Financing of Innovation under Information Asymmetry

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Xuelin Li

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Andrew Winton and
Martin Szydlowski, Advisers

June 2020

© Xuelin Li 2020

Dedication

To Meizi and my parents.

Abstract

My thesis studies the financing of innovations under information asymmetry. The growth of innovative companies requires sufficient fundings before a successful breakthrough. However, their assets are in general intangible and hard to evaluate by investors. Therefore, firms have private information about their technology and those of better qualities need to generate signals to potential funders. My thesis illustrates two of them: acquisition and contracts.

In Chapter 1, I study acquisitions between startups. I document that over one third of initial public offerings (IPOs) have acquired before filing. What determines whether startups can finish such transactions? How is IPO performance related to acquisition deals? To answer these questions, I develop a continuous-time real options model of two heterogeneous startups with information asymmetry as the key frictions. I first prove that private acquisitions are positive signals for startup qualities and will reduce IPO underpricing. However, signaling opportunity is not always available. Investors' overoptimism will lead to inefficient IPO waves that inhibit signaling acquisition transactions. I confirm the model predictions by empirically finding that IPOs with private acquisitions have both significantly less underpricing and better long-term operational performance.

In Chapter 2, my adviser Martin Szydlowski and I study venture capital finance with experimentation. An entrepreneur contracts with an investor and has private information about a project, which requires costly experimentation by both parties to succeed. In equilibrium, investors learn about the project from the arrival of exogenous information and from the entrepreneur's contract offers. The optimal contract features vesting and dilution, consistent with empirical evidence. Early payouts, pivots, and prestige projects emerge as signaling devices. Surprisingly, technological progress, which lowers the cost of experimentation or which increases the rate of learning, delays separation of types and worsens adverse selection. Liquidation rights for investors also delay separation.

Contents

List of Tables

List of Figures

Chapter 1

A Signaling Story of Private Acquisitions

1.1 Introduction

This paper develops a theoretical framework to study mergers and acquisitions between startups. Recently there is an emerging interest in the corporate decisions of private firms as abnormally few companies are publicly listed (Doidge et al., 2017). Existing literature has examined innovation activities (Budish et al., 2015; Bernstein, 2015), cash policies (Gao et al., 2013) and general investment decisions (Asker et al., 2011). However, one important aspect is missing, which is the merger and acquisition (M&A) transaction. Empirically, I document that over one third of initial public offering (IPO) firms have acquired before filing and find the following novel stylized facts. IPOs with private acquisitions have both significantly less underpricing and better long-term operational performance.

The empirical evidence leads to the following two questions. What determines whether startups could successfully acquire before going public? How are IPO underpricing and performance related to acquisition activities? The central contribution of this paper is to answer these questions theoretically. I first prove that private acquisitions send positive signals about startup qualities, which will reduce IPO underpricing. However, signaling opportunity is not always available. Sometimes investors can be over-optimistic about bad startups in the economy. This will make them reject takeover offers from more productive firms, which leads to inefficient IPO waves and inhibits signaling acquisition transactions.

The model is a real options model of two strategic players with information asymmetry as the key friction. Two startups seek IPO financing to cover the cost of investments. Each has some physical assets and an opportunity to invest on them. Investment net return is linear in the size of the assets, which stochastically changes before investments. Therefore, this is a real options framework, where firms choose the best investment timing. Startups differ in their managerial qualities. The high-quality firm's assets tend to grow faster before investments, and its marginal rate of investment return is larger. Before going public, both firms can acquire the other's physical assets. But managerial ability is not transferred through acquisition. The critical friction is that IPO investors initially cannot distinguish firm qualities. However, they observe firm sizes and whether acquisitions occurred. Firms optimally choose acquisition and stand-alone IPO timings for investment to maximize profits.

If the investors are perfectly informed about firm qualities, the low-quality firm will never be overpriced in a stand-alone IPO. In this case, the equilibrium is Pareto-optimal. The high-quality firm acquires the low-quality one since the former could utilize the assets with more productive investment technologies. Then firms could bargain over the positive synergy. As a result, each firm is better off with more profits than investing individually. This first-best benchmark replicates the neoclassical M&A literature (Jovanovic and Rousseau, 2002), which argues acquisitions represent efficient reallocation of assets. However, if investors cannot distinguish firm qualities, they will form a prior belief based on observed asset sizes. This belief is noisy, and investors can be too optimistic about the truly bad startup. Thus, a bad firm can be hugely overpriced when going public. This overpricing opportunity becomes an outside option for the low type to reject takeover offers, which inhibits acquisition transactions.

