Venture Capital and the Dynamism of Startups

Yang Yu
Beijing, China

M.S. Economics for Development, University of Oxford, 2016
B.A. Economics, Tsinghua University, 2015

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of Doctor of Philosophy

Department of Economics

University of Virginia
May, 2025

Abstract

I study the strategic investment in startups by competing venture capitals (VCs), focusing on the impact of uncertainty about startup quality on investment and startup dynamism. My framework considers multi-round funding requirements and VC optimization based on current information and projections of future success (M&A or IPO). Using a novel dataset on the "life-cycle" of biotech and software startups from *2000 to 2022*, I establish that my dataset identifies model parameters and propose a method to correct for dynamic selection to infer startups values and VCs information. Among several others, I find that (1) biotech investors initially possess more information than software investors but learn slower, reflecting sector-specific uncertainty, but in both cases the investors eventually learn the true quality; (2) uncertainty leads to underfunding of promising startups, causing welfare losses of 22% and 21% in biotech and software, respectively; and (3) positive "dynamic information externality" from early stage investors to late stage investors causes the former to invest less, leading to welfare losses of $10 billion in biotech and $3 billion in software. I then explore policies to mitigate these losses. I also estimate that an undervalued M&A exit reduces VC returns, significantly decreasing startup funding and exacerbating welfare losses.

JEL classification: C15, C32, D44, D83, G24, L26, M13

Keywords: Startup, Venture Capital, Information and Learning, Auction

Acknowledgements

The past six years have been a journey of persistent effort, often running into dead ends, in pursuit of those rare moments of meaningful progress. By its very nature, this process is often shaped by factors out of my hands. Looking back, all I feel is that I have been a fortunate person — someone who picked a topic that led to something in the end, met with many beautiful minds that supported me through the ups and downs, and got connected to exceptional researchers outside economics who expanded the scope of my work.

First, I am grateful to my primary advisors, Federico Ciliberto, Gaurab Aryal and Simon Anderson, for leading me on the academic career in empirical IO and for providing me with invaluable guidance, understanding and care throughout this journey. I have been inspired by their creativity, breadth of knowledge, profound understanding of their own field(s), and deep passion for science. I often felt off balance — yet very much enjoyed and valued our many conversations, during which I was intellectually challenged and pushed to my limits in thinking, in order to refine an idea. Their rigorous scientific attitude and working style — as I've observed — have shaped, and will continue to shape mine as I start my academic career. I am lucky to be their student.

I would like to acknowledge my outside committee member, Chris Yung, as well as other faculty members at UVa and UChicago — Kerem Cosar, Maxim Engers, Jack Fisher, Alex Mackey, Doron Ravid, Eric Richert, Peter Troyan, and Eric Young — for providing valuable comments that have significantly improved this dissertation. I would also like to thank Boning Zhou, my internship colleague at Amazon, for teaching me parallel computing — without it my code would likely still be running today. Besides, my gratitude goes to my co-authors and fellow students, especially Lingmin Bao, Anirban Chattopadhyaya, Moonju Cho, Xiaoyun Qiu, Max Schnidman, and Jiong Wu, for teaching me how to be a good collaborator, supporting me through difficult choices about the future, and for their companionship and friendship throughout several milestones of this journey and each and every bittersweet moment along the way.

A special thanks goes to Haifeng Xu for introducing me to the world of EconCS and for offering me precious opportunities to grow intellectually and build meaningful connections in this space. Through our many discussions over the years, I have been fortunate to benefit from his sharp and bold angles of thinking a problem, smoothness and beauty of academic writing and rebuttal, and genuine curiosity and desire to make disruptive progress and create new knowledge. These collaborations have injected new energy to my work in economics and broadened the scope of my research agenda. Because of them, I become (more) hopeful for an academic career and want to be part of the forefront of this interdisciplinary field. Without Haifeng's mentorship and collaboration, I would not be where I am today.

When I was an undergrad — knowing nothing and feeling lost about the future — I was fortunate to cross paths with Saiyin Sun (who probably doesn't remember me now) by taking her courses in *English Reading* and *Images of China in English Literature*. In her classroom, I was pushed to confront many ideas, perspectives, and concepts that either I had never heard of or were not aligned with (and even in the opposite of) the conventions that I had grown up with. This education opened my eyes and, after years of reflection and internalization, has emerged as a mindset of being explorative and open-minded that has guided many of the critical choices I've made since. I cannot thank her enough.

There are a few more people that I would like to mention and thank here. I am grateful to several economists and mathematicians that I met before starting graduate school, Doug Gollin, Peter Popkowski-Leszczyc, Xiaodong Ma, Mingyun Tang, and Binzhen Wu, for supporting my PhD application and helping me out at different points in time. I was also lucky to have met some lifelong friends so far. I send my thanks to Man Luo, Fei Guo, Yifei Xu, and Yuwei Cheng for being caring, supportive and inspiring friends all the time and for creating many lovely memories and bringing joy and color to my life.

Finally, I owe my deepest gratitude to my family. I am grateful for my parents, Wei Zhao and Zhende Yu, and my grandfather, Hongqing Zhao, for their unconditional and unreserved love that has made me feel secured and belonged in this world, and my sister, Xiao Yu, for sharing with me lots of fun, cheering me up in my down times, and always standing by me.

I dedicate this dissertation to my grandfather, who was a faculty member at a university in Beijing during the 1950s but was later expelled (along with many of his colleagues) amid the well-known historical events. It was not until three decades later — after the prime of his life had passed — that he was finally able to return. He would have been very happy to see this dissertation and to know that I have followed in his academic footsteps, but in what is (hopefully) a more peaceful era.

Contents

1 Introduction

Innovation is the cornerstone of economic growth and prosperity (e.g., Aghion et al., 2021; Akcigit and Van Reenen, 2023), and startups are important source of innovation (Kolev et al., 2022). For instance, 57% of all new drugs approved in the US in 2023 originated at startups (Economist, 2024). Startups are mostly funded by venture capitals (VCs) across multiple rounds. Although only 0.5% firms in the US receive VC financing, VC-backed firms make up over half of all IPOs and outperform other firms in market value, revenue, employment, and R&D (Lerner and Nanda, 2020). However, there is concern that the dynamism of startups is not at the level it could be, especially in the biotech and software sectors (Bloom et al., 2020; Park et al., 2023). One reason for this could be intrinsic to the imperfect competition among VCs over financing startups with uncertain qualities, leading to the underfunding of many viable startups.

If so, understanding the performance of this capital market and identifying the source(s) of inefficiency, if any, is an important first step to finding policies to support innovation through startups. For that, we have to answer the following questions: How does multi-round VC funding affect the dynamism of startups? How are the returns from startups shared across VCs at different stages? Is this capital market efficient? If not, what are the extent and causes of inefficiency? In this paper, I answer these questions in the context of the financing of biotech and software startups in the US, which, to my knowledge, has not yet been studied. To this end, I evaluate the strategic investment decisions of competing VCs, quantify inefficiency, and identify their cause(s) to understand the functioning of the current market and provide a new framework to evaluate policies meant to support the industry.

To understand the possibility of inefficiency, note that despite a potentially large upside, funding a startup entails tangible risk. On the one hand, most startups fail, some are average, and only a handful generate outstanding returns, but it is hard, if not impossible, to "pick the winner" at the outset.[1] As a result, VCs may under-invest, especially in early rounds when there is too little information about the startup. On the other hand, all else equal,

[1] For instance, Bessemer Venture Partners is one of the largest VC firms, but they declined the opportunities to invest in Airbnb, Apple, eBay, Facebook, Tesla, and Zoom, among others, on account of high uncertainties. Furthermore, Kerr et al. (2014) find that among several ex-ante equally ranked startups, 60% failed, while 10% generated a more than five times return on investment.

early-round investments generate higher returns than late-round investments. However, the former can generate informational spillover benefitting late-round VCs, which I refer to as positive "*dynamic informational externality.*" Which of these opposing forces dominates is an empirical question of interest that I answer.

To this end, first, I develop a flexible but tractable framework that captures uncertainty in startup quality, multi-round financing, dynamic information structure, and imperfect VC competition, building on the pure common value auction model of Milgrom and Weber (1982). Second, I establish conditions under which the model parameters can be identified from my dataset. In particular, I adapt the results from dynamic discrete choice by Heckman and Navarro (2007) that allow for serially correlated quality signals across multiple rounds in my setting to correct for dynamic selection and infer the startup's true quality and investor's interim information set about the startup. Third, I construct a rich and novel dataset on the universe of biotech and software startups in the US by combining several data sources to estimate this model.

The lack of such data has limited previous studies from incorporating uncertainty and dynamic information structure. I overcame this challenge and created the dataset by leveraging exit value data from Compustat and SEC filings, and VC investment data from Pitchbook. With these data, I can infer the true values of startups and the investors' interim information about them.

I have collected information on 5,591 US startups in the biotech and software sectors between 2000 and 2022 from the Pitchbook database. Besides being two of the most important sectors, I focus on them because they have distinct business models and risk profiles, which can shed light on the market performance. Software businesses are typically built on proven technologies, the primary source of uncertainty is the demand for a new product or service. So, the common practice is to develop and launch a minimum viable product quickly to test market demand.[2] Once validated, they refine the product and scale up the business, which usually comes with less uncertainty (Lerner and Nanda, 2020). In contrast,

[2]The most famous example is Amazon, which wanted to be an e-commerce platform, starting with an online bookstore and validating the demand for e-commerce, at least in the domain of books, before it expanded to become the "everything store" as it is today. See details at https://x.com/zackbshapiro/status/1780601872903979171

drug development consists of several distinct stages, each testing a different aspect of the drug — pre-clinical trial tests safety on animals, phase I assesses toxicity on a small group of healthy individuals, phase II evaluates efficacy on patients and phase III tests efficacy on a larger population (Food and Administration, 2018). Hence, unlike in software, the uncertainties in biotech are spread out over several stages. These sectoral differences are useful to better understand the interaction between VCs and startups with different risk profiles.

A few key data features suggest that startup quality is highly uncertain. To avoid an "all or nothing bet," startup financing is done in stages, enabling the "market" to learn its heretofore unobserved quality. Moreover, this uncertainty likely dissipates as startups mature and more information from previous rounds becomes publicly available. Furthermore, in the data, when funding a startup, it is common for a VC to lead only one round of investment in a startup, and they compete with other VCs.

So, in line with the data, I consider a multi-round game between a startup and several VCs. A startup of unknown (to both entrepreneur and investors) quality needs to secure several rounds of VC funding before exiting. In each round, a (new) set of VCs compete to invest based on their information about the startup. The unit of observation is a startup, and I follow it throughout its life cycle. In this dynamic investment model, VCs make investment decisions based on their current information and expectations on future investment decisions and, thereby, the startup's future trajectory.

In each round, the startup announces the amount of capital to be raised in this round. A set of VCs is matched and each obtains a private signal about the quality. Then, they compete to invest by making offers on how the future returns are shared. I model the competition "as if" VCs compete in an English auction with pure *common value* by bidding equity ownership. The VC that outbids the competitor by offering at least as much equity to the startup founder as the competitor is selected. If the startup fails to raise capital, it quits, and the capital already invested is sunk. If successful, it proceeds to the next round, and the market can infer VCs' signals from previous rounds. In the next round, another set of VCs is matched, and the game iterates until the last round, after which the exit value is realized, and returns are shared among the startup and all VCs according to their final equity share.

I characterize the equilibrium strategy of the game. For each round, I determine the funding rule and equity sharing rule between the VC and the startup based on current and past information. To isolate the source of inefficiency, I also determine the funding rule chosen by a "planner" that maximizes the net returns. To make these two investment strategies comparable, I impose that the planner has the same information as the VCs. I identify a potential "under-investment" in the VC market due to dynamic information spillover, i.e., once an early-round investor invests, it generates information that late-round investors can costlessly observe, helping them make more informed investment decisions. This feature leads to a positive information externality where late-round investors free-ride on the information produced by early-round investors. VCs cannot internalize the full benefit of their investment, leading to under-investment in even viable startups.[3]

I estimate the model using a simulated minimum distance estimator that matches moments in the empirical data to those predicted by the model (e.g., Agarwal, 2015; Aryal et al., 2024; McFadden, 1989; Pakes and Pollard, 1989). Some of the moments I match include summaries of transition patterns of the startups and within- and cross-round variation in capital investment, winning equity bid, and exit value.

The estimates suggest that while the uncertainties in software are primarily concentrated in the initial round, they extend across multiple rounds in biotech. These features are aligned with the distinct business models of each sector. As a result, software investors are initially less informed about the startup's quality but learn faster than biotech investors. Consequently, compared to biotech investors, software investors make more Type I errors — not funding good startups — but fewer Type II errors — funding bad startups. Additionally, in both sectors, early-stage investments generate two to three times returns than later-stage investments, which is comparable to the estimates in Kaplan and Schoar (2005) and Cochrane (2005). Unlike traditional financial models (e.g., CAPM), where risk premium arises from the assumption of risk-averse agents, in my model the uneven returns across stages stem from the sequential dissemination of information.

Using the estimates and the model, I characterize the efficiency level under alternative

[3]Some VC contracts may include anti-dilution protections, primarily to safeguard investors who enter at a high valuation. These protections are not designed to solve the dynamic information externalities that cause inefficiency, because they are triggered only if a startup's value drops in the next round.

information and market structures that provide new insights into the welfare consequences of information externalities and uncertainties. My main finding is that the current biotech and software sectors respectively yield 78% and 79% of the first-best efficiency where all uncertainties are resolved. Information externality and quality uncertainty each contribute to 18% and 73% of the efficiency loss in the biotech sector, translating to $10 billion and $39 billion, while in the software sector, they account for 5% and 54%, or $3 billion and $30 billion, respectively. Most efficiency losses occur in early rounds when the noise-to-signal ratio is the highest, which excludes some viable startups at the early stage. Thus, government-provided funds and grants for early-stage startups would increase market efficiency if allocated properly. Examples include Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR), which the US government runs.

Having identified the sources and magnitudes of inefficiencies in this market, I extend this framework to examine the effects of various policy interventions on innovation. Specifically, I study the relative importance of M&A and IPO exits on the "supply" side of innovation. I find that a 10% undervaluation in an M&A exit leads to a 10% drop in VC market net returns, with 30% attributed to extensive margins — fewer startups get funded, and 70% to intensive margins — funded startups are reduced in value.

Related Literature. My paper relates to the extensive literature on VC and innovation, e.g., Gornall and Strebulaev (2021); Howell et al. (2020); Lerner and Nanda (2020). Most of these papers, however, take the "reduced-form" approach and focus only on a small subset of startups that become publicly listed companies. I contribute by using the structural approach and curating new dataset on private and public companies to study the role of VC in innovation.

My paper is also related to the literature that considers VC-startup interaction, e.g., Ewens et al. (2022); Sørensen (2007). However, these papers only consider the first VC round and assume complete information about the quality of the startups; therefore, they ignore the dynamics involved in funding startups with uncertain qualities. Multi-stage financing and uncertainty about the quality of a startup are considered to be two of the most important features of VC financing (Ewens et al., 2018; Kerr et al., 2014; Lerner and Nanda, 2020). As such, these papers ignore the dynamic information externality and, therefore, underestimate the level of inefficiency. Furthermore, Sørensen (2007) considers a cooperative solution concept, which also ignores competition among VCs.

In terms of the modeling decisions, my paper is related to the literature on auctions with common value, in particular the papers on security design auctions of DeMarzo et al. (2005), Garmaise (2001), and Gorbenko and Malenko (2011), where bidders compete for an asset (i.e., startup) by bidding with securities whose payments are contingent on the asset's value to be realized in the future. Furthermore, given how I model the multi-round financing of startups, my paper is also loosely related to the literature in labor on symmetric employer learning (Altonji and Pierret, 2001; Aryal et al., 2022; Lange, 2007), where, just as employers learn about a worker's abilities over time, in my setting, VCs in later rounds can infer the signals of earlier-round VCs.

Lastly, my paper builds upon the identification results of discrete choice models by Heckman and Navarro (2007). In the context of startups, besides the right censoring of the data, we do not observe their true value because most startups fail. Furthermore, a startup can fail at different stages of financing, leading to dynamic selection, such that the "errors" are serially correlated. In this setting, identification is challenging as the standard results based on independent shocks, e.g., Rust (1987), do not apply. I show that we can adapt Heckman and Navarro (2007) to correct for the dynamic selection.

Overall, my paper contributes to the literature by providing a comprehensive framework for analyzing VC investments under uncertainty, offering new insights into market efficiency, and exploring the potential impacts of policy changes on startup outcomes.

2 Data

In my data, the unit of observation is a startup. My primary dataset on startup financing rounds and exit values comes from *Pitchbook*.[4] This database tracks each startup's VC rounds, providing detailed data on capital investment, valuation, equity ownership and investor identities at the round level. It also provides information on startup sector, patents and exits, including IPOs and M&A ("success" hereafter), and bankruptcies and business

[4]Pitchbook is a common tool to venture capitalists and has become more and more used in academia (e.g., Ewens et al., 2022; Jiang and Sohail, 2023).

closures ("fail" hereafter).[5] Pitchbook collects articles of incorporation filings from states like Delaware and California, and encode key contract terms from the financing rounds described in those documents (Ewens et al., 2022). Overall, Pitchbook has one of the best coverage and is one of the most accurate VC databases (Retterath and Braun, 2020). I limit my sample to companies established after 2000.

I complement Pitchbook data with several other sources. First, the acquisition prices for some M&A exits are not recorded in Pitchbook. I supplement these missing values by reviewing SEC filings (Form 8-K, 10-K, and 10-Q) of publicly traded acquirers, as they are required to report significant M&A transactions. Second, Pitchbook calculates IPO exit values by multiplying the number of outstanding common shares by the initial offering price, which often reflects strategic underpricing and may not represent the company's true value (Beatty and Ritter, 1986). To correct this, I replace the initial offering price with the closing price on the first IPO day, using data from Compustat. For firms listed outside North America, I obtain these closing prices from FactSet.

I collect all information on biotech and software companies that were founded between 2000 and 2022, headquartered in the US, and have raised VC or pre-VC rounds.

Biotech companies focus on drug discovery and the delivery of pharmaceuticals or biotechnology, such as Amgen, Moderna and Eli Lilly. Software companies design and develop softwares, including application softwares (Adobe, Oracle), business/productivity softwares (IBM, Salesforce), operating systems (Apple, Microsoft), social/platform softwares (LinkedIn, Meta) etc.

After dropping companies with missing data and focusing on firms that raise at most 7 rounds, my final sample consists of 5,591 companies, with 2,378 in biotech and 3,213 in software. In total, 2,072 were successful (589 IPO and 1,483 M&A) and 3,519 failed.[6]

[5] Unlike in IO literature where "exit" typically refers to a firm's (usually undesirable) leave from the market, in the context of VC and startups, an "exit" refers to a liquidity event like an M&A or IPO where investors realize returns on their investments.

[6] Excluding startups with more than 7 rounds does not weaken my sample's representativeness, as they comprise less than 5% of all startups.

2.1 Data Description

I begin with an example to introduce key concepts in this industry. See Appendix A for additional examples from the data.

Example 1. *ABC* is an app offering online English classes for kids. It started with 10 million outstanding shares, all owned by founders and employees. On Jun 19, 2008, it raised \$5 million in the first VC round and issued 5 million new shares to this round's investors.

Based on these terms, **price per share** equals to $\frac{\$5 \text{ million}}{5 \text{ million shares}} = \$1$ representing the dollar value of each share traded in this round; **equity ownership** is defined as the proportion of shares held by this round's investor relative to the total number of outstanding shares, representing a claim to a portion of the company's returns. It is calculated as $\frac{5 \text{ million shares}}{15 \text{ million shares}} = 33\%$; **PreVal** (pre-money valuation) equals to price per share multiplied by the total shares *before* the new issuance — \$1 per share $\times$ 10 million shares = \$10 million (or equivalently, $\frac{\$5 \text{ million}}{33\%} - \$5$ million), representing the "price" of the entire company in this round; **PostVal** (post-money valuation) is the same as PreVal except that it's *after* the new issuance, calculated as \$1 per share $\times$ 15 million shares = \$15 million (or equivalently, $\frac{\$5 \text{ million}}{33\%}$).

On Nov 24, 2009, *ABC* raised \$10 million in the second round and issued 5 million new shares to round 2 investors. Now, price per share, PreVal and PostVal are respectively \$2, \$30 million and \$40 million. Round 2 investor's equity ownership is 25%, while round 1 investor's equity ownership is *diluted* from 33% to 25% because more shares were issued. However, round 1 investor's estimated return still increases because the rise in *ABC*'s valuation offsets the dilution in the ownership.

Eventually, on Apr 29, 2010, *ABC* was acquired for \$100 million. Round 1 and 2 investors' returns were both 25% $\times$ \$100 million = \$25 million, and their multiple of money (MoM) were 5 and 2.5, respectively.

Section A and B in Appendix B describe the variables observed at startup and startup-round level, respectively. *ExitValue*, *CapitalAmount*, *PreVal*, and *PostVal* are converted to millions of USD as of September 2023. One limitation of the data is that Pitchbook does not provide exit values for failed companies, so I assign a value of 0 in such cases. Appendix C.1

and C.2 present summary statistics at startup-round and startup level, respectively. In general, as a startup proceeds to later rounds, *CapitalAmount* and *PostVal* tend to increase, while *EquityOwnership* tends to decrease. Additionally, *Patent*, *PostVal* and *ExitValue* are all highly skewed, indicating that while most startups are average, few are exceptional.

