Venture Capital and the Dynamism of Startups*

Yang Yu†

October 6, 2025

Abstract

I study how VCs make investment decisions and how that shape startup lifecycles. VCs make forward-looking investment choices, internalizing competition and anticipating that their information discoveries will benefit future investors — all while learning about startup quality. I develop a structural model that captures these dynamics and, using two decades of software and biotech data, propose a novel empirical strategy to identify startup quality, VCs' speed of learning, dynamic information spillovers, and VC markups. I find that information uncertainty leads VCs to systematically underfund promising startups, causing substantial welfare losses of 21% in software and 22% in biotech, despite software VCs learning considerably faster than biotech VCs. Because VCs cannot internalize the positive information externalities (spillovers) they create, early-round VCs strategically underinvest, resulting in a loss of $13 billion in welfare. I then explore policies to mitigate these losses.

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

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 an important source of inno-

*I thank my advisors Simon Anderson, Gaurab Aryal and Federico Ciliberto for their insightful comments. I also thank Chris Yung, Maxim Engers, Jack Fisher, Doron Ravid, Eric Richert, Peter Troyan, Boning Zhou and participants and discussants at the UVA, EGSC 2024, SEA 2024 and IIOC 2025 for their helpful suggestions. All errors are my own.

†Sloan & CSAIL, Massachusetts Institute of Technology. yyu7@mit.edu

vation (Kolev et al., 2022). For instance, 57% of all new drugs approved in the US in 2023 originated at startups (Economist, 2024), and many of today's leading tech companies (e.g., Microsoft, Google, Amazon, Apple, Meta, Nvidia, OpenAI) began as startups. Startups rely mainly on venture capital (VC) for their capital needs. Although only 0.5% *all* US firms receive VC funding, VC-backed firms make up over half of all IPOs and outperform others in market value, revenue, jobs, and R&D (Lerner and Nanda, 2020). However, there is a concern that startup dynamism falls short of its potential, especially in biotech and software sectors (Bloom et al., 2020; Park et al., 2023). Therefore, to support startup-led innovation, it is imperative first to understand how VC funding works and how it affects the "life-cycle" of startups.

VCs face uncertainty about startup qualities, making it hard to identify winners at the outset.[1] Therefore, to avoid an "all or nothing bet," VC funds are invested across multiple rounds. After each round, some information becomes public, which means the uncertainty that VCs face is dynamic. On the one hand, it is better to invest early because successful startups yield larger returns to earlier VCs than later VCs. On the other hand, it may be better to wait, because unlike early-round VCs, later-stage VCs face less dilution of their stake and, more importantly, have better information. Consequently, early-round investment creates a *dynamic informational spillover*, i.e., a positive externality, leading to potential underinvestment by early VCs who cannot internalize the full social value of their information. So, VCs have to balance these different forces while learning and competing with other VCs with affiliated information about the startup.

To understand how these forces interact and whether they lead to efficient outcomes, I study this startup capital market to understand how different forces (uncertainty, learning, competition, and information spillovers) shape funding decisions and startup outcomes. To this end, I build a rich but tractable model of VC investment,

[1]For example, Bessemer Venture Partners—one of the largest VC firms—passed on Airbnb, Apple, Facebook, Tesla, and Zoom, among others, due to high uncertainties. Furthermore , Kerr et al. (2014) show that among ex-ante equally ranked startups, 60% failed while 10% yielded more than five times return.

study its equilibrium, propose an identification strategy, and estimate the model parameters using newly constructed data on the universe of all US biotech and software startups over the past two decades.

While previous research (Ewens et al., 2022; Sørensen, 2007) has examined individual aspects of VC markets, a comprehensive analysis has been limited by data constraints and the analytical complexity of multi-round financing. To overcome these challenges, I have combined several databases that allow me to follow the lifecycles of startups along with their exit values, and I use ideas from the common value auctions (Milgrom and Weber, 1982) and employer learning (Altonji and Pierret, 2001) to propose an integrated model that enables systematic study of this capital market.

In each round, a startup announces its capital needs. It is then matched with a new set of competing VCs, reflecting the common pattern that VCs typically *lead* only one round per startup. Each VC receives a private signal about startup quality and competes by bidding on the equity stake they demand, with the startup selecting the VC requesting the smallest share. In other words, I model each round as an English auction with common value (Milgrom and Weber, 1982). The equilibrium bids determine round-by-round funding success and equity allocation across VCs, given available information. If the startup fails to raise capital, it terminates and prior investments are sunk; if successful, it proceeds to the next round, where the market infers information from previous outcomes. This process repeats until it is either acquired or goes public, and its exit value is realized and distributed according to each VC's final equity share.

I have collected data on the exit value of all successful startups. With this information plus detailed VC and startup data, I can estimate the model parameters. However, the identification is challenging and requires careful analysis, not least because of right-censoring, correlated information across VCs, and their learning, which lead to dynamic selection. For instance, pure common-value auctions cannot be nonparametrically identified when only winning bids are observed (Laffont and Vuong,

1996). So, I make a distributional assumption for identification. Then, I adapt the dynamic selection correction methods of Heckman and Navarro (2007), originally developed for identifying dynamic schooling decisions, to account for the fact that startup survival and VC participation decisions are jointly determined by information on unobserved quality. This allows me to infer both the startup's true quality and the VCs' interim information.

To estimate the model, I use Pitchbook database on the universe of US biotech and software startups from 2000 to 2022. For each startup, I observe the number of VC rounds and, for each round, the capital invested, the name of the VC, the startup valuation, and the equity ownership. I supplement this database with comprehensive data on exit values, combining information from Compustat and the firms' filings with the US Securities and Exchange Commission.

My final sample includes information on 5,591 startups, of which 2,378 are biotech and the rest are software firms. Besides being two of the most important sectors in the US, biotech and software have distinct risk profiles that generate different learning speeds for VCs. Software startups typically test market demand with minimum viable products, allowing rapid learning about commercial viability (Lerner and Nanda, 2020). In contrast, biotech development involves sequential stages (e.g., clinical trials), spreading uncertainty resolution over multiple years (Food and Administration, 2018). These sectoral differences in learning speed interact differently with VCs' strategic funding decisions, providing new insights into the role of uncertainty and information spillover on market performance, which we would otherwise overlook.

Next, I estimate the model parameters. For that I use the simulated minimum distance estimator (e.g., Agarwal, 2015; Aryal et al., 2024; McFadden, 1989; Pakes and Pollard, 1989). Some of the data moments that I match include summaries of transition patterns of the startups, within- and cross-round variation in investments, winning (equity) bid, and exit values. Moments based on transition and winning bids are informative about VCs' investment choices, their information, and

from it their speed of learning. Similarly, moments constructed from capital investment and exit value help estimate the parameters governing their interdependence.

The estimates suggest that while the uncertainties in software are primarily in the initial round, they extend across multiple rounds in biotech. As a result, even though software investors are initially less informed about the startup's quality, they learn faster than biotech investors. This difference in learning has important consequences: software VCs make more Type I errors (rejecting good startups) but fewer Type II errors (funding bad startups) than biotech VCs, as they act on less complete information initially but later gain greater clarity. Additionally, the early-stage investments generate two to three times returns than later-stage investments, which is comparable to Kaplan and Schoar (2005) and Cochrane (2005).

Using the estimates and the model, I characterize the efficiency level under alternative information and market structures that provide new insights into the welfare consequences of information externalities and uncertainties. I compare outcomes to a first-best benchmark where all startup information (including quality) is perfectly observable to all investors from the start. My main finding is that the current biotech and software sectors respectively yield 78% and 79% of the first-best. 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. These efficiency losses represent foregone value, calculated as the difference between actual and optimal startup funding and success rates.

The estimates suggest that 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.

I also explore a previously unexamined question: how important is M&A as an exit route for startup dynamism? Many startups are founded with the expectation of being acquired. To quantify this effect, I conduct counterfactual simulations in which I vary the probability of successful exit via an M&A. While not directly modeling regulatory changes, this exercise can be interpreted through the lens of stricter merger enforcement—a timely consideration given ongoing policy debates about tightening M&A rules. I find that a 10% decrease in acquisition probability 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.

In December 2023, the DOJ and FTC issued new merger guidelines signaling stricter M&A enforcement, particularly for incumbent (and possibly killer) acquisitions of startups. This shift could reduce exit opportunities, depress valuations, and make VCs more cautious about funding startups whose strategies rely on acquisition by large incumbents. Conversely, tougher M&A rules might push more startups to become standalone companies, potentially benefiting long-term innovation. A full assessment of these competing effects would require analyzing downstream competition dynamics, which I leave for future work.

Related Literature. My paper relates to and contributes to several strands of literature beyond those in empirical industrial organization. First, by focusing on VCs, my paper complements previous research on VC and innovation, see Gornall and Strebulaev (2021), Lerner and Nanda (2020), and Howell et al. (2020). At the same time, most of them focus on descriptive analysis using only startups that eventually went public. As such, among other things, they are silent about viable startups that fail because VCs choose not to fund, which is central to this market.

Second, my paper also complements the literature that considers VC-startup interaction, e.g., Ewens et al. (2022), Sørensen (2007), by going beyond just the first round of VC funding and allowing the market to be uncertain about startup qualities. Focusing only on the first round ignores the dynamic information externality in

this market. After all, Ewens et al. (2018), Kerr et al. (2014), and Lerner and Nanda (2020) highlight multi-round financing and quality uncertainty as two of the most salient features of VC financing.

Third, my paper is also related to the literature on auctions with common value, particularly the papers on security design auctions of DeMarzo et al. (2005), Garmaise (2001), and Gorbenko and Malenko (2011), where bidders compete for an asset (startup) by bidding with securities whose payments are contingent on the asset's value to be realized in the future. I combine ideas from these auction models with employer learning models in labor markets (Altonji and Pierret, 2001; Aryal et al., 2022; Lange, 2007) to capture the incentives and competition VCs face at different rounds.[2] Thus, my paper also contributes to a nascent but growing literature on the industrial organization of finance (Clark et al., 2021).

Lastly, my paper builds upon the identification results of discrete choice models by Heckman and Navarro (2007). I show that their identification strategy can be extended to dynamic VC financing of startups. Because VCs have correlated information, the classic identification arguments in Rust (1987) do not apply in my setting. My application of Heckman and Navarro (2007) provides a template for researchers to go beyond the conventional assumption of "i.i.d." information.

Overall, my paper contributes 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*.[3] This database tracks each startup's VC rounds, providing detailed data on capital investment, valuation and eq-

[2] In contrast, Sørensen (2007) uses cooperative solution concepts to study the market.

