Financing Innovation with Future Equity

Idan Hodor
Monash University

May, 2021*

Abstract

This article studies future equity financing in a continuous-time principal-agent setup whereby career concerns generate moral hazard tension. Our framework admits precise closed-form expressions. The higher firm value leading up to conversion, the fewer equity investors attain and the less risk the entrepreneur takes. We implement the contract using a convertible note with a valuation cap, that if set too high, developing innovation becomes suboptimal. Lastly, we introduce the implied probability of success: a novel measure allowing for risk comparison across different innovative technologies. We demonstrate its use to empirically estimate investors' skill and correct selection bias in realized returns.

1 Introduction

Future equity instruments have become very popular in financing early-stage firms. So much so that in recent years there has been a competition to standardize future equity agreements. In late 2013, startup accelerator Y Combinator unveiled its Simple Agreement for Future Equity ("SAFE") investment instrument with the stated goal of standardizing the funding terms between early-stage firms and investors. In mid-2014, another accelerator, 500 Startups, introduced a competing document, dubbed the Keep It Simple Security ("KISS"). Since then, these documents have become increasingly popular in financing early-stage ventures.[1] However, we know very little about the implications of future equity financing or the importance of its funding terms. This paper aims to bridge that gap.

In future equity financing, investors agree to transfer funds today against an undetermined amount of equity that will be permanently allocated in the future when certain conditions are met.

Contact: `idan.hodor@monash.edu`. We thank Jérôme Detemple, Marcel Rindisbacher, Zhiguo He, Darrel Duffie, Ilan Kremer, Lyndon Moore, Gabriele Lattanzio, Evgeny Lyandres, Guanming He, and seminar participants at Monash Business School, FMCG 2021, EFMA 2021, WFC 2021*.

[1] In a recent survey, Coyle and Green (2018) find that nearly all U.S. lawyers who specialize in startups signed at least one SAFE or KISS agreement.

In this paper, we provide a theory that studies these types of deals, and we do so in a continuous-time principal-agent setup. The continuous-time methodology is instrumental for this study. It allows us to analyze the repeated tension between exploration and exploitation, the risk-sharing properties of future equity financing deals, and better understand the importance of its two most critical provisions: the valuation cap and discount provisions. All of which has yet to be addressed. The empirical implication section provides a novel tool to compare the risk of different innovative technologies and empirically analyze investors' skill and return performance.

Specifically, investors want to invest in a high net present value innovative technology. The development of such innovative technology has a long-term and unknown maturity. The entrepreneur—the person with the skill to develop innovative technology—may avoid the developing time if she utilizes an already developed off-the-shelf technology instead. The tension between the two technologies comes from the cost of delaying the breakthrough. When the entrepreneur switches to utilize off-the-shelf technology, she postpones developing the innovative technology and extends the time it will eventually take to reach a breakthrough. This formulation is very well suited to study early-stage innovative projects; it is dynamically tractable and admits precise closed-form expressions. More importantly, it addresses the repeated, dynamic tension between exploration and exploitation.

Despite the low rewards from utilizing off-the-shelf technology, shortening the development time is valuable because it expedites the entrepreneur's successful future career. Our career concern agency friction is motivated by Graham, Harvey, and Rajgopal (2005)s' influential work. Their work emphasizes the impact of managers' career concerns on their decision-making and as the source of agency friction. They find that 78% of the executives in their sample admit to sacrificing long-term value due to career concerns.

Consistent with Burton, Sørensen, and Dobrev (2016), which view entrepreneurship as a career step and not as an end goal, we assume that the entrepreneur internalizes the future career benefits of fast technological deployment. Both developing innovative technology and utilizing off-the-shelf technology send a positive signal about the entrepreneur's characteristics. This positive signal advances the entrepreneur in her future career, and therefore, obtaining an early positive signal is valuable. It could be in better investment terms to finance her next idea. Alternatively, it could be in finding a high-paying corporate job.

Gompers, Kovner, Lerner, and Scharfstein (2010) document a manifestation of this channel. They find that entrepreneurs with a track record of success are much more likely to gain funding than first-time entrepreneurs and those who have previously failed. Further, Hall and Woodward (2010) show that entrepreneurs face a huge non-diversifiable risk and would benefit from selling some of the value they would obtain in a successful outcome in exchange for more wealth in the most likely failed outcome, suggesting that reducing the risk and the payoff is a beneficial strategic decision for entrepreneurs. Overall, empirical evidence suggests that, unlike investors, the entrepreneur prefers to utilize off-the-shelf technology. The extra career-related payoff creates a wedge between the entrepreneur and investors' preferred strategies, and it is the source of the agency friction in this model.

To fix ideas, unlike off-the-shelf technology, innovative technology takes time to develop. During the development phase, off-the-shelf technology produces a higher expected cash flow, and because of career benefits, it is valued more by the entrepreneur. If the entrepreneur develops innovative technology, it reaches a breakthrough at a random future date. However, because aligning incentives is potentially costly, the contract can be terminated before the breakthrough. Unlike the development phase, post-development, the innovative technology produces an expected cash flow high enough to offset the career benefits. From this point on, the agency friction disappears.

We resolve the conflict of interest with a financial contract that ensures the development of innovative technology all the time. The contract endogenously determines the maximal investors' equity while still providing sufficient incentives to the entrepreneur to develop the innovative technology. Since the optimal contract mandates the permanent allocation of equity conditional on future verifiable events, we coin this contract as a *future equity* contract.

For example, the entrepreneur can either develop a specialized artificial intelligence algorithm to tackle the speech recognition problem or, instead, can utilize an already developed picture recognition algorithm for that purpose. Investors prefer the specialized artificial intelligence algorithm because it has a higher net present value, while the entrepreneur prefers a fast deployment of the off-the-shelf algorithm, boosting her future career faster. The specialized algorithm pre-breakthrough net present value is \$16M, and the post-breakthrough net present value is \$91M. Investors agree to invest \$5M using a future equity contract, and there is no other source of funding. Accounting for the cost of motivating the entrepreneur to pursue the specialized algorithm, and given a 64%

likelihood that \$5M suffices to reach a breakthrough, firm value is \$10.26M, post-investment. As technology reaches breakthrough, firm value reaches \$14M, and investors' future equity converts to 44% equity; If the firm value had reached \$16M, investors' future equity would convert to 20% equity.

We obtain an exact closed-form characterization for the optimal contract, which leads to transparent and clear economic intuitions, and present our results in terms of firm value—an observable economic quantity.

Unconditional conversion to equity is not an equilibrium outcome. The optimal contract illustrates how investors and the entrepreneur share risk efficiently while motivating the entrepreneur to develop innovative technology all the time. In doing so, it gives rise to three important outcomes.

First, the optimal contract satisfies the main feature of convertible notes: the higher the firm value leading up to conversion, the fewer equity investors attain. Investors' maximal value is interior and attained such that the cost of early termination prior to conversion offsets the benefit of more equity upon conversion. This result is supported by Ewens, Gorbenko, and Korteweg (2019)s' empirical findings, documenting an optimal interior equity split between investors and entrepreneurs.

Second, we observe that positive pay-performance sensitivity discourages innovation, in line with Manso (2011)'s central insight and empirical evidence. A standard pay-for-performance contract would motivate the entrepreneur to pursue projects with the highest possible current expected cash flows. However, developing innovation produces low expected cash flow for some time, thus not aligning with a standard pay-for-performance scheme. The optimal contract punishes the entrepreneur for unexpected positive cash flows to prohibit deviations away from developing innovation.

Third, the optimal contract increases the entrepreneur risk exposure as firm value deteriorates. Intuitively, when the firm value is low, the entrepreneur takes on more risk as she reaps the benefits if things go well, but not the costs if things go poorly. Convex compensation, which in this model arises because the entrepreneur's future equity is always positive, is one of the leading explanations for managerial risk-taking, such as the one found by Brown, Harlow, and Starks (1996) in the mutual fund industry. We purport that entrepreneurs—similar to fund managers—take on more risk when firm values are low.

Investors' indirect utility is a superposition of a hyperbolic absolute risk aversion and a linear function. We obtain a convex absolute risk aversion in investors' future equity when the curvature parameter is less than two. In contrast, we obtain a decreasing absolute risk aversion in investors' future equity when the curvature parameter is higher than two. This result is in line with a typical financial investor, as Bitler, Moskowitz, and Vissing-Jørgensen (2005)s' empirical findings illustrated.

We implement the optimal contract with a convertible note with the following features. The note does not carry interest payments and does not have a maturity date. It contains a "valuation cap" provision that ensures investors' equity upon conversion cannot be too low and a time-varying "discount" provision, in which the discount increases with a lower investment. The note converts to equity under three verifiable conditions: (i) when technology reaches breakthrough, (ii) when investors' future equity reaches the valuation cap, (iii) when the firm defaults. Our implementation illustrates that both the SAFE and KISS instruments are in line with our model prediction.

A higher valuation cap is worse for investors. Although it may be tempting to reduce or remove the valuation cap during negotiation, our result highlights that it is crucial for developing innovation. That is, a contract with a valuation cap set too high is not optimal for developing innovation. We highlight this provision's importance and conjecture that it is widely used in convertible, SAFE, and KISS notes. When investors' future equity reaches the valuation cap, the contract reaches a first-best absorbing state, and the note converts to equity at the valuation cap. From this point forward, the contract guarantees that the innovative technology will reach a breakthrough. If the valuation cap provision was not triggered and the firm is still alive, the note unpredictably converts to equity when innovative technology reaches a breakthrough.

In addition to future equity, and as long as the convertible note is alive, the optimal contract mandates that the entrepreneur earns a compensation package for developing innovative technology. The package has two components—salary and incentives pay. The mixture of the two is state-dependent, and they counterbalance each other. As firm value grows, incentives pay converts to salary with a fixed conversion rate. The joint dynamics of the incentives pay and salary reveal the equilibrium's workings.

Finally, we introduce the implied probability of success—a novel tool for empirical analysis. It is hard to determine which innovative technology is a riskier investment opportunity because innovative technologies are different in many aspects. For example, would a \$2M investment in a new drug a riskeir investment than a \$3M in a new battery design? These two technologies have different characteristics, and the negotiated future equity terms could also be different. In some aspects, the first investment may be better, while in others, the second.

The implied probability of success transforms different investment opportunities to the same zero-one scale. It allows for a risk comparison across different innovative technologies. The implied probability of success measures the probability that technology reaches a breakthrough, as implied by the future equity contract and technology characteristics.

We illustrate two testable implications for the implied probability of success. In the first testable implication, we illustrate how to estimate investors' skills. Investors are skilled when they pay a low price for a given ex-ante success rate—when they consistently beat their odds. For example, if investment sizes are relatively big, probabilities of success approach one, and investors are unskilled if their realized success rate is below one. In contrast, when investment sizes are relatively small, probabilities of success approaches zero, and investors are skilled if their realized success rate is above zero. The implied probability of success is the ex-ante success rate. Therefore, for a portfolio of investments, skilled investors are investors with a realized success rate that is significantly higher than the average of the implied probabilities of success associated with those investments. Kaplan and Schoar (2005)s' seminal work pointed out that differential venture capital skill is the driving force behind funds' performance and persistence. Since then, the literature has fine-tuned the contribution of the fund's skill to return performance. The implied probability of success provides a direct way to measure this skill.

In the second testable implication, we illustrate how to correct investors' realized returns to account for selection bias. This bias arises because investors record successful investments more often than failed ones. If investors record only successful investments, the record occurs with probability given by the implied probability of success. Therefore, corrected returns can be calculated using the weighted average of realized returns with weights given by the implied probability of success. Since Cochrane (2005), the literature has emphasized the need to correct this selection bias to accurately measure investors' performance. The implied probability of success addresses this bias by transforming ex-post, realized returns to ex-ante, unbiased ones.

The paper is organized as follows. Section 2 summarizes the related literature; Section 3 setup the economy; Section 4 solves the optimal contract; Section 5 analyzes the optimal contract's implications; Section 6 implements the contract using a convertible note and a compensation package comprised of salary and incentives pay; Section 7 characterizes the implied probability of success and discusses its empirical implications; and lastly, Section 8 concludes and discusses future research avenues.

2 Related Literature

A recent interesting paper by Manso (2011) addressed a conceptually similar question to ours in a principal-agent context. The agent may either experiment with an innovative approach or exploit an already established approach. The main implication of his work is that pay-for-performance discourages the agent from experimenting with the new innovative approach. He utilizes a bandit problem in which experimentation affects the likelihood of success. That is, experimentation provides additional information.[2]

Unlike the experimentation literature, this paper has no learning; probability distributions and cash flows are known. We assume that going through the current iteration (experimenting) does not provide information about which iteration to try next. Instead of modelling the learning process, we model the time it takes to reach a breakthrough. This formulation is very well suited to study early-stage innovative projects, and, perhaps more importantly, it addresses the repeated, dynamic tension between exploration and exploitation, which has yet to be addressed in the literature.

Our article builds upon the multi-task agency literature, originated from Holmström and Milgrom (1991). Typically, in this literature, exerting effort on one task has a negative externality on another task. The myopic agency literature has expanded this focus by adding a time dimension. Specifically, exerting effort to increase today's output has a negative externality on future output. Recently, Zhu (2018) showed how an agent's persistent actions introduce a cliff-like contract that ties agent compensation today to high consecutive outputs leading up to today when the principal motivates the long-term action. Our model's main difference is that in Zhu (2018)'s model, manager actions have a persistent effect on output. Specifically, when an agent shirks today, the output today and tomorrow are affected. Varas (2018) studied a model of a project creation problem with

[2]There is a large and active literature on experimentation and learning, such as Bergemann and Hege (2005) and Halac, Kartik, and Liu (2016).

persistent effects on the firm's output.

Another related multi-task agency paper is Hoffmann and Pfeil (2018). In this paper, the agent either invests and reduces the expected cash flow today—and by doing so creates an opportunity to increase expected cash flow in the future—or does not invest and expected cash flow stays the same. Investment is the long-term action, which the principal prefers, but he does not observe whether the manager invests or steals. The main difference between our models is that in Hoffmann and Pfeil (2018) model, the technological shock occurs regardless of the manager's actions, whereas, in our model, a breakthrough is contingent on the entrepreneur's actions. Also related are Hellmann and Thiele (2011), and Hellmann (2007), which study how to motivate innovation in a multi-tasking agency model.