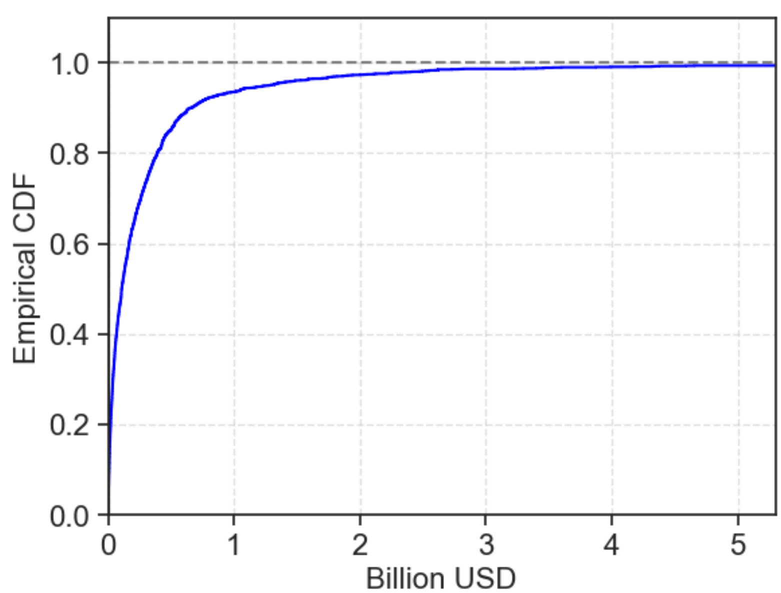
2.2 Empirical Observations

2.2.1 Skewed Exit Value

In the VC-startup context, exit value is a key object of interest, as it directly affects VC returns. By the nature of innovation, neither entrepreneurs nor investors can identify the biggest success ex-ante (e.g., Ewens et al., 2018; Kerr et al., 2014). Hence, the distribution of exit value defines how risky VC investments are and affects how well this market functions.

Figure 1 presents the distribution of exit values in the data. While 90% of startups have exit values below $1 billion, the top 1% reach as high as $31 billion, highlighting that VC investment combines skewed ex post returns with an inability to identify the biggest successes ex ante.

Figure 1: **Empirical CDF of exit value**



Note: Only IPO and M&A startups are included. All values are in Sep 2023 USD.

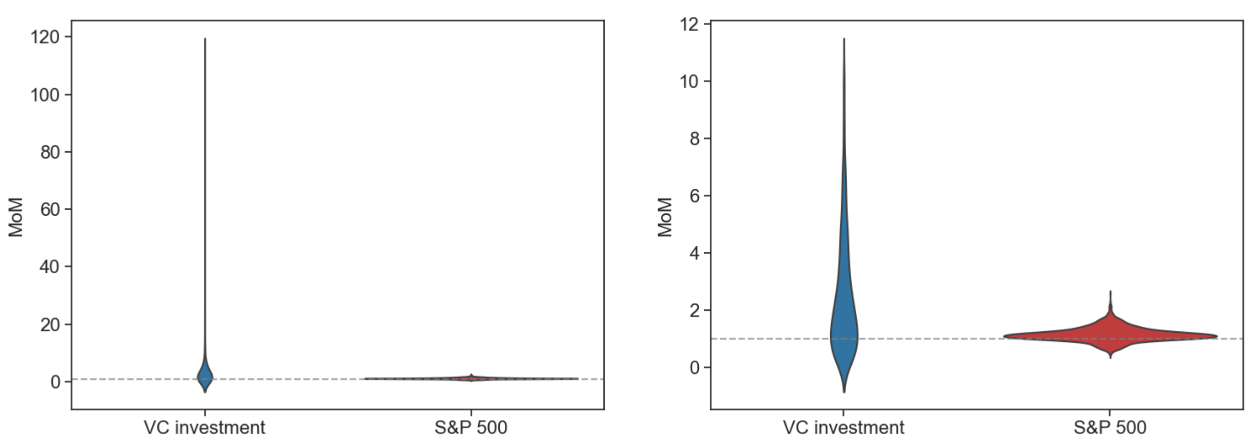
2.2.2 Risk in VC Investments

To evaluate the performance of VC investments, I follow Kaplan and Schoar (2005) by comparing a VC investment to an equivalently timed investment in a public market benchmark, such as the S&P 500 index.

I use Example 1 for illustration. Round 1 and 2 investors each invested $5 million and $10 million, respectively, and both received $25 million in return. Thus, the multiple of money (MoM) — a common metric for measuring VC investment performance —is calculated as $\frac{25+25}{5+10} = 3.33$. Had round 1 investor invested the same amount in S&P 500 index on the same day and sold his holdings at the time of acquisition, and similar for round 2 investor, the MoM would have been 1.02.

I replicate this exercise for each startup in my sample and compare the MoM of VC investment to its public market counterfactual. As shown in Fig. 2, VC investments tend to have higher expected returns and greater volatility compared to the public market counterfactuals, highlighting their high risk high reward nature.

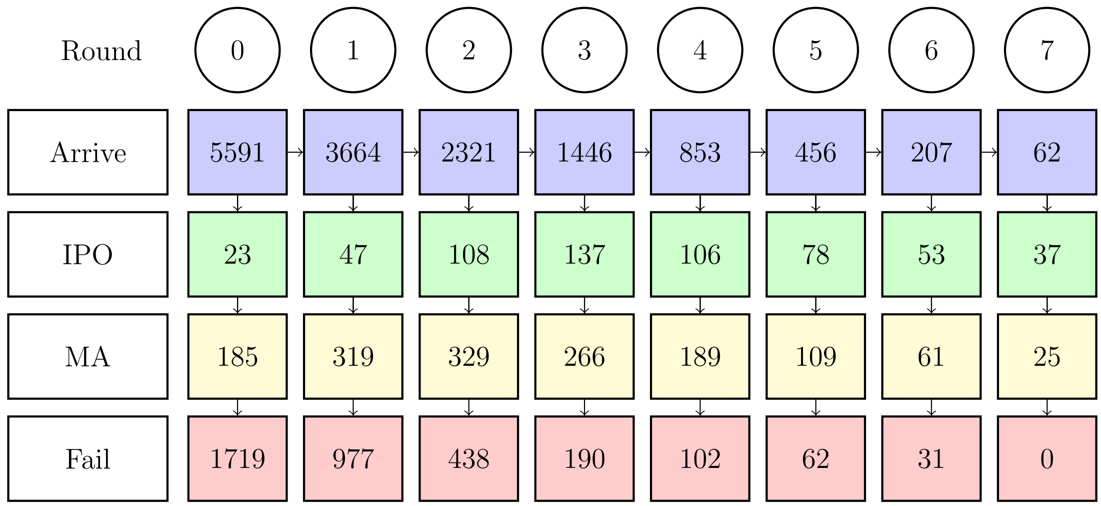
Figure 2: **MoM of VC investments and public market counterfactuals**



Note: Only IPO and M&A startups are included. Outliers (below 5 percentiles or above 95 percentiles) are kept in the left figure but removed from the right figure.

In response, venture capitalists phase their investments over multiple rounds, with each round generating some (new) information about the startup's exit value (Ewens et al., 2018), helping future investors make more informed decisions. Figure 3 summarizes the multi-round financing and exit of startups in my sample, based on which I derive the empirical transition probability matrix in Table 1.

Figure 3: **Multi-round financing and exit**



Note: The circles at the top row represent VC rounds, the rectangles on the leftmost column represent startup status in a given round. For example, among the 5,591 established startups, 3,664 raised the 1st VC round, 23 and 185 went IPO and got acquired directly, and the rest failed without raising any VC rounds. Round 0 = established.

Table 1: **Empirical transition probability matrix**

	Next state			
Current state	Next round	IPO	M&A	Fail
Established	0.66	0.00	0.03	0.31
Round 1	0.63	0.01	0.09	0.27
Round 2	0.62	0.05	0.14	0.19
Round 3	0.59	0.09	0.18	0.13
Round 4	0.53	0.12	0.22	0.12
Round 5	0.45	0.17	0.24	0.14
Round 6	0.30	0.26	0.29	0.15
Round 7	0.00	0.60	0.40	0.00

Note: The leftmost column shows the current states, including newly established and VC round 1 to 7. Each row represents the probability of transitioning from the current state to each possible next state, with the probabilities summing to 1. Take the first row for example. Among all the established startups, 66% secured the first VC round, 3% got acquired directly, 31% failed without raising any VC rounds.

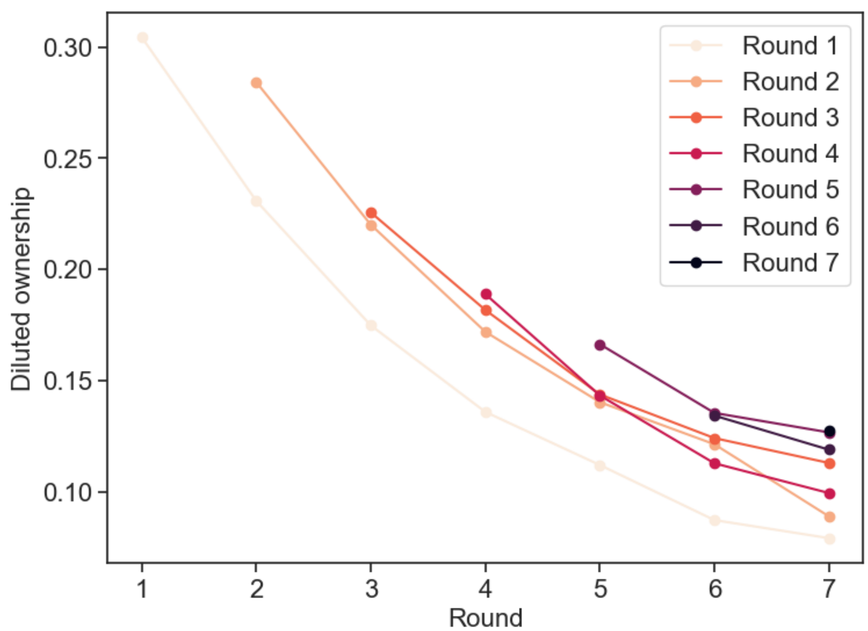
Each row in Table 1 shows the probability of transitioning from the current state to each of the possible next state. As startups progress to later rounds, success rates increase and failure rates decrease, as low-quality startups are continuously filtered out, leaving a stronger pool of companies.

In summary, VC investments are high risk high return compared to public market investments, but the risk decreases over rounds.

2.2.3 Dilution in Equity Ownership

As illustrated in Example 1, one of the features of VC financing is that equity ownership dilutes as a startup raises additional capital and issues more shares. Equity dilution is important because it determines how the returns from startups are shared across VCs at different stages.

Figure 4: **Diluted equity ownership**



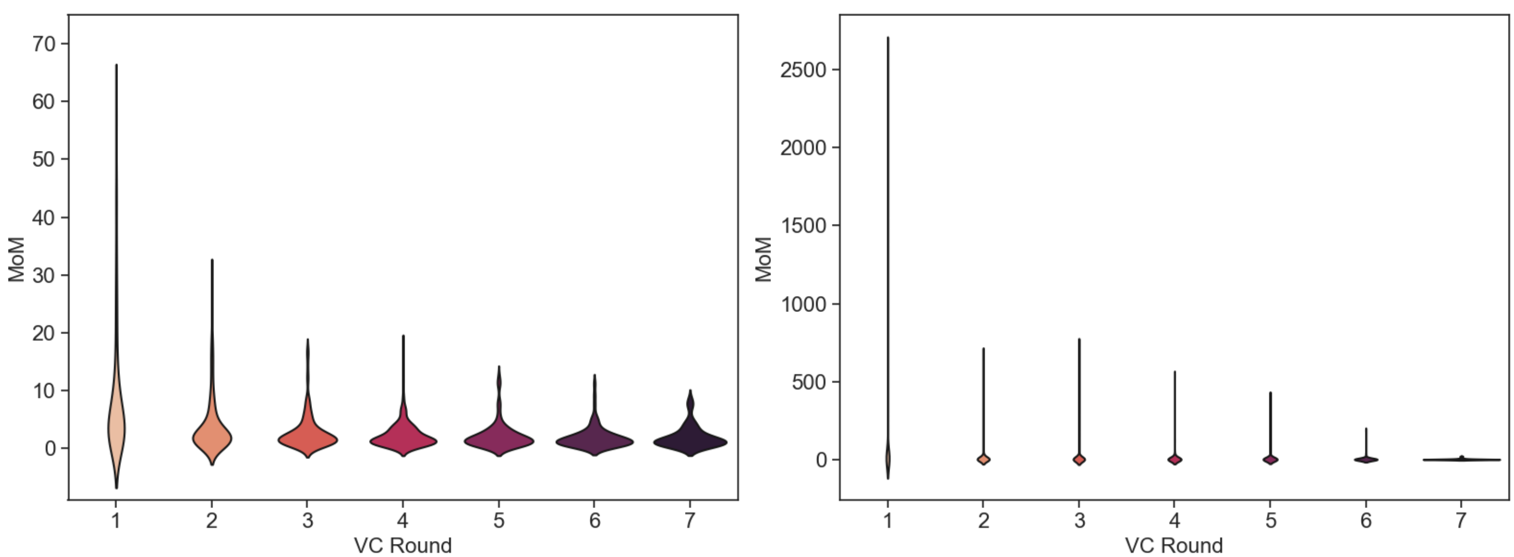
Note: This figure presents the median of each round investors' equity ownership in each future round after dilution. For instance, at the median, round 1 investors hold 30% equity at the close of round 1, but by round 7, their ownership is diluted to less than 5%.

As shown in Fig. 4, equity ownership dilution is significant, especially for early-stage investors. For instance, round 1 investors hold a median equity ownership of 30% by the end of the 1st round, but as new shares are issued in subsequent rounds, their ownership is diluted to less than 5% by round 7.

2.2.4 Round-specific Return

Despite significant dilution, early investors still earn more, provided that they invest in a company that is later proved to be successful. Fig. 5 presents the MoM returns to each round investors, conditional on investing in an ex-post IPO company. On average, early investors achieve higher MoMs than later investors. One reason could be intrinsic to the evolving information structure: due to significant uncertainties about startup quality in the early rounds, investors can buy equity shares at reduced prices; as the startup matures and moves to later rounds, its potential becomes more recognized, driving up equity prices and (over) offsetting the effects of dilution . Hence, an investor would have been better off investing early if he had known the ex post big success.

Figure 5: **MoM of IPO startups**



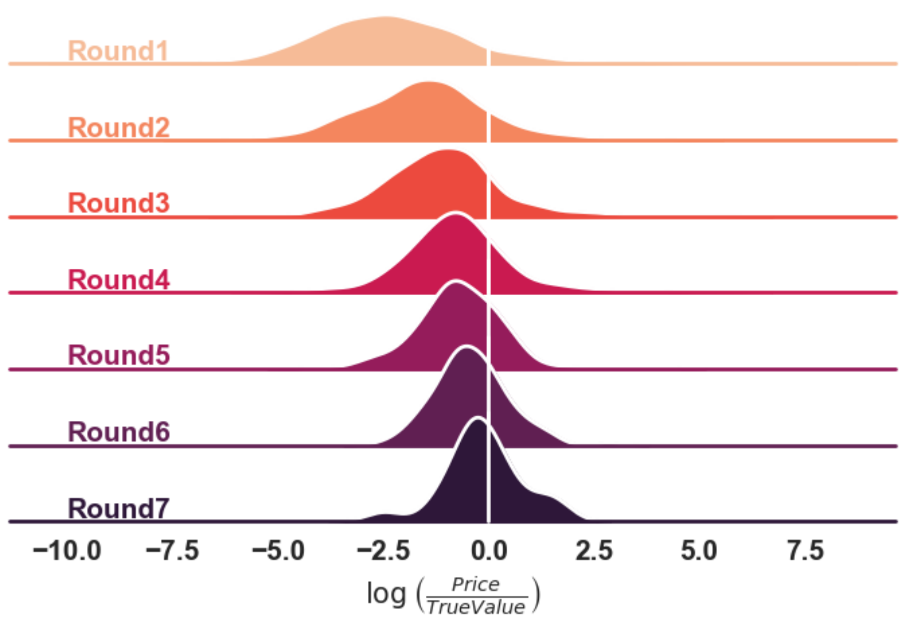
Note: Only IPO companies are included. The left (right) figure excludes (includes) outliers (below 5 percentiles or above 95 percentiles).

2.2.5 "Price" converging to true value

Venture capitalists learn about a startup's exit value from the information released in each round, which affects their "willingness-to-pay". Meanwhile, VCs compete with each other to secure the investment opportunity. Hence, the observed "price" (i.e., PreVal and PostVal) is an outcome of both VCs' learning about the true value and imperfect VC competition, the two effects cannot be disentangled without a model.

I use $\log\left(\frac{\text{PreVal}_k}{\text{ExitValue}}\right)$ to quantify the gap between the "price" paid by round k investor and the startup's true value. Failed startups are excluded due to missing exit values. Thus, the sample represents stronger companies. Initially, there is a large gap between the paid price and the true value, suggesting substantial uncertainties in startup quality at the outset. However, as the startup grows and releases more information, the price gradually converges toward its true value. Additionally, the gap is approximately normal for all rounds.

Figure 6: **Gap between price and true value**



Note: The horizontal axis is $\log\left(\frac{\text{PreVal}_k}{\text{ExitValue}}\right)$. For example, if a startup raised two rounds with PreVal at \$10 million and \$20 million each, and was later acquired at \$100 million, then this measure is $-\log 10$ and $-\log 5$ for each round. The white vertical line at 0 is when the paid price perfectly coincides with the true value. Only IPO and M&A companies are included.

2.2.6 Lead VC

When funding a startup, it's rare for a single VC to fund it throughout. Typically, there are several VCs, and each leads a different stage. One reason is that VC funds face a timeline constraint of 10 to 12 years, with the first 5 to 6 years deploying all capitals and the remaining years harvesting the returns, usually not enough time for a startup to reach a scale that attracts buyers or go public. Other reasons include 1) VC fund size is usually too small for funding a startup entirely, and 2) VCs tend to diversity investments, usually no single project takes up more than 60% of a fund.

To examine this pattern in the data, I count the number of rounds each VC has lead in each startup, and find that nearly 90% VCs lead only one round of investment in a startup.[7] So, in the model, I assume a new VC in each round.

Table 2: **Lead VC**

	1	2	3	4	5	6
Share	86.31%	10.44%	2.25%	0.69%	0.20%	0.07%

Note: This table summarizes the number of rounds a VC leads for a given startup. Nearly 90% of VCs lead just one round, although some VCs lead as many as 6 rounds.

2.2.7 Startup Heterogeneity

Next, I ask whether startups following different paths are inherently different. To address this, I group startups based on their exits and the number of VC rounds raised, then I compare the characteristics of each group.

Table 3: **Startup Heterogeneity**

Round		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Patents	*IPO*	6.91	27.68	27.23	23.57	33.59	53.92	152.25	77.54
	MA	2.12	2.28	2.98	5.12	9.2	12.04	16.13	14.2
	Fail	0.26	0.76	3.3	2.61	4.73	5.32	7.77	NA
Total Capital	*IPO*	0	29.84	74.15	93.77	103.72	151.36	224.95	208.33
	MA	0	12.25	26.29	47.5	82.04	94.3	128.53	141.54
	Fail	0	3.41	12.35	27.5	45.41	53.3	86.58	NA
Exit Value	*IPO*	203.03	308.25	361.46	437.02	399.54	1023.98	1947.91	596.47
	MA	147.52	129.58	193.93	249.69	320.28	440.27	595.18	515.3

Note: This table presents the mean of patent count, total capital raised from VC and exit value of startups in each group. Startups are grouped by their exit channel and number of rounds. E.g., 23.57 is the mean patent count among startups that went IPO after raising 3 VC rounds. Total capital raised and exit value are in million USD. Round 0 = established.

[7]In 38% of cases where Pitchbook does not identify the lead VC, I follow Ewens et al. (2022) and assume the lead investor to be the VC with the most experience, measured by the number of VC investments made up to the time of the current deal.

Table 3 shows that startups following different trajectories are heterogeneous in patents, total capital raised, and exit value.

Table 4 shows the correlations between patent, total capital and exit value, all of which are positively correlated with each other.

Table 4: **Correlation**

	Patents	Total Capital	Exit Value
Patents	1	.13	.13
Total Capital	.13	1	.54
Exit Value	.13	.54	1

Note: Exit values for failed companies are fill with 0.

3 Model

I consider a multi-round game between a startup and several VCs. The unit of observation is a startup, and I follow it throughout its life cycle. A startup of unknown (to both entrepreneur and investors) quality needs to secure several rounds of VC funding to be successful. In each round, a (new) set of VCs compete in equity shares to invest based on their information about the startup. VCs also form expectations about involvement of VCs in the future rounds before making their investment decisions. This leads to a dynamic, finite horizon decision problem.

Every round the amount of capital to be raised by the startup is public information. A set of VCs is matched, and each obtains a private signal on the unobserved quality. Then, they compete by making offers on how the future returns are shared. I model the competition "as if" it were an English auction with pure common value, and VCs "bid" on equity ownership.[8] The VC that outbids the competitor by offering at least as much equity to the startup as the competitor is selected. If the startup fails to raise capital, it fails and exits the market, and the capital already invested is sunk. If successful, it proceeds to the next round, and the VCs' signals in the current round become public. There is evidence that

[8]Throughout the paper, I assume that VCs 1) do not add values other than provide capital; and 2) only bid on equity ownership but not other contract dimensions. See Section 8 for discussions on this assumption.

late-round investor candidates in practice can assess the information early-stage investors used in their decisions. For example, during due diligence, prospective investors are typically provided with historical financial reports, product details, vendor and customer contracts, and management team backgrounds that have been reviewed by earlier investors (Ollar et al., 2021; Sannikov et al., 2016). In the next round, another set of VCs is matched, and the game iterates until the last round, after which the exit value is realized, and returns are shared.

English auction with pure common value delivers a tractable way of modeling the bargaining process while produces similar patterns as multi-lateral bargaining (McAfee and Vincent, 1997). Here bid is the equity ownership, and VCs favor larger equity ownership as it grants higher claim on returns, while startups prefer to offer less to the VCs.

I follow Milgrom and Weber (1982) and model the auction as follows. At the start, all bidders are active at an arbitrarily high equity ownership. As the equity ownership decreases, bidders drop out one by one. In a symmetric monotone equilibrium, the equity ownership at which each bidder drops out reveals her true signal, thus the bidders who are still active will incorporate the new information and update their bids accordingly. Any dropped-out bidder cannot be re-actived. The auction ends when only one bidder is left. She wins and is paid the equity ownership at which the second highest bidder quits.