[3] 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).

uity ownership at round level. It also provides information on startup sector, patents and exits, including IPOs and M&A ("success" hereafter), and bankruptcies and business closures ("failure" hereafter). Pitchbook collects company and financing rounds information from EDGAR online filings like Reg D filings, press release, and direct outreach to companies and investors. Overall, Pitchbook has one of the best coverage and is one of the most accurate VC databases (Retterath and Braun, 2020).

I complement Pitchbook with several other data. I supplement missing investment and valuation data with SEC filings (Forms 8-K, 10-K, and 10-Q) for startups that later went public, and fill in missing acquisition prices using the corresponding filings of publicly traded acquirers, who are required to disclose 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 biotech and software companies that were founded between 2000 and 2022, headquartered in the US, and have raised VC or pre-VC rounds. Biotech firms focus on drug discovery and delivery of pharmaceuticals or biotechnology (e.g., Amgen, Moderna, Eli Lilly). Software companies design and develop softwares, such as application softwares (e.g., Adobe, Oracle), business/productivity softwares (e.g., IBM, Salesforce), operating systems (e.g., Apple, Microsoft), social/platform softwares (e.g., LinkedIn, Meta).

I focus on these two sectors not only because they are among the largest and most innovative sectors in the US, but also because (as I will show later) they have distinct business models and risk profiles — this contrast allows me to explore how the learning process and welfare implications differ in sectors with fundamentally different risk profiles. 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. Out of them, 2,072 were successful (589 IPO and 1,483 M&A) and 3,519 failed. I exclude startups that raised more than 7 rounds, as their number is too small to produce reliable estimates for the parameters associated with rounds beyond 7.

2.1 Data Description

I use *Siri* as an example to introduce the data details (see Supplemental Appendices for additional examples from the data). 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. It was founded in 2007 and headquartered in San Jose, and its primary industry is *Application Software*. Table 1 presents its financing rounds and exit.

Table 1: **VC Financing and Exit of Siri**

Round	Date	Investment	Valuation	Initial Ownership	Diluted Ownership	VC
1	2008-06-19	$8.55M	$18.53M	46%	35%	Menlo
2	2009-11-24	$15.5M	$65.91M	24%	24%	SRI
M&A	2010-04-28	$200M	$200M	100%	-	-

On Jun 19, 2008, Siri raised $8.55 million from Menlo Ventures at a **valuation** of $18.53 million, which represents the "price" of the entire company in this round. In return, Menlo Ventures obtained $\frac{8.55}{18.53} = 46\%$ of Siri's **equity ownership**, representing a claim to a portion of the company's future returns.

On November 24, 2009, it raised $15.5 million in the second round from SRI Ventures at a valuation of $65.91 million. The company issued new shares such that SRI Ventures owned 24% ($=\frac{15.5}{65.91}$) equity stake. At this time, round 1 investor's equity ownership was *diluted* to 35% ($=46\% \times (1 - 0.24)$) because more shares were issued. However, his estimated return still increased because the rise in Siri's valuation offset the dilution in the ownership.

Eventually, on April 28, 2010, Siri was acquired for $200 million. Menlo Ventures and SRI Ventures each got 35%×$200 million = $70 million and 24%×$200 million = $48 million, respectively.

Section A and B in section A describe the variables observed at startup and startup-round level, respectively. *ExitValue*, *Investment* and *Valuation* are converted to millions of USD as of December 2023. One data limitation is that PitchBook does not report exit values for failed companies, that my estimation strategy has to address. See Section B for summary statistics at startup and startup-round level. In general, as a startup proceeds to later rounds, *Investment* and *Valuation* tend to increase, while *EquityOwnership* tends to decrease. Additionally, *Patent*, *Valuation* 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

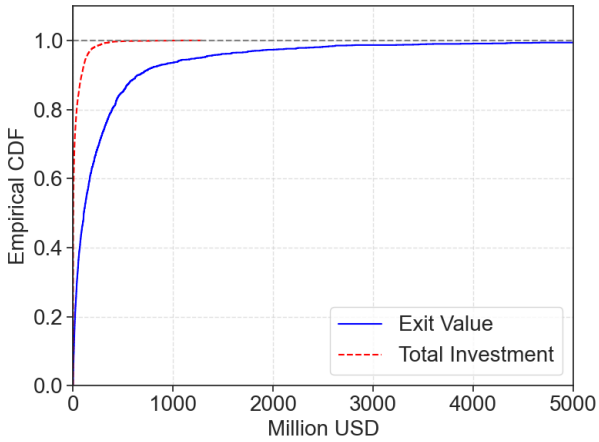
In the VC-startup context, exit value is a key variable of interest, as it affects VC returns. By the nature of innovation, neither entrepreneurs nor investors can identify the biggest success ex-ante (e.g., Ewens et al., 2018; Kerr et al., 2014). Hence, the distribution of exit value captures how risky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

Figure 1 presents the distribution of exit values in my sample. 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.

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

Figure 1: **Empirical CDF of exit value and total investment**



Note: Only IPO and M&A startups are included for exit values, while all startups are included for total investment. Total investment is defined at the startup level as the aggregate amount raised across all rounds. All values are in Dec 2023 USD.

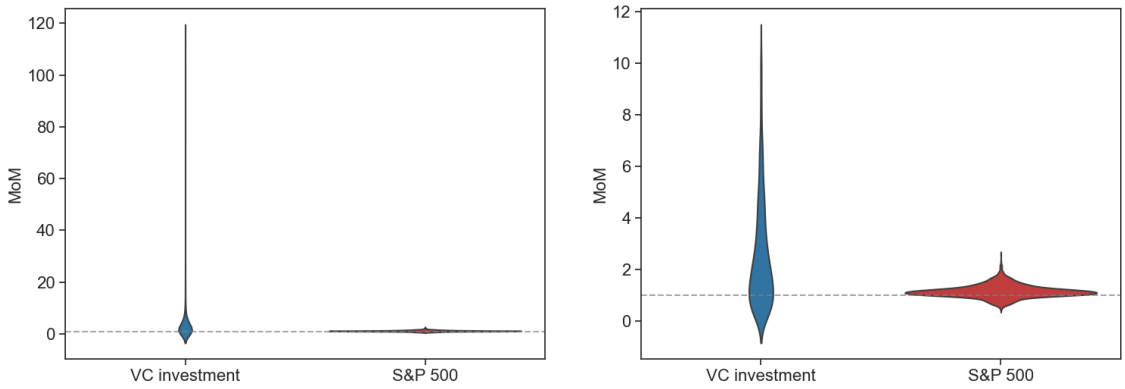
benchmark, such as the S&P 500 index.

I use the above-mentioned *Siri* example for illustration. Round 1 and 2 investors invested $8.55 million and $15.5 million, and received $70 million and $48 million in return, respectively. Thus, the multiple of money (MoM) — a common metric for measuring VC investment performance — is calculated as $\frac{70+48}{8.55+15.5} = 4.91$. Had round 1 investor invested the same amount in S&P 500 index on the day of round 1 and sold the holdings at the time of acquisition, and similar for round 2 investor, the MoM would have been 1.03.

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

In response to this risk, VCs phase their investments over multiple rounds, with each round generating some (new) information about the startup's exit value (Ewens et al., 2018). Figure 3 summarizes the multi-round financing and exit of startups in my sample. Round 0 is when the startup was established. The numbers are counts of startups reaching each state. For example, in total 5,591 startups were established,

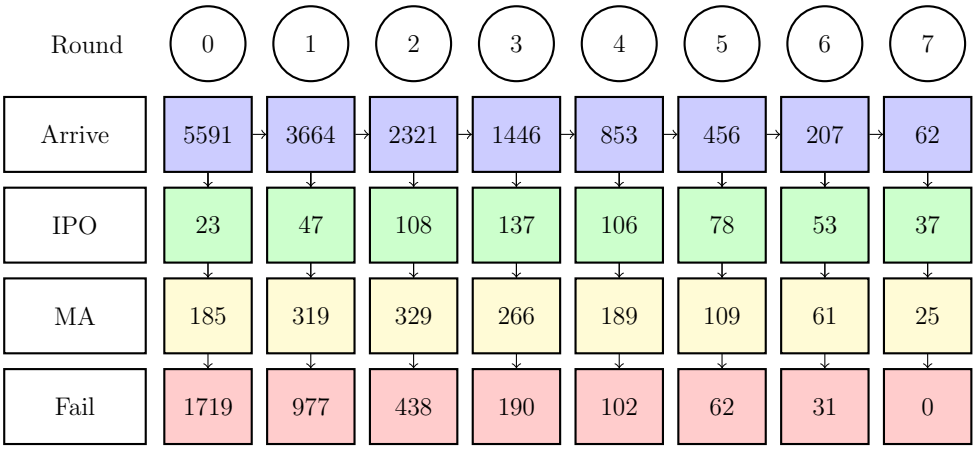
Figure 2: **MoM of VC investments and public market counterfactuals**



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. Investments and returns are in Dec 2023 US dollar.

out of which 3,664 raised 1st VC round, 23 and 185 went IPO and got acquired directly, and the rest failed without raising any VC rounds. Based on it I derive the empirical transition probability matrix in Table 2.

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. Round 0 = Established. The number in each box is the count of startups reaching that state. For example, Arrive & Round0 = 5591 means that I start with 5,591 startups, out of which 3,664 move to Round 1 etc.

Table 2: **Empirical transition probability matrix**

	Next state			
Current state	Next round	IPO	M&A	Fail
Round 0	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. Round 0 = Established.

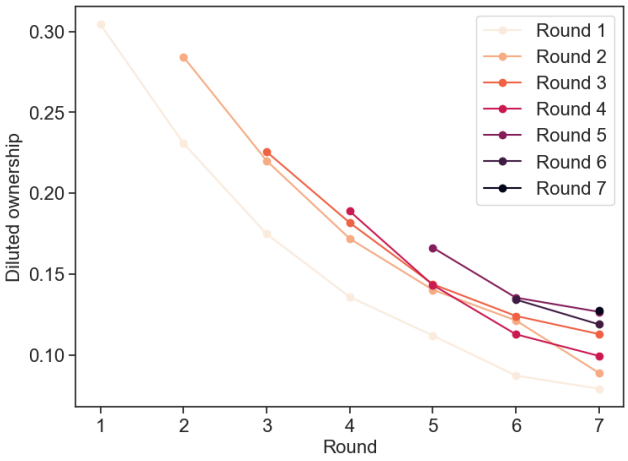
Each row in Table 2 shows the probability of transitioning from the current state to each of the possible next state. As startups progress to later rounds, success rates increase and failure rates decrease, because low-quality startups exit, leaving a stronger pool of startups.

