LIFE CYCLE OF STARTUP FINANCING

SVETLANA BOYARCHENKO*

2225 Speedway Stop C3100, Austin, TX 78712, U.S.A.
e-mail: sboyarch@utexas.edu

Abstract. I characterize an optimal, incentive compatible, and renegotiation proof contract of venture capital (VC) financing of a startup (that may be successful or not) whose rate of arrival of success is a function of the accumulated investment stock. The contract depends on the startup valuation, prior probability of success, and initial capital. Sufficient conditions for existence of such a contract are specified. The paper explains why the startup has to rely on different ways of financing in different stages of its life, and why VC financing is not feasible in early stages of development of the startup.

Keywords: startup financing, venture capital, hump-shaped distributions

JEL: C73, C61, D81, D86

1. Introduction

Startups are small companies, but they can play a significant role in an economy by creating jobs, spurring innovation and injecting competition. In different stages of their lives, startups depend on different modes of financing. Economic literature concentrates either on relatively late stages of development of startups when VC contracting and financing becomes possible, or on early stages when a startup has to raise money through crowdfunding. I present a model of a risky startup financing that explains why the startup has to rely on different ways of financing in different stages of its life and why VC financing is not possible until the startup becomes sufficiently profitable (in expected terms). To the best of my knowledge, this is the first paper that explains a life cycle of a startup financing.

I am thankful for useful comments and discussions of an earlier version of the paper to participants of The Econometric Society's 2020 World Congress, Bocconi University, Milan, Italy, Aug. 17-21, 2020 and Conference on Mechanism and Institution Design at the University of Klagenfurt, Germany, June 11-13, 2020. I am thankful for discussions of initial versions of the paper to Vasiliki Skreta, Dale Stalh, Max Stinchcombe, Juuso Välimäki and participants of the 18th SAET Conference on Current Trends in Economics, June 11-13, 2018, Taipei, Taiwan.

1.1. **Motivation.** The history of VC investment can be traced back as far as the financing of Columbus' expedition by the Spanish monarchy. Despite the fact that histories of successful startups are quite individual, they share at least two common features: at the very beginning, startups had to rely on family funding, seed funding[1], crowdfunding, or government money (SBA loans or SBIR grants). VC financing typically does not happen in early stages of a startup development (see Fig. 1 for illustration).

For example, the initial startup capital for Amazon.com came primarily from life savings of Jeff Bezos' parents in 1994. Amazon raised a series A[2] of \$8 million from Kleiner Perkins Caufield & Byers in 1995. In 1997, Amazon went public to raise additional capital. Uber was launched in August 2009 by seed funding. In October 2010, it switched to angel investment, that is a form of investment in exchange of convertible debt or ownership equity. In February 2011, Uber went through series A financing. This was followed many other different ways of financing until Uber went public in May 2019. Peloton, the company that sells internet-connected exercise bikes was founded in 2012, and during the first three years, the founder and his team resorted to angel financing because VCs turned their backs on the company despite of its evident success. In 2018, Peloton raised \$550 million in VC funding at a \$4.1 billion valuation; the business went public in August 2019.

1.2. **Related literature.** Examples presented in the previous Section demonstrate that all types of financing play important roles in the startup dynamics, but the literature in economics and finance usually concentrates on one particular stage of financing. There is a large and well-developed literature that provides important insights for VC financing and risks associated with entrepreneurial projects due to huge uncertainty about their outcomes, information asymmetries and lack of tangible collaterals. When a research project is started, nobody knows in advance which of raw ideas (if any) will generate a success and when. Therefore, models that exploit random times of arrival of success are most suitable for studying sponsored research. There is a large variety of models that study incentives for and delegation of experimentation, means of control, and information dissemination (see, e.g., Bergemann and Hege (1998, 2005), Bonatti and Hörner (2011, 2017), Grenadier et al. (2016), Guo (2016), Halac et al. (2016, 2017, 2020), Halac and Kremer (2020), Hörner and Samuelson (2013). All the aforementioned papers assume that breakthroughs of a good project arrive at jump times of the standard Poisson process. Unfortunately, models of innovation financing that use exponentially

[1]Seed funding is a form of securities offering in which an investor invests capital in exchange for an equity stake in the company.

[2]Series A is the first significant round of venture capital financing, when a preferred stock is sold to investors in exchange for their investment.

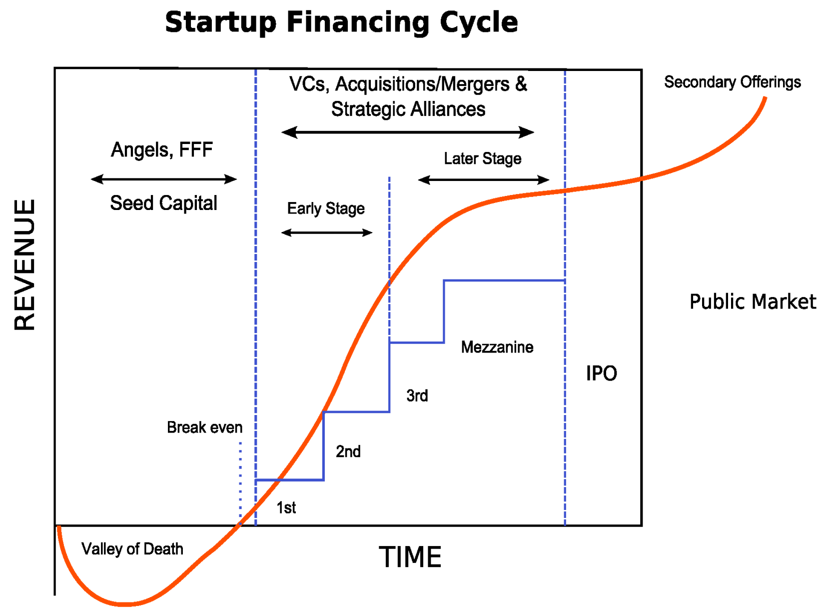


FIGURE 1. Source: Wikipedia https://commons.wikimedia.org/wiki/File:Startup_financing_cycle.svg

distributed arrival time of the first success do not capture the fact that success may be related to accumulated investment funds, despite the fact that capital constraints are usually regarded as the main obstacle to entrepreneurship.

The emergence of new sources of financing in the aftermath of the financial crisis has substantially increased the funding options available to new entrepreneurial ventures. As a result, there is a growing literature on crowdfunding (see, e.g., Strausz (2017) or Deb et al. (2019) and references therein). For example, Deb et al. (2019) consider a new class of dynamic contribution games. Halac et al. (2020) consider a model of funding a project by multiple investors in case when the investors are facing a random threshold that triggers the success of the project. Models of crowdfunding emphasize the fact that an innovative project cannot be successful unless a target level of pledge is reached, but do not consider the case when a fully funded research idea may be erroneous, and therefore unsuccessful.

1.3. **Contribution and main results.** The model presented in this paper has two novel features. First, as opposed to models of innovation financing that use exponentially distributed arrival time of the first success (see, for example, Bergemann and Hege (2005) or Hörner and Samuelson (2013)), the rate of arrival of success in my model depends on the stock of capital invested in the startup - it is increasing in the investment capital in early stages of life of the startup until it reaches a certain maximal level; after that point, the arrival rate decreases as the startup grows older (provided no success happened earlier).

Next, as opposed to the literature on crowd or multiple investors funding (see, for example, Strausz (2017), Deb et al. (2019), or Halac et al. (2020)), the success of a startup in my model is not guaranteed at any level of investment. This reflects the fact that innovation involves the exploration of uncertain and unknown terrain, so even some of the most highly touted innovators have had their share of failures (recall, for instance, Apple's MobileMe, Google Glass, and the Amazon Fire Phone).

The contribution of my paper is fourfold. First, my model not only demonstrates that VC investment into startups is not possible until a startup accumulates sufficient funding through other sources of financing but also explains why different kinds of sponsor-entrepreneur relations may be optimal in different stages.

Second, I study the optimal provision of incentives in a dynamic agency model. I consider startups whose valuations are high enough to generate positive expected gains in a certain region of the state space (otherwise, no kind of financing is possible), the boundaries of the region are the points where the startup breaks even. It is not socially optimal to continue financing of the startup after the right boundary was reached, because the startup will generate nothing but increasing losses if financing goes on. Before the left breakeven point is reached, the startup is in the "death valley" zone - it is making losses, but the losses become smaller as more capital is invested in the startup because the rate of arrival of success is increasing. I show that the socially optimal financing starts in the "death valley" zone, and the financing happens at the highest feasible rate. The main intuition behind the social optimality of the maximal rate financing is that it is better to pass through the "death valley" zone as fast as possible thereby minimizing the discount factor in the region where the instantaneous expected payoff is positive.

Similar consideration would have been true in case of VC financing were it not for a moral hazard problem - the entrepreneur may have incentives to divert invested money to alternative projects or personal use. Therefore, financing the startup at the highest possible rate is not always incentive compatible. Only high valuation startups get maximal rate VC financing, and this happens no earlier than they used alternative methods of financing to accumulate capital sufficient to leave the "death valley" zone, which is consistent with the typical picture of startup development in Fig. 1. Moreover, due to incentive compatibility restrictions, VC financing stops earlier than it is socially optimal.

Following Bergemann and Hege (2005), I consider two modes of VC financing - relationship financing (RF) with the entrepreneur's allocation of investment funds being observable; and arm's length financing with unobservable allocation of the capital. I show that for the same set of parameters, the highest incentive compatible rate of financing is achieved with observable actions when the rate of arrival of success is upward sloping, and with unobservable actions when the rate of arrival is downward sloping. Therefore, the entrepreneur who seeks the maximal rate of funds disbursement will find it optimal to change the financing mode and switch from RF to ALF. This result is qualitatively different from Bergemann and Hege (2005), where arm's length financing may be better or worse than relationship financing depending on a set of parameters of their model. In my model the entrepreneur who has an opportunity to divert investment money for alternative purposes may prefer her investment strategy to be observable in relatively early stages of VC financing in order to get the highest incentive compatible rate of financing. Interestingly, in Hörner and Samuelson (2013), for some parameter values, the principal who provides money prefers not to observe how the agent allocates investment funds.

Unlike Bergemann and Hege (2005) or Hörner and Samuelson (2013), I do not assume that one of the parties makes an offer to the other party repeatedly. Instead, contract provisions such as the rate of funds disbursement, termination in case of no success, regions of (un)observable funds allocation or shares due to the parties in case of success are specified at the time of initiation of the contract. This, in particular, alleviates problems related to modeling negotiations in continuous time. The contract which I design is renegotiation proof. If at any investment level, the startup decides to terminate the contract and change the sponsor, the new contract will be the continuation of the old one starting at the investment level accumulated by the time of termination.

My paper demonstrates that the optimal incentive compatible contract depends on the startup valuation and capital accumulated before the contract initiation. In particular, high valuation startups need a moderate level of the initial capital stock to obtain VC financing at the highest feasible rate. Startups with moderate valuation can get the highest rate of VC financing with a sufficiently high initial level of capital accumulated by alternative means of financing. If a moderate valuation startup has the initial capital insufficient for the highest rate of financing, VC financing starts at a smaller rate which keeps increasing as more capital is invested into the project. At the point where the capital stock reaches the threshold sufficient for the highest rate financing, the rate of financing jumps to this highest possible level. The jump happens due to long term effects of one-shot deviations. Had there been no long term effects, the incentive compatible rate of financing would have approached the maximal rate continuously. I call the effect of financing at an increasing rate the project "upsizing." The project "upsizing" is another qualitatively new prediction of my model. In the absence of success, financing of high and moderate valuation startups ends in finite time and at the highest rate.

As opposed to these, low valuation startups never get the highest rate of VC financing. The rate of financing of low valuation startups is a continuous and non-monotone function of the accumulated investment level - it is increasing until the investment level become sufficiently close to the termination barrier and starts rapidly decreasing in a left neighborhood of the barrier. VC financing in this case goes on, at a vanishing rate, until (if ever) the first success is achieved.

As far as the existence of the optimal incentive compatible contract is concerned, I prove that whenever VC financing is possible, such a contract always exists for high and moderate valuation startups if their initial capital stock is sufficiently high. For low valuation or moderate valuation startups with low levels of the initial capital, the optimal rate of financing may be not incentive compatible, so the optimal incentive compatible contract does not exist. However, financing at a suboptimal rate is incentive compatible.

The model presented here can be modified to study other problems where learning about long term potential effects (say, of a medication or pollutant), hidden actions and lack of commitment are essential components. For example, the mirror reflection of the problem of the startup financing is an application to the feedback loop in interactions of industry and nature: accumulation of greenhouse gases (GHG) and pollutants increase the unknown rate of arrival of a drastic climate change.

The last two contributions of my paper are methodological ones. The first one is the choice of a state variable that allows me to simplify the study and formulation of the results of the paper significantly and reduce the initial non-Markovian setting to a Markovian one. I use as the state variable the amount of VC funds invested in the startup. The second methodological contribution is the use of calculus of variations in application to the policy functions of the entrepreneur. This simple tool allows me to represent the value function of the entrepreneur so that all possible deviations from the contracted investment strategy are accounted for. Linearization of the value function of the entrepreneur around the equilibrium path gives necessary conditions for the optimal incentive compatible contract. I use the quadratic approximation of the entrepreneur's value function to derive sufficient conditions.

The rest of the paper is organized as follows. Main objects and their properties are introduced in Section 2. The primitives of the model are the rate of arrival of success of the startup which is successful for sure and the prior probability of success of a given startup. As a benchmark, I consider the social planner's problem and demonstrate that there is a critical startup valuation such that startups with less than the critical valuation do not get the government funding unless they raise a certain amount of initial capital elsewhere. Startups with sufficiently high valuation get the government funding at the highest feasible rate of financing even while the startup is still in the "death valley" zone.