I introduce some notations. Let K be the number of rounds a startup needs to complete to be successful. K is an exogenous random variable with discrete support $\{0, 1, ..., \bar{K}\}$, where $\bar{K}$ is the maximum possible round for raising funds. The realization of K is not observed by either the startup or the VCs ex-ante.[9] Let Z be a vector of round-unvarying observed characteristics of the startup, such as its sector (e.g., biotech or software), the number of assigned patents etc. In round $k \in \{1, ..., \bar{K}\}$, let $d_k \in \mathbb{R}_+$ be the amount of capital to be raised by the startup, it only becomes public at the beginning of that round, neither startup nor VCs know it ex-ante. To simplify the notations, I use $\boldsymbol{X}_k = (d_1, ..., d_k, Z)$ to denote all the observed characteristics of the startup in round k. $\boldsymbol{X}_K$ is exogenous and may be correlated with K, I assume that $(K, \boldsymbol{X}_K)$ is jointly distributed as $F(\cdot)$. Additionally, δ is the discount factor between two adjacent rounds.

Let ξ be the unobserved quality of a startup, capturing the hidden features of a startup

[9]This captures the randomness in when an acquisition opportunity might arise and how long it takes for a startup to reach the scale needed to go public.

that neither investors nor entrepreneur know ex-ante. Information on ξ only arrives with investment into the project. Some examples of the unobserved quality include the validity of a particular technology or product, and the potential market demand.[10] I make the following assumption.

Assumption 1. $\xi \sim N(0, \sigma_\xi^2)$, *and is independent of* $\boldsymbol{X}$ *and* K.

Under Assumption 1, agents cannot infer the value of ξ from observed characteristics $\boldsymbol{X}_k$ or current round k. This assumption is widely used to model the learning process that I will introduce later.[11]

Timeline. Fig. 7 shows the timeline. This game is composed of at most K rounds. At the beginning of round $k \in \{1, ..., K\}$, the startup samples the amount of capital to be raised in this round, d_k, from a known distribution, and announces it. Then a set of (new) VCs $\mathcal{J}_k$ is matched. They all observe $\boldsymbol{X}_k$ and each samples a private signal $s_{j,k}$ on the unobserved quality ξ. Then they bid on equity ownership in an English auction with pure common value. If the startup fails to raise capital, it quits, and the capital already invested is sunk; If successful, a winner is selected, the winning bid a_k is determined, and all private signals for this round, $\boldsymbol{s}_k$, are publicly disclosed. This process iterates until the last round, after which the exit value W_K is realized and shared.

Exit Value. A startup's exit value depends on both the observed characteristics and unobserved quality. The exit value after raising K rounds is

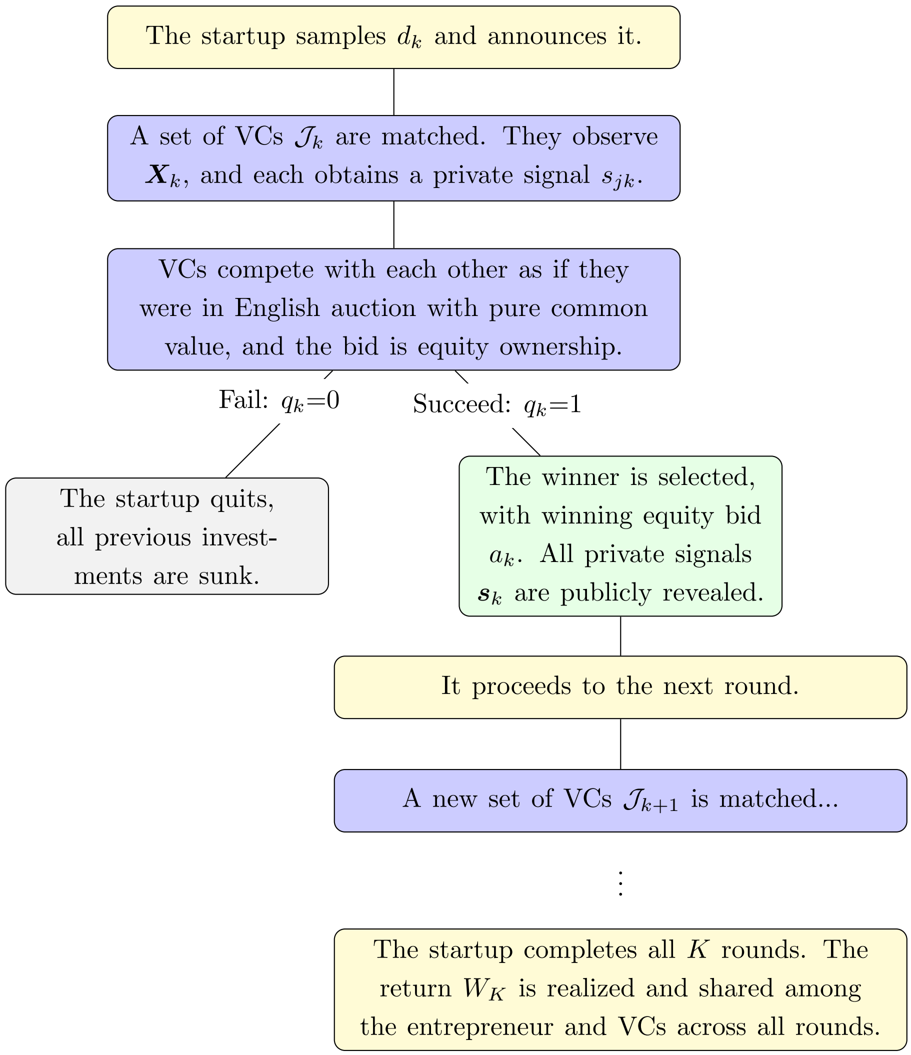
$$W_K = m_K(\boldsymbol{X}_K) \exp(\xi), \tag{1}$$

where $m_K(\boldsymbol{X}_K)$ is an arbitrary function supported on $\mathbb{R}_+$ and captures the contribution from observed characteristics, while $\exp(\xi)$ captures the contribution from unobserved quality. To

[10] For example, when Airbnb first came out in 2009, according to Fred Wilson of Union Square Ventures, it was an idea of "marketplace for air mattresses on the floors of people's apartments", and it was hard to evaluate the demand for this marketplace when it was only an idea. Investors (and entrepreneurs) can only learn about the demand by making the investment to build the platform and then see how much demand it attracts. See more details at https://avc.com/2011/03/airbnb/.

[11] The normality assumption on ξ can be generalized to Gaussian mixture distribution such that the key results of this paper still hold.

Figure 7: **Timeline**



Note: This schematic presents the timeline of the multi-round financing game. Yellow and blue boxes are startup and VC actions, respectively.

capture the skewed returns (see Fig. 1), I let ξ enter Eq. (1) through its exponential form. Furthermore, $m_K(\boldsymbol{X}_K)$ and $\exp(\xi)$ are multiplicatively separable, so that $\exp(\xi)$ is the unobserved marginal productivity of a "composite" observed characteristic that needs to be learnt across rounds.

Signal. VCs don't observe ξ directly, but can learn its value from their private signals. At round k, a matched VC $j \in \mathcal{J}_k$ receives a private, noisy and unbiased signal

$$s_{j,k} = \xi + \epsilon_{j,k}, \tag{2}$$

where

Assumption 2. *For any $k \in \{1, ..., \bar{K}\}$ and $j \in \mathcal{J}_k$, $\epsilon_{j,k} \overset{iid}{\sim} N(0, \sigma_k^2)$ and is independent of ξ, $\boldsymbol{X}$ and K.*

The mean zero assumption is needed for unbiased signals, the normality assumption is for tractability, and the round-specific variance allows signal accuracy to vary across rounds. For notation convenience, I define $\boldsymbol{\sigma} = (\sigma_k)_{k=1}^{\bar{K}}$ as the vector of standard deviations of each round's signal noise, $\boldsymbol{s}_k = (s_{j,k})_{j \in \mathcal{J}_k}$ as the set of signals generated at round k, and $\boldsymbol{S}_k = (\boldsymbol{s}_1, ..., \boldsymbol{s}_k)$ as the set of signals generated before (including) round k.

Ex-post payoff. When a company raises more capital, it issues additional shares, diluting the equity ownership of existing shareholders. Let a_k be the winning equity ownership (before dilution) in round k. When the startup exits after K rounds, ex-post payoff to round k investor depends on exit value W_K, her ownership a_k and dilution,

$$\underbrace{R_k}_{\text{ex-post payoff}} = \underbrace{a_k}_{\text{ownership}} \times \prod_{k'=k+1}^{K} \underbrace{(1 - a_{k'})}_{\text{dilution from round } k'} \times \underbrace{W_K}_{\text{exit value}}. \tag{3}$$

3.1 Equilibrium

I focus on symmetric equilibrium, where the identities of VCs do not matter. In each round, for any bidding strategies adopted by VCs, the funding rules can be summarized by a mapping $\phi_k := (\mathfrak{q}_k, \mathfrak{a}_k)$ that maps each $(\boldsymbol{X}_k, \boldsymbol{S}_k)$ in the characteristics and signals space to a pair of funding decisions (q_k, a_k), where $q_k \in \{0, 1\}$ is whether the startup secures round k investment ($q_k = 1$ if yes and 0 if no), and $a_k \in [0, 1]$ is the equity ownership to the winning investor of this round. I let $\boldsymbol{\phi}_{k+1} := (\phi_{k+1}, ..., \phi_{\bar{K}})$, $\boldsymbol{\mathfrak{q}}_{k+1} := (\mathfrak{q}_{k+1}, ..., \mathfrak{q}_{\bar{K}})$ and $\boldsymbol{\mathfrak{a}}_{k+1} := (\mathfrak{a}_{k+1}, ..., \mathfrak{a}_{\bar{K}})$, which come in handy later.

Ex-ante payoff. Since the startup's exit value is not realized when VCs bid for the investment opportunity, their bidding decisions are based on the expected payoff. Given the observed characteristics $\boldsymbol{X}_k$, the signals on the unobserved quality $\boldsymbol{S}_k$, and the future funding rules $\boldsymbol{\phi}_{k+1}$, the *maximum* expected payoff to round k investor is

$$V_k(\boldsymbol{X}_k, \boldsymbol{S}_k; \boldsymbol{\phi}_{k+1}) = \mathbb{E}\left[\delta^{K-k} \prod_{k'=k+1}^{K} (1 - a_{k'}) q_{k'} W_K \middle| \boldsymbol{X}_k, \boldsymbol{S}_k; \boldsymbol{\phi}_{k+1}\right], \tag{4}$$

The expectation is with respect to the number of ultimate rounds K, future capital requirements $d_{k+1}, ..., d_K$, signals $\boldsymbol{s}_{k+1}, ..., \boldsymbol{s}_K$, binary funding choices $q_{k+1}, ..., q_K$, winning equity ownerships $a_{k+1}, ..., a_K$, and the true quality ξ. In other words, this expectation is with respect to all possible future trajectories of the startup given the current information set and future funding rules.

The dimensionality of $\boldsymbol{S}_k$ increases rapidly with k. Under the independence and normality assumptions of quality and signal noises, $\boldsymbol{S}_k$ is only relevant for inferring the quality ξ. The posterior distribution of ξ is given by

$$\xi \mid \boldsymbol{S}_k \sim N(\mu_k, \tau_k^2), \tag{5}$$

where μ_k is a weighted sum of $\boldsymbol{S}_k$, and τ_k is independent of $\boldsymbol{S}_k$ (see Appendix D for the expressions of μ_k and τ_k^2). Since $\boldsymbol{S}_k$ impacts the posterior distribution only through μ_k, μ_k is a sufficient statistics for $\boldsymbol{S}_k$. Hence, Eq. (4) becomes

$$V_k(\boldsymbol{X}_k, \mu_k; \boldsymbol{\phi}_{k+1}) = \mathbb{E}\left[\delta^{K-k} \prod_{k'=k+1}^{K} (1 - a_{k'}) q_{k'} W_K \middle| \boldsymbol{X}_k, \mu_k; \boldsymbol{\phi}_{k+1}\right]. \tag{6}$$

Bidding strategy. VC j's strategy at round k specifies the equity level she should drop out as a function of the current round k, the observed characteristics $\boldsymbol{X}_k$, previous signals

$\boldsymbol{S}_{k-1}$, her own signal $s_{j,k}$, and the interim information revealed during the auction, including the number of bidders who have dropped out and the equity level at which they exited.

At the beginning of round k, each bidder doesn't know the exact value of μ_k because other bidders' signals are private, and thus can only infer μ_k from the revealed signals of previous rounds and her own signal. However, as bidders start to drop out, each equity level at which a competing VC withdraws discloses her true signal (because the strategies are revealing in equilibrium), based on which each remaining VC updates her guess of μ_k and uses it to determine the next dropping-out threshold, assuming no other VCs exit in the middle. See Milgrom and Weber (1982) for formal presentation of the bidding strategy.

Funding decision. I characterize the equilibrium binary funding rule $\mathfrak{q}_k^*$ and the equilibrium winning equity ownership rule $\mathfrak{a}_k^*$.

Proposition 1. *A VC in round $k \in \{1, ..., \bar{K}\}$ funds a startup with characteristics $\boldsymbol{X}_k$ and μ_k that needs capital d_k if $V_k^*(\boldsymbol{X}_k, \mu_k) \geq d_k$. In other words, the optimal funding rule in the VC market $\mathfrak{q}_k^*$ is given by*

$$\mathfrak{q}_k^*(\boldsymbol{X}_k, \mu_k) = \mathbb{1}\{V_k^*(\boldsymbol{X}_k, \mu_k) - d_k \geq 0\}, \tag{7}$$

where $V_k^*(\boldsymbol{X}_k, \mu_k) \coloneqq V_k(\boldsymbol{X}_k, \mu_k; \boldsymbol{\phi}_{k+1}^*)$. V_k is defined in Eq. (6) and $\boldsymbol{\phi}_{k+1}^*$ is the optimal funding decisions in all future rounds.

V_k^* is the maximum expected payoff to round k investor when all future rounds follow the same funding rules as defined in Eq. (7) and Eq. (8). Proposition 1 provides the necessary and sufficient condition for a startup to secure funding at round k. Intuitively, there should be at least one VC whose *maximum* expected payoff is larger than the necessary capital injection, akin to the entry cost in entry games. Hence, Eq. (7) determines the trajectory of a startup given its characteristics and signals.

Proposition 2. *In equilibrium, the winning equity ownership rule at round $k \in \{1, ..., \bar{K}\}$, $\mathfrak{a}_k^*$, satisfies*

$$\mathfrak{a}_k^*(\boldsymbol{X}_k, \hat{\mu}_k) \times V_k^*(\boldsymbol{X}_k, \hat{\mu}_k) = d_k, \tag{8}$$

where $\hat{\mu}_k$ *is the posterior mean of* ξ *given* $\boldsymbol{S}_{k-1}, s_{(J_k),k}, ..., s_{(3),k}, s_{(2),k}, s_{(2),k}$*, where* $s_{(J_k),k} \leq ... \leq s_{(1),k}$*. Moreover,* $V_k^*(\boldsymbol{X}_k, \mu_k)$ *is strictly increasing in* μ_k *for any* $\boldsymbol{X}_k$ *and* $k \in \{1, ..., \bar{K}\}$*.*

When only two bidders are left, all other signals (except for the top two) have been revealed. In Eq. (8), $\hat{\mu}_k$ represents the second-highest bidder's guess of μ_k conditional on all the revealing signals and *assuming* that the other active bidder holds the same signal as hers. Eq. (8) implies that the winning equity bid is such that the second-highest bidder's expected net payoff is driven down to 0 *assuming* that the other bidder receives her own signal. The formal proof of Proposition 2 is deferred in Appendix E, and it builds on Milgrom and Weber (1982).

3.2 Efficiency Analysis

Early-round investments can generate informational spillover that benefits late-round VC, but early-round VCs make decisions based on their private returns, potentially leading to positive "dynamic information externality". To examine the associated efficiency implications in the VC market, I consider a benchmark in which a planner who faces the same information as the VC investors chooses at each round whether to fund the startup, with the goal of maximizing total net payoff. I derive the planner's strategy and then compare it with the equilibrium in Section 3.1.

At round k, like the VC investors, the planner doesn't observe the true quality ξ and only knows $(\boldsymbol{X}_k, \mu_k)$. If he makes the investment, the expected net return (before investing capital d_k) is

$$\tilde{V}_k(\boldsymbol{X}_k, \mu_k; \mathbf{q}_{k+1}) = \mathbb{E}\left[\delta^{K-k} \prod_{k'=k+1}^{K} q_{k'} W_K - \sum_{k'=k+1}^{K} \delta^{k'-k} \left(\prod_{n=k+1}^{k'} q_n\right) d_{k'} \middle| \boldsymbol{X}_k, \mu_k; \mathbf{q}_{k+1}\right]. \quad (9)$$

Eq. (9) is different from Eq. (6) in two aspects: 1) there is no dilution from future rounds because the planner only cares about the *total* net return, hence only $\mathbf{q}_{k+1}$ but not $\boldsymbol{\alpha}_{k+1}$ enters $\tilde{V}_k$; 2) the planner also takes into account all the (discounted) possible future capital to inject, whereas VC investors only care about their own cost and do not consider future investors' costs.

Proposition 1'. *A planner in round* $k \in \{1, ..., \bar{K}\}$ *funds a startup with characteristics*

$\boldsymbol{X}_k$ and μ_k that needs capital d_k if $V_k^\dagger(\boldsymbol{X}_k, \mu_k) \geq d_k$. In other words, the planner's optimal funding rule $\mathfrak{q}_k^\dagger$ is

$$\mathfrak{q}_k^\dagger(\boldsymbol{X}_k, \mu_k) = \mathbb{1}\{V_k^\dagger(\boldsymbol{X}_k, \mu_k) - d_k \geq 0\}, \tag{10}$$

where $V_k^\dagger(\boldsymbol{X}_k, \mu_k) = \tilde{V}_k(\boldsymbol{X}_k, \mu_k; \mathfrak{q}_{k+1}^\dagger)$. $\tilde{V}_k$ is defined in Eq. (9) and $\mathfrak{q}_{k+1}^\dagger$ is the planner's optimal funding rules in all future rounds.

$V_k^\dagger$ is the expected net return (before injecting d_k) to the planner if he follows the same funding rules as defined in Eq. (10) in all future rounds. This is a cutoff strategy such that the planner will only fund the startup if the expected net payoff is non-negative.

Next, I compare the planner's strategy to the equilibrium in the VC market.

Proposition 3. *For any $(\boldsymbol{X}_k, \mu_k)$, $V_k^\dagger(\boldsymbol{X}_k, \mu_k) > V_k^*(\boldsymbol{X}_k, \mu_k)$.*

The proof is provided in Appendix F. Proposition 3 implies that for any startup, the planner would expect a higher return than the VC investors. Consequently, some startups, which the planner would fund, fail to secure financing in the VC market, suggesting "under"-investment in the VC market.

This inefficiency stems from dynamic information externalities. Consider an investor investing in an early round of a startup. This investment generates information that can be costlessly observed by later-stage investors, helping them make more informed investment decisions. However, early investors do not get fully paid for this — early investment pushes the project to the next stage and allows late-stage investor to draw his own signal from a (presumably) more informative distribution; however, early investor doesn't know the exact signal and thus the true value of the later investor, hence there doesn't exist a mechanism that achieves Pareto efficiency (i.e., at any given stage, the project is funded if and only if the total expected net payoff from the project conditional on all the signal realization by that time point is non-negative) while satisfying incentive compatibility and individual rationality (Myerson and Satterthwaite, 1983). This creates an information externality where later investors free-ride on the information produced by early investors. Since individual investors are not able to internalize the full benefit of their investment, they may stop investing even when the overall benefit is still positive. Hence, the equilibrium is suboptimal from a social point of view.

This inefficiency can be mitigated by government-provided funds and grants for early-stage startups. For example, the U.S. government has launched programs such as Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) to support startups at early stages.

4 Identification

The parameters to be identified are $\Gamma = (\boldsymbol{m}, \sigma_\xi, \boldsymbol{\sigma}, F)$. Recall that $\boldsymbol{m} = (m_1, ..., m_{\bar{K}})$ is a vector of functions defined in Eq. (1) and represents the effect of observed characteristics on the exit value; σ_ξ is the standard deviation of quality ξ; $\boldsymbol{\sigma}$ is the standard deviation of the signal noise in each round; F is the joint distribution of rounds K and observed characteristics $\boldsymbol{X}_K$.

I observe $(\boldsymbol{q}_K, \boldsymbol{X}_K, \boldsymbol{a}_K, W_K)$, where $\boldsymbol{q}_K = (q_k)_{k=1}^K$ is a vector of binary funding choices; $\boldsymbol{a}_K$ is a vector of equilibrium (winning) equity ownerships; $\boldsymbol{X}_K = (Z, d_1, ..., d_K)$ is the observed characteristics; and W_K is the exit value. The data is subject to threefold right-censoring: (1) $q_k, \boldsymbol{X}_k$ are only observed if $q_1 = ... = q_{k-1} = 1$; (2) a_k is only observed if $q_1 = ... = q_k = 1$; and (3) W_K is only observed if $q_1 = ... = q_K = 1$. That is, the future path and exit value are observed for only those startups that secured funding, but the funding decisions are endogenous and depend on the startup's unobserved quality, leading to endogeneous sample selection. Furthermore, a startup can fail at different stages of financing, and the "errors" that govern "passing" or "failing" each stage — signals about quality in the VC-startup context, are serially correlated, leading to dynamic selection. In this setting, identification is challenging as the standard results based on independent shocks across time, e.g., Rust (1987), do not apply.

To overcome this challenge it, I adapt Heckman and Navarro (2007) to my setting and allow for more general time series dependence in the unobservables. I correct the dynamic selection by leveraging VC's funding choices to explicitly account for the investor-side selection effect on startup outcomes. Hence, my choice-adjusted estimates of the model primitives are free from the influence of factors that affect both financing choices and startup outcomes.

