Startup Catering to Venture Capitalists*

Xiyue (Ellen) Li†

December 7, 2023

[Click here for the latest version]

Abstract

I show that information frictions in valuation can lead startups to select projects that align with the expertise of potential venture capital (VC) investors, a strategy I refer to as *catering*. First, I build a theoretical model where a startup trades off project quality with the informational benefits of catering. The startup selects catering when alternative information sources are limited or VC investors demonstrate proficiency in valuing projects close to their expertise. Second, using textual data from patent applications, I define catering projects as patent applications that deviate from the founders' experience toward VC's expertise. Consistent with model predictions, catering applications are more prevalent when patent examination is slow or VCs utilize past data to screen new deals. Catering applications are 19.3% less likely to get patent approval, suggesting low project quality. Overall, this paper shows that specialized financial intermediaries, such as VC, can broadly shape new technology developed by firms outside their portfolios.

JEL Codes: L26, D83, G24, O32

Keywords: Entrepreneurship, Information Frictions, Real Option, Venture Capital, Innovation, Patents

---

*The paper was previously circulated under the title "More Exploratory, More Opaque: Startup Project Choice in Pursuit of Venture Capital Financing". I am deeply grateful to my dissertation committee Jan Bena (Co-chair), Ali Lazrak (Co-chair), Lorenzo Garlappi, Will Gornall and Ralph Winter for their guidance and continuous support. I thank Isha Agarwal, Bo Bian, Markus Baldauf, Michael Ewens, Adlai Fisher, Thierry Foucault, Jack Favilukis, Joyce Guan, Zhichao Li (discussant), Xing Liu, Dongliang Lu, Peter Pham, Elena Pikulina, Hedieh Rashidi Ranjbar (discussant), Zhanbing Xiao, Chi Zhang (discussant), Yao Zeng, Keling Zheng and seminar participants at University of British Columbia, BFBRG Workshop at University of Sydney, and WEFI Fellow for helpful comments. All errors are my own.

†Sauder School of Business, University of British Columbia. Email: xiyue.li@sauder.ubc.ca. Website: http://www.ellenxiyueli.com

1 Introduction

A lack of information in valuation is a primary challenge for entrepreneurs of innovative startups. Startups commonly face a considerable time lag before generating first revenue or obtaining a patent review decision, during which the value of startups can be highly uncertain.[1] Information frictions lead to inefficient terminations by entrepreneurs and discourage external financiers from funding otherwise successful startups, both of which reduce entrepreneurial payoffs. Despite the critical role of information in entrepreneurial success, the existing literature rarely explores how entrepreneurs can actively manage the information available to them to attain higher payoffs.

In this paper, I show that entrepreneurs can alleviate information frictions in assessing project value by initiating projects close to the expertise of potential venture capital (VC) investors, a strategy I refer to as *catering*. First, I develop a theoretical model to show that entrepreneurs select catering at the cost of project quality when (i) information sources beyond venture capitalists (VCs) are limited or (ii) VCs are highly skilled in valuing projects close to their expertise. Second, I use first-time patent applications by US startups to construct an empirical measure of startup catering. I identify catering projects as patent applications that are far from the entrepreneurs' previous experience but align closely with projects funded by VCs. I provide empirical evidence supporting the model's predictions and show that catering projects are 19.3% less likely to obtain patent approval, suggesting lower project quality.

The key model assumption is that an entrepreneur selects and experiments with one of her two entrepreneurial ideas. The project selection involves a trade-off between project quality and information. One project (*exploration*) generates a higher payoff conditional on success but is far beyond VCs' expertise. The other project (*catering*) faces less information friction when valued by VCs but generates a low payoff. Project selection occurs before the entrepreneur engages with VCs for investment and project valuation.

In a frictionless world, the entrepreneur optimally selects exploration to achieve the first-best payoff. However, in the presence of information frictions, the entrepreneur may decide to cater. Catering yields informational benefits through two channels. First, catering leads to efficient

[1] Lack of information is prominently mentioned in the literature on entrepreneurial finance as a key defining feature of this important sector of the economy (Hall, 2002; Hall and Lerner, 2010).

investment decisions. In particular, VCs can make their investment decisions more efficiently by identifying and funding successful projects. This, in turn, helps the entrepreneur recognize and abandon unsuccessful projects. Second, the entrepreneur exerts low effort to engage with VCs to achieve efficient investments. In equilibrium, the entrepreneur caters when VCs are highly skilled in screening projects close to their expertise or when facing frictions in accessing alternative information sources beyond VCs.

I test the model predictions using 62,088 first-time patent applications by US startups from 2007 to 2018. Following Kelly et al. (2021), I measure the proximity of a startup's project to VC expertise, hereafter *VC Informativeness*, by textual similarity between the startup's first-time patent application and previous applications from firms in the VCs' portfolio.[2] In line with the model predictions, I find that startups of high VC informativeness are more likely to secure VC investments. One standard deviation increase in VC informativeness predicts a 27% increase in the probability of investment compared to the sample mean. I also show that startups of high VC informativeness experienced a shorter time on capital seeking. In particular, for the subsample of startups that accessed VC financing in the sample period, high VC informativeness associated with a shorter time to first VC investments. Furthermore, for those startups who failed to be funded and eventually terminated their business, an increase in VC informativeness predicts faster closure. Taken together, these results support the benefits of having projects close to VCs' expertise consistent with my theoretical model.

Next, I develop an empirical measure for startup catering. In particular, I define a patent application as a catering project if it (i) differs from the patenting experience of startup inventors and (ii) aligns closely with VCs' expertise. I show that, on average, catering applications are of lower quality in the sense that they underperform compared to other patent applications filed within the same industry and year. Notably, catering applications are 19.3% less likely to receive patent approval, and I find no evidence that they result in higher citations conditional on approval. These findings validate the model assumption that catering projects may come at the expense of project quality.

I further provide evidence regarding the determinants of startup catering. The model predicts

[2] This notion stems from the idea that VCs adopt expertise developed from monitoring portfolio firms to evaluate new deals (Sorensen, 2008).

that startup catering is more prevalent when alternative information sources are limited. I employ patent examination as a proxy for the alternative information source and use the speed of patent examination to measure its efficiency.[3] I find that startup catering is more prevalent in technology classes with slower patent examination, which supports the model prediction. One standard deviation increase in examination speed is associated with a 33% decrease in startup catering.

The model also predicts that startups are more likely to cater when VCs are highly skilled in valuing projects close to their expertise. I use VCs' data technology adoption to measure their skills in valuing projects. Upon adopting data technology, VCs are notably more adept at assessing the value of projects akin to their previous investments, while this advantage might not extend to exploration projects (Bonelli, 2023). Consequently, the catering project brings more informational benefits when VCs are data-driven. I define a VC firm as data-driven if the VC firm hires data scientists for deal screening. My results suggest that startup catering is more common in technology classes invested by data-driven VCs, which supports the informational channel proposed by the model.

In sum, this paper illustrates how specialized financial intermediaries, such as VC, can shape the nature of new technology. Specialized financial intermediaries help to mitigate information friction in valuation and increase entrepreneurial payoff, but only for entrepreneurs with projects close to their expertise. Hence, entrepreneurs may strategically cater to VC's expertise in project selection to obtain informational benefits at the expense of project quality. While VCs' adoption of data technology can reduce information friction in screening, it *cannot* lead to the first-best outcome because it causes more ex-ante catering behavior by entrepreneurs.

My paper adds to the literature on the real effects of venture capital. Existing literature finds that VCs spur innovation of portfolio startups through active monitoring (Kortum and Lerner, 2001; Bernstein et al., 2016), value-added services (Lerner, 1995, Hellmann and Puri, 2002, Hsu, 2004, Chemmanur et al., 2011), tolerance for failure (Manso, 2011; Tian and Wang, 2014), and security design (Hellmann, 1998; Cornelli and Yosha, 2003; Kaplan and Strömberg, 2003; Kaplan and Strömberg, 2004). This strand of literature highlights *ex-post* effects of VCs on their portfolio

[3] Existing literature shows that patent approval helps reduce information frictions in assessing startup values (Farre-Mensa et al., 2020), and its delay affects startup financing (Farre-Mensa et al., 2016).

firms *after* investments.[4] In contrast, my paper illustrates potential *ex-ante* real effects by VCs. Startups may strategically opt for projects close to VCs' expertise to take advantage of VCs' information for future learning. Through this channel, VCs can potentially shape the startup landscape by influencing the project selection of startups outside their portfolios.[5]

This paper also contributes to the real options literature, with a specific emphasis on entrepreneurial firms. The option feature of investment is especially relevant for innovative startups given their inherent uncertainty and skewed return distributions (Scherer and Harhoff, 2000; Hall and Woodward, 2010). Kerr et al. (2014) and Manso (2016) argue that experimentation is a key feature of entrepreneurship. In particular, entrepreneurial payoffs crucially depend on abandoning poorly performing projects and capitalizing on projects revealing good outcomes. Existing literature finds that entrepreneurial entry and financing are influenced by the cost of experimentation (Ewens et al., 2018), financing risk (Nanda and Rhodes-Kropf, 2013; Nanda and Rhodes-Kropf, 2016), downside insurance (Hombert et al., 2020; Gottlieb et al., 2022), and information frictions (Gompers, 1995; Bergemann and Hege, 1998; Bergemann et al., 2009). My paper focuses on information frictions in project valuation and shows that entrepreneurs can manage information through project selection.

Finally, my paper relates to the literature showing that financial investors, such as accelerators, crowdfunding investors, and VCs, serve as a source of information that guides entrepreneurs in investment decisions.[6] For example, Yu (2020) and Howell (2021) show that startups that receive negative feedback from accelerators or new venture competitions are more likely to close down, suggesting that entrepreneurs learn about the value of their project from the feedback. Xu (2018) and Chemla and Tinn (2020) conclude that entrepreneurs can acquire information about market demand through crowdfunding platforms. Gonzalez-Uribe et al. (2023) find evidence that VC due diligence yields informational benefits to entrepreneurs, irrespective of investment outcomes. My paper builds on this literature and shows that the informational role of investors can lead to

[4]See Lerner and Nanda (2020) for the review of literature on VC's role in financing innovation.

[5]Samila and Sorenson (2011) show that the supply of VC encourages entrepreneurial entry, which also aligns with the idea that VCs may affect startups outside of their portfolio. My paper, instead, focuses on the information channel through which VCs affect startup project selection.

[6]This paper focuses on the informational role of investors in the private market. See Bond et al. (2012) and Goldstein (2023) for the review of feedback effect literature, which shows the real effects of information from public market investors contained in stock prices.

catering behaviors by entrepreneurs. Initiating projects close to VCs' expertise helps to achieve efficient investments and enhance entrepreneurial payoff, though it may come at the cost of project quality.[7][8]

The remaining of this paper is organized as follows. In Section 2, I develop the theoretical model and discuss the corresponding empirical predictions. In Section 3, I consider comprehensive empirical tests to confirm model predictions. Section 4 concludes. All proofs are in the Appendix.

2 Model

This section introduces a theoretical model to illustrate the benefits and determinants of startup catering. Section 2.1 outlines the model's setup. I start by assuming that the entrepreneur possesses a project to experiment with. The entrepreneur optimally determines the effort spent on engaging with VCs in the financial market for investment and project valuation. Section 2.2 analyzes the equilibrium strategy of the entrepreneur and shows the channel through which project proximity to VCs' expertise impacts entrepreneurial payoffs. Section 2.3 endogenizes the project choice of the entrepreneur and discusses the determinants of startup catering to VCs' expertise. Section 2.4 discusses model extensions. Section 2.5 summarizes the empirical predictions of the model.

2.1 Setup

The economy has four dates $t \in \{0, 1, 2, 3\}$ (See Figure 1). At date $t = 0$, the entrepreneur initiates a project to experiment. The project has an unknown type for the entrepreneur upon its initiation. It is of high type ($T = H$) with probability p, in which case it generates a payoff of R, or of low type ($T = L$) with probability $1 - p$, in which case it pays out 0. To produce cash flow, the entrepreneur requires a one-time investment of k at date $t = 2$ from the VCs in the financial market to finalize the project. To capture the experimental nature of entrepreneurship, I assume

[7] Foucault and Frésard (2019) shows that firm product differentiation reduces stock price informativeness by restricting cross-asset learning. Such informational cost induces firm conformism. My paper also highlights the information-payoff trade-off by firms, while my focus is on the private market. In particular, I consider endogenous learning effort and limited cross-asset learning with VCs as intermediaries.

[8] Yang and Zeng (2019) discusses how entrepreneurs optimally design securities to leverage financial investors' expertise for efficient investments. Different from their work, my paper emphasizes project choice as a means to alleviate information frictions.

that $R > k$ and $pR < k$. A high-type project has a positive NPV. Unconditionally, however, the project has a negative NPV.

Figure 1: [INSERT FIGURE HERE]

At date $t = 1$, the entrepreneur decides on the effort level devoted to engaging with VCs, which is denoted as $n \in \mathbb{N}$. We can interpret the parameter n as the number of VCs the entrepreneur contacts. Each VC generates a valuation of the entrepreneur's project, which reflects the VC's information regarding the project type. Those valuations, from the entrepreneur's perspective, act as signals that help determine project type. The informativeness of those signals depends on whether the contacted VC possesses the requisite knowledge to evaluate the entrepreneur's project.

For simplicity, I assume the financial market consists of two types of VCs. A fraction of q among the VCs possess knowledge evaluating the project, enabling them to provide accurate signals regarding project type. The remaining VCs, however, lack experience with the entrepreneur's project and therefore cannot provide any information about the project type. In sum, each random VC possesses information with a probability of q. To better illustrate the key intuition of the model, I assume that the information at the hands of the VCs is costless.

All signals from VCs can be observed at $t = 2$. In addition to the signals from VCs, there is a probability of λ that the entrepreneur receives another publicly available signal that discloses the project's type. This signal can be interpreted as an event such as patent approval or profit generation. Notice that this signal is a perfect substitute for the information obtained from VCs.

The project requires a one-time investment of k at $t = 2$ to be completed. This investment can be interpreted as scaling project size or the cost required to bring technology to the product market. According to signals received, the startup makes investment decisions $I_2 \in \{1, 0\}$. $I_2 = 1$ represents raising capital of k from the VCs in the financial market to invest in the project, while $I_2 = 0$ represents abandonment. Abandon has no cost and yields zero payoff. If invested, the project generates a cash flow at $t = 3$.

One feature of the model is that the entrepreneur learns about project type through interacting with VCs. There are two assumptions behind this feature: First, VCs possess the skills necessary to produce information about the project type, and their information, at least to some extent, complements the knowledge of the entrepreneurs. Second, VCs can deliver this information to entrepreneurs through pre-investment screening. In other words, VCs bring additional insights and expertise that augment the entrepreneur's understanding of the project, irrespective of the investment outcome. Those two assumptions are supported by González-Uribe et al. (2022).[9] VC pre-investment screening is a multi-stage selection process involving intensive information exchange with entrepreneurs. The information provided by VCs can potentially enhance the entrepreneur's understanding of the project's value, facilitating informed investment decisions.

The parameter q describes the extent to which the entrepreneurial project is close to the expertise of VCs. A higher value of q indicates high proximity between the project and VCs' expertise and implies that VCs can provide precise project valuation to the entrepreneur. In other words, a higher q suggests a lower level of information frictions faced by VCs in valuing the project. This parameter can be interpreted as a characteristic of the entrepreneur's project. For example, if a project involves radical technologies, its value can be challenging for VC to assess due to the lack of comparable firms as references. Such a project is thus characterized by a small value of q, indicating the entrepreneur with the project may encounter difficulties in learning about project values from VCs. In the rest of the paper, I refer to q as *VC informativeness*.

To highlight the informational role of VCs, I make two simplifying assumptions. First, I assume all agents are risk-neutral. Second, there is no information asymmetry between the entrepreneur and VCs in the market. In particular, both parties (i) observe all signals and share the same information on project cash flow at any time point t, and (ii) observe the cash flow upon project completion. This assumption helps avoid agency problems associated with adverse selection. Both the entrepreneur and VCs agree with the investment surplus from the project and bargain over it. I start by analyzing the case where the financial market is competitive, and the entrepreneur obtains all investment surplus. I generalize these assumptions in Section 2.4.

Project cash flow (t=3). If the project is abandoned ($I_2 = 0$), the cash flow of the project at date $t = 3$ is zero. Otherwise, if the project is invested ($I_2 = 1$), the cash flow depends on the

[9] While González-Uribe et al. (2022) primarily focuses on the *type improvement* channel, my assumption aligns more closely with *type discovery* during the pre-investment screening stage. However, their findings support the notion of information exchange between VCs and entrepreneurs during this stage.

type of the project. Specifically, the cash flow is $\tilde{R} = R$ if the project is of high type ($T = H$) and $\tilde{R} = 0$ if the project is of low type ($T = L$).

The investment decision (t=2). At date 2, the entrepreneur has two sources of information. First, the entrepreneur receives a publicly available signal $s_e = \{T, \emptyset\}$. In particular, the signal follows a binary distribution. With a probability λ, the signal is of infinite precision and reflects the true type of the project ($s_e = T$). Otherwise, the signal has a precision of 0 and provides no information regarding project type ($s_e = \emptyset$). Second, the entrepreneur observes all signals collected from VCs $s_{vc_i} = \{T, \emptyset\}$, where $i \in \{1, ..., n\}$ and n is the number of VCs contacted by the entrepreneur. Each signal reflects project type $s_{vc_i} = T$ with probability q. $s_{vc_i} = \emptyset$ represents no signal from VC i, which occurs with probability $1 - q$. Hence, the entrepreneur's information set at $t = 2$ is $\mathcal{S} = \{s_e, s_{vc_1}, ..., s_{vc_n}\}$. I assume that the signals within the information set are independent of each other in terms of their precision. The net present value of the entrepreneur's project for given information set $\mathcal{S}$ and investment decision $I_2 \in \{0, 1\}$ is

$$V_2(I_2; \mathcal{S}) = I_2 \mathbb{E}(\tilde{R} - k|\mathcal{S}) = \max\{0, \mathbb{E}(\tilde{R} - k|\mathcal{S})\}. \tag{1}$$

The number of VCs to contact (t=1). At date 1, the entrepreneur chooses the number of VCs to contact, which is denoted as n. This decision affects entrepreneur payoff through changing information set $\mathcal{S}$. Increasing n leads to a high probability that *at least one* signal in $\mathcal{S}$ reflects project type T. I assume each signal comes at a constant cost of c. The cost c captures the entrepreneur's time and effort in interacting with VCs. The entrepreneur chooses n to solve

$$\max_{n \in \mathbb{N}} u(n) = \mathbb{E}(V_2|n) - cn. \tag{2}$$

where $u(n)$ represents utility of the entrepreneur.

Project Initiation (t=0). The entrepreneur initiates the project.

Definition of equilibrium. The equilibrium consists of two strategies by the entrepreneur. (i) At $t = 2$, the entrepreneur makes investment decision I_2 according to information set $\mathcal{S}$. (ii) At $t = 1$, the entrepreneur optimizes the number of VCs to contact (n) to maximize the expected payoff, which is described by equation (2).

2.2 Analysis

In this section, I analyze the equilibrium of the model. I start the analysis from the last decision node at $t = 2$. The entrepreneur makes investment decision I_2 to maximize V_2 as illustrated in equation (1). The entrepreneur's strategy is a function of the information set $\mathcal{S}$, which contains signals revealed at $t = 2$. Notice that

$$\mathbb{E}(\tilde{R} - k|\mathcal{S}) = \begin{cases} R - k & \text{if } \exists s \in \mathcal{S} \text{ such that } s = T \text{ and } T = H \\ -k & \text{if } \exists s \in \mathcal{S} \text{ such that } s = T \text{ and } T = L \\ pR - k & \text{otherwise,} \end{cases}$$

where $R - k > 0$ and $pR - k < 0$ by assumption. Hence, the entrepreneur project has a positive NPV only if (i) at least one signal reviews project type ($\mathcal{S}$ is informative), and (ii) the signal shows that the project is of high type. Consequently, the entrepreneur's optimal strategy follows

$$I_2^*(\mathcal{S}) = \begin{cases} 1 & \text{if } \exists s \in \mathcal{S} \text{ such that } s = T \text{ and } T = H \\ 0 & \text{otherwise.} \end{cases}$$

Notice that in the first-best scenario without information frictions ($\lambda = 1$), i.e., there always exist signals that reveal project type, the entrepreneur is able to efficiently invest in high type projects and abandon low-type ones, resulting in an expected payoff (at $t = 1$) of $p(R - k)$. Let us define $V := p(R - k)$ and discuss how the information frictions prevent the entrepreneur from achieving the first-best payoff. When the entrepreneur relies on signals in $\mathcal{S}$ to make investment decisions, the efficiency of investment is a function of the informativeness of $\mathcal{S}$. I define the informativeness of $\mathcal{S}$, from the point of view of $t = 1$, as the probability of knowing the project type by observing $\mathcal{S}$. In particular, it is denoted as

$$P(n; \lambda, q) := \mathbb{P}(\exists s \in \mathcal{S} \text{ s.t. } s = T) = 1 - (1 - \lambda)(1 - q)^n.$$

P represents the probability that the entrepreneur possesses complete information about the project value and can make efficient investment decisions, investing in the project when it is of high type and abandoning it otherwise. With a probability of $1 - P$, the entrepreneur has no information about the project type, abandons the project, and obtains zero payoff. In sum, the expected project value at $t = 1$ can be written as

$$\mathbb{E}(V_2|n) = PV, \tag{3}$$

and the entrepreneur's utility at $t = 1$ can be written as

$$u(n) = PV - cn. \tag{4}$$

Notice that P increases in n because the probability of having at least one informative signal is larger when more signals are collected. This illustrates the benefit of signal collection: more signals indicate higher informativeness of the information set $\mathcal{S}$, and, thereby, the expected payoff from the project. Meanwhile, each signal is associated with a fixed cost of c. Taken together, the number of VCs to contact affects the entrepreneur's utility through changing informativeness of set $\mathcal{S}$, namely P, and the total cost of signal collection cn.

The entrepreneur trades off the benefit of informativeness and the cost of signals to choose n. The entrepreneur's optimal choice of n^* is formalized by Proposition 1. Without affecting the key intuition of the model, I simplify the discussion by analyzing a case where the entrepreneur maximizes the expected payoff over a continuous set $n \in \mathbb{R}_0^+$.

Proposition 1. *At $t = 1$, the optimal number of signals to collect is given by*

$$n^*(V, q, \lambda, c) = \begin{cases} \frac{log(1-\lambda)\alpha V - \log c}{\alpha} & if\ (1-\lambda)\alpha V \geq c \\ 0 & otherwise, \end{cases} \tag{5}$$

where $\alpha = -log(1 - q)$.

Corollary 1. *There exists q^* such that*

$$\partial n^* / \partial q = \begin{cases} \geq 0 \ \ if\ q \leq q^* \\ < 0 \ \ if\ q > q^*. \end{cases}$$

The value of q^* is given by equation (12) in the Appendix.

The optimal number of VCs to contact is a function of project payoff V, VC informativeness q, the efficiency of alternative information source λ, and the cost c of contacting each VC. We are particularly interested in how VC informativeness q affects the entrepreneur's strategy in contacting VCs. Corollary 1 shows that the equilibrium n^* exhibits an inverse-U relationship with VC informativeness q. This pattern emerges because the marginal value of an additional signal beyond the first one initially rises and then declines in relation to q. When q is relatively low ($q < q^*$), an increase in q signifies higher signal quality, which incentivizes signal collection. Conversely, when q is relatively large ($q > q^*$), the first signal is sufficiently informative to determine the project type. As a result, the additional benefit derived from additional signals diminishes, and the optimal n^* decreases.

Proposition 2. *The probability of achieving efficient investment, which is denoted by $P(n^*)$, increases in VC informativeness (q).*

Proposition 2 concludes the impact of VC informativeness q on the probability of revealing project value after considering the entrepreneur's strategic choice of n^*. Even though the entrepreneur has the choice to collect signals to reduce information friction, she optimally selects the number of signals considering the associated cost. Overall, the probability of revealing project type $P(n^*)$ increases in VC informativeness q.

Recall that the project is invested if and only if two conditions are satisfied: (i) $\mathcal{S}$ is informative, and (ii) the project is of high type. These two events, which occur independently, have probabilities of P and p, respectively. Consequently, the probability of the project being invested is given by pP and increases in q. Corollary 2 summarizes this result.

Corollary 2. *The probability of project investment increases in VC informativeness q.*

From the discussion above, we can conclude that VC informativeness may affect the entrepreneur's utility through the total cost of signals collection (n^*) and the informativeness of information set $\mathcal{S}$. Figure 2 provides a numerical example and visualizes these two channels.

Proposition 3. *The entrepreneur's utility is*

$$u(n^*) = \begin{cases} V - \underbrace{\frac{c}{\alpha}}_{\text{Loss of low informativeness}} - \underbrace{\frac{c}{\alpha} log \frac{(1-\lambda)\alpha V}{c}}_{\text{Cost of signal collection}} & \text{if } (1-\lambda)\alpha V \geq c \\ \lambda V & \text{otherwise,} \end{cases} \tag{6}$$

where $\alpha = -log(1-q)$.

Corollary 3. *The entrepreneur's utility* $u(n^*)$ *(weakly) increases in VC informativeness* q.

Proposition 3 shows the effects of the informativeness of $\mathcal{S}$ and the cost of signal collection on the utility of the entrepreneur. Interestingly, the loss of low informativeness is not a function of the expected payoff V. This is because the entrepreneur strategically chooses the number of VCs to contact according to the project payoff. Overall, the entrepreneur is (weakly) better off when q is large, as concluded by Corollary 3.

In sum, this section presents a framework in which an entrepreneur experiments with a project. The VCs play two roles throughout this process. First, they provide the necessary funding to complete the project. Second, they generate signals that assist the entrepreneur in assessing the quality of the project. The analysis demonstrates that the entrepreneur benefits from the VC's knowledge of the project through two channels: the probability of conducting efficient investment and the cumulative cost of contacting VCs.

2.3 Entrepreneur Project Choice

In this section, I endogenize entrepreneur project choice and illustrate that the benefits of VC informativeness can lead to entrepreneur catering. I assume that the entrepreneur at $t = 0$ can select between one of her ideas to experiment: *exploration* or *catering*. These two project choices are denoted as $I_0 = I_E$ and $I_0 = I_C$ respectively. The two projects have different payoff distributions and degrees of VC informativeness. In particular, I assume that

$$R_E > R_C; q_E < q_C,$$

where R_i represents the payoff conditional on being high type and q_i stands for VC informativeness for project $i \in \{E, C\}$.[10]

Exploration involves unique technologies or projects without substitutes, which, if successful, generates high value.[11] However, due to their distinct nature compared to existing projects, including those in the VCs' portfolios, VCs are unable to provide precise estimations regarding the project type. Conversely, the catering project yields low value when of high type, but VCs possess adequate knowledge to identify the project type accurately. Lemma 1 demonstrates the optimal strategy for the entrepreneur under the first-best scenario without information frictions.

Lemma 1. *When information source beyond VCs is fully informative ($\lambda = 1$), the entrepreneur optimally chooses the exploration project ($I_0^* = I_E$).*

In a frictionless world, the entrepreneur exerts zero effort in the signal collection and obtains a utility of V based on the selected project, as suggested by Proposition 1. The entrepreneur optimally selects exploration as it provides a higher payoff conditional on being high type. When $\lambda < 1$, the entrepreneur encounters information frictions and is motivated to collect signals from VCs. As Proposition 4 shows, the entrepreneur selects exploration only under a certain set of parameter values.