In Section 3, I describe a VC financing contract, participation and incentive constraints and provide informal characterization of an optimal contract. I show that the social planner's termination barrier is higher than in the case of VC financing. The trigger barrier for the government financing is lower than the one for VC financing. Furthermore, for some low valuation projects VC financing is not feasible but the government financing at the maximal rate may be optimal. In Section 4, the unique candidate for the optimal incentive compatible contract is constructed. Sufficient conditions for existence of the optimal incentive compatible contract are provided. Contract provisions depend on the startup valuation, the initial level of capital, and the model primitives. Main results in Sections 3 and 4 are derived for the case when the entrepreneur has full bargaining power. At the end of Section 4, I show that my model is robust if both the investor and the entrepreneur have some bargaining power. Section 5 concludes. Technical proofs are relegated to the appendix.

2. Economic environment and first best policy

2.1. **Basic objects and their properties.** As a benchmark model of innovations financing I use a continuous time modification of Bergemann and Hege (2005) (BH from now on). As in BH, I consider the financing of a risky startup (or a research project) under uncertainty about the time of completion and the probability of the eventual success. If the project is bad, it never generates any success. If the startup is good, it can generate at most one success which produces a fixed instantaneous payoff $D > 0$ that can be viewed as a proxy for the startup valuation. Let $r > 0$ be the discount rate, and let $\pi_0 \in [0, 1]$ be the prior probability of the eventual success of the startup.

Time $t \in \mathbb{R}_+$ is continuous[3]. If the startup is good, the success arrives at a random time $\tilde{t} > 0$, and the rate of arrival of success at any $t < \tilde{t}$ depends on how much capital was invested in the startup up to time t, which is denoted by $\Gamma(t)$. The function Γ is called the *financing policy.* Introduce $\mathcal{M}$ - the class of increasing functions $\Gamma : \mathbb{R}_+ \to \mathbb{R}_+$ such that the measure $d\Gamma(t) = \Gamma'(t)dt$ is absolutely continuous. Thus, $\Gamma'(t)$ is the rate of financing of the startup at t.

[3]BH set their model in discrete time. In continuous time, the counterpart of their model is the exponential distribution for the random time of arrival of the first breakthrough (provided the innovation is successful), the rate of arrival being controlled by the choice of the investment rate.

Assume that $0 < \Gamma'(t) \le 1$ for all t. Note that, in a different setting, $\Gamma(t)$ can also be viewed as the total amount of effort spent on the project up to time t. If $\Gamma'(t) \equiv 1$, then $\Gamma(t) = t$.

To specify the random time of success arrival, I use a model of experimentation (first published in Boyarchenko (2020) and used in Boyarchenko (2020, 2021) to study strategic experimentation) which is almost as tractable as popular models based on exponential distributions, produces qualitatively new results, and agrees better with the data than exponential models. The primitive of the model is the rate of arrival (hazard rate) of success at the investment level g when $\pi_0 = 1$. Denote this hazard rate by $\Lambda_0(g)$. If $\pi_0 = 1$, the probability of the event that no success is observed before the investment level g is achieved is $p_0(g) = \exp\left(-\int_0^g \Lambda_0(g')dg'\right)$. Assume that the hazard rate is continuous and increasing in g. This property introduces a natural dependence of the rate of arrival of success on the level of accumulated investment. We have $p_0(0) = 1$; and $\lim_{g\to\infty} p_0(g) = 0$.

Example 2.1. Linear model[4] - unlimited growth: $\Lambda_0(g) = \lambda g$, where $\lambda > 0$. The probability of no success before g is reached is $p_0(g) = \exp(-\lambda g^2/2)$, and the expected level of investment at the first success is $\sqrt{2\pi}/(2\lambda)$.

Example 2.2. Erlang-(k,λ) model - decreasing returns to scale: let $f(g) = \lambda^k g^{k-1} e^{-\lambda g}/(k-1)!$ and $p_0(g) = \exp(-\lambda g)\sum_{j=0}^{k-1}(\lambda g)^k/k!$, where $k \ge 1$ is an integer and $\lambda > 0$, then $\Lambda_0(g) = f(g)/p_0(g)$. The expected level of investment at the first success is k/λ.

If $\pi_0 < 1$, the probability of the event that no success is observed before the investment level $g = \Gamma(t)$ is reached is $p(g) = 1 - \pi_0 + \pi_0 p_0(g)$. Obviously, $p(0) = 1$, because even a good project cannot bring a success without any initial investment, and $\lim_{g\to\infty} p(g) = 1 - \pi_0$. Furthermore, $p(g)$ is decreasing in g, because

$$p'(g) = \pi_0 p_0'(g) = -\pi_0 \Lambda_0(g) p_0(g). \tag{2.1}$$

If no news was observed until the investment level g was achieved, the posterior belief assigns the probability $\pi(g)$ to the project being successful, and, by the Bayes rule,

$$\pi(g) = \pi_0 p_0(g)/p(g). \tag{2.2}$$

Straightforward differentiation shows that

$$\pi'(g) = -\Lambda_0(g)\pi(g)(1-\pi(g)). \tag{2.3}$$

Thus, the posterior belief is decreasing in g - the more capital was invested into the startup without any observations, the lower is the probability that the project will ever achieve a success.

[4]I am thankful to Juuso Välimäki for the suggestion to use the linear model.

The rate of arrival of success at the investment level g of the project of unknown quality is $\Lambda(g) = -p'(g)/p(g)$. Equations (2.1) and (2.2) imply that

$$\Lambda(g) = \Lambda_0(g)\pi(g). \tag{2.4}$$

Equation (2.4) indicates that, for $\pi_0 \in (0,1)$ the hazard rate $\Lambda(g)$ is a product of an increasing and a decreasing function of Γ. I impose the following conditions on $\Lambda_0(g)$: (i) $\Lambda_0(0) = 0$; (ii) $\Lambda_0(g)$ is of the class C^2 and increasing; (iii) there exists $\Lambda_0'(0+)$, and it is finite; (iv) $2(\Lambda_0'(g))^2 > \Lambda_0''(g)\Lambda_0(g)$, $\forall\, g > 0$. Condition (iv) is a very mild restriction on $\Lambda_0(g)$. In particular, all concave and linear functions, including the ones in Examples 2.1 and 2.2, satisfy this condition. All increasing power and exponential functions also satisfy condition (iv). Boyarchenko (2020) shows that $\Lambda(g)$ is hump-shaped if assumptions (i) - (iv) hold.

2.2. **Necessary condition for startup financing.** Let the hazard rate $\Lambda_0(g)$ satisfy assumptions (i)-(iv) listed at the end of the previous Section. As before, let $\tilde{t} > 0$ denote the (random) time of the first success. Suppose the startup is financed, and $0 < \Gamma'(t) \leq 1$ is the rate of financing at time $t \in [0, \tilde{t})$. At each moment t, the probability of success in the interval $(t, t + dt)$ is $\Lambda(\Gamma(t))d\Gamma(t)$, and the expected gain over the same time interval is $(D\Lambda(\Gamma(t)) - 1)d\Gamma(t)$. Let $\hat{g} = \arg\max_g \Lambda(g)$. Evidently, the instantaneous gain is non-positive for all t if $D\Lambda(\hat{g}) \leq 1$. In this case, it is never optimal to finance the startup, because it will generate nothing but losses. Therefore, the necessary condition for the startup financing is

$$D\Lambda(\hat{g}) > 1. \tag{2.5}$$

From now on, I will focus in startups whose valuation D satisfies the necessary condition (2.5). If the inequality (2.5) holds, the equation $D\Lambda(g) = 1$ has two solutions $g_*(D) < \hat{g} < g^*(D)$. It is easy to see that the instantaneous gain $D\Lambda(g) - 1$ is positive iff $g_*(D) < g < g^*(D)$. Note that $g_*(D)$ is a decreasing function of the startup valuation D, and $g^*(D)$ is increasing in D. Since $\Lambda(g)$ is decreasing in g for all $g > \hat{g}$, and $g^*(D) > \hat{g}$, the instantaneous loss $1 - D\Lambda(g)$ is increasing for $g > g^*(D)$, hence financing the startup after $g^*(D)$ had been reached is never optimal. If $g_*(D) < g < g^*(D)$, then it is not optimal to stop, because the instantaneous gain is positive. If $0 < g < g_*(D)$, the instantaneous gain is negative, so the startup is in the "death valley" zone, but the loss is decreasing as more and more capital is invested in the startup, because the rate of arrival of success is increasing. The first question I ask is whether financing in the "death valley" zone optimal.

As it will be seen later, many results can be obtained easier or become more intuitive if they are represented not in terms of the policy function Γ, but in terms of its inverse $\tau(\Gamma; \cdot)$. This inverse function characterizes the time by which a given level of capital was invested in the startup. By definition, $\Gamma \in \mathcal{M}$

is an increasing function therefore, there exists an inverse function $\tau(\Gamma;\cdot)$: $\mathbb{R}_+ \to \mathbb{R}_+$ which is an increasing and continuous function s.t. $\partial\tau(\Gamma;g)/\partial g \ge 1$ a.e. Let $\mathcal{T}$ denote the class of inverses of functions of the class $\mathcal{M}$. Set $T^*(\Gamma) = \tau(\Gamma; g^*(D))$, and $T_*(\Gamma) = \tau(\Gamma; g_*(D))$.

2.3. **Socially optimal financing.** As in BH, I start with the first best solution. Consider the social planner (SP) who assigns prior belief π_0 to the success of the startup and begins financing the project at time $t_0 \ge 0$, when the startup had accumulated the investment level $g_0 = \Gamma(t_0) \ge 0$. Let $S(g)$ denote the value function of SP at the level of accumulated investment g. SP solves the following optimization problem:

$$S(g_0) = \sup_{\Gamma\in\mathcal{M}:\Gamma(t_0)=g_0} \int_{t_0}^{T^*(\Gamma)} e^{-r(t-t_0)}\frac{p(\Gamma(t))}{p(g_0)}(D\Lambda(\Gamma(t))-1)d\Gamma(t). \tag{2.6}$$

Changing the variable $t = \tau(\Gamma; g)$, rewrite SP's problem as

$$S(g_0) = \sup_{\tau\in\mathcal{T}:\tau(\Gamma;g_0)=t_0} \int_{g_0}^{g^*(D)} e^{-r(\tau(\Gamma;g)-\tau(\Gamma;g_0))}\frac{p(g)}{p(g_0)}(D\Lambda(g)-1)dg. \tag{2.7}$$

Consider also the present value of financing the startup at the highest rate $\Gamma'(t) = 1$ during the same time interval as above:

$$V_f(g_0) \;=\; \int_{g_0}^{g^*(D)} e^{-r(g-g_0)}\frac{p(g)}{p(g_0)}\,(D\Lambda(g)-1)\,dg. \tag{2.8}$$

For $g_0 \in [g_*(D), g^*(D))$, $V_f(g_0) > 0$ because $D\Lambda(g)-1 > 0$ if $g \in (g_*(D), g^*(D))$. Furthermore, V_f is a continuous function which is increasing on $[0, g_*(D)]$. If $V_f(0) < 0$, there exists $\underline{g}(D) < g_*(D)$ s.t. $V_f(\underline{g}(D)) = 0$, and $V_f(g_0) > 0$ (respectively, $V_f(g_0) < 0$) if $g_0 \in (\underline{g}(D), g^*(D))$ (respectively, $g_0 \in [0, \underline{g}(D))$). Hence $\underline{g}(D)$ is the minimal level of initial investment needed so that a startup of valuation D financed at the maximal rate is not a loss-maker. If $V_f(0) \ge 0$, then $\underline{g}(D) = 0$.

If the standing assumption (2.5) holds and $\underline{g}(D) \le g_0$, then

$$\tau(\Gamma; g) = \tau(\Gamma; g_0) + g - g_0,\;\; g \in [g_0, g^*(D)), \tag{2.9}$$

is the first best solution (see Section A.1 for the proof). Equivalently, $\Gamma(t) = \Gamma(t_0) + t - t_0$ for any $\tau(\Gamma; \underline{g}) \le t_0 \le t < T^*(\Gamma)$. The socially optimal value of financing is $S(g_0) = V_f(g_0)$. Hence, SP uses financing at the maximal rate if the startup has at least $\underline{g}(D)$ as the initial investment level. Notice that the rate of financing affects only how fast the startup leaves the "death valley" region $g < g_*(D)$ and reaches the critical investment level $g^*(D)$, but not the values $g_*(D)$ and $g^*(D)$. So, the main intuition behind the social optimality of the maximal rate financing is that it is better to pass through the "death valley" zone as fast as possible thereby minimizing the discount factor in the region $(g_*(D), g^*(D))$ where the instantaneous expected payoff is positive.

The social financing becomes optimal the first time the accumulated investment crosses $\underline{g}(D)$ although the instantaneous gain $D\Lambda(g) - 1 < 0$, i.e., when the startup is still in the "death valley" zone. As opposed to this, the instantaneous gains in BH and in HS are positive at the beginning of financing, and the gains keep decreasing as a function of time until it is optimal to stop financing.

If $g_0 < \underline{g}(D)$, there is no socially optimal policy to finance the startup. To see why, consider SP's problem for $g_0 < \underline{g}(D)$. If the investment level $\underline{g}(D)$ is ever reached, the financing will happen at the maximal rate until $g^*(D)$ is achieved, provided no breakthrough happens earlier. Suppose there exists an optimal financing policy $\tau_{\text{opt}}(g_0)$. Then, on the strength of (2.9), the restriction of $\tau_{\text{opt}}(g)$ on $[\underline{g}(D), g^*(D)]$ is $\tau_{\text{opt}}(g) = \tau_{\text{opt}}(\underline{g}) + g - \underline{g}$. Therefore, we can write the social value of startup financing as

$$S(g_0) = \int_{g_0}^{\underline{g}(D)} e^{-r(\tau_{\text{opt}}(g)-\tau_{\text{opt}}(g_0))} \frac{p(g)}{p(g_0)}(D\Lambda(g)-1)dg + e^{-r(\tau_{\text{opt}}(\underline{g})-\tau_{\text{opt}}(g_0))} V_f(\underline{g}(D)).$$

By definition of $\underline{g}(D)$, the last term on the RHS is zero, and the first term is negative, because the integrand is negative, hence $S(g_0) < 0$, which contradicts the optimality of τ_{opt}.

