the system” (Ferrary and Granovetter 2009).

In other words, the frequent interaction between companies, employees, venture capitalists, and founders contributes to CNT. By extension, the firms that participate successfully in the network will benefit from its advantages.

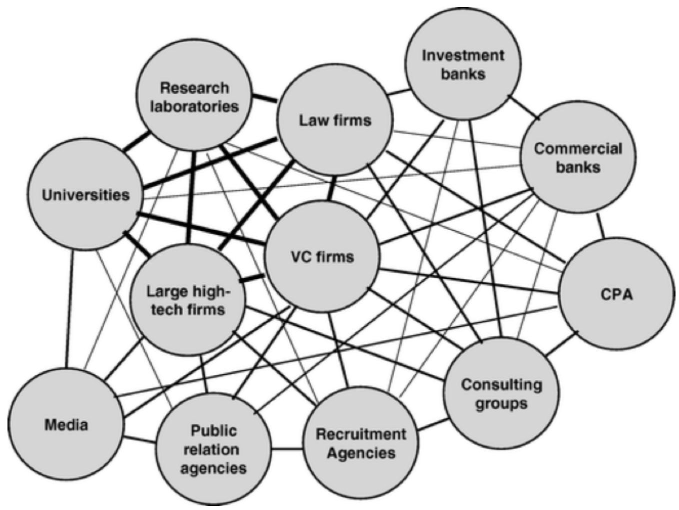

Figure 3: Critical Network Theory Actors (Ferrary and Granovetter 2009)

This graph demonstrates all of the active stakeholders under CNT and how they interact with one another (Ferrary and Granovetter 2009).

Based on the above network graph, we can specifically examine how participation with VCs can benefit start-ups. In the following graphs, we see how once a firm enters a VC's portfolio, it benefits from its interaction with the other actors. As designated by the circled "L," there is an active relationship between VCs and large incumbent firms. These firms could be successful former start-ups that the VC invested in, to whom they communicate learnings in certain markets or whether they plan on entering a new industry (Ferrary and Granovetter 2009). In the subsequent graph, we see how start-ups can benefit from the learnings of the VC's other portfolio companies. In both cases, it affirms the necessity of attaining venture capital.

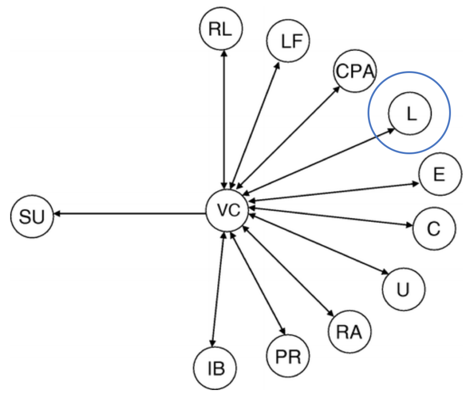

Figure 4: Relationship Between Incumbents and Venture Capitalists
(Ferrary and Granovetter 2009)

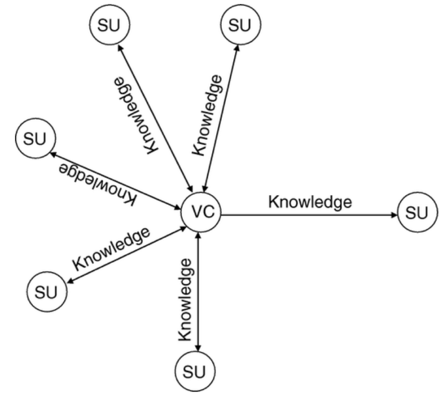

Figure 5: Relationship Between Entrant Firms and a Venture Capitalist's Portfolio
(Ferrary and Granovetter 2009)

While the literature credits venture capital to have a sizable outcome on start-up performance, it is (in part) insufficient in regards to CNT (Gompers and Lerner 1999). For instance, research focuses on the start-ups that *successfully receive funding*. Of the firms that *don't*, it is difficult to systematically measure and therefore articulate why. Generally, the start-ups that fail to raise are considered in the context of the 'Lemons Problem.' Historically, if a top venture capital firm forgoes an investment, "the information is quickly spread … and it becomes very difficult for the start-up to raise funding from other VC firms" (Ferrary and Granovetter 2009). We therefore assume that having worked in Big Tech alleviates the industry pervasive Lemons Problem, giving additional credibility to founders (deserved or otherwise).

Hypothesis and Confounding Variables

The underlying hypothesis is that working at a Big Tech firm is a significant advantage when starting a new company. The reasons for which are not limited to meeting new people, access to capital, and learning world-class engineering and process skills. Compared to non-alumni founders, we expect alumni founders to have far greater success rates, as determined by probability that a firm will survive a "round," achieve the next round, or achieve a successful exit. Other factors like amount of funding raised, acquisition price, and duration between rounds will be considered when applicable.

For the purposes of the research, we will identify five phases: 1) founding , 2) seed round, 3) early stage venture round, 4) late stage venture, and 5) exit (initial public offering or acquisition).

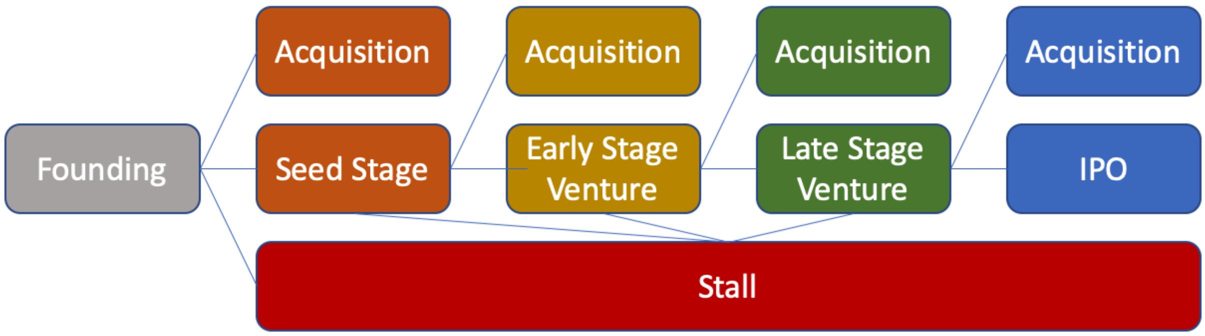

Figure 6: Funding Pipeline Considering what Firms Qualify for Each Successive Risk Set

All of which are defined in the methodology. In each, there is a significant amount of literature on which to build. As mentioned above, there was little insight as to how leaving a Big Tech firm would particularly affect outcomes. Throughout the start-up's lifecycle, we can begin to analyze Big Tech's halo effect under CNT. If alumni founders have higher odds of achieving each successive round when controlling for other factors, it would suggest there is some outsize influence relating to working in Big Tech as a founder. From the literature review above, we do not explore if this is a result of skills learned or signal (as it relates to the Lemons Problem in VC).

Methodology

` In our methodology, we incorporate the stages mentioned above. At each stage, our risk set includes the firms that achieved the *previous* stage in *this stage's* model. For clarity this means that our *seed* model will include the entire universe of firms, whereas the *early stage venture* model will only include the firms that received seed funding. The two models we run are a series of logistic regression models and multivariate logistic regression models. The latter of which is to accommodate for the "hump" identified in Figure 2. If a firm does not canonically "fail" nor does it achieve the *next* stage, we classify that as a stall. A stalled firm may be a "success" depending on the expected returns of both the founder and venture capitalist, but, as evidenced by the data, it is no longer on a viable path to acquisition or initial public offering (IPO).

Sources of Data

Within the literature concerning start-up outcomes, there are significant gaps in available data. In part, this is because the definition of a start-up is contingent upon the scope of the research. For instance, a large firm like Tesla (competing in the automotive industry) fits the definition of start-up for some literature but not others. In our case, we define a start-up as a firm within our sampling frame below. This, however, limits our available data. Small side-projects may scale into start-ups before securing the legally identifying certifications such as incorporation documents or employer identification numbers. Conversely a start-up may conclude operations after receiving such documentation but before raising capital or leaving the 'stealth' growth phase (MIT Enterprise Forum). In both cases, start-ups may not hire employees (or do so in an ad hoc manner), precluding us from capturing relevant Census data. To combat this, researchers often partner with venture capital firms, relying on data used internally to make financing decisions. However, with the goal of examining founder backgrounds across industries and funding rounds, this makes it difficult to perform longitudinal analysis.

CrunchBase Data Set

The methodology of this research is built upon the CrunchBase database. CrunchBase originated as a project of former parent organization TechCrunch to catalog the start-ups mentioned throughout the media outlet's articles. It has since become a valuable prospecting tool for investors and start-ups alike, describing itself as a "LinkedIn for company" data (Lunden 2019). CrunchBase originates data from its "CrunchBase Venture Program," a partnership that began with 11 venture capital firms agreeing to contribute data regarding "funding updates, staffing changes, product launches, and acquisitions" (Lunden 2019). This data is supplemented by moderated user submissions totalling over 1.6 million data points across 120,000 users. Such data points include the number of employees, relevant dates, and participating venture firms. Today, the partnership has grown to over 3,500 firms self-reporting portfolio data and CrunchBase's internal artificial intelligence and machine learning teams.