Proposition 4. *There exists $\bar{q_C}(q_E, R_E, R_C, p)$ such that the entrepreneur selects exploration at date $t = 0$ when $q_C < \bar{q_C}(q_E, R_E, R_C, p)$. If $q_C > \bar{q_C}(q_E, R_E, R_C, p)$, the entrepreneur selects exploration when $\lambda > \bar{\lambda}(q_C, q_E, R_C, R_E, p)$ and catering when $\lambda \leq \bar{\lambda}(q_C, q_E, R_C, R_E, p)$.*

Exploration provides a higher expected payoff, but it comes at the cost of low VC informativeness. Conversely, the catering project is featured by a lower expected payoff and high VC informativeness. A smaller value of λ indicates a higher importance placed on acquiring information from VCs, resulting in a larger informational benefit derived from catering to VCs. As a result, the entrepreneur chooses catering over exploration when $\lambda < \bar{\lambda}$. As Figure 3 (a)

[10] I assume an equal probability of being a high-type for both projects. This is equivalent to either assuming that the exploration project is more likely to be high-type ($p_E > p_C$) or that the two projects have the same expected value ($p_E R_E = p_C R_C$) with the exploration project having higher risk ($p_E < p_C$, $R_E > R_C$). Both alternative assumptions also result in a higher value for the exploration project ($V_E > V_C$).

[11] This assumption reflects the idea that differentiated project faces lower competition and generates more profit (Tirole, 1988).

shows a numerical example that catering can provide a higher utility than exploration when λ is sufficiently small.

The value of $\bar{\lambda}$ increases in q_C, which denotes the VC informativeness with the catering project. When q_C is large, catering provides a greater informational benefit, which incentivizes the entrepreneur to opt for the catering project. Hence, a large λ is required to motivate exploration. The discussion of q_C becomes particularly relevant as more VCs adopt data technology for pre-investment screening. Data technology enables VCs to leverage past investments to predict the potential payoff of new investment opportunities. This increased reliance on past information enhances the VC's informativeness when evaluating projects similar to those in their existing portfolio,[12] which corresponds to the catering project in my model. This may result in a larger informational benefit for entrepreneurs to choose the catering projects and thereby encourage catering, as shown by Figure 3 (b).

Note that increasing either λ or q_C can enhance the entrepreneur's utility, but their roles in welfare differ. Specifically, while an efficient alternative information source ($\lambda = 1$) can result in the first-best welfare,[13] enhancing VCs' expertise (a large q_C) can only contribute to achieving the sub-optimal outcome. This is because VCs' expertise in screening encourages startup catering at the expense of project quality.

2.4 Extensions

In the baseline model, I made several simplifying assumptions to highlight the role of VC informativeness. Specifically, I assumed that (i) the entrepreneur randomly contacts VCs for investment and valuation, (ii) all VCs are contacted simultaneously, (iii) all signals from VCs are common knowledge, and (iv) the financial market is competitive. In this section, I discuss model extensions that relax these assumptions. I show that the entrepreneur's payoff-information trade-off is robust to those extensions.

[12]For more details, refer to Bonelli (2023).

[13]This is suggested by Lemma 1.

2.4.1 Private Signals and VCs' Bargaining with the Entrepreneur

In the baseline model, I assume that signals generated by VCs are publicly available and that the financial market is competitive. Section G.1 considers the extension that the entrepreneur might be unable to deliver information from one VC to another in the financial market. Conditional on having a positive NPV, only informed VCs offer investment to the entrepreneur, and they bargain with the entrepreneur to share investment surplus. I show that the positive effect of VC informativeness on entrepreneur payoff is robust to the assumption of private signals. Consequently, the entrepreneur has an incentive to cater to VCs to enjoy informational benefits, the same as the implication of the baseline model.

2.4.2 Heterogeneous VCs

The baseline model describes the case where the entrepreneur randomly meets n VCs from the same probability distribution. I consider an extension that the entrepreneur strategically approaches VCs that are more likely to be informed about the project value. As shown in Section G.2 of the Appendix, the informational benefit of VC informativeness still holds under this setting. Similar to the baseline setting, the entrepreneur optimally caters to VCs when the alternative information source is less effective or the catering project is close to the expertise of average VCs.

2.4.3 An Optimal Stopping Framework

In the baseline model, I assume the entrepreneur contacts multiple VCs simultaneously, and all signals from VCs are observable at the same time. In Section H, I consider a dynamic framework where signals are realized sequentially. The entrepreneur pays a constant cost and receives a signal on project value at each date. The entrepreneur updates her belief on project value according to signals and strategically make continuation decision. The entrepreneur who stops signal collection either receives VC investment or abandons the project, depending on the distribution of posterior belief upon stopping. In this framework, VC informativeness is reflected by signal precision. I show that VC informativeness affects entrepreneur payoff through the probability of achieving efficient investments and the cumulative cost of signal collection, exactly as in the baseline model.

2.5 Empirical Predictions

The model generates empirical predictions regarding the benefit of s initiating projects of high VC informativeness and the determinants of startup catering. First, The model predicts the benefits of VC informativeness as below.

Prediction 1. *Startups with projects close to VCs' expertise (i.e., high VC informativeness) are associated with a higher probability of VC investments.*

Prediction 2. *Startups' effort in signal collection is an inverse U function of its proximity to the expertise of VCs.*

Prediction 1 arises from Corollary 2. Entrepreneurs with projects of high VC informativeness are more likely to identify project value utilizing information from VCs. Consequently, such projects are more likely to be invested conditional on being high-type.

Prediction 2 shows the optimal signal collection strategy illustrated by Corollary 1. The entrepreneur trades off the benefit and cost of signal collection to determine the optimal level of effort. A high VC informativeness indicates a higher value of each signal, which encourages signal collection. However, if the first signal provides sufficient information, the marginal value of subsequent signals diminishes, which discourages signal collection. This leads to an inverted U-shaped relationship between startup effort in signal collection and VC informativeness. If projects on the market are, on average, of high VC informativeness,[14] we expect to observe that entrepreneurs' effort in signal collection decreases in VC informativeness.

Second, The model generates predictions on the cross-sectional variation of startup project choice between exploration and catering, as illustrated below.

Prediction 3. *Startup catering is more prevalent when information source beyond VCs is limited.*

Prediction 4. *Startups are more likely to initiate catering projects when VCs exhibit proficiency in valuing projects close to their expertise.*

Prediction 3 shows that information frictions entrepreneurs encounter outside the financial market are associated with startups catering to VC's expertise. Startups can determine project

[14] This corresponds to the scenario $q > q^*$.

type utilizing information from several resources, including VCs and other exogenous events such as patent approval. These informational sources substitute each other in revealing project type. As shown by Proposition 4, when information outside the financial market is sufficiently informative, entrepreneurs rely less on VC's informational role and hence reduce catering behavior.

Prediction 4 shows that startups select catering when VCs demonstrate expertise in valuing projects close to their domain, i.e., catering projects exhibit high VC informativeness. Startups opt to initiate catering projects to take advantage of VCs' expertise to reduce information friction in valuation and enjoy informational benefits. In the empirical tests, I use VCs' data technology adoption to proxy for their proficiency in value projects close to their expertise. More details are discussed in Section 3.3.3.

3 Empirical Tests

In this section, I present empirical tests corresponding to the predictions outlined in Section 2.5. I employ patent applications as a proxy to identify projects undertaken by startups. This approach considers two key factors. First, VCs predominantly invest in highly innovative industries characterized by startups engaged in extensive R&D activities. For example, in 2021, the software sector and life science sector accounted for 37% and 14% of the overall VC deal count, respectively.[15] Second, startups often have limited tangible assets, making innovation projects an important source of firm value. Using patent applications as a proxy allows me to capture the importance of these innovation-driven projects in the startup ecosystem.

Patent applications provide a standardized format of innovation output through a credible system. Existing literature shows that patent applications play an informational role that boosts knowledge diffusion (Baruffaldi and Simeth, 2020; Kim and Valentine, 2021), licensing transactions (Hegde and Luo, 2018), and inventor reallocation (Melero et al., 2020; Zhao, 2022). I focus on the scenario where patent applications reveal information to external investors, as suggested by Saidi and Žaldokas (2021). In particular, patent applications can serve as a source of information for VCs when estimating the values of technologies developed by startups.

One assumption of this proxy is that startups pursuing VC financing file patent applications

[15] Source: NVCA Yearbook.

for their innovation projects. One might raise concerns about startups being reluctant to file patent applications due to potential risks, such as the leakage of valuable information that could benefit competitors.[16] I argue that this concern is less relevant for startups relying on external financing. The reason is that once a patent application is granted, it significantly enhances the fundraising prospects for the startup. Farre-Mensa et al. (2020) find that the approval of the first patent increases a startup's likelihood of receiving VC investment over the next three years by 47%. Such benefits provide startups with a strong incentive to file patent applications for their innovation projects.

This section is organized as follows. Section 3.1 introduces the primary data source and sample selection. Section 3.2 examines Predictions 1 and 2, which show that startups with projects close to VCs' expertise demonstrate a higher likelihood of securing VC investments and pay lower cost in signal collection. In Section 3.3, I show that startup catering is associated with VC data technology adoption and efficiency of the patent system, as summarized by Prediction 3 and 4.

3.1 Data and Sample Selection

The primary data of patent applications comes from PatentsView. PatentsView is a patent data visualization and analysis platform supported by the United States Patent and Trademark Office (USPTO). It provides comprehensive data on published patent applications filed after 2001.[17] Due to the low coverage of the earliest years, I only include applications filed after 2004 in my analysis.[18] For each patent application, PatentsView gives the filing date, technology classes, and textual information such as the title and the abstract. In addition, PatentsView provides patent assignees, from which I identify startups potentially seeking VC financing.

Since PatentsView does not indicate whether an assignee is a startup, I select the sample following the procedure below. First, I restrict the sample to assignees that are identified as US firms.[19] Second, I focus on firms that file at least one patent application between 2007-2018 but

[16] A startup's disclosures can facilitate rivals' innovation as Kim and Valentine (2021) suggest.

[17] Prior to the enactment of the American Inventors Protection Act (AIPA) in 2000, patent applications are published only when the patents are granted. The rejected applications before 2000 are not available in Patentsview. AIPA requires applications to be published after 18 months after the filing date, whether approved or not.

[18] The sample period starts in 2004 since PatentsView provides a total number of patent applications that is arguably complete compared to the number from USPTO statistics since 2004.

[19] PatentsView provide types of assigness, where *type*= 2 indicates US firms. In addition, I restrict the sample to firms that file the first patent applications with an address in the US.

no applications before this period.[20] Third, I remove firms that are publicly traded within three years since the first patent filing.[21] Fourth, I restrict the sample to firms that submit fewer than five patent applications in their first filing year. In addition, I also remove firms that received VC investments before the first filing year. This selection criteria helps to zero in 62,088 US firms in the main sample. For those sample firms, I use the first-time applications to proxy for the firms' projects and regard the first filing years as the year of project initiation.[22]

To test model predictions, we need to (i) measure the VC informativeness of each startup project and (ii) identify startups that are invested by VCs, both requiring us to match data on VC investments to patent assignees in PatentsView. I obtain data on VC investments and VC-backed startups from Pitchbook. Pitchbook is one of the primary VC data sources available. It provides detailed information on VC deals, such as the investors, invested startups, and deal dates. The database also contains startup information such as locations, industries, current status, etc. I match Pitchbook startups to PatentsView assignees on names and locations. Then, I verify the matching using other information, such as the first application dates and startup founding dates. I further conduct a manual verification of the results when multiple VC-backed firms were matched to a single assignee, as well as the reverse scenario.[23] More details of the matching procedure are provided in the Appendix C. The matching procedure gives 2,679 sample firms invested by VCs between 2007 and 2018. In the following section, I demonstrate how I use the matched sample to construct the measure of VC informativeness.

3.2 VC Informativeness: Measure Construction and Empirical Tests

3.2.1 Measure of VC Informativeness

In this section, I introduce the empirical definition of VC informativeness, a measure at the patent application level that quantifies the degree to which a patent application aligns with VC expertise. Conceptually, a patent application is of high VC informativeness if it is similar to applications

[20] In particular, I remove firms that (i) have any pre-grant application between 2001 and 2006 or (ii) have any granted patents between 1976 and 2006.

[21] I identify public firms with the help of patent-firm linkage provided by Kogan et al. (2017).

[22] 86% of sample firms have only one patent filing in the first filing year.

[23] For example, I check the matching when (i)standardized firm names in two databases are highly similar but not identical, or (ii) an assignee is matched with more than one VC-backed firms in Pitchbook. I validate my matching by checking startup patents provided by the Pitchbook web portal.

developed by VC portfolio firms in the past. The idea is that VCs learn about technology from existing portfolio firm and apply that knowledge in evaluating potential investment opportunities. This assumption is supported by existing literature. For example, Sorensen (2008) shows that VCs learn from past investments and anticipate evaluating future ones. Bonelli (2023) finds that VCs adopt data technology to utilize data from past investments for screening. Consequently, VCs are likely to possess more expertise in valuing projects that resemble their past portfolio firms.

I identify VC portfolio projects following the procedure below. First, I match patent assignees to VC-backed startups in Pitchbook, as described in the previous section and Appendix C. This gives a list of VC-backed startups with patent applications. Second, I identify the VC holding period of each VC-backed startup. For each VC-backed firm, the VC holding period starts from its first VC investment and ends in the year of startup closure or successful exit, defined as IPO or M&A. Third, all patent applications filed by VC-backed startups during the VC holding period are considered applications in the VCs' knowledge set. This reflects the idea that VCs are informed about the value of patent applications filed under their monitoring. In sum, I identify 55,671 patent applications filed by 10,455 startups satisfying the criteria above.[24] This set of applications is referred to as VC's knowledge set in the rest of the paper.

An application filed by a startup is considered high VC informativeness when it is close to at least one of the patent applications in VCs' knowledge set. In particular, to identify VC informativeness of application i filed in year t, I compare it with applications $p_1^{vc}, ..., p_N^{vc}$ filed by VC portfolio firms from $t-3$ to $t-1$, where N is the total number of such applications. First, I compute the cosine similarity between application i and each of applications p_j^{vc} where $j \in \{1, ..., N\}$. I construct the similarity measure using textual information from patent applications, following Kelly et al. (2021). Particularly, two patent applications are considered similar when they contain similar terms and corresponding frequencies.[25]

I obtain textual information from abstracts of patent applications. A patent application

[24] Patent applications after successful exits, such as M&A and IPOs, are excluded from the sample. The sample only contains patent applications filed by VC-backed startups before and during VC holding periods.

[25] Existing literature measures similarity between firms or captures catering behaviors using product similarity (Hoberg and Phillips, 2016), firm descriptions (Bonelli, 2023) or overlap in technology classes of innovation (Wang, 2018). In this paper, I focus on the textual similarity in patent applications, given my focus on the universe of high-tech startups that potentially seek VC financing.

typically contains a title, an abstract, descriptions, and claims. Abstracts are required to provide a summary that "*indicate the technical field to which the invention pertains and shall be drafted in a way which allows the clear understanding of the technical problem, the gist of the solution of that problem through the invention, and the principal use or uses of the invention.*" Also, abstracts shall be concise, "*preferably 50 to 150 words*".[26] Overall, abstracts provide a complete description of the technology developed by a patent application in a standard length.

To capture technology associated with a patent application, I convert the abstract of the patent application into a vector consisting of (i) a list of terms used in the abstract, and (ii) weighted frequencies of those terms. Here, the weight considers the uniqueness of the term in the patent universe, and unique terms are assigned higher weights.[27] I repeat this exercise for all applications i and $j \in \{p_1^{vc}, ..., p_N^{vc}\}$ to obtains a set of vectors $\{V_i, V_{p_1^{vc}}, ..., V_{p_N^{vc}}\}$. The similarity between patent applications i and each of $j \in \{p_1^{vc}, ..., p_N^{vc}\}$ is described by cosine similarities between the vectors

$$\rho_{i,j} = V_i \cdot V_j.$$

Overall, a pair of applications using the same set of terms has a high similarity score, especially when the distribution of frequencies across the terms is closer. This similarity measure maps textual information of a pair of applications to a number between 0 and 1. A similarity of 0 indicates that the two applications have no terms in common, while a similarity of 1 signifies that the applications are identical. VC's informativeness to application i is measured by

$$\textit{VC Informativeness}_i = \max_{j=p_1^{vc},...,p_N^{vc}} \rho_{i,j}. \tag{7}$$

The idea is that application i is familiar to VC if it is close to at least one patent application in VC's knowledge set.

The measure *VC Informativeness* is an empirical measure of parameter q in the model. When the first-time patent application of a startup is more similar to at least some applications in VC's portfolio, VCs are more knowledgeable in determining the value of the startup. Hence, the startup is more likely to learn about project value from interacting with VCs. Figure 4 shows the variation

[26] Source: USPTO 1826: The abstract.

[27] More details about measure construction are shown in Appendix D.

of *VC Informativeness* across patent applications (Figure (a)) and industries (Figure (b)).

3.2.2 Empirical Tests: Benefits from VC Informativeness

This section discusses empirical tests on the benefits of VC informativeness, as shown by Prediction 1 and 2. I first test the relationship between VC informativeness and VC investment by examining the linear model

$$VC\ Investment_{ijt} = \beta_1 VC\ Informativeness_{ijt} + X_{it} + \delta_{jt} + \epsilon_{ijt} \tag{8}$$

where i, j, t index startup, industry, and first filing year, respectively. $VC\ Investment_{ijt}$ is a dummy variable that equals 1 if startup i filed first-time patent application in year t in industry j obtained VC investment in the sample period. $VC\ Informativeness_{ijt}$ measures to what extent the first-time application of startup i is close to the expertise of VCs, as defined by equation (7). X_{it} controls for the quality of firm inventors, measured by the total number of patents granted to inventors of firm i before year t. δ_{jt} includes industry (1-digit Cooperative Patent Classification (CPC) patent classification) and year fixed effects. Standard errors are clustered at the industry level.

Table 3: [INSERT TABLE HERE]

Table 3 presents the results. The coefficient estimate of $VC\ Informativeness_{ijt}$ is positive and statistically significant, suggesting that startups with projects close to VCs' expertise are associated with VC investments. One standard deviation increase in VC informativeness is associated with 1.16% higher probability of getting invested, which represents a 27% increase compared to the sample mean. This result is consistent with Prediction 1 that startups closer to VCs' expertise are more likely to get project value revealed and obtain VC investments.

Next, I provide empirical evidence regarding the correlation between VC informativeness and the total cost in signal collection for startups, suggested by Prediction 2. The proposition shows that the total cost of signal collection is an inverse U function of VC informativeness. If VCs are, on average, informed ($q > q^*$), we would observe a decrease in the signal collection effort in VC informativeness.

I use the time spent on signal collection as a proxy for the total effort in signal collection. Specifically, I focus on a sub-sample of startups that fall into one of two categories: (i) For startups that secure VC financing during the sample period, I calculate the time spent on signal collection from the first filing year to the year of their first VC investment. (ii) For closed startups that were never invested by any VCs, I regard the time between the first patent filing and firm closure as the signal collection time.[28]

I examine the linear model (8) with a dependent variable of signal collection time described above. Results are reported in Table 4. Column (1) and (2) reports regression results using a sub-sample of startups that satisfy either condition (i) or (ii) above. Overall, VC Informativeness is associated with a shorter time for signal collection. Column (3) and (4) shows that for startups that are invested by VCs in the sample period, VC informativeness predicts faster VC investments. When VC informativeness increases from 0 to 1, the first VC investment comes about 0.58 years faster. One standard deviation increase in VC informativeness is associated with about 0.09 years decrease in time till the first VC investment, which represents about 5% decrease compared to the sample mean. Results in Columns (5) and (6) show that uninvested startups with high informativeness close faster. This is consistent with the explanation that entrepreneurs with low-type projects can effectively learn from VCs and abandon the projects when their projects are close to the expertise of VCs. Overall, results support a negative correlation between VC informativeness and signal collection cost, which is consistent with the parameter value $q > q^*$ in the theoretical model.

Table 4: [INSERT TABLE HERE]

Taken together, evidence suggests two benefits of VC informativeness: a higher probability of being invested and a small total cost in signal collection. As the theoretical model shows, such informational benefits may incentivize startups to cater to VCs' expertise. In section 3.3, I introduce the measure of startup catering and present empirical tests on the determinants of startup catering (Prediction 3 and 4).

[28]Pitchbook provides firm closure time. For this analysis, I focus on a subset of PatentsView startups that are matched with closed Pitchbook startups.

3.3 Startup Catering: Measure Construction and Empirical Tests

3.3.1 Measure of Startup Catering

This section introduces the empirical definition of startup catering. Conceptually, catering projects are those that (i) deviate from entrepreneurs' experience and (ii) highly align with VCs' expertise.

I identify whether an application i filed in year t is new to startup entrepreneurs following the procedure below. First, I define startup entrepreneurs as the inventors associated with first-time patent applications by the startup. Second, I define entrepreneurs' knowledge set at time t as all patent applications $p_1, ..., p_M$ filed by entrepreneurs from $t-3$ to $t-1$, where M denotes the total number of such applications.[29] Third, I compute pairwise similarities between application i and each application in the entrepreneurs' knowledge set as described in Section 3.2. The similarities are denoted as ρ_j where $j \in \{p_1, ..., p_M\}$. Application i is considered within entrepreneurs' knowledge if it is similar to at least one patent application in the inventors' knowledge set. In particular, the extent to which entrepreneurs are experienced with application i is defined as

$$\textit{Startup Experience}_i = \max_{j=p_1,...,p_M} \rho_{i,j}.$$

The application i is considered familiar to entrepreneurs if $\textit{Startup Experience}_i$ is sufficiently large.

Considering both *Startup Experience* and *VC Informativeness*, applications with (i) *Startup Experience* below the median and (ii) *VC Informativeness* above the median are classified as catering applications.[30] In sum, 7,669 sample applications satisfy both criteria and are identified as catering applications.[31]

In the following analysis, our focus narrows to a sub-sample of startups deviating from their entrepreneurs' previous experience in their first-time patent applications (*Startup Experience* below median). Within this subset, startups that pivot their innovation toward VCs' expertise are classified as exhibiting catering behavior. Conversely, other startups that deviate both from

[29] Startups whose inventors do not have prior patent applications are excluded from the analysis.

[30] Figure 5 shows that there is no clear correlation between *Startup Experience* and *VC Informativeness*.

[31] Empirical findings are robust to different thresholds. In Appendix Section F.1, I present results where catering applications are defined as those with *Startup Experience* in the lowest quartile and *VC Informativeness* in the highest quartile.

their own experience and VC experience are seen as engaging in exploratory projects.[32] We then investigate whether factors such as alternative information sources (e.g., patent examination speed) and VCs' adoption of data technology are associated with the choice of startup catering over exploration.

3.3.2 Empirical Tests: Patent Examination Speed and Startup Catering

Prediction 3 shows that the effectiveness of alternative information sources can influence startup project choices. In the empirical analysis, I use patent examination as a proxy for an alternative information source and the speed of patent examination to represent the effectiveness of the patent system in providing information. As Farre-Mensa et al. (2016) shows, patent examination results act as an information source for assessing startup quality, and their delay substantially impacts startups' ability to secure external capital.

Patent examination speed captures the time between patent filing and issuance of the examination result, either approval or rejection. Since patent rejection decisions are not observable in PatentsView, I utilize a sub-sample of approved applications to capture patent examination speed. In particular, for each technology class (4-digit CPC classification) c at year t, the speed of patent examination is defined as the fraction of patents granted within three years among all patents in class c and granted in year t. Notice that this empirical measure captures parameter λ in the theoretical model.

Figure 7 illustrates the substantial cross-sectional variance in patent approval speed among various technology classes, both within an industry (Panel A) and across industries (Panel B).[33] The model predicts that inventors who work on technology classes with slower patent approval can be forced to learn about project value from VCs and thereby initiate catering applications. I test the association between patent examination speed and startup catering by estimating the linear model below.

$$Startup\ Catering_{icjt} = \beta_1 Approval\ Speed_{c,t-1} + X_{it} + \delta_{jt} + \alpha_{c,t-1} + \epsilon_{icjt}, \tag{9}$$

[32] In Appendix F.3, I show that informational benefits are robust to this sub-sample of patent applications.

[33] In Appendix A, I provide more figures showing that the variation is robust to alternative definitions of approval speed.

where i represents patent application, c technology class, j industry, and t filing year. $Startup\ Catering_{icjt}$ is a dummy variable that equals one if the application i is defined as a catering application. $Approval\ Speed_{c,t-1}$ is the fraction of patents that are granted within three years in class c at year $t-1$. X_{it} contains firm controls, including the quality of firm inventors measured by the total amount of patents granted to firm inventors. δ_{jt} represents industry and year fixed effects. Considering potential endogeneity issues caused by other technology class-year level omitted variables such as competitions and complexity, I control for two technology-year level characteristics contained by $\alpha_{c,t-1}$: number of patents and the fraction of breakthrough patents.[34]

Panel A of Table 5 illustrates the results. A technology class with faster patent approval speed is associated with fewer catering applications, as predicted by the theoretical model. One standard deviation increase in examination speed is associated with 11.7% decrease in the fraction of catering applications, which represents a 33% decrease compared to the sample mean. This result supports Prediction 3.

Table 5: [INSERT TABLE HERE]

3.3.3 Empirical Tests: VCs' Data Technology Adoption and Startup Catering

In this subsection, I examine whether startup catering is associated with VCs' proficiency in valuing projects close to expertise, in accordance with Prediction 4. Empirically, I capture VCs' proficiency in valuation using VCs' data technology adoption in screening activity. Data technology enables VCs to leverage data from past investments to predict the value of potential deals. Bonelli (2023) shows that data technology enhances VCs' ability to assess projects resembling their prior investments. However, this proficiency in valuation may not extend to startups developing exploratory technologies that differ significantly from those in the VC portfolio. As a result, startups operating in technology classes supported by data-driven VCs may benefit from greater informational advantages when pursuing catering projects.

I construct a technology class-year level variable to capture whether a technology class c is supported by data-driven VCs at year t. As the first step, I identify data-driven VCs following

[34] The list of breakthrough patents is provided by Kelly et al. (2021).

Bonelli (2023). In particular, a VC firm is identified as data-driven starting from the year it hires the first employee to apply data technologies in screening activities. More details are available in Appendix E. Next, I identify patent applications supported by data-driven VCs. A patent application is defined to be supported by data-driven VCs if it is (i) filed by a VC portfolio firm during the VC holding period and (ii) the VC investor is data-driven when the application is filed. Then, I construct a technology class-year level variable $\textit{Invested by data-driven VCs}_{ct}$. This variable takes the value of 1 if at least one patent application filed in technology class c in year t is supported by data-driven VCs. In the final sample, we observe that 11.2% of technology-year observations are labeled as supported by data-driven VCs.

I test the correlation between data technology adoption by VCs and startup catering by examining model

$$\textit{Startup Catering}_{ict} = \beta_1 \textit{Invested by data-driven VCs}_{c,t-1} + X_{it} + \delta_c + \alpha_t + \epsilon_{ict}. \tag{10}$$

$\textit{Invested by data-driven VCs}_{c,t-1}$ is a dummy variable indicating technology classes invested by data-driven VCs, as defined above. X_{it} contains firm controls, including the quality of firm inventors measured by the total number of patents granted to firm inventors before year t. I control for technology class (δ_c) and year (α_t) fixed effects, respectively. Standard errors are clustered at technology class level.

Results are reported in Panel B of Table 5. The estimates coefficient of $\textit{Invested by data-driven VCs}_{ct}$ is positive and significant, which supports the hypothesis that startups initiate more catering applications in technologies classes invested by data-driven VCs. One standard deviation increase in $\textit{Invested by data-driven VCs}_{ct}$ is associated with 1.2% increase in startup catering, which represents 3% increase compared to the sample mean.