In equilibrium, private acquisitions occur when investors' belief is sufficiently accurate. In such transactions, assets flow from the less productive type to the startup with better technologies, but not vice versa. In other words, conditional on acquisitions happening, they still represent efficient reallocations. Since only the high-quality startup will acquire in equilibrium, private acquisitions are a self-selection mechanism that guarantees only high-quality firms go public afterwards. As a result, undertaking acquisitions before IPO reduces quality uncertainties and generates a positive signal.

However, the occurrence of such acquisitions is significantly delayed and even sometimes replaced by inefficient IPO waves. When investors have an extremely optimistic prior belief about the low-quality firm, the bad startup resists takeover offers and force the more productive firm to give up acquiring. In the end, both startups go public at the same time, which results in pooling IPOs with mixed firm qualities. Therefore, signaling opportunity is not always available. Lastly, if the investors' belief is intermediate, both startups will delay investments and wait.

The above results hinge on how investors form their prior beliefs about firm qualities. At any moment, they observe the two startups' asset sizes. Since the high type grows faster in expectation, they will rationally believe that the realized larger firm is more likely to have high productivities. However, assets growths are noisy processes. There is a chance that the low-quality firm has a sequence of positive shocks and is sufficiently larger. Given this, investors are overly optimistic about the low type. Then if both firms go public at the same time, investors will price firms' IPO shares with their mistaken beliefs. As a result, in pooling IPOs, the high-quality firm "loses" with an underpricing cost. Meanwhile, the low-quality firm "wins" by being overpriced.

On the contrary, acquisitions are too costly for the low-quality firm to imitate.

This is because managerial ability cannot be transferred in M&A.[1] A low-quality acquirer generates substantially fewer profits after mergers, and a high-quality target will charge a hefty price for being acquired. Therefore, when acquisitions do occur, investors rationally believe that the acquiring startup is more productive with certainty. So the high-quality firm "wins" with being priced correctly, but the low-quality firm "loses" without overpriced stock issuance.

Firms are making trade-offs between accepting the losing payoff right away and waiting for winning in the future. When investors are optimistic about the low-quality firm, the more productive type will give up acquisitions and go public, anticipating the underpricing cost. The reason is that a sufficiently mistaken belief takes a prolonged amount of time to revert. While waiting, the high-quality firm bears the cost of delaying growth opportunities. This long-term cost ultimately outweighs the short-term underpricing loss. On the contrary, when investors almost correctly distinguish firm qualities through assets, the low-quality firm loses its possibility of deceiving investors. The potential overpricing opportunity is too low, so that bad firm is willing to accept takeover offers. Lastly, when the difference between asset sizes is moderate, no firm is willing to give up. So the game falls into a region of delaying.

This equilibrium structure sheds light on the stylized facts. First, IPOs preceded by acquisitions will have significantly less underpricing because, given the positive signal, investors are willing to offer a higher price. Second, IPOs preceded by acquisitions will have significantly better long-term operational performance after going public because, in non-acquiring IPOs, the existence of bad startups will decline the average performance of new public firms.

Lastly, I empirically test these hypotheses by merging stock issuance data with

[1] Interpreting productivity as managerial and organizational talent can be dated at least back to Lucas (1978). It is commonly assumed as non-transferable in the neoclassical view of M&As (Jovanovic and Rousseau, 2002; Maksimovic and Phillips, 2002). However, the opposite assumption of technology complementarity as in Rhodes-Kropf and Robinson (2008) can also be accommodated and discussed in the model section.

private M&A data. In a sample of more than 1,500 IPOs, the supporting evidence is both statistically and economically significant. For underpricing, a typical IPO with private acquisitions has a 3.315% lower first-day return compared to a non-acquiring IPO. For long-term performance, a typical IPO with private acquisitions a significant 5.5% larger return of assets three years after going public. With these two findings, the signaling effect is validated.[2]

Related Literature

This paper adds to the literature on M&As as follows. First, I offer a theoretical framework that generates empirical predictions closely related to the patterns of private acquisitions (Maksimovic et al., 2013; Netter et al., 2011). The model result implies that private acquisitions still represent the efficient reallocation of assets, which is in line with models under perfect information (Jovanovic and Rousseau, 2002; Maksimovic and Phillips, 2002).[3] However, the results do not merely echo previous findings. The existence of information imperfection significantly delays and prevents assets transactions, which is the economic rationale that drives the distinct patterns of private deals.