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

2.2.3 Dilution in Equity Ownership

As illustrated in the *Siri* case, one feature of VC financing is that equity ownership of early-round VCs dilute as a startup raises additional capital and issues more shares in subsequent rounds. Equity dilution is important because it determines how the returns from startups are shared across VCs. As shown in Fig. 4, equity ownership dilution is significant, especially for early-stage investors.

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%.

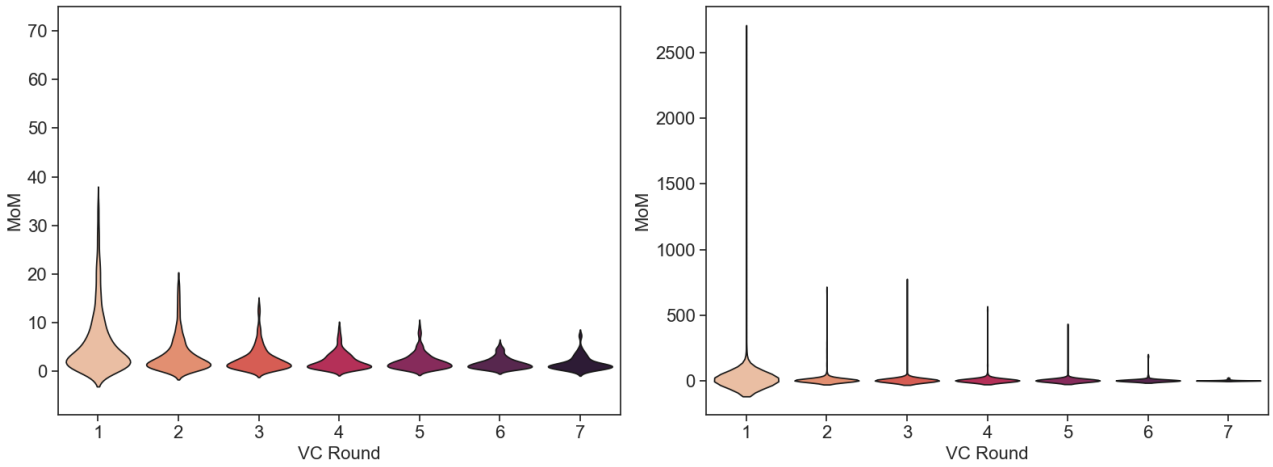
2.2.4 Round-specific Return

Despite significant dilution, early investors still earn more if the startup is successful. Fig. 5 presents the MoM returns to each round investors, conditional on the startup going public. On average, early investors achieve higher MoMs than later investors. This observation aligns with the underlying dynamic 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 leading late-stage investors to pay a higher price.

2.2.5 "Price" converging to true value

Venture capitalists learn about a startup's exit value from the information that becomes public after 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., valuation) is an outcome of both VCs' learning about the exit value and imperfect VC competition, the two effects cannot be disentangled without a model.

Figure 5: **MoM of successful startups**



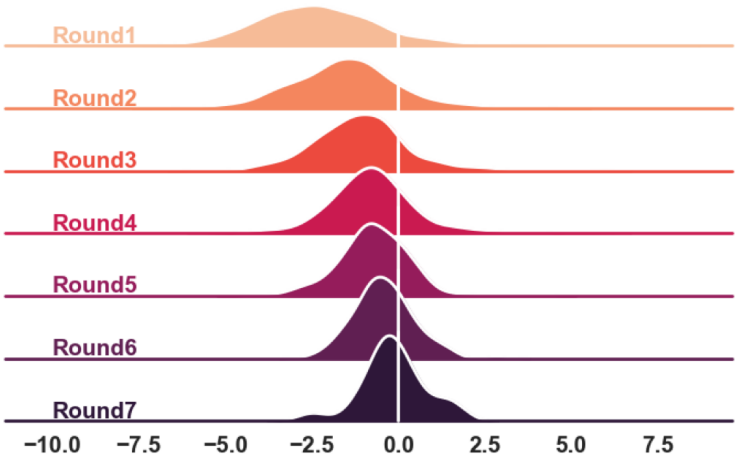
Note: IPO and M&A companies are included. The left (right) figure excludes (includes) outliers (below 5 percentiles or above 95 percentiles).

I use $\log\left(\frac{\text{Valuation}_k}{\text{ExitValue}}\right)$ to quantify the gap between the "price" paid by round k investor and the startup's exit value. Failed startups are excluded from this exercise. Thus, the sample represents stronger companies. Initially, the price as a percentage of the exit value is dispersed, suggesting substantial uncertainty in startup quality at the outset. However, as the startup grows and releases more information, the price gradually converges toward its exit value, suggesting a market-wide dynamic learning. Additionally, the log of this ratio is approximately normal for all rounds, justifying the specification choice in the model section.

2.2.6 Lead VC

In each funding round, although multiple VCs may co-invest, there is typically a lead VC who recruits other investors to fill out the round, leads the due diligence process, negotiates with the startup, and invests the largest amount. Moreover, each round tends to be led by a different VC. This is because VCs often specialize in different stages of a startup's lifecycle, reflecting their heterogeneity in expertise and fund size. 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

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



Note: The horizontal axis is $\log\left(\frac{\text{Valuation}_k}{\text{ExitValue}}\right)$. For example, a startup raised two rounds with valuations 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 exit value. Only IPO and M&A companies are included.

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. So, in the model, I assume that in each round, there will be only one VC that "wins" — the one that leads the deal — and that this leading VC is new.

Table 3: **Lead VC**

Number of rounds	1	2	3	4	5	6
Percentage of VCs	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. The numbers are in percentages.

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 and compare the characteristics of each group.

Table 4: **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	-
Total	*IPO*	-	29.84	74.15	93.77	103.72	151.36	224.95	208.33
Investment	*MA*	-	12.25	26.29	47.5	82.04	94.3	128.53	141.54
	Fail	-	3.41	12.35	27.5	45.41	53.3	86.58	-
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 investment raised from VC and exit value of startups in each group. Startups are grouped by their exit type and number of rounds, e.g., 93.77 is the mean total investment in startups that went IPO after raising 3 VC rounds. Total investment and exit value are in million USD (as of Dec 2023). Round 0 = Established.

Table 4 shows that startups following different trajectories are heterogeneous in patents, total capital raised, and exit value. In addition, I find that exit value is positively correlated with both patent and investment (Table 5), suggesting that these factors may influence exit outcomes and motivating their inclusion in the later exit value specification.

Table 5: **Correlation**

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

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

3 Model

I consider a multi-round game between a startup and several VCs. As described in Section 1, a startup of unknown (to both entrepreneur and investors) quality needs to secure VC funding to be successful. In each round, the startup announces its capital needs, it is then matched with a (new) set of VCs. They observe the characteristics revealed up to this point, and each obtains a private signal on the unobserved quality.

VCs then compete in equity stakes based on their information about the startup and their expectations about future VC participation. If the startup fails to raise capital, it exits the market, and the capital already invested is lost. If successful, it proceeds to the next round, and all VCs' signals in the current round become public.[4] This process repeats until it is either acquired or goes public, and its exit value is realized and distributed according to each VC's final equity share. Fig. 7 presents the timeline of the game, with some notations to be introduced below.

The competition among VCs for a startup is likely a multilateral bargaining process. I model it "as if" it were an English auction, which generates patterns similar to multilateral bargaining while preserving tractability (McAfee and Vincent, 1997). The implicit assumption behind it is that the startup holds all bargaining power vis-à-vis the VCs — a reasonable characterization in VC markets where capital is abundant but entrepreneurial talent is scarce, enabling startups to capture the entire surplus under intense VC competition.

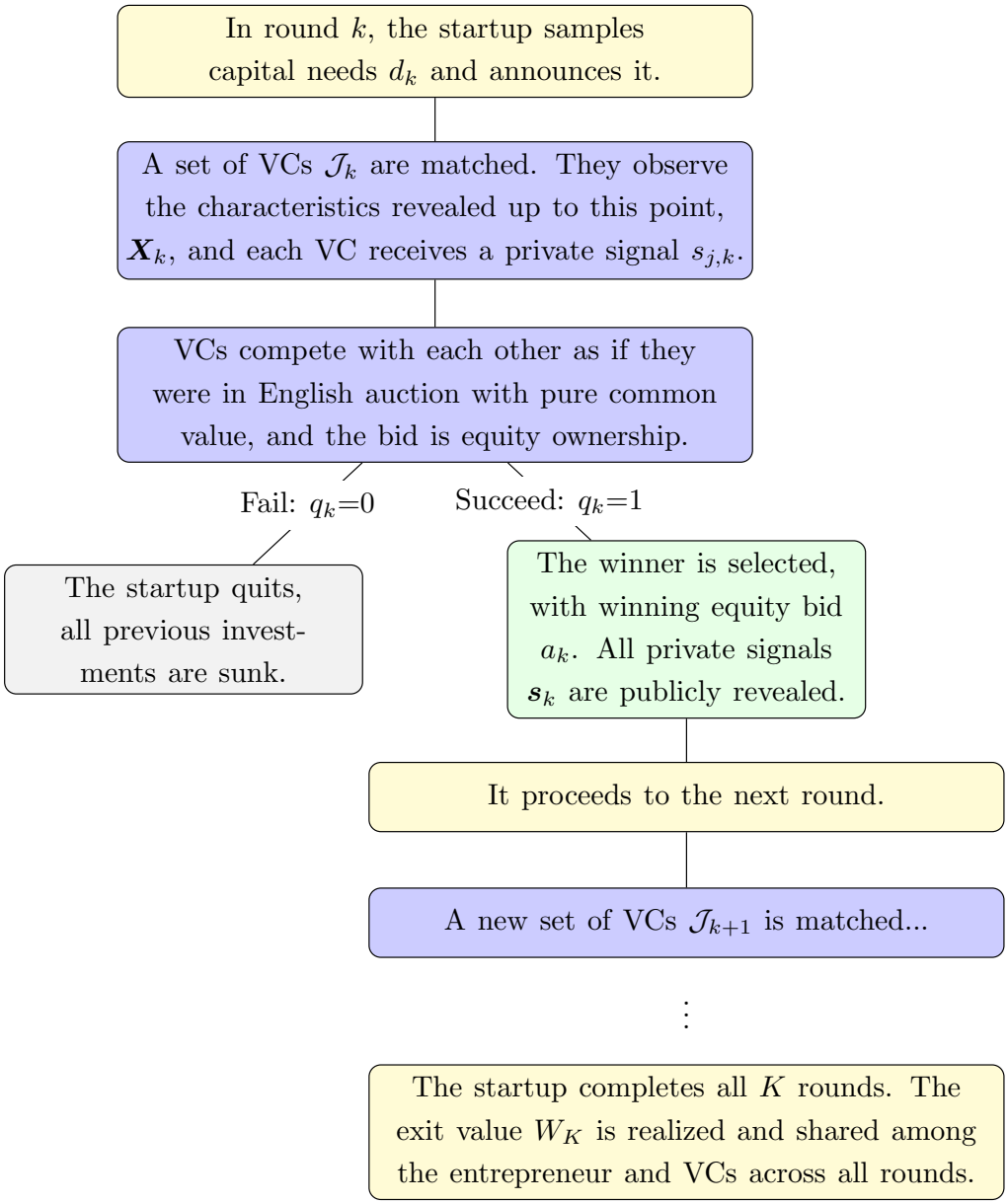
VCs bid on the required equity ownership for them to invest.[5] VCs seek larger stakes for higher returns, while startups aim to give away as little as possible, so the VC demanding the least equity wins. To characterize the equilibrium bids, I follow Milgrom and Weber (1982) and model the auction as follows. At the start, all VCs are active at an arbitrarily high equity ownership. As the equity ownership decreases, VCs drop out one by one. In a symmetric monotone equilibrium, the equity ownership at which each VC drops out reveals her true signal, thus the VCs who are still active will update their information and bids accordingly. Any VC that drops out cannot be re-enter. The auction ends when only one bidder is left. She wins and is paid the equity ownership at which the second lowest bidder quits.