3.4 Discussion and Additional Tests

3.4.1 Validation of Model Assumptions

One key assumption of the model is that the entrepreneur chooses between a high payoff, low VC informativeness project (exploration) versus a low payoff, high VC informativeness project (catering). The assumption establishes a payoff-informativeness trade-off faced by the entrepreneur upon project selection. To validate the assumption that catering applications have a low quality on average, I estimate a linear model as below:

$$Quality_{ijt} = \beta_1 Startup\ Catering_{ijt} + X_{it} + \delta_{jt} + \epsilon_{ijt}$$

where i represents patent applications, j industry, and t filing year. $Startup\ Catering_{ijt}$ is a dummy variable that equals one if application i is identified as a catering application, which was described in the previous section. $Quality_{ijt}$ is the quality of application i. I consider two quality measures below. $Patent\ Grant_{ijt}$ is a dummy variable that indicates whether a patent application is granted within three years of application. $Patent\ Citations_{ijt}$ means log citations received in three years since the approval of a patent application. It only applies to applications that are granted.[35] X_{it} represents the quality of inventors. δ_{jt} includes industry fixed effects (1-digit CPC classification) and year fixed effects. Standard errors are clustered at the industry level. Again, I apply this model to a sub-sample of startups that deviate from their past innovation experience (*Startup Experience* below median) to rule out the impact of inventor experience on project quality.

Table 6: [INSERT TABLE HERE]

As shown in column (1)-(3) of Table 6, catering applications are 9.3% less likely to be granted within three years. Given that the unconditional likelihood of approval is 48.2%, this estimate represents a 19.3% decrease in the likelihood of getting the application granted compared to the sample mean. Meanwhile, there is no evidence that catering applications, conditional on approval, are associated with more citations.[36] In summary, the results suggest that catering applications are of relatively low quality.

[35]The model assumes exploration generates a higher economic value in expectation. Since the dollar values of patent applications (or patents) are not observable in my datasets, I use citations to proxy for their value. This exercise is supported by existing literature showing that economic values and patent citations are positively correlated (Hall et al., 2005; Harhoff et al., 1999; Fischer and Leidinger, 2014; Kogan et al., 2017).

[36]In an unreported table, I show that this conclusion is robust to Poisson regression, which produces consistent and reasonably efficient estimate as Cohn et al. (2022) show.

3.4.2 VC Informativeness and Startup Success

In general, VC informativeness reflects the ability of VCs to assess project value accurately and make well-informed investment decisions. When considering startups that received VC investments, it is expected that VC informativeness is positively associated with high project quality. I test this hypothesis by estimating the model

$$\textit{Successful Exit}_{ijt} = \beta_1 \textit{VC Informativeness}_{ijt} + X_{it} + \delta_{jt} + \epsilon_{ijt}$$

where i represents patent applications, j industry, and t filing year. The dependent variable $\textit{Successful Exit}_{ijt}$ is a dummy variable that equals one if the startup assignee of patent application i experienced a successful exit (IPO or M&A) in the sample period. Again, I control for the quality of firm inventors (X_{it}) and industry and year fixed effects (δ_{jt}) to account for potential confounding factors at firm level and industry level. The results in Table A7 in Appendix F.3 show that a change in VC informativeness from 0 to 1 is associated with a 13.9% increase in the probability of successful exits, which is consistent with the hypothesis.

3.4.3 Empirical Measure of VC Informativeness

One concern regarding the VC informativeness measure is that it may capture 'success' rather than 'information.' In other words, some of the current results can be explained by (i) VCs learning from their past successful investments when making investment decisions and (ii) startups selecting projects similar to those of successful VC portfolio firms.

To address this concern, I adopt an alternative measure of VC informativeness that considers only pre-grant patent applications. Specifically, to construct VCs' knowledge set in year t, I identify patent applications filed by VC portfolio firms between years $t-3$ and $t-1$, and exclude those that were granted before or in year t. This alternative measure captures VCs' knowledge from past investments beyond projects that have been confirmed as successful.

I replicate all empirical analyses using this alternative measure and present the results in Appendix F.2. All results are robust to new measures, confirming information's role in VC investment and startup catering.

4 Conclusion

Existing literature shows that entrepreneurship is essentially about experimentation. Following project initiation, startups learn about project value over time and make informed decisions regarding subsequent investments. This paper demonstrates that startups can strategically cater to VCs' expertise in project selection to reduce information frictions in project valuation. The theoretical model shows that catering brings benefits through increasing investment efficiency and reducing the cumulative cost in engaging with VCs for investment and valuation. In equilibrium, startups select catering when alternative information sources are limited or when VCs demonstrate high skills in valuing projects close to their expertise.

I employ textual data from patent applications to capture startup catering and test the model predictions. Evidence suggests that startups with projects close to VCs' expertise are associated with VC financing and a shorter signal collection procedure, which supports the information benefits of catering. Furthermore, I find that startups initiate more catering projects in technology classes characterized by (i) slow patent examination and (ii) financing from data-driven VCs. These findings are consistent with the explanations that catering brings higher informational benefits when alternative information sources are less effective or when VCs adopt data technology to enhance ability screening projects close to their expertise.

References

Baruffaldi, S. H. and M. Simeth (2020). Patents and knowledge diffusion: The effect of early disclosure. *Research Policy 49*(4), 103927.

Bergemann, D. and U. Hege (1998). Venture capital financing, moral hazard, and learning. *Journal of Banking & Finance 22*(6-8), 703–735.

Bergemann, D., U. Hege, and L. Peng (2009). Venture capital and sequential investments.

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

Bond, P., A. Edmans, and I. Goldstein (2012). The real effects of financial markets. *Annual Review of Financial Economics 4*(1), 339–360.

Bonelli, M. (2023). Data-driven investors.

Chemla, G. and K. Tinn (2020). Learning through crowdfunding. *Management Science 66*(5), 1783–1801.

Chemmanur, T. J., K. Krishnan, and D. K. Nandy (2011). How does venture capital financing improve efficiency in private firms? a look beneath the surface. *The Review of Financial Studies 24*(12), 4037–4090.

Cohn, J. B., Z. Liu, and M. I. Wardlaw (2022). Count (and count-like) data in finance. *Journal of Financial Economics 146*(2), 529–551.

Cornelli, F. and O. Yosha (2003). Stage financing and the role of convertible securities. *The Review of Economic Studies 70*(1), 1–32.

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

Farre-Mensa, J., D. Hegde, and A. Ljungqvist (2016). The bright side of patents. Technical report, National Bureau of Economic Research.

Farre-Mensa, J., D. Hegde, and A. Ljungqvist (2020). What is a patent worth? evidence from the us patent "lottery". *The Journal of Finance 75*(2), 639–682.

Fischer, T. and J. Leidinger (2014). Testing patent value indicators on directly observed patent value—an empirical analysis of ocean tomo patent auctions. *Research Policy 43*(3), 519–529.

Foucault, T. and L. Frésard (2019). Corporate strategy, conformism, and the stock market. *The Review of Financial Studies 32*(3), 905–950.

Goldstein, I. (2023). Information in financial markets and its real effects. *Review of Finance 27*(1), 1–32.

Gompers, P. A. (1995). Optimal investment, monitoring, and the staging of venture capital. *The Journal of Finance 50*(5), 1461–1489.

Gonzalez-Uribe, J., R. Klingler-Vidra, S. Wang, and X. Yin (2023). The broader impact of venture capital on innovation: Reducing entrepreneurial capability constraints through due diligence. *Available at SSRN 4516863*.

González-Uribe, R., J. Klingler-Vidra, S. Wang, and X. Yin (2022). The value-add of venture capital due diligence for venture performance.

Gottlieb, J. D., R. R. Townsend, and T. Xu (2022). Does career risk deter potential entrepreneurs? *The Review of Financial Studies 35*(9), 3973–4015.

Hall, B. H. (2002). The financing of research and development. *Oxford Review of Economic Policy 18*(1), 35–51.

Hall, B. H., A. Jaffe, and M. Trajtenberg (2005). Market value and patent citations. *The RAND Journal of Economics*, 16–38.

Hall, B. H. and J. Lerner (2010). The financing of r&d and innovation. In *Handbook of the Economics of Innovation*, Volume 1, pp. 609–639. Elsevier.

Hall, R. E. and S. E. Woodward (2010). The burden of the nondiversifiable risk of entrepreneurship. *American Economic Review 100*(3), 1163–94.

Harhoff, D., F. Narin, F. M. Scherer, and K. Vopel (1999). Citation frequency and the value of patented inventions. *Review of Economics and Statistics 81*(3), 511–515.

Hegde, D. and H. Luo (2018). Patent publication and the market for ideas. *Management Science 64*(2), 652–672.

Hellmann, T. (1998). The allocation of control rights in venture capital contracts. *The RAND Journal of Economics*, 57–76.

Hellmann, T. and M. Puri (2002). Venture capital and the professionalization of start-up firms: Empirical evidence. *The Journal of Finance 57*(1), 169–197.

Hoberg, G. and G. Phillips (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy 124*(5), 1423–1465.

Hombert, J., A. Schoar, D. Sraer, and D. Thesmar (2020). Can unemployment insurance spur entrepreneurial activity? evidence from france. *The Journal of Finance 75*(3), 1247–1285.

Howell, S. T. (2021). Learning from feedback: Evidence from new ventures. *Review of Finance 25*(3), 595–627.

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

Kaplan, S. N. and P. Strömberg (2003). Financial contracting theory meets the real world: An empirical analysis of venture capital contracts. *The Review of Economic Studies 70*(2), 281–315.

Kaplan, S. N. and P. E. Strömberg (2004). Characteristics, contracts, and actions: Evidence from venture capitalist analyses. *The Journal of Finance 59*(5), 2177–2210.

Kelly, B., D. Papanikolaou, A. Seru, and M. Taddy (2021). Measuring technological innovation over the long run. *American Economic Review: Insights 3*(3), 303–20.

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

Kim, J. and K. Valentine (2021). The innovation consequences of mandatory patent disclosures. *Journal of Accounting and Economics 71*(2-3), 101381.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics 132*(2), 665–712.

Kortum, S. and J. Lerner (2001). Does venture capital spur innovation? In *Entrepreneurial inputs and outcomes: New studies of entrepreneurship in the United States*, pp. 1–44. Emerald Group Publishing Limited.

Lerner, J. (1995). Venture capitalists and the oversight of private firms. pp. 267–284.

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

Manso, G. (2011). Motivating innovation. *The Journal of Finance 66*(5), 1823–1860.

Manso, G. (2016). Experimentation and the returns to entrepreneurship. *The Review of Financial Studies 29*(9), 2319–2340.

Melero, E., N. Palomeras, and D. Wehrheim (2020). The effect of patent protection on inventor mobility. *Management Science 66*(12), 5485–5504.

Nanda, R. and M. Rhodes-Kropf (2013). Investment cycles and startup innovation. *Journal of Financial Economics 110*(2), 403–418.

Nanda, R. and M. Rhodes-Kropf (2016). Financing risk and innovation. *Management Science 63*(4), 901–918.

Puterman, M. L. (1994). *Markov decision processes: discrete stochastic dynamic programming.* John Wiley & Sons.

Saidi, F. and A. Žaldokas (2021). How does firms' innovation disclosure affect their banking relationships? *Management Science 67*(2), 742–768.

Samila, S. and O. Sorenson (2011). Venture capital, entrepreneurship, and economic growth. *The Review of Economics and Statistics 93*(1), 338–349.

Scherer, F. M. and D. Harhoff (2000). Technology policy for a world of skew-distributed outcomes. *Research Policy 29*(4-5), 559–566.

Sorensen, M. (2008). Learning by investing: Evidence from venture capital. In *AFA 2008 New Orleans Meetings Paper*.

Tian, X. and T. Y. Wang (2014). Tolerance for failure and corporate innovation. *The Review of Financial Studies 27*(1), 211–255.

Tirole, J. (1988). *The theory of industrial organization*. MIT press.

Wang, X. (2018). Catering innovation: Entrepreneurship and the acquisition market. *Kenan Institute of Private Enterprise Research Paper* (18-27).

Xu, T. (2018). Learning from the crowd: The feedback value of crowdfunding. *Available at SSRN 2637699*.

Yang, M. and Y. Zeng (2019). Financing entrepreneurial production: security design with flexible information acquisition. *The Review of Financial Studies 32*(3), 819–863.

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

Zhao, Y. (2022). Asymmetric information, patent publication, and inventor human capital reallocation.

Figures and Tables

Figure 1: **Timeline of the Model**

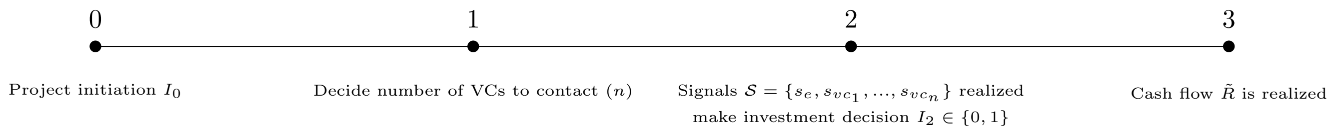
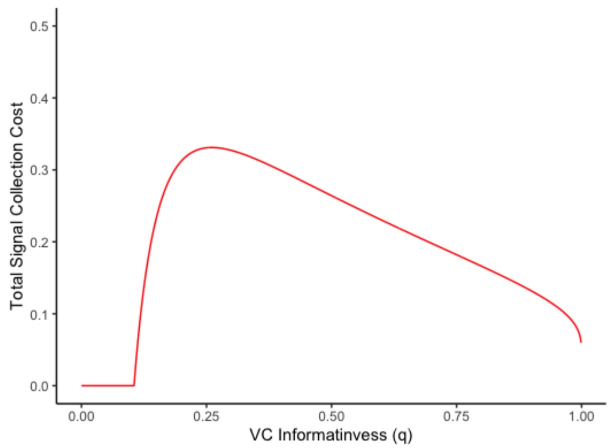
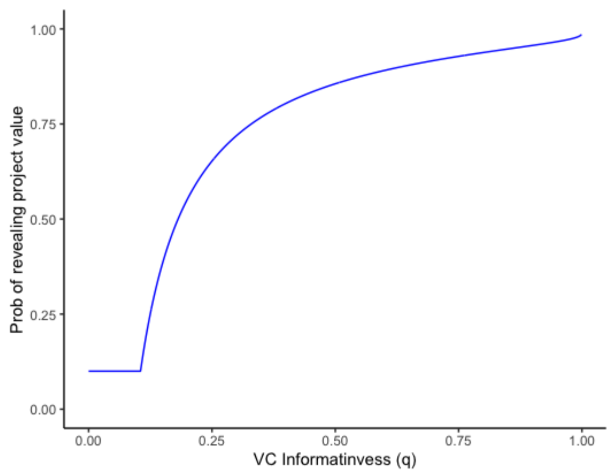


Figure 2: **Effects of VC Informativeness (q) on Entrepreneur's Utility**

(a) Total signal collection cost



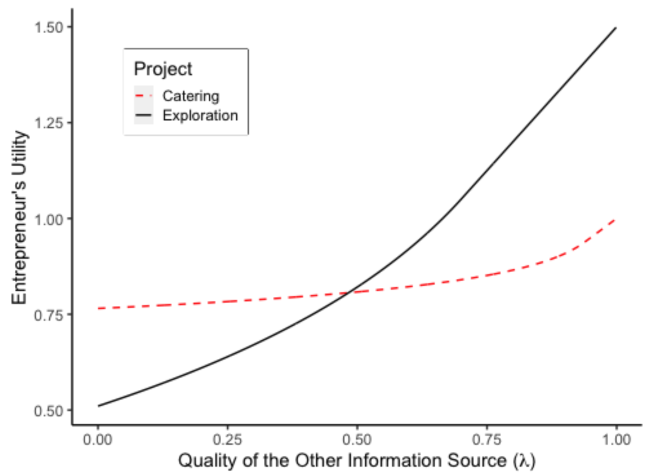
(b) Probability of revealing project value at $t = 2$



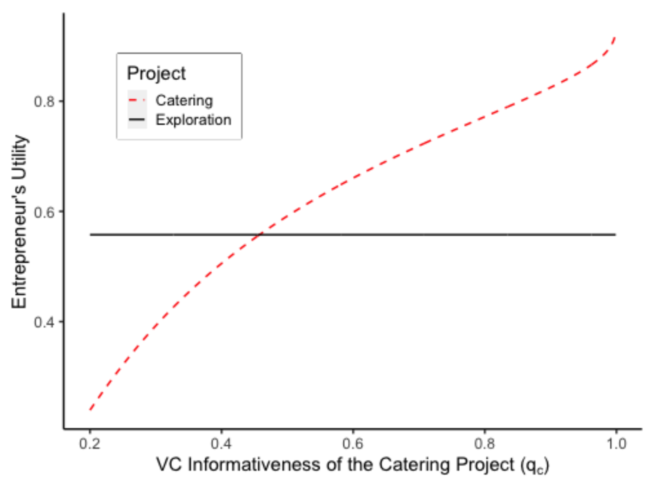
Note: This figure provides a numerical example showing the influences of VC informativeness (q) on the cost of contacting VCs for signals and the probability of revealing project value. Figure (a) shows that the cost of signal collection exhibits an inverse U shape in q. Figure (b) shows that the probability of revealing project value increases in q. I assumed $c = 0.1, \lambda = 0.1, V = p(R - k) = 1$.

Figure 3: **The Entrepreneur's Utility from Exploration and Catering**

(a) Impact of other information source on the entrepreneur's utility



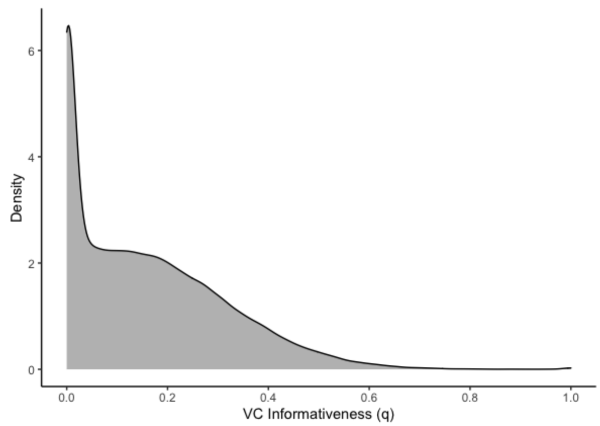
(b) Impact of VC Informativeness of catering on the entrepreneur's utility



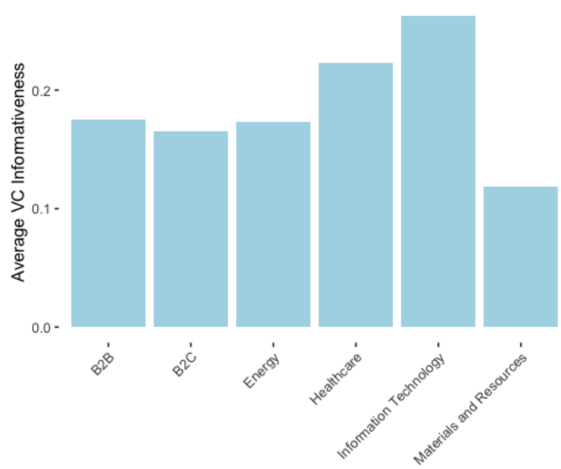
Note: This figure shows a numerical example of the entrepreneur's utilities from exploration (black solid line) and catering (red dashed line). Figure (a) shows that both exploration and catering provide a higher utility when the alternative information source is more informative (λ is large). The utility of exploration increases faster in λ and provides a higher utility when λ is sufficiently large. I assumed $c = 0.1, q_E = 0.2, q_C = 0.8, V_E = p(R_E - k) = 1.5, V_C = p(R_C - k) = 1$. Figure (b) shows that the catering project provides a higher utility when it is closer to VCs' area of expertise (q_c is large). I assumed $c = 0.1, q_E = 0.2, \lambda = 0.1, V_E = p(R_E - k) = 1.5, V_C = p(R_C - k) = 1$.

Figure 4: **Variation in VC Informativeness**

(a) Distribution of VC Informativeness

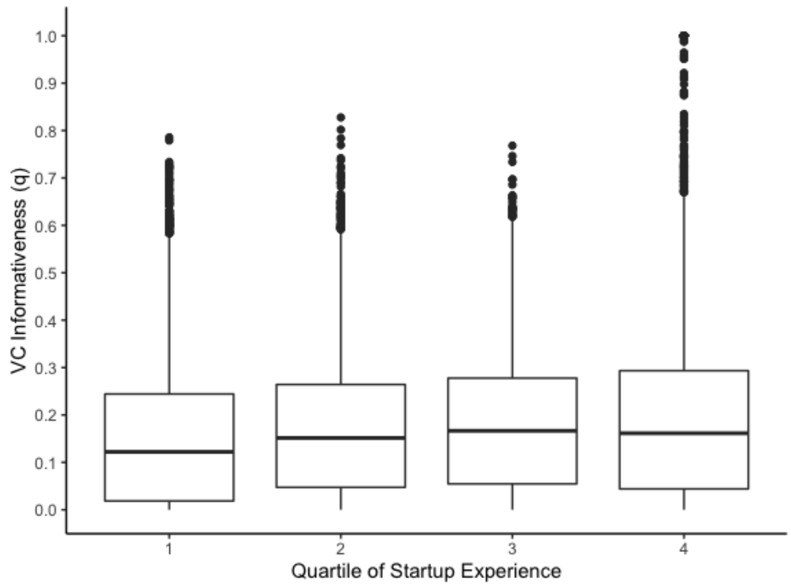


(b) Average VC Informativeness by Industry



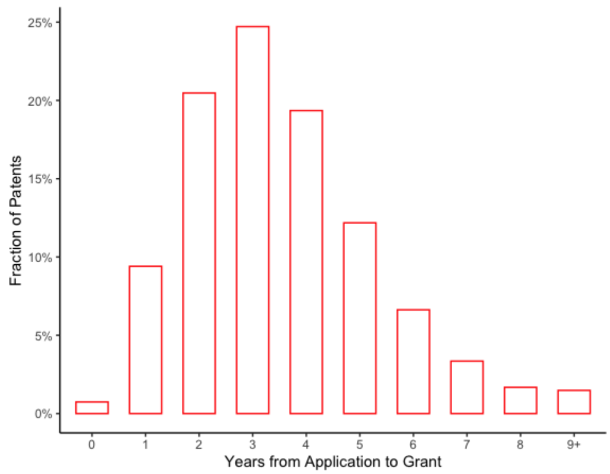
Note: This figure illustrates variations in *VC informativeness*. *VC Informativeness* of each first-time patent application by a startup firm is a number between 0 and 1, as defined by equation (7) in Section 3.2. The measure equals 0 when the patent application is distinct from all patent applications filed by VC portfolio firms, while it equals 1 if the patent application is identical to at least one patent application by VC portfolio firms. Figure (a) shows the distribution of *VC informativeness* across patent applications in the sample. Figure (b) shows average *VC informativeness* by industry.

Figure 5: **Correlation between *Startup Experience* and *VC Informativeness***



Note: This figure illustrates the correlation between *Startup Experience* and *VC Informativeness*. I divide patent applications into four quartiles according to the value of *Startup Experience*. The black dots plot the average of *VC informativeness* for patent applications that fall into the quartile. The gray bars plot the minimum and maximum value of *VC informativeness* for patent applications in the given quartile.

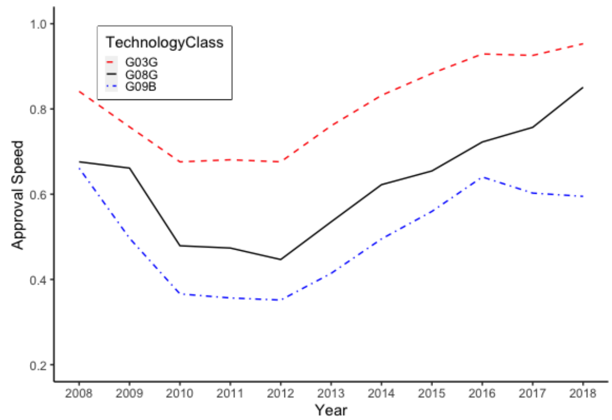
Figure 6: **Distribution of Patent Approval Speed**



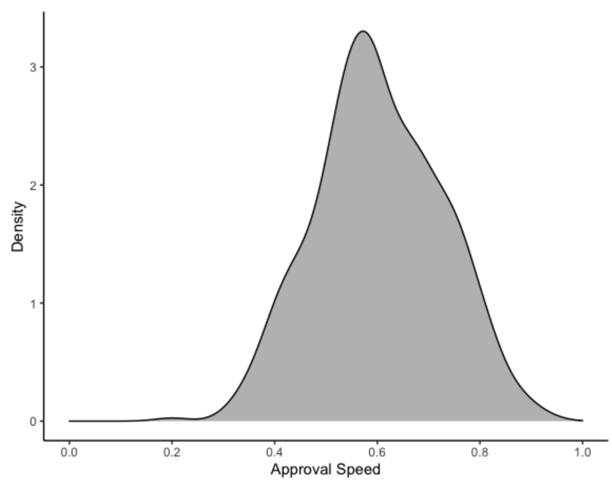
Note: This figure displays the distribution of patent approval speed for USPTO patents granted from 2007 to 2021. The x-axis represents the number of calendar years between patent application and patent grant, while the y-axis shows the fraction of patents in each group. The sample mean is 3.54 years with a standard deviation of 1.82.

Figure 7: **Variation in Speed of Approval across Technology Classes**

(a) Patent Approval Speed: Examples



(b) Distribution of Approval Speed across Technology Classes



Note: This figure illustrates the variation in patent approval speed across technology classes, where approval speed is measured by fraction of patents granted within three years since application filings. Figure (a) plots approval speed of three technology classes (G03G: electrography, electrophotography, magnetography; G08G: traffic control system; G09B: educational or demonstration appliances) in the sample period. Figure (b) plots approval speed of each technology class across the sample period.

Table 1: **Variable Definitions**

Variable	Definition
Patent Grant	A patent application level dummy variable indicating whether the application is approved within three calendar years (=1) or not (=0).
Patent Citations	Three-year forward citations to the granted patent corresponding to a patent application.
VC Informativeness	The extent to which the value of a patent application is informed to VCs. It reflects the proximity between an a patent application and VCs' area of expertise. Detailed definition is introduced in Section 3.2.
Startup Experience	The extent to which a patent application is within expertise of startup inventors. Detailed definition is introduced in Section 3.3.
Startup Catering	A dummy variable indicating whether a patent application is defined as a catering project. In the baseline setting, patent applications with (i) *Startup Experience* below median and (ii) *VC Informativeness* above median are classified as catering projects. See more details in Section 3.3 and Appendix D.
VC Investment	A dummy variable indicating whether the firm is invested by VCs in the sample period (=1) or not (=0).
Time to Investment	The number of years between first patent filing and the first VC investment.
Successful Exit	A dummy variable indicating whether a firm has successfully exited through IPO or M&A (=1) or not (=0).
Approval Speed	A technology class-year level variable that equals to the fraction of patents granted within three years, considering all patents granted in the technological class-year.
Invested by Data-driven VC	A technology class-year level dummy variable indicating whether any patent application in the technology class-year is developed by portfolio firms of data-driven VCs.

Table 2: **Summary Statistics**

	Obs	Mean	St. Dev.	Min	.25	.50	.75	Max
Panel A. First Patent Applications of Startups								
Patent Grant	62,088	0.483	0.500	0	0	0	1	1
Patent Citations	29,960	1.481	5.320	0	0	0	1	246
VC Informativeness	62,088	0.162	0.152	0	0.022	0.132	0.257	1
Startup Experience	32,555	0.263	0.335	0	0.016	0.101	0.383	1
VC Investment	62,088	0.043	0.202	0	0	0	0	1
Panel B. Startups with First Patent Applications Invested by VCs								
Time to Investment	2,643	1.930	1.881	0	1	1	3	14
Successful Exit	2,643	0.174	0.379	0	0	0	0	1
Panel C. Characteristics of Fields of Technological Innovation								
Approval Speed	4,252	0.637	0.189	0	0.500	0.629	0.770	1
Invested by data-driven VC	4,270	0.093	0.290	0	0	0	0	1

Note: This table presents summary statistics of patent applications, startup firms, and fields of technological innovation for a sample of startup firms with the filing year of the first patent applications between 2007 and 2018. Panel A. shows statistics of first-time patent applications filed by startups between 2007 and 2018. Note that *VC Informativeness* is the empirical measure of parameter q in the theoretical model. Panel B. gives statistics of a sub-sample of startups that obtained VC investment between 2007 and 2018 after their first patent filings. *Time to Investment* is the empirical analog of parameter n^* in the model. Panel C. illustrates characteristics of fields of technological innovation (4-digt CPC level) startups are involved in. *Approval Speed* is the empirical measure corresponding to parameter λ, capturing the probability startups learn about project value from the patent system. *Invested by data-driven VC* describes whether the field of technological innovation is supported by VCs that adopt data technology in screening new deals. It is the empirical measure of parameter q_C in the model. Variable definitions can be found in Table 1.