Second, this paper adds to the literature on the role of growth opportunities or generally intangible assets in M&A. Closely related is Lambrecht (2004), which analyzes how the profit-sharing terms impact the timing of mergers in a two-player real options model. In his model, mergers can generate economies of scale, but firms have no differences in their qualities. Therefore, Lambrecht (2004) does not allow the possibility of signaling. Instead, his model focuses on endogenously solving a profit-

[2]Rational investors will not reward private acquisitions if they merely reduce competitions. In fact, IPOs that make acquisitions simultaneously or soon afterward significantly underperform in the long run (Brau et al., 2012; Brown et al., 2005; Ritter, 2015). This distinguishes the quality of a private acquirer from that of a new public acquirer.

[3]The empirical evidence in support of this view include, but are not limited to, plant-level productivity (Li, 2013; Maksimovic and Phillips, 2001; Maksimovic et al., 2011), product quality (Sheen, 2014) and investment expenditure (Devos et al., 2008).

sharing rule to rationalize when mergers occur. On the contrary, my model focuses on how investors' learning process affects the occurrence of M&As. Conditional on an acquisition, profit division method is exogenously assumed. Therefore, our papers complement each other. Alternatively, Levine (2017) models growth opportunities as "seeds" to constrain capital investment.[4]

This paper generates a new perspective to the extensive literature on M&A motives, including but not limited to managerial hubris and empire building (Jensen, 1986; Roll, 1986), stock misvaluation (Shleifer and Vishny, 2003), market power (Kim and Singal, 1993) and complementarity (Rhodes-Kropf and Robinson, 2008). I show that private acquisitions generate valid signals in IPOs, and startups are motivated to resolve the adverse selection problem through acquiring. I also add to the literature of the relationships between IPOs and M&As. Going public and being acquired are typically assumed to be substitutes as exit choices for startups. For example, in Bayar and Chemmanur (2012), these two choices have different costs and benefits. In this paper, I do not assume they are mutually exclusive. As an equilibrium result, the high-quality firm will sometimes both acquire and go public. The implications are also consistent with the empirical evidence on the timing when IPO exits are relatively more frequent (Ball et al., 2011), and firm-level characteristics that predict higher chances of IPO compared to being acquired (Brau et al., 2003; Poulsen and Stegemoller, 2008).[5]

Theoretically, this paper belongs to the literature of real options models with information imperfection. As in Grenadier and Malenko (2011), the option's payoff has two parts: a direct project payoff and an indirect belief component depending on a third party's assessment of types. Gorbenko and Malenko (2017) consider a model

[4]Empirically acquisitions of intangible assets, especially in innovative industries, are frequent and important to firm growth, e.g., Bena and Li (2014); Cunningham et al. (2018); Higgins and Rodriguez (2006); Krieger et al. (2018).

[5]Some IPOs are motivated by a desire to acquire, for example either in a roll-up IPO (Brown et al., 2005; Ritter, 2015), or soon afterwards (Brau et al., 2012; Celikyurt et al., 2010).

where two acquirers with different marginal costs of using cash, privately knowing their synergy, decide when to approach the target and the method of payment. High-valuation player signals by paying costly cash. This paper differs from Gorbenko and Malenko (2017) in two ways. First, in their model, players have one common state variable, which is orthogonal to the target's belief.[6] In my model, both startups have their assets sizes evolving with different drifts. These two state variables are informative and exogenously move the investors' prior. Second, in Gorbenko and Malenko (2017), there are no direct interactions between players. For example, one bidder cannot pay the other to give up acquiring and signal its high valuation of M&As. Instead, my paper reveals the signaling effect of interactions, i.e., acquisitions.

This paper also relates to the dynamic adverse selection models (Daley and Green, 2012) in the sense that, the investors' belief stochastically moves and serves as the key state variable to influence the decision makers. The literature help explain, for example, liquidity dry-up (Daley and Green, 2016), misallocation (Fuchs et al., 2016), entry decision (Zryumov, 2015), market freezes (Fishman and Parker, 2015) and the recovery of them (Chiu and Koeppl, 2016). In Strebulaev et al. (2016), cash flow plays a dual role in signaling and loosening financial constraint, which both benefit the high type. In my model, assets sizes also send signals and impact the synergy in M&As. The difference is that the dual roles work in the opposite direction for the high type. When the low type's asset is larger, it generates a higher synergy, but investors are more optimistic about the less efficient firm, which makes it reluctant to accept takeover offers.