However, CrunchBase is still a profit-seeking organization and its database is curated towards the goals of investors, founders, and business development prospectors (as opposed to academic researchers). Since the data collection algorithms are not public, it is difficult to factor these into the regression models. For instance, founders' education history may list only a graduate school or be missing entirely. Upon inspection, many founders did complete undergraduate education, but others forewent undergraduate degrees for paths including military training. Others began their venture during undergrad. Additionally, since data is a combination of self-reported and moderated, it is difficult to differentiate factors like "Primary Organization" against "Current Organization" when they vary. This introduces ambiguity into which founders we should consider in our 'universe' of firms, as it is unclear when a founder left their organization (if at all). Over thousands of firms with tens of thousands of founders it is difficult to systematically correct these shortcomings.

Of the available data, biographical characteristics that relate to the firm (IPO status, acquisition status, and geographic location) are easily and repeatedly verifiable. By extension, certain founder characteristics (gender, number of co-founders, number of exits, and number of founded organizations) are also helpful. Finally, data relating to firms' funding rounds is available. As a result, we will be clear with the assumptions made in the models as well as the limiting factors in performing regression analysis.

Available Variables

Firm Biographical Data:

Variable Name	Variable Description	Type of Variable
Organization Name	Name of the firm (ex: "Contix")	String
Operating Status	Reported operating status (Active, Closed)	Factor

Industries	Software sub industry (“Analytics, Big Data”)	Vector of Strings
Funding Status	Current (last) *stage* of funding round	Factor
Number of Funding Rounds	Number of rounds	Integer
Price	Dollar amount of acquisition	Integer
Acquired by	Name of the acquiring firm	String
Total Funding Amount	Total amount of funding across all stages	Integer
Number of Founders	Factored into Solo, Co-Founded, Three, Many (>3)	Factor
Founders	Names of founders	Vector of Strings
Number of Employees	Present number of employees	Integer
IPO Status	Private, Public, or Delisted	Factor
Top 5 Investors	Five largest investors (by amount raised)	Vector of Strings
Number of Investors	Number of investors across all stages	Integer
Number of Acquisitions	Number of firms *this* firm acquired	Integer
Acquisition Status	None, Made Acquisition, and Was Acquired	Factor
Date Founded	Date that *this* firm was founded	Date
Cardinal	Factorized headquarters location into regions	Factor
Cardinal*	Regions are segmented into Atlanta, Denver, the DMV, Florida, Greater East, Greater South, Greater West, Los Angeles, Midwest, [New England, Boston, Northeast], New York, Philadelphia, Phoenix, San Diego, San Francisco, Seattle, Texas, and West Coast. This represents the metropolitan area. San Jose, CA is “San Francisco” whereas Fresno, CA is “West Coast,” for example.	
VC Hub	San Francisco Bay Area, Greater Boston, or Other	Factor
IsAlum	Whether any founder is an alumnus	Factor

Funding Round Data (across all of Seed, Early Stage Venture, and Late Stage Venture):

Variable Name	Variable Description	Type of Variable
IsAlum	Whether any founder is an alumnus	Factor
Investors	Investors at *this* funding stage	Vector of Strings

Date	Date when *this* funding round was received	Date
Type	Actual type of factorized stage (ie: PreSeed, Angel, Seed)	Factor
Stage	Stage of *this* funding stage (Seed, ESV, LSV)	Factor
Amount Raised	Amount of funding raised at *this* stage	Integer
Number of Investors	Number of investors at *this* stage	Factor
Number of Investors*	We consider None, One, Few [2, 5], and Many (>5) as our factors	
Normalized Amount	Amount of funding at *this* stage normalized to [0,1]	Integer
Got*stage*	Whether the firm achieved *this* stage (Seed, ESV, LSV)	Factor

Founder Biographical Data

Variable Name	Variable Description	Type of Variable
Full Name	Founder's full name	String
Organization Name	Name of founder's founded organization	String
Number of Exits	Number of founder's exits (acquisition or IPO)	Integer
Number of Founded Organizations	Number of organizations *this* founder founded	Integer
Gender	Founder Gender	Factor

Notice we have three sets of data that we will join to from one database. Unfortunately, we do not have all of the key data points. Most of the missing variables relate to founder biographical data. As mentioned above, we cannot systematically collect data relating to the founder's education. Ideally, we would like to control for the institution, field of study, as well as level of education (between undergraduate, masters, and beyond). Another variable is age. The literature regards the average age of founders to be 41.9, with founders at age 40 being 1.3 times more likely than founders at age 25 to build a successful startup (Azoulay et al. 2018). Fortunately, we do have the number of exits and the number of founded organizations as variables. Since the dominant literature about founder age and success is observational, their conclusion is that age leads to more experience. Our two variables therefore serve as a useful proxy for age.

We are also missing nationality. Regarding country of origin, the start-up visa problem is well studied. Generally, immigrants have higher self-employment rates, founding 25% of American start-ups overall (Vandor 2021). Additionally, it is common for tech companies to have the capital to sponsor visas for employees. Within our Big Tech sampling frame, firms applied for over 40,000 H1-B visas in 2019, becoming the largest firms by proportion of total employees (U.S. Citizenship and Immigration Services). In 2022, the House of Representatives advanced the "America COMPETES Act," removing visa limits for qualified STEM immigrants and loosening visa requirements for start-ups generally. Previously, the visa requirements were strict for founders, making starting a company an additionally risky endeavor for immigrants. This would have been a compelling variable for our study.

Sampling Frame: Universe of Firms

We consider our start-up sampling frame to be firms founded in the United States between 2012 and 2014. There are a few reasons for this. Firstly, the Obama Administration "Jumpstart Our Startups" (JOBS) Act was signed into law in 2012. The JOBS Act had the goal of creating opportunities for "emerging growth companies" by limiting regulatory constraints and allowing financing from non-accredited investors (The White House Office of the Press Secretary 2012). While the consequences of the JOBS Act are still uncertain, it allowed start-ups to raise large sums of money (not exceeding $1 million) from 'angel' investors and (not exceeding $5 million) from crowdfunding (Chaplinsky et al. 2017). As a point of clarity, this 'crowdfunding' is otherwise regulated unlike unregulated (non-equity) platforms like Indiegogo or Kickstarter.

Next we limit our sampling frame to American companies self-described as in the "software" industry. This is because our literature review revealed that start-ups are built on the innovations from existing technologies (examples including Apple's App Store and Amazon's web hosting services).

Between 2012-14, mobile penetration reached a critical point where it became feasible to create mobile-first products and dominant platform economics were established. Industries within the 'software' umbrella are not limited to apps, augmented reality, billing, cloud computing, cryptocurrency, data storage, developer tools, EdTech, file sharing, image recognition, speech recognition, video conferencing, and web development. Firms are also organized across industries, such that start-ups solving issues in human resources and debt collection (for example) are considered, despite being traditionally low-tech fields. In this analysis we consider 88 software sub industries.

Sampling Frame: Big Tech and Relevant Alumni

We will now consider what firms we regard as "Big Tech." Within which we have two categories: 1) the set of legacy, large public tech companies and 2) the set of (then private or IPOing) tech companies that are, at present, large and public. These companies are as follows: AirBnb, Alibaba, Amazon, Apple, Cisco, Dell, Dropbox, eBay, Meta (Facebook), Google (Alphabet), Microsoft, Netflix, Oracle, Palantir Technologies, PayPal, Salesforce, Samsung, SAP, Splunk, Twitter, VMWare, Workday, and Yelp. We consider an "alumnus" to be someone who worked at any of these firms. Due to the previously mentioned limitations of CrunchBase, this is limited to their professional work history. This means we make the assumption that tenure at these firms was longer than that of an internship, but seniority at the firm cannot be considered. Moreover we do not limit the duration between a founder leaving and starting their venture. Since we have no way to quantify the timeline between having the vague idea for a start-up and founding it, there is no systematic way to control for its effect.

Exploration of the Data

After collecting our data, we have a universe of 6911 firms that fit our above criteria. Of which, we have 880 firms founded by our definition of an alumnus, for a proportion of 12.7%. At this point, we must consider the variation we hope to observe by classifying our universe of firms by alumni status. Of all the firms in our sample frame, 1386 were acquired and 70 underwent an IPO, representing 1456 firms

(21.06%) that had a successful exit of some kind. By proportion of exit relative to presence in our sample frame, alumni firms were overrepresented. An alumni firm had a 32.61% chance of being acquired, compared to 18.22% for non-alumni. Likewise, an alumni firm had a 1.93% chance of achieving an IPO, compared to 0.88% for non-alumni. For reference, our Big Tech sample frame acquired 136 of these ventures, the three largest being Google at 27, Microsoft at 19, and Amazon at 11 acquisitions. This pattern holds when considering failure, or the 1465 start-ups reported as having a "closed" operating status. Alumni firms have a 20.00% of failing, compared to 21.37% for non-alumni. Notice that it is more desirable to have a lower proportion in the case of failure. The following figure shows a graphical representation of these outcomes.