Table 3: **The Role of VC Informativeness in Startup Investment Decisions**

	VC Investment		
	(1)	(2)	(3)
VC Informativeness	0.076***	0.076***	0.075***
	(0.019)	(0.008)	(0.008)
Industry	Yes		
Year	Yes		
Industry-Year		Yes	Yes
Firm Controls			Yes
Observations	61,481	61,481	61,481
R^2	0.011	0.012	0.012

Note: This table shows results for regressions at the firm level, investigating whether startups with patent applications close to VCs' expertise are associated with VC investments. *VC Informativeness* measures the extent to which the first-time application of the startup is close to the expertise of VCs, defined as equation (7). *VC Investment* is a dummy variable that equals one if a startup obtains VC investment in the sample period and zero otherwise. *Industry* denotes 1-digit CPC classification of the first application by the startup. $Year$ represents the first filing year of the startup. *Firm Controls* contains the quality of inventors, measured by total number of patents granted to firm inventors. Standard errors are clustered at the industry level. *** denotes p-value < .01, ** denotes p-value < .05, and * denotes p-value < .1.

Table 4: **The Role of VC Informativeness in Startup Signal Collection**

	All Startups		VC Invested Startups		Failed Startups	
	Time to Investment/Closure		Time to Investment		Time to Closure	
	(1)	(2)	(3)	(4)	(5)	(6)
VC Informativeness	-0.646**	-0.641**	-0.583**	-0.582**	-1.89**	-1.91**
	(0.226)	(0.221)	(0.200)	(0.202)	(0.621)	(0.601)
Industry	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Firm Controls		Yes		Yes		Yes
Observations	2,877	2,877	2,615	2,615	262	262
R^2	0.078	0.078	0.060	0.060	0.434	0.434

Note: This table shows the relationship between *VC informativeness* and startups' efforts in signal collection. The unit of observation is a firm. *VC Informativeness* measures the extent to which the first-time application of the startup is close to the expertise of VCs, defined as equation (7). Effort is proxied by time spent in the signal collection, which is defined as the number of years between first patent filing and first VC investment (*Time to Investment*) or firm closure (*Time to Closure*), depending on startup outcomes. *Industry* denotes 1-digit CPC patent classification. $Year$ represents the first filing year of the startup. *Firm Controls* contains the quality of inventors, measured by total number of patents granted to firm inventors. Standard errors are clustered at the industry level. *** denotes p-value $<$ .01, ** denotes p-value $<$ .05, and * denotes p-value $<$ .1.

Table 5: **Determinants of Startup Catering**

Panel A. Patent Approval Speed by Fields of Technological Innovation

	Startup Catering		
	(1)	(2)	(3)
Approval Speed	-0.680**	-0.650***	-0.648***
	(0.200)	(0.127)	(0.127)
Industry	Yes		
Year	Yes		
Industry-Year		Yes	Yes
Firm Controls			Yes
Technology Controls	Yes	Yes	Yes
Observations	6,494	6,494	6,494
R^2	0.314	0.334	0.335

Panel B. VC's Data Technology by Fields of Technological Innovation

	Startup Catering	
	(1)	(2)
Invested by Data-driven VCs	0.041**	0.041**
	(0.018)	(0.018)
Technology Class	Yes	Yes
Year	Yes	Yes
Firm Controls		Yes
Observations	16,140	16,140
R^2	0.430	0.430

Note: This table shows the determinants of startup catering. Each observation is at the firm level. *Startup Catering* is a dummy variable equal to one if the firm's first-time application (i) deviates from entrepreneur's experience (*Startup Experience* below median) and (ii) aligns closely with VCs' expertise (*VC Informativeness* above median). Panel A. shows the relationship between patent approval speed and startup catering. *Approval Speed* is a technology class-year level variable. It equals the fraction of patents granted within three years since application, considering all patents granted in the technology class one year before application filing. *Technology Controls* includes the total number of patents and the fraction of breakthrough patents at technology class-year. Standard errors are clustered at the industry level. Panel B. shows that data technology adoption by VCs predicts startup catering. *Invested by Data-driven VCs* is a technology class-year level dummy variable equal to one if at least one patent application in the technological class is developed by portfolio firms of data-driven VCs in the past year. *Firm Controls* contains the quality of inventors, measured by the total number of patents granted to firm inventors. Standard errors are clustered at the technology level. *** denotes p-value $< .01$, ** denotes p-value $< .05$, and * denotes p-value $< .1$.

Table 6: **Consequences of Startup Catering**

	Patent Grant			Patent Citations		
	(1)	(2)	(3)	(4)	(5)	(6)
Startup Catering	-0.093***	-0.093***	-0.092***	-0.009	-0.003	-0.003
	(0.013)	(0.008)	(0.008)	(0.025)	(0.015)	(0.015)
Industry	Yes			Yes		
Year	Yes			Yes		
Industry-Year		Yes	Yes		Yes	Yes
Firm Controls			Yes			Yes
Observations	16,140	16,140	16,140	7,349	7,349	7,349
R^2	0.119	0.125	0.126	0.072	0.084	0.085

Note: This table shows the quality (*Patent Grant* and *Patent Citations*) of the catering applications. The unit of observation is a patent application. *Patent Grant* is a dummy variable that indicates whether the application is granted within three years after filing. *Patent Citations* represents the natural logarithm of one plus the count of patent citations over a three-year period for the patent associated with the application, conditional on patent approval. *Startup Catering* is a dummy variable equal to one if the firm's first-time application (i) deviates from entrepreneur's experience (*Startup Experience* below median) and (ii) aligns closely with VCs' expertise (*VC Informativeness* above median). *Firm Controls* contains the quality of inventors, measured by total number of patents granted to firm inventors. Standard errors are clustered at the industry level. *** denotes p-value < .01, ** denotes p-value < .05, and * denotes p-value < .1.

Internet Appendix to
Startup Catering to Venture Capitalists

Xiyue (Ellen) Li

University of British Columbia

December 7, 2023

Table of Contents

Section A. Additional Figures

Section B. Proofs

Section C. Matching Pitchbook with PatentsView

Section D. Empirical Measures

Section E. Identify Technology Classes Invested by Data-driven VCs

Section F. Additional Tables

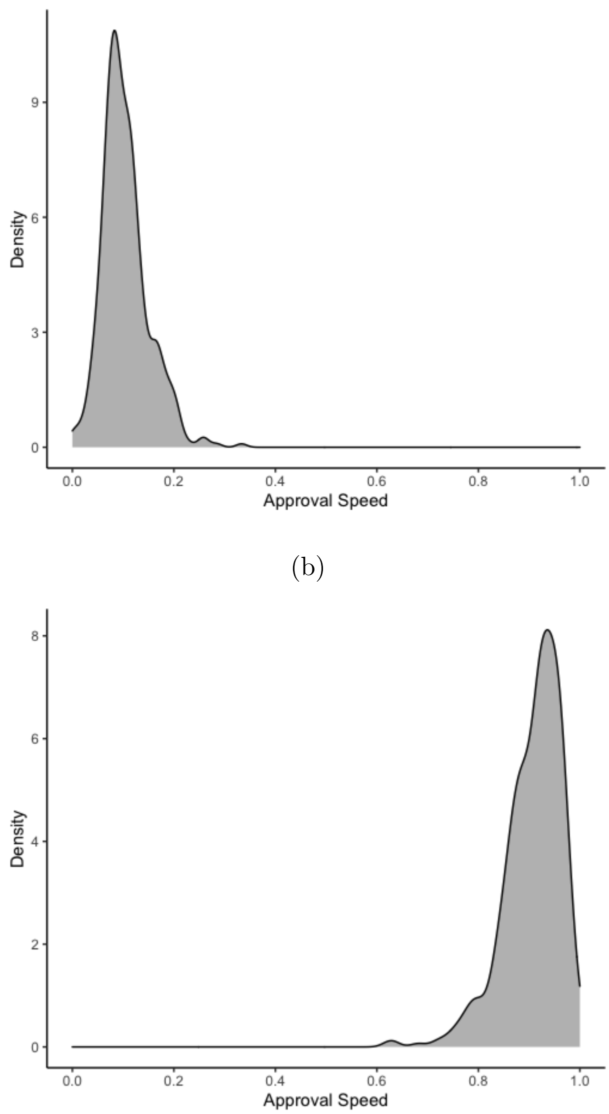
Section G. Model Extensions

Section H. Optimal Stopping Framework

A Additional Figures

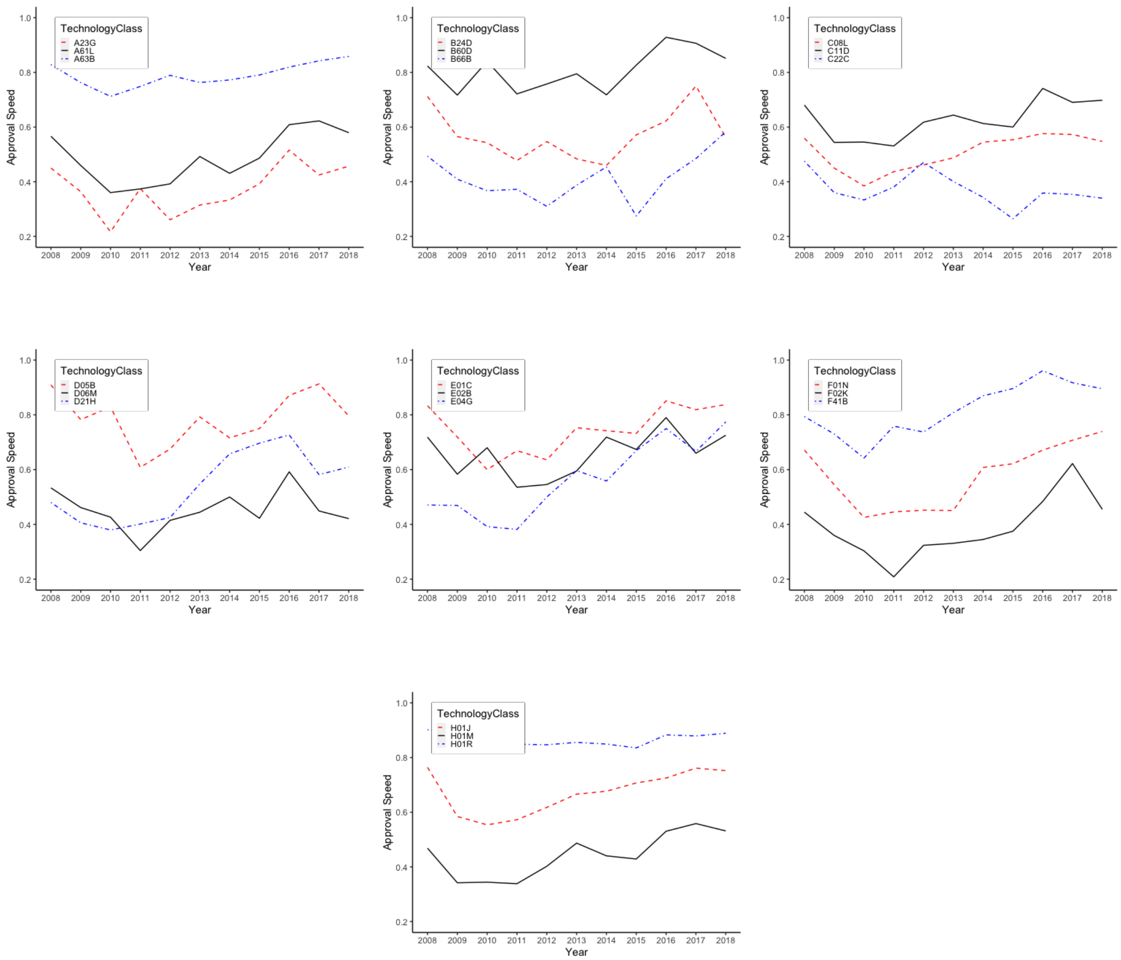
Figure A1: **Distribution of Patent Approval Speed**

(a)



Note: This figure displays the variation in approval speed for each technology class over the sample period. Approval speed is defined as the fraction of patents granted within one year (Figure (a)) and five years (Figure (b)), respectively.

Figure A2: **Variation in Approval Speed Within Industries**



Note: This figure illustrates the variation in patent approval speed across technology classes (4-digit CPC classification) within industries (1-digit CPC classification). Patent approval speed is defined as the fraction of patents granted within three years of filing.

B Proofs

Proof of Proposition 1 and Corollary 1. Notice that

$$\frac{du}{dn} = (1-\lambda)\alpha V(1-q)^n - c,$$

and $\frac{d^2u}{dn^2} < 0$ for any $n \in \mathbb{R}_0^+$, where $\alpha = -log(1-q) > 0$. If $(1-\lambda)\alpha V < c$, we have $\frac{du}{dn} < 0$ for any $n \in \mathbb{R}_0^+$. The optimal number of signal collection is then given by $n^* = 0$. Otherwise, n^* can be determined by first-order condition

$$\frac{du}{dn} = (1-\lambda)\alpha V(1-q)^{n^*} - c = 0. \tag{11}$$

The result is given by equation (5).

The discussion below proves Corollary 1. First, we show that n^* is an inverse-U function of q when $(1-\lambda)\alpha V \geq c \Leftrightarrow q \geq \exp\{-\frac{c}{(1-\lambda)V}\}$ is satisfied. Since

$$\frac{\partial n^*}{\partial q} = \frac{\partial n^*}{\partial \alpha}\frac{d\alpha}{dq} = \frac{1}{\alpha^2}(1 - log\frac{(1-\lambda)\alpha V}{c})\frac{1}{1-q},$$

we have

$$\frac{\partial n^*}{\partial q} = \begin{cases} \geq 0 \text{ if } \exp\{-\frac{c}{(1-\lambda)V}\} \leq q \leq q^* \\ < 0 \text{ if } q > q^* \end{cases}$$

where e represents Euler's number and q^* is given by

$$q^* = 1 - \exp\{-\frac{ec}{(1-\lambda)V}\} \tag{12}$$

Note that when $q < \exp\{-\frac{c}{(1-\lambda)V}\}$, we have $\frac{\partial n^*}{\partial q} = 0$ since $n^* = 0$. Taken together, we show that Corollary 1 holds. □

Proof of Proposition 2. First, we show that P increases in q when $(1-\lambda)\alpha V \geq c \Leftrightarrow q \geq$

$\exp\{-\frac{c}{(1-\lambda)V}\}$ holds. Notice that

$$P(n^*) = 1 - (1-\lambda)(1-q)^{n^*} = 1 - \frac{c}{\alpha V}, \tag{13}$$

where the second equality follows equation (11). Then we have

$$\frac{\partial P(n^*)}{\partial q} = \frac{c}{\alpha^2 V(1-q)} > 0.$$

The minimum value of P is given by

$$P(n^*|q = \exp\{-\frac{c}{(1-\lambda)V}\}) = \lambda.$$

When $(1-\lambda)\alpha V < c$, we have $P = \lambda$. Taken together, we have $P(n^*)$ (weakly) increases in q. Since $\mathbb{E}(V_2|n^*) = P(n^*)V$, we have $\mathbb{E}(V_2|n^*)$ increase in q as well. □

Proof of Proposition 3 and Corollary 3. We can plug equation (13) and equation (5) into equation (4) to get $u(n^*)$. To prove that $u(n^*)$ increases in q, it is sufficient to show that (i) $u(n^*)$ increases in q when q satisfies $(1-\lambda)\alpha V \geq c$; (ii) $u(n^*) = \lambda V$ when q gives $(1-\lambda)\alpha V = c$. Notice that when $(1-\lambda)\alpha V \geq c$ is satisfied, we have

$$\begin{aligned}\frac{\partial u(n^*)}{\partial q} &= \frac{\partial u(n^*)}{\partial \alpha}\frac{d\alpha}{dq} \\ &= [\frac{c}{\alpha^2} + \frac{c}{\alpha^2} log\frac{(1-\lambda)\alpha V}{C} - \frac{c}{\alpha}\frac{1}{\alpha}]\frac{d\alpha}{dq} \\ &= \frac{c}{\alpha^2} log\frac{(1-\lambda)\alpha V}{C}\frac{d\alpha}{dq} > 0.\end{aligned}$$

Hence, $u(n^*)$ increases in q when $(1-\lambda)\alpha V \geq c$ is satisfied. The infimum of $u(n^*) = \lambda V$ is obtained when q gives $(1-\lambda)\alpha V = c$. Therefore, both condition (i) and (ii) are satisfied, and we conclude that $u(n^*)$ weakly increases in q. □

Proof of Proposition 4. The entrepreneur selects the exploration project when $u_C^* < u_E^*$, where u_C^* and u_E^* are defined by equation (6), given the parameter values of two projects respectively. In the analysis below, we define $\alpha_i = -log(1-q_i)$ and $V_i = p(R_i - k)$ for $i \in \{C, E\}$.

First, I show that given $\lambda = 0$, there exists $\bar{q_C}$ such that the entrepreneur selects exploration if and only if $q_C < \bar{q_C}$. Notice that if q_C is sufficiently small such that $\alpha_C V_C < c$, the utility from catering is zero and the entrepreneur selects exploration. A non-trivial case is $\alpha_C V_C \geq c$ and $\alpha_E V_E \geq c$, where the entrepreneur selects exploration when $u_C^* < u_E^*$. There are two cases in total.

Case I. $\lim\limits_{q_C \to 1} u(n^*; q_C, V_C) = V_C \leq u(n^*; q_E, V_E)$. The entrepreneur selects exploration for any $q_C < 1$.

Case II. $\lim\limits_{q_C \to 1} u(n^*; q_C, V_C) = V_C > u(n^*; q_E, V_E)$. Since $u(n^*; q_E, V_C) < u(n^*; q_E, V_E)$ and $u(n^*; q_C, V_C)$ is continuous and increasing in q_C by Corollary 3, there exists $\bar{q_C} \in [q_E, 1]$ such that $u(n^*; q_C, V_C) < u(n^*; q_E, V_E)$ for any $q < \bar{q_C}$.

Taken together, there exists $\bar{q_C}$ such that the entrepreneur always selects exploration when $q_C < \bar{q_C}$.

Second, I restrict the discussion to $q_C \geq \bar{q_C}$, and show that there exists $\bar{\lambda}$ where the entrepreneur selects exploration if and only if $\lambda > \bar{\lambda}$. It is equivalent to prove that $u_C^* - u_E^*$ decreases in λ. Notice that $\alpha_c > \alpha_E$ and $V_C < V_E$ by model assumption. The discussion below contains two parts. First, I show that the conclusion holds when $\alpha_C V_C \geq \alpha_E V_E$. Second, I show that it also holds when $\alpha_C V_C < \alpha_E V_E$.

Case I. $\alpha_C V_C \geq \alpha_E V_E$**.** Notice that it is sufficient to show $u_C^* - u_E^*$ decreases in λ. Since

$$u_C^* - u_E^* = \begin{cases} V_C - \frac{c}{\alpha_C} - \frac{c}{\alpha_C} log \frac{(1-\lambda)\alpha_C V_C}{c} - [V_E - \frac{c}{\alpha_E} - \frac{c}{\alpha_E} log \frac{(1-\lambda)\alpha_E V_E}{c}] & \text{if } \lambda \leq 1 - \frac{c}{\alpha_E V_E} \\ V_C - \frac{c}{\alpha_C} - \frac{c}{\alpha_C} log \frac{(1-\lambda)\alpha_C V_C}{c} - \lambda V_E & \text{if } 1 - \frac{c}{\alpha_E V_E} < \lambda \leq 1 - \frac{c}{\alpha_C V_C} \\ \lambda V_C - \lambda V_E & \text{if } \lambda > 1 - \frac{c}{\alpha_C V_C}, \end{cases}$$

there are three scenarios in total.

- **Scenario i.** $\lambda \leq 1 - \frac{c}{\alpha_E V_E}$. We have

$$\frac{d(u_C^* - u_E^*)}{d\lambda} = (\frac{c}{\alpha_C} - \frac{c}{\alpha_E}) \frac{1}{1-\lambda} \leq 0.$$

- **Scenario ii.** $1 - \frac{c}{\alpha_E V_E} < \lambda \leq 1 - \frac{c}{\alpha_C V_C}$. We have

$$\begin{aligned}\frac{d(u_C^* - u_E^*)}{d\lambda} &= \frac{c}{\alpha_C}\frac{1}{1-\lambda} - V_E \\ &\leq \frac{c}{\alpha_C}\frac{\alpha_C V_C}{c} - V_E = V_C - V_E < 0.\end{aligned}$$

- **Scenario iii.** $\lambda > 1 - \frac{c}{\alpha_C V_C}$. We have

$$\frac{d(u_C^* - u_E^*)}{d\lambda} = V_C - V_E < 0.$$

This concludes that the $u_C^* - u_E^*$ decreases in λ when $\alpha_C V_C \geq \alpha_E V_E$. We next discuss the other case.

Case II. $\alpha_C V_C < \alpha_E V_E$**.** Similarly, we show that $u_C^* - u_E^*$ decreases in λ. Notice that

$$u_C^* - u_E^* = \begin{cases} V_C - \frac{c}{\alpha_C} - \frac{c}{\alpha_C} log\frac{(1-\lambda)\alpha_C V_C}{c} - [V_E - \frac{c}{\alpha_E} - \frac{c}{\alpha_E} log\frac{(1-\lambda)\alpha_E V_E}{c}] & \text{if } \lambda \leq 1 - \frac{c}{\alpha_C V_C} \\ \lambda V_C - [V_E - \frac{c}{\alpha_E} - \frac{c}{\alpha_E} log\frac{(1-\lambda)\alpha_E V_E}{E}] & \text{if } 1 - \frac{c}{\alpha_C V_C} < \lambda \leq 1 - \frac{c}{\alpha_E V_E} \\ \lambda V_C - \lambda V_E & \text{if } \lambda > 1 - \frac{c}{\alpha_E V_E}, \end{cases}$$

We again have three scenarios according to the value of λ above. Notice that the scenario (i) and (iii) satisfy $\frac{d(u_C^* - u_E^*)}{d\lambda} < 0$ as discussed in Case I. In the scenario ii., we have

$$\begin{aligned}\frac{d(u_C^* - u_E^*)}{d\lambda} &= V_C - \frac{c}{\alpha_E}\frac{1}{1-\lambda} \\ &< V_C - \frac{c}{\alpha_E}\frac{\alpha_C V_C}{c} \\ &= V_C(1 - \frac{\alpha_C}{\alpha_E}) < 0.\end{aligned}$$

Hence, the proposition holds for Case II as well. Since $u_C^* - u_E^* \geq 0$ when $\lambda = 0$, $u_C^* - u_E^* < 0$ when $\lambda = 1$ (by Lemma 1), and $u_C^* - u_E^*$ is continuous and decreasing in λ, there exists $\bar{\lambda}$ such that the $u_C^* - u_E^* = 0$. We have $u_C^* < u_E^*$ such that the entrepreneur selects exploration if and only if $\lambda > \bar{\lambda}$. □

C Matching Pitchbook with PatentsView

This appendix provides a detailed procedure for matching VC-backed startups in Pitchbook with patent assignees in PatentsView. The key identifiers are firm names in both databases. Supplementary information, including startup locations and founding years, is utilized to verify the accuracy of the matching results. In particular, I employ the following matching procedure:

Step 1. Selection of patent assignees and VC-backed startups for matching. I restrict sample assignees to be US startups[37] that filed their first patent applications after 2004.[38] For VC-backed startups, I narrow down the sample to include startups founded after 2004 and located in the US.

Step 2. Standardization of firm names and exact matching. I standardize firm names in both datasets by removing common company prefixes and suffixes. Then, I implement exact matching on firm names. This step results in a many-to-many matching between patent assignees and VC-backed startups.

Step 3. Verification. I implement two restrictions to verify the matching results. First, the filing dates of first-time patent applications by an assignee must fall within the timeframe defined by the startup's founding year and its *quit* year. The *quit* year can be either the year of closure or the year of acquisition. Second, the assignee is required to have at least one patent application filed in the same state where the headquarters of the VC-backed startup is situated. Among the matching results that survive the two criteria, I keep the one-to-one matchings and remove corresponding firms from the list that must be matched.

Step 4. Fuzzy matching on firm names. The remaining firms are subjected to a fuzzy matching procedure. Specifically, I match patent assignees with VC-backed startups using the first word of their names, as the first words in firm names are often unique and informative. Following this matching process, I apply the two restrictions outlined in Step 3 to validate these matching results.

Step 5. Manual check. The final step involves a manual verification process. First, I assess the similarity of startup names among the matching pairs generated in Step 3 and eliminate pairings

[37] Assignee *type=2* indicates US firms. Startups are identified following the procedure described in Section 3.

[38] I restrict the sample patent applications from 2004 because of the coverage issue. The number of patent applications covered by PatentsView is comparable to USPTO statistics starting from 2004.

with dissimilar names. Second, I validate the matches using patent data on the Pitchbook web portal. Pitchbook portal provides a comprehensive list of patents granted to each startup. I cross-reference the patents linked to the startups by matching those listed on Pitchbook to identify and remove any incorrect pairings.

I end up with 10,455 VC-backed startups matched with assignees from PatentsView. A total of 55,671 patent applications are identified as VC-backed applications since their filing dates are between the startups' first VC investments and exits.[39]

[39]Startup exists includes closing or successful exits such as IPO and M&A.

D Empirical Measures

This section introduces the methodology for constructing empirical measures using textual information from patent applications. In the first subsection, I introduce a measure named term-frequency-backward-inverse-document-frequency (hereafter *TFBIDF*), which helps to compute the pairwise similarity between patent applications. This measure is constructed following the methodology of Kelly et al. (2021). In the second and third subsections, I discuss the procedure for constructing the measures of *VC informativeness* and *Startup Catering* using pairwise similarities.

D.1 TFBIDF Measure and Pairwise Similarity

Conceptually, the similarity between a pair of patent applications measures the degree of term overlap within the application documents, considering the importance of each term to each application. The importance of a term w to patent application p depends on two factors: (i) The total frequency of term w in the patent application p. A higher frequency of term w implies a higher level of importance. (ii) The total frequency of term w in other patent applications. This criterion captures the informativeness of term w regarding the technology developed by patent application p. The importance of w diminishes if it is commonly adopted in other patent applications. For instance, terms like 'nuclear' may carry more distinctive information compared to generic terms such as 'device,' 'process,' or 'machine.'

The TFBIDF measure is a term-patent application level measure that assesses the importance of a term to a patent application. To be specific, the measured is defined as

$$TFBIDF_{pw} = TF_{pw} \times BIDF_{wt}.$$

The first component is defined as

$$TF_{pwt} = \frac{c_{pw}}{\Sigma_k c_{pk}},$$

where c_{pw} represents the count of term w in patent application p. The component equals the ratio of the count of term w in patent application p to the sum of counts of all terms in application p. It captures the relative importance of term w within application p in terms of frequency. The second component is defined as

$$BIDF_{wt} = log(\frac{\text{\# patent applications prior to } t}{1 + \text{\# patent applications prior to } t \text{ that include term w}}),$$

where t denotes the filing year of the patent application. Notice that $BIDF_{wt}$ decreases in the number of patent applications prior to t that include term w. This component aligns with the second criteria discussed earlier, capturing the informativeness of term w in distinguishing the specific features of application p. When term w is unique to patent application p, it highlights the distinctive characteristics of that application, and $BIDF_{wt}$ has a higher value. Conversely, when the term w is commonly used in other documents, $BIDF_{wt}$ has a smaller value. At any given year t, I considered all patent applications filed through USPTO from $t-3$ to $t-1$ when computing $BIDF_{wt}$.