To formalize this pricing, I first introduce some notations. Let K be the number

[4]Late-round investors can assess the information used by early-round investors in their decisions, since due diligence gives them access to financials, product details, contracts, and management backgrounds reviewed by previous investors (Ollar et al., 2021; Sannikov et al., 2016).

[5]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.

Figure 7: **Timeline**



Note: Yellow and blue boxes are startup and VC actions, respectively.

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. Neither the startup nor the VCs observe the realization of K ex-ante.[6] Let Z be a $1 \times z$ vector of observed characteristics of the startup, such as its sector (biotech or software), the number of assigned patents etc. $d_k \in \mathbb{R}_+$ is the startup's capital needs in round $k \in \{1, ..., \bar{K}\}$. It is drawn from a commonly known distribution, but its realization is observed only at the start of the round. 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 $F(\cdot)$. Let ξ be the unobserved quality of a startup, capturing the features of a startup that neither VCs nor entrepreneur know ex-ante. Information on ξ only arrives with investment into the project. Examples of unobserved quality include the validity of a technology or product and the potential market demand.[7] Additionally, $\delta \in (0, 1)$ is the discount factor between two adjacent rounds.

Exit Value. I model a startup's production function as a mapping from inputs—such as invested capital, sector, and startup quality—into output, namely the exit value. The exit value after raising K rounds, W_K, is

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

where $m_K(\boldsymbol{X}_K)$ represents a composite input in the startup's production process, determined by the *observed* characteristics $\boldsymbol{X}_K$, and $\exp(\xi)$ is the *unobserved* quality of a startup that must be learned across rounds. Because exit values are strongly skewed (Fig. 1), I model ξ in exponential form, allowing quality to exhibit a long

[6] 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.

[7] For example, when Airbnb launched in 2009, Fred Wilson of Union Square Ventures described it as a "marketplace for air mattresses on the floors of other people's apartments". At that stage, demand was highly uncertain and could only be revealed by investing to build the platform and observing adoption. See more details at https://avc.com/2011/03/airbnb/.

right tail.

This specification falls into the class of single-index models where the exit value only depends on $\boldsymbol{X}_K$ through the function $m_K : \mathbb{R}_+^{z+K} \to \mathbb{R}_+$ (Ichimura, 1993). This specification allows m_K to be nonparametric—so my results are not tied to a specific functional form—while limiting to one nonparametric dimension to avoid the curse of dimensionality. Multiplicative separability (equivalently, additive separability after a logarithmic transformation) between the observed and unobserved components is essential for identification. This condition has played a prominent role in econometric analysis, such as mean independence in linear models and separability of common and private values in affiliated-value auctions.

I make the following assumption on ξ.

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

Recall that one objective of this paper is to estimate the speed of learning. To maintain tractability, I follow the literature in assuming Gaussian prior beliefs. Combined with Gaussian signals, this ensures that posterior beliefs are also Gaussian, yielding a standard normal-normal learning structure. The choice of normality is also consistent with the "normal" beliefs observed in Fig. 6. In addition, I allow σ_ξ to be sector-specific in the estimation.

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

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

where $\epsilon_{j,k}$ is the signal noise subject to the assumption below.

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 guarantees that VCs' signals are unbiased, the normality assumption is for tractability, and the round-specific variance allows signal accuracy to vary across rounds. For notation convenience, let $\boldsymbol{\sigma} = (\sigma_k)_{k=1}^{\bar{K}}$ denote the vector of signal-noise standard deviations across rounds, $\boldsymbol{s}_k = (s_{j,k})_{j\in\mathcal{J}_k}$ the set of signals generated in round k, and $\boldsymbol{S}_k = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_k)$ the set of all signals up to and including round k.

Ex-post payoff. As is shown in Fig. 4, when a startup 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,

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

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 $\phi_k \coloneqq (\mathfrak{q}_k, \mathfrak{a}_k)$ 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 that round. For brevity, let $\boldsymbol{\phi}_{k+1} \coloneqq (\phi_{k+1}, ..., \phi_{\bar{K}})$, $\boldsymbol{\mathfrak{q}}_{k+1} \coloneqq (\mathfrak{q}_{k+1}, ..., \mathfrak{q}_{\bar{K}})$ and $\boldsymbol{\mathfrak{a}}_{k+1} \coloneqq (\mathfrak{a}_{k+1}, ..., \mathfrak{a}_{\bar{K}})$, which will come in handy later.

Ex-ante payoff. Since the startup's exit value is unknown at the time of bidding, VCs base their decisions on expected payoffs. Given the observed characteristics $\boldsymbol{X}_k$, the signals $\boldsymbol{S}_k$, and the future funding rules $\boldsymbol{\phi}_{k+1}$, the *maximum* expected payoff for

round k investor is

$$V_k(\boldsymbol{X}_k, \boldsymbol{S}_k; \boldsymbol{\phi}_{k+1}) = \mathbb{E}\left[\delta^{K-k} \prod_{k'=k+1}^{K} (1 - a_{k'}) q_{k'} W_K \middle| \boldsymbol{X}_k, \boldsymbol{S}_k; \boldsymbol{\phi}_{k+1}\right], \tag{2}$$

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 taken over all possible future trajectories of the startup, conditional on the current information set and future funding rules.

The dimensionality of $\boldsymbol{S}_k$ grows rapidly with k. To reduce it, note that under the independence and normality assumptions in Assumption 1 and 2, $\boldsymbol{S}_k$ matters only for inferring ξ, whose posterior distribution is also normal,

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

where μ_k is a weighted average of $\boldsymbol{S}_k$, and τ_k does not depend on $\boldsymbol{S}_k$ (see Section C for the explicit expressions of μ_k and τ_k). Since $\boldsymbol{S}_k$ impacts the posterior distribution only through μ_k, the latter becomes a sufficient statistics for $\boldsymbol{S}_k$. Hence, Eq. (2) is simplified to

$$V_k(\boldsymbol{X}_k, \mu_k; \boldsymbol{\phi}_{k+1}) = \mathbb{E}\left[\delta^{K-k} \prod_{k'=k+1}^{K} (1 - a_{k'}) q_{k'} W_K \middle| \boldsymbol{X}_k, \mu_k; \boldsymbol{\phi}_{k+1}\right]. \tag{3}$$

Bidding strategy. At round k, VC j's strategy specifies the equity level at which she should drop out, given the current round k, the observed characteristics $\boldsymbol{X}_k$, past 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 exited and the equity levels at which they did so.

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 it from revealed signals in earlier rounds and her own signal. However, as bidders start to drop out,

each equity level at which a competing VC exits discloses her true signal (because strategies are revealing in equilibrium), based on which the remaining VCs update their guess of μ_k and adjust their own dropout thresholds accordingly, assuming no other VCs exit in the middle. See Milgrom and Weber (1982) for formal presentation of the bidding strategy.

Equilibrium. Proposition 1 and 2 characterize the equilibrium binary funding rule $\mathfrak{q}_k^*$ and 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 and only if $V_k^*(\boldsymbol{X}_k, \mu_k) \geq d_k$. In other words, the equilibrium funding rule $\mathfrak{q}_k^*$ is given by*

$$\mathfrak{q}_k^*(\boldsymbol{X}_k, \mu_k) = \mathbb{1}\{V_k^*(\boldsymbol{X}_k, \mu_k) - d_k \geq 0\}, \tag{4}$$

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. (3) and $\boldsymbol{\phi}_{k+1}^*$ is the equilibrium funding decisions in all future rounds.

Proposition 1 provides a 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. (4) determines whether and when a startup fails.

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

$$\mathfrak{a}_k^*(\boldsymbol{X}_k, \hat{\mu}_k) \times V_k^*(\boldsymbol{X}_k, \hat{\mu}_k) = d_k, \tag{5}$$

$\hat{\mu}_k$ is the posterior mean of ξ after substituting the highest signal with the second-highest. Moreover, $V_k^*(\boldsymbol{X}_k, \mu_k)$ is strictly increasing in μ_k for any $\boldsymbol{X}_k$ and $k \in \{1, ..., \bar{K}\}$.

For example, in round 1, if the actual signals are $\boldsymbol{s}_1 = \{1, 2\}$, then $\hat{\mu}_1$ is computed from the modified set $\hat{\boldsymbol{s}}_1 = \{1, 1\}$, where the highest signal (2) is replaced by the

second-highest (1). To understand Proposition 2, when only two bidders are left, all signals (except for the highest two) have been revealed. Each of the two remaining bidders will drop out when their expected net payoff, conditional on winning, goes to zero. To calculate this threshold, a bidder asks: what would be the other bidder's signal if I were to win at a bid that makes me just break even? In a symmetric equilibrium, this occurs only if the other bidder exits at the same point — meaning the other bidder's signal must equal her own. The formal proof of Proposition 2 is deferred in Section D.

3.2 Efficiency Analysis

Having established the equilibrium, I examine its (in)efficiency—possibly due to information spillovers—by comparing it with the investment strategy of a benevolent planner who has access to the same information as VCs but internalizes the spillovers. Fixing information to match that of VCs but allowing the planner to internalize spillovers, I compare their investment choices to isolate the efficiency loss due to spillovers, which I quantify in Section 5.