Identifying the full structural model requires independence variation in the arguments of the binary funding rules. This is achieved through assumptions of independence between unobserved quality and signal noise, as well as functional restrictions imposed by the model. More specifically, I have four model-implied conditions that, along with the data, can be used to identify the parameters:

$$\ln W_K = \ln\left(m_K(\boldsymbol{X}_K)\right) + \xi, \tag{11}$$

$$q_k = \mathbb{1}\{V_k^*(\boldsymbol{X}_k, \mu_k) \geq d_k\}, \tag{12}$$

$$a_k = \frac{d_k}{V_k^*(\boldsymbol{X}_k, \hat{\mu}_k)}, \tag{13}$$

$$\begin{pmatrix} \xi \\ \mu_1 \\ \vdots \\ \mu_{\bar{K}} \end{pmatrix} \sim N(\mathbf{0}, \Sigma). \tag{14}$$

Eq. (11) is Eq. (1) after taking the log, Eq. (12) and Eq. (13) are the equilibrium conditions shown in Eq. (7) and Eq. (8). Eq. (14) describes the serial dependence among quality ξ and the posterior mean of it in all rounds. Under Assumption 1 and 2, they are jointly normal. The covariance matrix depends on the the variances of both quality and signal noises, as well as the number of VCs matched in each round. See Appendix D for its mathematical expression.

Proposition 4 shows that the unobservable μ_k can be isolated from the observables $\boldsymbol{X}_k$ in the funding choice described by Eq. (12). This separation relies on the monotonicity of V_k^* in μ_k, as established in Proposition 2. Hence, the choice equation (12') has the same structure as in Heckman and Navarro (2007), allowing me to directly apply their results to my setting.

Proposition 4. *Under Assumption 1 and 2, for any $k \in \{1, ..., \bar{K}\}$, $\exists$ a unique h_k such that Eq. (12) implies Eq. (12'):*

$$q_k = \mathbb{1}\{\mu_k \geq h_k(\boldsymbol{X}_k)\}, \text{ where} \tag{12'}$$

$$h_k(\boldsymbol{X}_k) = V_k^{*-1}(d_k; \boldsymbol{X}_k). \tag{15}$$

To summarize, the goal is to identify Γ with a right-censored and dynamically selected sample $(\boldsymbol{q}_K, \boldsymbol{X}_K, \boldsymbol{a}_K, W_K)$, along with four model implied conditions Eq. (11)(12')(13)(14), where ξ,μ_k,$\hat{\mu}_k$ are unobserved, $\boldsymbol{V}^* = (V_k^*)_{k=1}^{\bar{K}}$ satisfies Eq. (6) and $\boldsymbol{h} = (h_k)_{k=1}^{\bar{K}}$ is the inverse of $\boldsymbol{V}^*$. In addition, $\boldsymbol{V}^*, \boldsymbol{h}, \Sigma$ are known up to Γ.

One identification challenge is that the number of VCs competing for a startup in each round is not observed in the data; only the winner's identity, if any, is observed. While identifying the parameters $(\boldsymbol{m}, \sigma_\xi, \boldsymbol{h})$ does not rely on this information, to identify the other parameters, I assume there are only two VCs in each round.[12] If in each round the number of VCs is fixed (but not necessarily two), $\boldsymbol{\sigma}$ is identified up to the number of VCs because σ_k affects μ_k only through the ratio $\frac{\sigma_k^2}{J_k}$. Thus, this assumption is a normalization of $\boldsymbol{\sigma}$. However, without this assumption, we cannot point-identify $\boldsymbol{V}^*$ but can only identify its bound. Furthermore, the lower bound collapses onto the upper bound of $\boldsymbol{V}^*$ when I impose this assumption, suggesting that I am overestimating $\boldsymbol{V}^*$ if this assumption fails.

I identify the parameters in three steps. In the first step, I identify $\boldsymbol{m}, \sigma_\xi, \boldsymbol{\sigma}, \boldsymbol{h}$, followed by $\boldsymbol{V}^*$ in the second step and F in the third step.

Identification of $\boldsymbol{m}, \sigma_\xi, \boldsymbol{\sigma}, \boldsymbol{h}$. The first step identifies $\boldsymbol{m}, \sigma_\xi, \boldsymbol{\sigma}, \boldsymbol{h}$ by directly applying the identification results from Heckman and Navarro (2007) to my setting. The identification relies on the independent variation in the observed characteristics and functional restrictions on the binary funding rules $\boldsymbol{h}$.

Proposition 5. *Given the following conditions, $\boldsymbol{m}, \sigma_\xi, \boldsymbol{\sigma}, \boldsymbol{h}$ are identified with Eq.* (11) (12')(14).

1. *Independence:* $(\mu_1, ..., \mu_{\bar{K}}, \xi) \perp (\boldsymbol{X}_1, ..., \boldsymbol{X}_{\bar{K}})$.

2. $(\mu_1, ..., \mu_{\bar{K}}, \xi)$ *are continuous random variables with zero mean and finite variance, and with support* $Supp(\mu_1) \times ... \times Supp(\mu_{\bar{K}}) \times Supp(\xi)$.

[12] The empirical auction literature has proposed either using the measurement error approach (An et al., 2010) or using several order statistics (Luo and Xiao, 2023; Song, 2004) for identification when the number of bidders is not observed. However, these approaches do not apply in my setting because first I observe only one bid per round, and second, these approaches apply only to independent private value auctions with static competition. There is no known result for common value auctions with dynamic competition.

3. *Full rank:* $Supp(\ln(m_{\bar{K}}(\boldsymbol{X}_{\bar{K}})), h_1(\boldsymbol{X}_1), ..., h_{\bar{K}}(\boldsymbol{X}_{\bar{K}})) = Supp(\ln(m_{\bar{K}}(\boldsymbol{X}_{\bar{K}}))) \times Supp(h_1(\boldsymbol{X}_1)) \times ... \times Supp(h_{\bar{K}}(\boldsymbol{X}_{\bar{K}}))$

4. *Inclusion:* $Supp(h_k(\boldsymbol{X}_k)) \supseteq Supp(\mu_k), \forall k = 1, ..., \bar{K}$.

To prove Proposition 5, I show that the listed conditions are satisfied in my model. Hence, I can evoke Theorem 2, the sufficient condition for identifying the outcome equations, choice equations and joint distribution of unobservables, in Heckman and Navarro (2007) to identify $\boldsymbol{m}, \sigma_\xi, \boldsymbol{\sigma}, \boldsymbol{h}$. The complete proof of Proposition 5 is provided in Appendix G.

Identification of $\boldsymbol{V}^*$. In the second step, I leverage the variation in equity ownership to identify $\boldsymbol{V}^*$.

Proposition 6. *Given $\boldsymbol{h}$, using model implied Eq. (13), $\boldsymbol{V}^*$ is identified.*

The reason that $\boldsymbol{V}^*$ cannot be directly identified by inverting $\boldsymbol{h}$ is that as is shown in Eq. (15), d_k also shows up in $\boldsymbol{X}_k$ of V_k^{*-1}, thus the full rank condition on covariates is violated. However, Eq. (13) implies $\hat{\mu}_k = V_k^{*-1}(\boldsymbol{X}_k; \frac{d_k}{a_k})$, where a_k provides extra variation to d_k in the second argument, restoring the full rank condition. So that I can explore the variation in a_k to identify $\boldsymbol{V}^*$. The complete proof is deferred to Appendix H.

Identification of F. The third step identifies F by exploring the variation in the observed characteristics, accounting for dynamic selection based on the binary funding rules $\boldsymbol{h}$ identified in the first step.

Proposition 7. *Given $(\boldsymbol{m}, \sigma_\xi, \boldsymbol{\sigma}, \boldsymbol{h}, \boldsymbol{V}^*)$, F is identified.*

Recall that F is the joint distribution of $\boldsymbol{X}_K$ and K. The proof is composed of two steps. In the first step, I show that $F(\boldsymbol{X}_K \mid K)$ is identified. Specifically, consider the subsample of startups that complete all K rounds, the distribution of their observed characteristics, $F(\boldsymbol{X}_K \mid K, q_1 = ... = q_K = 1)$ is directly identified from data. Moreover, for any $\boldsymbol{X}_K$, the probability of passing all rounds can be identified given $(\sigma_\xi, \boldsymbol{\sigma}, \boldsymbol{h})$. Hence, $F(\boldsymbol{X}_K \mid K)$ is identified using Bayes rule. In the second step, I identify $F(K)$ from $\boldsymbol{V}^*$. This is because $\boldsymbol{V}^*$ are expectations with respect to the number of rounds and thus contain information about the marginal distribution of K. The complete proof is provided in Appendix I.

5 Estimation

In this section, I discuss the parameterization of the model and the estimation methodology (simulated minimum distance estimator). The parameterization balances the dimensionality of the parameters and the desired richness in the structure of startup attributes, and the estimation algorithm handles complex and nonlinear problems with large search spaces while limiting computing time by parallelization.

5.1 Parameterization

Recall that the parameters to be estimated is $\Gamma = (\boldsymbol{m}, \sigma_\xi, \boldsymbol{\sigma}, F)$. To capture the heterogeneous value-add of different round's capital injection while limiting the number of parameters to estimate, I use the following parameterization for $\boldsymbol{m}$:

Assumption 3.

$$
\begin{aligned}
m_0(\boldsymbol{X}_0) &= Z\beta_0, \\
m_1(\boldsymbol{X}_1) &= Z\beta_0 + d_1\beta_1, \\
&\cdots \\
m_{\bar{K}}(\boldsymbol{X}_{\bar{K}}) &= Z\beta_0 + d_1\beta_1 + \cdots + d_{\bar{K}}\beta_{\bar{K}},
\end{aligned}
\tag{16}
$$

where $\boldsymbol{\beta} = (\beta_0, ..., \beta_{\bar{K}})$ and $\bar{K} = 7$. This "triangular" specification in Eq. (16) homogenizes the capital's value-add within the same round for startups that exit at different times.

I assume that there is no dependence between $\boldsymbol{X}_K$ and K, which would drastically enhance the computational burden. With this assumption, the distribution of K is directly identified from the data because the censoring is independent of K. This means that startups failing at each round have equal probability of exiting at each subsequent rounds (including the current round) had they been funded all the way through. For example, in Fig. 3, the 1,719 startups failing to raise round 1 are equally likely to end up in each of the seven rounds if they received funding throughout.

Furthermore, to reduce state space dimensionality, I assume a Markov transition process on capital requirements:

Assumption 4.

$$d_{k+1} \sim Exp\left(\frac{1}{\lambda_0^k + \lambda_1^k d_k + \lambda_2^k \bar{W}_k}\right). \tag{17}$$

This distributional specification allows d_{k+1} to depend on all previous characteristics $\boldsymbol{X}_k$ through d_k and $\bar{W}_k$, where $\bar{W}_k = Z\beta_0 + d_1\beta_1 + \cdots + d_k\beta_k$ is the cumulative value from observed characteristics. For notation convenience, I define $\boldsymbol{\lambda} = (\lambda_0^k, \lambda_1^k, \lambda_2^k)_{k=0}^{\bar{K}-1}$.

Lastly, I choose the discounting factor between two adjacent rounds δ to be 0.95.

5.2 Simulated Minimum Distance Estimator

I use a simulated minimum distance estimator for Γ (Agarwal, 2015; Aryal et al., 2024; McFadden, 1989; Pakes and Pollard, 1989). The estimate $\hat{\Gamma}$ minimizes the distance between the empirical and simulated moments after properly weighted, i.e.,

$$\hat{\Gamma} = \arg\min_{\Gamma} \left(\hat{\boldsymbol{g}} - \hat{\boldsymbol{g}}^M(\Gamma)\right)' \Lambda \left(\hat{\boldsymbol{g}} - \hat{\boldsymbol{g}}^M(\Gamma)\right), \tag{18}$$

where $\hat{\boldsymbol{g}}$ is a vector of moments constructed from the empirical data, $\hat{\boldsymbol{g}}^M$ is the average of moments constructed from M simulations of startup trajectories given parameter Γ. I follow Agarwal (2015) and choose the weight matrix Λ as the inverse of the covariance matrix of 10,000 vectors of bootstrap moments, each is constructed from a resampling of the empirical data with replacement. The standard error of the simulated estimator is constructed following Gourieroux and Monfort (1997), which captures both the (conventional) error due to limited sample size and the simulation error. Agarwal (2015) provides a thorough explanation of the inference procedure in its Online Appendix B.2.

The vector $\hat{\boldsymbol{g}}$ is composed of five sets of empirical moments. $\hat{\boldsymbol{g}}^M$ contains the same moments but is derived from and averaged over M simulations. The five sets of moments include:

1. The fraction of startups that succeed (IPO or M&A) and fail (bankrupt/out of business) after raising k rounds, where $k = 0, 1, ..., 6$. For example, in Fig. 3, 1,719 out of 5,591 startups failed immediately after establishment, then the moment to match is $\frac{1719}{5591}$. Likewise, there are 14 moments to match in this first set.

2. The mean and standard error of observed capital injection at round $k \in \{1, ..., 7\}$. Let $\mathcal{N}_k = \{i \in N : \prod_{k'=1}^{k} q_{i,k'}^{*} = 1\}$, where N is the number of startups, be the set of startups that successfully secure round k, the moments are computed as

$$\hat{g}_{d,mean,k} = \frac{\sum_{i \in \mathcal{N}_k} d_{i,k}}{|\mathcal{N}_k|},$$
$$\hat{g}_{d,se,k} = \sqrt{\frac{\sum_{i \in \mathcal{N}_k} (d_{i,k} - \hat{g}_{d,mean,k})^2}{|\mathcal{N}_k|}}.$$

3. The mean and standard error of observed post-money valuation at round $k \in \{1, ..., 7\}$. Mathematically,

$$\hat{g}_{v,mean,k} = \frac{\sum_{i \in \mathcal{N}_k} v_{i,k}}{|\mathcal{N}_k|},$$
$$\hat{g}_{v,se,k} = \sqrt{\frac{\sum_{i \in \mathcal{N}_k} (v_{i,k} - \hat{g}_{v,mean,k})^2}{|\mathcal{N}_k|}},$$

where $v_{i,k} = \frac{d_{i,k}}{a_{i,k}}$ is the post-money valuation of startup i at round k.

4. The mean and standard error of observed exit value at round $k \in \{0, 1, ..., 7\}$. Let $\tilde{\mathcal{N}}_k = \{i \in N : K_i = k, i \in \mathcal{N}_k\}$ be the set of startups that secure round 1 to k and then exit successfully,

$$\hat{g}_{W,mean,k} = \frac{\sum_{i \in \tilde{\mathcal{N}}_k} W_{i,k}}{|\tilde{\mathcal{N}}_k|},$$
$$\hat{g}_{W,se,k} = \sqrt{\frac{\sum_{i \in \tilde{\mathcal{N}}_k} (W_{i,k} - \hat{g}_{W,mean,k})^2}{|\tilde{\mathcal{N}}_k|}}.$$

5. The covariance between capital investment and post-money valuation at round $k \in \{1, ..., 7\}$,

$$\hat{g}_{cov,k} = \frac{\sum_{i \in \mathcal{N}_k} (d_{i,k} - \hat{g}_{d,mean,k})(v_{i,k} - \hat{g}_{v,mean,k})}{|\mathcal{N}_k|}.$$

For empirical moments, the first set can be derived from Fig. 3, the second to fourth sets are presented in Table 13.

5.3 Estimation Algorithm

A startup can be fully characterized by a vector of variables $(K, \boldsymbol{X}_K, \xi, \boldsymbol{S}_K)$. K determines when it can successfully exit, $\boldsymbol{X}_K$ and ξ determine its exit value, $\boldsymbol{X}_K$ and $\boldsymbol{S}_K$ determine both the progression through the VC rounds and the winning equity bid in each round.

The key in simulating the path of a startup is to evaluate the *maximum* expected payoff, $V_k^*(\boldsymbol{X}_k, \mu_k)$, to the investors. There are two challenges: 1) the state space gets very large at later rounds; 2) V_k^* is essentially an expectation with respect to all possible future trajectories and hence is difficult to compute.

Under the Markov assumption in (17), the future capital injections, over which V_k^* takes expectation, are independent of the current characteristics $\boldsymbol{X}_k$ conditional on $\bar{W}_k$ and d_k, so that $V_k^*(\boldsymbol{X}_k, \mu_k)$ is reduced to $V_k^*(\bar{W}_k, d_k, \mu_k)$. Furthermore, by Eq. (7) and (8), $\mathfrak{q}_k^*(\boldsymbol{X}_k, \mu_k)$ and $\mathfrak{a}_k^*(\boldsymbol{X}_k, \hat{\mu}_k)$ are reduced to $\mathfrak{q}_k^*(\bar{W}_k, d_k, \mu_k)$ and $\mathfrak{a}_k^*(\bar{W}_k, d_k, \hat{\mu}_k)$. This technique simplifies the state variables of round k to $\Omega_k = (\bar{W}_k, d_k, \mu_k, \hat{\mu}_k)$, so that I only need to keep track of these four variables. To simplify notations, I write $V_k^*(\Omega_k)$, $\mathfrak{q}_k^*(\Omega_k)$ and $\mathfrak{a}_k^*(\Omega_k)$ while keeping in mind that μ_k is irrelevant to $\mathfrak{a}_k^*$ and $\hat{\mu}_k$ is irrelevant to both V_k^* and $\mathfrak{q}_k^*$.

To overcome the second challenge, I write V_k^* in its Bellman representation, which allows me to use backward induction to back out the value of V_k^* at any state for any parameter Γ.

$$\begin{aligned} V_k^*(\Omega_k) &= \Pr(K = k \big| K \geq k)\mathbb{E}\left[W_K \big| \Omega_k, K = k\right] \\ &\quad + \Pr(K > k \big| K \geq k)\delta\mathbb{E}\left[(1 - \mathfrak{a}_{k+1}^*(\Omega_{k+1}))\mathfrak{q}_{k+1}^*(\Omega_{k+1})V_{k+1}^*(\Omega_{k+1}) \big| \Omega_k, K > k\right]. \end{aligned} \tag{19}$$

To interpret the Bellman equation, for a startup at state Ω_k, there is a chance that immediately after this round, an acquisition opportunity arises or the board decides it's an ideal time to go public, leading to an exit and realization of value. Otherwise, it transitions to state Ω_{k+1} and proceeds to raise the next round, in which case the expectation is over the possibility of failing to raise money and ceasing operation and the dilution in equity ownership.

Therefore, to evaluate V_k^*, it is sufficient to know its transition point in the next round rather than its entire path.

Evaluate maximum expected payoff. Before introducing the steps for computing V_k^*, I discretize the space of Ω_k. Motivated by data, I discretize the support of d_k to grids $\{0, 10, 20, ..., 280, 290, 300\}$ and $\bar{W}_k$ to $\{0, 10, 20, ..., 290 + 100k, 300 + 100k\}$. I allow the maximum grid of $\bar{W}_k$ to expand over rounds to reflect the growth in a startup's value. Likewise, I discretize the space of μ_k and $\hat{\mu}_k$ to grids $\{-1.5 - 0.5k, -1.4 - 0.5k, ..., 1.4 + 0.5k, 1.5 + 0.5k\}$, with expanding support to reflect the increasing divergence in beliefs across startups. The set of grids for to be evaluated, $\mathbf{\Omega_k}$, encompass every permutation of the possible values for $\bar{W}_k, d_k, \mu_k, \hat{\mu}_k$.

Given parameter Γ, to compute V_k^* at each grid, I start from the last possible round.

1. At round $\bar{K}$, compute $V_{\bar{K}}^*(\Omega_{\bar{K}}), \mathfrak{q}_{\bar{K}}^*(\Omega_{\bar{K}}), \mathfrak{a}_{\bar{K}}^*(\Omega_{\bar{K}})$ for each discretized grid of $\Omega_{\bar{K}}$. Since $\bar{K}$ is for sure the last round and by the property of lognormal distribution, Eq. (19) is simplified to

$$V_{\bar{K}}^*(\Omega_{\bar{K}}) = \bar{W}_{\bar{K}} \exp\left(\mu_{\bar{K}} + \frac{1}{2}\tau_{\bar{K}}^2\right), \tag{20}$$

   and

$$\mathfrak{q}_{\bar{K}}^*(\Omega_{\bar{K}}) = \mathbb{1}\{V_{\bar{K}}^*(\Omega_{\bar{K}}) \geq d_{\bar{K}}\}, \tag{21}$$

$$\mathfrak{a}_{\bar{K}}^*(\Omega_{\bar{K}}) = \frac{d_{\bar{K}}}{V_{\bar{K}}^*(\hat{\Omega}_{\bar{K}})}, \tag{22}$$

   where $\hat{\Omega}_{\bar{K}} = (\bar{W}_{\bar{K}}, d_{\bar{K}}, \hat{\mu}_{\bar{K}}, \hat{\mu}_{\bar{K}})$.

2. At round $k < \bar{K}$, for each grid of Ω_k, simulate $(\Omega_{k+1}^l)_{l=1}^L$ transitioned from Ω_k, so that Eq. (19) becomes

$$\begin{aligned} V_k^*(\Omega_k) = {} & \Pr(K = k \big| K \geq k)\bar{W}_k \exp\left(\mu_k + \frac{1}{2}\tau_k^2\right) \\ & + \Pr(K > k \big| K \geq k)\delta\frac{1}{L}\sum_l \left[(1 - \mathfrak{a}_{k+1}^*(\Omega_{k+1}^l))\mathfrak{q}_{k+1}^*(\Omega_{k+1}^l)V_{k+1}^*(\Omega_{k+1}^l)\right]. \end{aligned} \tag{23}$$

   The two conditional probabilities are identified directly from data, the values of $\mathfrak{a}_{k+1}^*, \mathfrak{q}_{k+1}^*$ and V_{k+1}^* at the simulated grids are taken from the previous step.

3. Iterate the second step until all rounds have been visited.

Estimation algorithm. I use Algorithm 1 to compute the objective function in Eq. (18) given parameter Γ.