In summary, the *TFBIDF* describes the importance of a term to a patent application by considering both its importance within the application and its distinctiveness across prior applications. Notice that a vector that contains $TFBIDF_{wp}$ for all terms in a patent application p provides a clear scope of technologies related to p.

Next, I introduce the procedure to compute the similarity between a pair of patent application (i, j) utilizing $TFBIDF$ measure. I start by computing $TFBIDF$ measure for each term w in application i as

$$TFBIDF_{wit} = TF_{wi} \times BIDF_{wt}; t = \max\{\text{filing year for } i, \text{filing yera for } j\}. \quad (14)$$

I arrange $TFBIDT_{wit}$ into a W-Vector, where W represents the size of the set union for terms in pair (i, j). The vector is normalized to a length of 1 following

$$V_{it} = \frac{TFBIDF_{it}}{||TFBIDF_{it}||}.$$

Then, the similarity between i and j is defined as cosine similarity between vector V_{it} and V_{jt}.

In particular, the similarity equals to

$$\rho_{i,j} = V_{it} \cdot V_{jt}.$$

Overall, the measure captures the similarity between terms used in two patent applications, taking their frequencies and uniqueness into consideration. It maps textual information from patent application i and j into a numerical value between 0 and 1, where 0 means the two applications have no common terms and 1 means they are identical.

D.2 Measure of VC informativeness

The *VC informativeness* measure captures whether a patent application i is close to the expertise of VCs. It is constructed by comparing application i with each of the applications in VCs' knowledge set, where VCs' knowledge set consists of applications filed by VC-backed startups during the VC holding period. I compute *VC informativeness* measure following the procedures below.

Step 1. Define the VC holding periods of VC-backed startups. The VC holding period for a startup is defined based on two dates. The start of the holding period is determined by the deal date of the first VC investment in the startup. The end date is determined by either (a) the date of startup closure or (b) the date of successful exits, such as M&A or IPOs.

Step 2. Define VC's knowledge set. Patent applications filed by VC-backed startups during the VC holding periods are regarded as part of the VC's knowledge set. These applications are submitted while the startup is under the guidance and support of VC investors, encompassing the knowledge and expertise of the VCs.

Step 3. Compute similarities between application i and each application in VCs' knowledge set. Assuming application i is filed in year t, I determine patent applications that (i) filed by VC-backed startup from $t-3$ to $t-1$, and (ii) are classified to the same technology class (4-digit CPC classification) as application i. I denote such applications as $p_1^{vc}, ..., p_N^{vc}$, where N represents the total number of patent applications satisfying criteria (i) and (ii). I compute the similarity between each of $p_1^{vc}, ..., p_N^{vc}$ and application i following the procedure described in the last subsection to obtain a set of pairwise similarity $\{\rho_{i,1}^{VC}, ..., \rho_{i,N}^{VC}\}$, where $\rho_{i,j}^{VC}$ denotes the similarity between i and p_j^{vc}.

Step 4. Compute VC Informativeness of application i. The last step is to aggregate $\{\rho_{i,1}^{VC}, ..., \rho_{i,N}^{VC}\}$ to obtain an application-level measure of *VC informativeness.* I define *VC Informativeness* of application i to be

$$\text{\textit{VC Informativeness}}_i = \max\{\rho_{i,1}^{VC}, ..., \rho_{i,N}^{VC}\}.$$

The idea is that application i is close to VCs' expertise if there exists at least one similar patent application in VCs' knowledge set.

D.3 Measure of Startup Catering

In this subsection, I introduce the procedure for identifying catering applications filed by startups. Catering application satisfy two criteria: (i) it has a high *VC informativeness*, which was defined by the previous subsection; (ii) it is beyond startup inventors' innovation experience (hereafter *Startup Experience*). In particular, I use inventors' past patent applications to measure their experience and compute whether an application i is distinct from their past experience. The detailed procedure is listed below.

Step 1. Define startup inventors' knowledge set. I determine all inventors that file at least one patent application in year t, and trace back to their previous patent filing to capture their knowledge set. I define startup inventors' knowledge set as all applications filed by those inventors from $t-3$ to $t-1$. [40] These applications are denoted by $p_1, ..., p_M$, where M is the total number of such applications.

Step 2. Compute pairwise similarities between application i and all applications in the startup's knowledge set. Similarly, I compute the pairwise similarity between application i and each of $p_1, ..., p_M$ as introduced in the first subsection. Those similarities are denoted as $\rho_{i,1}, ..., \rho_{i,M}$.

Step 3. Compute Startup Experience with application i. The inventors are assumed to be equipped with knowledge to develop application i if at least one of their past applications is

[40] As a robustness check, I tried definitions that (i) consider inventors of application i only, or (ii) consider past applications in the same technology class (4-digit CPC) only. My main results are robust to those alternative definitions.

similar to i. Following this idea, I define *Startup Experience* of application i as

$$\textit{Startup Experience}_i = \max\{\rho_{i,1}, ..., \rho_{i,M}\}.$$

E Identify Technology Classes Invested by Data-driven VCs

In this section, I firstly provide a brief introduction to Crunchbase, the main database I rely on to identify data-driven VCs. Then I illustrate the detailed procedures in identifying data-driven VCs and technology classes invested by data-driven VCs.

E.1 Crunchbase Data

Crunchbase data provides comprehensive information regarding the activities of startups and VC firms. It includes detailed records of VC firms' investment histories, encompassing information about the firms they invest in (such as names, addresses, founding dates, industries, current statuses, and more), as well as the dates of investment announcements. This information allows me to identify the holding period of a VC firm for each of its portfolio firms.

In addition, Crunchbase provides information on the hiring activities of VC firms. It includes details about each firm employee, such as their starting date, job title, and job descriptions. This information is useful for tracking changes in the composition of employees within each VC firm, and it serves as a crucial data source for identifying data-driven VCs.

E.2 Identify Data-driven VCs

I identify data-driven VCs following the methodology of Bonelli (2023). Since the adoption of data technology relies on experts with corresponding skills, a VC firm is identified as a data-driven VC starting from the first year it hires an employee that utilizes data technology for pre-investment screening. In particular, I identify data-driven VCs following the procedure below.

Step 1. Determine employee composition of each VC firm. I sort out all organizations that are identified as VC firms in Crunchbase and link them to a list of employees who worked in those firms. Using the starting and end date of each employee in each VC firm, I can determine a VC firm-year level employee composition. I focus on VC firms that (i) are located in the US and (ii) have at least one investment in US startups from 2006-2017 for the analysis.

Step 2. Identify data scientists among employees. I identify data scientists using the job description provided in Crunchbase. In particular, an employee is identified as a data scientist if her job description contains data-related terms, such as 'data engineer', 'machine learning', 'data analytics', etc. I then adopt a few more filling requirements to pin down a list of employees who specifically utilize data skills for pre-stage screening. First, I focus on employees with a job type of "employee" and "executive" and drop those labeled as "advisor", "board member", or "board observer". Second, I manually drop employees who advise portfolio startups by job description. The remaining employees are regarded as data scientists who help with pre-investment screening.

Step 3. Identify data-driven VCs. A VC firm is regarded as data-driven starting from the year with at least one data scientist among the employees.

E.3 Identify Technology Classes Invested by Data-driven VCs

I identify technology classes invested by data-driven VCs following the procedure below.

Step 1. Identify patent applications in the portfolio of each VC firm-year. First, I match VCs to their portfolio firms and identify each startup's corresponding VC holding period. Second, I merge startups in Crunchbase to patent assignees in PatentsView to identify applications filed by VC-backed startups during the VC holding period. A VC firm's portfolio at year t consists of all patent applications filed by its portfolio firms at that year. This procedure is analogous to the matching procedure described in Section C and D.

Step 2. Identify technology classes invested by data-driven VCs. After establishing matches between startup patent applications and VC firms, I further (i) match technology classes to patent applications, and (ii) match data-driven indicator to VCs. This step helps to establish a link between technology classes and features of VCs in terms of the adoption of data technology. If at least one patent application was filed in year t and technology c is within the knowledge set of a VC firm regarded as data-driven before (and including) year t, the technology class is identified as invested by data-driven VCs in that year.

F Additional Tables

F.1 Alternative Definition of Startup Catering

Table A1: **Determinants of Startup Catering**

Panel A. Patent Approval Speed by Fields of Technological Innovation

	Startup Catering		
	(1)	(2)	(3)
Approval Speed	-0.278**	-0.257***	-0.255***
	(0.098)	(0.066)	(0.066)
Industry	Yes		
Year	Yes		
Industry-Year		Yes	Yes
Firm Controls			Yes
Technology Controls	Yes	Yes	Yes
Observations	3,239	3,239	3,239
R^2	0.136	0.154	0.155

Panel B. VC's Data Technology by Fields of Technological Innovation

	Startup Catering	
	(1)	(2)
Invested by data-driven VCs	0.069***	0.069***
	(0.024)	(0.024)
Technology Class	Yes	Yes
Year	Yes	Yes
Firm Controls		Yes
Observations	8,083	8,083
R^2	0.242	0.242

Note: This table shows the determinants of startup catering. Each observation is at the firm level. *Startup Catering* is a dummy variable equal to one if the firm's first-time application (i) deviates from entrepreneur's experience (*Startup Experience* in the bottom quartile) and (ii) aligns closely with VCs' expertise (*VC Informativeness* in the top quartile). Panel A. shows the relationship between patent approval speed and startup catering. *Approval Speed* is a technology class-year level variable. It equals the fraction of patents granted within three years since application, considering all patents granted in the technology class one year before application filing. *Technology Controls* includes the total number of patents and the fraction of breakthrough patents at technology class-year. Standard errors are clustered at industry level. Panel B. shows that data technology adoption by VCs predicts startup catering. *Invested by Data-driven VCs* is a technology class-year level dummy variable equal to one if at least one patent application in the technological class is developed by portfolio firm of data-driven VCs in the past year. *Firm Controls* contains the quality of inventors, measured by total number of patents granted to firm inventors. Standard errors are clustered at technology level. *** denotes p-value $< .01$, ** denotes p-value $< .05$, and * denotes p-value $< .1$.

Table A2: **Consequences of Startup Catering**

	Patent Grant			Patent Citations		
	(1)	(2)	(3)	(4)	(5)	(6)
Startup Catering	-0.073***	-0.073***	-0.073***	0.035	0.043	0.044
	(0.018)	(0.012)	(0.012)	(0.027)	(0.029)	(0.029)
Industry	Yes			Yes		
Year	Yes			Yes		
Industry-Year		Yes	Yes		Yes	Yes
Firm Controls			Yes			Yes
Observations	8,083	8,083	8,083	3,724	3,724	3,724
R^2	0.120	0.130	0.130	0.068	0.086	0.086

Note: This table shows the quality (*Patent Grant* and *Patent Citations*) of the catering applications. The unit of observation is a patent application. *Patent Grant* is a dummy variable that indicates whether the application is granted within three years after filing. *Patent Citations* represents the natural logarithm of one plus the count of patent citations over a three-year period for the patent associated with the application, conditional on patent approval. *Startup Catering* is a dummy variable equal to one if the firm's first-time application (i) deviates from entrepreneur's experience (*Startup Experience* in the bottom quartile) and (ii) aligns closely with VCs' expertise (*VC Informativeness* in the top quartile). *Firm Controls* contains the quality of inventors, measured by total number of patents granted to firm inventors. Standard errors are clustered at the industry level. *** denotes p-value < .01, ** denotes p-value < .05, and * denotes p-value < .1.

F.2 Alternative Definition of VC Informativeness

Table A3: **The Role of VC Informativeness in Startup Investment Decisions**

	VC Investment		
	(1)	(2)	(3)
VC Informativeness	0.075***	0.075***	0.074***
	(0.019)	(0.008)	(0.007)
Industry	Yes		
Year	Yes		
Industry-Year		Yes	Yes
Firm Controls			Yes
Observations	61,480	61,480	61,480
R^2	0.011	0.012	0.012

Note: This table shows results for regressions at the firm level, investigating whether startups with patent applications close to VCs' expertise are associated with VC investments. *VC Informativeness* measures the extent to which the first-time application of the startup is close to the expertise of VCs, defined as equation (7). Note that only pre-grant patent applications are considered when define VCs' knowledge set. *VC Investment* is a dummy variable that equals one if a startup obtains VC investment in the sample period and zero otherwise. *Industry* denotes 1-digit CPC classification of the first application by the startup. $Year$ represents the first filing year of the startup. *Firm Controls* contains the quality of inventors, measured by total number of patents granted to firm inventors. Standard errors are clustered at the industry level. *** denotes p-value < .01, ** denotes p-value < .05, and * denotes p-value < .1.

Table A4: **The Role of VC Informativeness in Startup Signal Collection**

	All Startups		VC Invested Startups		Failed Startups	
	Time to Investment/Closure		Time to Investment		Time to Closure	
	(1)	(2)	(3)	(4)	(5)	(6)
VC Informativeness	-0.634**	-0.630**	-0.580**	-0.580**	-1.81**	-1.82**
	(0.211)	(0.205)	(0.178)	(0.178)	(0.622)	(0.601)
Industry	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Firm Controls		Yes		Yes		Yes
Observations	2,877	2,877	2,615	2,615	262	262
R^2	0.078	0.078	0.060	0.060	0.433	0.433

Note: This table shows the relationship between *VC informativeness* and startups' efforts in signal collection. The unit of observation is a firm. *VC Informativeness* measures the extent to which the first-time application of the startup is close to the expertise of VCs, defined as equation (7). Note that only pre-grant patent applications are considered when define VCs' knowledge set. Effort is proxied by time spent in the signal collection, which is defined as the number of years between first patent filing and first VC investment (*Time to Investment*) or firm closure (*Time to Closure*), depending on startup outcomes. *Industry* denotes 1-digit CPC patent classification. *Year* represents the first filing year of the startup. *Firm Controls* contains the quality of inventors, measured by total number of patents granted to firm inventors. Standard errors are clustered at the industry level. *** denotes p-value < .01, ** denotes p-value < .05, and * denotes p-value < .1.

Table A5: **Consequences of Startup Catering**

	Patent Grant			Patent Citations		
	(1)	(2)	(3)	(4)	(5)	(6)
Startup Catering	-0.096***	-0.096***	-0.095***	-0.010	-0.006	-0.005
	(0.014)	(0.008)	(0.008)	(0.025)	(0.015)	(0.015)
Industry	Yes			Yes		
Year	Yes			Yes		
Industry-Year		Yes	Yes		Yes	Yes
Firm Controls			Yes			Yes
Observations	16,139	16,139	16,139	7,348	7,348	7,348
R^2	0.120	0.126	0.126	0.072	0.084	0.085

Note: This table shows the quality (*Patent Grant* and *Patent Citations*) of the catering applications. The unit of observation is a patent application. *Patent Grant* is a dummy variable that indicates whether the application is granted within three years after filing. *Patent Citations* represents the natural logarithm of one plus the count of patent citations over a three-year period for the patent associated with the application, conditional on patent approval. *Startup Catering* is a dummy variable equal to one if the firm's first-time application (i) deviates from entrepreneur's experience (*Startup Experience* below median) and (ii) aligns closely with VCs' expertise (*VC Informativeness* above median). *Firm Controls* contains the quality of inventors, measured by total number of patents granted to firm inventors. Standard errors are clustered at the industry level. *** denotes p-value < .01, ** denotes p-value < .05, and * denotes p-value < .1.

Table A6: **Determinants of Startup Catering**

Panel A. Patent Approval Speed by Fields of Technological Innovation

	Startup Catering		
	(1)	(2)	(3)
ApprovalSpeed	-0.689**	-0.659***	-0.657***
	(0.201)	(0.128)	(0.128)
Industry	Yes		
Year	Yes		
Industry-Year		Yes	Yes
Firm Controls			Yes
Technology Controls	Yes	Yes	Yes
Observations	6,494	6,494	6,494
R^2	0.316	0.336	0.336

Panel B. VC's Data Technology by Fields of Technological Innovation

	Startup Catering	
	(1)	(2)
Invested by data-driven VCs	0.039**	0.039**
	(0.018)	(0.018)
Technology Class	Yes	Yes
Year	Yes	Yes
Firm Controls		Yes
Observations	16,139	16,139
R^2	0.430	0.430

Note: This table shows the determinants of startup catering. Each observation is at the firm level. *Startup Catering* is a dummy variable equal to one if the firm's first-time application (i) deviates from entrepreneur's experience (*Startup Experience* below median) and (ii) aligns closely with VCs' expertise (*VC Informativeness* above median). Panel A. shows the relationship between patent approval speed and startup catering. *Approval Speed* is a technology class-year level variable. It equals the fraction of patents granted within three years since application, considering all patents granted in the technology class one year before application filing. *Technology Controls* includes the total number of patents and the fraction of breakthrough patents at technology class-year. Standard errors are clustered at industry level. Panel B. shows that data technology adoption by VCs predicts startup catering. *Invested by Data-driven VCs* is a technology class-year level dummy variable equal to one if at least one patent application in the technological class is developed by portfolio firm of data-driven VCs in the past year. *Firm Controls* contains the quality of inventors, measured by total number of patents granted to firm inventors. Standard errors are clustered at technology level. *** denotes p-value $<$ .01, ** denotes p-value $<$ .05, and * denotes p-value $<$ .1.

F.3 Additional Tests

Table A7: **VC Informativeness and Firm Outcomes for Invested Startups**

	Successful Exit		
	(1)	(2)	(3)
VC Informativeness	0.159**	0.139**	0.139**
	(0.063)	(0.065)	(0.065)
Industry	Yes		
Year	Yes		
Industry-Year		Yes	Yes
Firm Controls			Yes
Observations	2,615	2,615	2,615
R^2	0.071	0.103	0.103

Note: This table shows the relationship between VC informativeness and outcome of invested startups. Each observation is at startup level. *VC Informativeness* measures the extent to which the first-time application of the startup is close to the expertise of VCs, defined as equation (7). *Successful Exit* is a dummy variable that equals one if the startup successfully exits (through IPO or M&A) in the sample period. *Firm Controls* contains the quality of inventors, measured by total number of patents granted to firm inventors. Standard errors are clustered at industry level. *** denotes p-value < .01, ** denotes p-value < .05, and * denotes p-value < .1.

Table A8: **The Role of VC Informativeness in Startup Investment Decisions**

	VC Investment		
	(1)	(2)	(3)
VC Informativeness	0.082**	0.082***	0.082***
	(0.027)	(0.017)	(0.017)
Industry	Yes		
Year	Yes		
Industry-Year		Yes	Yes
Firm Controls			Yes
Observations	16,140	16,140	16,140
R^2	0.011	0.015	0.015

Note: This table shows results for regressions at the firm level, investigating whether startups with patent applications close to VCs' expertise are associated with VC investments. This table focus on a sub-sample of startups whose first patent application deviates from past experience of firm inventors. *VC Informativeness* measures the extent to which the first-time application of the startup is close to the expertise of VCs, defined as equation (7). Note that only pre-grant patent applications are considered when define VCs' knowledge set. *VC Investment* is a dummy variable that equals one if a startup obtains VC investment in the sample period and zero otherwise. *Industry* denotes 1-digit CPC classification of the first application by the startup. $Year$ represents the first filing year of the startup. *Firm Controls* contains the quality of inventors, measured by total number of patents granted to firm inventors. Standard errors are clustered at the industry level. *** denotes p-value $<$ .01, ** denotes p-value $<$ .05, and * denotes p-value $<$ .1.

G Model Extensions

This section provides more details of model extensions described in Section 2.4.

G.1 Private Signals and VCs' Bargaining with the Entrepreneur

In this section, I extend the baseline model by assuming that (i) signals from VCs are not publicly available and (ii) informed VCs bargain with the entrepreneur to share investment surplus. The entrepreneur can obtain a $w(x) \in [0, 1]$ fraction of investment surplus when x VCs are informed that the project is of good type. The fraction of surplus w concavely increases in x and $\lim_{x\to\infty} w(x) = 1$.

To simplify, I assume w is determined by Shapley value. In particular, Shapley value reflects the notion that each player's payoff depends on the player's marginal contribution to the total payoff. Consider the set of all players denoted as N. Let C be a subset of players from the set of all players engaged in bargaining. We use $\Pi(C)$ to denote total payoff that can be obtained by the players in C if they cooperate. Then the Shapley value of player i is determined by

$$w_i = \sum_{C \subseteq N-i} \frac{|C|!(|N| - |C| - 1)!}{|N|!} (\Pi(C \cup i) - \Pi(C)).$$

According to the definition of Shapley value, we conclude that the entrepreneur obtains $w(x) = \frac{x}{x+1}$ fraction of investment surplus when there exists x informed VCs.

Now we conclude the entrepreneur's utility function. There are two cases in total. If the alternative information source, which provides project value with probability λ, generates a signal on project value, all VCs on the market are informed about project type and the entrepreneur obtains all investment surplus if there is any ($w = 1$). Otherwise, the entrepreneur relies on signals from VCs and bargain with informed VCs over the investment surplus. The probability that x out of n VCs are informed is

$$f(x, n) = \frac{n!}{(n-x)!x!} q^x (1-q)^{n-x}.$$

The entrepreneur's utility function can be written as

$$u(n;q) = [\lambda + (1-\lambda)w_e(n;q)]V - cn,$$

where $V = p(R-k)$ denotes the first-best expected payoff and

$$w_e(n;q) = \mathbb{E}[w(x)|n] = \sum_{i=1}^{n} f(x;n)w(x)$$

represents expected fraction of investment surplus the entrepreneur can obtain when n signals are collected.

In the discussion below, I show that the main results of the model (summarized in Corollary A1) survives.

Corollary A1. *a. All else equal, the entrepreneur's utility $u(n^*)$ increases in q, where n^* represents the optimal effort on signal collection determined by the entrepreneur.*
b. The startup initiates the exploratory projects when q_C is sufficiently small or λ is sufficiently large.

Proof. **Proof of part a.**
First, I show that how the optimal number of signal collection changes with q. The entrepreneur collects one additional signal only when the marginal benefit of a new signal outweighs the cost of c. I show that the marginal value of a signal decreases. Notice that the marginal benefit of signal $n+1$ is the product of (i) the probability that the signal is informative and (ii) the marginal increases of investment surplus brought by the signal. In particular, it can be written as

$$(1-\lambda)Vq\sum_{x=1}^{n} f(x;n)(\frac{x+1}{x+2} - \frac{x}{x+1}).$$

In order to show that the marginal value of signals decreases, it is sufficient to show

$$g(n) := \sum_{x=1}^{n} f(x;n)(\frac{x+1}{x+2} - \frac{x}{x+1})$$

decreases. Notice that

$$
\begin{aligned}
g(n+1)-g(n)=&\sum_{x=1}^{n+1} f(x;n+1)(\frac{x+1}{x+2}-\frac{x}{x+1})-\sum_{x=1}^{n} f(x;n)(\frac{x+1}{x+2}-\frac{x}{x+1})\\
=&f(n+1;n+1)(\frac{n+2}{n+3}-\frac{n+1}{n+2})\\
&+\sum_{x=1}^{n} f(x;n+1)(\frac{x+1}{x+2}-\frac{x}{x+1})-\sum_{x=1}^{n} f(x;n)(\frac{x+1}{x+2}-\frac{x}{x+1})\\
=&q^{n+1}(\frac{n+2}{n+3}-\frac{n+1}{n+2})\\
&+(1-q)\sum_{x=1}^{n} f(x;n)(\frac{x+1}{x+2}-\frac{x}{x+1})-\sum_{x=1}^{n} f(x;n)(\frac{x+1}{x+2}-\frac{x}{x+1})\\
=&q^{n+1}(\frac{n+2}{n+3}-\frac{n+1}{n+2})-q\sum_{x=1}^{n} f(x;n)(\frac{x+1}{x+2}-\frac{x}{x+1})\\
=&q^{n+1}(\frac{n+2}{n+3}-\frac{n+1}{n+2})-qf(n;n)(\frac{n+1}{n+2}-\frac{n}{n+1})-q\sum_{x=1}^{n-1} f(x;n)(\frac{x+1}{x+2}-\frac{x}{x+1})\\
=&q^{n+1}(\frac{n+2}{n+3}-\frac{n+1}{n+2})-q^{n+1}(\frac{n+1}{n+2}-\frac{n}{n+1})-q\sum_{x=1}^{n-1} f(x;n)(\frac{x+1}{x+2}-\frac{x}{x+1})\leq 0,
\end{aligned}
$$

we have $g(n)$ decreases and thereby the marginal value of signals decreases. Hence, we conclude that the entrepreneur's strategy on signal collect follows

(i) if marginal value of the first signals $(1-\lambda)Vq \leq c$, the entrepreneur optimally choose n^* to maximize utility;

(ii) otherwise, there exists n_0 such that $(1-\lambda)Vqg(n_0-1) > c \geq (1-\lambda)Vqg(n_0)$, the entrepreneur selects $n^* \in \{n_0-1, n_0\}$ that gives a higher utility.

Since $g(n)$ increases in q, we have n^* (weakly) increases in q. Last, we show that the utility of the entrepreneur increases in q. For any given q_1, q_2 such that $0 < q_1 < q_2 < 1$, we have

$$
\begin{aligned}
u(n^*(q_2);q_2)-u(n^*(q_1);q_1) &\geq u(n^*(q_1);q_2)-u(n^*(q_1);q_1)\\
&=(1-\lambda)V[w_e(n^*(q_1);q_2)-w_e(n^*(q_1);q_1)] > 0.
\end{aligned}
$$

Proof of part b.

Follow the conclusion in part a., it is obvious that the entrepreneur is more likely to initiate the catering project when q_C is large. Next, we show that the entrepreneur chooses exploration when λ is sufficiently large. It is sufficient to show that there exists λ^* such that the entrepreneur chooses exploration when $\lambda > \lambda^*$. I show that $\lambda^* = \max\{0, 1 - \frac{c}{q_C V_C}, 1 - \frac{c}{q_E V_E}\}$ satisfies the condition. According to analysis in part a., the entrepreneur selects n^* no matter which project is initiated. Then the utilities of exploration and catering are λV_E and λV_C respectively. Exploration gives a higher utility by the assumption that $V_E > V_C$.

□

G.2 Heterogeneous VCs

In this section, I extend the baseline model by assuming that the entrepreneur can directly search towards the VCs that are more likely to be informed about the project value. In particular, I assume that VC i has a probability of $q_i = q + \delta_i$ to provide a precise signal about project value. q_i of each VC is known by the entrepreneur. The idea is that the entrepreneur can infer each VC's expertise by observing their past investment experience. When the entrepreneur collects n signals, the probability of revealing project value is

$$P = 1 - (1 - \lambda)\Pi_{i=1}^{n}(1 - q_i)$$

The entrepreneur's utility function can be written as

$$u(n) = P(n)V - cn.$$

The proof below shows that Corollary A1 holds under this set of assumptions.

Proof. **Proof of part a.**

For any q_1, q_2 such that $0 < q_1 < q_2 < 1$, we have

$$\begin{aligned} u(n^*(q_2), q_2) - u(n^*(q_1), q_1) \geq & u(n^*(q_1), q_2) - u(n^*(q_1), q_1) \\ = & [P(n^*(q_1), q_2) - P(n^*(q_1), q_1)] > 0, \end{aligned}$$

where the inequality is given by P increases in q for any given n.