Thus, if $\underline{g}(D) > 0$, and EN has the starting capital $g_0 < \underline{g}(D)$, financing of the startup is not socially optimal, and all the applicants for the government grants will be rejected if their starting capital $g_0 < \underline{g}(D)$. Alternatively, since the integrand and $g^*(D)$ in (2.8) are increasing in the value of the startup D, we can define $\underline{D}$ as the solution to the equation

$$\int_0^{g^*(\underline{D})} e^{-r(g-g_0)} p(g) \left(\underline{D}\Lambda(g) - 1\right) dg = 0. \tag{2.10}$$

Then, for all $D > \underline{D}$, $V_f(0) > 0$, and for all $D < \underline{D}$, $V_f(0) < 0$, so that all the applicants for the government grants will be rejected if the value of their projects is less than $\underline{D}$, unless they make the initial investment $\underline{g}(D)$.

Notice that if $\pi_0 = 1$, the case labeled as the certain project in BH, $g^* = \infty$, i.e., financing lasts until the first breakthrough is achieved. At the same time the instantaneous gain $D\Lambda_0(g) - 1 < 0$ for small g, so it is still necessary to define either the critical value of the project $\underline{D}$ s.t. (2.10) is satisfied, or the critical initial investment level $\underline{g}(D)$ for optimality of the social financing.

3. Venture capital financing

3.1. **Financing contract.** There are two agents - the entrepreneur (EN), who owns the risky startup project described in Section 2, and the venture capitalist (VC), who owns wealth that can be invested into the project. Both players discount the future at the rate $r > 0$ and assign the prior $\pi_0 \in [0,1]$ to the eventual success of the project. Time is continuous; the game starts at $0 \le t_0 < \tilde{t}$ (as before, $\tilde{t}$ denotes the random time of the first success), when EN had accumulated the capital stock $g_0 \ge 0$ through some ways of financing. At t_0, the entrepreneur (she) makes a take-it-or-leave it offer of a contract $\mathcal{C}$ to the venture capitalist (he). I define the financing contract as a list $\mathcal{C} = (g_0, t_0, T_f, \Gamma, \alpha, z_{rf}, z_{alf})$, where (i) $T_f \wedge \tilde{t}$ is the termination date of the contract; (ii) $\Gamma(t)$ is an increasing investment policy function s.t. the measure $d\Gamma(t) = \Gamma'(t)dt$ is absolutely continuous; $\Gamma(t)$ identifies the projected stock of investment accumulated by $t \in [t_0, T_f \wedge \tilde{t}]$; and $0 < \Gamma'(t) \le 1$ ($t \in [t_0, T_f \wedge \tilde{t}\,)$) is the rate of disbursement of funds by VC so that during an infinitesimal time interval $(t, t+dt)$ VC advances $\Gamma'(t)dt$ to EN; (iii) $0 \le \alpha(\tilde{t}\,) \le 1$ is a measurable function which assigns the share of the payoff D that VC gets at $\tilde{t}$ if $\tilde{t} \le T_f$. I assume that Γ' and α are left continuous on $(t_0, T_f]$. Lastly, z_{rf} denotes the subset of $(t_0, T_f \wedge \tilde{t}\,)$, where EN's investment capital allocations are observable, and z_{alf} denotes the subset of $(t_0, T_f \wedge \tilde{t}\,)$, where EN's actions are hidden. Obviously, $z_{rf} \cup z_{alf} = (t_0, T_f \wedge \tilde{t}\,)$.

Given the contract $\mathcal{C}$, VC makes a binary decision $d \in \{0,1\}$, where $d = 0$ means that $\mathcal{C}$ is rejected, and $d = 1$ means that $\mathcal{C}$ is accepted. If the offer is accepted and financing starts, EN may have an incentive to divert investment funds for alternative purposes. Let $I(t)$ be the actual stock of investment accumulated by $t \in [t_0, T_f \wedge \tilde{t}]$, where I is a non-decreasing function s.t. the measure $dI(t) = I'(t)dt$ is absolute continuous. Then $0 \le I'(t) \le \Gamma'(t)$ ($t \in [t_0, T_f \wedge \tilde{t}\,)$) is the rate of investment in the startup by EN so that during an infinitesimal time interval $(t, t+dt)$ EN invests $I'(t)dt$ into the project. Cases when $I'(t) < \Gamma'(t)$ on a full measure subset of $(t_0, T_f \wedge \tilde{t}\,)$, and consequently $I(T_f) < \Gamma(T_f)$, correspond to investment funds diversion by EN. BH separately consider two types of models for financing: relationship financing (RF) when the allocation decisions of the entrepreneur are observable, and arm's-length financing (ALF) when they are unobservable. In case of RF, the actual investment path $I(t)$ is observed by VC; in case of ALF, deviations (if any) from the contracted policy $\Gamma(t)$ are not observed by VC.

The contract is called *incentive compatible* (IC) if $d = 1$, and $I(t) = \Gamma(t)$ for $t \in (t_0, T_f \wedge \tilde{t}\,)$ a.e. The IC contract is called *optimal* if no other IC contract improves EN's value. For a given initial level $g_0 = \Gamma(t_0)$, the choice of a policy function Γ (which implies the choice of the total investment capital advanced to EN $\Gamma(T_f) = g_0 + \int_{t_0}^{T_f} \Gamma'(t)dt$) together with financing zones z_{rf} and z_{alf} is equivalent to the choice of the contract. In the paper I find candidates for optimal IC policy functions as solutions of ODEs of the form $\Gamma' = \rho(D;\Gamma)$ subject to $\Gamma(t_0) = g_0$, where the rate of financing ρ depends on (un)observability of

EN's actions. Then I say that Γ and the contract are defined by t_0, g_0, T_f and ρ.

Notice that BH and Halac et al. (2020) assume that the entrepreneur has the full bargaining power. HS assume that VC has the full bargaining power arguing that such an assumption agrees with empirical literature that identifies rates of return for venture capitalist in excess of the rates normally used for conventional investment as a basic feature of the venture capital market (see, e.g., Blass and Yosha (2003) or Hall (2005)). As opposed to this, Cochrane (2005) shows that investments in VC portfolio firms did not outperform investments in other NASDAQ stocks during the boom period of the 1990s. The VC industry experienced severe problems during the burst of the dot com bubble and the recent financial crisis, and it is unclear if it will ever be able to recover its former strength. New methods of financing such as crowdfunding or angel investment make the VC investment less profitable. According to the Cambridge Associates LLC U.S. Venture Capital Index, in the first decade of the millennium, the quarterly internal rates of return on VC investment were hovering in the single percentage points, dipping sometimes into negative territory (see Ghalbouni and Rouziés (2010) for details). This evidence suggests that assigning the bargaining power to EN is not that counterintuitive. It also agrees, for example, with the setting in a popular reality show "Shark Tank," where startups make offers to a panel of experts (VC investors). Of course, the real world VC financial contracting is much more complicated than in my model. Kaplan and Strömberg (2003) analyze 213 VC investments and find that financing contracts allow VCs to separately allocate cash flow rights, board rights, voting rights, liquidation rights, etc. Also, their findings indicate that VCs receive few or no private benefits of control.

3.2. **Participation constraints.** Consider any $t \in (t_0, T_f \wedge \tilde{t}\,)$. Suppose that the total volume of investment in the startup has reached $I(t) \le \Gamma(t)$ by t. The probability of the success in the time interval $(t, t+dt)$ is $\Lambda(I(t))$. VC evaluates the expected rate of return on investment as $\alpha(t)D\Lambda(I(t))\Gamma'(t)$ in case of RF, and $\alpha(t)D\Lambda(\Gamma(t))\Gamma'(t)$ in case of ALF. VC finances the project if and only if the rate of financing does not exceed the rate of expected gains. Therefore, the participation constraint in case of RF is $\alpha(t)D\Lambda(I(t))\Gamma'(t) \ge \Gamma'(t)$. Equivalently, $\alpha(t) \ge 1/(D\Lambda(I(t))$. Similarly, in case of ALF, $\alpha(t) \ge 1/(D\Lambda(\Gamma(t))$. Naturally, $\alpha(t)$ may not exceed 1. Thus, the necessary condition for RF (respectively, ALF) financing is $D\Lambda(I(t)) \ge 1$ (respectively, $D\Lambda(\Gamma(t)) \ge 1$) with strict inequality on a full measure subset of $(t_0, T_f \wedge \tilde{t}\,)$. Hence, VC financing is possible if and only if the necessary condition (2.5) for startup financing holds and $g_*(D) \le I(t) \le \Gamma(t) \le g^*(D)$ for every $t \in (t_0, T_f \wedge \tilde{t})$. In particular, VC financing is impossible before the startup has accumulated the investment funds $g_0 \ge g_*(D)$, that is before the startup's instantaneous expected payoff becomes positive. Thus, VC financing always starts later than the socially optimal financing.

In case of success, EN's payoff is $(1-\alpha(t))D$. Since EN has full bargaining power, and her payoff is decreasing in α, it is non-optimal for EN to share any surplus with VC. Therefore the participation constraint binds and VC is offered a break-even share

$$\alpha(t) = 1/(D\Lambda(I(t))), \quad \text{or} \quad \alpha(t) = 1/(D\Lambda(\Gamma(t))). \tag{3.1}$$

The first equation in (3.1) is the break-even share offered to VC in case of RF; it is a function of $I(t)$ - the actual capital stock accumulated by time $t < T_f \wedge \tilde{t}$. The second equation in (3.1) is the break-even share offered to VC in case of ALF; this share is a function of $\Gamma(t)$ - the projected capital stock accumulated by time $t < T_f \wedge \tilde{t}$. From now on, with a slight abuse of notation, I will write $\alpha(I(t))$ for RF and $\alpha(\Gamma(t))$ for ALF instead of $\alpha(t)$.

Let $g = \Gamma(t)$ be the current stock of capital invested in the startup. The break-even share $\alpha(g)$ is decreasing in g on the upward sloping part of the hazard rate Λ, and increasing in g on the downward sloping part. The minimum of the break-even share is achieved at $\hat{g}$, the point of maximum of Λ. Non-monotonicity of the break-even share is a novel prediction of my model, which can be explained as follows. First, consider the investment level g on the upward sloping part of Λ. As more capital has been invested in the startup, the rate of arrival of success in the next instant becomes higher and higher, therefore, VC would agree to a smaller share of the payoff since the expected payoff increases. On the other hand, on the downward sloping part of Λ, the rate of arrival of success decreases as the investment capital grows, so EN has to offer VC higher and higher shares of the project's payoff to keep investment afloat. In BH, the break-even share is independent of the volume of financing; it depends only on posterior beliefs in the success of the startup. The lower the beliefs, the higher is the break-even share. Since posterior beliefs are decreasing in time, EN in BH model always offers a higher share to VC as time goes by and no success is observed.

Non-monotonicity of VC's break-even share results in non-monotonicity of EN's propensity to divert investment funds. To see why, it suffices to compare break-even shares at the time of success $\alpha(\Gamma(\tilde{t}))$ (according to the contract) and $\alpha(I(\tilde{t}))$ (if there was funds diversion prior to $\tilde{t}$). Since $I(\tilde{t}) < \Gamma(\tilde{t})$, $\alpha(I(\tilde{t})) > \alpha(\Gamma(\tilde{t}))$ on the upward sloping part of Λ and $\alpha(I(\tilde{t})) < \alpha(\Gamma(\tilde{t}))$ on the downward sloping part of Λ.

If the investment funds allocation is observable, in the event of success EN has to pay VC $\alpha(I(\tilde{t}))D$, therefore she has less incentives for funds diversion while the investment stock is on the upward sloping part of Λ. EN has to pay VC a larger break-even share if she uses some of the investment funds not as intended, and the rising hazard rate of success also reduces the temptation of diverting funds for personal purposes. On the downward sloping part of Λ, it may be harder to provide incentives for EN when actions are observable. Reducing the investment volume by diverting the funds from the startup means offering a smaller share to VC; and decreasing hazard rate indicates that the startup is less likely to succeed, so EN would like to enjoy longer the flow of funds before the financing stops. As opposed to RF, in case of unobservable actions, EN pays the contracted share $\alpha(\Gamma(\tilde{t}))$ which is smaller than the actual break-even share $\alpha(I(\tilde{t}))$ if the success occurs on the upward sloping part of Λ, and bigger than $\alpha(I(\tilde{t}))$ on the downward sloping part of Λ. Hence, in case of ALF, EN has more (respectively, less) incentives to deviate from the policy $\Gamma(t)$ on the upward (respectively, downward) sloping part of Λ. It will be shown later that such differences in propensity to divert investment funds induces significant differences in incentive compatible rates of funds disbursement in RF and ALF.

3.3. **Termination barrier.** An optimal incentive compatible contract satisfies $D\Lambda(\Gamma(T_f))(1-\alpha(\Gamma(T_f))=1$. Indeed, if the LHS is larger than the RHS, the instantaneous gain for EN from the continuation of the contract is positive, hence, EN finds it optimal to suggest a larger T_f. If the LHS is smaller than the RHS, then, when t is sufficiently close to T_f, EN's gain from a deviation exceeds the continuation value of the contract. Substituting the RHS of (3.1) for the break-even share $\alpha(\Gamma(T_f))$, one derives the following equation for the termination barrier (BH have a similar condition for termination of VC financing):

$$D\Lambda(\Gamma(T_f))-2=0. \tag{3.2}$$

Notice that if $D\Lambda(\hat{g})\le 2$, then the LHS in (3.2) is negative a.e., and there is no optimal incentive compatible contract. If at the same time $D\Lambda(\hat{g})>1$, SP financing at the maximal rate may be optimal. For the termination condition (3.2) to be meaningful, the startup valuation must satisfy $D\Lambda(\hat{g})>2$. Then equation $D\Lambda(g)=2$ has two solutions $g_{*,vc}(D)$ and $g^*_{vc}(D)$ s.t. $g_*(D)<g_{*,vc}(D)<\hat{g}<g^*_{vc}(D)<g^*(D)$. Clearly, it cannot be optimal to stop before the peak $\hat{g}$ of the intensity $\Lambda(g)$ is reached, because the expected hazard rate is increasing in the investment capital. Therefore, the termination barrier of VC financing is $\Gamma(T_f)=g^*_{vc}(D)$.