Algorithm 1: Algorithm for computing the objective function given Γ

```
// Step 1: for each grid in each round, compute the expected payoff,
   binary funding rule and equity ownership rule
for k ← 1 to K̄ do
    foreach Ω_k in Ω_k do
        compute V*_k(Ω_k), q*_k(Ω_k), a*_k(Ω_k) using Eq. (20)(21)(22)(23);
    end
end
// Step 2: generate simulated data
for m ← 1 to M do
    generate simulated data D_m;
end
// Step 3: for each startup in each simulation, determine its
   trajectory
for m ← 1 to M do
    foreach startup in D_m do
        generate its trajectory and each round's winning equity ownership using
         results from Step 1;
    end
end
// Step 4: compute the objective function
compute the objective function using results from Step 3;
```

In Step 2, each simulation contains the N startups in the empirical data. If the startup exits successfully, then I only simulate its $\boldsymbol{S}_K$ and fix $K, \boldsymbol{X}_K$ and ξ as they are in the empirical data (ξ, though not observed directly, can be backed out from Eq. (11) given Γ). If the startup fails, then I simulate its K, unrevealed future capital requirements, and unobserved quality ξ, in additional to signals $\boldsymbol{S}_K$.

Considering the non-smoothness in the simulated objective function, I use genetic algorithm to find the global minimum.

6 Results

In this section, I present my estimation results. First, I discuss how my estimates shed light on VC learning across rounds in each sector. Second, I simulate startup trajectories based on the estimates and discuss the heterogeneity among startups following different paths. Third, I show how costs and returns are shared across VCs at different stages.

6.1 Information and learning

To get a sense of what the estimates of σ_ξ and $\boldsymbol{\sigma}$ imply about the learning process in each sector, I choose ξ at the 1st percentile — "low" quality startup — and simulate investors' posterior beliefs on ξ after each round for 10,000 times. Fig. 8 presents the averaged beliefs. Compared to biotech investors, software investors' beliefs are initially more dispersed, but converge faster in the subsequent rounds, suggesting that software investors start with less information but learn faster. In Appendix J, I provide the results for each quality level.

Statistically, I measure the convergence in beliefs by coefficient of variation (CV), defined as $\frac{\tau_k}{\mu_k}$. As shown in Table 5, in both sectors and across all quality levels, CV approaches 0 over rounds, reflecting the convergence of perceived quality to its actual value.

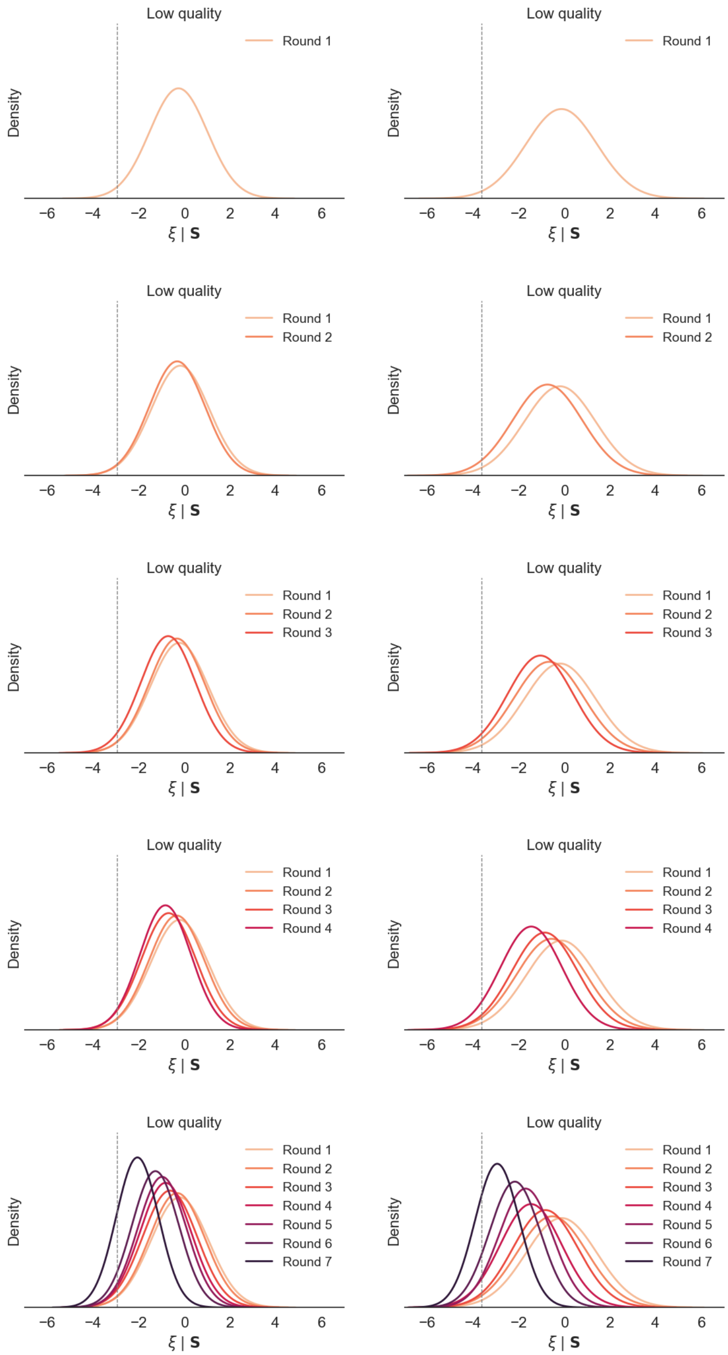
Another key observation is that in software, CV is significantly higher in round 1 compared to biotech, but thereafter it converges to 0 at a faster speed. This suggests that for software startups, most uncertainty is concentrated in the initial rounds. Once these are passed, signal precision improves rapidly, and startup quality is quickly revealed. In contrast, uncertainty in biotech startups is spread more evenly across multiple rounds, leading to slower learning.

Table 5: **Coefficient of variation**

Sector	**Quality**	**Round 1**	**2**	**3**	**4**	**5**	**6**	**7**
Biotech	Low	-5.53	-3.6	-1.72	-1.25	-1.0	-0.68	-0.31
	Medium-low	-18.64	-12.12	-5.91	-4.26	-3.4	-2.33	-1.05
	Medium-high	19.38	12.8	5.97	4.32	3.45	2.36	1.06
	High	5.6	3.63	1.73	1.25	1.0	0.68	0.31
Software	Low	-11.16	-2.52	-1.48	-0.75	-0.62	-0.44	-0.21
	Medium-low	-37.23	-8.54	-5.06	-2.55	-2.11	-1.49	-0.73
	Medium-high	39.34	8.96	5.14	2.58	2.14	1.51	0.74
	High	11.37	2.53	1.48	0.75	0.62	0.44	0.21

Note: This table displays the coefficient of variation for each sector, round and level of unobserved quality. The 1st, 25th, 75th, and 99th percentiles of ξ are selected to represent "low," "medium-low," "medium-high," and "high" levels.

Figure 8: **Learning process**



Note: This figure displays, for a low quality startup, the posterior belief on the quality after each round. The left and right figures are biotech and software, respectively. The x-axis represents the belief on ξ conditional on the signals $\boldsymbol{S}$. The grey dashed lines mark the true quality. The 1st percentile of ξ is selected to represent "low" quality.

These findings reflect sector-specific business models. Software businesses are typically built upon proven technologies, the primary source of uncertainty is the demand for a new product or service. So, the common practice is to develop and launch a minimum viable product quickly to test market demand (Lerner and Nanda, 2020). Once validated, the next steps involve product iteration and business expansion, which usually come with less uncertainty.

In contrast, drug development process consists of several distinct stages, each testing a different aspect of the drug (DiMasi and Grabowski, 2012). Specifically, pre-clinical trials involve creating a new compound and testing it on animals to answer basic safety questions. If successful, the firm proceeds to a three-phase clinical study involving human subjects. Phase I tests the drug on a small group of healthy volunteers to assess toxicity and determine safe dosages. Phase II tests the drug on patients with the targeted disease to evaluate efficacy and identify side effects. Phase III focuses on testing the drug's efficacy in a broad population (Food and Administration, 2018). Since each stage addresses a distinct aspect of the drug, uncertainties for biotech startups are distributed more evenly throughout the development process, resulting in slower learning rate compared to software startups.

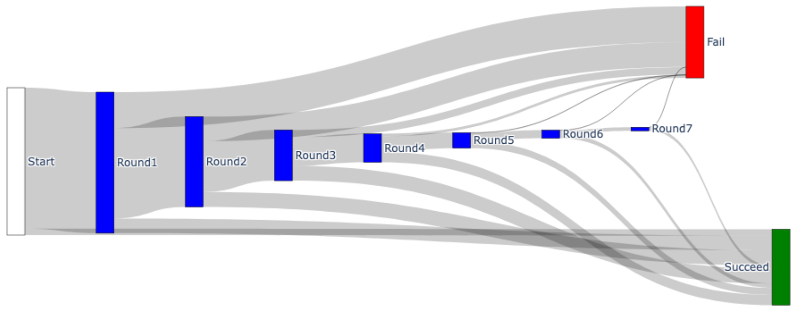
6.2 Exit outcome and startup heterogeneity

With my estimates and model, I simulate trajectories for any startup and determine the impact of inherent characteristics on the path it follows. Fig. 9 visualizes the simulated paths of 2,377 biotech startups and 3,203 software startups in my data. Compared to biotech sector, software sector has a lower success rate (33.0%) — the proportion of startups that successfully exit through IPO or M&A. This difference is primarily driven by the high failure rate in the initial round due to high initial uncertainties.

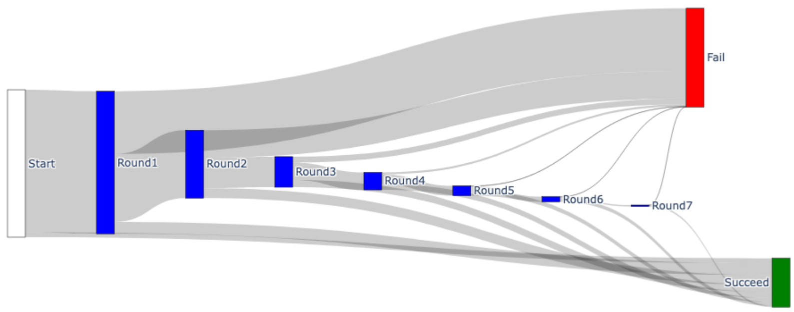
To better understand the factors driving a startup's fundraising journey, I compare the characteristics of startups on different paths. In Table 6, I present the average number of patents held by startups that either secure funding in a given round and exit immediately via IPO or M&A, or fail to secure funding in that round. I compare only startups that have undergone the same number of rounds, so that differences in patent counts are not due to additional time for development.

Figure 9: **Simulated trajectories for all startups**

(a) Biotech



(b) Software



Note: This figure presents the average of the simulated paths for 2,377 biotech startups and 3,203 software startups from 100 simulations using my estimates. In Fig. 9a, the percentages of startups that exit successfully after Start, Round 1,..., Round 7 are 4.3%, 10.0%, 10.1%, 10.1%, 7.0%, 4.7%, 3.0%, 2.3%, and that fail after Round 1 to 7 are 24.4%, 16.8%, 5.1%, 1.8%, 0.3%, 0.1%, 0.1%. In Fig. 9b, the percentages of startups that exit successfully after Start, Round 1,..., Round 7 are 3.3%, 8.1%, 6.8%, 4.9%, 3.8%, 2.9%, 2.3%, 1.0%, and that fail after Round 1 to 7 are 42.6%, 18.6%, 4.0%, 1.4%, 0.3%, 0.1%, 0.1%.

Table 6: **Average patents of successful and failed startups**

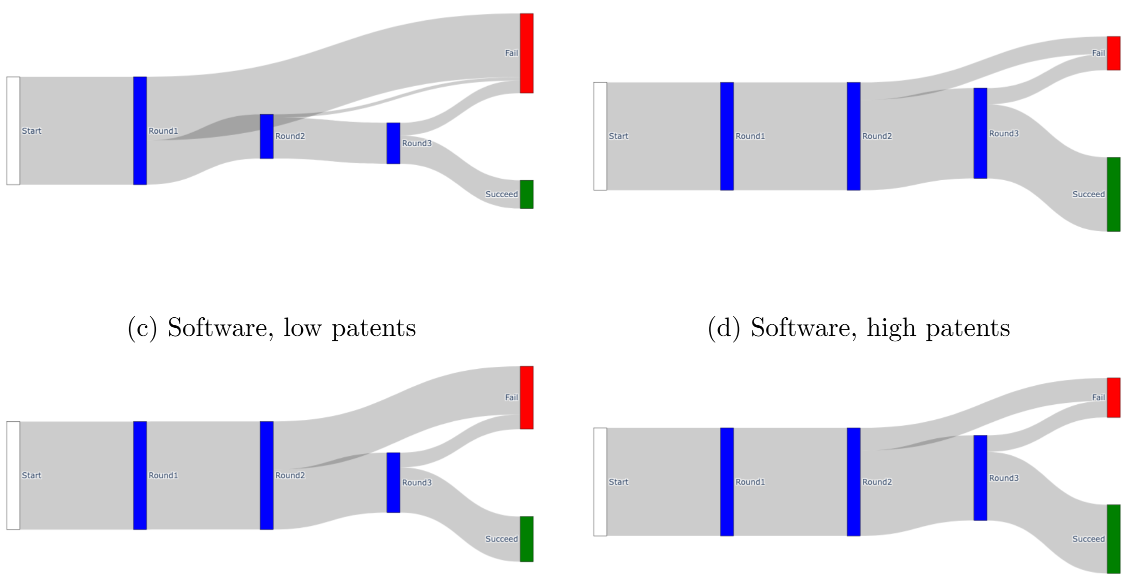
Sector	**Success/Fail**	**Round 1**	**2**	**3**	**4**	**5**	**6**	**7**
Biotech	Success	4.20	9.97	12.39	15.34	15.34	24.18	18.05
	Fail	1.18	2.00	4.65	6.67	6.40	5.23	5.87
Software	Success	1.71	3.41	6.13	9.70	17.70	13.25	18.69
	Fail	0.28	0.32	1.14	3.45	6.17	8.34	6.23

Note: This table he average number of patents held by startups that either secure funding in a given round and exit immediately via IPO or M&A, or fail to secure funding in that round. These patents consist of both in-house developed patents and those licensed from other entities.

Successful startups tend to have more patents than failed startups. In addition, software companies typically own fewer patents than biotech companies, consistent with the views that patents are crucial in biotech sector for protecting innovations and securing investment, while in software sector, patents are valuable but often less central due to the industry's fast-moving nature and alternative forms of IP protection.

Following this, I simulate the trajectories of a given startup by holding all other characteristics constant and only varying the number of patent. The results are presented in Fig. 10. In both sectors, startups with more patents are associated with a higher probability of success. This finding is consistent with previous research that documents the positive impact of innovation output on a firm's future performance (Kline et al., 2019; Kogan et al., 2017).

Figure 10: **Simulated trajectories for a given startup**



(a) Biotech, low patents

(b) Biotech, high patents

(c) Software, low patents

(d) Software, high patents

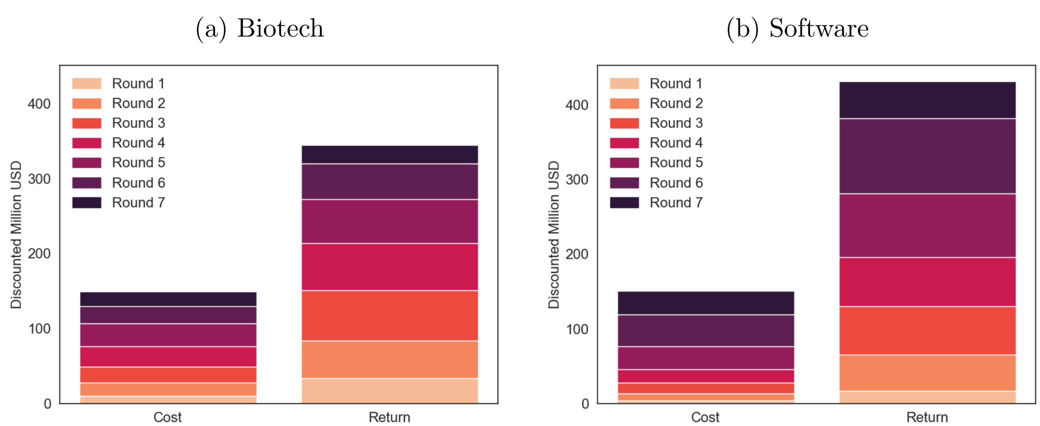
Note: For each sector, I construct a representative startup fixing all characteristics at the median level except for patents — "low" and "high" patents are chosen at 5th and 95th percentile, respectively. Then I simulate the trajectories for each representative startup for 1000 times with randomly sampled noises.

6.3 Cost and Return Sharing

While many studies have focused on evaluating average VC returns (Bygrave, 1992; Chen et al., 2012), surprisingly little is known about how (un)equal the return sharing is across stages in practice. Moreover, in traditional financial models (e.g., CAPM), the inequality on return-sharing, if any, arises from the assumption of risk-averse agents, yet little is known about the impact of evolving information structure on return sharing among VCs at different stages.

In this subsection, I analyze how the interaction between sequential dissemination of information and VC competition affects the return sharing across VCs at different stages. Using the simulated data, I find that early-round VCs obtain a significantly higher MoM than late-round VCs in both sectors (see Fig. 11). On average, round 1 VCs in biotech and software receive \$3.31 and \$4.02, respectively, for every dollar invested — roughly three times the returns of last-round VCs, consistent with estimates from Kaplan and Schoar (2005) and Cochrane (2005). Compared to biotech investors, software investors achieve higher MoM, particularly in the early rounds, due to greater initial uncertainties. However, in my model this unequal risk-sharing across stages stems from the sequential dissemination of information rather than risk-averse agents in standard financial models such as CAPM.

Figure 11: **Cost and return sharing**



Note: In Fig. 11a, the capital inputs and returns of each round's investors are 10.24, 17.42, 21.65, 27.28, 30.67, 22.5, 19.64, and 33.94, 50.33, 66.56, 62.83, 58.79, 47.29, 24.77; the MoMs are 3.31, 2.89, 3.07, 2.3, 1.92, 2.1, 1.26. In Fig. 11b, these values are 4.11, 9.06, 14.68, 17.95, 30.42, 42.62, 31.81; 16.53, 49.05, 65.01, 65.53, 85.51, 100.44, 49.03; and 4.02, 5.41, 4.43, 3.65, 2.81, 2.36, 1.54, respectively. All numbers except for MoMs are in million USD.

The two sectors exhibit distinct cost structure, with biotech distributing expenses evenly across rounds, while software pushes most costs to later stages, reflecting sector-specific business models. As explained in Section 6.1, in the biotech sector, each stage involves a steadily expanding pool of volunteers, matching the *steady* increase in the costs over rounds. However, in the software sector, initial rounds focus on developing a minimum viable product at the lowest possible cost in order to test the market demand, whereas later rounds focus on "cash-burning" activities like business expansion and product promotion. Hence, the cost structures are consistent with the prioritized activities over rounds.

7 Inefficiency and Welfare

Three sources of inefficiency are present in this market: 1) dynamic information externalities; 2) uncertainty in startup quality; and 3) uncertainty regarding the number of future rounds and future capital infusions. As discussed in Section 3.2, the first source results in under-investment. The second source leads to ex-post sub-optimal capital allocation, because the choice of which startup to support is made without knowing the true quality and only based on noisy signals. However, these choices are ex-ante optimal, making this type of inefficiency better referred as "constrained efficiency". This is also true for the third inefficiency. Notably, the first two inefficiencies are linked because the information externalities would be resolved if the startup's true quality were observable. In that case, VC bidders would have symmetric information and engage in Bertrand competition until the surplus is bid down to 0.

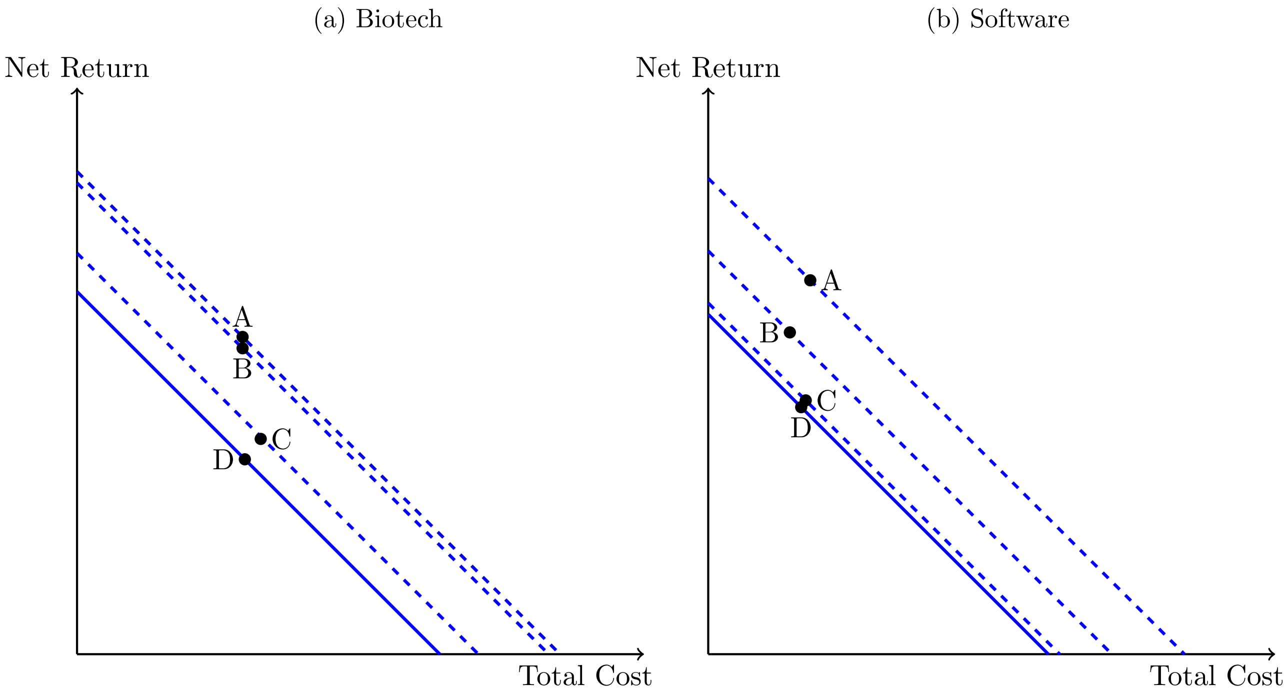
Moving forward, I refer to these inefficiencies as "information externalities", "quality uncertainty", and "other uncertainties". First, I run counterfactuals to quantify the efficiency losses from each source. Then, I compute Type I errors — not funding good startups, and Type II errors — funding bad startups, in order to identify the main reason(s) for efficiency loss and explore policies to mitigate these losses. Finally, I extend this framework to examine the effects of various policy interventions on innovation.