Proof of part b. We prove part b. following three steps. First, I show that the first signal provides a highest marginal benefit. The entrepreneur selects $n^* = 0$ if the marginal benefit of the first signal is smaller than c. Second, I show that the entrepreneur selects $n^* = 0$ when λ is sufficiently large. Third, I show that when λ is large enough to give $n^* = 0$ for both project choices, the entrepreneur selects exploration in equilibrium.

Step 1. In the discussion below, I show that the first signal provides a highest marginal benefit for any given project. Notice that the marginal benefit of signal $n + 1$ equals

$$[P(n+1) - P(n)]V = (1-\lambda)Vq_{n+1}\Pi_{i=1}^{n}(1-q_i)$$

The marginal value of a signal is larger when q_{n+1} is larger and $\Pi_{i=1}^{n}(1-q_i)$ is larger. Comparing marginal value of signal n and $n+1$, we have $q_n \geq q_{n+1}$ (since the entrepreneur contacts more informed VCs first) and $\Pi_{i=1}^{n} - 1(1-q_i) \geq \Pi_{i=1}^{n}(1-q_i)$. Hence, the marginal value of signal n is larger than that of signal $n+1$. The marginal value of signals decreases. The first signal provides largest marginal benefit that equals $(1-\lambda)Vq_1$.

Step 2. Next, I show that the entrepreneur selects $n^* = 0$ when λ is sufficiently large. The entrepreneur compares the marginal benefit of a signal with its cost (c) to determine the optimal number of signals to collect. When $(1-\lambda)Vq_1 < c$, the optimal choice is n^* for the entrepreneur.

Step 3. I show that the entrepreneur selects exploration when λ is sufficiently large. It is equivalent to show that there exists λ^* such that the entrepreneur selects exploration when $\lambda > \lambda^*$. I argue that $\lambda^* = \max\{0, 1 - \frac{c}{q_C V_C}, 1 - \frac{c}{q_E V_E}\}$ satisfies the condition. When $\lambda > \lambda^*$, the entrepreneur does not contact any VCs for signals no matter which project is initiated. Hence, the entrepreneur's utility from exploration and catering are λV_E and λV_C respectively. Exploration gives a higher utility given the assumption that $V_E > V_C$. $\square$

H Optimal Stopping Framework

H.1 Setup

Consider an economy with infinite periods. There is a skillful startup that is financially constrained. At date $t = 0$, the startup initiates a project. The startup requires a one-time investment of k from the market to complete the project. Once invested, the project can be immediately completed, generating a verifiable cash flow of θ. My key assumption is that θ is unobservable before project completion. The startup has the prior belief that $\log(\theta) \sim \mathcal{N}(\mu_0, \sigma_0^2)$. At each date $t \geq 1$, the startup collects a signal x_t on cash flow θ and update beliefs correspondingly. According to the posterior belief, the startup decides whether to stop collecting signals and raise capital from the market. Figure A3 shows the timeline of the model.

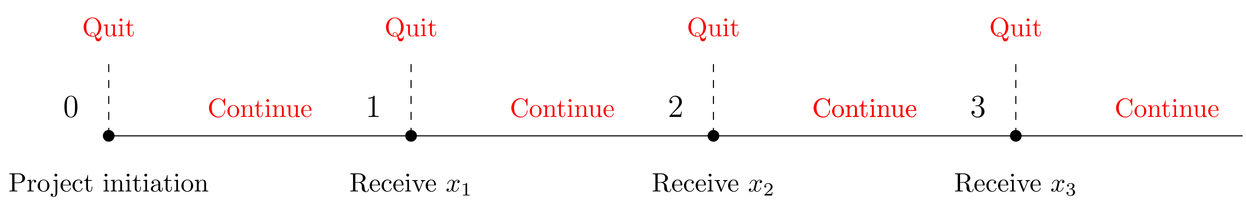


Figure A3: **Timeline of the Model**

The startup pays a fixed cost of $c > 0$ to collect one signal. The cost represents startup entrepreneurs' effort in interacting with VCs to collect signals.[41] At any time point t, the startup has a posterior belief that can be described by an element of state space $\mathcal{S}$, where $s = (\mu, \sigma^2) \in \mathcal{S}$ represents a Bayesian distribution of log-return $\log(\theta)|(\mu, \sigma^2) \sim \mathcal{N}(\mu, \sigma^2)$. According to s, the startup selects an action from the set $\mathcal{A} = \{C, Q\}$, where C and Q represent *continue* to collect signals and *quit* with existing signals, respectively. When taking action C, the startup pays a cost of c, receives another signal, and transit to another state $\tilde{S} \in \mathcal{S}$. After the action Q, the startup raises capital, completes the project if invested, and obtains a payoff of $g^*(s)$. I assume the startup is risk-neutral and does not discount future cash flow. In sum, the startup faces the

[41] In this extension, I assume VCs are the only information source and highlight the channels through which information from VCs benefits the startup. This assumption is equivalent to assuming $\lambda = 0$ in the baseline model.

optimal stopping problem

$$\mathcal{V}^d(s) := \mathbb{E}^d\left\{\sum_{t=0}^{\tau} g(S_t, A_t)\middle| S_0 = s\right\}, \tag{15}$$

where

$$g(s,a) = \begin{cases} g^\star(s), & s \in \mathcal{S},\ a = Q; \\ -c, & s \in \mathcal{S},\ a = C \end{cases} \tag{16}$$

denotes startup payoff, $d : \mathcal{S} \to \mathcal{A}$ characterizes the policy and $\tau = \inf\{t \geq 0 : a_t = Q\}$. The startup targets to find an optimal policy d^* such that $\mathcal{V}^{d^*}(s) \geq \mathcal{V}^d(s)$ for every initial state $s \in \mathcal{S}$ and every policy d. For ease of notation, we denote the optimal value function as $\mathcal{V}(s) = \mathcal{V}^{d^*}(s)$. Notice that $\mathcal{V}(s)$ can be characterized by a Bellman equation

$$\mathcal{V}(s) = \max\{g^*(s), -c + \mathbb{E}_s[\mathcal{V}(\tilde{S})]\}, \tag{17}$$

where $\mathbb{E}_s(\cdot) = \mathbb{E}(\cdot|S = s)$.[42]

The key friction of the model, namely *VC informativeness*, comes from the assumption that each signal received by the startup can be either precise or not, with a certain probability. In particular, I assume that the signal on project values satisfies $\log(x_t) = \log(\theta) + Z_t\epsilon_t$, where $\epsilon_t \sim \mathcal{N}(0, \nu^2)$ is a noise term. Z_t is a random variable that follows Bernoulli distribution, which equals 0 with a probability of $\alpha \in [0, 1]$ and 1 otherwise. The parameter α describes the probability that the signal is precise. A high α indicates less information frictions faced by the startup. I assume that Z_t and ϵ_t are independent, and both of them are independent across different time point t.

Parameter α captures the degree of VC informativeness. A high α indicates a high quality of VC's feedback, which means a high degree of VC informativeness. One example is to interpret α as the fraction of VCs that familiar with the startup's project. Imagine that there are two types of VCs in the financial market. Only α fraction of VCs are experienced in investing similar

[42]In Appendix, I show that the solution to this Bellman equation is the solution to the startup's payoff maximization problem shown by equation (15).

projects and thereby are able to provide precise signals on project cash flow. The other VCs can only observe a noisy signal. If the startup can interact with one random VC at each date to collect a signal, the signal is precise with a probability α. I assume the startup can distinguish precise signals from noisy ones, i.e. the realization of Z_t is observable to the startup. This assumption reflects that the startup knows which VCs are the source of a specific signal and whether those VCs are knowledgeable enough to value the project precisely.

The parameter α plays a role by affecting the transition probability between states. Notice that the transition distribution from S to $\tilde{S}$ conditional on (S, A, Z) is given by

$$\widetilde{S}|(S, A, Z) = (\widetilde{\mu}, \widetilde{\sigma}^2) : \begin{array}{l} \widetilde{\mu} \sim \mathcal{N}\left(\mu, \sigma^2(1 - Z) + \dfrac{\sigma^4 Z}{\sigma^2 + \nu^2}\right), \\ \widetilde{\sigma} = \sqrt{\dfrac{Z}{\sigma^{-2} + \nu^{-2}}}. \end{array} \tag{18}$$

A high α indicates a high probability of achieving the state of $\sigma = 0$, which indicates resolving uncertainty on project value.

I assume there is no information asymmetry between the startup and the market. In particular, both parties (i) share the same information on project cash flow θ at any time point t, and (ii) observe the cash flow θ upon project completion. This assumption makes security design irrelevant. Both the startup and the market agree with the investment surplus from the project and bargain over the surplus. Without loss of generality, I analyze the case where (i) the market is competitive and (ii) the startup raises capital by selling shares. In particular, once taking action Q, the startup obtains

$$g^*(s) = (\mathbb{E}_s(\theta) - k)^+ = (\exp\{\mu + \sigma^2/2\} - k)^+, \tag{19}$$

which is the expected investment surplus according to the state s. Notice that there are two possibilities. If $\mathbb{E}_s(\theta) < k$, the startup quits without investment; the project is abandoned and generates a payoff of $g^*(s) = 0$. If $\mathbb{E}_s(\theta) \geq k$, the market requires an ownership of $\lambda = \frac{k}{\mathbb{E}_s(\theta)}$ to break even. The startup completes the project and obtains a payoff of $(1 - \lambda)\theta$, which has an expectation value of $g^*(s) = \exp\{\mu + \sigma^2/2\} - k$.

H.2 Startup Strategy

In this section, I analyze the equilibrium strategy of the startup. Before concluding the optimal policy, I first analyze why the startup benefits from additional signals. Lemma A1 shows that the value of additional signals comes from reducing σ and obtaining a precise posterior belief.

Lemma A1. *Assuming that the startup is in state $s = (\mu, \sigma)$ at date t, where $\sigma > 0$, we have*

$$\underbrace{g^*(\mu, \sigma)}_{\text{Quit}} < \underbrace{\mathbb{E}[g^*(\tilde{\mu}, \tilde{\sigma})|Z_{t+1} = 1, a_{t+1} = Q]}_{\text{Noisy Signal}} < \underbrace{\mathbb{E}[g^*(\tilde{\mu}, 0)|Z_{t+1} = 0, a_{t+1} = Q]}_{\text{Precise Signal}},$$

where $\tilde{\sigma} = \sqrt{\frac{1}{\sigma^{-2}+\nu^{-2}}}$.

The first term in Lemma A1 is the payoff of stopping at date t, while the other two terms show the expected payoff of stopping at the following date with a noisy and precise signal, respectively. The first inequality demonstrates that an additional noisy signal benefits the startup by reducing the standard deviation from σ to $\tilde{\sigma}$. This reduction in uncertainty results in lower investment inefficiencies and higher investment surplus. The second inequality shows that the startup benefits from receiving a precise signal that identifies the exact value of θ. After observing a precise signal, the project is invested only if the NPV is positive ($\theta > k$) and no inefficiencies occur, resulting in the first-best investment surplus. Given the assumption of a competitive financial market, all investment surplus accrues to the startup. Consequently, the startup has an incentive to collect signals to reduce σ and increases investment surplus.

Though an additional signal increases financial payoff of the startup, it is associated with a fixed cost of c. The startup trades-off the benefit and the cost of signals to make continuation decisions. Proposition A1 and Corollary A2 establish that the startup continues to collect signals until σ is sufficiently small, at which point further signal collection is no longer profitable considering the cost of signal collection.

Proposition A1. *Define $\Sigma_0 := \left[0, \frac{2\pi c^2}{k^2}\right]$, the optimal policy satisfies $d^*(s) = Q$ for $s \in \mathbb{R} \times \Sigma_0$. In particular, the startup quits upon the arrival of the first precise signal.*

Corollary A2. *For any given μ, there exists σ^* such that $d^*(s) = Q$ if $\sigma < \sigma^*$.*

As σ decreases, the posterior belief converges to θ. When σ is sufficiently small, the posterior mean μ barely changes with additional signals, resulting in a tiny increase in the expected payoff of the startup to continue. When the payoff gain from additional signals are dominated by the cost of signal collection, the startup optimally quits from the financial market and takes investment surplus if there is any. Figure A4 shows the numerical solution to the optimal policy. As illustrated in Figure A4, for any given μ, the startup is more likely to continue when σ is large.

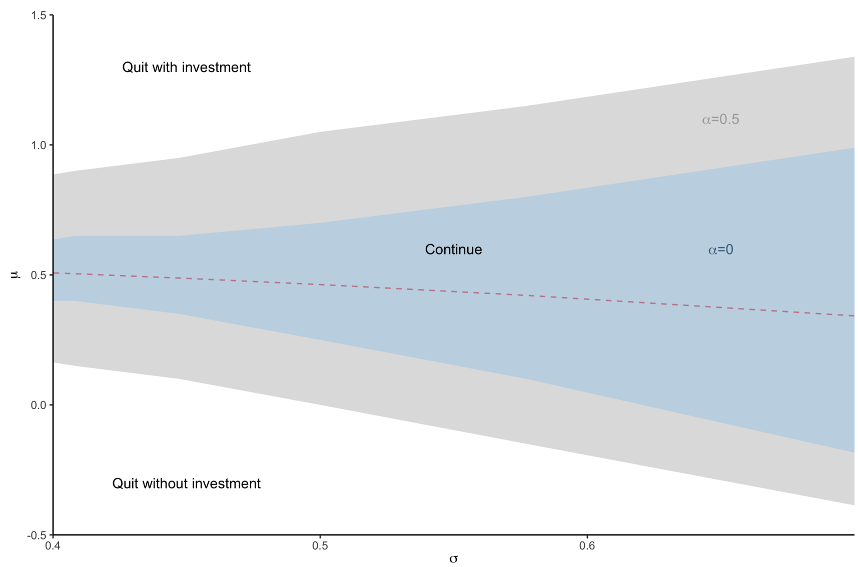


Figure A4: **Startup's Continuation Decision as a Function of Posterior Belief.** This figure shows the startup's continuation decision as a function of μ and σ. The startup continues signal collection in the shaded area and stops searching otherwise. The red lines plots a combination of μ and σ that satisfy $\mathbb{E}_s(\theta) = \exp\{\mu + \frac{1}{2}\sigma^2\} = k$. The blue shade and the gray shade correspond to $\alpha = 0$ and $\alpha = 0.5$ respectively. As the figure shows, the startup is more likely to search when (i) σ is higher; (ii) μ that provides $\mathbb{E}(\theta)$ around k; and (iii) α is larger. I assumed $\mu_0 = 0, \sigma_0 = 1, \nu = 1, k = 1.8, c = 0.05$.

In addition to σ, the optimal decision also depends on μ. Proposition A2 concludes the optimal policy as a function of μ.

Proposition A2. *For any given σ, there exists $\underline{\mu} \leq \bar{\mu}$ such that the optimal policy is $d^*(s) = Q$ if $\mu < \underline{\mu}$ or $\mu > \bar{\mu}$.*

Proposition A2 indicates the startup continues when μ is neither too small nor too large. As shown in Figure A4, the startup selects to continue when μ is around the red dashed line, which plots μ that gives $\mathbb{E}_s(\theta) = k$. This can be explained by the option feature of the project, which is summarized by Proposition A3.

Proposition A3. *For $s = (\mu, \sigma) \in \mathcal{S}$ where $\sigma > 0$, the expected value of the next date $\mathbb{E}_s[\mathcal{V}(\tilde{S})]$ is convex in μ and increasing in σ.*

Payoff of an option is featured by its convexity in the value of the fundamental asset, which is captured by the parameter μ in this model. Proposition A3 shows that the expected value from the following date $\mathbb{E}_s[\mathcal{V}(\tilde{S})]$ is convex in the posterior mean, which shows an option feature. The feature comes from the assumption that the startup obtains any positive investment surplus without bearing downside costs. The startup's project can be described as a real option below. The option has a premium of c. The log value of fundamental asset follows the distribution of $\mathcal{N}(\mu, \sigma^2)$. After purchasing the option, the startup has a right to make a trading decision at the next date. The trading involves in a strike price of k. Whenever the expected project value is larger than k, the startup can pay k to the market in equity to exercise the option and obtain the surplus. Otherwise, the startup can long another option to continue the process.

The option feature explains the optimal policy of the startups. Since the value of the option increases with the volatility of the fundamental asset, the startup is more likely to long the option when σ is large. This is consistent with the conclusion given by Proposition A1 and Corollary A2. Moreover, since the value of signal collection comes from the reduction of σ, the startup is more likely to long the option at a μ that makes option value sensitive to the change in σ. Therefore, the startup is more likely to continue when the underlying price is near the option's strike price.

This result is consistent with existing papers that highlights the option feature of entrepreneurial projects (Kerr et al., 2014; Manso, 2016). My model highlights that the realization of option value relies on the information from the financial market. The financial market helps reveal project value besides providing capital investments. As a result, the financial market's ability to provide information regarding project value affects the value of the real option, as shown by the next section.

H.3 Benefits from VC Informativeness

Given the informational role of the financial market, it is crucial that the financial market can precisely predict the cash flow generated from a startup project and then provide the startup with precise information. This section examines the benefits of having a project within VCs expertise, i.e. high VC informativeness.

H.3.1 VC Informativeness Enhances Startup Payoff

Recall that VC informativeness is captured by the probability of having a precise signal each round. In this section, I demonstrate that the startup benefits more from receiving a precise signal in the next stage, provided they choose to continue. First, notice that the expected value next date can be written as

$$\mathbb{E}_s[\mathcal{V}(\tilde{S})] = \alpha \mathbb{E}_s[g^*(\tilde{\mu}, 0)|Z_{t+1} = 0] + (1-\alpha)\mathbb{E}_s[\mathcal{V}(\tilde{\mu}, \tilde{\sigma})|Z_{t+1} = 1],$$

where $\tilde{\sigma} = \sqrt{\frac{1}{\sigma^{-2}+\nu^{-2}}}$. According to Proposition A1, the startup immediately quits once receiving a precise signal and obtains a payoff $g^*(\tilde{\mu}, 0)$. While receiving another noisy signal and update belief to $\tilde{S} = (\tilde{\mu}, \tilde{\sigma})$, the startup solves the Bellman equation (17) and obtain an expected payoff of $\mathcal{V}(\tilde{\mu}, \tilde{\sigma})$. Lemma A2 shows that the startup has a high expected payoff conditional on having a precise signal next stage.

Lemma A2. *For any given state $s = (\mu, \sigma)$ and $\tilde{\sigma} = \sqrt{\frac{1}{\sigma^{-2}+\nu^{-2}}}$, we have*

$$\mathbb{E}_s[\mathcal{V}(\tilde{\mu}, \tilde{\sigma})|Z_{t+1} = 1] \leq \mathbb{E}_s[g^*(\tilde{\mu}, 0)|Z_{t+1} = 0].$$

Lemma A2 shows that the startup benefits from a precise signal. Receiving a precise signal results in $\sigma = 0$, maximizing the investment surplus and dominating the case where a noisy signal is received. The option of continuation cannot undo the disadvantage of receiving a noisy signal since continuation comes with a cost for the startup. Considering this, the expected value from signal collection increases in the probability of having a precise signal, as shown by Lemma A3.

Lemma A3. *For any given state $s = (\mu, \sigma)$, $\mathbb{E}_s[\mathcal{V}(\tilde{S})]$ increases in α.*

Notice that a high $\mathbb{E}_s[\mathcal{V}(\tilde{S})]$ caused by a high α encourages continuation. Figure A4 shows that the startup is more likely to continue when α is large. The figure shows the startup strategy under $\alpha = 0$ (blue shade) and $\alpha = 0.5$ (gray shade), respectively. The blue shade is fully covered by gray, indicating that the startup tends to wait more when $\alpha = 0.5$. Driven by the increase of $\mathbb{E}_s[\mathcal{V}(\tilde{S})]$, the value function $\mathcal{V}(s)$ also increases in α, which is summarized below.

Proposition A4. *For any given state $s = (\mu, \sigma)$, $V(s)$ increases in α.*

Note that Proposition A4 applies to $\mathcal{V}(\mu_0, \sigma_0; \alpha)$ as well. In other words, the expected total payoff of the startup at $t = 0$ increases in α, i.e., decreases in opacity. Figure A5 illustrates this idea.

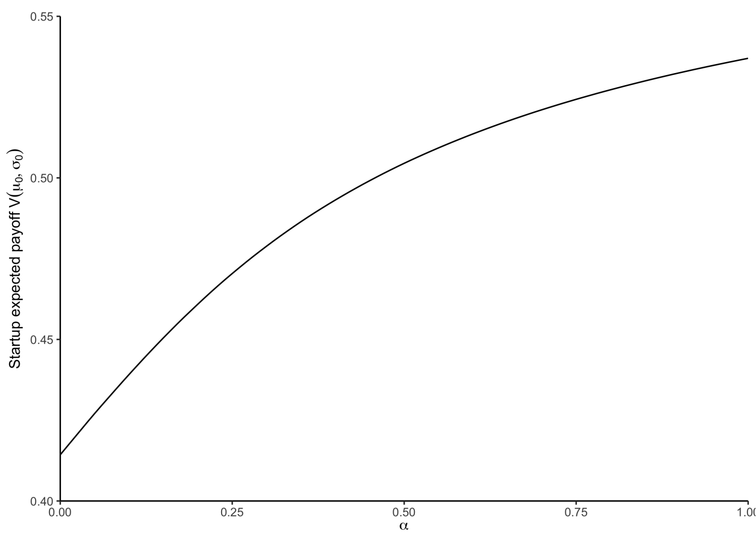


Figure A5: **Startup Expected Payoff.** This figure illustrates the the expected payoff of the startup $V(\mu_0, \sigma_0; \alpha)$. The startup has a higher expected payoff when opacity is lower (α is larger). I assumed $\mu_0 = 0, \sigma_0 = 1, \nu = 1, k = 1.8, c = 0.05$.

H.3.2 Channels through which VC Informativeness Enhances Startup Payoff

The previous section concludes that the startup's expected payoff increases in VC informativeness. In this section, I examine more details about the mechanism behind the conclusion. In particular, I analyze how VC informativeness impacts the probability of investment, the cost of capital, and the expected signal collection time of the startup. All figures in this section are produced by simulation, and details are provided in Section H.6.2.

As shown in the previous analysis, VC informativeness allows the startup to resolve the true project type, which leads to a high surplus from the investment. Figure A6 plots the probability that the startup resolves the true project type before quit as a function of VC informativeness α. The total probability, plotted by the black dots, consists of two parts. First, a high α mechanically means a higher probability of resolving the true type, as the red dashed line shows. Second, α also encourages searching as indicated by Lemma A3. This explains the difference between the black dots and the red dashed line. Overall, the total probability of resolving project type increases in α.

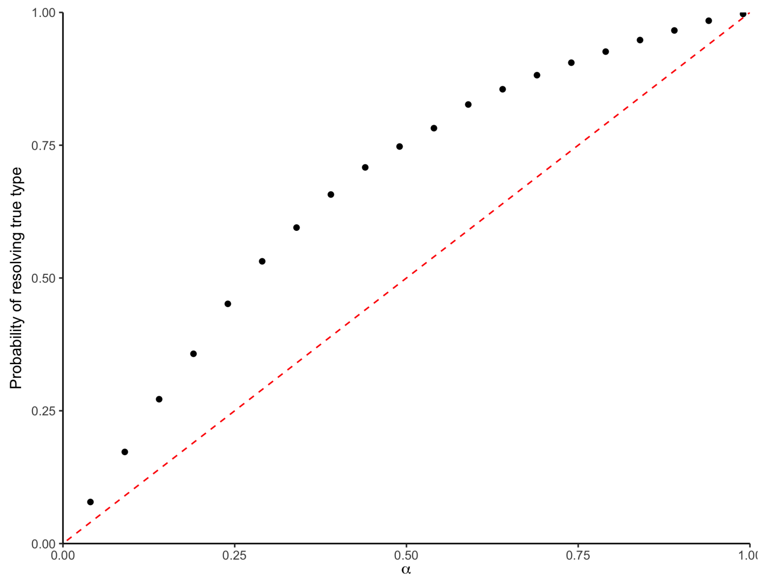
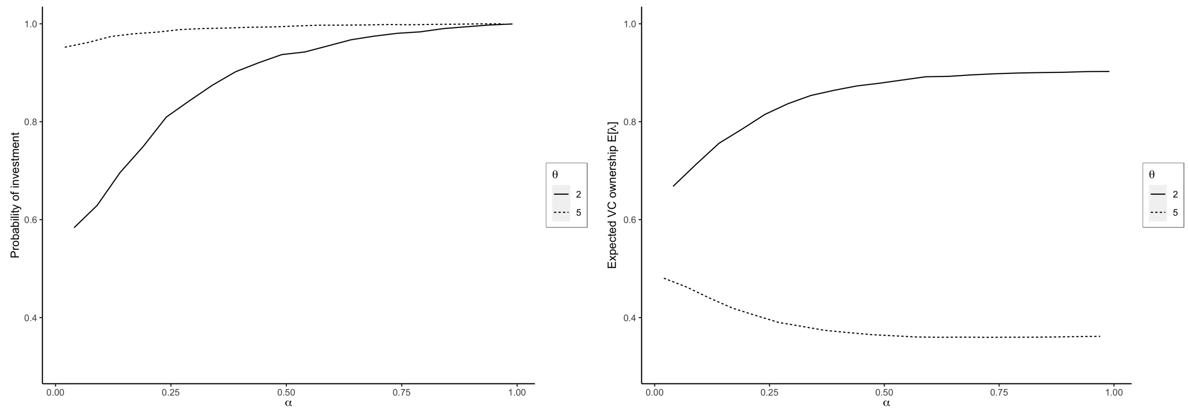


Figure A6: **Probability of Resolving Project Type in the Last Search Round.** This figure (black dots) shows simulation results on the probability of resolving true type, i.e., having a fully informed VC, in the last search round. The red dashed line is a 45 degree line. It shows the probability of resolving true type in the last round if the startup's continuation decision is independent to signal precision. I assume $\mu_0 = 0, \sigma_0 = 1, \nu = 1, k = 1.8, c = 0.05$.

As Figure A6 shows, a high α indicates that a startup is less likely to be misvalued and face inefficient investment decisions. I call this *misvaluation channel*. Misvaluation impacts the startup's expected payoff through the probability of investment and the cost of capital, and the effects differ according to the true project value. Figure A7 shows the simulation results, taking two positive NPV projects as examples.[43] $\theta = 2$ and $\theta = 5$ represent a marginal project and

[43] I mainly analyze the case where parameter value satisfies $\mathbb{E}(\theta) = \mu_0 + \sigma_0^2/2 < k$ since most startups create zero value (Hall and Woodward, 2010).



(a) Probability of investment

(b) VC's required ownership

Figure A7: **Probability of Investment and Cost of Capital.** This figure shows simulation results on the probability of investment and cost of capital for given types of projects. $\theta = 2$ is an example of a marginal project that has a positive NPV, and $\theta = 5$ represents a high value project. Figure (a) show that the probability of investment increases in α for both projects. Figure (b) shows that cost of capital increases (decreases) in VC informativeness for the startup if the project has marginal (high) value. I assumed $\mu_0 = 0, \sigma_0 = 1, \nu = 1, k = 1.8, c = 0.05$.

a high-value project, respectively. For a positive NPV project, the probability of investment increases in α, as shown in Figure A7 (a). This effect is extremely pronounced for a marginal project ($\theta = 2$). Figure A7 (b) shows the expected ownership required by the VC conditional on investment. A marginal project ($\theta = 2$) is less likely to be pooled with high-value projects when α is larger, so its cost of capital increases in α. On the contrary, a high-type project has a decreasing cost of capital when α increases. To conclude, an increase in α has heterogeneous impacts on the projects depending on the true type.

Notice that α also impacts startup payoff through the *signal collection channel*. There are two driving forces through which α affects time spent on signal collection. First, a higher α increases the probability of receiving the precise signal, which causes quits by Proposition A1. Second, according to Lemma A3, α encourages waiting conditional on having $\sigma > 0$. The two forces drive the expected waiting time in different directions, and the former dominates when α is sufficiently large. Figure A8 shows the expected signal collection time as a function of VC informativeness α. When α is sufficiently large, the startup can quickly resolve project value and quit from the market.