3.4. **Incentive compatibility.** If EN uses the investment funds as intended, i.e., $I(t)=\Gamma(t)$ on $(t_0,T_f\wedge\tilde{t}\,]$ a.e., the payoff over the interval $(t,t+dt)$ is

$$\Pi(\Gamma;t)=(1-\alpha(\Gamma(t)))D\Lambda(\Gamma(t))d\Gamma(t)=(D\Lambda(\Gamma(t))-1)d\Gamma(t). \tag{3.3}$$

Here I used (3.1) for the break-even share $\alpha(\Gamma(t))$.

EN may have an incentive to misuse the investment funds because there is a trade-off between the uncertain possibility of winning a share of the large prize D and an at hand opportunity of diverting the flow of investment to alternative purposes. Suppose EN diverts investment funds on a full measure subset of $(t_0, T_f \wedge \tilde{t}\,]$. Then the actual investment level is $I(t) = (\Gamma + \delta\Gamma)(t)$, where $\delta\Gamma$ is a non-positive function that represents deviations from the investment policy Γ. It is important to notice that even a local deviation on a small interval (t_1, t_2) has long-term consequences because for any $t \in [t_2, T_f]$, $I(t) = \Gamma(t) + \delta\Gamma(t_2) < \Gamma(t)$. If EN's actions are observable, the payoff over the interval $(t, t + dt)$ is

$$\Pi(\Gamma; \delta\Gamma; t) = (1 - \alpha((\Gamma + \delta\Gamma)(t)))D\Lambda((\Gamma + \delta\Gamma)(t))d(\Gamma + \delta\Gamma)(t) - d(\delta\Gamma)(t), \quad (3.4)$$

where the first term on the RHS describes changes to EN's instantaneous payoff resulting from the diversion, and the second term is the instantaneous gain from the diversion of funds.

In case of hidden actions, EN's payoff is

$$\Pi(\Gamma; \delta\Gamma; t) = (1 - \alpha(\Gamma(t)))D\Lambda((\Gamma + \delta\Gamma)(t))d(\Gamma + \delta\Gamma)(t) - d\delta\Gamma(t). \quad (3.5)$$

The difference in expressions on the RHS of (3.4) and (3.5) comes from the fact that under RF, EN has to pay VC the actual break-even share in case of success, and under ALF, EN has to pay VC what the latter believes to be the break-even share. Unlike BH, I will assume that EN may use different modes of VC financing during her lifetime. I will show that the optimal VC financing indeed involves switching between RF and ALF. In this case, (3.4) and (3.5) can be written in terms of financing zones z_{rf} and z_{alf} as

$$\begin{aligned} \Pi(\Gamma; \delta\Gamma; t) \;\; = \;\; & \left[1 - \mathbb{1}_{z_{rf}}(t)\alpha((\Gamma + \delta\Gamma)(t)) - \mathbb{1}_{z_{alf}}(t)\alpha(\Gamma(t))\right] \\ & \times D\Lambda((\Gamma + \delta\Gamma)(t))d(\Gamma + \delta\Gamma)(t) - d\delta\Gamma(t). \end{aligned}$$

Since any local diversion of investment capital has a long-term impact on the investment stock, it is necessary to compare EN's value functions with and without misuse of funds over all time interval $(t_0, T_f \wedge \tilde{t}\,]$. If EN never diverts investment funds, her expected value is

$$W(\Gamma; g_0) = \int_{t_0}^{T_f} e^{-r(t-t_0)} \frac{p(\Gamma(t))}{p(g_0)} (D\Lambda(\Gamma(t)) - 1) d\Gamma(t). \quad (3.6)$$

Notice that the integrand in (3.6) is exactly the same as in the social planner's objective function (2.6). In Section 2.3, it was shown that the socially optimal rate of financing is the highest possible rate. By the same reasoning, EN would prefer to get the highest possible rate of funds disbursement. As it will become clear later, the rate $\Gamma'(t) = 1$ is not always incentive compatible, so EN will have to use the highest incentive compatible rate of financing.

If EN decides to divert investment funds on any full measure subset of $(t_0, T_f \wedge \tilde{t}\,]$, her expected value is

$$\begin{aligned} W_{div}(\Gamma; g_0; \delta\Gamma) \;\; = \;\; & \int_{t_0}^{T_f} e^{-r(t-t_0)} \frac{p((\Gamma+\delta\Gamma)(t))}{p(g_0)} \left[1 - \mathbb{1}_{z_{rf}}(t)\alpha((\Gamma+\delta\Gamma)(t))\right. \\ & \left. - \mathbb{1}_{z_{alf}}(t)\alpha(\Gamma(t))\right] D\Lambda((\Gamma+\delta\Gamma)(t)) d(\Gamma+\delta\Gamma)(t) \\ & - \int_{t_0}^{T_f} e^{-r(t-t_0)} \frac{p((\Gamma+\delta\Gamma)(t))}{p(g_0)} d\delta\Gamma(t). \end{aligned} \tag{3.7}$$

Equation (3.7) captures all possible diversions of funds over the funding period as well as long term consequences of such diversions. Funds diversion is not profitable iff the gain from the deviation $W_{div}(\Gamma; g_0; \delta\Gamma) - W(\Gamma; g_0) \leq 0$.

3.5. **Optimal incentive compatible contracts: an informal characterization.** Let $g_1 \in [g_0, g^*_{vc}(D))$ and let function $\rho(D; \cdot)$ be piecewise Lipschitz continuous, positive on $(g_1, g^*_{vc}(D))$ and bounded away from zero on any $(g_1, g_2] \subset (g_1, g^*_{vc}(D))$. We say that investment policy Γ is defined by $t_1 \in [t_0, T_f)$, g_1 and $\rho(D; \cdot)$ if Γ is the solution of the Cauchy problem $\Gamma'(t) = \rho(D; \Gamma(t)), t \in (t_1, T_f)$, subject to $\Gamma(t_1) = g_1$.

On the interval $(g_1, g^*_{vc}(D))$ define function $\tau(\Gamma; \cdot)$ by

$$\tau(\Gamma; g) = \int_{g_1}^{g} \frac{dg'}{\rho(D; g')}. \tag{3.8}$$

To simplify the notation, I suppress dependence of τ on g_1. Evidently, τ and Γ are mutual inverses.

First, I find conditions that make local deviations non-profitable. To this end I introduce the *local incentive compatibility (IC) index* at the spot level of investment $g = \Gamma(t)$ and the rate of financing $\Gamma'(t) = \rho(D; \Gamma(t))$:

$$G(D, \rho; g) = rD\Lambda(g)(1-\alpha(g)) - r - \Lambda(g)\rho(D; g) + \mathbb{1}_{z_{alf}}(\tau(\Gamma; g))D\alpha'(g)\Lambda(g)\rho(D; g).$$

The underlying assumption here is that both the rate of financing $\rho(D; \cdot)$ and break-even share α depend on the spot level of investment $g = \Gamma(t)$ only. Clearly, $rD\Lambda(g)(1-\alpha(g))dt$ is the infinitesimal contribution to EN's value due to the possible success during the infinitesimal time interval $(t, t+dt)$, $-rdt$ is the loss caused by the discounting of the future, and $-\Lambda(g)\rho(D; g)dt = dp(g)/p(g)$ is the loss of the continuation value of the contract due to the success during the same infinitesimal time interval. Finally, $D\alpha'(g)\Lambda(g)\rho(D; g)dt = -D\alpha'(g)dp(g)/p(g)$ is EN's instantaneous gain/loss from VC's share adjustment if success happens during the infinitesimal time interval $(t, t+dt)$. Since VC financing starts at $g_0 \geq g_*(D)$ and stops at $g^*_{vc}(D)$ unless a success happened earlier, the local IC index is defined for $g \in (g_0, g^*_{vc}(D)) \subset [g_*(D), g^*_{vc}(D))$.

Differentiation of (3.1) gives $D\alpha'(g) = -\Lambda'(g)/\Lambda(g)^2$. Substituting the latter for $D\alpha'(g)$ and the RHS of (3.1) for $\alpha(g)$ and simplifying, I rewrite the local IC index $G(D, \rho; g)$ as

$$G(D, \rho; g) = r(D\Lambda(g) - 2) - \left(\Lambda(g) + \mathbb{1}_{z_{alf}}(\tau(\Gamma; g))\frac{\Lambda'(g)}{\Lambda(g)}\right)\rho(D; g). \quad (3.9)$$

The linear approximation to the expected gain from deviations is given by

$$W_{div}(\Gamma; g_0; \delta\Gamma) - W(\Gamma; g_0) = \int_{t_0}^{T_f} e^{-r(t-t_0)}\frac{p(\Gamma(t))}{p(g_0)}G(D, \rho; \Gamma(t))\delta\Gamma(t)dt \quad (3.10)$$

(see Section A.2). It follows from (3.10) that if $G(D, \rho; g) > 0$, then EN finds it non-optimal to divert the funds at $g = \Gamma(t)$, because the local gain $G(D, \rho; \Gamma(t))\delta\Gamma(t)dt < 0$.

Next, I consider the case when a local deviation may be profitable. Study of the case when $G(D, \rho; g) \leq 0$ is more subtle, because changes in the value of the expected profit stream are non local, if funds diversion happens on a full measure subset $(t_1, t_2) \subset [t_0, T_f \wedge \tilde{t}]$. Had the changes in the value of the expected profit stream caused by a deviation been local, then the standard variational considerations would have lead to the conclusion that the deviation must be optimal if $G(D, \rho; g) < 0$: EN appropriates the funds and forfeits a part of the expected future gains from the contract in the current form. However, a deviation changes Γ after the moment of deviation forever. If one takes into account the effects of one-shot deviation (diversion of funds during an infinitesimal time interval) on the profit stream in the future, then deviation at time t and capital level $g = \Gamma(t)$ may be suboptimal even if $G(D, \rho; g) < 0$. The intuition behind is similar to the one behind SP financing of the startup in the "death valley" zone – if the startup has a sufficiently high valuation, then the best policy is to finance it at the highest possible rate in order to leave the "death valley" as soon as possible.

In order to study global effects of a one-shot deviation on $(g_1, g_2) \subset (g_0, g^*_{vc}(D))$, for $g \in [g_0, g^*_{vc}(D)]$, define the *accumulated incentive compatibility index*

$$J(D, \rho; g) = \int_{g}^{g^*_{vc}(D)} \frac{e^{-r\tau(\Gamma; g')}}{\rho(D; g')}p(g')G(D, \rho; g')dg'. \quad (3.11)$$

Evidently, $J(D, \rho; \Gamma(t))$ calculates the present value of the stream $G(D, \rho; \Gamma(s))$ on the interval $[t, T_f)$ evaluated at t. By definition, $J(D, \rho; g^*_{vc}(D)) = 0$. Using (3.11) and the fact that $\delta\Gamma(t_0) = 0$ (because VC financing starts at t_0), integrate by parts the RHS in (3.10), then (3.10) becomes

$$W_{div}(\Gamma; g_0; \delta\Gamma) - W(\Gamma; g_0) = \frac{e^{rt_0}}{p(g_0)}\int_{t_0}^{T_f} J(D, \rho; \Gamma(t))d\delta\Gamma(t). \quad (3.12)$$

Suppose there exists $g_1 \in [g_0, g^*_{vc}(D))$ s.t. $J(D, \rho; g_1) < 0$, then, by continuity, $J(D, \rho; g) < 0$ in a small neighborhood of g_1. Choose $d\delta\Gamma(t) < 0$ in the corresponding neighborhood of $\tau(g_1)$ and zero everywhere else, then it follows from (3.12) that the gain from deviation is positive at the investment level g_1. Hence, if the contract is incentive compatible at any investment level $g \in [g_0, g^*_{vc}(D))$, then

$$J(D, \rho; g) \geq 0, \forall\ g \in [g_0, g^*_{vc}(D)]. \tag{3.13}$$

Next, I argue that if the contract is incentive-compatible and optimal, then the constraint (3.13) binds for all policies $\Gamma(t)$ s.t. the rate of financing $\rho(D; \Gamma(t)) < 1$. Suppose (3.13) holds with the strict inequality on a non-empty interval $(g_1, g_2) \subset (g_0, g^*_{vc}(D))$ for some incentive compatible $\Gamma(t)$ s.t. $\rho(D; \Gamma(t)) < 1$. Let for the same $\Gamma(t)$, $J(D, \rho; g) \geq 0$ for all $g \in [g_2, g^*_{vc}(D))$. If EN proposes another policy defined by the rate of financing $\tilde{\rho}(D; \cdot)$ which differs from $\rho(D; \cdot)$ only on (g_1, g_2) so that $\rho(D; g) < \tilde{\rho}(D; g) \leq 1$ for $g \in (g_1, g_2)$ and $J(D, \tilde{\rho}; g) \geq 0$ for all $g \in (g_1, g_2)$, then the contract remains incentive compatible. To see why, recall that changing the rate of financing changes only how fast the level g_2 can be reached. Since the rate $\tilde{\rho}$ is the same as ρ from the moment when g_2 is reached, $J(D, \tilde{\rho}; g) \geq 0$ for all $g \in [g_2, g^*_{vc}(D))$. The RHS in (3.6) is the same as in the social planner's objective function. The same argument as in Section 2.3 shows that EN's payoff function is maximized at the highest possible rate of financing, hence any IC contract such that (3.13) does not bind and $\rho(D; \Gamma(t)) < 1$ is not optimal.