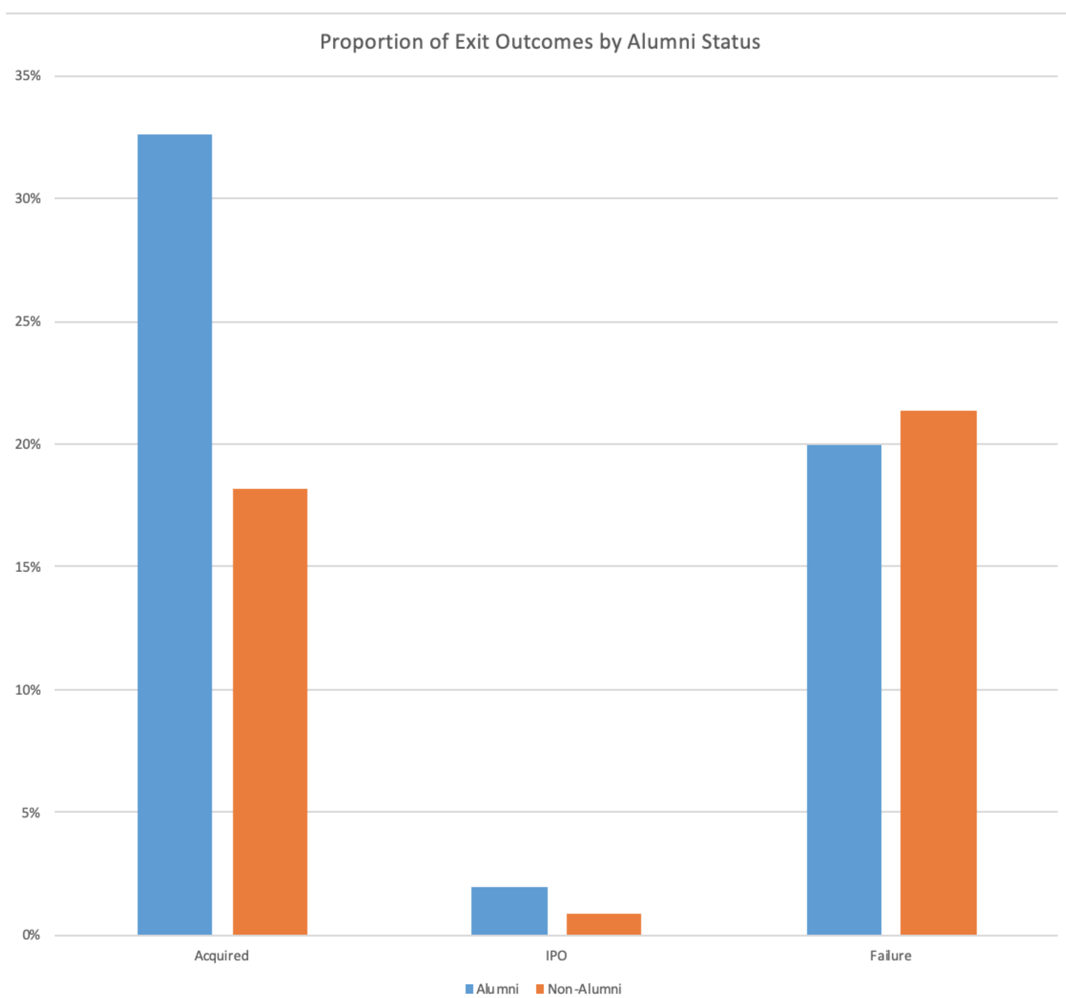

Figure 7: Proportions of Outcomes by Sampling Criteria

Although exit and failure are helpful binary metrics of success, they do not fully capture the lifecycle of a start-up. Rather, we hope to examine the start-up's outcome at each successive round of funding. Beginning at founding, we detail three funding milestones: seed, early stage venture, and late stage venture. Seed rounds include any angel, pre-seed, seed, or analogous type; early stage venture (ESV) is either Series A or Series B, and late stage venture (LSV) is any subsequent round after Series B (beginning at Series C) until IPO.

We first examine if being an alumnus has any effect on the *amount* of funding raised at each successive stage. When controlling for firm and biographical data, the presence of an alumnus founder has no apparent effect on the amount of funding raised. Even after normalizing the amount of funding between an inclusive range of [0, 1] we see no significant change. In part, we can contribute this to the skew of funding data. For each of the three funding stages we have the following respective mean and medians: seed ($1,606,889, $580,000), ESV ($13,667,224, $8,500,000), and LSV ($83,211,082, $4,700,000). Here is the magnitude of these differences.

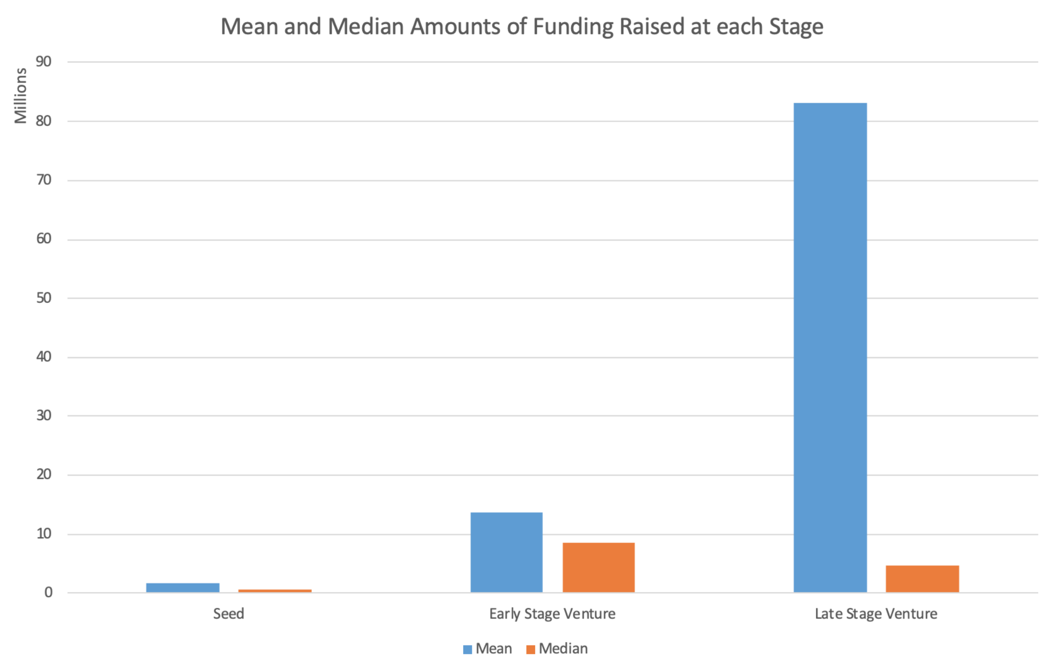

Figure 8: Skew of Funding in Each Stage of Venture Funding Pipeline

These following histograms show the distribution of funding amounts at each stage. Since our firms are strictly in the software industry, it is not clear why the variance is this high. Since we do not have access to cost structure or equity partitions, we do not further consider the amount of funding raised.

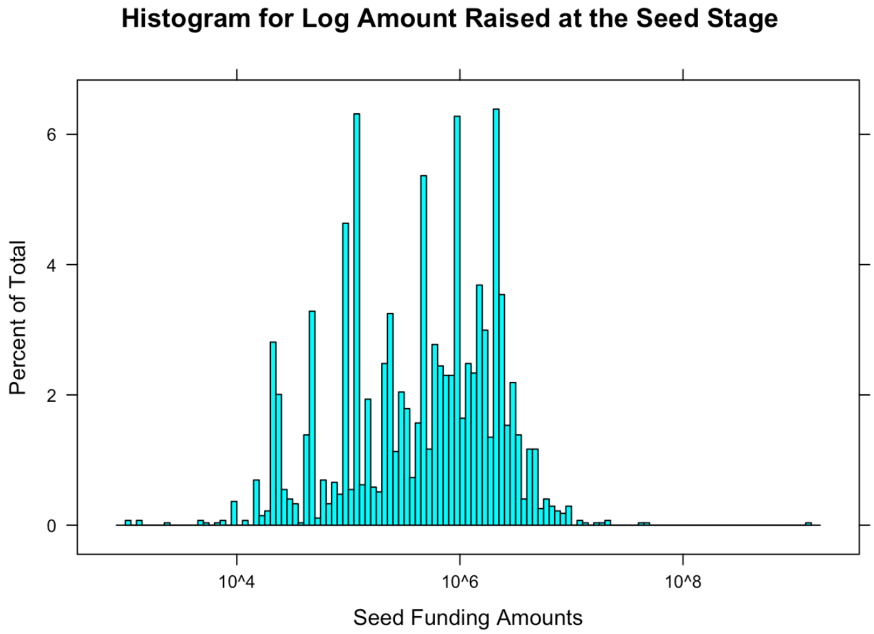

Figure 9: Skew of Funding at the Seed Stage (100 Partition Bins)