Our article also contributes to the literature on agency conflict between the entrepreneur and venture capitalists. To the best of our knowledge, we are the first to formalize an optimal continuous-time future-equity contract and analyze its dynamics and implications, including the inverse relation between investors' future equity and firm value and the necessity of the valuation cap provision. Further, we develop the implied probability of success and utilize it to estimate investors' skills and correct selection bias in reported returns, which provides new empirical predictions.

Focusing on future equity contracts, in a double moral hazard setup, Casamatta (2003) analyzed the optimal security design depending on the investment size. She found that if the entrepreneur's monetary contribution is relatively low, it is harder to induce her to work. In this case, the venture capitalist should get convertible bonds or preferred equity, which gives him relatively better remuneration in bad states. Further, Cornelli and Yosha (2003) analyzed a window-dressing problem. The entrepreneur has incentives to manipulate a short-term signal to reduce the probability that the project will be liquidated, even at the expense of reducing the project's long-term prospects. They show that convertible debt can prevent the entrepreneur from manipulating a short-term signal.

The literature on agency conflict between the entrepreneur and venture capitalists is long and active. One strand in this literature builds on the value-enhancing capabilities of venture capitalists. It is based on the presumption that venture capitalists contribute to firm value, for example, through the definition of strategy and financial policy, the professionalization of their internal organization, and the recruitment of key employees. Models in this strand of literature have mainly built upon a double moral hazard problem and address the extensive use of convertible securities to finance firms and include Cestone (2014), Hellmann (2006), Repullo and Suarez (2004), Schmidt (2003), Casamatta (2003) and Marx (1998). Another strand of literature has built upon venture capitalists' role as informed investors, which naturally gives rise to stage investing. In every stage, an investor may acquire information about a project's quality and decide whether to continue or abandon it. Admati and Pfleiderer (1994) were the first to address the moral hazard induced by this channel. They showed that when a venture capitalist has a fixed claim, it prevents strategic trading and induces optimal-continuation decisions. Later, Bergemann and Hege (1998), Habib and Johnsen (2000), Cornelli and Yosha (2003), and Dessí (2005) analyzed how financial contracts elicit information revelation and are useful in discriminating across projects and inducing efficient continuation or liquidation decisions. Recently, Ewens et al. (2019) employed a dynamic search and matching model to estimate the split of value between an entrepreneur and an investor. Similar to our model, they find an internally optimal equity split.

Methodologically, our article is most closely related to Hoffmann and Pfeil (2010) and Biais, Mariotti, Rochet, and Villeneuve (2010). The former embedded luck shocks in a continuous-time principal-agent model of DeMarzo and Sannikov (2006) and Biais, Mariotti, Plantin, and Rochet (2007). The latter analyzes the effects of accident risk, and similar to our model, in their model, the agent's actions induce changes in the likelihood of an accident occurring. Both papers use methodologies found in Sannikov (2008).

3 The Economic Setup

This section aims to develop a simple and tractable contract in a dynamic, continuous-time setting to allow for clear implications. The entrepreneur (agent) requires external capital to develop technology that investors (principal) contribute.[3] The contract endogenously determines the equity investors attain for their initial investment. Since the optimal contract mandates the permanent allocation of equity conditional on future events, we coin this contract as a *future equity* contract.

The entrepreneur privately decides whether to develop innovative technology or to utilize off-

[3]For clarity, investors are in the plural, and the entrepreneur is in the singular. Throughout this paper, it is always the same group of investors unless otherwise is stated explicitly.

the-shelf technology, a decision that can be reevaluated and changed at every instant of time and no cost.[4] For simplicity, there are two technologies: innovative technology produces cash flow characterized by

$$dX_t^1 = (\mu + \alpha N_t)\, dt + \sigma dZ_t, \tag{1}$$

and off-the-shelf technology produces cash flow characterized by

$$dX_t^0 = \mu_0 dt + \sigma dZ_t, \tag{2}$$

where $\mu, \alpha, \mu_0, \sigma$ are strictly positive parameters, and (Z_t) is a Brownian motion on the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The parameters μ and μ_0 capture the two technologies' expected cash flows, while α, referred to as potential, captures the jump in the innovative technology's expected cash flow upon reaching a breakthrough. We assume that both technologies have the same cash flow volatility to focus on the technologies' expected cash flows first-order effects.

To capture the tradeoff between the two technologies, we impose that

$$\mu + \alpha > \mu_0 > \mu. \tag{3}$$

This condition states that the off-the-shelf technology has higher expected cash-flow than the innovative technology before the breakthrough, but then, after the breakthrough the inequality flips and the innovative technology has higher expected cash-flow.

The process $(N_t) \equiv \mathbf{1}_{\tau \leq t}$ captures the breakthrough event, where time τ is exponentially distributed with intensity $\nu > 0$. Intuitively, τ is the (random) time required to develop innovative technology consistently, and t measures the amount of time the entrepreneur spends doing so. The innovative technology materializes only if the entrepreneur develops that technology. Once $N_t = 1$, the innovative technology reaches breakthrough, and from that time onward, the expected cash flow jumps to $\mu + \alpha$.

For tractability and expositional simplicity, the entrepreneur and investors discount future cash

[4]Since the optimal strategy is never to switch from developing the innovative technology, we can treat both technologies' expected values as net switching costs, as long as switching costs are not verifiable.

flow at the same rate r.[5] Let us denote the entrepreneur's actions by $\pi_t \in \{0, 1\}$, for $t \geq 0$, where $\pi_t = 1$ implies the development of innovative technology. Firm cash flow, therefore, evolves according to

$$dX_t^\pi = [\pi_t (\mu + \alpha N_t) + (1 - \pi_t) \mu_0] \, dt + \sigma dZ_t. \tag{4}$$

The desired "working" action is $\pi_t = 1$ and is referred to as developing innovation. The undesired "shirking" action is $\pi_t = 0$ and is referred to as utilizing off-the-shelf. Investors do not observe the entrepreneur action π_t, but only firm cash flow X_t^π; expected cash flow is internal and unobservable. We assume that breakthrough, captured by N_t, is observable and contractable, and once innovative technology reaches breakthrough, investors know right away. Thus, the only source of asymmetric information is that investors cannot disentangle the entrepreneur's hidden action π_t from the Brownian component in (4).

In utilizing off-the-shelf technology, the net present value is $\mathbb{E}\left[\int_0^\infty e^{-rs} dX_t^0\right] = \frac{\mu_0}{r}$, and due to career benefits of fast technological deployment, the entrepreneur obtains an additional value of $\frac{\lambda}{r}$.

In this reduced-form approach, the entrepreneur's objective function accounts for the effects of today's action on her future career; fast technological deployment expedites the entrepreneur path towards successful career. It could be in better investment terms to financing her next idea or finding a high-paying corporate job. Our reduced-form approach is motivated by Graham et al. (2005)s' influential work. They find that 78% of the executives in their sample admit to sacrificing long-term value due to career concerns. Further, Gompers et al. (2010) find that entrepreneurs with a track record of success are much more likely to gain funding than first-time entrepreneurs and those who have previously failed.

In contrast, in developing innovative technology, the net present value is

$$\mathbb{E}\left[\int_0^\infty e^{-rs} dX_t^1\right] = \frac{1}{r}\left(\mu + \frac{\alpha\nu}{r + \nu}\right) \equiv V_1^h. \tag{5}$$

[5]See Remark 3 for a further discussion about the entrepreneur's impatience and justification of this point.

Post-development, the net present value of innovative technology equals

$$\frac{\mu + \alpha}{r} \equiv V_0^h. \tag{6}$$

The subscripts 0, 1 stand for during and post-development, respectively.

Agency friction arises during the development phase because the entrepreneur prefers to utilize off-the-shelf technology while investors prefer to develop innovative technology. We conjecture that the entrepreneur's future equity is sufficiently low, that the entrepreneur value under off-the-shelf plus the career benefits are higher than the entrepreneur value under developing innovation: $q_t V_1^h \leq \frac{q_t \mu_0}{r} + \frac{\lambda}{r}$. That is, we conjecture that $q_t \leq \bar{q}$, and $\bar{q}$ is defined as

$$\bar{q} \equiv \min \left\{ \frac{\lambda}{r V_1^h - \mu_0}, 1 \right\}. \tag{7}$$

Post-development, the agency friction disappears. We conjecture that the entrepreneur's future equity, q_t, is sufficiently high, that the entrepreneur value under off-the-shelf plus the career benefits are not higher than the entrepreneur value under developing innovation: $q_t \frac{\mu_0}{r} + \frac{\lambda}{r} \leq q_t V_0^h$. That is, we conjecture that $\underline{q} \leq q_t$, and $\underline{q}$ is defined as

$$\underline{q} \equiv \frac{\lambda}{r V_0^h - \mu_0}. \tag{8}$$

These two conjectures capture the idea that if the entrepreneur's future equity is high, her incentives are aligned with investors, and innovation is developed without incentive provision. Alternatively, if the entrepreneur's future equity is low, her incentives are never aligned with investors, and innovation is never developed. Thus, we conjecture and verify in the optimal contract, Theorem 1, that indeed q_t satisfies

$$\underline{q} \leq q_t \leq \bar{q}. \tag{9}$$

The entrepreneur internalizes the future career benefits of fast technological deployment. Gompers et al. (2010) find that entrepreneurs with a track record of success are much more likely to gain funding than first-time entrepreneurs and those who have previously failed. Further, Hall and

Woodward (2010) show that entrepreneurs face a huge non-diversifiable risk and would benefit from selling some of the value they would obtain in a successful outcome in exchange for more wealth in the most likely failed outcome, suggesting that reducing the risk and the payoff is a beneficial strategic decision for entrepreneurs.

To develop innovation, the contract dynamically allocates future equity to the entrepreneur: $\mathcal{Q} \equiv \{q_t \in [0,1] : 0 \leq t < T\}$, an allocation that depends on the whole history of observable firm cash flow. The time it takes to develop innovation is random and unpredictable. Therefore, the entrepreneur's consumption flow at time t is given by $q_t dX_t^\pi + \lambda (1 - \pi_t) \, dt$. The endogenous termination time T is random and can take the value ∞ if the contract never terminates. If termination is triggered, the entrepreneur has an outside option given by $RL \geq 0$. Thus, the total expected entrepreneur value at $t = 0$ equals

$$W_0 = \mathbb{E}^\pi \left[\int_0^T e^{-rs} \Big\{ q_s dX_s^\pi + \lambda (1 - \pi_s) \, ds \Big\} + e^{-rT} RL \right]. \tag{10}$$

$\mathbb{E}^\pi$ represents the expectation under the probability measure induced by the entrepreneur's strategy $\pi = \{\pi_t : 0 \leq t < T\}$. L is the firm residual value when termination is triggered, and $0 \leq R \leq 1$ is the entrepreneur's share of the residual value L. On the other hand, investors receive a flow of payoff at time t equal to $(1 - q_t) \left(dX_t^\pi \right)$, and if the contract is terminated they receive a residual value $(1 - R) \, L$. We assume it is inefficient to terminate the contract, so $L < V_1^h$ and $(1 - R)L < (1 - \bar{q})V_1^h$. Liquidation is required and a part of an ex-ante optimal contract because the entrepreneur wealth is bounded from below. Investors' total expected payoff, therefore, equals

$$V_0 = \mathbb{E}^\pi \left[\int_0^T e^{-rs} \left(1 - q_s\right) \Big\{ dX_s^\pi \Big\} + e^{-rT} \left(1 - R\right) L \right]. \tag{11}$$

Investors and the entrepreneur commit to a long-term contract $\{\mathcal{Q}, T\}$. This contract specifies the entrepreneur's future equity $\mathcal{Q}$ and a termination time T that are both based on the complete history of observable cash flow $\{X_t^\pi, t \geq 0\}$ and innovation development $\{N_t, t \geq 0\}$. For a given contract $\{\mathcal{Q}, T\}$, the entrepreneur then decides on a feasible strategy π to maximize her value W_0. A strategy π is, therefore, incentive-compatible if it maximizes W_0 for a given contract $\{\mathcal{Q}, T\}$. An incentive-compatible contract is a triplet $\{\pi, \mathcal{Q}, T\}$ that induces the entrepreneur to pursue the innovation strategy $\pi_t = 1$ all the time.

To deepen our understanding, we describe the chain of events for any infinitesimal time interval $[t, t + dt]$ before the breakthrough event. First, the entrepreneur decides on her action π_t, and investors decide on the entrepreneur's future equity q_t. The realization of N_t is then revealed, which equals one only if the entrepreneur develops innovation, and lastly, investors decide whether to continue to operate $(t < T)$ or terminate the firm $(t \geq T)$.[6]

Remark 1 (Alternative Interpretation). *One may alternatively interpret the career concern benefits as a lower cost of effort, in line with the moral hazard literature. Utilizing off-the-shelf technology requires less effort, and by doing so, the entrepreneur gains the differential cost of effort between utilizing the off-the-shelf technology and the innovative technology.*

4 Optimal Contract

The contracting problem is to find an incentive-compatible contract that maximizes investors' expected profits (11) subject to delivering the entrepreneur her value (10) while also satisfying the constraint: $W_t \geq RL$, for all $0 \leq t \leq T$. We consider different devisions of bargaining power between investors and the entrepreneur, and allow W_0 to vary. For instance, if there is a competition between different investors, the entrepreneur enjoys all the bargaining power and receives the maximal value for W_0 under the optimal contract conditions.

We solve for the optimal contract using backward induction. Initially, we characterize the optimal contract after the technology reaches breakthrough (Section 4.1), and in a second step, we derive the optimal contract before the technology reaches breakthrough (Section 4.2). We assume that incentivizing the entrepreneur to develop innovation all the time is optimal, and verify its optimality later. We define $W_{t^-} \equiv \lim_{s \uparrow t} W_s$ to capture the entrepreneur value before the technology reaches a breakthrough. We omit the time subscript t^- when it does not confuse.

[6] Formally, this chain of events is captured by the fact that T is (X^π, N)-measurable stopping time, $\mathcal{Q}$ is predictable with respect to the filtration generated by (X^π, N), and π_t is predictable with respect to the filtration generated by (Z, N).