7.1 Quantifying inefficiency

I define *first-best* efficiency as the one in which all NPV-positive startups are funded through and all NPV-negative startups are not funded from the beginning. It is the highest possible efficiency that can be achieved, and is attained when all of the three sources of inefficiency are eliminated. *Second-best* efficiency is the one in which a net return-maximizing planner, with full knowledge of startups' true qualities, makes the funding choices, it corresponds to the efficiency level when dynamic information externalities and quality uncertainties are resolved. *Third-best* efficiency is similar to second-best efficiency, except the planner does not observe the startups' true qualities directly; instead, he learns from the same signals available to VC investors. This corresponds to the efficiency level when only dynamic information externalities are eliminated. *Status quo* is the efficiency under the current market practice when all the three inefficiencies present.

In Fig. 12, I present the net return and total cost with the sequential elimination of each source of inefficiency. Point D denotes the total cost and net return in status quo, based on the simulated results using my estimates. Point C corresponds to the third-best efficiency. In both sectors, investors at Point D invest less than what the planner invests at Point C, aligning with the theoretical results of under-investment in Proposition 3. The vertical distance between Point C and D — \$10 billion in biotech and \$3 billion in software — quantifies the efficiency loss due to dynamic information externality.

Figure 12: **Quantifying inefficiency**



Note: A, B, C, D represent first-best, second-best, third best and status quo scenario, respectively. The coordinates for A, B, C, D are respectively (73.2, 239.7), (73.1, 234.6), (80.9, 195.7), (73.6, 186.3) in Fig. 12a, and (44.8, 264.9), (35.82, 241.9), (43.1, 212.0), (40.7, 209.0) in Fig. 12b, all in billion USD. All results are based on the same simulated data as before.

Point B represents second-best efficiency; it incurs lower costs than Point C but yields a higher net return, due to improved identification of NPV-positive projects. The vertical difference between B and C — \$39 billion in biotech and \$30 billion in software — measures the efficiency loss caused by uncertainties in quality. This loss is smaller in software due to faster learning. Point A is the first-best efficiency, its vertical distance to B — \$5 billion in biotech and \$23 billion in software — quantifies the efficiency loss due to "other uncertainties".

In addition to the aggregated and ex-post outcomes in Fig. 12, Table 7 presents the average expected value of individual startups across the four scenarios.

Table 7: **Expected Value of Startups**

(a) Biotech

Round		**1**	**2**	**3**	**4**	**5**	**6**
Scenarios	A 1st-best	105.43	113.65	126.71	130.89	148.35	168.40
	B 2nd-best	62.44	99.97	126.62	123.68	133.73	168.67
	C 3rd-best	14.04	28.59	45.98	58.84	74.49	104.07
	D status quo	11.00	21.75	38.49	52.28	68.11	96.21
Gaps	gap between D and A	0.9	0.81	0.7	0.6	0.54	0.43
	gap between D and C	0.22	0.24	0.16	0.11	0.09	0.08
	gap between C and B	0.78	0.71	0.64	0.52	0.44	0.38

(b) Software

Round		**1**	**2**	**3**	**4**	**5**	**6**
Scenarios	A 1st-best	85.53	101.59	124.28	166.85	253.54	307.0
	B 2nd-best	48.46	78.78	94.14	134.0	227.41	283.64
	C 3rd-best	11.51	21.0	32.46	62.13	101.97	165.75
	D status quo	7.98	16.31	27.52	57.39	97.95	162.59
Gaps	gap between D and A	0.91	0.84	0.78	0.66	0.61	0.47
	gap between D and C	0.31	0.22	0.15	0.08	0.04	0.02
	gap between C and B	0.76	0.73	0.66	0.54	0.55	0.42

Note: This table presents the mean of the expected values of individual startups in each round (before the capital is injected) across all scenarios. All expected values are in million USD. In order to remove the selection effect, the expected value of a startup is filled in with 0 under the first-best scenario if it's NPV-negative, and under the other scenarios if it fails to raise or reach that round. The gap between two scenarios are calculated as one minus the ratio of the mean expected values in each scenario.

On average, the expected value rises with each round in both sectors across all scenarios, with the status quo and first-best scenarios persistently yielding the lowest and highest expected values, respectively. This gap is largest in the initial round, where individual startups are undervalued by up to 90% under the status quo, primarily because outstanding startups are not recognized at the beginning. However, this "under-valuation" narrows to around 40% by the sixth round, as more information about startup quality becomes available.

Similar to Fig. 12, information externalities drive the gap between D and C, which narrows over rounds as there are fewer future VCs benefitting from late-stage externalities. Additionally, uncertainties in quality drive the gap between C and B, which narrows as more information is released to the market.

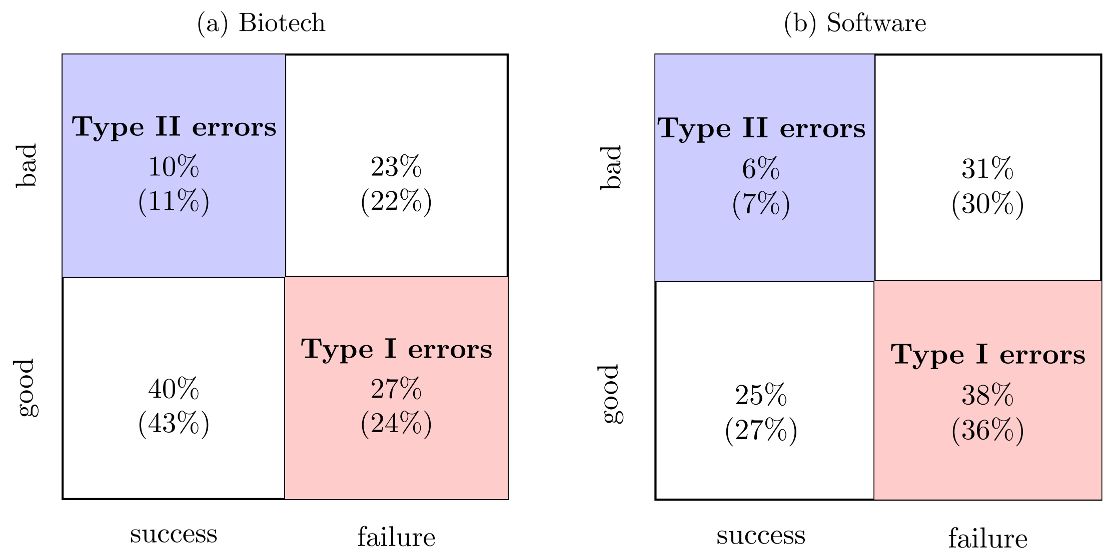
7.2 Type I and Type II errors

In this subsection, I dig deeper by quantifying the losses associated with Type I errors — not funding good startups, and Type II errors — funding bad startups. The results have important policy implications: if Type I errors dominate, the government should complement the private VC market; if Type II errors dominate, the government should carry out stricter regulations on investment activities.

I call NPV-positive startups *good* startups and NPV-negative startups *bad* startups. Good startups that exit successfully (i.e., securing all VC rounds and then either getting acquired or going public) are labeled as *good-success*, bad startups that fail are labeled as *good-fail*. Similarly for *bad-success* and *bad-fail*.

Fig. 13 presents the proportions of startups in each group in status quo and 3rd-best scenario. These percentages are based on the same simulations as before. In both sectors, Type I errors occur more frequently. Compared to biotech, software makes more Type I errors due to less informative initial signals, but fewer Type II errors due to faster learning. This result is consistent with the finding of initial high failure rate for software startups in Fig. 9. Furthermore, the planner in the 3rd-best scenario make fewer Type I errors and more Type II errors than status quo, as he assigns higher value to each individual startup (Proposition 3) and thus invests more aggressively.

Figure 13: **Count of Type I and Type II errors**



Note: The values with (without) parentheses are the percentages of startups in each group in the 3rd-best scenario (status quo).

The loss from Type I errors is the gap in net returns for good startups between the status quo and a benchmark scenario, chosen as either the 1st-best or 3rd-best scenario. Similarly, the loss from Type II errors corresponds to the gap for bad startups.

The results are shown in Table 8. Take the first row in Table 8a for example. The total expenses on all good startups under 1st-best, 3rd-best and status quo scenarios are 73.28, 60.91, 56.36 billion USD, respectively. Similar for the net returns in each scenario. The gap in net return between status quo and 1st- or 3rd-best are respectively -44.64 and -10.88 billion USD, quantifying the dollar losses from Type I errors — not funding good startups. Using either benchmark, the losses from Type I errors are significantly higher than that from Type II errors, suggesting that "not funding good startups" is the main reason for efficiency losses in this market.

Table 8: **Quantifying Type I and Type II errors**

(a) Biotech

	Cost			**Net return**			**Δ Net Return**	
	1st-best	3rd-best	status quo	1st-best	3rd-best	status quo	1st-best	3rd-best
Good	73.28	60.91	56.36	239.75	205.99	195.11	-44.64	-10.88
Bad	0.00	20.08	17.25	0.00	-10.28	-8.78	-8.78	1.50
Total	73.28	80.99	73.61	239.75	195.72	186.33	-53.42	-9.38

(b) Software

	Cost			**Net return**			**Δ Net Return**	
	1st-best	3rd-best	status quo	1st-best	3rd-best	status quo	1st-best	3rd-best
Good	44.86	31.81	30.64	264.98	219.03	215.20	-49.78	-3.83
Bad	0.00	11.43	10.06	0.00	-6.96	-6.13	-6.13	0.83
Total	44.86	43.24	40.70	264.98	212.07	209.07	-55.91	-3.00

Note: Each row presents the cost and net return for good and bad startups under 1st-best, 3rd-best and status quo. Additionally, the last two columns show the gap in net return between status quo and 1st-/2nd-best. All the numbers are in billion USD.

Following this, I explore when Type I errors occur. Table 9 presents the timing of failure for good-fail startups. For example, 53% out of the 623 good-fail biotech startups fail in the first round, and 37% in the second, indicating that most Type I errors occur in the early stages when information is least reliable. Once surpassing the initial phases, good startups are almost guaranteed to secure additional fundings. Software startups follow similar patterns, but a larger proportion fail in the first round due to inaccurate initial signals. These results reinforce the insights presented in Table 5.

Table 9: **Timing of failure for good-fail startups**

Sector	**Round 1**	**2**	**3**	**4**	**5**	**6**	**7**
Biotech,%	53	37	9	1	0	0	0
Software, %	68	26	5	1	0	0	0

Note: The sample size is 623 for biotech and 1,184 for software.

Type I errors could be corrected by government funds for early-stage startups if allocated properly. For example, the U.S. government has launched programs such as Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) to support startups at early stages. A systematic study of the mechanism design for allocating such funds is beyond the scope of this paper, but indeed intriguing.

7.3 Relative Exit Importance

In this subsection, I extend the framework to examine the relative importance of the two types of exits: M&A and IPO. Using the model and estimates, I run counterfactuals to study the impact of an undervalued M&A exit on the financing of innovation.

Assume that conditional on success, a startup exits via M&A with probability 54.4% in biotech and 91.5% in software, based on empirical data. Using the same simulated data as in Fig. 12, I reduce the realized value of M&A exits by 10% to 50% to reflect stricter M&A policies, while keeping the value of IPO exits unchanged. In other words, the value of M&A exits is now drawn from a different, left-shifted distribution compared to IPO exits. Table 10 presents the cost and return estimates under each level of "devaluation". In addition, the change in net return is the difference in net returns between 0% devaluation and a given devaluation level; it is decomposed into extensive margin — fewer startups get funded, and intensive margin — funded startups are reduced in value.

For example, in biotech, when M&A exits are devalued by 10%, the overall return and cost become \$236.3 billion and \$69.6 billion. The net return, compared to 0% devaluation, is \$19.6 billion less, which is decomposed to 1) *extensive margin*: fewer projects get funded because VC investors expect lower returns and thus become more selective in the startups they choose to fund; and 2) *intensive margin*: reduced values of startups funded both before and after the implementation of an M&A devaluation. Extensive and intensive margins each attribute to \$6 billion and \$13.6 billion of the total drop in net return. In addition, M&A devaluation leads to a larger net return decrease in software compared to biotech, because M&A is a more common exit channel in software.

Table 10: **Counterfactuals of an undervaluation in M&A exit**

(a) Biotech

M&A	**Overall**			**Change in net return**		
Devaluation	Return	Cost	Net return	Sum	Extensive margin	Intensive margin
0%	259.9	73.6	186.3	-	-	-
10%	236.3	69.6	166.7	-19.6	-6.0	-13.6
20%	210.6	64.1	146.4	-39.9	-14.2	-25.7
30%	183.9	58.1	125.8	-60.5	-24.7	-35.9
40%	157.8	51.8	105.9	-80.4	-36.5	-43.9
50%	132.1	44.4	87.7	-98.6	-49.3	-49.3

(b) Software

M&A	**Overall**			**Change in net return**		
Devaluation	Return	Cost	Net return	Sum	Extensive margin	Intensive margin
0%	249.8	40.7	209.1	-	-	-
10%	222.2	38.7	183.5	-25.6	-3.2	-22.4
20%	192.5	35.9	156.6	-52.5	-9.3	-43.1
30%	160.7	31.9	128.8	-80.3	-19.4	-60.8
40%	127.4	27.0	100.4	-108.7	-35.1	-73.6
50%	91.5	20.4	71.0	-138.1	-60.8	-77.2

Note: This table presents the returns, costs and net returns if the realized value of an M&A exit drops by 10% to 50%. 0% devaluation is when M&A exits are not devalued, and is the same as point D in Fig. 12. All values are in billion USD.

8 Discussion

I have shown that understanding VCs' investment strategies under several opposing incentives requires assumptions on what they bid on in the equity space. I set aside second-order issues to keep the framework tractable and convey the core messages. In this section, I briefly touch on a few of these issues, which may serve as pointers for future research.

8.1 Multi-dimensional contract

Throughout this paper, I have assumed that VCs only bid on equity ownership. This is mainly due to data limitation — a non-trivial portion of contract terms other than equity ownership is missing in Pitchbook. In addition, incorporating multi-dimensional contracts significantly complicates the bidding process, thus it is challenging to derive meaningful implications for financing innovations. That said, my estimated level of inefficiency due to information spillover is likely an upper bound on the true level, as other contract terms may step in to bring the agreement closer to, if not exactly coincide with, the optimal contract.

Although the actual VC contracts tend to be multi-dimensional, to the best of my knowledge, there is not a clause that specifically compensates early investors for the information externalities they generate. For example, anti-dilution protection is commonly used to address the concern of ownership dilution. However, it's important to point out that it does *not* imply a type of stock that is immune to dilution under all circumstances — as a company issues additional equity to raise more capital in subsequent fundraising rounds, the equity ownership of existing shareholder, whether founders or investors, dilutes *naturally*. Hence, the concept of an undilutable equity share is effectively non-existent in this context. What anti-dilution protection really means is a downside protection that is only triggered in down rounds when a startup raises new capital at a lower valuation than previous funding round. If this happens, the protected investors are entitled to additional shares in the subsequent financing or exiting events (Kaplan and Strömberg, 2003). While it might somewhat compensate early investors during challenging times, in other times it's often seen as a cosmetic clause. Hence, anti-dilution protection is not fundamentally intended to address information externalties. Beyond that, these downside protections are sometimes waived even when triggered — the triggering event typically means that the company is in distress, enforcing these protections often exacerbates the company's situation and makes future fundraising even harder, which are not in the best interest of current investors (Gornall and Strebulaev, 2023).[13]

8.2 Value-added by VC

I also assume that VCs do not add values in addition to capital input, so that VCs' bids are solely based on their private signals but not their (perhaps) heterogeneous ability to add value. While convenient, this assumption is at odds with some research suggesting that

[13] https://x.com/twistartups/status/1821607995207537038

VCs add values by bringing networks of connections, recruiting key personnel, and providing operational support (Bernstein et al., 2016; Sørensen, 2007). In this case, my estimated β and $\boldsymbol{\sigma}$ are likely to be biased up as they pick up the effects of VC value-add. However, the impact of value-add tends to be case-dependent. For example, underrepresented and first-time founders may gain significant benefit from VC support in building an initial customer base and recruiting talents, while there may be less opportunity for VCs to add value to experienced founders and startups at later stages apart from the capital they contribute.[14] However, the value-add to heterogeneous founders and projects is at present not well understood. A potential extension is to model VC competition as a scoring auction in which each bidder's value is a function of both private signal and value-add ability. The major challenge is the data limitation, as this extension also requires information on the identity of the second strongest competitor rather than just the winner. This issue can be solved either with better data or reasonable assumptions that help identify the losing bidders.

8.3 Risk-neutral VC

While individual investors and certain institutional investors (e.g., banks) may exhibit risk-averse tendencies, VCs differ significantly in two key aspects. First, most — if not all — VCs employ strong diversification strategies to manage risk (e.g., Harris et al., 2014; Proksch et al., 2016). Large VCs, such as Sequoia, typically hold portfolios with hundreds of companies spanning multiple sectors. In contrast, mid-sized and smaller VCs often specialize in one or two sectors but diversify within them by investing in multiple startups — typically 20 to 30 — across different verticals (e.g., 3D printing and autonomous driving). Such diversified portfolios allow VCs to tolerate individual failures in pursuit of outsized returns.

Second, startup returns are highly skewed, compelling VCs to chase the outliers (Kerr et al., 2014). VCs understand very well that most of the startups they invest will fail and only a small number will generate outsized returns. For example, Kerr et al. (2014) show that 7% investments from an average VC fund generates nearly 60% of the final returns. Therefore, VCs focus on startups with massive upside potential and may reject "safe" businesses with predictable growth — if just one or two investments yield returns in the hundreds or thou-

[14] https://www.vcstack.io/blog/vc-value-add

sands of times, they can offset all the fund's losses while still delivering substantial overall profits (Ewens et al., 2018).

Due to these reasons, modeling VCs as risk-neutral may not be as unreasonable as it seems. Zhao et al. (2015) estimate VC-specific risk-aversion level using Crunchbase data and find that VC firms holding a large number of investments (e.g., Sequoia Capital, Accel Partners, Tiger Global) tend to be risk-neutral. To understand the sensitivities of my results to this assumption, I could follow their method to estimate the risk-aversion level for each VC in my dataset, and then run robustness check on the subsample of startups invested only by risk-neutral VCs.

However, it's still useful to understand how my results will be biased if VCs are risk-averse but I assume they are risk-neutral. In my current work, I use the gap between valuation and exit value (after correcting for selection) as a measure of the uncertainty in unobserved quality. For example, if a startup's valuation at round 1 is \$1M and its exit value at IPO a year later is \$2M, then the \$1M gap (\$2M - \$1M) reflects the level of uncertainty — since, in the absence of uncertainty, the valuation would have already been \$2M at round 1. However, if VCs are risk-averse, this gap reflects not only uncertainty but also a risk premium demanded by risk-averse investors. As a result, I might overestimate the level of uncertainty in current work (in other words, my estimate is an upper bound for the level of uncertainty).

8.4 Follow-up investment

Another assumption I make is that a VC leads a single round. This assumption is motivated by the data observation in Section 2.2.6 that nearly 90% VCs lead only one round of investment in a startup, and that the further apart two rounds are, the less overlap there is between the identities of their syndicate VC groups. Furthermore, this assumption meaningfully improves the tractability of this work — without it, I need to explicitly model each VC's choice of following-up or not and the price for an internal VC candidate, likely adding substantial computation complexity to it.

That said, there are two extreme cases: one where each round is led by a different VC, and another where a single VC leads all rounds (i.e., the planner's case), with reality lying somewhere in between. Hence, the estimated level of inefficiency is likely an upper bound on the true value as the threshold for investing is likely lower for a VC who is given the opportunity to reinvest in the future (option value).

9 Conclusion

In this paper, I develop a tractable empirical framework to understand the strategic investment decisions of competing VCs under uncertainties and the efficiency gap compared to alternative information and market structures. My framework captures all the salient features of VC-startup interaction, including uncertainty in startup quality, staged financing, dynamic information structure, and imperfect VC competition.

I estimate the model using a novel dataset of biotech and software startups that includes both investment and exit data. I find that most uncertainties in software businesses are concentrated in the initial round, whereas in the biotech sector, they are spread across multiple rounds. Consequently, software investors, though initially less informed, learn about the startup quality at a faster pace.

Next, through several counterfactual exercises, I use the estimates to explore the efficiency gap under alternative information and market structures. I find substantial efficiency loss relative to the first-best scenario. In particular, the current biotech and software sectors respectively yield 78% and 79% of the first-best efficiency. By isolating the role of different sources of constraints in determining efficiency, I find that information externalities, quality uncertainty, and other uncertainties contribute to 18%, 73%, 9% of the efficiency loss in biotech, and 5%, 54%, 41% in software.

My paper is the first to incorporate uncertainty, staged financing and imperfect competition in the VC-funding space. I see this work as opening up several important areas for future research on related topics. I hope that future researchers studying VC-startup interaction in similar contexts will explore the possibilities of capturing VC value-add and other important contract dimensions, and consider what new insights their estimates provide in understanding the role of VC.