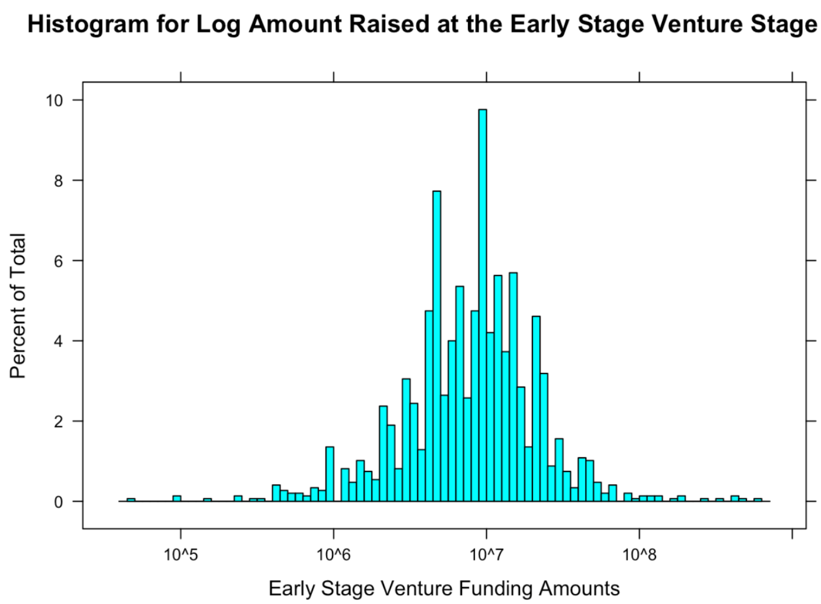

Figure 10: Skew of Funding at the Early Stage Venture Stage (100 Partition Bins)

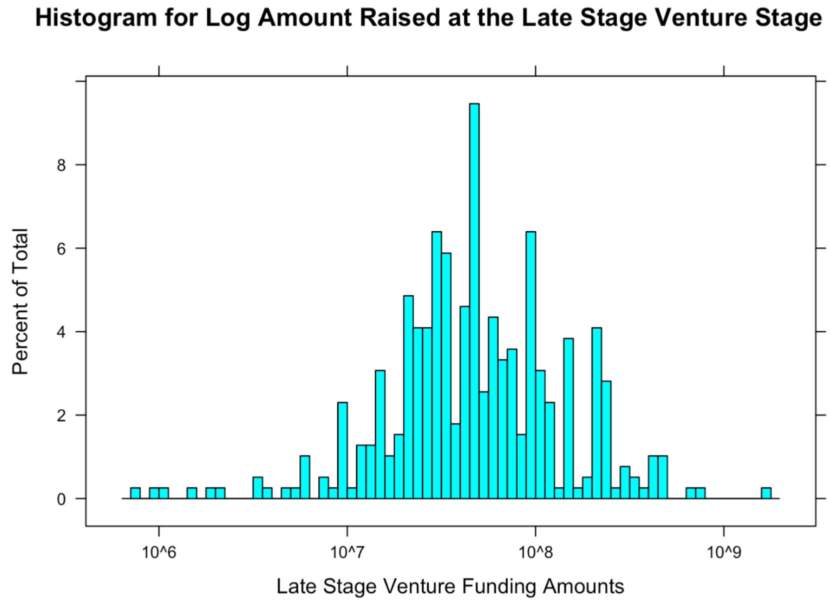

Figure 11: Skew of Funding at the Late Stage Venture Stage (100 Partition Bins)

Next we perform the same analysis on the price that acquiring firms pay for start-ups in our sampling frame. For reference, the average price for an acquisition was $7,736,106. Again, we see that the presence of an alumni founder has no significant effect on the exit purse of an acquisition. This finding, in combination with the amounts raised in each stage, is interesting when considering CNT.

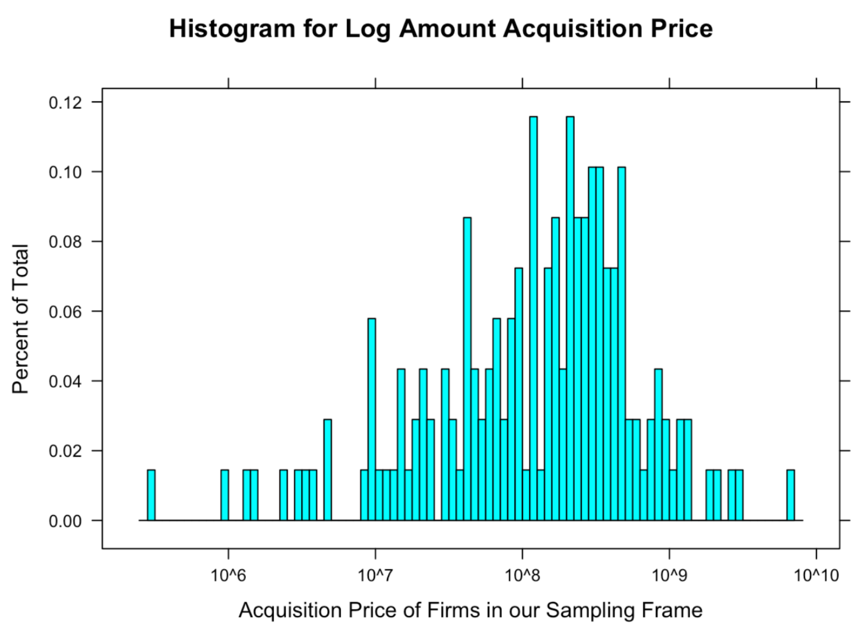

Figure 12: Skew of Acquisition Prices (100 Partition Bins)

As above, here is the histogram showing the skew of acquisition prices. From our network graphs, we remember that VCs communicate knowledge within that specific innovation cluster. Yet, the literature review revealed there are still two factors that determine venture funding: the entrepreneur and the market. Since the amount of funding raised nor the exit purse is affected by founder background, it suggests that both VCs and incumbents are assessing the start-up based on the strength of its technology and competitive advantage, rather than the founder.

Speed is also a critical factor for start-ups. The ability to raise funding more quickly allows firms to go to market sooner and use capital to grow in industries with first mover advantage. For all the firms that received seed funding, alumni did so 517.69 days after being founded on average (596.30 days for non). This pattern holds for both ESV and LSV stages. For ESV, alumni and non-alumni had respective averages of 1156.63 days and 1424.38 days and for LSV, respective averages of 2371.22 days and 2515.83 days. This indicates for each round, alumni firms were 13%, 19%, and 6% faster in raising funding. Notice this means that non-alumni took approximately 828.08 days between the Seed and ESV round, compared to only 638.94 days for alumni. This indicates non-alumni had over a six month penalty between rounds. Figure 13 shows a timeline of these differences.

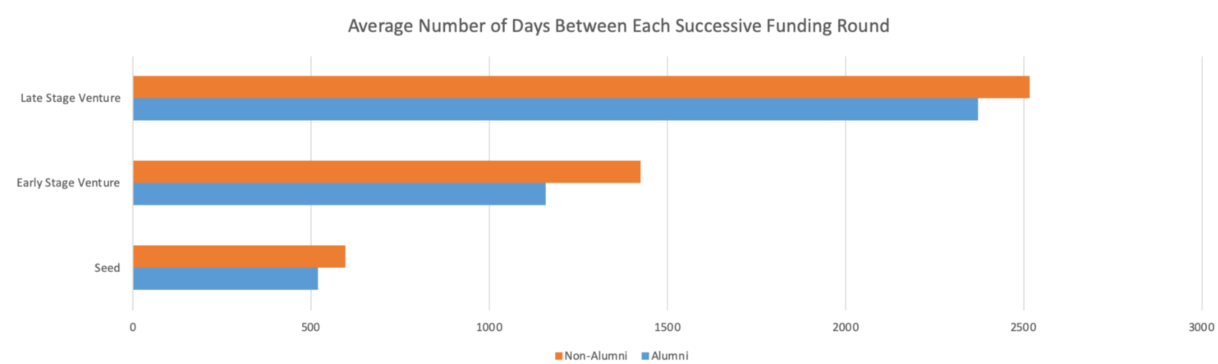

Figure 13: Timeline Between Funding Stages

<u>Approach by Successive Funding Rounds</u>

For the following sections, we will consider our seed, ESV, and LSV stages. After founding, alumni firms have a 59.09% of reaching seed (52.18% non-alumni), a 37.39% of reaching ESV (21.82% non-alumni), and a 12.16% of reaching LSV (5.12% non-alumni). This is legibly represented as Figure 14. This is in contrast to Figure 7, where we show the final outcomes. In Figure 14, we notice the proportion of alumni that stall is lower than that of non-alumni. For reference, of all American start-ups (loosely defined), between 0.25% to 2% receive any venture funding (Rose 2012).