Specifically, I consider a benchmark in which a planner with 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 Proposition 1 and 2. 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 invests, the expected net return (before investing capital d_k) is

$$\tilde{V}_k(\boldsymbol{X}_k, \mu_k; \mathbf{q}_{k+1}) = \mathbb{E}\left[\delta^{K-k}\prod_{k'=k+1}^{K} q_{k'} W_K - \sum_{k'=k+1}^{K} \delta^{k'-k}\left(\prod_{n=k+1}^{k'} q_n\right) d_{k'} \middle| \boldsymbol{X}_k, \mu_k; \mathbf{q}_{k+1}\right]. \tag{6}$$

Eq. (6) is different from Eq. (3) 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 $\mathbf{a}_{k+1}$ 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*

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

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. (6) and $\mathfrak{q}_{k+1}^\dagger$ is the planner's optimal funding rules in all future rounds.

Proposition 1' provides a necessary and sufficient condition for a startup to secure funding from the planner at round k. 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 section E. 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.

Early investors are not fully compensated for the spillovers they generate because, while their investment advances the startup and enables later investors to draw more informative signals, they cannot observe those signals and therefore cannot contract upon them. Hence, there doesn't exist a mechanism that achieves Pareto efficiency (i.e., at any stage, the project is funded if and only if the total expected net payoff conditional on the current information set is non-negative) while satisfying incentive compatibility and individual rationality (Myerson and Satterthwaite, 1983). Consequently, they may stop investing even when the overall benefit is still positive, resulting in suboptimal equilibrium 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

I provide the intuition for leveraging variations across startup and over time to identify the model parameters, while deferring the formal proof of identification to Section F. This discussion also guides my choice of moments used in estimation.

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), representing the effect of observed characteristics on the exit value; σ_ξ is the standard deviation of unobserved quality ξ; $\boldsymbol{\sigma}$ is the vector of signal noise standard deviations across rounds; and F is the joint distribution of the last round K and observed characteristics $\boldsymbol{X}_K$.

In the data, I observe $\{\boldsymbol{q}_K, \boldsymbol{X}_K, \boldsymbol{a}_K, W_K\}$ for all startups where $\boldsymbol{q}_K = (q_k)_{k=1}^K$ is a vector of binary funding choices, $\boldsymbol{a}_K$ is a vector of winning equity ownerships, $\boldsymbol{X}_K = (Z, d_1, ..., d_K)$ is the startup characteristics, and W_K is the exit value. However, 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 endogenous 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 it, I adapt the dynamic selection correction methods of Heckman and Navarro

(2007), which allow for correlated unobservables over time (correlated signals across rounds in the VC-startup context), to account for the correlation in VC participation decisions across rounds even after conditional on observables. Because exit values are observed only for successful startups, this method corrects for investor-side selection, ensuring that the choice-adjusted estimates of model primitives are free from the influence of unobserved quality that affect both VC participation and exit values.

Identifying the full structural model requires independent variation in observed and unobserved characteristics (achieved by independence assumptions on observed characteristics $\boldsymbol{X}$, unobserved quality ξ, and signal noise ϵ), as well as functional restrictions of the model. More specifically, I have four model-implied conditions that, along with the data, can be used to identify the parameters:

$$\ln W_K = \ln\left(m_K(\boldsymbol{X}_K)\right) + \xi, \tag{7}$$

$$q_k = \mathbb{1}\{V_k^*(\boldsymbol{X}_k, \mu_k) \geq d_k\}, \tag{8}$$

$$a_k = \frac{d_k}{V_k^*(\boldsymbol{X}_k, \hat{\mu}_k)}, \tag{9}$$

$$(\xi, \mu_1, ..., \mu_{\bar{K}})^\top \sim N(\mathbf{0}, \Sigma). \tag{10}$$

Eq. (7) is Eq. (1) after taking the log, Eq. (8) and Eq. (9) are the equilibrium conditions shown in Eq. (4) and Eq. (5). Eq. (10) 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 section C for its mathematical expression.

To understand the sources of variation for identification, consider two identical startups, both raised one funding round and then exited via IPO. Everything — exit value, capital investment, and other characteristics — is the same except for the winning equity ownership. This variation must stem from VCs receiving different

signals about the same underlying quality. This is reflected in the model-implied equilibrium bidding strategy in Eq. (9) that maps the observed winning bid a_k to $\hat{\mu}_k$, an order statistics of the underlying signals. Hence, the variation in a_k allows me to trace out the distribution of $\hat{\mu}_k$.

In the data, however, winning equity ownership is subject to selection, as it is only observed for startups that successfully secure funding. I address this by using the equilibrium funding choices in Eq. (8) to correct for selection, and then recover $\boldsymbol{\sigma}$ from the joint distribution specification in Eq. (10), which links $\boldsymbol{\sigma}$ to μ_k. Similarly, I identify $\boldsymbol{m}, \sigma_\xi, F$ using data and the exit value specification in Eq. (7) while correcting the selection by Eq. (8).

One thing to note is that the number of VCs competing for a startup in each round is not observed in the data; only the winner's bid, if any, is observed. I assume there are only two VCs in each round.[8] This assumption is only a normalization of the noise standard deviations $\boldsymbol{\sigma}$ when the number of VCs in a round is fixed (not necessarily equal to two), and thus does not bias the estimates of learning speed (e.g., observing four signals from a normal distribution with standard deviation 2 yields the same learning speed as observing two signals from the same distribution but with standard deviation $\sqrt{2}$).

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 nonlinear problems with

[8]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.

large search spaces while limiting computing time by parallelization. Hereafter, I use one sector to illustrate the estimation strategy, but in practice, I estimate the two sectors separately.

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.

$$\begin{aligned} m_0(\boldsymbol{X}_0) &= Z\beta_0, \\ m_1(\boldsymbol{X}_1) &= Z\beta_0 + d_1\beta_1, \\ &\dots \\ m_{\bar{K}}(\boldsymbol{X}_{\bar{K})} &= Z\beta_0 + d_1\beta_1 + \cdots + d_{\bar{K}}\beta_{\bar{K}}, \end{aligned} \tag{11}$$

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

I assume $\boldsymbol{X}_K$ and K are independent, as allowing for such dependence would drastically enhance 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 exit after each of the seven rounds if they received funding throughout.

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

Assumption 4.

$$d_{k+1} \sim Exp\left(\frac{1}{\lambda_0^k + \lambda_1^k d_k + \lambda_2^k \bar{W}_k}\right), \tag{12}$$

where $\bar{W}_k = m_k(\boldsymbol{X}_k)$.

This distributional specification is chosen for three reasons. First, the exponential distribution is skewed and thus reflects the skewness of investment shown in Fig. 1. Second, it implies that the expected investment in the next round, d_{k+1}, depends linearly on the current round's investment, d_k, and the cumulative value of observed characteristics to date, $\bar{W}_k$, thereby capturing the (positive) correlation among round-to-round characteristics. Third, the Markov specification, commonly used in dynamic settings, frees me from tracking the history of all past investments, whose dimensionality grows polynomially with rounds and quickly becomes computationally intractable. 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.,

$$\hat{\Gamma} = \arg\min_{\Gamma} \left(\hat{\boldsymbol{g}} - \hat{\boldsymbol{g}}^M(\Gamma)\right)' \Lambda \left(\hat{\boldsymbol{g}} - \hat{\boldsymbol{g}}^M(\Gamma)\right), \tag{13}$$

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 three sets of empirical moments. $\hat{\boldsymbol{g}}^M$ contains the

same moments but is derived from and averaged over M simulations of the startup-VC market. These moments include (1) the fractions of startups that succeed and fail after raising $k \in \{0, ..., 6\}$ rounds; (2) the mean and standard error of observed capital investment, valuation, and exit value at round $k \in \{1, ..., 7\}$; (3) the covariance between capital investment and valuation at round $k \in \{1, ..., 7\}$. The first set corresponds to Fig. 3, the second set is reflected in Section B.1.

5.3 Estimation Algorithm

A startup is fully characterized by the vector $(K, \boldsymbol{X}_K, \xi, \boldsymbol{S}_K)$. 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 Eq. (12), the future capital investments, 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. (4) and (5), $\mathfrak{q}_k^*(\boldsymbol{X}_k, \mu_k)$ 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 Bellman form, which allows me to use backward induction to back out its value at any state for any parameter Γ.

$$\begin{aligned} V_k^*(\Omega_k) &= \Pr(K = k \big| K \geq k) \mathbb{E}\left[W_K \big| \Omega_k, K = k\right] \\ &+ \Pr(K > k \big| K \geq k) \delta \mathbb{E}\left[(1 - \mathfrak{a}_{k+1}^*(\Omega_{k+1})) \mathfrak{q}_{k+1}^*(\Omega_{k+1}) V_{k+1}^*(\Omega_{k+1}) \big| \Omega_k, K > k\right]. \end{aligned} \quad (14)$$

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, $\mathbf{\Omega}_k$, encompasses every permutation of the possible values for $\bar{W}_k, d_k, \mu_k, \hat{\mu}_k$.

Given Γ, I compute V_k^* at each grid by starting 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 $\Omega_{\bar{K}} \in \mathbf{\Omega}_{\bar{K}}$. Since $\bar{K}$ is the final round, by the property of lognormal distribution, Eq. (14) is simplified to Eq. (15), and the funding and bidding policies are Eq. (16) and (17).

$$V_{\bar{K}}^*(\Omega_{\bar{K}}) = \bar{W}_{\bar{K}} \exp\left(\mu_{\bar{K}} + \frac{1}{2}\tau_{\bar{K}}^2\right), \tag{15}$$

$$\mathfrak{q}_{\bar{K}}^*(\Omega_{\bar{K}}) = \mathbb{1}\{V_{\bar{K}}^*(\Omega_{\bar{K}}) \geq d_{\bar{K}}\}, \tag{16}$$

$$\mathfrak{a}_{\bar{K}}^*(\Omega_{\bar{K}}) = \frac{d_{\bar{K}}}{V_{\bar{K}}^*(\hat{\Omega}_{\bar{K}})}, \tag{17}$$

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 $\Omega_k \in \mathbf{\Omega}_k$, simulate $(\Omega_{k+1}^l)_{l=1}^L$ transitioned from Ω_k, so that Eq. (14) becomes

$$V_k^*(\Omega_k) = \Pr(K = k \big| K \geq k)\bar{W}_k \exp\left(\mu_k + \frac{1}{2}\tau_k^2\right)$$
$$+ \Pr(K > k \big| K \geq k)\delta \frac{1}{L} \sum_l \left[(1 - \mathfrak{a}_{k+1}^*(\Omega_{k+1}^l))\mathfrak{q}_{k+1}^*(\Omega_{k+1}^l)V_{k+1}^*(\Omega_{k+1}^l)\right]. \quad (18)$$

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. (13) given parameter Γ.