If $\rho(D; g) = 1$ on some interval(s) of investment levels, then (3.13) may hold as a strict inequality over the same interval(s), because increasing the rate of financing above $\Gamma'(t) = 1$ is not feasible. On the complement to the union of such intervals, Γ is found as the solution of the equation $J(D, \rho; g) = 0$. Differentiating the integral on the RHS of (3.11) w.r.t. g, we find that on an interval where the global IC index equals zero, the local IC index is zero as well, and, therefore, $\Gamma'(t) = \rho_0(D; \Gamma(t))$, where function $\rho_0(D; g)$ is defined by

$$\rho_0(D; g) = \frac{r(D\Lambda(g) - 2)}{\Lambda(g) + \mathbb{1}_{z_{alf}}(\tau(g))\Lambda'(g)/\Lambda(g)}. \tag{3.14}$$

We need to find subsets of $(g_{*,vc}(D), g^*_{vc}(D))$, where $\rho(D; g) = 1$ and (3.14) cannot be used to define $\rho(D; g)$ unless the RHS in (3.14) equals one. Since EN prefers the highest possible IC rate of financing, and $\Lambda'(g) > 0$ (respectively, $\Lambda'(g) < 0$) on the upward (respectively, downward) sloping part of $\Lambda(g)$, it is optimal to start with RF if $g_0 < \hat{g}$ and switch to ALF as soon as $g = \hat{g}$. Hence, $z_{alf} = (\hat{t}, T_f)$, where $\hat{t} = \tau(\hat{g})$. If $g \in (g_{*,vc}(D), \hat{g})$, then $\rho_0(D; g) > 0$. If $g \in (\hat{g}(D), g^*_{vc}(D))$, the numerator in (3.14) is positive, but the denominator may become negative. Let g_+ denote a solution to

$$\Lambda'(g)/\Lambda(g)^2 = -1. \tag{3.15}$$

In Section A.3, I prove that there exists a unique solution to (3.15), and moreover the denominator in (3.14) is positive for $g < g_+$ and negative for $g > g_+$. Thus, $\rho_0(D; g)$ is positive and bounded away from zero on any interval $[g_1, g_2] \subset (g_{*,vc}(D), \min\{g^*_{vc}(D), g_+\})$. Note that g_+ depends only on the primitives $\Lambda_0(\cdot)$ and π_0, and it is independent of the startup valuation D. The termination barrier $g^*_{vc}(D)$ is increasing in D, hence for low valuation startups s.t. D is very close to $2/\Lambda(\hat{g})$, $g^*_{vc}(D) \le g_+$. For relatively high D, $g^*_{vc}(D) > g_+$. On $(g_+, g^*_{vc}(D))$, the local IC index is positive, hence, the accumulated IC index is positive as well. Hence, $\rho(D, g) = 1$ for $g \in (g_+, g^*_{vc}(D))$. Substitute $\rho(D, g) = 1$ in (3.9) and consider $G(D, 1; g)$. It is easy to see that $G(D, 1; g)$ is monotone on $(g_{*,vc}(D), \hat{g}]$: it is increasing in g if $rD > 1$, and decreasing if $rD < 1$. It is difficult to characterize behavior of $G(D, 1; g)$ on $[\hat{g}, g^*_{vc}(D))$ in terms of primitives. However, in a number of numerical experiments, I verified that for reasonable and even mildly reasonable parameter values in several models for Λ_0, $G(D, 1; g)$ is increasing on $(\hat{g}, g^*_{vc}(D))$ (see Section A.3 for details). From now on I will focus only on the case when $G(D, 1; g)$ given by (3.9) is increasing on $(\hat{g}, g^*_{vc}(D))$.

If $G(D, 1; g) \ge 0$, then $\rho(D; g) \le 1$ binds. Evidently, $G(D, 1; g_{*,vc}(D)) = -2/D < 0$, and

$$G(D, 1; g^*_{vc}(D))) = -\frac{2}{D}\left(1 + \frac{\Lambda'}{\Lambda}\right)(g^*_{vc}(D)) > 0 \;\Leftrightarrow\; g^*_{vc}(D) > g_+.$$

Thus, for the startup with a sufficiently high valuation, $g^*_{vc}(D) > g_+$, and the constraint $\rho(D; g) \le 1$ binds in a left neighborhood of the termination barrier $g^*_{vc}(D)$. If $g^*_{vc}(D) \le g_+$, then, for all $g \in (g_0, g^*_{vc}(D))$, $G(D, 1; g) < 0$ and $\rho(D, g) = \rho_0(D, g)$.

Now I am in a position to prove that if an optimal incentive compatible contract exists, then it is unique, and derive an explicit representation for the unique candidate.

Lemma 3.1. *Let $g_1 \in [g_0, g^*_{vc}(D))$, and let $\mathcal{C} = (g_1, t_1, T_f, \Gamma, \alpha, z_{rf}, z_{alf})$ be an optimal incentive compatible contract. Then, in this contract,*

(a) α is given by (3.1), $z_f = (t_0, \tau(\hat{g}))$, and $z_{alf} = (\tau(\hat{g}), T_f)$;
(b) VC financing does not stop until $\Gamma(T_f) = g^*_{vc}(D)$ is reached;
(c) If $g^*_{vc}(D) > g_+$, VC financing stops in finite time $T_f = \tau(g^*_{vc}(D))$, and the rate of financing is $\rho = 1$ in a left neighborhood of the termination barrier. If $g^*_{vc}(D) \le g_+$, $T_f = +\infty$, hence, VC financing never ends but the rate of financing tends to 0 as the time goes by and no success is observed;

(d) Γ *is the solution of the following Cauchy problem*

$$\Gamma'(t) = \rho(D; \Gamma(t)), \text{ s.t. } \Gamma(t_1) = g_1, \tag{3.16}$$

where

$$\rho(D; g) = \begin{cases} 1, & \text{if } J(D, 1; g) > 0, \\ \rho_0(D; g), & \text{otherwise.} \end{cases}$$

See Section A.4 for the proof.

4. Optimal incentive compatible contract

4.1. Construction of the unique candidate for financing policy Γ.

Lemma 3.1 indicates that the highest possible disbursement rate $\rho(D; \Gamma(t)) = 1$ can be attained only when the accumulated IC index is positive. It neither provides conditions for existence of state space regions where the index is positive, nor specifies how to find the boundaries of such regions if they exist. In this Section, for a fixed startup valuation D, I construct the unique candidate for the optimal incentive compatible Γ and specify the intervals where the highest rate of financing can be achieved. I start with the classification of startups according to their valuations.

The first category are low valuation startups such that $G(D, 1; g^*_{vc}(D)) \leq 0$. This implies $G(D, 1; g) < 0$ on $(g_{*,vc}(D), g^*_{vc}(D))$ (see Section A.3 for details), hence low valuation startups never achieve the maximal rate of financing $\rho = 1$. BH arrive at a similar conclusion for low profitability projects. The next two categories are moderate and high valuation startups with $G(D, 1; g^*_{vc}(D)) > 0$. Given the startup valuation D, the construction of financing policy Γ can be described as follows.

Consider a low valuation startup. In the proof of (c) in Lemma 3.1, I show that if $G(D, 1; g^*_{vc}(D)) \leq 0$, then in a left neighborhood of $g^*_{vc}(D)$, $\Gamma'(t) = \rho_0(D; \Gamma(t)) \sim c(g^*_{vc}(D))(\Gamma(t) - g^*_{vc}(D))$, where $c(g^*_{vc}(D)) < 0$ and $\rho_0(D; g)$ is given by (3.14). Hence, a low valuation startup will eventually be financed at a decreasing disbursement rate unless a success arrives earlier. This agrees with findings in BH and the downsizing effect in HS. If the startup has low valuation, then the disbursement policy Γ is defined by $\rho_0(D; \cdot) < 1$ on (g_0, g^*_{vc}). Observe that in BH, the rate of financing is a decreasing function of time. In my model, the disbursement rate $\rho_0(D; \cdot)$ is a non-monotone function of the capital investment level. The rate $\rho_0(D; \cdot)$ given by (3.14) is increasing in g on $(g_0, \hat{g})$ and decreasing in a left neighborhood of $g^*_{vc}(D)$. Since $\rho_0(D; \cdot)$ is bounded away from zero on any subset of $(g_{*,vc}(D), g^*_{vc}(D))$, VC financing of a low valuation startup can become feasible only if $g_0 > g_{*,vc}(D)$. The accumulated IC index $J(D, \rho_0; g) = 0$ for all $g \in [g_0, g^*_{vc}(D))$. Due to the latter fact, the contract may not be incentive compatible on the whole interval $(g_{*,vc}(D), g^*_{vc}(D))$. Existence of IC strategies will be examined in Section 4.2.

If $g^*_{vc}(D) > g_+$, there exists at least one point $g_{\rm loc}(D) \in (g_{*,vc}(D), g^*_{vc}(D))$ s.t. $G(D,1; g_{\rm loc}(D)) = 0$, because $G(D,1; g^*_{vc}(D)) > 0$, and $G(D,1; g_{*,vc}(D)) < 0$. Under the assumption that $G(D,1;g)$ is increasing on $(\hat{g}, g^*_{vc}(D))$ the investment level $g_{\rm loc}(D)$ defined above is unique. On the interval $(g_{\rm loc}(D), g^*_{vc}(D))$, $G(D,1;g) > 0$, hence the rate of financing $\rho(D;g) = 1$ on $[g_{\rm loc}(D), g^*_{vc}(D))$. Therefore, the accumulated IC index $J(D,1;g)$ is positive on $[g_{\rm loc}(D), g^*_{vc}(D))$. By continuity, the index will remain positive for $\rho(D;g) = 1$ in a left neighborhood of $g_{\rm loc}(D)$. Let $g_{\rm acc}(D)$ denote the minimal level of capital s.t. the startup is financed at the maximal rate $\rho(D;g) = 1$ for all $g > g_{\rm acc}(D)$ (equivalently, $J(D,1;g) > 0$ for all $g > g_{\rm acc}(D)$). $J(D,1;g)$ is increasing in g on $(g_*(D), g_{\rm loc}(D))$ because

$$\partial J(D,1;g)/\partial g = -e^{-rg}p(g)G(D,1;g) > 0 \text{ for } g \in (g_*(D), g_{\rm loc}(g)).$$

If $J(D,1;g_*(D)) \geq 0$, then $J(D,1;g) > 0$ on $(g_*(D), g^*_{vc}(D))$, hence the startup is financed at the highest rate $\rho(D;g) = 1$ during the whole life of VC financing contract, and VC financing becomes possible the first time the initial capital $g_0 \geq g_*(D)$. If $J(D,1;g_*(D)) < 0$, then there exists a unique $g_{\rm acc}(D) \in (g_*(D), g_{\rm loc}(D))$ s.t. $J(D,1;g_{\rm acc}(D)) = 0$, and $J(D,1;g) > 0$ for all $g \in (g_{\rm acc}(D), g^*_{vc}(D))$.

It is necessary to consider separately the cases when $g_{\rm acc}(D) \in [g_*(D), g_{*,vc}(D)]$ (high valuation startup) and $g_{\rm acc}(D) \in (g_{*,vc}(D), g_{\rm loc}(D))$ (moderate valuation startup). If the startup has high valuation, then VC financing is impossible for $g_0 < g_{\rm acc}(D)$. If $g_0 \in [g_{\rm acc}(D), g^*_{vc}(D))$, construction of the optimal funds disbursement policy is complete, and $\rho(D;g) = 1$ for all $g \in [g_0, g^*_{vc}(D))$.

For a moderate valuation startup, the construction of the optimal disbursement policy Γ is more subtle unless $g_0 \geq g_{\rm acc}(D)$. In this case, $\rho(D;g) = 1$ for all $g \in [g_0, g^*_{vc}(D))$. Thus, the startup with a moderate valuation can get the highest rate of financing if they apply for VC financing with a sufficiently high initial capital accumulated by other means of financing.

Consider now the case when $g_0 < g_{\rm acc}(D)$. VC financing is impossible if $g_0 \leq g_{*,vc}(D)$, because $\rho_0(D;g) \leq 0$ on $[g_0, g_{*,vc}(D)]$. If $g_0 \in (g_{*,vc}(D), g_{\rm acc}(D))$, then, on the one hand, $\rho_0(D;g)$ given by (3.14) is positive on $(g_{*,vc}(D), g_{\rm acc}(D)]$, on the other hand, starting VC financing at rate $\rho_0(D;g) < 1$ may be not incentive compatible because the accumulated IC index $J(D,\rho_0,g_0) = 0$. I study sufficiency conditions for existence of IC strategies in Section 4.2. If the contract is incentive compatible for $g_0 \in (g_{*,vc}(D), g_{\rm acc}(D))$, then construction of the optimal funds disbursement policy is complete, and

$$\rho(D;g) = \begin{cases} 1, & \text{if } g \in [g_{\rm acc}(D), g^*_{vc}(D)), \\ \rho_0(D;g), & \text{if } g \in [g_0, g_{\rm acc}(D)). \end{cases} \tag{4.1}$$

Notice that in this case, VC financing starts at the rate lower than the highest possible one. Let $\bar{g}(D) = \arg\max_{g\in[g_0, g_{\mathrm{acc}}(D)]} \rho_0(D; g)$. The rate of financing $\rho_0(D; g)$ increases on $[g_0, \bar{g}(D)]$, as more capital is invested in the startup. At the level $g_{\mathrm{acc}}(D)$, the rate of financing jumps up to one and remains this high until the total investment reaches $g^*_{vc}(D)$ unless a success arrives earlier. I call the effect that a low or moderately valued startup is financed at less than the maximal rate in early stages of its life *project upsizing*. Project upsizing is a qualitatively new effect which cannot be observed in a model of startup financing based on exponentially distributed times of arrival of success. Interestingly, in the case of upsizing of a moderate valued startup, the rate of financing is discontinuous at $g_{\mathrm{acc}}(D)$: $\rho_0(D; g_{\mathrm{acc}}) < 1$ because $G(D, 1; g_{\mathrm{acc}}) < 0$. The jump happens due to long term effects of a one-shot deviation. Had there been no long term effects, $\rho_0(D; g)$ would have been used until the total investment reached $g_{\mathrm{loc}}(D)$, and the rate of financing would have been continuous because $\rho_0(D; g_{\mathrm{loc}}(D)) = 1$.