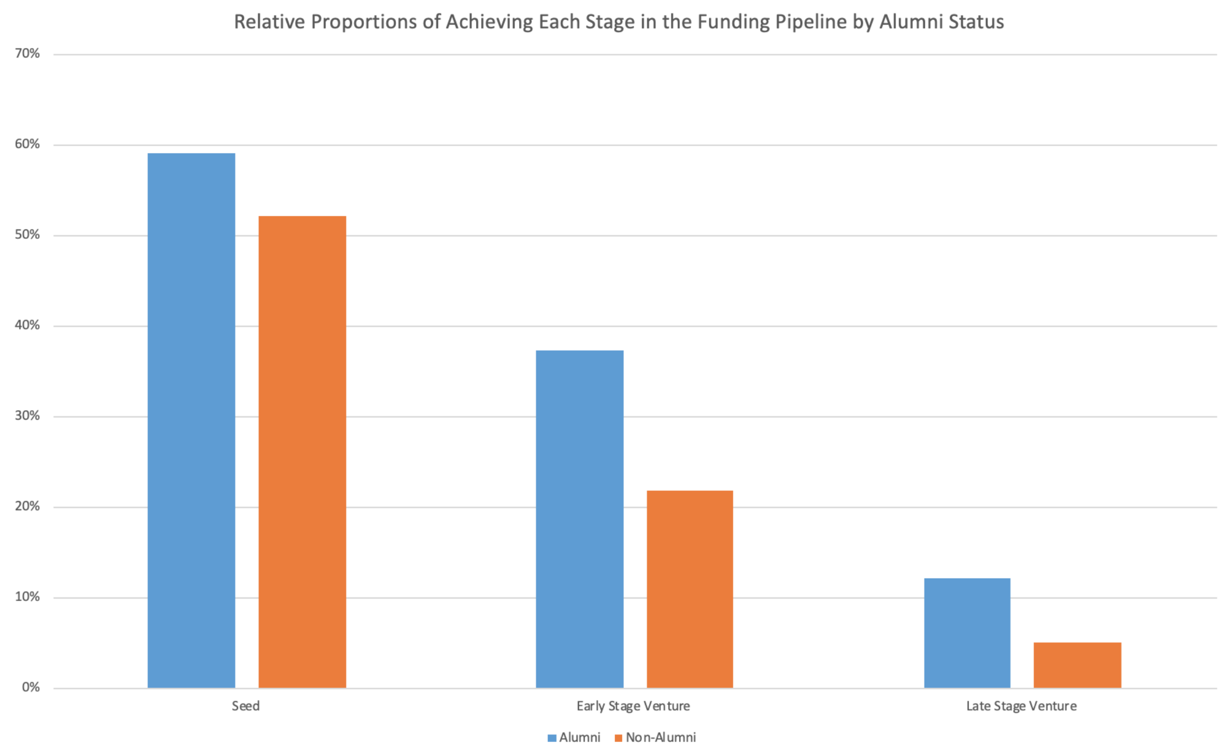

Figure 14: Proportions of Sample Frame that Achieve Each Successive Stage

Results from Data

After cleaning the data, we identified 7 factors that we will incorporate into our models. Figure 15 shows the variables, as encoded by firm, founder, and round characteristics along with their relevant descriptions. In the data we notice there are 768 firms with many (>3) founders. In our universe, the average is 2.41 founders for alumni and 2.12 founders for non-alumni. This means that alumni teams were slightly larger, consistent with our presumed advantage that working in Big Tech leads to hiring from a larger pool of talent.

Characteristic	Variable	Description
Firm	Alumni Status	Whether or not founder worked in Big Tech
Firm	Number of Founders	Number of founders at this firm
Firm	VC Hub	Whether firm is in Silicon Valley, Boston, or Other
Founder	Number of Exits	Number of exits the founder achieved
Founder	Number of Founded Organizations	Number of previously founded organizations (including this one)
Founder	Gender	Gender of the (non-)alumni founder
Round	Number of *stage* Investors	Number of Investors at the previous stage

Figure 15: Factors Considered in our Models

Logistic Regression Models

The series of logistic regression models will consider founding to each successive funding stage. In these models, we treat the Got*stage* variable as binary, where *stage* is our seed, ESV, and LSV stages. We perform two sets of regressions, "firm and founder view" and "firm, founder, and round view." The latter is to analyze how the Lemons Problem signal affects alumni status. In the firm, founder, and round view, we simply include the number of participating investors at the *previous* stage.

Logistic Regression: Firm and Founder View

At all of the seed, ESV, and LSV stages, alumni status is a significant positive predictor for the likelihood of achieving that round. Alumni status is weakest at the seed stage with a coefficient of 0.13 and is strongest at ESV with a coefficient of 0.625. The number of founders is also significant at each stage, with having *many* co-founders as having the most positive effect, while being a solo founder has a strong negative effect. There are factors that only affect the seed stage, which are gender and VC Hub. Being a male founder and being in Silicon Valley are strong positive predictors. All other factors, like number of exits or founded organizations are insignificant. This could be because we consider the alumni founder specifically, whereas their co-founders may have significant start-up experience. Since we identified the ability to draw from a strong pool of colleagues, it could explain why the number of founders is significant, whereas other founder characteristics are not.

	Dependent variable: GotSeed
IsAlum1	0.130* (0.076)
Number.of.FoundersMany	0.451*** (0.086)
Number.of.FoundersSolo	−0.754*** (0.060)
Number.of.FoundersThree	0.358*** (0.069)
Number.of.ExitsNone	0.433 (0.342)
Number.of.ExitsOne	0.583 (0.388)
Number.of.Founded.OrganizationsOne	0.014 (0.099)
Number.of.Founded.OrganizationsThree	0.093 (0.124)
Number.of.Founded.OrganizationsTwo	−0.020 (0.106)
Gender1	0.145* (0.086)
VCHubOther	0.241** (0.105)
VCHubSilicon Valley	0.408*** (0.109)
Constant	−0.642* (0.376)
Observations	6,911
Log Likelihood	−4,579.966
Akaike Inf. Crit.	9,185.932
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable: GotESV
IsAlum1	0.625*** (0.100)
Number.of.FoundersMany	0.576*** (0.107)
Number.of.FoundersSolo	−0.501*** (0.108)
Number.of.FoundersThree	0.206** (0.091)
Number.of.ExitsNone	0.458 (0.655)
Number.of.ExitsOne	0.572 (0.702)
Number.of.Founded.OrganizationsOne	0.005 (0.146)
Number.of.Founded.OrganizationsThree	0.049 (0.180)
Number.of.Founded.OrganizationsTwo	0.003 (0.155)
Gender1	−0.126 (0.129)
VCHubOther	−0.048 (0.167)
VCHubSilicon Valley	0.230 (0.170)
Constant	−1.384** (0.691)
Observations	3,667
Log Likelihood	−2,177.541
Akaike Inf. Crit.	4,381.082
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable: GotLSV
IsAlum1	0.320** (0.141)
Number.of.FoundersMany	0.293* (0.156)
Number.of.FoundersSolo	−0.329* (0.190)
Number.of.FoundersThree	0.145 (0.147)
Number.of.ExitsNone	−0.153 (0.828)
Number.of.ExitsOne	−0.320 (0.931)
Number.of.Founded.OrganizationsOne	0.174 (0.245)
Number.of.Founded.OrganizationsThree	0.367 (0.296)
Number.of.Founded.OrganizationsTwo	0.383 (0.257)
Gender1	−0.013 (0.200)
VCHubOther	−0.349 (0.242)
VCHubSilicon Valley	0.142 (0.241)
Constant	−1.226 (0.903)
Observations	1,645
Log Likelihood	−891.546
Akaike Inf. Crit.	1,809.093
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 16: Output for Seed, Early Stage Venture, and Late Stage Venture Logistic Regression Models

The acquisition model also shows alumni status and *many* co-founders has a strong positive effect with coefficients of 0.65 and 0.39 respectively. Likewise being a solo founder strongly decreases acquisition chances with a coefficient of -0.45. Being outside of both Boston and Silicon Valley is also significant, with a negative coefficient of -0.33. All other factors are not statistically significant. The IPO model shows no significant predictors in any variable. This is consistent with our earlier findings that the exit purse is not affected by the same variables. To reiterate, it seems that the market is "accurately" assessing and pricing start-ups, indiscriminately of founder background. These models show that the closer start-ups get to the innovation cluster, the more outsize the impact of founder background. Whereas when start-ups can no longer benefit from VC network effects, founder background becomes insignificant.