Algorithm 1: Algorithm for computing the objective function given Γ

```
// Step 1:for each grid and round, compute expected payoff,
   funding rule and ownership rule.
for k ← K̄ to 1 do
    foreach Ω_k in Ω_k do
        compute V_k^*(Ω_k), q_k^*(Ω_k), a_k^*(Ω_k) using Eq. (15)(16)(17)(18);
    end
end
// Step 2:simulate data in market m
for m ← 1 to M do
    generate simulated data D_m;
end
// Step 3:simulate each startup's trajectory in each simulation
for m ← 1 to M do
    foreach startup in D_m do
        compute its state in each round, then generate its trajectory and each
          round's winning equity bid using results from Step 1;
    end
end
// Step 4:compute 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. (7) 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.

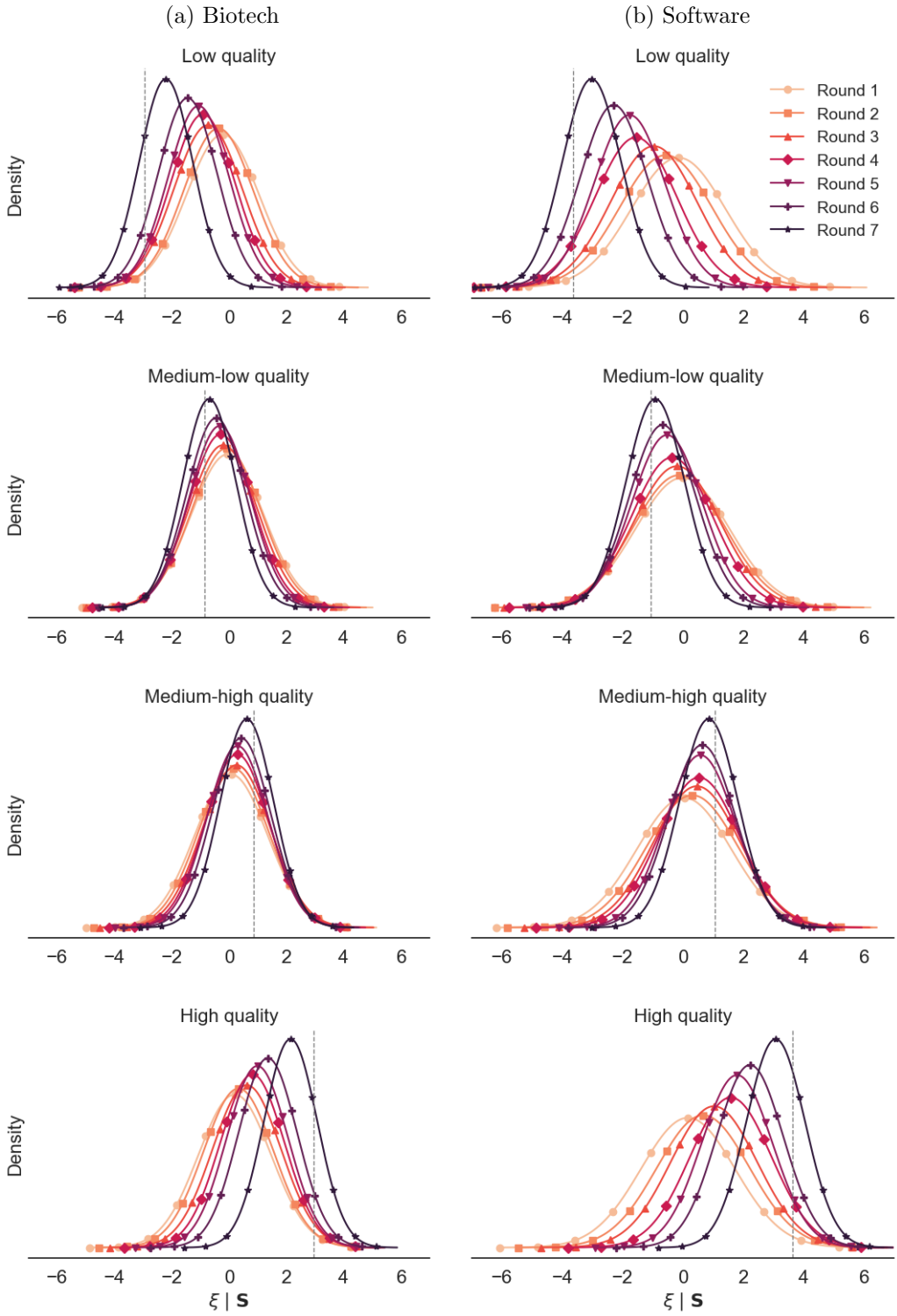
6.1 Information and learning

To get a sense of what the estimated distributions of unobserved quality and signal noise imply about the learning process in each sector, I choose low-, medium-low-, medium-high-, and high-quality startup (ξ at the 1st, 25th, 75th, 99th percentile) and simulate investors' posterior beliefs about its quality 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.

Statistically, I measure the convergence in beliefs by coefficient of variation (CV), defined as $\frac{\tau_k}{\mu_k}$. As shown in Table 6, 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 reinforces

Figure 8: **Learning process**



Note: This figure shows, for startups of different quality, the posterior belief about quality after each round. The x-axis represents the belief on ξ conditional on the signals $\boldsymbol{S}$. The grey dashed lines mark the true quality. The 1st, 25th, 75th, 99th percentile of ξ is selected to represent low, medium-low, medium-high, high quality.

the observations in Fig. 8 and 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 6: **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.

These findings reflect sector-specific business models: software businesses are typically built on proven technologies, so the main uncertainty lies in market demand for a new product or service, which is usually concentrated at the outset. In contrast, drug development involves several distinct stages, each testing a different aspect of the drug; thus, uncertainties for biotech startups are distributed more evenly throughout the process, leading to a slower learning rate than in software startups.

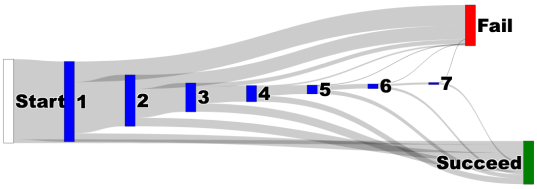
6.2 Exit outcome and startup heterogeneity

With my estimates, 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.

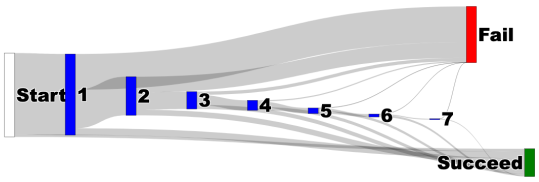
To better understand the factors driving a startup's fundraising journey, I compare the characteristics of startups on different paths. In Table 7, I present the average number of patents held by startups that either secure funding in a given round and exit immediately via IPO or M&A, or fail to secure 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 in 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 in Round 1 to 7 are 42.6, 18.6, 4.0, 1.4, 0.3, 0.1, 0.1.

Table 7: **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 shows 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. These patents consist of both in-house developed patents and those licensed from other entities.

Successful startups tend to have more patents than failed ones. 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. See

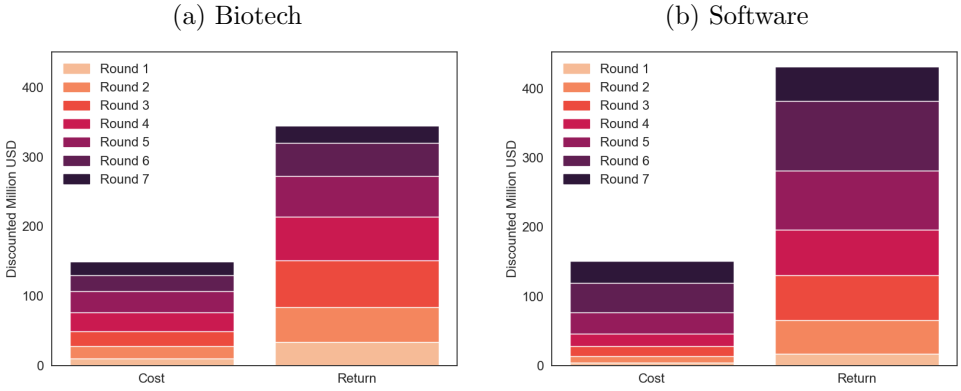
Supplemental Appendix for a robustness check that simulates a startup's trajectory holding all other characteristics constant and varying only the number of patents.

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. I analyze how interaction between sequential dissemination of information and VC competition shape return sharing across 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. 10). 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 return-sharing across stages stems from the sequential dissemination of information rather than risk-averse agents in standard financial models such as CAPM.

Figure 10: **Cost and return sharing**



Note: In Fig. 10a, the costs (capital investments) 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. 10b, 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 more evenly across rounds, while software pushes most costs to later stages, consistent with the prioritized activities over rounds — in biotech, each stage involves a steadily expanding pool of volunteers, matching the *steady* increase in the costs over rounds; however, in software, 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.

7 Inefficiency and Welfare

Having understood the learning process in the current market, I then quantify the gap between current efficiency and that achievable under alternative information and market structures. There are three sources of inefficiency in the current market: 1) dynamic information externalities; 2) uncertainty in startup quality; and 3) uncertainty regarding the number of future rounds and future capital needs. The first source, as discussed in Section 3.2, 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 ones) 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

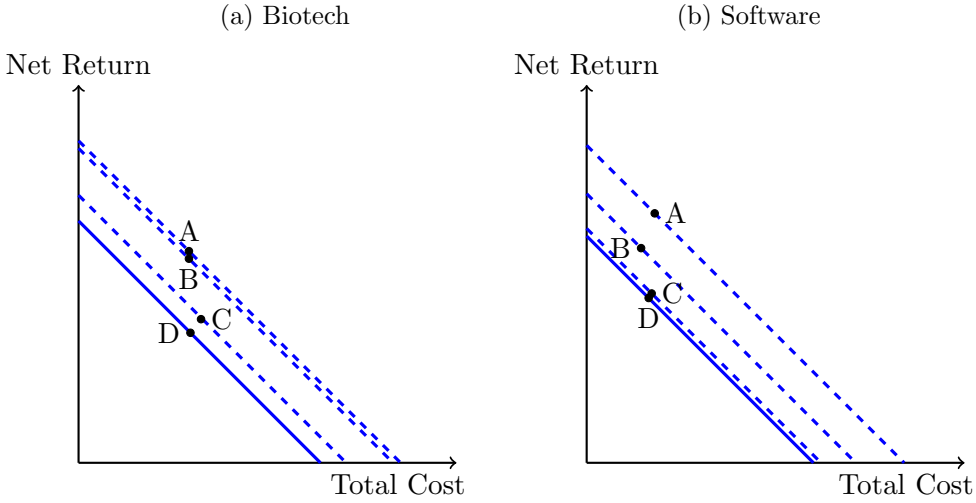
I define *first-best* efficiency as the benchmark where all NPV-positive startups are funded through and all NPV-negative startups are rejected from the start — the maximum efficiency attainable when all sources of inefficiency are removed. *Second-best* is achieved when a net return-maximizing planner, fully informed of startups' true qualities, makes funding decisions — i.e., when both information externalities and quality uncertainty are absent. *Third-best* is similar to second-best, except the planner doesn't know the true quality and only sees all VCs' signals, so only information externality is eliminated. *Status quo* is the efficiency under the current market practice when all the three inefficiencies present.

In Fig. 11, point A, B, C, D respectively show the net return and total cost of 1st-, 2nd-, 3rd-best, and status quo efficiency. Two things are worth noting: (1) in both sectors, investors at D invest less than the planner at C, consistent with the theoretical result of under-investment in Proposition 3; the vertical gap — \$10 billion in biotech and \$3 billion in software — measures the efficiency loss from dynamic information externality; (2) B has lower costs but higher net returns than C, due to better identification of NPV-positive projects; the vertical gap — \$39 billion in biotech and \$30 billion in software — captures the efficiency loss from quality uncertainty.

In addition to the aggregated and ex-post outcomes in Fig. 11, Table 8 presents the average expected value of individual startups across the four scenarios.

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% in 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

Figure 11: **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. 11a, and (44.8, 264.9), (35.82, 241.9), (43.1, 212.0), (40.7, 209.0) in Fig. 11b, all in billion USD. All results are based on the same simulated data as before.

Table 8: **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

(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

Note: This table reports the mean expected value of pre-investment startups (million USD) by round across scenarios. To remove selection effects, expected value is set to 0 for NPV-negative startups under the first-best scenario, and for those that fail to raise or reach that round under the other scenarios. Gaps between scenarios are calculated as one minus the ratio of mean expected values.

about startup quality becomes available. Information externalities drive the gap between D and C, which narrows over rounds as there are fewer future VCs benefitting from late-stage externalities.

7.2 Type I and Type II errors

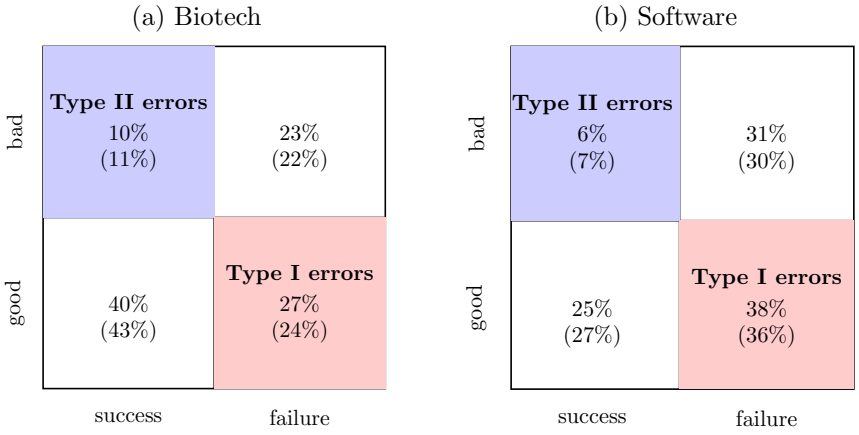
In this subsection, I quantify the losses associated with Type I errors — not funding good startups, and Type II errors — funding bad startups. The results have important policy implications: if Type I errors dominate, the government should complement the private VC market; if Type II errors dominate, 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 (getting acquired or going public) are labeled as *good-success*, good startups that fail are labeled as *good-fail*. Similarly for *bad-success* and *bad-fail*.

Fig. 12 presents the proportions of startups in each group in status quo and 3rd-best scenario. 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.

Next, I quantify the monetary loss from Type I (II) errors, defined as the gap in net returns for good (bad) startups between status quo and a benchmark, chosen as either the 1st- or 3rd-best scenario (see Table 9). Take the first row in Table 9a for example. The net returns on all good startups under 1st-best, 3rd-best and status quo scenarios are 239.75, 205.99, 195.11 billion USD, thus the gap in net return between status quo and 1st- or 3rd-best are respectively 44.64 and 10.88 billion USD — the

Figure 12: **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). These percentages are based on the same simulations as before.

dollar losses from Type I errors. Using either benchmark, for both sectors, Type I losses far exceed Type II losses, suggesting that "not funding good startups" is the main reason for efficiency losses in this market.

Table 9: **Quantifying Type I and Type II errors**

(a) Biotech

	Net return			**Loss**	