4.1 Post-development Phase

If the entrepreneur followed the recommended strategy and developed innovation, the firm expected cash flow jumps from μ to $\mu + \alpha$. Under the alternative strategy of utilizing off-the-shelf, expected cash flow remains at μ_0.

Looking forward, exactly when innovation reaches breakthrough, the agency friction disappears as the innovative technology generates the highest expected cash flow for both the entrepreneur and investors. As a result, precisely at that time, investors attain their permanent equity allocation $1 - q_\tau$, the entrepreneur attains the residual claim, q_τ, and the total surplus becomes V_0^h. Looking backward, exactly when innovation reaches breakthrough, the entrepreneur value W_{t^-} jumps by δ_t.

The promise-keeping condition ensures that the entrepreneur's *promised* value $(W_{t^-} + \delta_t)$ equals the *expected* value $(q_t V_0^h)$; therefore, investors set q_t such that

$$q_t V_0^h = (W_{t^-} + \delta_t), \qquad \underline{q} \leq q_t \leq \bar{q}. \tag{12}$$

The following proposition summarizes our findings.

Proposition 1 (Post-development). *Investors choose the entrepreneur's future equity q_t to satisfy the promise-keeping condition such that (12) always holds. Once innovative technology reaches breakthrough ($t = \tau$), both the entrepreneur value and investors' value function remain constant indefinitely; the entrepreneur attains $q_\tau V_0^h$, and investors $(1 - q_\tau)V_0^h$.*

4.2 Development Phase

We solve for investors' value function using the dynamic programming approach. We can adopt this approach because the intensity of developing innovation is constant and independent of the entrepreneur's past actions. Following the dynamic contracting literature, we summarise the state space using the entrepreneur's continuation value, which we refer to as *the entrepreneur value* throughout the paper. Specifically, the entrepreneur value at time t given that the entrepreneur develops innovation from time t onward equals

$$W_t = \mathbb{E}_t \left[\int_t^T e^{-r(s-t)} q_s dX_s^\pi + e^{-rT} RL \right]. \tag{13}$$

Next, we apply the Martingale Representation Theorem and characterize the evolution of the entrepreneur value. The following proposition summarizes our key findings.

Proposition 2 (**Incentive Compatibility**). *The dynamics of the entrepreneur value can be represented by*

$$dW_t = r(W_{t^-})dt - q_t dX_t^\pi + \beta_t \Big(dX_t^\pi - \mu dt\Big) + \delta_t \Big(dN_t - \nu dt\Big) \tag{14}$$

for some predictable processes $\beta = \{\beta_t : 0 \le t \le T\}$, $\delta = \{\delta_t : 0 \le t \le T\}$ *and* $q = \{q_t : 0 \le t \le T\}$. *Furthermore, the contract is incentive-compatible if and only if*

$$\beta_t (\mu_0 - \mu) \le \delta_t \nu - \lambda \qquad \text{for all } t \in [0, T]. \tag{15}$$

When investors set the cash flow risk exposure to zero ($\beta_t = 0$) and the benefits of developing innovation to zero ($\delta_t = 0$), the entrepreneur remains with a negative exposure to cash flow shocks $-q_t dX_t^\pi$ because $q_t \ge 0$. This result is not surprising, as the entrepreneur's future equity (q_t) represents a claim on all future cash flow. Like an asset with claims on future dividends, a positive cash flow shock (dividend inflow) reduces the entrepreneur value, while a negative cash flow shock (dividend outflow) increases it.

Investors set incentives such that the entrepreneur is discouraged from utilizing off-the-shelf technology. They achieve their goal by setting cash flow risk exposure (β_t) to offset what the entrepreneur would have obtained by switching from developing innovation to utilizing off-the-shelf. By switching to off-the-shelf, the entrepreneur gains the off-the-shelf expected cash flow ($\mu_0 \beta_t dt$) and additional career related benefits (λdt). However, the entrepreneur loses the expected cash flow obtained by developing innovative technology ($-\mu \beta_t dt$) and the expected benefits of developing innovation ($-\nu \delta_t dt$). Overall, the entrepreneur is discouraged from utilizing off-the-shelf technology when $(\mu_0 - \mu) \beta_t + \lambda - \delta_t \nu \le 0$. Unlike the traditional principal-agent setup such as DeMarzo and Sannikov (2006), in our setup, shirking entails a higher expected cash flow ($\mu_0 > \mu$), which implies that investors incentivize the entrepreneur against over-reporting as opposed to under-reporting.

Investors choose the benefits of developing innovation (δ_t) to set the marginal value of increasing δ_t to the marginal cost. On the value side, increasing δ_t facilitates investors with more control over exposure to cash flow risk since β_t can take higher values and further reduce the entrepreneur's exposure, as (15) illustrates. However, on the cost side, increasing δ_t also increases q_t (12), which increases the entrepreneur's exposure to cash flow risk ($q_t dX_t^{\pi}$). There is more cash flow risk to control, which indirectly reduces investors' control over the cash flow risk exposure.

Inefficiency arises through the risk of early termination, which occurs when the entrepreneur value falls to RL. To lower this probability, investors would like to increase β_t as much as possible; by doing so, they shrink the entrepreneur's overall exposure to cash flow shocks ($\beta_t - q_t$). However, increasing β_t too much violates the incentive compatibility restriction, and thus investors increase β_t just enough for incentive compatibility to bind, a characteristic similar to DeMarzo and Sannikov (2006).

We denote investors' value function as $V_1(W)$. We utilize Itô's Lemma for jump processes and find that investors' value function must satisfy the following Hamilton-Jacobi-Bellman (HJB) equation

$$rV_1(W_{t^-}) = \max_{q_t = \frac{W_{t^-}+\delta_t}{V_0^h}, \delta_t, \beta_t \leq \frac{\delta_t \nu - \lambda}{\mu_0 - \mu}} \Bigg\{ (1-q_t)\,\mu + V_1'\,(W_{t^-}) \Big(rW_{t^-} - q_t \mu - \delta_t \nu \Big)$$
$$+ \frac{1}{2}\sigma^2 V_1''\,(W_{t^-})\,(\beta_t - q_t)^2 + \nu \Big[(1-q_t)V_0^h - V_1(W_{t^-}) \Big] \Bigg\}. \tag{16}$$

Investors choose controls $\{q_t, \delta_t, \beta_t\}$ to maximize their value function, subject to the incentive compatibility restriction (15), the entrepreneur's future equity constraint (9), and the promise-keeping condition at breakthrough, (12). We conjecture and verify in the proof of Theorem 1 that (i) the value function $V_1(W_{t^-})$ is strictly concave for $RL \leq W < \bar{q}V_1^h$; (ii) the incentive compatibility restriction always binds, $\beta_t = \frac{\delta_t \nu - \lambda}{\mu_0 - \mu}$, which also implies that $\frac{\delta_t \nu - \lambda}{\mu_0 - \mu} < q_t$; (iii) the range of entrepreneur values $[RL, \bar{q}V_1^h]$ satisfies our future equity (q_t) restriction (9).

By plugging both the incentive compatibility and the promise-keeping conditions into the HJB and taking the first-order condition with respect to δ_t, we obtain

$$\delta_t = \frac{1 + V_1'(W_{t^-})}{V_1''(W_{t^-})} \frac{(\mu_0 - \mu)^2}{\sigma^2} \frac{V_0^h \left(V_0^h \nu + \mu\right)}{\left(V_0^h \nu + \mu - \mu_0\right)^2} + \frac{\lambda V_0^h + W_{t^-}\,(\mu_0 - \mu)}{\left(V_0^h \nu + \mu - \mu_0\right)}. \tag{17}$$

We plug the first-order condition (17) back to the HJB equation and find that investors' value function satisfies the following second-order differential equation:

$$(r+\nu)\, V_1(W_{t-}) = \left(V_0^h - W_{t-}\right)\left(\frac{\nu V_0^h + \mu}{V_0^h}\right) - \frac{1}{2}\frac{(\mu_0-\mu)^2}{\sigma^2}\left(\frac{\nu V_0^h + \mu}{\nu V_0^h + \mu - \mu_0}\right)^2 \frac{(1+V_1'(W_{t-}))^2}{V_1''(W_{t-})}$$
$$- \left(1+V_1'(W_{t-})\right)\left(\frac{\nu V_0^h + \mu}{V_0^h}\right)\left(\frac{\lambda V_0^h + W_{t-}\,(\mu_0-\mu)}{\nu V_0^h + \mu - \mu_0}\right) + V_1'(W_{t-}) W_{t-}\left(\frac{r V_0^h - \mu}{V_0^h}\right). \qquad (18)$$

We express this differential equation in closed-form, show that it maximizes investors' value, and delivers the value $W_0 \in [RL, \bar{q}V_1^h]$ to the entrepreneur. The following theorem summarizes our key findings.

Theorem 1 (**Optimal Contract**). *For a given entrepreneur value $W_0 \in [RL, \bar{q}V_1^h]$, investors' value function can be expressed in closed-form and equals to*

$$V_1(W_{t-}) = \left(V_1^h - W_{t-}\right) - \left(V_1^h - L\right)\left(\frac{\bar{q}V_1^h - W_{t-}}{\bar{q}V_1^h - RL}\right)^{\frac{1}{a}}, \quad RL \le W_{t-} \le \bar{q}V_1^h. \qquad (19)$$

The parameter V_1^h denotes the net present value of developing innovation. The curvature parameter satisfies $0 < a < 1$ and is given in (45). The incentive compatibility condition always binds, (15), the entrepreneur's future equity restrictions are satisfied, (9), and W_t evolves according to (14). Investors' controls are linear functions of W_{t-} and equal to

$$\delta_t = V_0^h q_t - W_{t-}, \qquad (20)$$
$$\beta_t = \frac{W_{t-}\nu + \lambda}{V_1^h \nu + \frac{\lambda}{\bar{q}}}(1-\phi_\beta) + \frac{W_{t-}}{V_1^h}\phi_\beta, \qquad (21)$$
$$q_t = \frac{W_{t-}\nu + \lambda}{V_1^h \nu + \frac{\lambda}{\bar{q}}}(1-\phi_q) + \frac{W_{t-}}{V_1^h}\phi_q. \qquad (22)$$

The parameters ϕ_q and ϕ_β are given in (52) and (53), respectively.

The optimal contract reveals that starting with $W_0 \in [RL, \bar{q}V_1^h]$ the entrepreneur's future equity, q_t, never goes strictly below $\underline{q}$ or strictly above $\bar{q}$. Clearly, when $W_0 < RL$, the entrepreneur is better off exercising the outside option and innovation is never developed. More importantly, the next section reveals that $W_0 = \bar{q}V_1^h$ is a first best absorbing state, implying that the entrepreneur develops the innovative technology without incentives. Thus, it is clearly not in the investors' best interest to offer $W_0 > \bar{q}V_1^h$, as the entrepreneur develops the innovative technology for less. In the next section we analyze the equilibrium mechanism and derive the optimal contract economic implications.

5 Implications

We refer to $W + V_1(W)$ as the *firm value* because the entrepreneur and investors' joint future equity is always a hundred percent. Interestingly, as (19) reveals, firm value monotonically increases with W:

$$\frac{\partial \left(V_1(W) + W\right)}{\partial W} > 0, \qquad \underline{q} \leq q_t < \bar{q}, \tag{23}$$

which allows us to present our intuitions and implications in terms of the empirically observable firm value.

We observe that unconditional conversion to equity is not an equilibrium outcome. It is more costly to unconditionally convert future equity to equity because it reduces investors' control over the entrepreneur's actions. The optimal contract mandates that future equity will convert to equity under well-defined verifiable conditions, described at length in the next section. The following proposition verifies that future equity satisfies the main convertible note feature.

Proposition 3 (**Future Equity**)**.** *The higher the firm value leading up to conversion, the fewer equity investors attain upon conversion,*

$$\frac{\partial(1 - q_t)}{\partial W_{t^-}} < 0, \qquad \underline{q} \leq q_t < \bar{q}. \tag{24}$$

The convertible note feature arises from the economic tradeoff investors face: balancing the cost early termination prior to conversion and benefits of more equity upon conversion. As firm value improves, the probability of early termination decreases; however, investors' future equity, $1 - q_t$, also decreases. For instance, at the extreme, if $q_t = 1$, investors would not attain any equity upon conversion — certainly not a desirable outcome. Instead, as firm value deteriorates, the probability of early termination increases; however, investors' future equity also increases. For instance, at the other extreme, if $q_t = 0$, investors would attain all the equity upon conversion but the contract terminates prior to breakthrough, again, not a desirable outcome. Consequently,

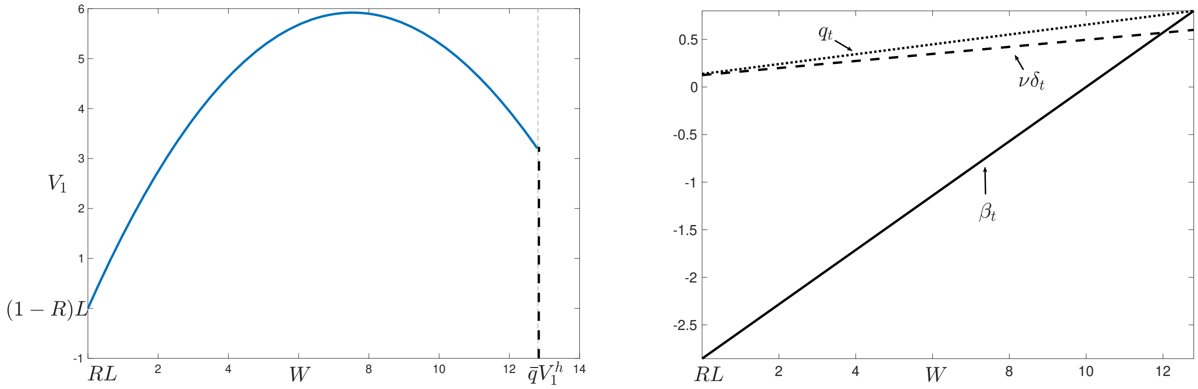

Figure 1. On the left-hand side, we have **investors' value function,** $V_1(W)$, as characterized in (19). On the right-hand side, we have the controls β_t, $\nu\delta_t$ and q_t as linear increasing functions of W. We observe that as the entrepreneur value deteriorates, the gap between q_t and β_t widens and increases the entrepreneur's overall exposure to cash flow shocks. As the entrepreneur value improves, the gap shrinks, and q_t and β_t converge to $\bar{q}$ and eliminate cash flow shocks' overall exposure. Plots are typical. Parameter values are $\mu_0 = 0.18$, $\mu = 0.05$, $\alpha = 4.5$, $\nu = 0.01$, $r = 0.05$, $\sigma = 0.8$, $\bar{q} = 0.8$, $\underline{q} = 0.1135$, and $L = 0$.

investors' maximal value is interior and attained such that the cost of early termination prior to conversion offsets the benefit of more equity upon conversion. This result is supported by Ewens et al. (2019)s' empirical findings, documenting an optimal interior equity split between investors and entrepreneurs. Figure 1 illustrates the tradeoff investors face.