A Case Studies

A.1 Siri

Siri is a developer of a web-based personal assistant designed to answer questions, make recommendations and perform actions by delegating requests to a set of Internet services. The application software supports a wide range of user commands, including performing phone actions, checking basic information, scheduling events and reminders, handling device settings, searching the Internet, navigating areas, finding information on entertainment, and is also able to engage with iOS-integrated apps, enabling users to get personalized results for their actions.

Siri was founded in 2007 and headquartered in San Jose. Its primary industry is *Application Software*, and it belongs to the verticals of *Artificial Intelligence & Machine Learning* and *Mobile*. Siri was assigned 3 patents.

Table 11 presents the financing history of Siri. Siri raised $8.55 million of Series A venture funding from Morgenthaler Ventures and Menlo Ventures on June 19, 2008. Before Series A, there were 26,995,354 outstanding shares. Siri issued 23,135,294 new shares to Series A investors at the price of $0.37 per share, putting the company's post-money valuation at $18.53 million. After the deal was closed, Series A investors own 46.15% of the company's shares.

Then it raised $15.5 million of Series B venture funding from SRI Ventures, Morgenthaler Ventures and Menlo Ventures on November 24, 2009. Li Ka-shing also participated in the round. Siri issued 15,500,000 new shares to Series B investors at the price of $1 per share, putting the company's post-money valuation at $65.91 million. After the deal was closed, Series B investors own 23.62% of the company's shares, and Series A investors' ownership was diluted to 35.25%.

Siri was acquired by Apple (NAS: AAPL) for $200 million on April 28, 2010. Previous investors sold all their shares and exited in full.

#	Deal Type	Date	Amount	Pre-Val	Post-Val	Ownership, %	Diluted Ownership, %
1	Series A	2008-06-19	8.55	9.98	18.53	0.46	0.35
2	Series B	2009-11-24	15.5	50.41	65.91	0.24	0.24
3	M&A	2010-04-28	200	200	200	1.00	1.00

Table 11: This table presents the VC financing rounds and exit outcome of the software company Siri. *Amount, Pre-Val, Post-Val* are converted to 2010 Million USD.

A.2 DogRadar

DogRadar is a developer of a gaming application designed to offer dog-friendly adventure games. The company's application offers a series of mysterious puzzles and brain teasers, enabling users to organize park playdates for their dogs with like-minded people.

DogRadar was founded in 2017 in Los Angeles. Its primary industry is *Entertainment Software*, it belongs to the verticals of *Pet Technology* and *Mobile*. DogRadar was not a patent assignee.

The company joined the accelerator of Hiventures on June 27, 2018 and received $9 million in funding, puttiing the company's post-money valuation at $100 million. After the deal, Hiventures owned 9% of the company's shares. The company went out of business on November 12, 2021.

A.3 Talphera

Talphera is a specialty pharmaceutical company focused on the development and commercialization of therapies for use in medically supervised settings. Its product portfolio includes DSUVIA and Zalviso for moderate-to-severe acute pain.

Talphera was founded in 2005 in Hayward, CA. Its primary industry is *Pharmaceuticals* with keywords *Pain Therapeutic*. So far, it has been assigned 12 patents.

The company raised $21.14 million of Series A venture funding from undisclosed investors on August 30, 2006. Before Series A, there were 4,256,739 outstanding shares. The company issued 8,456,581 new shares to Series A investors at the price of $2.5 per share, putting the company's post-money valuation at $31.14 million. After the deal was closed, Series A investors own 66.56% of the company's shares.

The company raised $20.22 million of Series B venture funding from Three Arch Partners, Skyline Venture Partners, Alta Partners and Kaiser Foundation Hospitals on February 19, 2008. The company issued 5,054,544 new shares to Series B investors at the price of $4 per share, putting the company's post-money valuation at $72.22 million. After the deal was closed, Series B investors own 28.0% of the company's shares.

Then the company raised $14.81 million of Series C venture funding from Three Arch Partners, Pinnacle Ventures, Skyline Venture Partners, Alta Partners and Kaiser Foundation Hospitals on November 23, 2009. The company issued 15,028,106 new shares to Series B investors at the price of $0.99 per share, putting the company's post-money valuation at $52.27 million. After the deal was closed, Series C investors own 28.33% of the company's shares.

The company raised $40 million in its initial public offering on the NASDAQ under the ticker symbol of ACRX on February 11, 2011. A total of 8,000,000 shares were sold at a price of $5 per share. All VC investors' shares were converted to common stock. After the offering, there was a total of 19,371,750 outstanding shares (excluding the over-allotment option) priced at $5 per share, valuing the company at $96.85 million. The underwriters were granted an option to purchase up to an additional 1,200,000 shares from the company to cover over-allotments, if any. The close price on the first IPO day was $91.00 per share.

A.4 3D Biomatrix

3D Biomatrix is a manufacturer of three-dimensional systems designed to facilitate cellular assays in drug discovery. The company's systems uses 3D sheroid culture through bone-marrow, hepatocytes and thymus applications, enabling researchers and drug discovery experts to treat cancer and tumor.

3D Biomatrix was founded in 2010 in Ann Arbor, MI. Its primary industry is *Pharmaceuticals* and it belongs to the vertical of *Oncology*. It was assigned 1 patent.

The company received $545,000 of grant funding from Department of Health & Human Services in 2013.

The company raised $1.47 million of Series A funding from Ann Arbor Spark, Biosciences

Research and Commercialization Center and other investors on February 5, 2013. The company issued 789,536 new shares to Series A investors at the price of \$1.86 per share, putting the company's post-money valuation at \$3.49 million. After the deal was closed, Series A investors own 42.1% of the company's shares.

The company went out of business in October 2015.

B Variable Description

Table 12: **Variable Description**

Variables	**Description**	**Source**
A. Startup level		
Sector	1 if a biotech company, 0 if a software company	PB
Patent	The number of patents assigned to the startup	PB
IPO	1 if the startup exited via IPO, 0 otherwise	PB
MA	1 if the startup exited via M&A, 0 otherwise	PB
Fail	1 if the startup went bankrupt or out of business, 0 otherwise	PB
ExitValue	The realized value of the startup. It is the close price on the first IPO day multiplied by outstanding common shares immediately after IPO for IPO companies, acquisition price for M&A companies, and 0 for failed companies	PB, SEC, CT, FS
B. Startup-round level		
Round	The chronological order of the VC round, e.g., for the 3rd VC round, *Round*=3	PB
CapitalAmount	The capital investment injected to a startup in a round	PB
PreVal	Price per share multiplied by total outstanding shares before the new issuance	PB
PostVal	Price per share multiplied by total outstanding shares after the new issuance	PB
EquityOwnership	The proportion of shares held by this round's investor relative to the total number of outstanding shares, it also equals to $\frac{CapitalAmount}{PostVal}$.	PB

Note: PB, SEC, CT, FS are abbreviations for Pitchbook, SEC Filings, Compustat and FactSet. Startup is indexed by i, round is indexed by k.

C Summary Statistics

C.1 Startup-round level

Table 13: **Summary statistics at startup-round level**

	Round	count	mean	std	min	25%	50%	75%	max
	1	3,664	7.97	15.62	0.001	1.18	3.08	8.50	270.47
	2	2,321	16.19	28.36	0.001	3.19	8.59	19.05	810.09
	3	1,446	22.70	30.83	0.02	5.24	13.55	29.87	400.78
CapitalAmount	4	853	26.99	39.69	0.02	6.68	17.01	35.89	700.97
	5	456	35.54	43.68	0.004	7.90	21.79	49.92	333.15
	6	207	48.50	76.89	0.14	9.72	23.50	57.27	742.54
	7	62	44.87	54.72	0.31	9.11	29.63	55.52	317.54
	1	2,942	22.35	62.89	0.03	5.81	11.33	23.14	1,797.41
	2	1,959	57.94	119.03	0.84	15.90	31.93	64.42	2,960.44
	3	1,204	108.81	185.83	1.85	28.63	59.85	122.78	3,295.58
PostVal	4	701	171.71	309.12	0.66	48.53	99.86	182.48	5,252.16
	5	391	266.91	456.77	0.96	65.60	132.43	279.55	4,757.84
	6	171	474.27	914.57	1.86	86.21	159.56	381.66	5,728.13
	7	48	484.40	685.77	26.23	103.28	212.83	520.30	3,825.49
	1	2,942	0.34	0.19	0.001	0.20	0.30	0.46	0.97
	2	1,959	0.31	0.17	0.004	0.19	0.28	0.40	0.92
	3	1,204	0.26	0.16	0.003	0.15	0.23	0.35	0.93
EquityOwnership	4	701	0.23	0.16	0.001	0.11	0.19	0.30	0.88
	5	391	0.21	0.15	0.01	0.10	0.17	0.28	0.87
	6	171	0.18	0.13	0.02	0.08	0.13	0.25	0.68
	7	48	0.15	0.10	0.02	0.08	0.13	0.20	0.48
	0	208	153.66	925.54	0.01	4.37	17.46	55.65	12875.85
	1	366	152.52	396.51	0.003	13.68	44.02	137.45	4588.70
	2	437	235.33	593.15	0.21	28.33	95.61	241.81	7938.39
ExitValue	3	403	313.37	621.38	0.16	49.89	145.92	360.10	8820.16
	4	295	348.76	703.51	0.01	68.25	197.06	388.53	8442.44
	5	187	683.75	1656.25	0.53	92.00	284.54	539.81	15501.80
	6	114	1224.08	3473.23	2.13	113.30	353.53	736.37	31318.33
	7	62	563.74	756.68	7.17	63.31	267.08	625.04	3226.83

Note: *CapitalAmount*, *PostVal* and *ExitValue* are in million USD as of Sep 2023. *EquityOwnership* is between 0 and 1.

C.2 Startup level

Table 14: **Summary statistics at startup level**

Variable	Category	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Sector		0.43	0.49	0	0	0	1	1
Patent		6.86	95.37	0	0	0	2	6,515
	IPO	0.11	0.31	0	0	0	0	1
Exit	*MA*	0.27	0.44	0	0	0	1	1
	Fail	0.63	0.48	0	0	1	1	1
ExitValue	*Unadjusted*	348.53	1,149.49	0.003	28.90	107.50	318.04	31,318.33
	Adjusted	129.16	719.63	0.00	0.00	0.00	45.13	31,318.33

Note: *Sector, IPO, MA, Fail* are dummy variables. *ExitValue* is in million USD as of September 2023. Unadjusted (adjusted) *ExitValue* is before (after) filling 0 for failed companies. The minimum exit value comes from *AND CO*, a software company acquired in Jan 2018 at $3,250. The number of observations are all 5,591 except for *(unadjusted) ExitValue* that has 2,072 observations.

D Covariance Matrix

The posterior mean and variance are

$$\mu_k = \omega_0^k \cdot 0 + \sum_{l=1}^{k} \omega_l^k \bar{s}_l \tag{24}$$

$$\tau_k^2 = \omega_0^k \sigma_\xi^2 \tag{25}$$

where $\bar{s}_k$ is the mean of $\boldsymbol{s}_l$ and ω_l^k is the weight of round l's mean signal at round k. To characterize the weights, let $\alpha_0 = 1$ and $\alpha_l = \frac{\frac{1}{J_l}\sigma_l^2}{\frac{1}{J_l}\sigma_l^2 + (\prod_{l'=0}^{l-1}\alpha_{l'})\sigma^2}, l = 1, ..., k$. Then

$$\omega_l^k = \begin{cases} \prod_{l=1}^{k} \alpha_l & l = 0 \\ (1-\alpha_l)\prod_{l'=l+1}^{k} \alpha_{l'} & 1 \le l < k \\ 1 - \alpha_k & l = k \end{cases}$$

Note that $\sum_{l=0}^{k} \omega_l^k = 1$.

The unobserved quality ξ and posterior mean μ_k are correlated. By Eq. (24),

$$\mu_k = \sum_{l=1}^{k} \omega_l^k (\xi + \bar{\epsilon}_l) = (1 - \omega_0^k)\xi + \sum_{l=1}^{k} \omega_l^k \bar{\epsilon}_l$$

Hence,

$$\begin{pmatrix} \xi \\ \mu_1 \\ \vdots \\ \mu_{\bar{K}} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1-\omega_0^1 & \omega_1^1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1-\omega_0^{\bar{K}} & \omega_1^{\bar{K}} & \cdots & \omega_{\bar{K}}^{\bar{K}} \end{pmatrix}}_{\Omega} \begin{pmatrix} \xi \\ \bar{\epsilon}_1 \\ \vdots \\ \bar{\epsilon}_{\bar{K}} \end{pmatrix}$$

So that the covariance matrix

$$\Sigma = \Omega \begin{pmatrix} \sigma_\xi^2 & & & \\ & \frac{\sigma_1^2}{J_1} & & \\ & & \ddots & \\ & & & \frac{\sigma_1^2}{J_{\bar{K}}} \end{pmatrix} \Omega'$$

E Proof of Proposition 2

Assuming that $V_k(\boldsymbol{X}_k, \mu_k)$ is increasing in μ_k, we can evoke Theorem 10, the existence of a symmetric equilibrium in an English auction game, in Milgrom and Weber (1982). In this equilibrium, each bidder bids her maximum gain knowing the true signals of the dropped-out bidders and assuming that all other active bidders receive the same signal as herself. Because the strategies are revealing in equilibrium, the true signals of the dropped-out bidders become known to the remaining bidders. Whenever a bidder drops out, all remaining bidders back out her true signal by observing the equity level at which she drops out, and update their bids accordingly. This process iterates until only one bidder is left.

Since $V_k(\boldsymbol{X}_k, \mu_k)$ is increasing in μ_k, and as is shown in Eq. (24), μ_k is increasing in each bidder's signal, in equilibrium, bidders drop out one by one in the ascending order of their signals. When only two bidders remain, they are the ones who received the top two signals. The bidder with the second-highest signal exits first at the equity level that makes her expected net return zero, assuming that the other bidder received the same signal as she did. Hence, the winning equity bid a_k^* satisfies Eq. (8).

The winner knows all signals when the second-highest bidder drops out, and would only fund the startup if the net expected return is non-negative conditional on all her information. Hence, the choice equation satisfies Eq. (7)

Next, we prove that under Assumption 1 and 2, V_k is strictly increasing in μ_k for any

$k \in \{1, ..., \bar{K}\}$. With a sequence of English auctions, $V_k(\boldsymbol{X}_k, \mu_k)$ not only depends on the current information but also the equilibrium strategies of future investors, thus it is not necessarily increasing in μ_k for any possible strategies of future investors. Hence, we prove by backward induction.

When $k = \bar{K}$. By Eq. (6),

$$\begin{aligned} V_{\bar{K}}(\boldsymbol{X}_{\bar{K}}, \mu_{\bar{K}}) &= \mathbb{E}\left[W_{\bar{K}} \middle| \boldsymbol{X}_{\bar{K}}, \mu_{\bar{K}}\right] \\ &= \mathbb{E}\left[m_{\bar{K}}(\boldsymbol{X}_{\bar{K}}) \exp(\xi) \middle| \boldsymbol{X}_{\bar{K}}, \mu_{\bar{K}}\right] \\ &= m_{\bar{K}}(\boldsymbol{X}_{\bar{K}}) \exp\left(\mu_{\bar{K}} + \frac{1}{2}\sigma_{\bar{K}}^2\right) \end{aligned} \tag{26}$$

The first equality is because $\bar{K}$ is the maximum possible round, the third equality is by the property of log-normal distribution. Hence, it's straightforward that $V_{\bar{K}}(\boldsymbol{X}_{\bar{K}}, \mu_{\bar{K}})$ is strictly increasing in $\mu_{\bar{K}}$.

When $k < \bar{K}$. We assume that $V_{k+1}(\boldsymbol{X}_{k+1}, \mu_{k+1})$ is strictly increasing in μ_{k+1}. As a start, we write V_k in its Bellman representation.

$$\begin{aligned} V_k(\boldsymbol{X}_k, \mu_k) =& \Pr(K = k \big| K \geq k, \boldsymbol{X}_k, \mu_k) \mathbb{E}\left[W_K \big| \boldsymbol{X}_k, \mu_k, K = k\right] \\ &+ \Pr(K > k \big| K \geq k, \boldsymbol{X}_k, \mu_k) \mathbb{E}\left[(1 - a_{k+1}) q_{k+1} V_{k+1}(\boldsymbol{X}_{k+1}, \mu_{k+1}) \big| \boldsymbol{X}_k, \mu_k, K > k\right] \\ =& \Pr(K = k \big| K \geq k, \boldsymbol{X}_k) m_k(\boldsymbol{X}_k) \exp\left(\mu_k + \frac{1}{2}\sigma_k^2\right) \\ &+ \Pr(K > k \big| K \geq k, \boldsymbol{X}_k) \delta \mathbb{E}\left[(1 - a_{k+1}) q_{k+1} V_{k+1}(\boldsymbol{X}_{k+1}, \mu_{k+1}) \big| \boldsymbol{X}_k, \mu_k, K > k\right] \end{aligned} \tag{27}$$

By Assumption 1 and 2, K doesn't depend on μ_k, thus we drop μ_k from the probability. From the above expression, a sufficient condition for $V_k(\boldsymbol{X}_k, \mu_k)$ to be strictly increasing in μ_k is that the following function is strictly increasing in μ_k.

$$\mathbb{E}\left[(1 - a_{k+1}) q_{k+1} V_{k+1}(\boldsymbol{X}_{k+1}, \mu_{k+1}) \big| \boldsymbol{X}_k, \mu_k, K > k\right] \tag{28}$$

As we have shown above, if $V_{k+1}(\boldsymbol{X}_{k+1}, \mu_{k+1})$ is strictly increasing in μ_{k+1}, then the equilibrium strategy at round $k+1$ follows Eq. (7) and (8). Plugging it to Eq. (28),

$$\mathbb{E}\left[\underbrace{\left(1-\frac{d_{k+1}}{V_{k+1}(\boldsymbol{X}_{k+1},\hat{\mu}_{k+1})}\right)\mathbb{1}\{V_{k+1}(\boldsymbol{X}_{k+1},\mu_{k+1})>d_{k+1}\}V_{k+1}(\boldsymbol{X}_{k+1},\mu_{k+1})}_{\dot{V}_{k+1}(\boldsymbol{X}_{k+1},\mu_{k+1},\hat{\mu}_{k+1})}\Big|\boldsymbol{X}_k,\mu_k,K>k\right] \tag{29}$$

The expectation is with respect to $\boldsymbol{X}_{k+1},\mu_{k+1},\hat{\mu}_{k+1}$. In the following, we ignore $\boldsymbol{X}$ because we only care about how Eq. (29) changes with μ_k, and by Assumption 1 and 2, μ is independent of $\boldsymbol{X}$.

Under the assumption that $V_{k+1}(\boldsymbol{X}_{k+1},\mu_{k+1})$ is strictly increasing in μ_{k+1}, it's easy to show that $\dot{V}_{k+1}(\boldsymbol{X}_{k+1},\mu_{k+1},\hat{\mu}_{k+1})$ is weakly increasing in μ_{k+1} and $\hat{\mu}_{k+1}$. We prove that Eq. (29) is strictly increasing in μ_k with the help of the following two lemmas.

Lemma 1. *For any $\mu_k>\mu'_k$, $(\mu_{k+1},\hat{\mu}_{k+1})$ first order stochastically dominates $(\mu'_{k+1},\hat{\mu}'_{k+1})$, that is, for any real number a and b,*

$$\Pr(\mu_{k+1}\le a,\hat{\mu}_{k+1}\le b\mid\mu_k)<\Pr(\mu'_{k+1}\le a,\hat{\mu}'_{k+1}\le b\mid\mu'_k)$$

Lemma 2. *Consider two random variables $x,y\in\mathbb{R}^2$. Suppose $u(x,y)$ is weakly increasing in x and y in general, but there exists a region of non-zero measure where $u(x,y)$ is strictly increasing in x and y. Moreover, (x,y) first order stochastically dominates (x',y'), then*

$$\mathbb{E}\left[u(x,y)\right]>\mathbb{E}\left[u(x',y')\right]$$

Combining these two lemmas, it's straightforward that Eq. (29) is strictly increasing in μ_k. Hence, $V_k(\boldsymbol{X}_k,\mu_k)$ is strictly increasing in μ_k.

F Proof of Proposition 3

For notation simplicity, we write V_k and $\hat{V}_k$ as the shorthand for $V_k(\boldsymbol{X}_k,\mu_k)$ and $V_k(\boldsymbol{X}_k,\hat{\mu}_k)$. Hence, by Eq. (27) and (29), the Bellman representation for the *maximum* possible payoff to round k investor, V_k^*, is as follows:

$$V_k^* = \Pr(K = k \big| K \geq k, \boldsymbol{X}_k) m_k(\boldsymbol{X}_k) \exp\left(\mu_k + \frac{1}{2}\sigma_k^2\right) + \Pr(K > k \big| K \geq k, \boldsymbol{X}_k) \delta \mathbb{E}\left[\mathbb{1}\{V_{k+1}^* \geq d_{k+1}\}\left(V_{k+1}^* - \frac{V_{k+1}^*}{\hat{V}_{k+1}^*} d_{k+1}\right) \big| \boldsymbol{X}_k, \mu_k, K > k\right] \quad (30)$$

The Bellman representation for $V_k^\dagger$ when the social planner makes decisions is

$$V_k^\dagger = \Pr(K = k \big| K \geq k, \boldsymbol{X}_k) m_k(\boldsymbol{X}_k) \exp\left(\mu_k + \frac{1}{2}\sigma_k^2\right) + \Pr(K > k \big| K \geq k, \boldsymbol{X}_k) \delta \mathbb{E}\left[\mathbb{1}\{V_{k+1}^\dagger \geq d_{k+1}\}\left(V_{k+1}^\dagger - d_{k+1}\right) \big| \boldsymbol{X}_k, \mu_k, K > k\right] \quad (31)$$

When $k = \bar{K}$, $V_k^* = V_k^\dagger$ is straightforward. Next, we show that for any $(\boldsymbol{X}_k, \mu_k)$, $V_k^* < V_k^\dagger$. If $V_{k+1}^* \leq V_{k+1}^\dagger$ for any $(\boldsymbol{X}_{k+1}, \mu_{k+1})$, then

$$\begin{aligned} \mathbb{1}\{V_{k+1}^* \geq d_{k+1}\}\left(V_{k+1}^* - \frac{V_{k+1}^*}{\hat{V}_{k+1}^*} d_{k+1}\right) &\leq \mathbb{1}\{V_{k+1}^* \geq d_{k+1}\}\left(V_{k+1}^* - d_{k+1}\right) \\ &\leq \mathbb{1}\{V_{k+1}^\dagger \geq d_{k+1}\}\left(V_{k+1}^\dagger - d_{k+1}\right) \end{aligned}$$

The first inequality is because V_k^* is strictly increasing in μ_k for any k as is suggested by proposition 2. The second inequality is by $V_{k+1}^* \leq V_{k+1}^\dagger$. Moreover, $\exists(\boldsymbol{X}_{k+1}, \mu_{k+1})$, such that

$$\mathbb{1}\{V_{k+1}^* \geq d_{k+1}\}\left(V_{k+1}^* - \frac{V_{k+1}^*}{\hat{V}_{k+1}^*} d_{k+1}\right) < \mathbb{1}\{V_{k+1}^\dagger \geq d_{k+1}\}\left(V_{k+1}^\dagger - d_{k+1}\right)$$

Hence, $V_k^* < V_k^\dagger$ for any $(\boldsymbol{X}_k, \mu_k)$. By backward induction, it's straightforward that for any $(\boldsymbol{X}_k, \mu_k)$, $V_k^* = V_k^\dagger$ when $k = \bar{K}$, $V_k^* < V_k^\dagger$ when $k < \bar{K}$.