	1st-best	3rd-best	status quo	1st-best	3rd-best
Good	239.75	205.99	195.11	44.64	10.88
Bad	0.00	-10.28	-8.78	8.78	-1.50

(b) Software

	Net return			**Loss**	
	1st-best	3rd-best	status quo	1st-best	3rd-best
Good	264.98	219.03	215.20	49.78	3.83
Bad	0.00	-6.96	-6.13	6.13	-0.83

Note: This table shows the loss in net return in status quo compared with 1st- and 3rd-best scenario for good and bad startups, respectively. All the numbers are in billion USD.

Following this, I explore when Type I errors occur. Table 10 presents the timing of failure for good-fail startups. For example, 53% out of all 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 few phases, good startups are almost guaranteed to secure additional funding. Software startups show similar patterns, but more fail in the first round due to (more) inaccurate initial signals. These results reinforce the insights presented in Table 6.

Table 10: **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 intriguing but beyond the scope of this paper.

7.3 Role of Acquisition as an Exit Option

In practice, VCs typically exit a startup when it either goes IPO or is acquired by an established company. Using my estimates, I study the role of M&A as an exit route for startup dynamism through counterfactuals that vary the probability of successful exit via an M&A. While not explicitly modeling regulatory changes, this exercise can shade light on the potential consequences of stricter merger enforcement, marked by the DOJ and FTC's new merger guidelines that intensify antitrust scrutiny of incumbent acquisition of startups (or "killer acquisitions").

Specifically, I 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. 11, I model tougher M&A exits by reducing the realized value (or equivalently, the probability) of successful M&A exits by 10 to 50%, 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 11 reports the cost and return estimates under each level of "devaluation". *Change in net return* is measured relative to the 0% devaluation case and is decomposed into extensive margin — fewer startups receiving funding, and intensive margin — lower valuation for funded startups.

Table 11: **Counterfactuals of Stricter M&A**

(a) Biotech

M&A	**Overall**			**Change in net return**		
Devaluation	Return	Cost	Net return	Total	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	Total	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. 11. All values are in billion USD.

For example, in biotech, when M&A exits are devalued by 10%, overall return falls to $236.3 billion and cost to $69.6 billion. The net return is $19.6 billion lower than under 0% devaluation, driven by (1) *extensive margin* — fewer projects get funded as VCs expect lower returns and become more selective in the startups they choose to fund, and (2) *intensive margin* — lower valuations for startups funded both before and after tougher M&A exits. Of the $19.6 billion decline in net return, $6 billion is attributable to extensive margin and $13.6 billion to intensive margin. Moreover, tougher M&A exits reduce net returns more sharply in software than in biotech, since

M&A is a more prevalent exit channel in software.

These results attest to the concerns with tougher M&A enforcement, including fewer exit opportunities, depressed valuations, and VCs being more conservative about funding startups that rely on acquisition by large incumbents, yet tougher M&A policies may push more startups to become standalone companies, potentially strengthening long-term innovation. A full assessment of these competing effects requires analyzing downstream competition dynamics, which I leave for future work.

8 Discussion

I have shown that understanding VCs' investment strategies requires assumptions on what they bid in the equity space and how they add value. I set aside important 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 Value-added by VC

Throughout this paper, I have assumed that VCs do not add value beyond their capital contributions, so that their bids are based solely on their private signals but not any potentially heterogeneous ability to add value. This is because in the data, I observed only the winning VC, with no information on bids or identities of those that made an offer but were rejected. Given this, any VC value-added cannot be separately identified from the exit value.

To explain, consider two identical startups, both raised one funding round from the same VC and then exited via IPO. Everything is the same except for the investor's ownership, which depends not only on the signals (reflecting unobserved quality) and value-added of the winning VC, but also on that of the strongest competitor. Without knowing the identity of this competitor, I'm not able to distinguish how much of the winning bid is attributable to VC's assessment of the firm versus the pressure from the second strongest VC candidate, whose value-added ability influences the competition

dynamics but remains unknown due to the lack of information about his identity.

While convenient, this assumption is at odds with some research suggesting that VCs add values through networks, talent recruitment, and operational support (Bernstein et al., 2016). Hence, I might have over-estimated the level of uncertainty as it also picks up the variation in VC value-added ability.

To incorporate VC value-added, at the very least, I need to find the identities of losing bidders. One way is to manually construct the set of competitors using text data from SEC filings, press release, investor comments (VC websites, twitter, podcasts) and startup's old pitch decks. Similar idea has been used previously by Slattery (2025) to study subsidy competition among US states to attract firms. She used newspaper articles to find the top two competitive states, and from there the identity of a losing bidder. Using this richer data, I could extend the previous model of English auction with common value to affiliated value, where the startup's value to each bidder depends on both his value-added ability and an unknown common quality. VC-specific value-added could then be identified by observing the same VC investing across different startups.

8.2 Bidding space

I also assume that VCs bid only on equity ownership, whereas in reality they may structure contracts contingent on realized exit values. This simplification reflects both data limitation — a non-trivial portion of contract terms other than equity ownership is missing in Pitchbook — and model (in)tractability — the bidding process becomes significantly more complicated when allowing for multi-dimensional bids and even more so in a multi-agent dynamic setting. While some studies examine optimal contract design in single-agent static settings (DeMarzo et al., 2005), much less is understood when contracts extend across dynamic rounds. Thus, it is currently challenging to derive meaningful implications for VC financing in such settings.

8.3 Follow-up investment

Another assumption I make is that each VC leads a single round. This implies that, when evaluating a startup's value, a VC does not consider the option value of future reinvestments. Some justifications for this assumption include the fact that VCs often specialize in different stages of a startup due to distinct expertise and fund sizes (Brown et al., 2023), as well as the empirical observation in Section 2.2.6 that nearly 90% of VCs lead only a single round in a given startup (there is notable overlap in syndicate members across rounds, but this overlap diminishes as the rounds grow further apart). 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, as well as the pricing trade-off between new ownership and dilution of existing ownership for internal VCs, likely adding substantial computation complexity to it.

To see how my estimates might be biased if this assumption does not hold, consider two extreme cases: one where each round is led by a different VC, and another where a single VC leads all rounds. The first case reflects my current setup, while the second aligns with the planner benchmark and thus achieves ex-ante Pareto efficiency. If reality lies somewhere in between, then my estimated inefficiency is likely an upper bound on the true value, since the ability to follow up on the same startup brings the model structure closer to the planner benchmark.

9 Conclusion

In this paper, I develop a tractable empirical framework to understand the strategic investment decisions of competing VCs under uncertainties, and to quantify the efficiency gap relative 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 competition among VCs.

Using a novel dataset of biotech and software startups, together with the model, I identify and quantify a new source of market failure due to dynamic information externality. My estimates suggest that currently the two sectors both yield around 80% of the first-best efficiency, and information externalities take up 18% of this loss in biotech and 5% in software, our estimates also suggest that this loss in value is mainly due to not funding good startups.

My paper is the first to incorporate uncertainty, staged financing, dynamic learning, and imperfect competition into the VC-funding space. To do so requires taking a stand on how VCs compete for a startup. I hope that future researchers studying VC-startup interactions will explore how to capture VC value-add and other important contract dimensions, and what new insights their estimates provide in understanding the role of VC.

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.