The results of this Section are summarized in the following theorem.

Theorem 4.1. *Let D be the startup valuation. Let $g_0 \in [g_*(D), g^*_{vc}(D))$, and let Γ be the unique candidate for an optimal IC strategy that starts at the accumulated capital stock g_0. Then Lemma 3.1 holds, and, in the case of*
(i) **High valuation startup***: $G(D, 1; g^*_{vc}(D)) > 0$, and there exists a unique point $g_{\mathrm{acc}}(D) \in [g_*(D), g_{*,vc}(D)]$ s.t. VC financing is possible iff $g_0 \geq g_{\mathrm{acc}}(D)$; the startup is financed at the highest possible rate $\rho(D; g) = 1$ for all $g \in [g_0, g^*_{vc}(D))$. VC financing stops in finite time, at $T_f = \tau(g^*_{vc}(D))$, unless a success arrives earlier.*
(ii) **Moderate valuation startup***: $G(D, 1; g^*_{vc}(D)) > 0$, VC financing is impossible if $g_0 \leq g_{*,vc}(D)$; there exists a unique point $g_{\mathrm{acc}}(D) \in (g_{*,vc}(D), g_{\mathrm{loc}}(D))$ s.t. $\rho(D; g) = 1$ for all $g \in [g_{\mathrm{acc}}(D), g^*_{vc}(D))$. On the interval $(g_{*,vc}(D), g_{\mathrm{acc}}(D))$ the rate of financing is $\rho_0(D; g) < 1$. VC financing stops in finite time, at $T_f = \tau(g^*_{vc}(D))$, unless a success arrives earlier.*
(iii) **Low valuation startup***: $G(D, 1; g^*_{vc}(D)) \leq 0$, VC financing is impossible if $g_0 \leq g_{*,vc}(D)$, and the rate of financing is $\rho_0(D; g) < 1$ for all $g \in (g_{*,vc}(D), g^*_{vc}(D))$. The start-up is financed until the first observation of success (with probability $1 - \pi_0$, forever). The rate of financing tends to 0 as $t \to +\infty$; the accumulated investment tends to $g^*_{vc}(D)$ from below and never reaches the limit.*

4.2. **Sufficiency conditions and existence of the optimal strategy in the class of incentive compatible strategies.** For high valuation startups and moderate valuation startups with the invested capital stock $g_0 \geq g_{\mathrm{acc}}(D)$, the financing at the rate $\rho(D, g) = 1$ is incentive compatible because the accumulated IC index is positive on (g_0, g^*_{vc}). If there exists an interval where

$\rho(D,g) < 1$ (the cases of moderate valuation startups with $g_0 < g_{\rm acc}(D)$ and low valuation startups), then, on this interval, the linear approximation to the accumulated IC index is zero, hence, one needs to consider the second order approximation.

Calculating the quadratic term in the Taylor expansion of the value of diversion $W_{div}(\Gamma; g_0; \delta\Gamma) - W(\Gamma; g_0)$ (details available on request), I find the second order term in the above approximation

$$SOT = \frac{e^{rt_0}}{2p(g_0)} \int_{t_0}^{T_f} \mathcal{G}(D,\rho;\Gamma(t)) d(\delta\Gamma(t)^2) + \frac{e^{-r(T_f - t_0)}}{2p(g_0)} \frac{p\Lambda'}{\Lambda}(g^*_{vc}(D))\delta\Gamma(T_f)^2,$$

where

$$\begin{aligned}\mathcal{G}(D,\rho;g) \;\; &= \;\; \int_g^{g^*_{vc}(D)} \frac{e^{-r\tau(\Gamma;g')}}{\rho(D;g')\Lambda(g')} \left[-p''(g')G(D,\rho;g') \right. \\ & \qquad \left. + (1 + \mathbb{1}_{(g,\hat{g})}(g'))p(g')\Lambda'(g') \right] dg'.\end{aligned}$$

Let $g \in (g_{*,vc}(D), g_{\rm acc}(D))$. The standard argument shows that if

$$\mathcal{G}(D,\rho;g) + e^{-rT_f}(p\Lambda'/\Lambda)(g^*_{vc}(D)) > 0, \tag{4.2}$$

then deviation is optimal at the capital level g. If (4.2) holds with the opposite sign, then deviation at g is non-optimal. In Section A.5, I prove that (1) in case of a low valuation startup, here exists a unique $g_{SO}(D) \in (g_{*,vc}(D), \hat{g}(D))$ s.t. for $g \in (g_{*,vc}(D), g_{SO}(D))$, (4.2) holds, and for any $g \in (g_{SO}(D), g^*_{vc}(D))$, (4.2) with the opposite sign holds; (2) in case of a moderate valuation startup, either there exists $g_{SO}(D) \in (g_{*,vc}, g_{\rm acc}(D))$ such that (4.2) holds for $g \in (g_{*,vc}, g_{SO}(D))$, and the inequality of the opposite sign is valid for $g \in (g_{SO}(D), g_{\rm acc}(D))$, or (4.2) holds for all $g \in (g_{*,vc}(D), g_{\rm acc}(D))$. In the latter case, I set $g_{SO}(D) = g_{\rm acc}(D)$. This leads to the following theorem.

Theorem 4.2. *(A)* **High valuation startup**. *The financing strategy $\Gamma(t)$ defined by $g_0 \in [g_{\rm acc}(D), g^*_{vc}(D))$, $t_0, T_f = \tau(g^*_{vc}(D))$, and $\rho(D;g) = 1$ for all $g \in [g_0, g^*_{vc}(D))$ is a unique optimal IC strategy.*
(B) **Moderate valuation startup**. *If $g_0 \in [g_{\rm acc}(D), g^*_{vc}(D))$, the financing strategy $\Gamma(t)$ defined by $g_0, t_0, T_f = \tau(g^*_{vc}(D))$, and $\rho(D;g) = 1$ for all $g \in [g_0, g^*_{vc}(D))$ is a unique optimal IC strategy. If $g_0 \in (g_{SO}(D), g_{\rm acc}(D))$, the financing strategy $\Gamma(t)$ defined by $g_0, t_0, T_f = \tau(g^*_{vc}(D))$, and $\rho(D;g)$ given by (4.1) is a unique optimal IC strategy.*

If $g_0 \in (g_{*,vc}(D), g_{SO}(D))$, then $\Gamma(t)$ defined by g_0, t_0 and $\rho(D,g)$ given by (4.1) is not IC. However, for any $\varepsilon > 0$, a suboptimal strategy $\Gamma_\varepsilon(t)$ defined by $g_0, t_0, T_f = \tau(g^*_{vc}(D))$ and

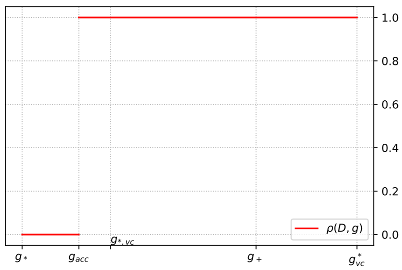
$$\rho_\varepsilon(D;g) = \begin{cases} \rho_0(D;g) - \varepsilon, & \text{if } g \in [g_0, g_{SO}(D)], \\ \rho(D;g), & \text{if } g \in (g_{SO}(D), g^*_{vc}(D)) \end{cases}$$

is IC, and $W(\Gamma_\varepsilon; g_0) \uparrow W(\Gamma; g_0)$ *as* $\varepsilon \downarrow 0$.
(C) **Low valuation startup**. *If* $g_0 \in (g_{SO}(D), g^*_{vc}(D))$*, the financing strategy* $\Gamma(t)$ *defined by* $g_0, t_0, T_f = +\infty$*, and* $\rho_0(D; g)$ *given by* (3.14) *is a unique optimal IC strategy. If* $g_0 \in (g_{*,vc}(D), g_{SO}(D))$*, the aforementioned financing strategy is not IC. However, for any* $\varepsilon > 0$*, a suboptimal strategy* $\Gamma_\varepsilon(t)$ *defined by* $g_0, t_0, T_f = +\infty$ *and* $\rho_\varepsilon(D; g)$ *which equals* $\rho_0(D; g) - \varepsilon$ *on* $[g_0, g_{SO}(D))$ *and* $\rho_0(D; g)$ *on* $[g_{SO}(D), g^*_{vc}(D))$ *is IC, and* $W(\Gamma_\varepsilon; g_0) \uparrow W(\Gamma; g_0)$ *as* $\varepsilon \downarrow 0$.

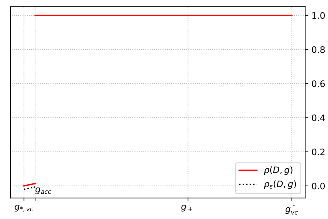
See Section A.5 for the proof.

4.3. **Numerical examples.** All illustrations here are produced for the fixed parameters of Erlang distribution k, λ, discount rate r, and prior π_0. The maximal possible rate of financing $\rho = 1$ is measured in mln. of \$ US. Only the startup valuation D (also measured in mln. of \$ US) varies. Nahata (2019) uses the data from SDC VentureXpert database that identifies VC-backed companies and indicates that the average time between founding of the sample startup and the prior company is 7 years for previously successful entrepreneurs. Therefore, I use the Erlang-(k, λ) distribution with the average waiting time to the first observation equal to 7 (equivalently, the expected total investment at the fist success is \$US 7mln. at the maximal rate of financing). I set $k = 5$, then $\lambda = 5/7$. The most difficult task is to calibrate π_0, the initial belief in the success of the startup. First of all, it is difficult to access the data on startups. Second, VCs tend to emphasize successful investments and "bury their dead" quietly. Lastly, there are different definitions of success or failure. If failure means liquidating all assets, with investors losing all their money, an estimated 30 to 40 per cent of high potential U.S. startups fail. If failure is defined as failing to see the projected return on investment, say, a specific revenue growth rate or date to break even on cash flow, then more than 95 per cent of startups fail. These estimations are based on research of S. Ghosh at the Rock Center for Entrepreneurship at Harvard Business School. According to separate studies by the U.S. Bureau of Labor Statistics and the Ewing Marion Kauffman Foundation, a nonprofit that promotes U.S. entrepreneurship, of all companies, about 60 per cent of startups survive to age three and roughly 35 per cent survive to age 10. I set $\pi_0 = 0.35$, and $r = 0.01$. All the thresholds in Fig. 2 are measured in mln. of \$ US and conditioned on no earlier observation of a success.

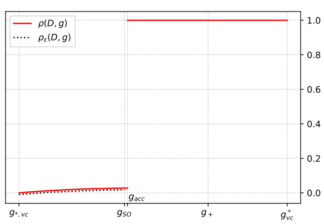
4.4. **Dependence on the bargaining power.** Assume that VC has a certain bargaining power, and, for simplicity, all competitive investors have the same reservation rate of return $\beta > 0$, and, therefore, accept the contract (which, as they believe, will force the entrepreneur to use the strategy $\Gamma(t)$) only if they receive the stream of expected gains $\Gamma'(t)(1 + \beta)dt$ from the project in return for their investment. For the entrepreneur, the stream of expected gains is



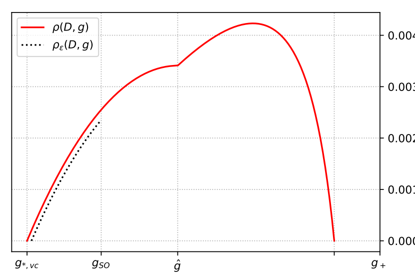
(A) High valuation startup ($D = 44$). VC financing is possible when the stock of investment reaches $g_{\text{acc}} = 3.72$. After that, the startup is financed at rate $\rho = 1$. Termination barrier is $g^*_{vc} = 8.30$.



(B) Moderate valuation startup ($D = 42$). VC financing is possible when the startup's stock of capital is above $g_{*,vc} = 4.43$. At $g_{\text{acc}} = 4.58$, the rate of financing jumps to $\rho = 1$, and the optimal contract becomes IC. On $(g_{*,vc}, g_{\text{acc}})$, only suboptimal rate $\rho_\varepsilon(D; g)$ is IC. Termination barrier is $g^*_{vc} = 8.04$.



(C) Moderate valuation startup ($D = 38$). VC financing is possible when the startup's stock of capital is above $g_{*,vc} = 4.95$. Optimal contract becomes IC at $g_{SO} = 5.89$. On $(g_{*,vc}, g_{SO})$, only suboptimal rate $\rho_\varepsilon(D; g)$ is IC. At the $g_{\text{acc}} = 5.92$, the rate of financing jumps to $\rho = 1$. Termination barrier is $g^*_{vc} = 7.35$.



(D) Low valuation startup ($D = 35.5$). VC financing is possible when the startup's stock of capital is above $g_{*,vc} = 5.67$. The optimal contract is IC if $g > g_{SO} = 5.87$. On $(g_{*,vc}, g_{SO})$, only suboptimal rate $\rho_\varepsilon(D; g)$ is IC. At $\hat{g} = 6.08$, $\rho(D; \cdot)$ has a kink, because the startup switches from RF to ALF. Termination barrier is $g^*_{vc} = 6.52$.