Though my model generates a two-threshold equilibrium as in Daley and Green (2012), the fact that it has two strategic players changes how separation works. In Daley and Green (2012), there is a partial-separating threshold where the low type drops with some probability to make belief reflect to the boundary. Yet in my model,

[6]This makes the timing of acquisition initiation have no signaling ability.

Figure 1.1: Stage Game Timeline at time t

the high type will initiate acquisition with certainty below this boundary since this is its dominant strategy motivated by the efficiency gain, and low type will accept. Therefore I have full separations due to the possibility of interactions between players. This also distinguishes the paper from Gul and Pesendorfer (2012).

The paper is organized as follows. Section 2 develops the full model. Section 3 analyzes the model and characterizes. Section 4 extends the model in several ways. Section 5 tests and discusses the model's implications empirically. Section 6 concludes.

1.2 Model

Figure 1.1 summarizes the sequence of events at time t. First, assets sizes of the two startups are updated. Then each of them decides whether to takeaway the competitor's assets by acquiring it. If there is a completed merger, game ends and the merged firm invests by going public. Otherwise, players arrive at the IPO stage. They individually decide whether to utilize the investment opportunity and complete stand-alone IPOs. If they do, game ends. Otherwise, game continues to time $t + dt$. New assets sizes are updated and the above procedures repeat.

Assets

The game has two strategic players. They are startups, h (the high type) and l (the low type), operating in time $t \in [0, +\infty)$. Both firms are risk-neutral and have the identical discount rate r. Each firm i has assets under management, whose size x_{it} follows a geometric Brownian motion,

$$\frac{dx_{it}}{x_{it}} = \mu_i dt + \frac{1}{\sqrt{2}}\sigma dB_{it}. \tag{1.1}$$

Firms initially have identical sizes, $x_{h0} = x_{l0} = 1$. B_{ht} and B_{lt} are two independent standard Brownian motions on the canonical probability space $\{\Omega, \mathcal{F}, \mathcal{Q}\}$. μ_i is the expected growth rate. Firm h has a higher expected growth rate of assets. Following Dixit and Pindyck (1994), I assume $r > \mu_h > \mu_l$. This regulates finite solutions and implies that delaying is costly.[7]

Real Options

Each firm has a costly real option to speed up growth. It is interpreted as an investment opportunity whose payoff and cost are both linear in a firm's assets size. The stand-alone NPV is $(H - \alpha)\, x_{ht}$ for firm h and $(L - \alpha)\, x_{lt}$ for firm l. $Z_i \in \{H, L\}$ is the marginal rate of return for firm i. It satisfies (i) $H - \alpha > 0$ and (ii) $H > L$. The first restriction implies that investment NPV for the high type is strictly positive. The second implies that the high type generates a strictly higher return than the low type. I interchangeably call firm h as the more efficient or productive type, motivated by the fact that it has better investment technology and asset growth.

αx_{it} is the cost of investment. Since startups in general lack internal cash and rely heavily on external financing, firms will raise fundings through IPOs to cover the

[7] The assumption captures in reality delaying IPO and positive NPV projects are costly for startups. For VC-backed startups, VC funds have predetermined investment horizon around 10 years (Gompers, 1996; Barrot, 2016). Delaying exit beyond that scope is generally not feasible.

cost. They endogenously choose the timing of IPO to exercise the real options.

I assume that if a single player exercises option at τ, firm types become public information at τ^+. Thus, the game effectively ends after one execution.

Acquisition

Before stand-alone IPOs, each firm i can make an acquisition offer to the other one $-i$. If M&A succeeds, a merged firm, indexed by m, will go IPO and exercise the option.[8] The merged firm expands the combined assets of size $x_{mt} = x_{it} + x_{-it}$, using the acquirer i's technology, which is not transferable in acquisitions. If the high type is the acquirer, the NPV is $(H - \alpha)\, x_{mt}$ and a positive synergy $(H - L)\, x_{lt}$ is generated. If the low type is the acquirer, the NPV is $(L - \alpha)\, x_{mt}$ and an efficiency loss $(H - L)\, x_{ht}$ is generated. Therefore, in an economy with complete information, only the more efficient type will acquire as in the neoclassical M&A literature.