	Dependent variable:
	Acquired
IsAlum1	0.659***
	(0.082)
Number.of.FoundersMany	0.397***
	(0.095)
Number.of.FoundersSolo	−0.454***
	(0.080)
Number.of.FoundersThree	0.179**
	(0.080)
Number.of.ExitsNone	−0.152
	(0.403)
Number.of.ExitsOne	−0.256
	(0.463)
Number.of.Founded.OrganizationsOne	0.187
	(0.126)
Number.of.Founded.OrganizationsThree	0.120
	(0.156)
Number.of.Founded.OrganizationsTwo	0.155
	(0.134)
Gender1	0.078
	(0.108)
VCHubOther	−0.337***
	(0.123)
VCHubSilicon Valley	−0.061
	(0.127)
Constant	−1.314***
	(0.445)
Observations	6,911
Log Likelihood	−3,360.803
Akaike Inf. Crit.	6,747.605
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable:
	IPOd
IsAlum1	0.512
	(0.408)
Number.of.FoundersMany	−0.213
	(0.470)
Number.of.FoundersSolo	−16.404
	(912.116)
Number.of.FoundersThree	−0.343
	(0.466)
Number.of.ExitsNone	15.746
	(4,608.763)
Number.of.ExitsOne	16.454
	(4,608.763)
Number.of.Founded.OrganizationsOne	−0.259
	(0.803)
Number.of.Founded.OrganizationsThree	0.452
	(0.898)
Number.of.Founded.OrganizationsTwo	−0.076
	(0.825)
Gender1	−0.669
	(0.538)
VCHubOther	−0.571
	(0.721)
VCHubSilicon Valley	−0.162
	(0.683)
Constant	−17.169
	(4,608.763)
Observations	416
Log Likelihood	−102.695
Akaike Inf. Crit.	231.389
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 17: Output for Acquisition and IPO Logistic Regression Models

Logistic Regression: Firm, Founder, and Round View

Notice in this model, we cannot analyze the seed stage, because there is no previous investor signal. When including the number of investors at the previous stage, it affects the results of the Firm and Founder View Model. In both the ESV and LSV stages, alumni status is significant, but the coefficient is lesser at 0.35 and 0.28 respectively. The impact of the number of founders has not changed, with *many* founders being the most significant. Interestingly, being a solo founder is not penalized at the LSV stage, when factoring in the number of investors. This could be that the signal from investors overwhelms the negative signal of having a smaller founding team.

In the ESV model, having *many* seed stage investors has a strong positive impact, with a coefficient of 0.35. Naturally, having *no* previous investors strongly decreases the chances of achieving ESV, but some firms do leapfrog the seed stage and directly enter the ESV stage. In the LSV model, having *many* ESV investors is not significant. This may indicate that the firm is entering maturity and can be assessed in other ways. In both models, having *one* previous stage investor is a strong negative predictor with coefficients of -1.45 at the ESV stage and -0.87 at the LSV stage. The literature suggests two reasons: the first is that only having one participatory firm is a negative signal or that the single firm monopolizes the start-up's cap table to a point where it makes further dilution infeasible. Since we do not have access to equity partitions, we can only observe the results of the model.

	Dependent variable:
	GotLSV
IsAlum1	0.280** (0.142)
Number.of.FoundersMany	0.295* (0.159)
Number.of.FoundersSolo	−0.274 (0.192)
Number.of.FoundersThree	0.109 (0.148)
Number.of.ExitsNone	−0.128 (0.839)
Number.of.ExitsOne	−0.327 (0.942)
Number.of.Founded.OrganizationsOne	0.234 (0.248)
Number.of.Founded.OrganizationsThree	0.401 (0.299)
Number.of.Founded.OrganizationsTwo	0.449* (0.261)
Gender1	0.032 (0.202)
VCHubOther	−0.328 (0.246)
VCHubSilicon Valley	0.128 (0.245)
NumESVInvestors2	0.202 (0.142)
NumESVInvestors3	−0.559** (0.261)
NumESVInvestors4	−0.870*** (0.178)
Constant	−1.197 (0.914)
Observations	1,645
Log Likelihood	−872.196
Akaike Inf. Crit.	1,776.392
Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable:
	GotESV
IsAlum1	0.350*** (0.104)
Number.of.FoundersMany	0.531*** (0.112)
Number.of.FoundersSolo	−0.358*** (0.112)
Number.of.FoundersThree	0.183* (0.095)
Number.of.ExitsNone	0.436 (0.668)
Number.of.ExitsOne	0.616 (0.718)
Number.of.Founded.OrganizationsOne	−0.017 (0.152)
Number.of.Founded.OrganizationsThree	0.005 (0.188)
Number.of.Founded.OrganizationsTwo	−0.023 (0.162)
Gender1	−0.119 (0.134)
VCHubOther	0.073 (0.173)
VCHubSilicon Valley	0.187 (0.176)
NumSeedInvestorsMany	0.357*** (0.125)
NumSeedInvestorsNone	−1.450*** (0.107)
NumSeedInvestorsOne	−0.396*** (0.095)
Constant	−0.875 (0.705)
Observations	3,667
Log Likelihood	−2,041.092
Akaike Inf. Crit.	4,114.184
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 18: Output of Funding Pipeline Logistic Regression Models Considering Number of Previous Investors

Multivariate Logistic Regression Models

The multivariate logistic regression models will consider a risk set where founders that get *this* stage will be considered in the *next* stage's model. There are three possible outcomes: got *stage*, failed, or stalled. Due to database constraints, it is difficult to measure the timelines between founding and each round, so we designate a stall as a firm that cannot achieve the next stage of venture funding but has not ceased operations. As mentioned in the methodology section, these firms may be somewhat successful but are no longer on a path to exit. In the final model, after which firms received late stage venture funding, we expand our possibilities to include failure, acquisition, and IPO. All of the following models use stall as a baseline.

Seed Stage Multivariate Logistic Regression Results

At the seed stage, we see that alumni status has no statistical effect. The number of founders is again significant, with many having the highest positive impact and being a solo founder has the lowest negative impact. VC Hub is significant, where Silicon Valley firms have higher likelihood. This model shows male founders have a higher chance of achieving seed funding. While it is unclear why the seed stage does not share the results of previous models, the literature suggests a few explanations. Namely that seed stage investors are likely not institutional venture capitalists and therefore may be more tight pursed than VC firms (Bachher & Guild 1996). This model indicates that founder background is not sufficient in identifying what start-ups achieve seed funding.

	Dependent variable:	
	Failed	GotSeed
	(1)	(2)
IsAlum1	−0.046	0.131
	(0.103)	(0.085)
Number.of.FoundersMany	0.252**	0.504***
	(0.119)	(0.097)
Number.of.FoundersSolo	−0.366***	−0.821***
	(0.077)	(0.068)
Number.of.FoundersThree	0.065	0.322***
	(0.094)	(0.076)
Number.of.ExitsNone	0.164	0.326
	(0.429)	(0.383)
Number.of.ExitsOne	0.223	0.439
	(0.492)	(0.434)
Number.of.Founded.OrganizationsOne	−0.066	0.047
	(0.129)	(0.112)
Number.of.Founded.OrganizationsThree	−0.057	0.151
	(0.163)	(0.140)
Number.of.Founded.OrganizationsTwo	−0.088	0.035
	(0.138)	(0.119)
Gender1	0.134	0.163*
	(0.114)	(0.097)
VCHubOther	0.293**	0.192*
	(0.144)	(0.115)
VCHubSilicon Valley	0.564***	0.363***
	(0.149)	(0.120)
Constant	−1.075**	−0.599
	(0.477)	(0.420)
Akaike Inf. Crit.	14,376.020	14,376.020
Note: 1	*p<0.1; **p<0.05; ***p<0.01	

Figure 19: Seed Stage Multivariate Regression Model Output

Early Stage Venture Stage Multivariate Logistic Regression Results

The ESV stage results are consistent with those found in our initial logistic regression models. The alumni effect is less significant but significant nonetheless, with a coefficient of 0.555. The number of founders is the same, with *many* having the highest positive effect and *solo* having the lowest negative effect.

	Dependent variable:	
	Failed	GotESV
	(1)	(2)
IsAlum1	0.096	0.550***
	(0.126)	(0.112)
Number.of.FoundersMany	0.096	0.624***
	(0.136)	(0.119)
Number.of.FoundersSolo	0.069	−0.515***
	(0.106)	(0.120)
Number.of.FoundersThree	−0.011	0.176*
	(0.107)	(0.101)
Number.of.ExitsNone	0.328	0.342
	(0.670)	(0.677)
Number.of.ExitsOne	0.454	0.575
	(0.730)	(0.731)
Number.of.Founded.OrganizationsOne	−0.177	−0.070
	(0.161)	(0.163)
Number.of.Founded.OrganizationsThree	−0.242	−0.005
	(0.203)	(0.200)
Number.of.Founded.OrganizationsTwo	−0.201	0.009
	(0.172)	(0.173)
Gender1	−0.013	−0.045
	(0.148)	(0.144)
VCHubOther	0.274	−0.022
	(0.198)	(0.179)
VCHubSilicon Valley	0.523***	0.290
	(0.203)	(0.184)
Constant	−1.266*	−1.133
	(0.719)	(0.719)
Akaike Inf. Crit.	7,537.844	7,537.844
Note: 2	*p<0.1; **p<0.05; ***p<0.01	

Figure 20: Early Stage Venture Stage Multivariate Regression Model Output

Late Stage Venture Stage Multivariate Logistic Regression Results

As mentioned earlier, our possible outcomes are acquired, failed, got LSV, and (a baseline of) stall. Here we see alumni status has a significant effect in the outcomes of acquisition and LSV, with coefficients of 0.44 and 0.35 respectively. The amount of founders is no longer significant, nor is the effect of geography. Uniquely, having no exits and one previous exit has a significant effect in the chance of failure, whereas it makes no appearance in other models. In this model, it suggests founder background has a high likelihood of predicting all three of acquisition, failure, and getting LSV funding.