Further, we observe that positive pay-performance sensitivity discourages innovation, in line with Manso (2011)'s central insight and empirical evidence. A standard pay-for-performance contract would motivate the entrepreneur to pursue projects with the highest possible current expected cash flows. However, by design, developing innovation produces low expected cash flow for some time, thus not aligning with a standard pay-for-performance scheme. The optimal contract punishes the entrepreneur for unexpected positive cash flows to prohibit deviations away from developing innovation. In other words, the contract motivates the entrepreneur to under-report cash flows until the breakthrough event, implying a negative pay-performance sensitivity.

Proposition 4 (Pay-Performance Sensitivity). *The entrepreneur pay-performance sensitivity is always negative*

$$\beta_t - q_t < 0, \qquad \underline{q} \le q_t < \bar{q}. \tag{25}$$

The optimal contract illustrates how to reduce the probability of early termination while still keeping the entrepreneur incentivized. In doing so, the optimal contract mandates that the entrepreneur loads on more risk in bad states.

As firm value deteriorates, the probability of early termination increases. In these bad states, the optimal contract increases investors' future equity. Doing so reduces the entrepreneur's benefits from developing innovation, as δ_t decreases with q_t (12). To keep the entrepreneur motivated, the optimal contract reduces the cash flow risk exposure β_t, (15). The decrease in cash flow risk exposure is steeper than the decrease in q_t, resulting in an overall increase in the entrepreneur's exposure to cash flow shocks when firm value is low.

Proposition 5 (**Risk-Exposure**). *The optimal contract increases the entrepreneur's risk exposure when firm value is low,*

$$\frac{\partial\left(q_t-\beta_t\right)}{\partial W_{t^-}}<0, \qquad \underline{q} \leq q_t<\bar{q}. \tag{26}$$

The reduction in cash flow risk exposure is steeper because a one unit reduction in cash flow risk exposure decreases the benefits of developing innovation (δ_t) by $(\mu_0-\mu)/\nu$ units, which then translates to a reduction by $(\mu_0-\mu)/\left(V_0^h \nu\right)$ units in q_t. Now, because the value of developing innovation always integrates to be higher than the value of utilizing off-the-shelf, $\nu V_0^h+\mu>\mu_0$, the decrease in cash flow risk exposure is steeper.

The optimal contract's tendency to increase the entrepreneur's risk exposure when firm value is low is in line with managers' risk-taking behavior when facing a convex compensation structure. A convex compensation structure, which in this model arises because firm value is non-negative, is one of the leading explanations of managerial risk-taking behavior, such as the one found by Brown et al. (1996) in the mutual fund industry.

The contract closed-form characterization highlights that investors' indirect utility, $V_1(W)$, is a superposition of a hyperbolic absolute risk aversion and a linear function. The risk aversion—curvature parameter $1/a$—is endogenously determined. Interestingly, when the curvature parameter is below 2, investors' indirect utility exhibits a state-dependent absolute risk aversion, yet another manifestation of the tradeoff investors face. When firm value is low, the risk of termination is high. In order to sustain such a high termination risk, investors' risk appetite increases. Instead, when firm value is high, investors' future equity may reach its lowest point, $1-\bar{q}$. In order to sustain this type of risk, investors' risk appetite also increases. Overall, investors' risk aversion is concave with respect to firm value. The inverse relation between investors' future equity, $1-q_t$, and firm value implies that investors' risk aversion is convex with respect to their future equity.

In contrast, when the curvature parameter is above 2, investors' indirect utility exhibits an increasing absolute risk aversion with respect to firm value, which implies a decreasing absolute risk aversion with respect to their future equity. This result is in line with Bitler et al. (2005)s' empirical findings, and a common assumption for a typical financial investor. The following proposition summarizes our key findings.

Proposition 6 (**Risk Aversion**). *Investors' indirect utility exhibits a state dependent absolute risk aversion, defined in (60). When the curvature parameter is high, $2 \leq 1/a$, absolute risk aversion decreases with investors' future equity. When the curvature parameter is low, $2 > 1/a$, it decreases for $(1-q_t) \geq 1-\tilde{Q}$ and increases for $(1-q_t) < 1-\tilde{Q}$, excluding the maximal argument and the high-boundary points in which the absolute risk aversion is not defined. The parameter $1-\tilde{Q}$ is given in (62).*

We rewrite the dynamics of the entrepreneur value, (14), as

$$dW_t = (r+\nu)\left(W_{t^-} - q_t V_1^h\right)dt + (\beta_t - q_t)\left(dX_t^\pi - \mu dt\right) + \delta_t dN_t \tag{27}$$

by plugging δ_t from (20) and using the identity $\nu V_0^h + \mu = (\nu + r) V_1^h$. This representation illustrates that as firm value approaches V_1^h, both the dt and dZ_t terms vanish. The dt vanishes because the entrepreneur value approaches $\bar{q}V_1^h$, and the dZ_t term vanishes because q_t and the cash flow risk exposure, β_t, both approach $\bar{q}$. The benefits of developing innovation, δ_t, approach $\bar{q}\left(V_0^h - V_1^h\right)$. Putting it all together, the dynamics of the entrepreneur value becomes

$$dW_t = 0dt + 0dZ_t + \bar{q}\left(V_0^h - V_1^h\right)dN_t,$$

and the highest entrepreneur value, $\bar{q}V_1^h$, is the first-best absorbing state. Once the entrepreneur value reaches $\bar{q}V_1^h$, the contract permanently allocates $\bar{q}$ equity to the entrepreneur and never terminates. Intuitively, when future equity reaches $\bar{q}$, the entrepreneur value from developing innovation equals the value from utilizing off-the-shelf (7). As a result, investors do not benefit from further increasing the future equity, and thus $\bar{q}V_1^h$ becomes the first-best absorbing state. When technology reaches a breakthrough, the entrepreneur value jumps to $\bar{q}V_0^h$, and investors remain with $(1-\bar{q})\,V_0^h$.

An important question remains. Are there any restrictions on $\bar{q}$? Investors prefer to reduce $\bar{q}$ as much as possible, while the entrepreneur prefers to increase it. A higher $\bar{q}$ means lower investors' minimal payoff, $(1-\bar{q})\,V_1^h$. As it turns out, $\bar{q}$ has an upper bound, and investors' minimal equity cannot be too low. If it is too low, developing innovation all the time becomes suboptimal.

Proposition 7 (Optimality). *Developing innovation is optimal if and only if*

$$\bar{q} \leq Q, \tag{28}$$

where $0 < Q < 1$ is a constant implicitly given in (65). The existence of Q is guaranteed if innovative technology is valuable enough relative to off-the-shelf technology, as given in (66).

Remark 2 (Equity Rights). *The dynamic allocation of future equity does not depend on the breakthrough event because it is unpredictable. The optimal contract dynamically allocates the future equity leading up to the breakthrough, and at the breakthrough event, converts the future equity to equity. The contract decides on the optimal split of future equity based on the whole history of cash flow.*

Remark 3 (Equally Patient Entrepreneur). *In the traditional setup such as DeMarzo and Sannikov (2006), the optimal contract balances the principal's costs and benefits of backloading payments to the agent. In their model, backloading is costly because the agent is impatient relative to the principal. When the agent is equally patient, backloading becomes costless. The principal finds it optimal to postpone payments indefinitely, and the optimal contract does not exist because the promise-keeping condition is violated. In contrast, in this model, the optimal contract allows for an equally patient agent because the tradeoff investors face is not due to backloading payments but rather due to a shrinking share of their future equity, as Proposition 3 and the pursuing analysis illustrate. Allowing the entrepreneur to be impatient would not eliminate this tradeoff, but rather superimpose the costly traditional channel with our new costly channel. We leave this extension* *to future research. He (2009) derived a similar conclusion when firm size evolves according to a geometric Brownian motion.*[7]

6 Convertible Note Implementation

So far, we have characterized the optimal contract between the entrepreneur and investors. In this section, we implement the contract using a convertible note with a "valuation cap" provision and a time-varying "discount" provision, without maturity date or interest payments. Our implementation illustrates that both SAFE and KISS instruments, variants of convertible notes, align with our model predictions.

Convertible debt and its recent variants have become very popular in financing early-stage firms, mostly through seed and first rounds. In late 2013, Y Combinator, one of the leading seed round ventures, introduced its convertible debt version with a SAFE note instrument. It is similar to a typical convertible debt except that it does not carry interest payments and does not have a maturity date. In 2014, in an attempt to standardize the use of convertible notes, 500 Startups, an early-stage venture fund, introduced KISS, another similar security, which is available in both a debt version and an equity version. The debt version has a maturity date and fixed interest payments, while the equity version is similar to a SAFE note.[8]

Generally, there are two important provisions negotiated in every SAFE and KISS: a valuation cap provision and a discount provision. The valuation cap provision ensures the minimum equity investors attain in the future, and the discount provision ensures that investors receive more equity than their investment otherwise suggests.

For instance, investors agree to invest \$5M using a future equity contract, and there is no other source of funding. Accounting for the cost of motivating the entrepreneur to pursue the innovative technology, and given a 64% likelihood that \$5M suffices to reach a breakthrough, firm value is \$10.26M, post-investment. As technology reaches breakthrough, firm value reaches

[7] In related work, He (2009) contrasted the optimal contract for two different cases of the agent's impatience. He analyzes the optimal contract when firm size evolves according to a geometric Brownian motion. He then shows that when the agent is equally patient, the optimal contract achieves a first-best absorbing state because a portion of incentives is costless when firm size is time-varying. While when the agent is impatient, the costless incentives are superimposed with a mechanism similar to the one analyzed in DeMarzo and Sannikov (2006).

[8] More information about SAFE can be found at the Y-Combinator webpage: https://www.ycombinator.com/documents/. More information about KISS can be found at the 500 Startups webpage original announcement: https://500.co/kiss/

\$14M, and investors' future equity converts to 44% equity. The contract provides a discount of roughly $(0.44/0.36-1)=22\%$ since without a discount, investors should attain $5/14=36\%$ equity. Had investors invested \$4M instead, the discount would have been more considerable. The note's valuation cap ensures that investors' equity is at least 20%.

The discount is the ratio between investors' equity under the contract and their equity implicit in their initial investment:

$$\frac{1-q_t}{K_0/\left(V_1\left(W\right)+W\right)}. \tag{29}$$

The representation reveals that the discount is bigger for smaller investments. Intuitively, a small investment is riskier because it implies a higher probability of termination before breakthrough. The contract compensates investors for riskier investments by providing bigger discounts upon conversion. For a bigger investment, the discount can become negative, which reflects a premium.

The optimal contract reveals that the higher the firm value leading up to conversion, the fewer equity investors attain, which means that a higher valuation cap is worse for investors. Increasing the valuation cap too much or removing it completely is not optimal for developing innovation, as Proposition 7 reveals.

The note in our implementation carries a "valuation cap" benefit, ensuring investors get at least $1-\bar{q}$ equity, and a "discount" relative to equity implicit in their investment, given by (29). The note converts to equity under three predetermined conditions: (i) If firm value reaches its bottom threshold L, the firm defaults, and the note converts to $1-R$ shares.[9] (ii) If firm value reaches its top threshold of V_1^h, the valuation cap provision is triggered, and the note converts to $1-\bar{q}$ shares. (iii) Alternatively, when neither valuation cap nor default has been triggered, technology unpredictably reaches breakthrough. At that time, the note converts to $1-q_\tau$ shares. The note does not have a maturity date, and until one of these three predetermined conditions is met, the note is alive.

Note that the convertible note does not depend on the amount of external capital K_0 investors inject. However, the entrepreneur and investors' initial payoffs depend upon K_0 and the parties' relative bargaining power. In exchange for the convertible note agreement, investors inject K_0

[9] Our default definition is similar to Merton (1974) in which firm value falls to the point that equity holders get nothing.

capital into the firm's bank account, which we assume equals the entrepreneur's initial value, $W_0 = K_0$. The bank account keeps track of the remaining cash the firm promises to the entrepreneur after accounting for future equity distributions. That is,

$$dK_t = dX_t^\pi + dW_t - (1 - q_t)\, dX_t^\pi \tag{30}$$

gives the dynamics of the bank account. In this formulation, we assume that cash inflows are paid immediately and saved by the entrepreneur. An offsetting outflow of $q_t dX_t^\pi$ is paid to the entrepreneur, and another offsetting outflow of $(1 - q_t)dX_t^\pi$ is paid to investors, for every dX_t^π inflow.

Further, to keep the entrepreneur incentivized, and for as long as the convertible note is alive, the firm pays the entrepreneur a compensation package for developing innovation. The package has two components: a continuous salary and incentive pay. The mixture of the two is state-dependent, and they counterbalance each other. Plugging dW_t, the bank account balance then becomes

$$dK_t = (r + \nu)\left(K_t - \int_0^t q_s dX_s^\pi\right) dt - q_t\Big((r + \nu)\left(V_1^h\right) - \mu\Big)dt + \beta_t\Big(dX_t^\pi - \mu dt\Big) + \delta_t dN_t. \tag{31}$$

Proposition 8 (Compensation). *The entrepreneur earns a continuous and positive salary of $q_t \left[(r + \nu)\, V_1^h - \mu\right]$, incentive pay of $-\beta_t$, and her future equity. Incentive pay converts to salary as firm value increases, with a fixed conversion rate given by $(r + \nu)\, V_1^h \frac{\partial q}{\partial \beta}$. The parameter $\frac{\partial q}{\partial \beta}$ is the elasticity of q_t with respect to β_t, which equals*

$$\frac{\partial q}{\partial \beta} = \frac{\phi_q \left(rV_1^h - \mu_0\right) + \nu V_1^h}{\phi_q \left(rV_1^h - \mu_0\right) \frac{\nu V_0^h}{\mu_0 - \mu} + \nu V_1^h} > 0. \tag{32}$$

The bank account balance (31) is bounded from below, and the parameter ϕ_q is defined in (52).