FIGURE 2. Rates of financing for startups of different valuations

$(D\Lambda(\Gamma(t))\Gamma'(t) - (1+\beta)\Gamma'(t))\,dt = (1+\beta)(D_\beta\Lambda(\Gamma(t))-1)\Gamma'(t)dt$, where $D_\beta = D/(1+\beta) < D$. Hence, all the results above are valid with $D_\beta < D$ in place of D. The new interval $(g_*(D_\beta), g^*_{vc}(D_\beta))$ where VC financing is possible is either empty or a proper subinterval of the initial financing interval $(g_*(D), g^*_{vc}(D))$. The rates of financing $\rho_\beta(D; g) < \rho(D; g)$ on all intervals where $\rho_\beta(D; g) < 1$;

otherwise, $\rho_\beta(D; g) \leq \rho(D; g) = 1$. It is optimal to switch from RF to ALF at $\hat{g} = \arg\max_g \Lambda(g)$, independent of the bargaining power.

5. Conclusion

I constructed a model of innovation financing that explains life-cycles of startups. The novelty of my approach in the assumption that the time of arrival of the first breakthrough of a successful project is a jump time of a time-inhomogeneous Poisson process whose intensity is increasing in time. As a result, the arrival rate of success of the startup of unknown quality is hump-shaped - it grows from zero to the point of maximum, and then starts decreasing in time. Non-monotone hazard rate is one of the reasons why at early stages, the entrepreneur has to rely on family funds or government money, because the rate of arrival of success is too low for VC investment to be feasible. After sufficient investment funds have been accumulated, the arrival rate becomes sufficiently high, and VC investment becomes optimal. If VC financing starts very close to the boundary of the region where both the participation constraint for the investor and incentive constraint for the entrepreneur can be met, then the rate of financing increases starting from almost zero. This, in particular, implies that stopping the government financing and switching to the market financing as soon as the latter becomes available may be inefficient.

Qualitative results in my model are different from the ones in models of innovation financing that use the standard Poisson process to model the first time of success of a good startup. In particular, BH show that, *depending on the set of parameters*, the project may get full funding or restricted funding. I show that for the *same set of parameters*, the project may receive full or restricted funding in different stages of development. Next, BH show that, *depending on the set of parameters*, ALF may be preferable to RF. I show that for the *same set of parameters*, EN will find it optimal to start RF on the upward sloping part of the hazard rate and switch to ALF on the downward sloping part. Qualitative predictions of the model remain valid if both players have some bargaining power.

References

Bergemann, D. and U. Hege, "Dynamic venture capital financing, learning and moral hazard," *Journal of Banking and Finance*, 1998, *22*, 703–735.

——— **and** ———, "The Financing of Innovation: Learning and Stopping," *RAND Journal of Economics*, 2005, *36*, 719–752.

Blass, A.A. and O. Yosha, "Financing R&D in mature companies: an empirical analysis," *Economics of Innovation and New Technology*, 2003, *12*, 425–448.

Bonatti, A. and J. Hörner, "Collaborating," *American Economic Review*, 2011, *101*, 349–374.

——— **and** ———, "Career concerns with exponential learning," *Theoretical Economics*, 2017, *12*, 425–475.

Boyarchenko, S., "Super and submodularity of stopping games with random observations," *Economic Theory*, 2020, *70* (4), 983–1022.

———, "Inefficiency of sponsored research," *Journal of Mathematical Economics*, 2021. https://doi.org/10.1016/j.jmateco.2021.102497.

Cochrane, J.H., "The Risk And Return Of Venture Capital," *Journal of Financial Economics*, 2005, *75*, 3–52.

Deb, J., A. ́Ory, and K. Williams, "Aiming for the goal:contribution dynamics of crowdfunding," 2019. NBER Working Paper No. w25881.

Ghalbouni, J. and D. Rouziés, "The VC shakeout," *Harvard Business Review*, 2010, *July-August*, 21–23.

Grenadier, S., A. Malenko, and N. Malenko, "Timing decisions in organizations: communication and authority in a dynamic environment," *American Economic Review*, 2016, *106*, 2552–2581.

Guo, Y., "Dynamic delegation of experimentation," *American Economic Review*, 2016, *106*, 1968–2008.

Halac, M. and I. Kremer, "Experimenting with career concerns," *American Economic Journal: Microeconomics*, 2020, *12*, 260–288.

———, ———, and E. Winter, "Raising capital from heterogeneous investors," *American Economic Review*, 2020, *110*, 889–921.

———, N. Kartik, and Q. Liu, "Optimal contracts for experimentation," *Review of Economic Studies*, 2016, *83*, 1040–1091.

———, ———, and ———, "Contests for experimentation," *Journal of Political Economy*, 2017, *125*, 1523–1569.

Hall, B.H., "The financing of innovation," in S. Shane, ed., *Handbook of Technology and Innovation Management*, Oxford, U.K.: Blackwell Publishers, 2005.

Hörner, J. and L. Samuelson, "Incentives for experimenting agents," *RAND Journal of Economics*, 2013, *44*, 632–663.

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

Nahata, R., "Success is good, but failure is not bad either:serial entrepreneurs and venture capital contracting," *Journal of Corporate Finance*, 2019, *58*, 624–649.

Strausz, R., "A theory of crowdfunding: a mechanism design approach with demand uncertainty and moral hazard," *American Economic Review*, 2017, *107*, 1430–1476.

Appendix A. For online publication only

A.1. Proof of equation (2.9).

Lemma A.1. *Let* $f : \mathbb{R}_+ \to \mathbb{R}$ *be a continuous function satisfying the following properties: there exist* $0 < k_* < k^*$ *s.t.* $f(k_*) = f(k^*) = 0$*; and* $f(k) > 0$ *iff* $k \in (k_*, k^*)$*. Let* $\underline{k} \in [0, k_*]$ *be defined by*

$$\int_{\underline{k}}^{k^*} e^{-rk} f(k) dk = 0,$$

and let $k_0 \in [\underline{k}, k^*]$*. Then the maximization problem*

$$V(k_0) = \sup_{\tau \in \mathcal{T}} \int_{k_0}^{k^*} e^{-r(\tau(k)-\tau(k_0))} f(k) dk \tag{A.1}$$

has the unique solution τ *defined by* $\tau(k) = \tau(k_0) + k - k_0$ *for* $k \in [\underline{k}, k^*]$.

Proof. I start the proof by reducing the set of candidates for an optimal policy. For every $k \in [\underline{k}, k^*]$, let $\tau^0(k) = k$ and let

$$W(k; \tau) = \int_{k}^{k^*} e^{-r(\tau(k')-\tau(k))} f(k') dk'.$$

By assumptions of the Lemma, $W(k; \tau^0) > 0$, for any $k \in (\underline{k}, k^*]$ hence any optimal policy τ has to satisfy

$$W(k, \tau) \geq W(k; \tau^0), \quad \forall\, k \in [\underline{k}, k^*]. \tag{A.2}$$

Let $\mathcal{T}_0 \subset \mathcal{T}$ be the subset of policies that satisfy (A.2). I look for a solution to the problem (A.1) in the class of policies $\tau \in \mathcal{T}_0$. Since the function f changes the sign at $k = k_*$, I need to consider the optimal policy on $[k_*, k^*]$ and $[0, k_*]$ separately. Let $\tau^*(k)$ denote the restriction of the optimal policy on $[k_*, k^*]$, and $\tau_*(k)$ denote the restriction of the optimal policy on $[\underline{k}, k_*]$.

Consider, first, problem (A.1) for $k_0 \in [k_*, k^*]$. Since $f(k) > 0$ for all $k \in (k_*, k^*)$ and $\tau'(k+) \geq 1$, $V(k_0)$ is maximized with $\tau(k) = \tau(k_0) + k - k_0$. Hence, an optimal solution is in the subclass $\mathcal{T}_1 \subset \mathcal{T}_0$ of strategies satisfying

$$\tau(k) = \begin{cases} \tau_*(k), & \text{if } k \in [\underline{k}, k_*] \\ \tau^*(k_*) + k - k_*, & \text{if } k \in [k_*, k^*]. \end{cases} \tag{A.3}$$

Fix $k_0 \in (\underline{k}, k_*]$ and $\tau \in \mathcal{T}_1$. Then $W(k_0, \tau) \geq W(k_0, \tau_0) > 0$. Since $f(k) \leq 0$ on $[\underline{k}, k_*]$, I conclude that the last term on the RHS of

$$W(k_0, \tau) = \int_{k_0}^{k_*} e^{-r(\tau(k)-\tau(k_0)} f(k) dk + e^{-r(\tau(k_*)-\tau(k_0))} V(k_*) \tag{A.4}$$

is greater than or equal to $W(k_0,\tau_0)>0$. Hence, $r(\tau(k_*)-\tau(k_0))\le \ln\frac{V(k_*)}{W(k_0,\tau_0)}$. Clearly, $r(\tau(k_*)-\tau(k_0))\downarrow 0$ as $k_0\uparrow k_*$, uniformly in $\tau\in\mathcal{T}_1$, therefore it follows from (A.4) that

$$W(k_0,\tau)=o(r(\tau(k_*)-\tau(k_0)))-r(\tau(k_*)-\tau(k_0))V(k_*),$$

where $o(r(\tau(k_*)-\tau(k_0)))/r(\tau(k_*)-\tau(k_0))\to 0$ also uniformly in $\tau\in\mathcal{T}_1$. It follows that if $\delta>0$ is sufficiently small, then the choice $\tau=\tau_0$ maximizes $W(k_*-\delta,\tau)$.

As long as $k_*-\delta>\underline{k}$, $V(k_*-\delta)>0$. Hence, one can replace k_* with $k_*-\delta$ and repeat the same argument.

□

Function $p(g)(D\Lambda(g)-1)$ satisfies conditions of Lemma A.1, therefore the supremum in (2.7) is attained at (2.9) for all $g\in[\underline{g},g^*)$.

A.2. **Proof of equation** (3.10)**.** To find the necessary conditions for an optimal IC contract, I linearize the RHS in (3.7) w.r.t. $\delta\Gamma$.

$$\begin{aligned}
&e^{-rt_0}p(g_0)W_{div}(\Gamma;g_0;\delta\Gamma)\\
&=\int_{t_0}^{T_f} e^{-rt}\left[p(\Gamma(t))+p'(\Gamma(t))\delta\Gamma)(t))\right]D\left[\Lambda(\Gamma(t))+\Lambda'(\Gamma(t))\delta\Gamma(t)\right]\\
&\quad\times\left[1-\alpha(\Gamma(t))-\mathbb{1}_{z_{rf}}(t)\alpha'((\Gamma(t))\delta\Gamma(t)\right]d(\Gamma+\delta\Gamma)(t)\\
&\quad-\int_{t_0}^{T_f} e^{-rt}\left[p(\Gamma(t))\right]d\delta\Gamma(t)\\
&=e^{-rt_0}p(g_0)W(\Gamma;g_0)\\
&\quad+\int_{t_0}^{T_f} e^{-rt}\left[(p\Lambda)'D(1-\alpha)-\mathbb{1}_{z_{rf}}(t)p\Lambda D\alpha'\right](\Gamma(t))\delta\Gamma(t)d\Gamma(t) &\text{(A.5)}\\
&\quad+\int_{t_0}^{T_f} e^{-rt}\left[p\Lambda D(1-\alpha)-p)\right](\Gamma(t))d\delta\Gamma(t) &\text{(A.6)}
\end{aligned}$$

Integrate the integral (A.6) by parts using the fact that $\delta\Gamma(t_0)=0$ because VC financing starts at t_0:

$$\begin{aligned}
&\int_{t_0}^{T_f} e^{-rt}\left[p\Lambda D(1-\alpha)-p)\right](\Gamma(t))d\delta\Gamma(t)\\
&=e^{-rT_f}(p(D\Lambda(1-\alpha)-1))(g^*_{vc}(D))\delta\Gamma(T_f)\\
&\quad+\int_{t_0}^{T_f} e^{-rt}r\left[p\Lambda D(1-\alpha)-p)\right](\Gamma(t))\delta\Gamma(t)dt\\
&\quad-\int_{t_0}^{T_f} e^{-rt}\left[(p\Lambda)'D(1-\alpha)-p\Lambda D\alpha'-p')\right](\Gamma(t))\delta\Gamma(t)d\Gamma(t).
\end{aligned}$$

Using (3.1) for the break-even share, I conclude that, on the strength of (3.2), the boundary term is zero. Substituting the results for (A.6) and calculating the sum of (A.5) and (A.6) and using the fact that $z_{rf} \cup z_{alf} = (t_0, T_f \wedge \tilde{t}\,)$, I derive

$$\begin{aligned}
&e^{-rt_0} p(g_0) \left(W_{div}(\Gamma; g_0; \delta\Gamma) - W(\Gamma; g_0) \right) \\
&= \int_{t_0}^{T_f} e^{-rt} r \left[p(D\Lambda - 2) \right] (\Gamma(t)) \delta\Gamma(t) dt \\
&\quad + \int_{t_0}^{T_f} e^{-rt} \left[p' + \mathbb{1}_{z_{alf}}(t) p \Lambda D \alpha' \right] (\Gamma(t)) \delta\Gamma(t) d\Gamma(t).
\end{aligned}$$

Differentiate (3.1):

$$\alpha'(g) = -\frac{\Lambda'}{D\Lambda^2}(g). \tag{A.7}$$

Substituting (A.7) for $\alpha'(\Gamma(t))$ and $-(p\Lambda)(g)$ for $p'(g)$, rewrite the linearization as

$$\begin{aligned}
&e^{-rt_0} p(g_0) \left(W_{div}(\Gamma; g_0; \delta\Gamma) - W(\Gamma; g_0) \right) \\
&= \int_{t_0}^{T_f} e^{-rt} r \left[p(D\Lambda - 2) \right] (\Gamma(t)) \delta\Gamma(t) dt \\
&\quad - \int_{t_0}^{T_f} e^{-rt} \left[p\Lambda \left(1 + \mathbb{1}_{z_{alf}}(t) \frac{\Lambda'}{\Lambda^2} \right) \right] (\Gamma(t)) \delta\Gamma(t) d\Gamma(t)
\end{aligned} \tag{A.8}$$

It remains to multiply both sides of the last equation by $e^{rt_0}/p(g_0)$ and use (3.9) to arrive at (3.10).