	Dependent variable:		
	Acquired	Failed	Got LSV
	(1)	(2)	(3)
IsAlum1	0.441***	−0.092	0.353**
	(0.150)	(0.290)	(0.175)
Number.of.FoundersMany	0.210	−0.038	0.347*
	(0.161)	(0.318)	(0.190)
Number.of.FoundersSolo	−0.261	0.545**	−0.258
	(0.175)	(0.266)	(0.219)
Number.of.FoundersThree	0.125	0.118	0.252
	(0.148)	(0.273)	(0.176)
Number.of.ExitsNone	−0.028	10.832***	0.521
	(0.771)	(0.335)	(1.130)
Number.of.ExitsOne	−0.567	10.920***	0.093
	(0.883)	(0.485)	(1.239)
Number.of.Founded.OrganizationsOne	0.263	0.134	0.121
	(0.236)	(0.404)	(0.276)
Number.of.Founded.OrganizationsThree	0.374	0.092	0.385
	(0.294)	(0.514)	(0.340)
Number.of.Founded.OrganizationsTwo	0.337	0.026	0.253
	(0.251)	(0.437)	(0.294)
Gender1	0.245	0.422	0.039
	(0.203)	(0.390)	(0.231)
VCHubOther	−0.165	0.514	−0.138
	(0.238)	(0.542)	(0.299)
VCHubSilicon Valley	0.098	0.760	0.417
	(0.242)	(0.548)	(0.300)
Constant	−0.753	−13.834***	−1.791
	(0.850)	(0.535)	(1.207)
Akaike Inf. Crit.	4,050.986	4,050.986	4,050.986
Note:	3	*p<0.1; **p<0.05; ***p<0.01	

Figure 21: Late Stage Venture Stage Multivariate Regression Model Output including Acquisition Outcomes

Post Late Stage Venture Stage Multivariate Logistic Regression Results

The final risk set includes all of the firms that received LSV funding. Here we consider the outcomes of acquisition, failure, and IPO. Intuitively, there are no firms that IPO'd without raising LSV funding or before doing so. As found in the previous logistic regression considering the likelihood of IPO, being an alumni founder has no significant effect. Since there are so few firms in our sample frame that IPO'd, we see interesting results as it relates to previous exit status. The model shows that having no or one exit leads to high probability of achieving an IPO. However this must be considered in the context of the previous models. Since the founders that IPO'd may be first time founders, this would explain the strong coefficient. However since the number of founded organizations is insignificant, this explanation is inconclusive.

	Dependent variable:		
	Acquired	Failed	IPOd
	(1)	(2)	(3)
IsAlum1	0.430	0.970**	0.676
	(0.300)	(0.486)	(0.417)
Number.of.FoundersMany	0.123	−0.002	−0.177
	(0.332)	(0.626)	(0.478)
Number.of.FoundersSolo	−0.535	−0.994	−41.166
	(0.470)	(1.108)	
Number.of.FoundersThree	−0.387	0.421	−0.389
	(0.341)	(0.550)	(0.473)
Number.of.ExitsNone	−1.286	14.136***	10.283***
	(1.433)	(0.689)	(0.542)
Number.of.ExitsOne	−0.760	−20.093***	11.080***
	(1.676)	(0.000)	(0.844)
Number.of.Founded.OrganizationsOne	32.461***	−1.010	−0.121
	(0.429)	(0.730)	(0.811)
Number.of.Founded.OrganizationsThree	32.453***	−1.714	0.556
	(0.502)	(1.222)	(0.909)
Number.of.Founded.OrganizationsTwo	32.761***	−0.693	0.156
	(0.447)	(0.763)	(0.835)
Gender1	0.534	−0.534	−0.610
	(0.513)	(0.686)	(0.551)
VCHubOther	−0.653	−0.143	−0.735
	(0.491)	(1.133)	(0.738)
VCHubSilicon Valley	−0.651	0.271	−0.297
	(0.486)	(1.090)	(0.700)
Constant	−32.404***	−15.872***	−11.498***
	(1.188)	(0.689)	(0.818)
Akaike Inf. Crit.	790.103	790.103	790.103
Note:	4		*p<0.1; **p<0.05; ***p<0.01

Figure 22: Post-Late Stage Venture Stage Multivariate Regression Model Output including Acquisition, Failure, and IPO Outcomes

Discussion and Further Review

In our models, we see that alumni status has some positive effect on the likelihood of raising venture capital. This is consistent with the assumptions found in literature as well as our overall hypothesis. The missing data points severely limit our ability to draw further conclusions. For instance, there was no way to control for position (engineer, sales, etc.) or seniority at the Big Tech firm. In certain outcomes, alumni status had no effect. This indicates that a firm's success is dependent on both the founding team as well as its overall viability in a given market.

A more holistic approach would be to partner with our sample of Big Tech firms, securing anonymized internal performance data as well as data on who leaves the firm. Of the latter, we could follow them to see if they found start-ups or work elsewhere. If they choose to start a firm, we can observe the industry, founding teams' provenance, and the biographical data captured in our report. Next, we could partner with venture capital firms that invest in these companies, comparing their scorecards against that of the founder's former company. Finally we can benchmark the skill of employees at the Big Tech firm against the criteria of the firms that get funding. From which, we can analyze the universe of Big Tech employees, noting whether higher-scored employees tend to leave to start companies and (if so) whether VCs are over-investing in founders when their credentials might be lesser than those of employees who didn't choose to start companies. While this is infeasible for a myriad of reasons, it would more effectively capture the effect of working in Big Tech as it relates to receiving venture capital.

Conclusion and Interpretation

The history of Silicon Valley informs the norms of modern technological innovation. This is important to understand to examine the network structure of incumbent technology firms, new firms, and their relationship to venture capital. The formation of start-ups is one way large, incumbent technology firms are displaced. We define Big Tech to be a unique set of firms upon which new entrants build their products and ultimately compete against. As a result, we hypothesize that "alumni" of these firms have certain advantages over non-alumni when starting ventures.

As raising venture capital is a key determinant for start-up success, we measure how alumni status affects the likelihood of achieving each stage in our funding pipeline. Overall, alumni have a statistically significant advantage in raising venture funding. In each of our Seed, Early Stage Venture, and Late Stage Venture stages, alumni status is a significant predictor. However, depending on the model, this advantage may be nuanced or not exist. A few examples are in the amount of funding raised, exit purse (of an acquisition), and the overall likelihood of an IPO.

We interpret the results based on the literature review concerning critical network theory and venture capital criteria. Based on our results, we see that the closer a founder is to the innovation cluster, the greater effect it has on the likelihood that they raise venture funding. As the set of outcomes goes beyond the venture capitalists' control, the effects are no longer present. Finally, we do not make an attempt to measure what about these alumni makes them successful (whether it be technical expertise, personal network, or skills learned at the Big Tech firm.

References

Abdelkafi, N., Raasch, C., Roth, A. *et al.* (2019). Multi-sided platforms. *Electron Markets* 29, 553–559 https://doi.org/10.1007/s12525-019-00385-4 https://link.springer.com/article/10.1007/s12525-019-00385-4

Abrahamson, E., and Rosenkopf, L. (1997). Social network effects on the extent of Innovation Diffusion: A Computer Simulation. *Organization Science*, *8*(3), 289–309. https://doi.org/10.1287/orsc.8.3.289

Apple Newsroom (2021, June). “Meeting pandemic challenges, Apple developers grow total billings and sales in the App Store ecosystem by 24 percent to $643 billion in 2020” *Apple Newsroom* https://www.apple.com/newsroom/2021/06/apple-developers-grow-app-store-ecosystem-billings-and-sales-by-24-percent-in-2020/

Anatolijs Prohorovs, Julija Bistrova & Daria Ten (2019) Startup Success Factors in the Capital Attraction Stage: Founders’ Perspective, Journal of East-West Business, 25:1, 26-51, DOI: 10.1080/10669868.2018.1503211 https://www.tandfonline.com/action/showCitFormats?doi=10.1080%2F10669868.2018.1503211

Azoulay, P., Jones, B., Kim, D.J., and Miranda J. (2018 May 18) “How Old Are Successful Tech Entrepreneurs?” *Kellogg Insight* https://insight.kellogg.northwestern.edu/article/younger-older-tech-entrepreneurs

Bachher, J & Guild, Paul. (1996). "Financing early stage technology based companies: investment criteria used by investors." Frontiers of Entrepreneurship Research. Frontiers of Entrepreneurship Research. https://www.researchgate.net/publication/256420976_Financing_early_stage_technology_based_companies_investment_criteria_used_by_investors_Frontiers_of_Entrepreneurship_Research

Bachher, J.S. & Diaz de Leon Lopez, Enrique & Guild, Paul. (1999). Decision criteria used by investors to screen technology-based ventures. International Journal of Thermal Sciences - INT J THERM SCI. 269 - 273 vol.2. 10.1109/PICMET.1999.787815.