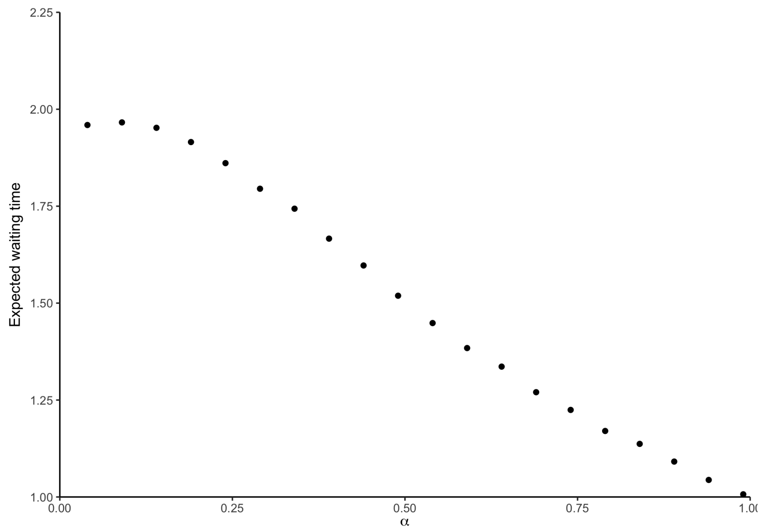


Figure A8: **Expected Signal Collection Time.** This figure shows simulation results on the startup's expected signal collection time as a function of α. I assume $\mu_0 = 0, \sigma_0 = 1, \nu = 1, k = 1.8, c = 0.05$.

In sum, VC informativeness affects startup payoff through the probability of being invested, the cost of capital, and the signal collection time. The influence differs in true project type.

H.4 Startup Project Choice

In this section, I endogenize startup project choice. Similar to the baseline model, I assume that the startup at $t = 0$ selects between a high-payoff, low-VC informativeness project (exploration) and a low-payoff, high-VC informativeness project (catering). These actions are denoted by $I_0 = \{I_E, I_C\}$. In particular, I assume that

$$\sigma_{0,E} > \sigma_{0,C}; \alpha_E < \alpha_C,$$

where $\sigma_{0,i}$ represents the standard deviation of the prior belief on θ for $i \in \{E, C\}$. The parameter α_i represents VC informativeness of project i.

The startup faces a payoff-information trade-off similar to the baseline model. All else equal, a high α_0 project brings higher expected payoff by Proposition A3, which illustrates potential benefits of exploration. However, low VC informativeness associated with exploration can reduces startup payoff through the information channels, as shown by Lemma A3. The startup considers the expected project payoffs and VC informativeness when selecting a project. Figure A9 show that the startup is more likely to choose the catering project when its value is sufficiently informed to VCs.

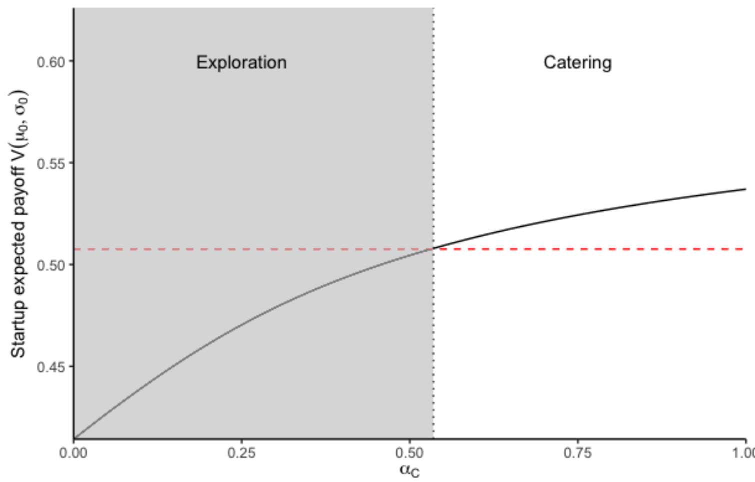


Figure A9: **Startup Project Choice.** This figure shows simulation results on the startup's project choice as a function α_C. The black line shows startup expected payoff from the catering project, while the red dashed line plots startup expected payoff from exploration. Ib the shaded area where α_C is relatively small, the startup selects exploration in equilibrium. When α_C is sufficiently large (unshaded region), the startup optimally initiates the catering project. In I assume $\mu_0 = 0, \sigma_{0,C} = 1, \sigma_{0,E} = 1.05, \nu = 1, k = 1.8, c = 0.05$.

H.5 Proofs

This section is organized as below. First, I show that the underlying Markov decision process is a positive bounded model (Puterman, 1994, Section 7.2) so that I can leverage established results in Puterman (1994) to prove lemmas and properties. Next, I provide proofs of lemmas, properties and corollaries. The last subsection, I provide additional lemmas used in the second subsection.

H.5.1 Proof of Positive Boundedness

In order to characterize the value function, we define the bellman operator of a general state-value function $\widetilde{\mathcal{V}}$ as:

$$\mathsf{B}(\widetilde{\mathcal{V}})(s) := \begin{cases} \max \begin{Bmatrix} g^\star(s), & a = Q \\ -c + \mathbb{E}_{\widetilde{\mu} \sim \mathcal{N}\left(\mu, \sigma^2 Z + \frac{\sigma^4}{\sigma^2+\nu^2}(1-Z)\right)} \left[\widetilde{\mathcal{V}} \left(\widetilde{\mu}, \dfrac{1-Z}{\sigma^{-2}+\nu^{-2}} \right) \right], & a = C \end{Bmatrix}, & s = (\mu, \sigma^2) \in \mathcal{S}; \\ 0, & s = \Delta. \end{cases}$$

We are interested in the following two properties.

- (Puterman, 1994, Bellman Equation, Theorem 7.2.3) A general state-value function $\widetilde{\mathcal{V}} : \mathcal{S}_0 \to \mathbb{R}^+$ satisfies the Bellman equation $\widetilde{\mathcal{V}} = \mathsf{B}(\widetilde{\mathcal{V}})$ if and only if $\widetilde{\mathcal{V}}$ is the optimal value function. In particular, $\mathcal{V} = \mathsf{B}(\mathcal{V})$.
- (Puterman, 1994, Value Iteration, Corollary 7.2.13) Suppose an initial value function $\mathcal{V}^{(0)} : \mathcal{S}_0 \to \mathbb{R}^+$ satisfies $0 \le \mathcal{V}^{(0)}(s) \le \mathcal{V}(s)$. Consider value iteration as follows: for $t = 1, 2, \cdots$, let $\mathcal{V}^{(t)} := \mathsf{B}(\mathcal{V}^{(t-1)})$. Then for every fixed $s \in \mathcal{S}_0$, $\mathcal{V}^{(t)}(s)$ increases to $\mathcal{V}(s)$ as $t \to \infty$.

To better understand the dynamic of state transition, especially before the stopping state Δ is hit, we study a never-stopping policy and characterize the dynamic in the following proposition.

Proposition A5 (Equivalent Data Generation)**.** *Consider $\mathbb{P}^C$ as the probability measure corresponding to the never-stopping policy, that is, $\mathbb{P}^C(A_t = C | S_t = s) = 1$ for any $s \in \mathcal{S}$. Fix $(\mu_0, \sigma_0^2) \in \mathbb{R} \times \mathbb{R}_+$. Suppose $\{\zeta_t\}_{t=1}^\infty \overset{iid}{\sim} \mathcal{N}(0,1)$. For $t = 1, 2, \cdots$, let*

$$\sigma_t^{-2} := \sigma_{t-1}^{-2} + \nu^{-2}; \quad \mu_t := \mu_{t-1} + \sqrt{\sigma_{t-1}^2 - \sigma_t^2}\,\zeta_t; \quad \mu_t^\Delta := \mu_{t-1} + \sigma_{t-1}\zeta_t.$$

Also let $\tau_\gamma := \inf\{t \in \mathbb{N}_0 : \gamma_t = 1\}$. Then for $t \in \mathbb{N}$, the followings hold.

- $\sigma_t^{-2} = \sigma_0^{-2} + t\nu^{-2}$, $\mu_t = \mu_0 + \sum_{u=1}^{t} \sqrt{\sigma_{u-1}^2 - \sigma_u^2}\zeta_u$, $\mu_t^\Delta = \mu_0 + \sum_{u=1}^{t-1} \sqrt{\sigma_{u-1}^2 - \sigma_u^2}\zeta_u + \sigma_{t-1}\zeta_t$.

- *Conditional on $S_0 = (\mu_0, \sigma_0^2)$ and $\{\gamma_t \geq t\}$, we have $\{S_u\}_{u=0}^t \stackrel{\mathcal{D}}{=} \{(\mu_u, \sigma_u^2)\}_{u=0}^t$, and in particular, $\mu_t \sim \mathcal{N}(\mu_0, \sigma_0^2 - \sigma_t^2)$,*

- *Conditional on $S_0 = (\mu_0, \sigma_0^2)$ and $\{\tau_\gamma \leq t-1\}$, we have $\{S_u\}_{u=0}^t \stackrel{\mathcal{D}}{=} \{(\mu_u, \sigma_u^2)\}_{u=0}^{\tau_\gamma} \cup \{(\mu_{\tau_\gamma+1}^\Delta, 0)\}_{u=\tau_\gamma+1}^t$, and in particular, $\mu_{\tau_\gamma+1}^\Delta = \mu_{\tau_\gamma+2}^\Delta = \cdots = \mu_t^\Delta \sim \mathcal{N}(\mu_0, \sigma_0^2)$.*

Next, I show that the underlying MDP is a positive bounded model (Puterman, 1994, Section 7.2).

Lemma A4 (Positivity)**.** *For each $s \in \mathcal{S}_0$, there exists $a \in \mathcal{A}_s$, such that $g(s,a) \geq 0$.*

Proof of Lemma A4. If $s = \Delta$, then $g(\Delta, Q) = 0$. If $s = (\mu, \sigma^2) \in \mathcal{S}$, then $g(s, Q) = \left[\exp\left(\mu + \frac{\sigma^2}{2}\right) - k\right]^+ \geq 0$. □

Lemma A5 (Finiteness)**.** *For any policy d and initial state $s \in \mathcal{S}_0$, we have*

$$\mathcal{V}_+^d(s) := \mathbb{E}^d\left\{\sum_{t=0}^{\infty}[g(S_t, A_t)]^+ \middle| S_0 = s\right\} < +\infty$$

Proof of Lemma A5. The finiteness holds trivially at $s = \Delta$ for all policy d, since $d(\Delta) = Q$ and $\mathcal{V}^d(\Delta) = 0$. Then we can assume $s = (\mu, \sigma^2) \in \mathcal{S}$ without loss of generality. By definition, we have

$$\mathcal{V}_+^d(s) = \mathbb{E}^d[g^\star(S_{\tau-1})\mathbb{I}(\tau < +\infty)|S_0 = s].$$

Let $\mathcal{S}^d := \{s \in \mathcal{S} : \mathcal{V}^d(s) = g^\star(s)\}$ be the stopping state space corresponding to the policy d, and $\tau(\mathcal{S}^d) := \inf\{t \in \mathbb{N}_0 : S_t \in \mathcal{S}^d\}$ be the corresponding stopping time. Then under $\mathbb{P}^d$, we have $\tau = \tau(\mathcal{S}^d) + 1$. Let $\mathbb{P}^C$ be the probability measure corresponding to an never-stopping policy, that is, $\mathbb{P}^C(A_t = C|S_t = s) = 1$ for any $s \in \mathcal{S}$. Then $\{g^\star(S_{\tau(\mathcal{S}^d)\wedge t})\}_{t=0}^\infty$ under $\mathbb{P}^d$ has the same distribution as $\{g^\star(S_{\tau(\mathcal{S}^d)\wedge t})\}_{t=0}^\infty$ under $\mathbb{P}^C$. By Lemma A6, $\{g^\star(S_t)\}_{t=0}^\infty$ is a non-negative $\mathbb{P}^C$-sub-martingale. Then the stopped sequence $\{g^\star(S_{\tau(\mathcal{S}^d)\wedge t})\}_{t=0}^\infty$ is a non-negative $\mathbb{P}^C$-sub-martingale as well. For $p > 1$, we have

$$
\begin{aligned}
&\sup_{t\in\mathbb{N}} \mathbb{E}^C \left\{ [g^\star(S_{\tau(\mathcal{S}^d)\wedge t})]^p \middle| S_0 = s \right\} \\
&\le \sup_{t\in\mathbb{N}} \mathbb{E}^C \left\{ \max_{1\le u\le t} [g^\star(S_u)]^p \middle| S_0 = s \right\} \\
&\le \left(\frac{p}{p-1}\right)^p \sup_{t\in\mathbb{N}} \mathbb{E}^C \left\{ [g^\star(S_t)]^p | S_0 = s \right\} \quad \text{(by Doob's } L^p\text{-maximal inequality)} \\
&< +\infty \quad \text{(by Lemma A7)}.
\end{aligned}
$$

Then by Doob's Martingale Convergence Theorem, we further have

$$
g^\star\big(S_{\tau(\mathcal{S}^d)\wedge t}\big) \overset{t\to\infty}{\longrightarrow} g^\star\big(S_{\tau(\mathcal{S}^d)}\big) \ a.s.\ \mathbb{P}^C \text{ and in } L^p(\mathbb{P}^C); \quad \mathbb{E}^C \left[g^\star\big(S_{\tau(\mathcal{S}^d)}\big) \middle| S_0 = s \right] < +\infty.
$$

Therefore,

$$
\begin{aligned}
\mathcal{V}_+^d(s) &= \mathbb{E}^d[g^\star(S_{\tau-1})\mathbb{I}(\tau < +\infty)|S_0 = s] \\
&= \mathbb{E}^d \left[g^\star\left(S_{\tau(\mathcal{S}^d)}\right) \mathbb{I}[\tau(\mathcal{S}^d) < +\infty] \middle| S_0 = s \right] \\
&= \mathbb{E}^C \left[g^\star\left(S_{\tau(\mathcal{S}^d)}\right) \mathbb{I}[\tau(\mathcal{S}^d) < +\infty] \middle| S_0 = s \right] \\
&< +\infty.
\end{aligned}
$$

□

Lemma A6 (Sub-Martingale)**.** *Consider the filtration $\mathscr{F} = \{\mathcal{F}_t\}_{t=0}^\infty$ generated by $\{S_t, A_t, \gamma_t\}_{t=0}^\infty$. Suppose $\mathbb{P}^C$ is the probability measure corresponding to the never-stopping policy, that is, $\mathbb{P}^C(A_t = C|S_t = s) = 1$ for any $s \in \mathcal{S}$. Then the sequence $\{g^\star(S_t)\}_{t=0}^\infty$ is a $\mathbb{P}^C$-$\mathscr{F}$-sub-martingale.*

Proof of Lemma A6. Without loss of generality, consider $s \ne \Delta$ and $s = (\mu, \sigma^2) \in \mathcal{S}$. Under the initial condition $S_0 = s \ne \Delta$ and the never-stopping policy, we have $S_t \ne \Delta$ for all $t \in \mathbb{N}_0$. For ease of notation, we write $\mathbb{P}^C$ as $\mathbb{P}$ in this proof. Suppose $S_t = (\mu_t, \sigma_t^2)$ and $S_{t+1} = (\mu_{t+1}, \sigma_{t+1}^2)$.

- If $\sigma_t^2 = 0$, then $S_{t+1} = S_t$, and we have $\mathbb{E}[g^\star(S_{t+1})|\mathcal{F}_t] = g^\star(S_t)$.

- If $\sigma_t^2 > 0$ and $\gamma_t = 1$, then $\mu_{t+1}|\mathcal{F}_t \sim \mathcal{N}(\mu_t, \sigma_t^2)$ and $\sigma_{t+1}^2 = 0$. We have

$$\begin{aligned}
&\mathbb{E}[g^\star(S_{t+1})|\mathcal{F}_t] = \mathbb{E}[(e^{\mu_{t+1}} - k)^+|\mathcal{F}_t] \\
&\geq [\mathbb{E}(e^{\mu_{t+1}}|\mathcal{F}_t) - k]^+ && \text{(by Jensen's inequality)} \\
&= \left[\mathbb{E}_{\xi\sim\mathcal{N}(\mu_t,\sigma_t^2)}(e^\xi) - k\right]^+ \\
&= g^\star(S_t).
\end{aligned}$$

- If $\sigma_t^2 > 0$ and $\gamma_t = 0$, then $\sigma_{t+1}^{-2} = \sigma_t^{-2} + \nu^{-2}$ and $\mu_{t+1}|\mathcal{F}_t \sim \mathcal{N}(\mu_t, \sigma_t^2 - \sigma_{t+1}^2)$. We have

$$\begin{aligned}
&\mathbb{E}[g^\star(S_{t+1})|\mathcal{F}_t] = \mathbb{E}\left\{\left[\mathbb{E}_{\xi\sim\mathcal{N}(\mu_{t+1},\sigma_{t+1}^2)}(e^\xi) - k\right]^+ \middle| \mathcal{F}_t\right\} \\
&\geq \left\{\mathbb{E}\left[\mathbb{E}_{\xi\sim\mathcal{N}(\mu_{t+1},\sigma_{t+1}^2)}(e^\xi)|\mathcal{F}_t\right] - k\right\}^+ && \text{(by Jensen's inequality)} \\
&= \left[\mathbb{E}_{\xi\sim\mathcal{N}(\mu_t,\sigma_t^2)}(e^\xi) - k\right]^+ \\
&= g^\star(S_t).
\end{aligned}$$

  Here, the third equality follows from the fact that, if $\xi|\mu_{t+1} \sim \mathcal{N}(\mu_{t+1}, \sigma_{t+1}^2)$, and $\mu_{t+1}|\mathcal{F}_t \sim \mathcal{N}(\mu_t, \sigma_t^2 - \sigma_{t+1}^2)$, then conditional on $\mathcal{F}_t$, ξ is normally distributed, with

$$\begin{aligned}
&\mathbb{E}(\xi|\mathcal{F}_t) = \mathbb{E}[\mathbb{E}(\xi|\mu_{t+1})|\mathcal{F}_t] = \mathbb{E}(\mu_{t+1}|\mathcal{F}_t) = \mu_t; \\
&\mathsf{Var}(\xi|\mathcal{F}_t) = \mathbb{E}[\mathsf{Var}(\xi|\mu_{t+1})|\mathcal{F}_t] + \mathsf{Var}[\mathbb{E}(\xi|\mu_{t+1})|\mathcal{F}_t] \\
&= \sigma_{t+1}^2 + \mathsf{Var}(\mu_t|\mathcal{F}_t) = \sigma_{t+1}^2 + \sigma_t^2 - \sigma_{t+1}^2 = \sigma_t^2.
\end{aligned}$$

These three cases conclude the proof that $\mathbb{E}[g^\star(S_{t+1})|\mathcal{F}_t] \geq g^\star(S_t)$. □

Lemma A7 (L^p-Boundedness)**.** *Suppose $\mathbb{P}^C$ is the probability measure corresponding to the never-stopping policy, that is, $\mathbb{P}^C(A_t = C|S_t = s) = 1$ for any $s \in \mathcal{S}$. For any $p \geq 1$ and initial condition $s \in \mathcal{S}_0$, we have*

$$\sup_{t\in\mathbb{N}} \mathbb{E}^C\big\{[g^\star(S_t)]^p|S_0 = s\big\} < +\infty.$$

Proof of Lemma A7. Without loss of generality, consider $s \neq \Delta$ and $s = (\mu, \sigma^2) \in \mathcal{S}$. Under the initial condition $S_0 = s \neq \Delta$ and the never-stopping policy, we have $S_t \neq \Delta$. Denote $S_t = (\mu_t, \sigma_t^2)$.

$$
\begin{aligned}
&\sup_{t\in\mathbb{N}} \mathbb{E}^C \left\{[g^\star(\mu_t,\sigma_t^2)]^p \middle| S_0 = s\right\} \leq 2^{p-1}\left\{e^{p\sigma_t^2/2}\mathbb{E}(e^{p\mu_t}|S_0=s) + k^p\right\} && \text{(by AM-GM inequality)}\\
&\leq 2^{p-1}\left\{e^{p\sigma_t^2/2} \times e^{p\mu+p^2\sigma^2/2} + k^p\right\} && \text{(by Proposition A5)}\\
&< +\infty.
\end{aligned}
$$

□

H.5.2 Proof of Lemmas, Propositions and Corollaries

Proof of Lemma A1. Notice that $g^*(\mu,\sigma)$ is convex in μ. Lemma A1 holds by Jensen's inequality.

□

Proof of Proposition A1. It is equivalent to prove that $\mathcal{V}(s) = g^*(s)$ for any $s \in \mathbb{R}\times\Sigma_0$. In particular, let $\mathcal{V}^{(0)}(s) = g^*(s)$ for $s \in \mathcal{S}$ and $\mathcal{V}^{(t)} := B(\mathcal{V}^{(t-1)})$ for $t = 1, 2, ...,$ I show that for every $s \in \mathbb{R}\times\Sigma_0$, $\mathcal{V}(s) = \lim_{t\to\infty}\mathcal{V}^{(t)}(s) = g^*(s)$.

First of all, I use mathematical induction to prove that $\mathcal{V}^{(t)}(s) = g^*(s)$ holds for all $t \in \mathbb{N}_0$ and $s \in \mathbb{R}\times\Sigma_0$. Notice that $\mathcal{V}^{(0)}(s) = g^*(s)$ for any $s \in \mathbb{R}\times\Sigma_0 \subset \mathcal{S}$ holds by assumption. Next, I show that $\mathcal{V}^{(t+1)}(s) = g^*(s)$ for any $s \in \mathbb{R}\times\Sigma_0$ if $\mathcal{V}^{(t)}(s) = g^*(s)$ holds for any $s \in \mathbb{R}\times\Sigma_0$. By definition, for a fixed $s \in \mathbb{R}\times\Sigma_0$,

$$
\mathcal{V}^{(t+1)}(s) := \left\{\max\begin{cases} g^\star(s), & a = Q \\ -c + \mathbb{E}_S[\mathcal{V}(\tilde{S})], & a = C\end{cases}\right\},
$$

where

$$
\mathbb{E}_S[\mathcal{V}(\tilde{S})] = \mathbb{E}_{\widetilde{\mu}\sim\mathcal{N}\left(\mu,\sigma^2 Z + \frac{\sigma^4}{\sigma^2+\nu^2}(1-Z)\right)}\left[\mathcal{V}^{(t)}\left(\widetilde{\mu}, \frac{1-Z}{\sigma^{-2}+\nu^{-2}}\right)\right]
$$

Since $\frac{1-Z}{\sigma^{-2}+\nu^{-2}} \leq \sigma^2 \leq \frac{2\pi c^2}{k^2}$, we have $(\widetilde{\mu}, \frac{1-Z}{\sigma^{-2}+\nu^{-2}}) \in \mathbb{R}\times\Sigma_0$ and $\mathcal{V}^{(t)}(\widetilde{\mu}, \frac{1-Z}{\sigma^{-2}+\nu^{-2}}) = g^*(\widetilde{\mu}, \frac{1-Z}{\sigma^{-2}+\nu^{-2}})$.

Then $\mathcal{V}^{(t+1)} = g^*(s)$ is equivalent to

$$
\begin{aligned}
g^*(s) &\geq -c + \mathbb{E}_{\widetilde{\mu}\sim\mathcal{N}\left(\mu,\sigma^2 Z+\frac{\sigma^4}{\sigma^2+\nu^2}(1-Z)\right)} g^*(\widetilde{\mu}, \frac{1-Z}{\sigma^{-2}+\nu^{-2}}) \\
&= -c + \alpha \underbrace{\mathbb{E}_{\tilde{\mu}\sim\mathcal{N}(\mu,\sigma^2)} g^*(\tilde{\mu}, 0)}_{(I)} + (1-\alpha) \underbrace{\mathbb{E}_{\tilde{\mu}\sim\mathcal{N}(\mu,\frac{\sigma^4}{\sigma^2+\nu^2})} g^*(\tilde{\mu}, \frac{1}{\sigma^{-2}+\nu^{-2}})}_{(II)}
\end{aligned}
$$

Notice that

$$
\begin{aligned}
&(II) = \mathbb{E}_{\tilde{\mu}\sim\mathcal{N}\left(\mu,\frac{\sigma^4}{\sigma^2+\nu^2}\right)} g^*\left(\tilde{\mu}, \frac{1}{\sigma^{-2}+\nu^{-2}}\right) = \mathbb{E}_{\tilde{\mu}\sim\mathcal{N}\left(\mu,\frac{\sigma^4}{\sigma^2+\nu^2}\right)} \left\{\mathbb{E}_{\xi\sim\mathcal{N}\left(\tilde{\mu},\frac{1}{\sigma^{-2}+\nu^{-2}}\right)} e^{\xi} - k\right\}^+ \\
&\leq \mathbb{E}_{\xi\sim\mathcal{N}(\mu,\sigma^2)}(e^{\xi}-k)^+ \quad \text{(by Jensen's inequality)} \\
&= (I),
\end{aligned}
$$

where the second equality follows from the fact that ξ is normally distributed with center μ and variance

$$
\mathsf{Var}(\xi) = \mathsf{Var}[\mathbb{E}(\xi|\widetilde{\mu})] + \mathbb{E}[\mathsf{Var}(\xi|\widetilde{\mu})] = \frac{\sigma^4}{\sigma^2+\nu^2} + \frac{1}{\sigma^{-2}+\nu^{-2}} = \sigma^2.
$$

Therefore,

$$
-c + \alpha(I) + (1-\alpha)(II) \leq -c + \mathbb{E}_{\widetilde{\mu}\sim\mathcal{N}(\mu,\sigma^2)}(e^{\widetilde{\mu}} - k)^+
$$

It is sufficient to prove

$$
g^*(s) \geq -c + \mathbb{E}_{\widetilde{\mu}\sim\mathcal{N}(\mu,\sigma^2)} g^*(\widetilde{\mu}, \frac{1-Z}{\sigma^{-2}+\nu^{-2}})
$$

when $\sigma^2 \leq \frac{2\pi c^2}{k^2}$. This conclusion is given in the lemma A8.

To conclude, $\mathcal{V}^{(t+1)}(s) = g^*(s)$ holds for any $s \in \mathbb{R}\times\Sigma_0$. By mathematical induction, $\mathcal{V}^{(t)}(s) = g^*(s)$ holds for any $t \in \mathbb{N}_0$ for any $s \in \mathbb{R}\times\Sigma_0$. By corollary 7.2.13 of Puterman (1994), $\mathcal{V}(s) = \lim_{t\to\infty}\mathcal{V}^{(t)}(s) = g^*(s)$ for any $s \in \mathbb{R}\times\Sigma_0$. □

Proof of Lemma A2, Lemma A3 and Proposition A4 . Let $\mathcal{V}^{(0)}(s) = g^*(s) = (e^{\mu+\frac{\sigma^2}{2}} - k)^+$ and $\mathcal{V}^{(t)} := B(\mathcal{V}^{(t-1)})$ for $t \in \mathbb{N}_+$. I use mathematical induction to show that $\mathcal{V}^{(t)}(s)$

(i) increases in α;

(ii) for any $s = (\mu, \sigma) \in \mathcal{S}$, $\mathbb{E}_{\tilde{\mu} \sim \mathcal{N}(\mu, \frac{\sigma^4}{\sigma^2+\nu^2})}[\mathcal{V}^{(t)}(\tilde{\mu}, \tilde{\sigma}^2)] \leq \mathbb{E}_{\tilde{\mu} \sim \mathcal{N}(\mu, \sigma^2)}[g^*(\tilde{\mu}, 0)]$.