The joint dynamics of the incentive pay and salary are attractive and revealing the equilibrium's workings. When the firm value is low, the risk of termination is high. To mitigate that risk, investors convert a portion of the entrepreneur's salary to incentive pay. Reducing the entrepreneur's salary decreases the probability of termination, and increasing the incentives pay, increases the entrepreneur's risk exposure. A higher risk exposure increases firm value because the firm value is bounded from below. The firm pays the entrepreneur $-\beta_t \geq 0$ more than her future equity in these bad states, reducing the bank account balance after a positive cash flow.

In contrast, as firm value increases, the probability of reaching the valuation cap increases. At that point, investors are not concerned about termination but rather about a shrinking share of their future equity. Investors convert the incentive-pay back to salary to increase their future equity share and mitigate the probability of reaching the valuation cap. Investors reduce the incentive pay because reaching the valuation cap is better than reaching termination. The costs associated with a higher risk exposure are more significant than the benefits, resulting in an incentive pay reduction. The firm pays the entrepreneur $-\beta_t \leq 0$ less than her future equity in these good states, increasing the bank account balance after a positive cash flow.

Investors might consider selling their convertible notes before shares are issued. Potential reasons to do so are a liquidity shock that forces liquidation or a predefined strategy whereby investors source deals and sell them off. We assume the note's market price is $V_1(W)$, and the entrepreneur continues to develop innovation with the new investors. From the investors' perspective, $K_0^* = \arg\max V_1(W)$ is the optimal investment without considering a liquidity event as it minimizes the probability of termination and maximizes investors' future equity. However, if investors consider a liquidity event, they may consider investing less than K_0^* so that their value is higher than the initial investment, $V_1(K_0) > K_0$. By doing so, investors can source the deal and sell it right away with a positive return. Figure 2 illustrates this idea.[10]

[10] Additional equity financing rounds that do not give rise to agency issues and known at time t does not affect our qualitative results. The innovative technology's net present value is revised downward to QV_1^h, where $Q < 1$ captures the amount of equity that remained to be shared between investors and the entrepreneur. The entrepreneur's future equity is $q_t Q$, and investors' is $(1 - q_t)\, Q$.

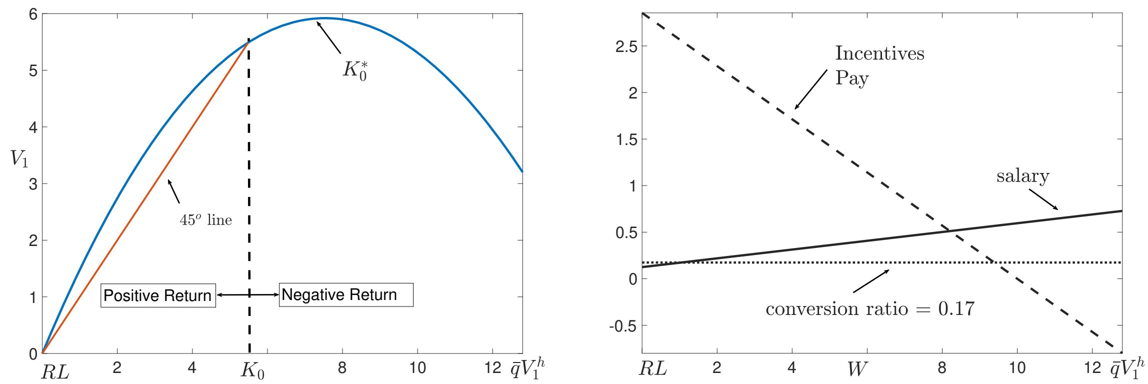

Figure 2. Example: Suppose $K_0 = \$3M$. On the right graph, salary moves between $\$125K$ to $\$728K$ per unit time. The incentive pay is 285% when firm value is low and deleverages to -80% when firm value is high; incentive pay converts to salary as firm value increases; a 1% decrease in incentive pay increases salary by 0.17%. On the left graph, investors' value at the time of investment is $V_1(3) = \$3.79M > \$3M$, which implies that investors' return is roughly 26.5% if sold right at the time. If instead, investors choose to retain their note until conversion to equity is triggered, they make \$0 if development was unsuccessful (termination) and from $\$18.2M$ to $\$80.7M$ if development was successful. Thus, investors return between $6X$ to roughly $27X$ times their initial investment. The parameter values are the same as in Figure 1.

7 The Implied Probability of Success

In this section, we formalize the implied probability of success—a novel tool for empirical analysis. We then show how to utilize it for two important cases.

It is hard to decide which innovative technology is a riskier investment opportunity because innovative technologies are different in many aspects. For example, would a \$2M investment in a new drug a riskier investment than a \$3M in a new battery design? These two technologies have different characteristics, such as breakthrough likelihood (ν) and potential (α), among other aspects. The negotiated future equity terms could also be different, such as different investment size, valuation cap, and residual value.

The implied probability of success transforms different investment opportunities to the same zero-one scale. It allows for a risk comparison across different innovative technologies. The implied probability of success measures the probability that technology reaches a breakthrough, as implied by the future equity contract and technology characteristics. This section formalizes the implied probability of success.

We begin by characterizing the map between the entrepreneur value and the implied probability of success and show that it is a one-to-one map. We then utilize the implied probability of success to compare different innovative technologies.

To characterize the implied probability of success, we split all possible trajectories of W into two mutually exclusive sets: a trajectory of W either ends up hitting RL, and the firm defaults ($T < \tau$), or technology reaches breakthrough first, and default never occurs ($T \geq \tau$). We can split the trajectories because the breakthrough intensity is strictly positive, $\nu > 0$, implying that technology eventually reaches a breakthrough ($\tau < \infty$).

Consider the following thought experiment. Suppose that investors choose their optimal termination time T to maximize their $t = 0$ payoff, conditional on either one of the two mutually exclusive sets just described. What would be the maximal (conditional) payoff?

The answer to this question is straightforward once we realize that investors cannot predict the time it takes technology to reach a breakthrough. Consider the first event. To ensure that termination occurs before a breakthrough ($\tau > T$), investors have to terminate the contract immediately ($T = 0$). More precisely, at any $T > 0$, there exists a positive measure set of trajectories of W that violates the conditional event ($\tau > T$). As a result, conditional on termination occurring before the breakthrough, firm value equals the residual value L. Consider the complementary event. To ensure that breakthrough occurs before termination ($\tau \leq T$), investors never terminate the contract ($T = \infty$). More precisely, at any finite termination time ($T < \infty$), there exists a positive measure set of trajectories of W that violates the conditional event ($\tau \leq T$). As a result, conditional on the breakthrough occurring before default, firm value equals the net present value (V_1^h).

The two conditional firm values are constant and can be represented as the probability of the conditional event multiplied by the corresponding constant. Thus, we represent the (unconditional) value function ($V_1(W)$) as a function of the implied probability of success:

$$V_1(W) + W = \mathbb{E}^\pi \left[L \mathbf{1}_{\tau > T} \right] + \mathbb{E}^\pi \left[V_1^h \mathbf{1}_{\tau \leq T} \right] = L \Pr\left[\tau > T\right] + V_1^h \Pr\left[\tau \leq T\right].$$

Solving for $\Pr\left[\tau \leq T\right]$, plugging $V_1(W)$ from (19), we find that

$$\Pr\left[\tau \leq T\right](W) = 1 - \left(\frac{\bar{q} V_1^h - W}{\bar{q} V_1^h - RL} \right)^{\frac{1}{a}}. \tag{33}$$

The following proposition formalizes our intuition and summarizes our findings.

Proposition 9 (**Implied Probability of Success**). *For a given entrepreneur value, the implied probability of success admits a closed-form characterization given in (33). It is a one-to-one map that increases with firm value,* $\frac{\partial \Pr[\tau \leq T](W)}{\partial W} > 0$*, starts at zero,* $\Pr[\tau \leq T](RL) = 0$*, and ends at one,* $\Pr[\tau \leq T](\bar{q}V_1^h) = 1$*. Also, it decreases with the valuation cap and the residual value*

$$\frac{\partial \Pr[\tau \leq T](W)}{\partial \bar{q}} \leq 0, \qquad \frac{\partial \Pr[\tau \leq T](W)}{\partial L} \leq 0. \tag{34}$$

The three items are intuitive. As the entrepreneur value deteriorates, the chance of hitting the lower threshold increases, and thus the probability that technology reaches breakthrough decreases. When the entrepreneur value reaches its lowest value, the firm defaults, and the implied probability of success becomes zero. In contrast, as the entrepreneur value improves, the chance of hitting the lower threshold decreases, and thus the implied probability of success increases. When the entrepreneur value reaches its upper bound, the drift and volatility of dW_t vanish. The risk of default completely disappears, and the implied probability of success becomes one. It is now clear that more significant capital investment is mapped to a higher implied probability of success, and vice versa.

The contract characteristics affect the implied probability of success. When the residual value increases, the risk of default increases, reducing the implied probability of success. Further, when the valuation cap increases, the chances of hitting the upper bound $\bar{q}V_1^h$ decreases, reducing the implied probability of success.

Both the implied probability of success and investors' value function contains the same information. As such, the implied probability of success summarizes the information investors need to make an informed decision about investments across different technologies. We illustrate how intuitive implied probabilities of success are in an example given in Figure 3.

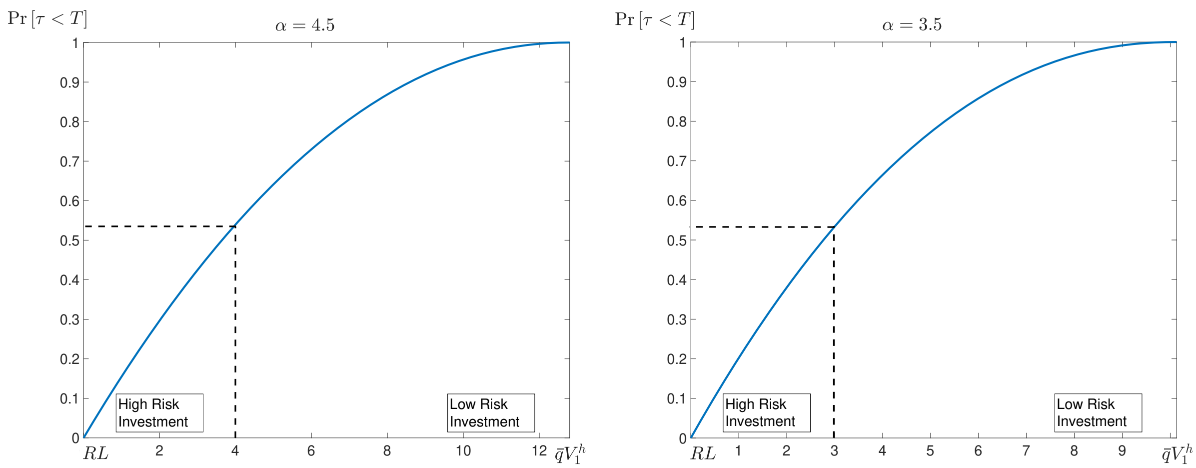

Figure 3. These two figures capture the implied probability (y-axis) for any given investment size (x-axis). Investing $K_0 = 4$ in the left graph technology is as risky as investing $K_0 = 3$ in the right graph technology. Both have an implied probability of success of approximately $\Pr[\tau \leq T] \approx 54\%$. The two technologies differ only in the expected cash flow jump size α. In the left figure, $\alpha = 4.5$, while in the right figure, $\alpha = 3.5$. We emphasize that even though α has no direct effect on the intensity of developing innovation ($\nu = 0.01$), it indirectly affects the implied probability of success through equilibrium. The rest of the parameter values are the same as in Figure 1.

7.1 Empirical Implications

One of the main contributions of the implied probability of success is that it is empirically easier to measure. This section illustrates this idea in two central cases. We assume that investment investors make equals the entrepreneur's initial value, $K_0 = W_0$.

In the first case, we illustrate how to estimate the skill of an individual investor. Investors are skilled when they pay a low price for a given success rate—when they consistently beat their odds. The implied probability of success is the priced success rate because it measures the probability of success for a given investment amount exactly when the deal is signed. Specifically, consider an investor who made N different investments that materialized in the past. Ex-post, M of those investments are considered successful, which gives this investor a realized success rate of M/N. At the time of investment, we back-out firm value from investment size and then infer the implied probability of success from this past firm value. Therefore, we obtain an estimator for skilled investors:

Empirical Implication 1 (**Skill**). *Skilled investors are those with a realized success rate that is consistently higher than their average of implied success rates,*

$$\frac{M}{N} > \frac{1}{N}\sum_{i=1}^{N} \Pr\left[\tau < T\right]_i . \tag{35}$$

The left-hand side is the realized success rate, while the right-hand side is the average success rate implied by investment size; both the realized and the implied success rates account for failed and successful investments. For example, if investment sizes are relatively big, probabilities of success approach one, and investors are considered unskilled if their realized success rate is below one. In contrast, when investment sizes are relatively small, probabilities of success approaches zero, and investors are considered skilled if their realized success rate is above zero. Our implied probability of success creates a new testable implication contributing to the analysis of venture capitalists' skills. Kaplan and Schoar (2005)s' seminal work pointed out that differential venture capital skill is the driving force behind funds' performance and persistence. Since then, the literature has fine-tuned the contribution of the fund's skill to return performance. The implied probability of success provides a direct way to measure this skill.