The acquirer transfers part of the expansion profit to the target as acquisition offers. This can be realized in a stock-exchange transaction such that the target receives shares of the merged company.[9] In the baseline model, players have exogenously given reservation value, which is the target's stand-alone NPV plus an exogenous markup. Specifically, there exist parameters $\gamma_h \geq 0$ and $\gamma_l \geq 0$ such that the offer has to be $(H - \alpha + \gamma_h)\, x_{ht}$ for firm h being the target and $(L - \alpha + \gamma_l)\, x_{lt}$ for firm l being the target.[10]

An acquisition deal is successful only if both parties are willing to participate, and

[8] I assume that after mergers, firm m has to exercise at once. This setup is a parsimonious way to highlight the signaling effect of acquisition. Alternatively, firm m can grow its assets x_{mt} following the acquirer's Brownian motion and optimally choose the exercising time. Notice this would further increase the efficiency advantage of firm h as it has higher assets growth rate, which makes signaling cheaper.

[9] This simplifies the acquisition process and avoid alternative signaling concerns through methods of payments. See Hege and Hennessy (2010); Lambrecht (2004) for a similar assumption.

[10] In Section 1.4.2, I endogenize the offer value through Nash bargaining where the threat point is to go public alone while being regarded as the low type.

acquisition is observable to investors. Firms endogenously choose the timing of the acquisition, jointly with the decision of a stand-alone IPO.

Timeline and Strategy

The sequence of events during the infinitesimal time interval $[t, t+dt]$ can be heuristically illustrated as follows:

Step 1: Asset sizes x_{ht} and x_{lt} are updated, observable to both firms.

Step 2: Nature flips a coin. If it is head, firm h moves first in Step 3 and 4. Otherwise, firm l moves first.

Step 3: Acquisition stage: Sequentially firm i decides whether to acquire $-i$. If $-i$ accepts, i expands total assets by raising $\alpha(x_{ht}+x_{lt})$ from investors. Otherwise game continues to IPO stage.

Step 4: IPO stage: Sequentially firm i decides whether to file for IPO. If i goes public without acquisitions, it raises αx_{it} from investors.

Step 5: Game continues if both firms delay.

Step 2 introduces a randomization device that stipulates the moving order of players in Step 3 and 4. At each stage, the second mover can observe the decision of the first one. The order is indistinguishable from the investors' perspective.

Following the real options literature, I consider Markov stopping strategies, which is a 3-tuple $\sigma_{it} = \left(\sigma_{it}^{A}\left(x_{ht}, x_{lt}\right), \sigma_{it}^{T}\left(x_{ht}, x_{lt}\right), \sigma_{it}^{I}\left(x_{ht}, x_{lt}\right)\right)$. $\sigma_{it}^{A} : X_t^h \times X_t^l \to [0,1]$ is the probability of firm i making an acquisition offer at Step 3. $\sigma_{it}^{T} : X_t^h \times X_t^l \to [0,1]$ is the probability of firm i accepting an offer, conditional on it receives one. σ_{ht}^{I} : $X_t^h \times X_t^l \to [0,1]$ is the probability of firm h's stand-alone IPO when it moves first in the IPO stage. $\sigma_{lt}^{I} : X_t^h \times X_t^l \to [0,1]$ is the probability of firm l's stand-alone IPO when it moves secondly.

Important interpretations of the strategies are in order. First, I do not need to consider the strategy when firm l moves first in the IPO stage. This is because then the low type will never go public. If it did, the high type had a dominant strategy to wait an infinitesimal amount of time and let its quality revealed. Investors can then update that a single stand-alone IPO is the low type with probability one. Second, the low type has a dominant strategy to mimic in the IPO stage when moving secondly, since it observes firm h's decision. Therefore I can equivalently focus on σ_{lt}^{I} as a mapping from the size space, rather than make it a function of the realized action of firm h, with the restriction that $\sigma_{ht}^{I}(x_{ht}, x_{lt}) = \sigma_{lt}^{I}(x_{ht}, x_{lt})$. Lastly, I will focus on the equilibria that acquisition offer equals to the reservation value of the target in equilibrium.[11] Hence offer value is not in the strategy space.