A.3. **Solution to** (3.15) **and properties of** $G(D, 1; g)$**.** First, I show that there exists a unique solution to (3.15). Let $\Lambda_0(g)$ satisfy assumptions (i)-(iv) specified in the last paragraph of Section 2.1. Differentiation of $\Lambda(g)$ gives $\Lambda'(g) = \Lambda_0(g)^2 \pi(g) f(g)$, where

$$f(g) := \Lambda_0'(g)/\Lambda_0(g)^2 + \pi(g) - 1$$

is strictly decreasing because assumption (iv) holds and π is decreasing in g. It follows from assumptions (i) and (iii) that $\lim_{g \to 0+} f(g) = +\infty$; and from assumptions (ii) and (iv) that $\lim_{g \to \infty} f(g) = -1 < 0$. Using (2.4) and $\Lambda'(g)$ calculated earlier, one can write $\Lambda'(g)/\Lambda(g)^2 = f(g)/\pi(g)$. Hence, (3.15) is equivalent to

$$f(g) = -\pi(g). \tag{A.9}$$

By properties of $\pi(g)$, the RHS in (A.9) is a strictly increasing function s.t. $-\lim_{g \to 0+} \pi(g) = \pi_0$, and $\lim_{g \to \infty} \pi(g) = 0$. Hence, there exists a unique solution g_+ to (A.9) (equivalently, to (3.15)).

Next, I study properties of the function $G(D,1;g)$ defined by (3.9). As it was noted earlier, $G(D,1;g_{*,vc}(D)) = -2/D < 0$, and $G(D,1;g^*_{vc}(D)) > 0$ iff $g^*_{vc}(D) > g_+$. Consider behavior of $G(D,1;\cdot)$ on $(g_{*,vc}(D),\hat{g})$. Calculate the derivative $\partial G(D,1;g)/\partial g = (rD-1)\Lambda'(g)$. Since $\Lambda'(g) > 0$, $\partial G(D,1;g)/\partial g > 0$ iff $rD > 1$, hence $\hat{g}$ is the local maximum (respectively, minimum) of $G(D,1;\cdot)$ on $(g_{*,vc}(D),\hat{g})$ if $rD > 1$ (respectively, $rD < 1$). Thus, $G(D,1;\cdot)$ is a monotone function on $(g_{*,vc}(D),\hat{g})$. If $(rD-1)\Lambda(\hat{g}) > 2r$, then $G(D,1;\hat{g}) \geq 0$ (the case of high or moderately high valuation firms). In the latter case, there exists a unique $g_{\rm loc}(D) \in (g_{*,vc}(D),\hat{g}]$ s.t. $G(D,1;g_{\rm loc}(D)) = 0$.

Suppose, $G(D,1;\hat{g}) < 0$ and $G(D,1;g^*_{vc}(D)) > 0$. Since $G(D,1;g^*_{vc}(D)) > 0$ iff $g^*_{vc}(D) > g_+$, it follows that $G(D,1;g_+) = r(D\Lambda(g_+)-2) > 0$. Hence there exists a $g_{\rm loc}(D) \in (\hat{g},g_+)$ s.t. $G(D,1;g_{\rm loc}(D)) = 0$. Finally, if $G(D,1;\hat{g}) < 0$ and $G(D,1;g^*_{vc}(D)) \leq 0$ (equivalently, $g^*_{vc} \leq g_+$), then either $G(D,1;\cdot)$ remains negative on $(g_{*,vc}(D),g^*_{vc}(D))$ or it changes the sign on the interval $(\hat{g},g^*_{vc}(D))$ an even number of times. In general, $G(D,1;g)$ can be non-monotone on $(\hat{g},g^*_{vc}(D))$. In a number of numerical experiments, I verified that for reasonable and even mildly reasonable parameter values in several models for Λ_0, $G(D,1;g)$ is increasing on $(\hat{g},g^*_{vc}(D))$, hence the point $g_{\rm loc}(D)$, where $G(D,1;g_{\rm loc}(D)) = 0$ is unique. Writing $G(D,1;g) = \Lambda(g)\Psi(g)$, where

$$\Psi(g) = r\left(D - 2/\Lambda(g)\right) - 1 - \Lambda'(g)/\Lambda(g)^2,\ g \in (\hat{g},g^*_{vc}(D)],$$

one can see that $G(D,1;g) > 0$ iff $\Psi(g) > 0$. By definition, $\Psi(g_+) = r\left(D - 2/\Lambda(g_+)\right)$, and

$$\Psi'(g_+) = 2r\frac{\Lambda'(g_+)}{\Lambda(g_+)^2} - \frac{\Lambda''(g_+)}{\Lambda(g_+)^2} + 2\left(\frac{\Lambda'(g_+)}{\Lambda(g_+)^2}\right)^2 = 2(1-r) + \frac{\Lambda''(g_+)}{\Lambda'(g_+)}.$$

Since $r < 1$, the first term in the above expression is positive. Recall that $\hat{g} = \arg\max \Lambda(g)$, hence $\Lambda''(g)/\Lambda'(g) > 0$ in a right neighborhood of $\hat{g}$. Since $\hat{g} = \hat{g}(\pi_0)$ is given by $\Lambda_0'(g)/\Lambda_0(g)^2 + \pi(g) - 1 = 0$, and $g_+ = g_+(\pi_0)$ satisfies $\Lambda_0'(g)/\Lambda_0(g)^2 + 2\pi(g) - 1 = 0$, both $\hat{g}(\pi_0)$ and $g_+(\pi_0)$ are increasing in π_0, and the difference $g_+(\pi_0) - \hat{g}(\pi_0)$ is also increasing. Numerical experiments show that $\Lambda''(g_+)/\Lambda'(g_+) > 0$ unless the probability of success of a startup becomes unnaturally high (say, $\pi_0 = 0.8$ or larger), but even in this case, $\Psi'(g_+) > 0$ for $r < 1$. Hence, I conclude that $\Psi(g)$ approaches g_+ from below, and the same holds for $g^*_{vc}(D)$ when $g^*_{vc}(D)$ approaches g_+ from the left. As, it was mentioned above, for reasonable values of π_0, the points $\hat{g}(\pi_0)$ and $g_+(\pi_0)$ are very close to each other. Therefore, $G(D,1;g^*_{vc}(D)) < 0$ implies $G(D,1;\hat{g}) < 0$, hence a low valuation startup never achieves the maximal rate of financing $\rho = 1$.

A.4. **Proof of Lemma 3.1.** Part (a) was already proved.

(b) If $\Gamma(T_f) < g^*_{vc}(D)$, then the value of the contract can be increased continuing the financing. If $g^*_{vc}(D) > g_+$, then the constraint $\rho(D;g) \leq 1$

binds in a left neighborhood of the termination barrier, and by definition, the contract remains IC because the local IC index is positive in the same neighborhood. If $g^*_{vc}(D) \le g_+$, the value of the contract can still be increased continuing the financing at a sufficiently low rate $\Gamma'(t) = \varepsilon > 0$ during a sufficiently short time interval $(T_f, T_f + \delta)$. Indeed, if $\varepsilon > 0$ and $\delta > 0$ are sufficiently small, then the local IC index is positive on $(T_f, T_f + \delta)$, hence, the contract is IC, and the value of the contract for EN at T_f is positive.

(c) If $g^*_{vc}(D) > g_+$, then the rate of financing is bounded away from zero, hence financing ends in finite time, at $T_f = \tau(g^*_{vc}(D))$. If $g^*_{vc}(D) \le g_+$, then at $g^*_{vc}(D)$, the RHS of (3.14) is zero, hence, there exists a left neighborhood of $g^*_{vc}(D)$ such that the local incentive compatibility index is 0 in this neighborhood, and $\Gamma'(t) = \rho_0(D; \Gamma(t))$ for $\Gamma(t)$ in this neighborhood. As $\Gamma(t)$ approaches $g^*_{vc}(D)$ from the left, $\Gamma'(t) = \rho_0(D; \Gamma(t)) \sim c(g^*_{vc}(D))(\Gamma(t) - g^*_{vc}(D))$, where $c(g^*_{vc}(D)) < 0$. Hence, $\Gamma(t)$ is well-defined for all $t > t_1$, and never reaches $g^*_{vc}(D)$; but $\Gamma(+\infty) = g^*_{vc}(D)$.

(d) The arguments preceding Lemma imply that Γ satisfies the ODE (3.16). In the explicit construction of Γ in Section 4.1, it is shown that the RHS in (3.16) is a piecewise Lipschitz continuous positive function of $\Gamma(t)$ uniformly bounded away from zero on any $(g_1, g_2] \subset (g_1, g^*_{vc}(D))$. Hence, there exists a solution of problem (3.16), and it is unique.

A.5. **Proof of Theorem 4.2.** (A) For a high valuation startup, $J(D, 1, g) > 0$ for all $g \in (g_0, g^*_{vc}(D)) \subset (g_{\mathrm{acc}}(D), g^*_{vc}(D))$, hence the policy $\Gamma(t)$ is incentive compatible.

(C) For a low valuation startup, $G(D, \rho, g) = 0$, hence $J(D, \rho, g) = 0$ for all $g \in (g_0, g^*_{vc}(D)) \subset (g_{*,vc}(D), g^*_{vc}(D))$, and

$$\mathcal{G}(D, \rho; g) = \int_g^{g^*_{vc}(D)} \frac{e^{-r\tau(\Gamma; g')}}{\Lambda(g')\rho(D; g')}(1 + \mathbb{1}_{(g,\hat{g})})p(g')\Lambda'(g')dg'. \tag{A.10}$$

On $(\hat{g}(D), g^*_{vc}(D))$ $\Lambda'(g) < 0$, therefore $\mathcal{G}(D, \rho; g) < 0$. Hence, inequality (4.2) holds with the opposite sign, and no deviation from the financing strategy $\Gamma(t)$ can be optimal after the investment level reached $\hat{g}(D)$. In a neighborhood of $g_{*,vc}(D)$,

$$\Lambda(g)\rho(D, g) = \Lambda(g)\rho_0(D, g) = r(D\Lambda(g) - 2).$$

Hence, as $g \downarrow g_{*,vc}(D)$, the integrand on the RHS of (A.10) behaves as $c(g - g_{*,vc}(D))^{-1}(1+o(1))$, where $c > 0$, therefore, $\mathcal{G}(D, \rho; g)) \to +\infty$ as $g \downarrow g_{*,vc}(D)$.

It is easy to see that $\mathcal{G}(D, \rho; g)$ decreases on $(g_{*,vc}(D), \hat{g})$. Hence, there exists a unique $g_{SO}(D) \in (g_{*,vc}(D), \hat{g}(D))$ s.t. inequality (4.2) holds with the opposite sign iff $g \in (g_{SO}(D), g^*_{vc}(D))$. Thus, in case of a low valuation startup, the financing strategy $\Gamma(t)$ defined by $\rho(D; g)$ satisfying (3.14) is incentive compatible iff $g_0 \in (g_{SO}(D), g^*_{vc}(D))$.

(B) In case of a moderate valuation startup, $J(D,1,g) > 0$ for all $g \in (g_0, g^*_{vc}(D)) \subset (g_{\mathrm{acc}}(D), g^*_{vc}(D))$, hence the policy $\Gamma(t)$ is incentive compatible.

If $g_0 \in (g_{*,vc}(D), g_{\mathrm{acc}}(D)]$, then for $g \in (g_0, g_{\mathrm{acc}}(D)]$, $G(D,\rho,g) = 0$, $J(D,\rho,g) = 0$, and

$$\mathcal{G}(D,\rho;g) = \int_g^{g_{\mathrm{acc}}(D)} \frac{e^{-r\tau(\Gamma;g')}}{\Lambda(g')\rho(D;g')}(1 + \mathbb{1}_{(g,\hat{g})}(g'))p(g')\Lambda'(g')dg'$$
$$+ \int_{g_{\mathrm{acc}}(D)}^{g^*_{vc}(D)} \frac{e^{-rg'}}{\Lambda(g')}\left[-p''(g')G(D,1;g') + \left((1 + \mathbb{1}_{(g,\hat{g})}(g'))p(g')\Lambda'(g')\right)\right] dg'.$$

By the same argument as in (C), $\mathcal{G}(D,\rho;g)) \to +\infty$ as $g \downarrow g_{*,vc}(D)$. Hence inequality (4.2) holds in a right neighborhood of $g_{*,vc}(D)$. If

$$\mathcal{G}(D,\rho;g) > -e^{-rT_f}(p\Lambda'/\Lambda)(g^*_{vc}(D)) \;\forall\, g \in (g_{*,vc}(D), g_{\mathrm{acc}}(D)),$$

then the policy $\Gamma(t)$ is not incentive compatible on $(g_{*,vc}(D), g_{\mathrm{acc}}(D))$. If there exists a unique $g_{SO}(D) \in (g_{*,vc}(D), g_{\mathrm{acc}}(D))$ s.t.

$$\mathcal{G}(D,\rho;g) < -e^{-rT_f}(p\Lambda'/\Lambda)(g^*_{vc}(D)) \;\forall\, g \in (g_{SO}(D), g_{\mathrm{acc}}(D)),$$

then the financing strategy $\Gamma(t)$ is incentive compatible on $(g_{SO}(D), g_{\mathrm{acc}}(D))$. To finish the proof, it suffices to note that if the rate of funds disbursement $\rho(D;g)$ is replaced by $\rho_\varepsilon(D;g)$, then the equality $J(D,\rho,g) = 0$ becomes the strict inequality $J(D,\rho_\varepsilon,g) > 0$, and the sufficient F.O.C. is satisfied. Hence, a deviation from the investment Γ_ε, where $\varepsilon > 0$ is small, is non-optimal.