In the second case, we illustrate how to correct investors' realized returns to account for selection bias. This bias arises because investors record successful investments more often than failed ones, which implies that their reported realized returns are much higher than their actual returns. Consider an investor who realized a positive net return $R > 0$ on a particular investment:

$$R = \frac{\left(1 - q_\tau\right) V_0^h - K_0}{K_0}.$$

If the investor records only successful investments, the record occurs with probability $\Pr\left[\tau < T\right]$. Therefore, we obtain a direct measure to correct realized returns:

Empirical Implication 2 (**Selection Bias**). *The corrected return $\hat{R}$ is a weighted average of the realized, reported return, and hundred percent loss. The implied probability of success gives the weights,*

$$\hat{R} = R\big(\Pr\left[\tau < T\right]\big) - 1\big(1 - \Pr\left[\tau < T\right]\big), \tag{36}$$

assuming that $L = 0$.

The corrected return considers that with probability $1 - \Pr[\tau < T]$, the investor would have made a hundred percent loss and would not record it.

Our implied probability of success creates a new testable implication contributing to the analysis of venture capitalists' return performance. The literature on venture capital performance has focused on return-on-investment as a tool to compare across different investments. However, it is an ex-post measure that suffers from selection bias. As Cochrane (2005) pointed out, we are likely to see returns of those successful firms. Since then, the literature has emphasized the need to correct this selection bias to accurately measure investors' performance. Our measure mitigates this bias by transforming ex-post returns to ex-ante ones.

8 Conclusion

This article studies the equilibrium implications of future equity financing in a continuous-time principal-agent setup. The entrepreneur's career concerns generate a moral hazard tension, which investors address with a financial contract that promises equity in the future. The optimal contract admits precise closed-form expressions and guarantees that the entrepreneur develops innovative technology all the time.

The optimal contract satisfies the main feature of convertible notes: the higher the firm value leading up to conversion, the fewer equity investors attain. We point that positive pay-performance sensitivity discourages innovation, in line with Manso (2011)'s central insight, and the entrepreneur risk exposure increases as firm value deteriorates.

We implement the contract with a convertible note similar to a SAFE and KISS. The convertible note has no interest payment, no maturity date, and it contains a valuation cap provision that ensures investors' future equity cannot be too low, and a time-varying discount provision that increases with a lower investment. The valuation cap benefit is necessary; if it is set too high, developing innovative technology all the time is suboptimal. Our model emphasizes the importance of these two provisions. It would be interesting to investigate if the valuation cap provision is more prevalent in firms pursuing the innovative technology, and if investors attain a higher discount when their investments are lower.

Lastly, we introduce the implied probability of success, an investor-specific performance measure that can be assessed empirically. We show how to utilize it to estimate investors' skills and correct selection bias in realized returns.

Our model generates new testable predictions. Most notably, those introduced by the implied probability of success. It would be interesting to take the implied probability of success to the data and contrast its outcomes against other methods to estimate investors' skills and correct selection bias in realized returns. Further, our model predicts that future equity and convertible notes are optimal to finance innovation. It would be interesting to contrast firms with equity seed round financing with future equity ones. Are future equity ones more innovative? Do firms with future equity financing reach a breakthrough with higher probability? These and other questions yet to be answered.

A Proofs

Proof of Proposition 1 (Post-development). Post-development, for any $q_t \in [\underline{q}, \bar{q}]$, the innovative technology generates the highest value for both the entrepreneur and investors, as our assumption (8) illustrates. Consequently, for any feasible q_t, both the entrepreneur value and investors' value function remain constant indefinitely. Investors choose the entrepreneur's future equity q_t to satisfy the promise-keeping condition such that (12) always holds; the entrepreneur attains $q_\tau V_0^h$, and investors the residual $(1 - q_\tau)V_0^h$. This result is a direct consequence of the promise-keeping condition, which requires that at every point in time, the entrepreneur's (promised) continuation payoff exactly matches her expected continuation payoff, as DeMarzo and Fishman (2007) illustrated.

The entrepreneur's promised continuation payoff coming-in to time t is $(W_{t^-} + \delta_t)$, for some feasible entrepreneur's future equity, q_t. Therefore, if $q_t \neq \frac{W_{t^-}+\delta_t}{V_0^h}$, then the promise-keeping condition is violated. We verify that $q_t \in [\underline{q}, \bar{q}]$ in Theorem 1. □

Proof of Proposition 2, (Incentive Compatibility). Let us define G_t^e as the entrepreneur lifetime wealth evaluated conditional on information at time t, $G_t^e \equiv \mathbb{E}_t \left[\int_0^T e^{-rs} \left(q_s dX_s^\pi\right) + e^{-rT} LR\right]$. By using (13), we write G_t^e such that

$$G_t^e = \int_0^t e^{-rs} q_s dX_s^\pi + e^{-rt} W_t. \tag{37}$$

By construction, G_t^e is a martingale with respect to the filtration generated by (Z, N) under the probability measure induced by innovative strategy π, $\mathbb{P}^\pi$. Therefore, by the Martingale Representation Theorem,

$$G_t^e = G_0^e + \int_0^t e^{-rs} \beta_s \sigma dZ_s + \int_0^t e^{-rs} \delta_s \left(dN_s - \nu ds\right),$$

where we express the Brownian increments as a function of observable performance, $dZ_t = \frac{dX_t^{\pi} - (\mu + \alpha N_t)dt}{\sigma}$. By taking a derivative of (37) and using the above martingale representation, we obtain (14).

We now show that the contract $\{\{q_t : 0 \leq t \leq T\}, T\}$ is incentive-compatible if and only if $\beta_t (\mu_0 - \mu) \leq \delta_t \nu - \lambda$. Consider an arbitrary strategy $\hat{\pi}$ and define

$$\hat{G}_t^e = \hat{G}_0^e + \int_0^t e^{-rs} \left(q_s dX_s^{\hat{\pi}} + (1 - \hat{\pi}_s) \lambda ds \right) + e^{-rt} W_t \tag{38}$$

as the lifetime expected utility evaluated conditional on information at time t when the entrepreneur follows $\hat{\pi}$ up to time t and then switches to π after time t. We find that

$$\begin{aligned}
e^{rt} d\hat{G}_t^e &= q_t dX_t^{\hat{\pi}} + (1 - \hat{\pi}_t) \lambda dt + e^{rt} d\left(e^{-rt} W_t\right) \\
&= q_t dX_t^{\hat{\pi}} + (1 - \hat{\pi}_t) \lambda dt - q_t dX_t^{\hat{\pi}} + \beta_t \Big(dX_t^{\hat{\pi}} - \mu dt \Big) + \delta_t \Big(dN_t^{\hat{\pi}} - \nu dt \Big) \\
&= q_t \left(dX_t^{\hat{\pi}} - dX_t^{\hat{\pi}} \right) + \left(\beta_t \left(\mu^{\hat{\pi}} - \mu \right) + (\lambda - \delta_t \nu) (1 - \hat{\pi}_t) \right) dt + \beta_t \left(dX_t^{\hat{\pi}} - \mu^{\hat{\pi}} dt \right) + \delta_t \Big(dN_t^{\hat{\pi}} - \nu \hat{\pi}_t dt \Big) \\
&= \Big(\beta_t (\mu_0 - \mu) + \lambda - \delta_t \nu \Big) (1 - \hat{\pi}_t) dt + \beta_t \left(dX_t^{\hat{\pi}} - \mu^{\hat{\pi}} dt \right) + \delta_t \Big(dN_t^{\hat{\pi}} - \nu \hat{\pi}_t dt \Big),
\end{aligned}$$

where $\mu^{\hat{\pi}} = [\pi_t (\mu + \alpha N_t) + (1 - \pi_t) \mu_0]$ and $N_t = 0$ during the development phase, by taking a derivative and plugging (14). We obtain the first equality by taking the derivative of (38) and multiplying by e^{rt}; the second by applying Itô's Lemma for $e^{-rt} W_t$ and plugging (14); the third by changing measures to identify the martingale part and the drift part; and the fourth by direct manipulation. The drift is non-positive if and only if

$$\beta_t (\mu_0 - \mu) + \lambda - \delta_t \nu \leq 0,$$

which implies that $\hat{G}_t^e$ is a supermartingale under any feasible strategy and a martingale only under the recommended innovative strategy of $\pi_t = 1$. If the strategy π_t does not satisfy (15) on a set with positive measure, during $[0, T)$, let us choose $\hat{\pi}_t = 0$ when (15) is not satisfied and $\hat{\pi}_t = \pi_t = 1$ otherwise. Then, $\{\pi_t = 1; 0 \leq t < T\}$ is suboptimal. We verify the restriction (9) in Theorem 1. $\square$

Proof of Theorem 1 (Optimal Contract). First, from (5) and (6), we find that $\nu V_0^h + \mu = V_1^h (\nu + r)$ and $r V_0^h - \mu = \alpha$. By plugging these simplifications back to (18), our second-order differential equation becomes

$$\begin{aligned}
V_1(W_{t-}) = \Big(V_0^h - W_{t-} \Big) \left(\frac{V_1^h}{V_0^h} \right) - \frac{1}{2} \frac{(\mu_0 - \mu)^2}{\sigma^2} \left(\frac{V_1^h (\nu + r)}{V_1^h (\nu + r) - \mu_0} \right)^2 \left(\frac{1}{r + \nu} \right) \frac{(1 + V_1'(W_{t-}))^2}{V_1''(W_{t-})} \\
- \Big(1 + V_1'(W_{t-})\Big) \left(\frac{V_1^h}{V_0^h} \right) \left(\frac{\lambda V_0^h + W_{t-} (\mu_0 - \mu)}{V_1^h (\nu + r) - \mu_0} \right) + V_1'(W_{t-}) W_{t-} \left(\frac{\alpha}{V_0^h} \right) \left(\frac{1}{r + \nu} \right).
\end{aligned} \tag{39}$$

Next, we verify that our suggested solution (19) satisfies the second-order differential equation (39). Toward that goal, let us define $k_1 \equiv \tilde{k}_1 (a)^{\frac{1}{a}}$ and $b \equiv \bar{q} V_1^h a$. By using these definitions, our suggested solution then becomes

$$V(W_{t-}) = V_1^h - W_{t-} - \tilde{k}_1 (b - a W_{t-})^{\frac{1}{a}}, \quad RL \leq W_{t-} \leq \bar{q} V_1^h. \tag{40}$$

We proceed by characterizing the constants a, b/a and k_1. We first show that they match our characterization for

V_1, δ_t, β_t, and q_t. We then show that these characterizations satisfy our conjectures: (i) $V''(W_{t-}) \leq 0$, (ii) $\beta_t \leq q_t$, (iii) $\underline{q} \leq q_t \leq \bar{q}$, for any $W_{t-} \in [RL, \bar{q}V_1^h]$, with strict inequalities everywhere except for $W_{t-} = \bar{q}V_1^h$. Plugging the derivatives of the value function (40) back to the second-order differential equation (39), we obtain the following equation:

$$(b - aW)\left[-1 - \frac{1}{2}\left(\frac{\mu_0 - \mu}{\sigma}\right)^2 \left(\frac{V_1^h(\nu + r)}{V_1^h(\nu + r) - \mu_0}\right)^2 \frac{1}{r + \nu}\frac{1}{1 - a}\right]$$
$$= \left(\frac{V_1^h}{V_0^h}\right)\left(-\frac{\lambda V_0^h + W_{t-}(\mu_0 - \mu)}{V_1^h(\nu + r) - \mu_0} + \frac{\alpha W}{V_1^h(r + \nu)}\right). \tag{41}$$

This last equation has to hold for any W, which leads to two equations—one for the free terms and one for the W term

$$b\left[-1 - \frac{1}{2}\left(\frac{\mu_0 - \mu}{\sigma}\right)^2 \left(\frac{V_1^h(\nu + r)}{V_1^h(\nu + r) - \mu_0}\right)^2 \frac{1}{r + \nu}\frac{1}{1 - a}\right] = \left(\frac{V_1^h}{V_0^h}\right)\left(-\frac{\lambda V_0^h}{V_1^h(\nu + r) - \mu_0}\right), \tag{42}$$

$$-a\left[-1 - \frac{1}{2}\left(\frac{\mu_0 - \mu}{\sigma}\right)^2 \left(\frac{V_1^h(\nu + r)}{V_1^h(\nu + r) - \mu_0}\right)^2 \frac{1}{r + \nu}\frac{1}{1 - a}\right] = \left(\frac{V_1^h}{V_0^h}\right)\left(-\frac{(\mu_0 - \mu)}{V_1^h(\nu + r) - \mu_0} + \frac{\alpha}{V_1^h(r + \nu)}\right), \tag{43}$$

respectively. Dividing b by $-a$, we obtain

$$-\frac{b}{a} = \frac{\left(-\frac{\lambda V_0^h}{V_1^h(\nu + r) - \mu_0}\right)}{\left(-\frac{(\mu_0 - \mu)}{V_1^h(\nu + r) - \mu_0} + \frac{\alpha}{V_1^h(r + \nu)}\right)} = \frac{-\lambda V_0^h}{\mu + \alpha - \frac{\mu_0}{V_1^h}\left(V_1^h + \frac{\alpha}{r + \nu}\right)} = -\frac{\lambda}{r - \frac{\mu_0}{V_1^h}} = -\bar{q}V_1^h. \tag{44}$$

We obtain the first equality by dividing (42) by (43), the second by simplifying the fraction, the third by observing that $V_1^h + \frac{\alpha}{r+\nu} = V_0^h$, and the fourth by observing that $rV_1^h - \mu_0 = \frac{\lambda}{\bar{q}}$. Therefore, we conclude that $\frac{b}{a} = \bar{q}V_1^h$. Next, (43) is quadratic in a. Our desired solution is the lower solution of this quadratic; that is, a satisfies

$$a = \frac{1 + X + Y - \sqrt{(1 + X + Y)^2 - 4Y}}{2}, \tag{45}$$

where

$$X = \frac{1}{2}\left(\frac{\mu_0 - \mu}{\sigma}\right)^2 \left(\frac{V_1^h(\nu + r)}{V_1^h(\nu + r) - \mu_0}\right)^2 \frac{1}{r + \nu} > 0, \tag{46}$$

$$Y = \left(\frac{V_1^h}{V_0^h}\right)\left(-\frac{(\mu_0 - \mu)}{V_1^h(\nu + r) - \mu_0} + \frac{\alpha}{V_1^h(r + \nu)}\right) = \frac{rV_1^h - \mu_0}{(r + \nu)V_1^h - \mu_0} = \frac{\lambda}{\bar{q}V_1^h\nu + \lambda} > 0. \tag{47}$$

The second-to-last equality in Y is obtained by simplifying the fraction and observing that $V_1^h + \frac{\alpha}{r+\nu} = V_0^h$, and the last by observing that $\bar{q}\left(rV_1^h - \mu_0\right) = \lambda$. Furthermore, $X > 0$ and $Y > 0$ because investors always prefer to develop innovation over utilizing off-the-shelf (6). We now show that a, given in (45), always exists and satisfies $0 < a < 1$ for any set of parameters. It is straight forward to verify that the parameter a exists, it is strictly less than one, and strictly bigger than zero because $X, Y > 0$. Therefore, our suggested solution in (19) solves the differential equation in (18) for any given k_1.