Investors

Investors provide the rationale for pricing. They are non-strategic and assumed to be short-term players who earn zero profit in expectation. The types of firms are unknown by the investors, who initially believe both firms have an equal chance to be the high type.[12] When IPO market opens, they observe firm assets sizes and whether acquisition occurred. They know the data generating process in equation (1.1) and form prior beliefs about firm qualities based realized sizes. In addition, when acquisition happens, they update a posterior belief based on equilibrium acquiring strategies of firms. The investors make pricing decisions based on

$$s_{it}\left(HP_{it} + L\left(1 - P_{it}\right)\right) = \alpha. \tag{1.2}$$

s_{it} are the shares issued to investors. $1 - s_{it}$ shares belong to the startup owners. The price of the issuing shares is a function of P_{it}, which is firm i's probability of being a

[11] It does not rule out off-equilibrium deviations with offers higher than the reservation value. A full-fledged model with endogenous offer value can be pinned down by a commonly preferred stopping threshold of both types. See page 50 in Lambrecht (2004).

[12] The investors know there is exactly one high type and one low type, but they do not know "who is who".

high type. $(HP_{it} + L(1 - P_{it}))$ is also the expected marginal return of investment in investors' belief. Equation (1.2) is a zero-profit condition. Given this, a stand-alone IPO has the following payoff for firm i:

$$R_i^I(x_{it}, x_{-it}) = (1 - s_{it}) Z_i x_{it}.$$

In a merged IPO, given that firm i is acquiring, the acquirer and target has the following payoff respectively:

$$R_i^m(x_{it}, x_{-it}) = (1 - s_{mt}) Z_i x_{mt} - (Z_{-i} - \alpha + \gamma_{-i}) x_{-it},$$
$$R_{-i}^m(x_{it}, x_{-it}) = (Z_{-i} - \alpha + \gamma_{-i}) x_{-it}.$$

Equilibrium Concept

Firm i's strategy σ^i maximizes its continuation $V_i(x_{it}, x_{-it})$ given its opponent strategy σ^{-i}. Formally, it solves the following (FP) given σ^{-i}:

$$V_i(x_{it}, x_{-it}) = \sup_{\sigma^i} \mathbb{E}^i \left(\int_t^\tau e^{-r\tau} \left(R_i^I(x_{i\tau}, x_{-i\tau}) \mathbb{1}_\tau^I + R_i^m(x_{i\tau}, x_{-i\tau}) \mathbb{1}_\tau^m \right) | x_{it}, x_{-it} \right). \tag{FP}$$

In (FP), τ is the random stopping time. $\mathbb{1}_{i\tau}^I$ is the indicating function of a stand-alone IPO and $\mathbb{1}_{i\tau}^m$ is the indicating function of mergers. σ^i and σ^{-i} together decide the expected probability of these events and whether the player is an acquirer or a target in acquisitions. A Markov Perfect Bayesian Equilibrium is a pair of (σ^h, σ^l) that solves (FP) for both firms. Investors use Bayes' rule whenever possible. Off-equilibrium belief is restricted by D1 refinement following Cho and Sobel (1990).

Discussion I now discuss a few assumptions embedded in the model setup. First, the strategies are Markovian. I do not consider a strategy in which today's action depends on past initiated but failed acquisitions. This is without loss of generality because investors are short-term players who will not observe historical acquisition attempts. Since past offers cannot be verified, restricting Markovian strategies help me rule out cheap-talk equilibria. In reality, most acquisition initiations are private before the announcements. Even public companies will only file a proxy statement for their investors once terms are successfully negotiated.

Second, technology is non-transferable through acquisition. Though this is commonly assumed in neoclassical models of M&As, my model can also work with the complementarity assumption in Rhodes-Kropf and Robinson (2008). A merged firm can have a fixed return Z_m regardless of the identity of the acquirer. The equilibrium structure will not change because the low type would still reject takeovers given the existence of pooling IPOs. This alternative case is also easier to solve, since the merged firm's marginal rate of return is not random, which automatically "signals", and the belief of investors when seeing M&As is degenerate.