Gregory W Brown, Celine Yue Fei, and David T Robinson. Portfolio management in private equity. Technical report, National Bureau of Economic Research, 2023.

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.

Robert Clark, Jean-François Houde, and Jakub Kastl. The industrial organization of financial markets. In *Handbook of Industrial Organization*, volume 5, pages 427–520. Elsevier, 2021.

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.

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.

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

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.

Hidehiko Ichimura. Semiparametric least squares (sls) and weighted sls estimation of single-index models. *Journal of econometrics*, 58(1-2):71–120, 1993.

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.

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

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.

Jean-Jacques Laffont and Quang Vuong. Structural analysis of auction data. *The American Economic Review*, 86(2):414–420, 1996.

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.

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.

Cailin Slattery. Bidding for firms: Subsidy competition in the united states. *Journal of Political Economy*, 133(8):000–000, 2025.

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.

A Variable Description

Table 12: **Variable Description**

Variables	Description	Source
A. Startup level		
Sector	1 for biotech, 0 for software	PB
Patent	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	Startup realized value — 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 N/A for failed companies	PB, SEC, CT, FS
B. Startup-round level		
Round	The chronological order of a VC round, e.g., for the 3rd VC round, *Round*=3	PB
Investment	Capital investment injected to a startup in a round	PB
Valuation	Price of the entire company in a given round (i.e., post-money valuation)	PB
EquityOwnership	Ratio of shares held by this round's investors to total outstanding shares, it also equals to $\frac{Investment}{Valuation}$.	PB

Note: PB, SEC, CT, FS are abbreviations for Pitchbook, SEC Filings, Compustat and FactSet.

B Summary Statistics

B.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
Investment	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
Valuation	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: Investment, *Valuation* and *ExitValue* are in million USD as of Dec 2023. *EquityOwnership* is between 0 and 1.

B.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
Exit	*IPO*	0.11	0.31	0	0	0	0	1
	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 Dec 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.

C Covariance Matrix

The posterior mean and variance are

$$\mu_k = \omega_0^k \cdot 0 + \sum_{l=1}^{k} \omega_l^k \bar{s}_l \tag{19}$$

$$\tau_k^2 = \omega_0^k \sigma_\xi^2 \tag{20}$$

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 $\alpha_l = \frac{\frac{1}{J_l}\sigma_l^2}{\frac{1}{J_l}\sigma_l^2 + \left(\prod_{l'=0}^{l-1}\alpha_{l'}\right)\sigma^2}, l = 1, ..., k$. 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. (19),

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

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

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

D 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. (19), 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. (5).

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. (4)

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. (3),

$$\begin{aligned} V_{\bar{K}}(\boldsymbol{X}_{\bar{K}}, \mu_{\bar{K}}) &= \mathbb{E}\left[W_{\bar{K}} \middle| \boldsymbol{X}_{\bar{K}}, \mu_{\bar{K}} \right] \\ &= \mathbb{E}\left[m_{\bar{K}}(\boldsymbol{X}_{\bar{K}}) \exp(\xi) \middle| \boldsymbol{X}_{\bar{K}}, \mu_{\bar{K}} \right] \\ &= m_{\bar{K}}(\boldsymbol{X}_{\bar{K}}) \exp\left(\mu_{\bar{K}} + \frac{1}{2}\sigma_{\bar{K}}^2 \right) \end{aligned} \tag{21}$$

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.

$$\begin{aligned}
V_k(\boldsymbol{X}_k, \mu_k) &= \Pr(K = k \big| K \geq k, \boldsymbol{X}_k, \mu_k) \mathbb{E}\left[W_K \big| \boldsymbol{X}_k, \mu_k, K = k\right] \\
&+ \Pr(K > k \big| K \geq k, \boldsymbol{X}_k, \mu_k) \mathbb{E}\left[(1 - a_{k+1}) q_{k+1} V_{k+1}(\boldsymbol{X}_{k+1}, \mu_{k+1}) \big| \boldsymbol{X}_k, \mu_k, K > k\right] \\
&= \Pr(K = k \big| K \geq k, \boldsymbol{X}_k) m_k(\boldsymbol{X}_k) \exp\left(\mu_k + \frac{1}{2}\sigma_k^2\right) \\
&+ \Pr(K > k \big| K \geq k, \boldsymbol{X}_k) \delta \mathbb{E}\left[(1 - a_{k+1}) q_{k+1} V_{k+1}(\boldsymbol{X}_{k+1}, \mu_{k+1}) \big| \boldsymbol{X}_k, \mu_k, K > k\right]
\end{aligned} \tag{22}$$

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.

$$\mathbb{E}\left[(1 - a_{k+1}) q_{k+1} V_{k+1}(\boldsymbol{X}_{k+1}, \mu_{k+1}) \big| \boldsymbol{X}_k, \mu_k, K > k\right] \tag{23}$$

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. (4) and (5). Plugging it to Eq. (23),

$$\mathbb{E}\left[\underbrace{\left(1 - \frac{d_{k+1}}{V_{k+1}(\boldsymbol{X}_{k+1}, \hat{\mu}_{k+1})}\right) \mathbb{1}\{V_{k+1}(\boldsymbol{X}_{k+1}, \mu_{k+1}) > d_{k+1}\} V_{k+1}(\boldsymbol{X}_{k+1}, \mu_{k+1})}_{\dot{V}_{k+1}(\boldsymbol{X}_{k+1}, \mu_{k+1}, \hat{\mu}_{k+1})} \big| \boldsymbol{X}_k, \mu_k, K > k\right] \tag{24}$$

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. (24) 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. (24) 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,*

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

See Supplemental Appendices for proof of Lemma 1.

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*

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

Lemma 2 is straightforward: under (x, y), more weight is shifted toward higher values of u due to its monotonicity and that the distribution of (x, y) places greater mass on higher values than does (x', y').

Combining these two lemmas, it's straightforward that Eq. (24) is strictly increasing in μ_k. Hence, $V_k(\boldsymbol{X}_k, \mu_k)$ is strictly increasing in μ_k.

E 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. (22) and (24), the Bellman representation for the *maximum* possible payoff to round k investor, V_k^*, is as follows:

$$\begin{aligned} V_k^* = {} & \Pr(K = k \big| K \geq k, \boldsymbol{X}_k) m_k(\boldsymbol{X}_k) \exp\left(\mu_k + \frac{1}{2}\sigma_k^2\right) \\ & + \Pr(K > k \big| K \geq k, \boldsymbol{X}_k) \delta \mathbb{E}\left[\mathbb{1}\{V_{k+1}^* \geq d_{k+1}\}\left(V_{k+1}^* - \frac{V_{k+1}^*}{\hat{V}_{k+1}^*} d_{k+1}\right) \Big| \boldsymbol{X}_k, \mu_k, K > k\right] \end{aligned} \tag{25}$$

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

$$V_k^\dagger = \Pr(K = k \big| K \geq k, \boldsymbol{X}_k) m_k(\boldsymbol{X}_k) \exp\left(\mu_k + \frac{1}{2}\sigma_k^2\right) + \Pr(K > k \big| K \geq k, \boldsymbol{X}_k) \delta \mathbb{E}\left[\mathbb{1}\{V_{k+1}^\dagger \geq d_{k+1}\}\left(V_{k+1}^\dagger - d_{k+1}\right) \big| \boldsymbol{X}_k, \mu_k, K > k\right] \tag{26}$$

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 $k < \bar{K}$.

F Identification

Proposition 4 shows that the unobservable μ_k can be isolated from the observables $\boldsymbol{X}_k$ in the funding choice described by Eq. (8). This separation relies on the monotonicity of V_k^* in μ_k, as established in Proposition 2. Hence, the choice equation (8') 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. (8) implies Eq. (8'):

$$q_k = \mathbb{1}\{\mu_k \geq h_k(\boldsymbol{X}_k)\}, \text{ where} \tag{8'}$$

$$h_k(\boldsymbol{X}_k) = V_k^{*-1}(d_k; \boldsymbol{X}_k). \tag{27}$$

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. (7)(8')(9)(10), where ξ,μ_k,$\hat{\mu}_k$ are unobserved, $\boldsymbol{V}^* = (V_k^*)_{k=1}^{\bar{K}}$ satisfies Eq. (3) 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.

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. (7) (8')(10).*

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* $Supp(\mu_1) \times ... \times Supp(\mu_{\bar{K}}) \times Supp(\xi)$.

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 section 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. (9), $\boldsymbol{V}^*$ is identified.*

The reason that $\boldsymbol{V}^*$ cannot be directly identified by inverting $\boldsymbol{h}$ is that as is shown in Eq. (27), d_k also shows up in $\boldsymbol{X}_k$ of V_k^{*-1}, thus the full rank condition on covariates is violated. However, Eq. (9) 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 Section 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 section I.

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. (21),

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

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

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

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. (9),

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

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 $K = k$.

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

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,

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

Note that the numerator is identified directly from data, the denominator can be computed given $\sigma_\xi, \boldsymbol{\sigma}, \boldsymbol{h}$. Therefore, $f(\boldsymbol{X}_k \big| K=k)$ is known up to c_k. With an additional condition of $\int f(\boldsymbol{X}_k \big| K=k) d\boldsymbol{X}_k$=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. (25) that connects V_k^* and V_{k+1}^*. Note that in Eq. (25),

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 $k' \geq k$.