To finish the value function's characterization, we pin down k_1 using the value-matching condition as follows.

Whenever $W_{t-} \leq RL$, the entrepreneur is better off exercising her outside option RL. Therefore, the value-matching condition is $V_1(RL) = (1-R)L$. When we plug this relation to (40), we obtain

$$k_1 = \frac{(V_1^h - L)}{\left(\bar{q}V_1^h - RL\right)^{\frac{1}{a}}}. \tag{48}$$

Observe that because terminating the contract is inefficient ($V_1^h > L$), the coefficient is always positive, $k_1 > 0$. We have completed the characterization of the value function.

Further, by plugging our suggested solution to the first-order condition for δ_t, (17), we find that

$$\delta_t = -\frac{b - aW_{t-}}{1-a} \frac{(\mu_0 - \mu)^2}{\sigma^2} \frac{V_0^h \left(V_0^h \nu + \mu\right)}{\left(V_0^h \nu + \mu - \mu_0\right)^2} + \frac{\lambda V_0^h + W_{t-} (\mu_0 - \mu)}{\left(V_0^h \nu + \mu - \mu_0\right)}, \tag{49}$$

$$\beta_t = \frac{\delta_t \nu - \lambda}{\mu_0 - \mu} = -\frac{b - aW_{t-}}{1-a} \frac{(\mu_0 - \mu)^2}{\sigma^2} \frac{V_0^h \left(V_0^h \nu + \mu\right)}{\left(V_0^h \nu + \mu - \mu_0\right)^2} \left(\frac{\nu}{\mu_0 - \mu}\right) + \frac{\nu W_{t-} + \lambda}{\left(V_0^h \nu + \mu - \mu_0\right)}, \tag{50}$$

$$q_t = \frac{\delta_t + W_{t-}}{V_0^h} = -\frac{b - aW_{t-}}{1-a} \frac{(\mu_0 - \mu)^2}{\sigma^2} \frac{V_0^h \left(V_0^h \nu + \mu\right)}{\left(V_0^h \nu + \mu - \mu_0\right)^2} \left(\frac{1}{V_0^h}\right) + \frac{\nu W_{t-} + \lambda}{\left(V_0^h \nu + \mu - \mu_0\right)}. \tag{51}$$

To obtain the representation for q_t, (22), we first define

$$\phi_q \equiv \frac{2X}{Y} \frac{a}{1-a}, \tag{52}$$

where X,Y, and a are given in (46),(47), and (45), respectively. Using simple algebra, we then rewrite

$$\begin{aligned}
q_t &= -\left(\bar{q}V_1^h - W_{t-}\right) \frac{a}{1-a} \frac{2X}{Y} \frac{1}{V_1^h} \frac{\lambda}{\nu\bar{q}V_1^h + \lambda} + \frac{\nu\bar{q}W_{t-}}{\nu\bar{q}V_1^h + \lambda} + \frac{\bar{q}\lambda}{\nu\bar{q}V_1^h + \lambda} \\
&= \frac{1}{\nu\bar{q}V_1^h + \lambda} \left(\lambda\bar{q}(1 - \phi_q) + W_{t-} \frac{\phi_q\lambda + \phi_q\nu\bar{q}V_1^h}{V_1^h} + W_{t-} \left(-\phi_q\nu\bar{q} + \nu\bar{q}\right)\right) \\
&= \frac{1}{\nu\bar{q}V_1^h + \lambda} \left(\lambda\bar{q}(1 - \phi_q) + \nu\bar{q}W(1 - \phi_q) + \frac{W\phi_q}{V_1^h} \left(\nu\bar{q}V_1^h + \lambda\right)\right),
\end{aligned}$$

which leads to our desired representation (22). We obtain the first equality from the representation of q_t above, (51), the second by plugging the definition of ϕ_q, and the third by adding and subtracting $\nu\bar{q}W$. Similarly, we obtain β_t. Let us define

$$\phi_\beta \equiv \frac{2X}{Y} \frac{a}{1-a} \frac{\nu V_0^h}{\mu_0 - \mu}. \tag{53}$$

We then rewrite β_t as follows:

$$\begin{aligned}
\beta_t &= -\left(\bar{q}V_1^h - W_{t-}\right) \frac{a}{1-a} \frac{2X}{Y} \frac{V_0^h}{V_1^h} \frac{\lambda}{\nu\bar{q}V_1^h + \lambda} \frac{\nu}{\mu_0 - \mu} + \frac{\nu\bar{q}W_{t-}}{\nu\bar{q}V_1^h + \lambda} + \frac{\bar{q}\lambda}{\nu\bar{q}V_1^h + \lambda} \\
&= \frac{1}{\nu\bar{q}V_1^h + \lambda} \left(\lambda\bar{q}(1 - \phi_\beta) + W_{t-} \frac{\phi_\beta\lambda + \phi_\beta\nu\bar{q}V_1^h}{V_1^h} + W_{t-} \left(-\phi_\beta\bar{q}\nu + \bar{q}\nu\right)\right) \\
&= \frac{1}{\nu\bar{q}V_1^h + \lambda} \left(\lambda\bar{q}(1 - \phi_\beta) + \nu\bar{q}W_{t-}(1 - \phi_\beta) + \frac{W_{t-}\phi_\beta}{V_1^h} \left(\nu\bar{q}V_1^h + \lambda\right)\right).
\end{aligned}$$

We obtain the representation for δ_t by rewriting it in terms of q_t.

We now turn to the second part of the proof and show that our conjectures are satisfied. First, the second derivative (40) immediately verifies strict concavity as long as $W_{t-} < \bar{q}V_1^h$:

$$V''(W_{t-}) = -(1-a)\,\tilde{k}_1\,(b - aW_{t-})^{\frac{1}{a}-2} < 0$$

because $k_1 > 0$ and $a < 1$. Second, by comparing β_t and q_t above, we find that $\beta_t < q_t$ if and only if

$$\frac{b-aW}{1-a}\frac{(\mu_0-\mu)^2}{\sigma^2}\frac{V_0^h\left(V_0^h\nu+\mu\right)}{\left(V_0^h\nu+\mu-\mu_0\right)^2}\left(\frac{1}{V_0^h}-\frac{\nu}{\mu_0-\mu}\right) < 0,$$

which holds for $W_{t-} < \bar{q}V_1^h$ because the left-hand side is strictly negative. To see this, observe that $V_0^h\nu + \mu = (r+\nu)V_1^h > (r)V_1^h > \mu_0$. Therefore, the incentive compatibility condition always binds.

Lastly, we verify our conjecture that, indeed, $\underline{q} \leq q_t \leq \bar{q}$. The lowest wealth that satisfies the entrepreneur's future equity constraint ($q_t = \underline{q}$), given in (51), is

$$\underline{W} = \frac{\underline{q}\left(V_0^h\nu+\mu\right) + \bar{q}V_1^h\frac{a}{1-a}\frac{(\mu_0-\mu)^2}{\sigma^2}\frac{\left(V_0^h\nu+\mu\right)}{\left(V_0^h\nu+\mu-\mu_0\right)} - \lambda}{\nu + \frac{a}{1-a}\frac{(\mu_0-\mu)^2}{\sigma^2}\frac{\left(V_0^h\nu+\mu\right)}{\left(V_0^h\nu+\mu-\mu_0\right)}} \leq RL, \tag{54}$$

which is below the outside option. Similarly, the highest wealth that satisfies the entrepreneur's future equity constraint ($q_t = \bar{q}$) is

$$\bar{W} = \frac{\bar{q}V_1^h\frac{a}{1-a}\frac{(\mu_0-\mu)^2}{\sigma^2}\frac{\left(V_0^h\nu+\mu\right)}{\left(V_0^h\nu+\mu-\mu_0\right)} + \bar{q}V_1^h\nu}{\nu + \frac{a}{1-a}\frac{(\mu_0-\mu)^2}{\sigma^2}\frac{\left(V_0^h\nu+\mu\right)}{\left(V_0^h\nu+\mu-\mu_0\right)}} = \bar{q}V_1^h. \tag{55}$$

It is easy to verify from (51) that $\frac{\partial q}{\partial W} > 0$. Both (54) and (55) imply that our future equity conjecture (9) is satisfied over the domain $[RL, \bar{q}V_1^h]$. To conclude, let us prove that V_1 represents investors' value function under the contract outlined in this proposition. The proof proceeds in a standard way. Let

$$G_t \equiv \int_0^t e^{-rs}\,(1-q_s)\,dX_s^\pi + e^{-rt}V_1\,(W_t)\,, \tag{56}$$

for time $t \leq T \wedge n$. Under an arbitrary incentive-compatible contract, the entrepreneur wealth (W_t) evolves according to (14). By applying Itô's Lemma, we find

$$\begin{aligned}
&e^{rt}dG_t = \\
&\Big((1-q_t)\,\mu + V_1'\,(W_{t-})\left(rW_{t-} - q_t\mu - \delta_t\nu\right) + \frac{1}{2}\sigma^2 V_1''\,(W_{t-})\,(\beta_t - q_t)^2 + \nu(1-q_t)V_0^h - (r+\nu)\,V_1\,(W_{t-})\Big)\,dt \\
&+ \Big((1-q_t) + V_1'\,(W_{t-})\,(\beta_t - q_t)\Big)\sigma dZ_t + \Big((1-q_t)V_0^h - V_1(W_{t-})\Big)\,(dN_t - \nu dt)\,.
\end{aligned} \tag{57}$$

The process G_t is a supermartingale because the dt term is non-positive (16). It becomes a martingale if and only if we set δ_t, β_t, and q_t according to (20),(21), and (22), respectively. Taking the expectation of (56) and utilizing the supermartingale property, we find

$$V_1(W_0) = G_0 \geq E[G_{T\wedge n}] = E\left[\int_0^{T\wedge n} e^{-rs}(1-q_s)\,dX_s^\pi + e^{-rT\wedge n}V_1(W_{T\wedge n})\right]$$
$$= E\left[\int_0^{T\wedge n} e^{-rs}(1-q_s)\,dX_s^\pi + e^{-rT\wedge n}(1-R)L\right] + E\left[e^{-rT\wedge n}(V_1(W_{T\wedge n}) - (1-R)L)\right]. \quad (58)$$

Taking the limit $n \to \infty$, and using the fact that $V_1(W_T) \leq V_0^h$, the bounded convergence theorem implies that

$$V_1(W_0) \geq E\left[\int_0^T e^{-rs}(1-q_s)\,dX_s^\pi + e^{-rT}(1-R)L\right]. \quad (59)$$

Under the optimal contract, G_t becomes a martingale until time T because both the dZ_t and $dN_t - \nu dt$ terms in (57) are bounded.

□

Proof of Proposition 4 (Pay-Performance Sensitivity). This result can be easily verified by subtracting q_t from β_t, (51) and (50), respectively. □

Proof of Proposition 5 (Risk-Exposure). Risk-taking implies that $\frac{\partial(q_t-\beta_t)}{\partial W} < 0$. This result can be easily verified by subtracting β_t from q_t, (50) and (51), respectively. □

Proof of Proposition 3 (Future Equity). The statement investors' future equity decreases with the entrepreneur value implies that $\frac{\partial(1-q_t)}{\partial W_{t^-}} < 0$. This result can be easily verified from (51). □

Proof of Proposition 6 (Risk Aversion). Absolute risk aversion is defined as

$$RA(W) \equiv -\frac{V_1^{''}(W)}{V_1^{'}(W)}, \quad (60)$$

where the number of primes signifies the number of derivatives taken. Thus, $RA'(W) < 0$ if and only if

$$\frac{1}{(V_1^{'})^2}\left(-V_1^{'''}V_1^{'} + V_1^{''}V_1^{''}\right) < 0.$$

By plugging the derivatives of V_1, (40), this inequality becomes

$$\frac{1}{(V_1^{'})^2}k_1\frac{1}{a}\left(\frac{1}{a}-1\right)\left(\bar{q}V_1^h - W\right)^{\frac{1}{a}-3}\left[\left(\frac{1}{a}-2\right) + k_1\frac{1}{a}\left(\bar{q}V_1^h - W\right)^{\frac{1}{a}-1}\right] < 0.$$

The function multiplying the square brackets is positive and well defined for $W < \bar{q}V_1^h$ and $W \neq \arg\max(V_1(W))$. Thus, plugging k_1, risk aversion increases if and only if

$$\left(\frac{V_1^h - L}{\bar{q}V_1^h - RL}\right)\left(\frac{\bar{q}V_1^h - W}{\bar{q}V_1^h - RL}\right)^{\frac{1}{a}-1} \geq 2a - 1. \quad (61)$$

First, observe this inequality is always satisfied when $a \leq 0.5$, so $RA(W)$ always increases. Second, for any $a > 0.5$, we apply the fixed point theorem. The right-hand side is strictly positive and below one. In contrast, the left-hand side starts strictly above 1, decreases with W, and becomes zero as W approaches $\bar{q}V_1^h$. It starts above one because it is inefficient to terminate the contract, $(1-\bar{q})V_1^h > (1-R)L$. There exists a fixed point $\tilde{W}$ such that (61) holds with equality, and when $W \geq \tilde{W}$, risk aversion decreases, (61) is not satisfied. In contrast, when $W \leq \tilde{W}$, risk aversion increases, (61) is satisfied.