First, I show that (i) and (ii) hold for $\mathcal{V}^{(1)}$. By definition, for a fixed $s \in \mathcal{S}$,

$$\mathcal{V}^{(1)}(\mu, \sigma) := \left\{ \max \begin{cases} g^\star(s), & a = Q \\ -c + \mathbb{E}_s[\mathcal{V}^{(0)}(\tilde{\mu}, \tilde{\sigma})], & a = C \end{cases} \right\},$$

where

$$\mathbb{E}_s[\mathcal{V}^{(0)}(\tilde{\mu}, \tilde{\sigma})] = \alpha \underbrace{\mathbb{E}_{\tilde{\mu} \sim \mathcal{N}(\mu, \sigma^2)} g^*(\tilde{\mu}, 0)}_{(I)} + (1-\alpha) \underbrace{\mathbb{E}_{\tilde{\mu} \sim \mathcal{N}(\mu, \frac{\sigma^4}{\sigma^2+\nu^2})} g^*(\tilde{\mu}, \frac{1}{\sigma^{-2}+\nu^{-2}})}_{(II)}$$

(I) $\geq$ (II) by Jensen's inequality. Therefore, $\mathbb{E}_s[\mathcal{V}^{(0)}(\tilde{\mu}, \tilde{\sigma})]$ increases in α. Since $\mathcal{V}^{(1)}(\mu, \sigma)$ increases in $\mathbb{E}_s[\mathcal{V}^{(0)}(\tilde{\mu}, \tilde{\sigma})]$, it increases in α as well. Notice that

$$\begin{aligned} \mathcal{V}^{(1)}(\mu, \sigma) &= \max\{g^*(s), -c + \mathbb{E}_s[\mathcal{V}^{(0)}(\tilde{\mu}, \tilde{\sigma})]\} \\ &\leq \max\{g^*(s), \alpha \mathbb{E}_{\tilde{\mu} \sim \mathcal{N}(\mu, \sigma^2)} g^*(\tilde{\mu}, 0) + (1-\alpha) \mathbb{E}_{\tilde{\mu} \sim \mathcal{N}(\mu, \frac{\sigma^4}{\sigma^2+\nu^2})} g^*(\tilde{\mu}, \frac{1}{\sigma^{-2}+\nu^{-2}})\} \\ &\leq \max\{g^*(s), \mathbb{E}_{\tilde{\mu} \sim \mathcal{N}(\mu, \sigma^2)} g^*(\tilde{\mu}, 0)\} \\ &\leq \mathbb{E}_{\tilde{\mu} \sim \mathcal{N}(\mu, \sigma^2)} g^*(\tilde{\mu}, 0) \end{aligned}$$

where the last inequality follows Jensen's inequality. Therefore,

$$\mathbb{E}_{\tilde{\tilde{s}}}\{\mathcal{V}^{(1)}(\mu, \sigma^2)\} \leq \mathbb{E}_{\tilde{\tilde{s}}}\{\mathbb{E}_{\tilde{\mu} \sim \mathcal{N}(\mu, \sigma^2)} g^*(\tilde{\mu}, 0) | (\mu, \sigma)\} = \mathbb{E}_{\tilde{\tilde{s}}}\{g^*(\mu, 0)\}.$$

In sum, $\mathcal{V}^{(1)}(s)$ satisfies properties (i) and (ii).

Next, I show that $\mathcal{V}^{(t)}$ satisfies properties (i) and (ii) if they apply to $\mathcal{V}^{(t-1)}$. By definition, for a fixed $s \in \mathcal{S}$,

$$\mathcal{V}^{(t)}(\mu, \sigma) := \left\{ \max \begin{cases} g^\star(s), & a = Q \\ -c + \mathbb{E}_s[\mathcal{V}^{(t-1)}(\tilde{\mu}, \tilde{\sigma})], & a = C \end{cases} \right\},$$

where

$$\mathbb{E}_s[\mathcal{V}^{(t-1)}(\tilde{\mu},\tilde{\sigma})] = \alpha \underbrace{\mathbb{E}_{\tilde{\mu}\sim\mathcal{N}(\mu,\sigma^2)}[\mathcal{V}^{(t-1)}(\tilde{\mu},0)]}_{(I)} + (1-\alpha) \underbrace{\mathbb{E}_{\tilde{\mu}\sim\mathcal{N}(\mu,\frac{\sigma^4}{\sigma^2+\nu^2})}[\mathcal{V}^{(t-1)}(\tilde{\mu},\tilde{\sigma}^2)]}_{(II)}$$

and $\tilde{\sigma}^2 = \frac{1}{\sigma^{-2}+\nu^{-2}}$. Therefore, we have

$$\begin{aligned}
\frac{d}{d\alpha}\mathbb{E}_s[\mathcal{V}^{(t-1)}(\tilde{\mu},\tilde{\sigma})] &= (I) - (II) + \frac{d}{d\alpha}\mathbb{E}_s[\mathcal{V}^{(t-1)}(\tilde{\mu},\tilde{\sigma})] \\
&= (I) - (II) + \mathbb{E}_s[\frac{d}{d\alpha}\mathcal{V}^{(t-1)}(\tilde{\mu},\tilde{\sigma})] \\
&\geq 0
\end{aligned}$$

where the equality follows Danskin's Theorem and the inequality follows the assumption that property (i) and (ii) apply to $\mathcal{V}^{(t-1)}$. Since $\mathcal{V}^{(t)}$ is an increasing function of $\mathbb{E}_s[\mathcal{V}^{(t-1)}(\tilde{\mu},\tilde{\sigma})]$, it increases in α as well. Notice that

$$\begin{aligned}
\mathcal{V}^{(t)}(\mu,\sigma) =& \max\{g^*(s), -c + \mathbb{E}_s[\mathcal{V}^{(t-1)}(\tilde{\mu},\tilde{\sigma})]\} \\
\leq& \max\{g^*(s), \alpha(I) + (1-\alpha)(II)\} \\
\leq& \max\{g^*(s), (I)\} \\
\leq& (I) = \mathbb{E}_{\tilde{\mu}\sim\mathcal{N}(\mu,\sigma^2)}[\mathcal{V}^{(t-1)}(\tilde{\mu},0)]
\end{aligned}$$

where the last inequality follows Jensen's inequality. Therefore,

$$\mathbb{E}_{\tilde{\tilde{s}}}\{\mathcal{V}^{(t)}(\mu,\sigma^2)\} \leq \mathbb{E}_{\tilde{\tilde{s}}}\{\mathbb{E}_{\tilde{\mu}\sim\mathcal{N}(\mu,\sigma^2)}[\mathcal{V}^{(t-1)}(\tilde{\mu},0)|(\mu,\sigma)]\} \leq \mathbb{E}_{\tilde{\tilde{s}}}[\mathcal{V}^{(t-1)}(\mu,0)] = \mathbb{E}_{\tilde{\tilde{s}}}\{g^*(\mu,0)\}.$$

In sum, $\mathcal{V}^{(t)}$ satisfies property (i) and (ii) if they apply to $\mathcal{V}^{(t-1)}$.

By mathematical induction, (i) and (ii) if they apply to $\mathcal{V}^{(t)}$ where $t \in \mathbb{N}_+$. Notice that $0 \leq \mathcal{V}^{(0)} \leq \mathcal{V}(s)$. By Puterman (1994), $\mathcal{V}(s) = \lim_{t\to\infty}\mathcal{V}^{(t)}(s)$. Therefore, $\mathcal{V}(s)$ satisfies property (i) and (ii). □

Proof of Proposition A2. By Lemma A2., $\mathbb{E}[\mathcal{V}(\tilde{S})] \leq \mathbb{E}_s g^*(\tilde{\mu},0)$. Therefore, a sufficient condition for the startup to quit is $g^*(\mu,\sigma) > -c + \mathbb{E}_{\tilde{\mu}\sim\mathcal{N}(\mu,\sigma^2)} g^*(\tilde{\mu},0)$, which is equivalent to

$$f(s) := \mathbb{E}_{\tilde{\mu}\sim\mathcal{N}(\mu,\sigma^2)} g^*(\tilde{\mu},0) - g^*(\mu,\sigma) \tag{20}$$

$$= \mathbb{E}[g(\mu+\sigma\zeta,0)] - g^*(\mu,\sigma) < c \tag{21}$$

where $\zeta \sim \mathbb{N}(0,1)$. First, we have

$$\begin{aligned}
&\frac{\partial}{\partial\mu}\mathbb{E}[g^\star(\mu+\sigma\zeta,0)] = \mathbb{E}\left\{\frac{\partial}{\partial\mu}\left(e^{\mu+\sigma\zeta}-k\right)\mathbb{I}\left(e^{\mu+\sigma\zeta}-k\geq 0\right)\right\} \quad \text{(by Danskin's Theorem)}\\
&= e^{\mu}\mathbb{E}\left\{e^{\sigma\zeta}\mathbb{I}\left(-\zeta \leq \frac{\mu-\log(k)}{\sigma}\right)\right\}\\
&= e^{\mu+\sigma^2/2}\Phi\left(\frac{\mu-\log(k)}{\sigma}+\sigma\right).
\end{aligned}$$

where the last equality follows from the moment generating function of the truncated normal distribution. Second, we have

$$\frac{\partial}{\partial\mu}g^\star(\mu,\sigma^2) = \frac{\partial}{\partial\mu}\left(e^{\mu+\sigma^2/2}-k\right)\mathbb{I}\left(e^{\mu+\sigma^2/2}-k\geq 0\right) = e^{\mu+\sigma^2/2}\mathbb{I}\left(\mu \geq \log(k)-\frac{\sigma^2}{2}\right).$$

Then for $\mu < log(k) - \frac{\sigma^2}{2}$, we have

$$\frac{\partial}{\partial\mu}f(s) = e^{\mu+\sigma^2/2}\Phi\left(\frac{\mu-\log(k)}{\sigma}+\sigma\right) > 0.$$

For $\mu > log(k) - \frac{\sigma^2}{2}$, we have

$$\frac{\partial}{\partial\mu}f(s) = e^{\mu+\sigma^2/2}\Phi\left(\frac{\mu-\log(k)}{\sigma}+\sigma\right) - e^{\mu+\sigma^2/2} < 0.$$

So $\max_\mu f(s) = f(log(k) - \frac{\sigma^2}{2}, \sigma)$. Also, notice that $lim_{\mu\to\infty} f(s) = 0$. There are two cases as below. If $\max\{f(s)\} < c$, the proposition holds for any $\underline{\mu} = \bar{\mu}$ in $\mathbb{R}$. If $\max\{f(s)\} > c$, there exists $\underline{\mu} < log(k) - \frac{\sigma^2}{2}$ and $\bar{\mu} > log(k) - \frac{\sigma^2}{2}$ such that $f(\underline{\mu},\sigma) = f(\bar{\mu},\sigma) = c$. Then we have $f(s) < c$ if $\mu < \underline{\mu}$ or $\mu > \bar{\mu}$. □

Proof of Proposition A3. Let $\mathcal{V}^{(0)}(s) = g^*(s) = (e^{\mu+\frac{\sigma^2}{2}} - k)^+$ and $\mathcal{V}^{(t)} := B(\mathcal{V}^{(t-1)})$ for $t \in \mathbb{N}_+$. I use mathematical induction to show that $\mathcal{V}^{(t)}(s)$ is convex in μ and increasing in σ for $t \in \mathbb{N}_+$

and $s \in \mathcal{S}$.

First, notice that $\mathcal{V}^{(0)}(s)$ is convex in μ and increasing in σ for $t \in \mathbb{N}_+$ and $s \in \mathcal{S}$. Next, I show that $\mathcal{V}^{(t)}$ is convex in μ and increasing in σ assuming they hold for $\mathcal{V}^{(t-1)}$. By definition, for a fixed $s \in \mathcal{S}$,

$$\mathcal{V}^{(t)}(\mu, \sigma) := \left\{ \max \begin{cases} g^{\star}(s), & a = Q \\ -c + \mathbb{E}_s[\mathcal{V}^{(t-1)}(\tilde{\mu}, \tilde{\sigma})], & a = C \end{cases} \right\},$$

where

$$\begin{aligned} \mathbb{E}_s[\mathcal{V}^{(t-1)}(\tilde{\mu}, \tilde{\sigma})] &= \alpha \mathbb{E}_{\tilde{\mu} \sim \mathcal{N}(\mu, \sigma^2)}[\mathcal{V}^{(t-1)}(\tilde{\mu}, 0)] + (1-\alpha)\mathbb{E}_{\tilde{\mu} \sim \mathcal{N}(\mu, \frac{\sigma^4}{\sigma^2+\nu^2})}[\mathcal{V}^{(t-1)}(\tilde{\mu}, \tilde{\sigma}^2)] \\ &= \alpha \underbrace{\mathbb{E}_{\epsilon \sim \mathcal{N}(0,\sigma^2)}[\mathcal{V}^{(t-1)}(\mu+\epsilon, 0)]}_{(I)} + (1-\alpha) \underbrace{\mathbb{E}_{\epsilon \sim \mathcal{N}(0, \frac{\sigma^4}{\sigma^2+\nu^2})}[\mathcal{V}^{(t-1)}(\mu+\epsilon, \tilde{\sigma}^2)]}_{(II)} \end{aligned}$$

and $\tilde{\sigma}^2 = \frac{1}{\sigma^{-2}+\nu^{-2}}$. By Lemma A9 and $\mathcal{V}^{(t-1)}(s)$ is convex in μ, (I) and (II) are also convex in μ. By $V^{(t-1)}(\mu+\epsilon, 0)$ is convex in ϵ and Lemma 6, we have (I) increases in σ. Given any $\sigma_1 < \sigma_2$, we have

$$\begin{aligned} (II)(\sigma_1) = \mathbb{E}_{\epsilon \sim \mathcal{N}(0, \frac{\sigma_1^4}{\sigma_1^2+\nu^2})}[\mathcal{V}^{(t-1)}(\mu+\epsilon, \tilde{\sigma}^2(\sigma_1))] &\leq \mathbb{E}_{\epsilon \sim \mathcal{N}(0, \frac{\sigma_1^4}{\sigma_1^2+\nu^2})}[\mathcal{V}^{(t-1)}(\mu+\epsilon, \tilde{\sigma}^2(\sigma_2))] && (22) \\ &\leq \mathbb{E}_{\epsilon \sim \mathcal{N}(0, \frac{\sigma_2^4}{\sigma_2^2+\nu^2})}[\mathcal{V}^{(t-1)}(\mu+\epsilon, \tilde{\sigma}^2(\sigma_2))] && (23) \\ &= (II)(\sigma_2) && (24) \end{aligned}$$

where the first inequality follows our assumption that $\mathcal{V}^{(t-1)}$ increases in σ and the second inequality follows the feature of second order stochastic dominance. Therefore, (II) also increases in σ. Given that (I) and (II) are convex in μ and increasing in σ, we conclude that $\mathbb{E}_s[\mathcal{V}^{(t-1)}(\tilde{\mu}, \tilde{\sigma})]$ satisfy both properties as well. Notice that $g^{*}(s)$ is also convex in μ and increasing in σ. Given $\mathcal{V}^{(t)}(s)$ is increasing and convex in $\mathbb{E}_s[\mathcal{V}^{(t-1)}(\tilde{\mu}, \tilde{\sigma})]$, which is convex in μ, $\mathcal{V}^{(t)}(s)$ is also convex in μ. By the feature of second order stochastic dominance, $\mathcal{V}^{(t)}(s)$ increases in σ. The completes the proof that $\mathcal{V}^{(t)}$ is convex in μ and increasing in σ assuming they hold for $\mathcal{V}^{(t-1)}$. By mathematical induction, for any $t \in \mathbb{N}_+$ and $s \in \mathcal{S}$, $\mathcal{V}^{(t)}$ is convex in μ and increasing in σ.

Notice that $0 \leq \mathcal{V}^{(0)} \leq \mathcal{V}(s)$. By Puterman (1994), $\mathcal{V}(s) = \lim_{t\to\infty} \mathcal{V}^{(t)}(s)$. Therefore, $\mathcal{V}(s)$ is convex in μ and increasing in σ. □

H.5.3 Additional Lemmas

Lemma A8. *Suppose $Z \sim \mathcal{N}(\mu, \sigma^2)$. If $\sigma^2 \leq \frac{2\pi c^2}{k^2}$, then*

$$\mathbb{E}(e^Z - k)^+ - (\mathbb{E}e^Z - k)^+ \leq c.$$

Proof of Lemma A8. Notice that

$$\begin{aligned}
[\mathbb{E}e^Z - k]^+ &= \left(e^{\mu+\sigma^2/2} - k\right)^+; \\
\mathbb{E}(e^Z - k)^+ &= \mathbb{P}[Z \geq \log(k)]\big\{\mathbb{E}_t[e^Z | Z \geq \log(k)] - k\big\} \\
&= \Phi\left(\frac{\mu - \log(k)}{\sigma} + \sigma\right) e^{\mu+\sigma^2/2} - \Phi\left(\frac{\mu - \log(k)}{\sigma}\right) k \\
&= \Phi\left(\frac{\mu - \log(k)}{\sigma} + \sigma\right)\left(e^{\mu+\sigma^2/2} - k\right) + \left[\Phi\left(\frac{\mu - \log(k)}{\sigma} + \sigma\right) - \Phi\left(\frac{\mu - \log(k)}{\sigma}\right)\right] k \\
&\leq \Phi\left(\frac{\mu - \log(k)}{\sigma} + \sigma\right)\left(e^{\mu+\sigma^2/2} - k\right) + \frac{\sigma k}{\sqrt{2\pi}}.
\end{aligned}$$

In particular, $\left(e^{\mu+\sigma^2/2} - k\right)^+ - \Phi\left(\frac{\mu-\log(k)}{\sigma} + \sigma\right)\left(e^{\mu+\sigma^2/2} - k\right) \geq 0$. Then

$$(\mathbb{E}e^Z - k)^+ - \mathbb{E}(e^Z - k)^+ + c \geq -\frac{\sigma k}{\sqrt{2\pi}} + c.$$

The right hand side ≥ 0 if and only if $\sigma^2 \leq \frac{2\pi c^2}{k^2}$. □

Lemma A9. *Assume $f(x)$ is a convex function of $x \in \mathbb{R}$. Then $g(x) = \mathbb{E}[f(x + \epsilon)]$ where $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is also convex in x.*

Proof of Lemma A9. For any $x_1, x_2 \in \mathbb{R}$ such that $x_1 \neq x_2$, we have

$$\begin{aligned}
g(tx_1 + (1-t)x_2) &= \mathbb{E}[f(tx_1 + (1-t)x_2 + \epsilon)] \\
&\leq \mathbb{E}[f(tx_1) + f((1-t)x_2 + \epsilon)] \\
&= f(tx_1) + \mathbb{E}[f((1-t)x_2 + \epsilon)] \\
&= f(\mathbb{E}[tx_1 + \epsilon]) + \mathbb{E}[f((1-t)x_2 + \epsilon)] \\
&\leq \mathbb{E}[f(tx_1 + \epsilon)] + \mathbb{E}[f((1-t)x_2 + \epsilon)] \\
&= g(tx_1) + g((1-t)x_2)
\end{aligned}$$

where the two inequalities follow the assumption that $f(x)$ is a convex function. Therefore, by definition of convex function, $g(x)$ is convex in x. □

H.6 Numerical Analysis and Simulation

This appendix illustrates the procedure of solving equation (17) and simulating results on the benefits of VC informativeness.

H.6.1 Numerical Analysis

First, I show the procedure of numerically solving $\mathcal{V}(s)$ and $d^*(s)$ for any $s = (\mu, \sigma) \in \mathcal{S}$. For each time point t, σ can take one of the two values depending on the realization of $Z_1, ..., Z_t$. If at least one of $Z_1, ..., Z_t$ equals zero (the startup receives at least one precise signal), we have $\sigma = 0$. The startup takes an action of $a = Q$ and obtains $g^*(s) = (e^\mu - k)^+$ according to Corollary A1. If, instead, $Z_1 = ... = Z_t = 1$ (the startup didn't receive any precise signal before t), we have σ expressed by

$$\sigma = \sqrt{\frac{1}{\sigma_0^{-2} + t\nu^{-2}}}, \tag{25}$$

under which case $\mathcal{V}(s)$ and $d^*(s)$ are determined by the procedure below.

I set up a two-dimensional matrix where the each element corresponds to a possible combination of μ and t. Notice that the value of σ for each element can be determined by equation (25).

Our goal is to compute $\mathcal{V}(s)$ and $d^*(s)$ for each $s = (\mu, \sigma)$ in the matrix. Recall that the startup quits when σ is sufficiently small by Proposition A2 and Corollary A1. Since σ monotonically decreases in t, there exists $t = T_{max}$ such that the startup quits before T_{max}. I start by assuming the value of T_{max} and use backward induction to compute $\mathcal{V}(s)$ and $d^*(s)$ of all possible states for $t < T_{max}$. In particular, I solve the model follow the steps as below.

Step 1. Set up values of exogenous parameters. I manually assign values to five exogenous parameters in the model. I first select the value of $\mu_0 = 0$, $\sigma_0 = 1$ and $\nu = 1$, and then determine the value of k which satisfies the assumption that $\mathbb{E}(\theta) = \exp\{\mu_0 + \frac{\sigma_0^2}{2}\} < k$. I select $k = 1.8$. The cost of one signal c is determined by trial-and-error. The assumption is that c is sufficiently small so that it does not prevent the startup from collecting the first signal. I choose $c = 0.05$ that satisfies this assumption.

Step 2. Set up ranges of state variables. The range of μ is from -2.8 to 2.8 to cover 99.5% of unconditional probability, where grid is set to be $w = 0.05$. I denote all possible values of μ as μ_k, where $k \in \{1, ..., k_{max}\}$ where $k_{max} = 113$. The maximum time T_{max} is determined by trial-and-error and I select $T_{max} = 50$.[44] The logic behind the trial-and-error is that selecting a larger T_{max} should not cause a change of optimal policy.[45] The value of σ at T_{max}, namely σ_{max}, is determined by equation (25). For any μ_k, I let $\mathcal{V}(\mu_k, \sigma_{max}) = g^*(\mu_k, \sigma_{max})$ and $d^*(\mu_k, \sigma_{max}) = Q$.

Step 3. Compute the transition matrix. The transition matrix is determined by the discretization of equation (18) conditional on $Z = 1$. The transition probability from μ_j to μ_i for a given σ is determined by

$$
\begin{aligned}
p_{ij} =& \mathbb{P}(\tilde{\mu} \in (\mu_i - \frac{w}{2}, \mu_i + \frac{w}{2}) | \mu \in (\mu_j - \frac{w}{2}, \mu_j + \frac{w}{2})) \\
\approx& \mathbb{P}(\tilde{\mu} \in (\mu_i - \frac{w}{2}, \mu_i + \frac{w}{2}) | \mu = \mu_j) \\
=& \Phi(\frac{\mu_i + \frac{w}{2} - \mu_j}{\sqrt{\frac{\sigma^4}{\sigma^2+\nu^2}}}) - \Phi(\frac{\mu_i - \frac{w}{2} - \mu_j}{\sqrt{\frac{\sigma^4}{\sigma^2+\nu^2}}}).
\end{aligned}
$$

[44] T_{max} can also be determined given the upper bar of σ given by Corollary A1 and equation (25). Here I select a smaller value to improve computational efficiency.

[45] Note that the choices of T_{max} and c may affect each other.

where Φ is CDF of standard normal distribution and $1 < i, j < k_{max}$. In addition, we have

$$p_{1j} = \mathbb{P}[\tilde{\mu} \in (-\infty, \mu_1 + \frac{w}{2})|\mu = \mu_j],$$
$$p_{k_{max}j} = \mathbb{P}[\tilde{\mu} \in (\mu_{k_{max}} - \frac{w}{2}, \infty)|\mu = \mu_j].$$

Step 4. Solve dynamic programming problem by backward induction. Let $t = T_{max} - 1$, I determine σ according to equation (25). Then for any μ_k, we have

$$\mathcal{V}(\mu_k, \sigma) = \max\{g^*(\mu_k, \sigma), -c + \alpha h(\mu_k, \sigma) + (1-\alpha)\Sigma_{i=1}^{k_{max}} p_{ik} \mathcal{V}(\mu_i, \sigma_{max})\},$$

where

$$\begin{aligned} h(\mu_k, \sigma) &= \mathbb{E}_{log(x) \sim \mathcal{N}(\mu_k, \sigma^2)}(x-k)^+ \\ &= \exp\{\mu_k + \frac{\sigma^2}{2}\}\Phi(\frac{\mu_k + \sigma^2 - \log(k)}{\sigma}) - k(1 - \Phi(\frac{\log(k) - \mu_k}{\sigma})). \end{aligned}$$

is the expected payoff conditional on having a precise signal at T_{max}. The optimal policy is

$$d(\mu_k, \sigma) = \begin{cases} Q, & \text{if } g^*(\mu_k, \sigma) \geq -c + \alpha h(\mu_k, \sigma) + (1-\alpha)\Sigma_{i=1}^{k_{max}} p_{ik} \mathcal{V}(\mu_i, \sigma_{max}); \\ C, & \text{Otherwise.} \end{cases}$$

Then I repeat the procedure to compute $\mathcal{V}(s)$ and $d^*(s)$ for any μ_k and and σ that satisfies $t \in \{1, 2, ..., T_{max} - 2\}$.

H.6.2 Simulation

For a given $\alpha \in [0, 1]$, I simulate the model by drawing $n_1 = 1,000$ value of θ where $log(\theta) \sim \mathcal{N}(\mu_0, \sigma_0^2)$ and running the process for $n_2 = 10,000$ iterations. For each iteration i of each value θ_j at each time $t \geq 1$,

- If $Z_t = 0$, the startup quits and obtains a payoff of $g_{ij} = g^*(log(\theta_j), 0)$; the market requires an ownership of $\lambda_{ij} = k/\theta_j$ if $g_{ij} > 0$;

- If $Z_t = 1$ and the startup receives a signal x_t, the startup updates belief according to

$$\mu_t = \frac{\nu^2}{\sigma_{t-1}^2 + \nu^2}\mu_{t-1} + \frac{\sigma_{t-1}^2}{\sigma_{t-1}^2 + \nu^2}x_t$$

and σ_{t-1} by equation (25) at $t-1$. Denote $\mu^* = \arg\min_{\mu_k \in \{\mu_1, \ldots, \mu_{max}\}} |\mu_k - \mu_t|$, and σ given by equation (25), the startup takes action $d(\mu^*, \sigma)$. If $d(\mu^*, \sigma) = Q$, the startup quits with a payoff $g_{ij} = g^*(\mu_t, \sigma)$; and the market requires an ownership of $\lambda_{ij} = k/\exp\{\mu_t + \frac{\sigma_t^2}{2}\}$ if $g_{ij} > 0$. Otherwise, the startup moves to $t+1$.

For each iteration i of value θ_j, I document (i) the final payoff g_{ij}; (ii) stopping time t_{ij}; (iii) the value of Z at the stopping time (Z_{ij}); (iv) a dummy variable $I_{ij} = \mathbb{I}_{g_{ij}>0}$, which indicates whether an investment is made; (v) the ownership fraction required by VCs λ_{ij}. I utilize those values to plot the figures below.

- Figure A6 the probability of resolving true value in the last stage

$$\frac{1}{n_1 n_2}\Sigma_i \Sigma_j (1 - Z_{ij}).$$

- Figure A8 plots expected signal collection time

$$\frac{1}{n_1 n_2}\Sigma_i \Sigma_j t_{ij}.$$

- Figure A7 plots the probability of investment

$$\frac{1}{n_2}\Sigma_i I_{ij}.$$

and the expected ownership required by VC conditional on investment

$$\frac{\Sigma_i \lambda_{ij} I_{ij}}{\Sigma_j I_{ij}}$$

for θ_j^* that satisfies $\theta_j^* = \arg\min_{\theta_j} |\theta_j - 2|$ and $\theta_j^* = \arg\min_{\theta_j} |\theta_j - 5|$, respectively.