Beard, A. (2021, December). “Time to Rein In Big Tech?” Harvard Business Review, November-December 2021. https://hbr.org/2021/11/time-to-rein-in-big-tech

Brin, S., Page, L. (1998) The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Stanford Infolab* http://infolab.stanford.edu/~backrub/google.html

Buckland J. (2021) AllBirds Going to Zero. *Shopify Plus* https://www.shopify.com/plus/allbirds

Chaplinsky, S., Hanley, K.W. and Moon, S.K. (2017), The JOBS Act and the Costs of Going Public. Journal of Accounting Research, 55: 795-836. https://doi.org/10.1111/1475-679X.12172

Choi, H. Kim, S., Lee, J. (2010) Role of network structure and network effects in diffusion of innovations, *Industrial Marketing Management*, Volume 39, Issue 1, Pages 170-177, ISSN 0019-8501, https://doi.org/10.1016/j.indmarman.2008.08.006 .(https://www.sciencedirect.com/science/article/pii/S0019850108001557)

Cockayne, D. (2019, December) “What is a startup firm? A methodological and epistemological investigation into research objects in economic geography” Geoforum, Volume 107, 2019, Pages 77-87, ISSN 0016-7185, https://doi.org/10.1016/j.geoforum.2019.10.009.

Deighton J. (1999, November) “Webvan: Groceries on the Internet” https://www.hbs.edu/faculty/Pages/item.aspx?num=26728

Domonoske C. (2014, June 20). Stephanie Kwolek, Chemist Who Created Kevlar, Dies at 90. *National Public Radio* *https://www.npr.org/sections/thetwo-way/2014/06/20/323951708/stephanie-kwolek-chemist-who-created-kevlar-dies-at-90#:~:text=More%20Podcasts%20%26%20Shows-,Stephanie%20Kwolek%2C%20Chemist%20Who%20Created%20Kevlar%2C%20Dies%20At%2090%20%3A,world%20in%20bulletproof%20body%20armor*

Efrati, A. (2015, April 6) The Ascension of Google’s Sridhar Ramaswamy. *The Information* https://www.theinformation.com/articles/the-ascension-of-google-s-sridhar-ramaswamy

Elon University (1990). “1870s - 1940s: Telephone” Imagining the Internet: a History and Forecast. https://www.elon.edu/u/imagining/time-capsule/150-years/back-1870-1940/

Engel, J. S. (2015). Global Clusters of Innovation: Lessons from Silicon Valley. California Management Review, 57(2), 36–65. https://doi.org/10.1525/cmr.2015.57.2.36 https://journals.sagepub.com/doi/pdf/10.1525/cmr.2015.57.2.36

Ferrary, Michel and Granovetter, Mark. (2009). The Role of Venture Capital Firms in Silicon Valley's Complex Innovation Network. Economy and Society - ECON SOC. 38. 326-359. 10.1080/03085140902786827.

Gompers, P. and Lerner, J. (1999, September) *The Venture Capital Cycle*. MIT Press. *https://mitpress.mit.edu/books/venture-capital-cycle*

Google. (2020) *From 10-k 2020 Annual Report*. https://www.sec.gov/Archives/edgar/data/1652044/000165204421000010/goog-20201231.htm

Hartmans, A. (2020, October 6) The rise of Kevin Systrom, who founded Instagram 10 years ago and built it into one of the most popular apps in the world. Business Insider https://www.businessinsider.com/kevin-systrom-instagram-ceo-life-rise-2018-9

Harvard Business School. (2020, November 12). *What are network effects?: HBS Online*. Business Insights - Blog. Retrieved November 24, 2021, from https://online.hbs.edu/blog/post/what-are-network-effects.

Henton, D., and Held, K. (2013). The dynamics of Silicon Valley: Creative destruction and the evolution of the innovation habitat. Social Science Information, 52(4), 539–557. https://doi.org/10.1177/0539018413497542 https://journals.sagepub.com/doi/pdf/10.1177/0539018413497542

Hershovitz, M. (1995, October). “Unhitching the Trailer Clause: The Rights of Inventive Employees and Their Employers” Journal of Intellectual Property Law Vol. 3 No. 1 (Article 7).

https://digitalcommons.law.uga.edu/cgi/viewcontent.cgi?referer=andhttpsredir=1andarticle=1110andcontext=jipl

Kantrowitz, A. (2021, May 16). "Why former Google ads Boss Sridhar Ramaswamy is building an ad-free search engine." CNBC Tech. https://www.cnbc.com/2021/05/16/sridhar-ramaswamy-ex-google-building-ads-free-search-engine-neeva-.html

Katz, M. and Shapiro, C. (1994). "Systems Competition and Network Effects" Journal of Economic Perspectives Vol. 8, No. 2 (pp. 93-115). https://www.aeaweb.org/articles?id=10.1257/jep.8.2.93

Lunden, I. (2019 October 31). "Crunchbase raises $30M more to double down on its ambition to be a 'LinkedIn for company data'" *TechCrunch* https://techcrunch.com/2019/10/31/crunchbase-raises-30m-more-to-double-down-on-its-ambition-to-be-a-linkedin-for-company-data/

Loyalty Conversion Systems Corporation v. American Airlines Inc (E.D. Tex. 2014). https://law.justia.com/cases/federal/district-courts/texas/txedce/2:2013cv00655/146668/127/

MIT Enterprise Forum CEE "What is a Stealth Startup?" *MIT Enterprise Forum CEE* https://mitefcee.org/what-is-a-stealth-startup/

Neufeld, D. (2018, March 26). "Shopify's Infrastructure Collaboration with Google" Shopify Engineering. https://shopify.engineering/shopify-infrastructure-collaboration-with-google

Nicas, J. (2016, April 19). "Google Alumni Pitch Investors at Xoogler Demo Days" Wall Street Journal. https://www.wsj.com/articles/google-alumni-pitch-investors-at-xoogler-demo-days-1461092062

Okrah, J., Nepp, A., & Agbozo, E. (2018 April). "Exploring the factors of startup success and growth" *The Business and Management Review* Vol. 9 Num. 3 https://cberuk.com/cdn/conference_proceedings/2019-07-14-09-58-17-AM.pdf

Olson, P. (2018, September 26). "Exclusive: WhatsApp Cofounder Brian Acton Gives The Inside Story On #DeleteFacebook And Why He Left $850 Million Behind." *Forbes* https://www.forbes.com/sites/parmyolson/2018/09/26/exclusive-whatsapp-cofounder-brian-acton-gives-the-inside-story-on-deletefacebook-and-why-he-left-850-million-behind/?sh=6b9fd7b03f20

Parc, a Xerox Company. Parc History. https://www.parc.com/about-parc/parc-history/

Peal, G (2018, June 19). "Sunsetting React Native" *The AirBnB Tech Blog from AirBnB Engineering* https://medium.com/airbnb-engineering/sunsetting-react-native-1868ba28e30a

Rubens, N. et al. (2011). "Alumni network analysis." IEEE Global Engineering Education Conference (EDUCON) pp. 606-611. doi: 10.1109/EDUCON.2011.5773200

Rose, D. 2012. How many start-ups in the US get seed/VC funding per year?. http://blog.gust.com/how-many-start-ups-in-the-us-get-seedvc-funding-per-year (accessed January 2, 2018). [Google Scholar]

Saxenian, A. L. (1996). *Regional advantage: Culture and competition in Silicon Valley and Route 128*. Cambridge, Mass: Harvard University Press.

Schumpeter, J. A., and Opie, R. (1934). "The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle." Cambridge, Mass: Harvard University Press.

Simmons, J. (2012, November 26). "Inventions Made for Hire" The NYU Journal of Intellectual Property and Entertainment Law Vol. 2 No. 1 https://jipel.law.nyu.edu/vol-2-no-1-1-simmons/#_ftn46

Statt, N. (2016, February 1). "WhatsApp has grown to 1 billion users: More than Doubling Since Facebook Acquired It." The Verge. https://www.theverge.com/2016/2/1/10889534/whats-app-1-billion-users-facebook-mark-zuckerberg

The White House Office of the Press Secretary. (2012). "President Obama To Sign Jumpstart Our Business Startups (JOBS Act). *The White House Office of the Press Secretary* https://obamawhitehouse.archives.gov/the-press-office/2012/04/05/president-obama-sign-jumpstart-our-business-startups-jobs-act

U.S. Citizenship and Immigration Services. (2020). "H-1B Employer Data Hub" *U.S. Citizenship and Immigration Services Reports and Studies* https://www.uscis.gov/tools/reports-and-studies/h-1b-employer-data-hub

Vandor, P. (2021, August 4). "Research: Why Immigrants Are More Likely to Become Entrepreneurs" *Harvard Business Review* https://hbr.org/2021/08/research-why-immigrants-are-more-likely-to-become-entrepreneurs#:~:text=A%202012%20study%20found%20that,and%2025%25%20of%20startup%20 founders.

Wakabayashi, D. (2020, June 19) A former Google Executive Takes Aim at His Old Company With a Start-Up. *The New York Times* *https://www.nytimes.com/2020/06/19/technology/google-neeva-executive.html*

Watts, D., and Strogatz, S. (1998) Collective dynamics of 'small-world' networks Nature 393 (6684): 440--442 https://www.bibsonomy.org/bibtex/02e54cb4f5d01567f9d4d2bfd5ca6587#:~:text=D

Zider, B. (1998, December). “How Venture Capital Works” *Harvard Business Review* November-December 1998. Reprint 98611 http://mengwong.com/school/HarvardBusinessReview/how_venture_works.pdf