Observe that $(1-q_t)$ is inversely and linearly related to W_{t^-}, (51). Thus, we can write $W(1-q_t)$ as a function of $1-q_t$ and derive $RA(1-q_t)$. Lastly, we find $1-\tilde{Q}$ by plugging $\tilde{W}$ in (51),

$$1-\tilde{Q} = 1 + \frac{b-a\tilde{W}}{1-a}\frac{(\mu_0-\mu)^2}{\sigma^2}\frac{V_0^h\left(V_0^h\nu+\mu\right)}{\left(V_0^h\nu+\mu-\mu_0\right)^2}\left(\frac{1}{V_0^h}\right) - \frac{\nu\tilde{W}+\lambda}{\left(V_0^h\nu+\mu-\mu_0\right)}. \tag{62}$$

□

Proof of Proposition 7 (Optimality). There are no agency frictions during the post-development phase as both the entrepreneur and investors prefer the innovative strategy. Formally, when innovation becomes fully developed, $t=\tau$, investors' expected value equals their share of total value $(1-q_\tau)V_0^h$. If the entrepreneur utilizes off-the-shelf instead, investors' expected value is $(1-q_\tau)\frac{\mu_0}{r}$. Since $V_0^h > \frac{\mu_0}{r}$, (3), the innovative strategy is optimal post-development.

During the development phase, $t<\tau$, we follow the technique of DeMarzo and Sannikov (2006) and show the optimality of developing innovation. If the entrepreneur utilizes off-the-shelf, her wealth is not sensitive to innovation ($\pi\nu = 0$), and due to concavity of the value function, investors set $\beta_t = q_t$. To maximize their share, investors set $q_t = 0$ and the entrepreneur's wealth evolves according to

$$dW_t = rW_{t^-}dt - \lambda dt. \tag{63}$$

The new strategy leads to a lower value than the original contract if $rV_1(W) \geq \mu_0 + V_1'(W)(rW-\lambda)$ for all W, which holds if

$$\frac{\mu_0}{r} \leq \min_W \left\{V_1(W) + V_1'(W)\left(\frac{\lambda}{r} - W\right)\right\}. \tag{64}$$

Plugging V_1 and its derivative, given in Theorem 1, we find that the right-hand side attains the global minimum at $\frac{\lambda}{r}$. Thus, (64) is satisfied if and only if

$$\frac{\mu_0}{r} \leq V_1\left(\frac{\lambda}{r}\right).$$

By plugging V_1 and using (7), the inequality becomes

$$(1-\bar{q})\left(V_1^h - \frac{\mu_0}{r}\right) \geq \left(V_1^h - L\right)\left(\frac{\bar{q}\frac{\mu_0}{r}}{\bar{q}V_1^h - RL}\right)^{\frac{1}{a}}. \tag{65}$$

It is easy to verify that the left-hand side decreases with $\bar{q}$, starts strictly positive, and ends at 0. The right-hand side decreases with $\bar{q}$ when $RL > 0$, and strictly bigger than 0 because the denominator is strictly positive, $\bar{q}V_1^h > RL$. Thus, when $\bar{q} = 1$, the left-hand side is strictly lower than the right-hand side, implying that it must satisfy $Q < 1$ if a fixed point exists. We guarantee the existence of the fixed point by assuming

$$\begin{aligned} &V_1^h \geq 2\tfrac{\mu_0}{r}, && && \text{when } L = 0, \\ &V_1^h \geq \tfrac{2-R}{1-R}\tfrac{\mu_0}{r}, && \tfrac{\mu_0}{r} > L && \text{when } L > 0. \end{aligned} \tag{66}$$

For $L = 0$, we find

$$V_1^h - \frac{\mu_0}{r} \geq \frac{\mu_0}{r} \geq \left(V_1^h\right)\left(\frac{\frac{\mu_0}{r}}{V_1^h}\right)^{\frac{1}{a}},$$

where the left inequality holds due to (66), and the right due to $\frac{\mu_0}{r} < V_1^h$ and $0 < a < 1$. For $L > 0$, we look at $\bar{q} = R$ and find

$$(1-R)\left(V_1^h - \frac{\mu_0}{r}\right) \geq \frac{\mu_0}{r} = \left(V_1^h - L\right)\left(\frac{\frac{\mu_0}{r}}{V_1^h - L}\right) > \left(V_1^h - L\right)\left(\frac{\frac{\mu_0}{r}}{V_1^h - L}\right)^{\frac{1}{a}} = \left(V_1^h - L\right)\left(\frac{\bar{q}\frac{\mu_0}{r}}{\bar{q}V_1^h - RL}\right)^{\frac{1}{a}},$$

where the left and right inequalities hold because of (66). There exists a fixed point $0 < Q < 1$, which solves (65) with equality. □

Proof of Proposition 8 (Compensation). The salary is always positive because

$$(r+\nu)V_1^h > (r)V_1^h > \mu_0 > \mu, \qquad q_t \geq 0.$$

The elasticity (32) is given by

$$\frac{\partial q}{\partial \beta} = \frac{\partial q}{\partial W}\frac{\partial W}{\partial \beta} = \frac{\frac{a}{1-a}\frac{(\mu_0-\mu)^2}{\sigma^2}\frac{V_0^h\left(V_0^h\nu+\mu\right)}{\left(V_0^h\nu+\mu-\mu_0\right)^2}\left(\frac{1}{V_0^h}\right) + \frac{\nu}{\left(V_0^h\nu+\mu-\mu_0\right)}}{\frac{a}{1-a}\frac{(\mu_0-\mu)^2}{\sigma^2}\frac{V_0^h\left(V_0^h\nu+\mu\right)}{\left(V_0^h\nu+\mu-\mu_0\right)^2}\left(\frac{\nu}{\mu_0-\mu}\right) + \frac{\nu}{\left(V_0^h\nu+\mu-\mu_0\right)}} > 0,$$

which is strictly positive because $\frac{\partial q}{\partial W}$ and $\frac{\partial \beta}{\partial W}$ are strictly positive. Plugging ϕ_q leads to our characterization (32). Suppose $K_t = W_t$, the bank account balance is higher than the entrepreneur wealth after positive shock ($K_t + dK_t \geq W_t + dW_t$), while it is lower than the entrepreneur wealth after negative shock ($K_t + dK_t \leq W_t + dW_t$). Negative shock decreases the bank account if and only if $\beta_t > 0$, and positive shock if and only if $\beta_t < 0$. The biggest negative shock that leads the entrepreneur wealth to the upper bound $\bar{q}V_1^h$, and the biggest positive shock that leads the entrepreneur wealth to the lower bound RL are bounded. Thus, the biggest drop in the bank account must also be bounded.

□

Proof of Proposition 9 (Implied Probability of Success). This result is a direct consequence of the law of iterated expectation as follows:

$$
\begin{aligned}
V_1(W) + W &= \mathbb{E}^{\pi}\left[\int_0^T e^{-rs}\Big\{dX_s^{\pi}\Big\} + e^{-rT}L\right] \\
&= \mathbb{E}^{\pi}\left[\mathbb{E}^{\pi}\left[\int_0^T e^{-rs}\Big\{dX_s^{\pi}\Big\} + e^{-rT}L \mid \mathbf{1}_{\tau\geq T}\right]\mathbf{1}_{\tau\geq T}\right] \\
&+ \mathbb{E}^{\pi}\left[\mathbb{E}^{\pi}\left[\int_0^T e^{-rs}\Big\{dX_s^{\pi}\Big\} + e^{-rT}L \mid \mathbf{1}_{\tau<T}\right]\mathbf{1}_{\tau<T}\right].
\end{aligned} \tag{67}
$$

The time τ is exponentially distributed over $(0,\infty)$ and is unpredictable when investors choose T (T is (X^{π}, N)-measurable). In (67), in the first conditional expectation, to guarantee that $\mathbf{1}_{\tau\geq T}$ holds almost surely, T must be set equal to zero. By contradiction, for a given $T > 0$, there is a positive measure set such that $\mathbf{1}_{\tau\geq T}$ is violated. Similarly, in the second conditional expectation, to guarantee that $\mathbf{1}_{\tau\leq T}$ holds almost surely, T must be set equal to ∞. By contradiction, for a given $T < \infty$, there is a positive measure set such that $\mathbf{1}_{\tau<T}$ is violated. Plugging these two results back into (67), we obtain

$$
V_1(W) + W = \mathbb{E}^{\pi}\left[L\mathbf{1}_{\tau>T}\right] + \mathbb{E}^{\pi}\left[V_1^h \mathbf{1}_{\tau\leq T}\right] = L\Pr\left[\tau > T\right] + V_1^h \Pr\left[\tau \leq T\right], \tag{68}
$$

and by plugging V_1 from Theorem 1, we obtain our desired result (33).

We verify that $\Pr[\tau \leq T]$ is between zero and one. The lower bound on W, $W \geq RL$, immediately implies that $0 \leq \Pr[\tau \leq T]$. The upper bound on W, $W \leq \bar{q}V_1^h$, implies that $1 \geq \Pr[\tau \leq T]$, and because $\frac{\partial \Pr[\tau\leq T]}{\partial W_0} > 0$, these are the minimum and maximum values. Lastly, $\Pr[\tau \leq T]$ admits a one-to-one map because it is a continuous and increasing function of W. Verifying the derivatives in (34) is straightforward.

□

References

Admati, A. R., and P. Pfleiderer. 1994. Robust Financial Contracting and the Role of Venture Capitalists. *The Journal of Finance* 49:371–402.

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

Bergemann, D., and U. Hege. 2005. The Financing of Innovation: Learning and Stopping. *The RAND Journal of Economics* 36:719–752.

Biais, B., T. Mariotti, G. Plantin, and J.-C. Rochet. 2007. Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications. *The Review of Economic Studies* 74:345–390.

Biais, B., T. Mariotti, J. Rochet, and S. Villeneuve. 2010. Large Risks, Limited Liability, and Dynamic Moral Hazard. *Econometrica* 78:73–118.

Bitler, M. P., T. J. Moskowitz, and A. Vissing-Jørgensen. 2005. Testing Agency Theory with Entrepreneur Effort and Wealth. *The Journal of Finance* 60:539–576.

Brown, K. C., W. V. Harlow, and L. T. Starks. 1996. Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry. *The Journal of Finance* 51:85–110.

Burton, M. D., J. B. Sørensen, and S. D. Dobrev. 2016. A Careers Perspective on Entrepreneurship. *Entrepreneurship Theory and Practice* 40:237–247.

Casamatta, C. 2003. Financing and Advising: Optimal Financial Contracts with Venture Capitalists. *The Journal of Finance* 58:2059–2085.

Cestone, G. 2014. Venture Capital Meets Contract Theory: Risky Claims or Formal Control? *Review of Finance* 18:1097–1137.

Cochrane, J. H. 2005. The risk and return of venture capital. *Journal of Financial Economics* 75:3 – 52.

Cornelli, F., and O. Yosha. 2003. Stage Financing and the Role of Convertible Securities. *The Review of Economic Studies* 70:1–32.

Coyle, J. F., and J. Green. 2018. The Safe, the KISS, and the Note: A Survey of Startup Seed Financing Contracts. *Minn. L. Rev. Headnotes* 103.

DeMarzo, P., and Y. Sannikov. 2006. Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model. *The Journal of Finance* 61:2681–2724.

DeMarzo, P. M., and M. J. Fishman. 2007. Optimal Long-Term Financial Contracting. *The Review of Financial Studies* 20:2079–2128.

Dessí, R. 2005. Start-up Finance, Monitoring, and Collusion. *The RAND Journal of Economics* 36:255–274.

Ewens, M., A. S. Gorbenko, and A. Korteweg. 2019. Venture Capital Contracts. *NBER Working Paper No. 26115* .

Gompers, P., A. Kovner, J. Lerner, and D. Scharfstein. 2010. Performance Persistence in Entrepreneurship. *Journal of Financial Economics* 96:18–32.

Graham, J. R., C. R. Harvey, and S. Rajgopal. 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40:3–73.

Habib, M. A., and D. B. Johnsen. 2000. The Private Placement of Debt and Outside Equity as an Information Revelation Mechanism. *The Review of Financial Studies* 13:1017–1055.

Halac, M., N. Kartik, and Q. Liu. 2016. Optimal Contracts for Experimentation. *The Review of Economic Studies* 83:1040–1091.

Hall, R. E., and S. E. Woodward. 2010. The Burden of the Nondiversifiable Risk of Entrepreneurship. *American Economic Review* 100:1163–1194.

He, Z. 2009. Optimal Executive Compensation when Firm Size Follows Geometric Brownian Motion. *The Review of Financial Studies* 22:859–892.

Hellmann, T. 2006. IPOs, acquisitions, and the use of convertible securities in venture capital. *Journal of Financial Economics* 81:649–679.

Hellmann, T. 2007. When Do Employees Become Entrepreneurs? *Management Science* 53:919–933.

Hellmann, T., and V. Thiele. 2011. Incentives and Innovation: A Multitasking Approach. *American Economic Journal: Microeconomics* 3:78–128.

Hoffmann, F., and S. Pfeil. 2010. Reward for Luck in a Dynamic Agency Model. *The Review of Financial Studies* 23:3329–3345.

Hoffmann, F., and S. Pfeil. 2018. Dynamic Multitasking and Managerial Investment Incentives. *Working Paper* .

Holmström, B., and P. Milgrom. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization* p. 24.

Kaplan, S. N., and A. Schoar. 2005. Private Equity Performance: Returns, Persistence, and Capital Flows. *The Journal of Finance* 60:1791–1823.

Manso, G. 2011. Motivating Innovation. *The Journal of Finance* 66:1823–1860.

Marx, L. M. 1998. Efficient venture capital financing combining debt and equity. *Review of Economic Design* 3:371–381.

Merton, R. C. 1974. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *The Journal of Finance* 29:449–470.

Repullo, R., and J. Suarez. 2004. Venture Capital Finance: A Security Design Approach. *Review of Finance* 8:75–108.

Sannikov, Y. 2008. A Continuous-Time Version of the Principal-Agent Problem. *The Review of Economic Studies* 75:957–984.

Schmidt, K. M. 2003. Convertible Securities and Venture Capital Finance. *The Journal of Finance* 58:1139–1166.

Varas, F. 2018. Managerial Short-Termism, Turnover Policy, and the Dynamics of Incentives. *The Review of Financial Studies* 31:3409–3451.

Zhu, J. Y. 2018. Myopic Agency. *The Review of Economic Studies* 85:1352–1388.