G Proof of Proposition 5

Proof. We start by proving that the listed conditions are satisfied in my model. The first two conditions are satisfied under Assumption 1 and 2.

To prove the third condition, it's straightforward that $h_1(\boldsymbol{X}_1), ..., h_{\bar{K}}(\boldsymbol{X}_{\bar{K}})$ have full rank because a new variable, d_k, enters $\boldsymbol{X}_k$ in each round, and d_k cannot be deterministically inferred from $\boldsymbol{X}_{k-1}$. Likewise, $h_1(\boldsymbol{X}_1), ..., h_{\bar{K}-1}(\boldsymbol{X}_{\bar{K}-1}), \ln(m_{\bar{K}}(\boldsymbol{X}_{\bar{K}}))$ have full rank. Next, we show that $\ln(m_{\bar{K}}(\boldsymbol{X}_{\bar{K}}))$ and $h_{\bar{K}}(\boldsymbol{X}_{\bar{K}})$ have full rank. To see this, by Eq. (26),

$$V_{\bar{K}}^*(\boldsymbol{X}_{\bar{K}}, \mu_{\bar{K}}) = m_{\bar{K}}(\boldsymbol{X}_{\bar{K}}) \exp\left(\mu_{\bar{K}} + \frac{1}{2}\tau_{\bar{K}}^2\right)$$

Then it's easy to derive $h_{\bar{K}}$

$$h_{\bar{K}}(\boldsymbol{X}_{\bar{K}}) = \ln d_{\bar{K}} - \ln(m_{\bar{K}}(\boldsymbol{X}_{\bar{K}})) - \frac{1}{2}\tau_{\bar{K}}^2$$

Since $d_{\bar{K}} \in \mathbb{R}_+$, then $h_{\bar{K}}(\boldsymbol{X}_{\bar{K}})$ can attain its full support conditional on $\ln(m_{\bar{K}}(\boldsymbol{X}_{\bar{K}}))$. Hence, condition 3 is established.

To show condition 4, note that h_k is continuous and μ_k is supported on $\mathbb{R}$. So we only need to show that (1) $\exists \boldsymbol{X}_k$ such that for any μ_k, $h_k(\boldsymbol{X}_k) \leq \mu_k$, i.e., a startup will never be funded regardless of its μ_k; (2) $\exists \boldsymbol{X}_k$ such that for any μ_k, $h_k(\boldsymbol{X}_k) \geq \mu_k$, i.e., a startup will always be funded regardless of its μ_k. It's easy to show that (1) is established when $d_k \to \infty$, and (2) is established when $d_k \to 0$. Hence, condition 4 is established.

We can evoke Theorem 2, the sufficient condition for identifying the outcome equations, choice equations and joint distribution of unobservables, in Heckman and Navarro (2007) to identify $\boldsymbol{m}, \sigma_\xi, \boldsymbol{\sigma}, \boldsymbol{h}$.

□

H Proof of Proposition 6

Proof. I show that $\boldsymbol{V}^*$ is identified non-parametrically. By Eq. (13),

$$V_k^*(\boldsymbol{X}_k, \hat{\mu}_k) = \frac{d_k}{a_k^*}$$

We have shown in Proposition 2 that V_k^* is increasing in the second argument (here $\hat{\mu}_k$). Hence, for any $\boldsymbol{X}_k$, we can identify the χ quantile of a_k^*, which we denote as $a_{k,\chi}^*$ with $\chi \in [0, 1]$. Moreover, since $\sigma_\xi, \boldsymbol{\sigma}$ are identified, then we are able to back out the distribution for $\hat{\mu}_k$. Likewise, we denote the χ quantile of $\hat{\mu}_k$ as $\hat{\mu}_{k,\chi}$. Hence,

$$V_k^*(\boldsymbol{X}_k, \hat{\mu}_{k,\chi}) = \frac{d_k}{a^*_{k,1-\chi}}$$

By varying $\hat{\mu}_{k,\chi}$, we are able to non-parametrically identify V_k^* on its full support . □

I Proof of Proposition 7

The proof is composed of two steps.

Step 1: We first identify the distribution of $\boldsymbol{X}_k$ conditional on $K = k$.

$$\begin{aligned}
&\Pr\left(\boldsymbol{X}_k \middle| \mu_1 \geq h_1(\boldsymbol{X}_1), ..., \mu_k \geq h_k(\boldsymbol{X}_k); K = k\right) \\
&= \frac{\Pr\left(\mu_1 \geq h_1(\boldsymbol{X}_1), ..., \mu_k \geq h_k(\boldsymbol{X}_k) \middle| \boldsymbol{X}_k; K = k\right) f(\boldsymbol{X}_k | K = k)}{\Pr\left(\mu_1 \geq h_1(\boldsymbol{X}_1), ..., \mu_k \geq h_k(\boldsymbol{X}_k) \middle| K = k\right)} \\
&= \frac{\Pr\left(\mu_1 \geq h_1(\boldsymbol{X}_1), ..., \mu_k \geq h_k(\boldsymbol{X}_k) \middle| \boldsymbol{X}_k; K = k\right) f(\boldsymbol{X}_k | K = k)}{c_k}
\end{aligned}$$

where $c_k = \Pr\left(\mu_1 \geq h_1(\boldsymbol{X}_1), ..., \mu_k \geq h_k(\boldsymbol{X}_k) \middle| K = k\right)$ is a constant. Hence,

$$f(\boldsymbol{X}_k | K = k) = \frac{\Pr\left(\boldsymbol{X}_k \middle| \mu_1 \geq h_1(\boldsymbol{X}_1), ..., \mu_k \geq h_k(\boldsymbol{X}_k); K = k\right)}{\Pr\left(\mu_1 \geq h_1(\boldsymbol{X}_1), ..., \mu_k \geq h_k(\boldsymbol{X}_k) \middle| \boldsymbol{X}_k; K = k\right)} c_k$$

Note that the numerator is identified directly from data, the denominator can be computed given $\sigma_\xi, \boldsymbol{\sigma}, \boldsymbol{h}$. Therefore, $f(\boldsymbol{X}_k | K = k)$ is known up to c_k. With an additional condition of $\int f(\boldsymbol{X}_k | K = k) d\boldsymbol{X}_k$=1, c_k is identified.

Step 2: Next, we prove that $\Pr(K = k)$ is identified $\forall k \in \{1, ..., \bar{K}\}$. Recall the Bellman equation in Eq. (30) that connects V_k^* and V_{k+1}^*. Note that in Eq. (30),

1. $\Pr(K = k \big| K \geq k, \boldsymbol{X}_k)$ is known up to $\Pr(K = k')$ where $k' \geq k$. This is because

$$
\begin{aligned}
\Pr(K = k \big| K \geq k, \boldsymbol{X}_k) &= \frac{f(\boldsymbol{X}_k \big| K = k) \Pr(K = k \mid K \geq k)}{\sum_{k' \geq k} f(\boldsymbol{X}_{k'} \big| K = k') \Pr(K = k' \mid K \geq k)} \\
&= \frac{f(\boldsymbol{X}_k \big| K = k) \Pr(K = k)}{\sum_{k' \geq k} f(\boldsymbol{X}_{k'} \big| K = k') \Pr(K = k')}
\end{aligned}
$$

   $f(\boldsymbol{X}_k \big| K = k)$ is identified from **Step 1**, hence $\Pr(K = k \big| K \geq k, \boldsymbol{X}_k)$ is known up to $\Pr(K = k')$ where $k' \geq k$.

2. $\mathbb{E}\left[\mathbb{1}\{V^*_{k+1} \geq d_{k+1}\} \left(V^*_{k+1} - \frac{V^*_{k+1}}{\hat{V}^*_{k+1}} d_{k+1}\right) \big| \boldsymbol{X}_k, \mu_k, K > k\right]$ is known up to $\Pr(K = k')$ for all $k' > k$. This is because given $\beta, \sigma_\xi, \boldsymbol{\sigma}, \boldsymbol{V}^*$, it's easy to show that this expectation is known up to $f(\boldsymbol{X}_{k+1} \big| \boldsymbol{X}_k, K > k)$, which is known up to $\Pr(K = k')$ where $k' > k$. To see why,

$$
f(\boldsymbol{X}_{k+1} \big| \boldsymbol{X}_k, K > k) = \sum_{k' > k} f(\boldsymbol{X}_{k+1} \big| \boldsymbol{X}_k, K = k') \Pr(K = k' \big| \boldsymbol{X}_k, K > k)
$$

   $f(\boldsymbol{X}_{k+1} \big| \boldsymbol{X}_k, K = k')$ is identified from **Step 1**. $\Pr(K = k' \big| \boldsymbol{X}_k, K > k)$ can be rewritten as below:

$$
\Pr(K = k' \big| K > k, \boldsymbol{X}_k) = \frac{f(\boldsymbol{X}_k \big| K = k') \Pr(K = k')}{\sum_{k'' > k} f(\boldsymbol{X}_{k''} \big| K = k'') \Pr(K = k'')}
$$

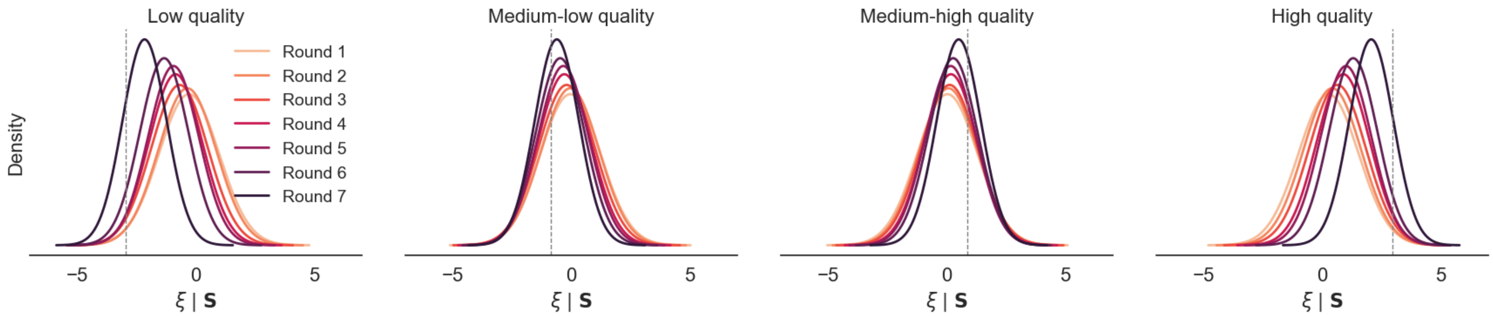
   Hence it's known up to $\Pr(K = k')$ for all $k' > k$.

Hence, by varying the values of $\boldsymbol{X}_k$ and μ_k in V^*_k, we are able to trace back the values of $\Pr(K = k')$ for all $k' \geq k$.

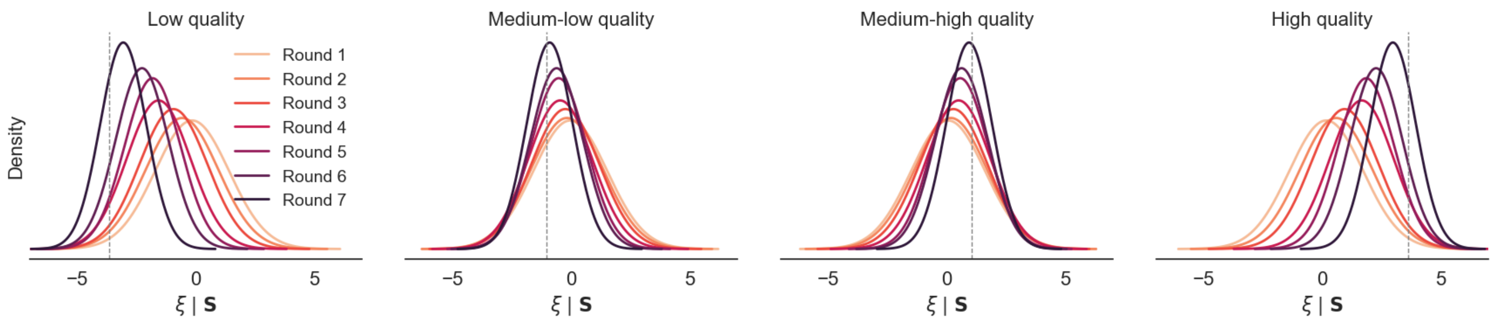
J Learning Process

Figure 14: **Learning process**

(a) Biotech



(b) Software



Note: This figure displays, for different levels of ξ, the posterior distribution after each round. The x-axis represents the belief on ξ conditional on the signals $\boldsymbol{S}$. The grey dashed lines mark the true values of ξ. The 1st, 25th, 75th, and 99th percentiles of ξ are selected to represent "low," "medium-low," "medium-high," and "high" levels.

References

Nikhil Agarwal. An empirical model of the medical match. *American Economic Review*, 105 (7):1939–1978, 2015.

Philippe Aghion, Céline Antonin, and Simon Bunel. *The power of creative destruction: Economic upheaval and the wealth of nations*. Harvard University Press, 2021.

Ufuk Akcigit and John Van Reenen. *The Economics of Creative Destruction*. Cambridge: Harvard University Press.[644], 2023.

Joseph G Altonji and Charles R Pierret. Employer learning and statistical discrimination. *The quarterly journal of economics*, 116(1):313–350, 2001.

Yonghong An, Yingyao Hu, and Matthew Shum. Estimating first-price auctions with an unknown number of bidders: A misclassification approach. *Journal of Econometrics*, 157 (2):328–341, 2010.

Gaurab Aryal, Manudeep Bhuller, and Fabian Lange. Signaling and employer learning with instruments. *American Economic Review*, 112(5):1669–1702, 2022.

Gaurab Aryal, Charles Murry, and Jonathan W Williams. Price discrimination in international airline markets. *Review of Economic Studies*, 91(2):641–689, 2024.

Randolph P Beatty and Jay R Ritter. Investment banking, reputation, and the underpricing of initial public offerings. *Journal of financial economics*, 15(1-2):213–232, 1986.

Shai Bernstein, Xavier Giroud, and Richard R Townsend. The impact of venture capital monitoring. *The Journal of Finance*, 71(4):1591–1622, 2016.

Nicholas Bloom, Charles I Jones, John Van Reenen, and Michael Webb. Are ideas getting harder to find? *American Economic Review*, 110(4):1104–1144, 2020.

WD Bygrave. Venture capital at the crossroads, 1992.

Peng Chen, Gary T Baierl, and Paul D Kaplan. Venture capital and its role in strategic asset allocation. *Frontiers of Modern Asset Allocation*, pages 179–190, 2012.

John H Cochrane. The risk and return of venture capital. *Journal of financial economics*, 75(1):3–52, 2005.

Peter M DeMarzo, Ilan Kremer, and Andrzej Skrzypacz. Bidding with securities: Auctions and security design. *American economic review*, 95(4):936–959, 2005.

Joseph A DiMasi and Henry G Grabowski. R&d costs and returns to new drug development: a review of the evidence. *The Oxford Handbook of the economics of the biopharmaceutical industry*, pages 21–46, 2012.

The Economist. Can biotech startups upstage eli lilly and novo nordisk? 2024. URL https://www.economist.com/business/2024/04/30/can-biotech-startups-upstage-eli-lilly-and-novo-nordisk.

Michael Ewens, Ramana Nanda, and Matthew Rhodes-Kropf. Cost of experimentation and the evolution of venture capital. *Journal of Financial Economics*, 128(3):422–442, 2018.

Michael Ewens, Alexander Gorbenko, and Arthur Korteweg. Venture capital contracts. *Journal of Financial Economics*, 143(1):131–158, 2022.

Food and Drug Administration. The drug development process. 2018. URL https://www.fda.gov/patients/learn-about-drug-and-device-approvals/drug-development-process.

Mark J Garmaise. Informed investors and the financing of entrepreneurial projects. *Available at SSRN 263162*, 2001.

Alexander S Gorbenko and Andrey Malenko. Competition among sellers in securities auctions. *American Economic Review*, 101(5):1806–1841, 2011.

Will Gornall and Ilya A Strebulaev. The economic impact of venture capital: Evidence from public companies. *Available at SSRN 2681841*, 2021.

Will Gornall and Ilya A Strebulaev. The contracting and valuation of venture capital-backed companies. In *Handbook of the Economics of Corporate Finance*, volume 1, pages 3–76. Elsevier, 2023.

Christian Gourieroux and Alain Monfort. *Simulation-based econometric methods*. OUP Oxford, 1997.

Robert S Harris, Tim Jenkinson, and Steven N Kaplan. Private equity performance: What do we know? *The Journal of Finance*, 69(5):1851–1882, 2014.

James J Heckman and Salvador Navarro. Dynamic discrete choice and dynamic treatment effects. *Journal of Econometrics*, 136(2):341–396, 2007.

Sabrina T Howell, Josh Lerner, Ramana Nanda, and Richard R Townsend. Financial distancing: How venture capital follows the economy down and curtails innovation. 2020.

Helu Jiang and Faisal Sohail. Skill-biased entrepreneurial decline. *Review of Economic Dynamics*, 48:18–44, 2023.

Steven N Kaplan and Antoinette Schoar. Private equity performance: Returns, persistence, and capital flows. *The journal of finance*, 60(4):1791–1823, 2005.

Steven N Kaplan and Per Strömberg. Financial contracting theory meets the real world: An empirical analysis of venture capital contracts. *The review of economic studies*, 70(2): 281–315, 2003.

William R Kerr, Ramana Nanda, and Matthew Rhodes-Kropf. Entrepreneurship as experimentation. *Journal of Economic Perspectives*, 28(3):25–48, 2014.

Patrick Kline, Neviana Petkova, Heidi Williams, and Owen Zidar. Who profits from patents? rent-sharing at innovative firms. *The quarterly journal of economics*, 134(3):1343–1404, 2019.

Leonid Kogan, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. Technological innovation, resource allocation, and growth. *The quarterly journal of economics*, 132(2): 665–712, 2017.

Julian Kolev, Alexis Haughey, Fiona Murray, and Scott Stern. Of academics and creative destruction: Startup advantage in the process of innovation. Technical report, National Bureau of Economic Research, 2022.

Fabian Lange. The speed of employer learning. *Journal of Labor Economics*, 25(1):1–35, 2007.

Josh Lerner and Ramana Nanda. Venture capitalâs role in financing innovation: What we know and how much we still need to learn. *Journal of Economic Perspectives*, 34(3): 237–261, 2020.

Yao Luo and Ruli Xiao. Identification of auction models using order statistics. *Journal of Econometrics*, 236(1):105457, 2023.

R. Preston McAfee and Daniel Vincent. Sequentially Optimal Auctions. *Games and Economic Behavior*, 18(2):246–276, February 1997. URL https://ideas.repec.org/a/eee/gamebe/v18y1997i2p246-276.html.

Daniel McFadden. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society*, pages 995–1026, 1989.

Paul R Milgrom and Robert J Weber. A theory of auctions and competitive bidding. *Econometrica: Journal of the Econometric Society*, pages 1089–1122, 1982.

Roger B Myerson and Mark A Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of economic theory*, 29(2):265–281, 1983.

Mariann Ollar, Marzena J Rostek, and Ji Hee Yoon. Privacy in markets. *Available at SSRN 3071374*, 2021.

Ariel Pakes and David Pollard. Simulation and the asymptotics of optimization estimators. *Econometrica: Journal of the Econometric Society*, pages 1027–1057, 1989.

Michael Park, Erin Leahey, and Russell J Funk. Papers and patents are becoming less disruptive over time. *Nature*, 613(7942):138–144, 2023.

Dorian Proksch, Wiebke Stranz, Andreas Pinkwart, and Michael Schefczyk. Risk management in the venture capital industry: Managing risk in portfolio companies. *The Journal of Entrepreneurial Finance (JEF)*, 18(2):1–33, 2016.

Andre Retterath and Reiner Braun. Benchmarking venture capital databases. *Available at SSRN 3706108*, 2020.

John Rust. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, pages 999–1033, 1987.

Yuliy Sannikov, Andrzej Skrzypacz, et al. Dynamic trading: price inertia and front-running. *Preprint*, 2016.

Unjy Song. *Nonparametric estimation of an eBay auction model with an unknown number of bidders*. Citeseer, 2004.

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

Xiaoxue Zhao, Weinan Zhang, and Jun Wang. Risk-hedged venture capital investment recommendation. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 75–82, 2015.