Lastly, I assume a sequential sub-game in each stage at time t. This is because pooling IPOs are not sustainable in equilibrium if the IPO stage has simultaneous moves. When the high type expects the low type to go public with non-zero probability, it will always delay by an infinitesimal time. However, in a discrete-time model, each stage can be assumed as a simultaneous subgame and pooling IPOs are sustainable. The limit of such a model, when the time interval goes to 0, matches the current setup. The sequence is not observable to investors. In reality, it is hard to distinguish the initiation time of IPO, which usually starts with negotiating with investment banks. After the JOBS Act, emerging growth companies (companies with less than $1 billion in annual revenue) can hide their prospectus until 15 days before the roadshow.

1.3 Equilibrium

1.3.1 Belief

Investors form prior beliefs on firms based on realized assets sizes. The idea is that the high-quality firm grows its assets faster in expectation before investment. The investors will rationally believe a realized larger firm is more likely to be the high type. However, since assets growths are noisy processes, there is a chance that sizes are misleading so that the investors are optimistic about the low-quality firm. Firm i's probability of being the high type, based on assets sizes, is

$$P_i\left(x_{it}, x_{-it}\right) = \frac{f_t^h(x_{it})f_t^l(x_{-it})}{f_t^h(x_{it})f_t^l(x_{-it}) + f_t^h(x_{-it})f_t^l(x_{it})},$$

where f_t^h (f_t^l) is the size distribution of the high (low) type at time t, derived from equation (1.1). In the case of pooling IPOs, no further updates are taken as both firms are making identical actions. It is of particular importance to explicitly derive investors' "mistaken belief", which is firm l's log likelihood ratio of being the high type, $\rho_{lt} = log(\frac{P_{lt}}{1-P_{lt}})$:

$$\rho_{lt} = log(\frac{x_{lt}}{x_{ht}})\frac{\mu_h - \mu_l}{\sigma^2}.^{13} \tag{1.3}$$

ρ_{lt} measures how optimistic that the investors are about the low type through observing asset sizes in pooling IPOs. A large value of ρ_{lt} implies a wrong belief. By equation (1.1), $d\rho_{lt} = -\frac{(\mu_h-\mu_l)^2}{\sigma^2}dt + \frac{\mu_h-\mu_l}{\sigma}dB_t$, where $dB_t = \frac{1}{\sqrt{2}}dB_{lt} - \frac{1}{\sqrt{2}}dB_{ht}$. So the mistaken belief is expected to decrease strictly over time. If players are sufficiently

[13] This is because:

$$log\left(\frac{f_t^h(x_{lt})f_t^l(x_{ht})}{f_t^h(x_{ht})f_t^l(x_{lt})}\right) = log(\frac{x_{lt}}{x_{ht}})\frac{\mu_h - \mu_l}{\sigma^2}.$$

patient, assets sizes will truthfully reveal types as $\lim_{t\to\infty} E(\rho_{lt}) = -\infty$. In equation 1.3, the first part is the logarithm of the size ratio. Intuitively, a larger realized size of the low type will induce a more optimistic belief by investors. The second part measures how precise asset sizes are for discerning companies. Firm h's log-likelihood would just be the opposite of ρ_{lt} and henceforth neglected for saving the notation.[14] Given that the initial sizes are $x_{h0} = x_{l0} = 1$, $\rho_{lo} = 0$, which matches the fact that both firms are equally likely to be the high type.

Then consider the case when merger happens. Investors generate a posterior belief based on the equilibrium acquiring probabilities. Suppose that firm i acquires, the probability that it is a high-quality firm follows

$$P_i^A\left(x_{it}, x_{-it}\right) = \frac{P_i \sigma_h^A \sigma_l^T}{P_i \sigma_h^A \sigma_l^T + \left(1 - P_i\right) \sigma_l^A \sigma_h^T}.$$

Recall that P_i, σ_i^A, and σ_i^T are all functions of assets sizes and I suppress the expression to reduce clutter. Lemma 1 implies that I can focus on equilibria in which only firm h acquires. The intuition is that given a pair of (x_h, x_l), there exists no equilibria in which both types can be acquirers with non-zero probability. If so, in such a mixed strategy, both players must be indifferent between being an acquirer or a target. Otherwise, they will only participate in the deal that they strictly prefer. However, the high type generates strictly higher synergy than the low type, which means the total profit of players must be different when the identity of acquirer changes. This is contradictory.

Lemma 1. *There exist no equilibria in which firm l is acquiring with a strictly positive probability, i.e., $\sigma_l^A \sigma_h^T = 0$ for all (x_h, x_l).*

Proof. All proofs are omitted and shown separately in Appendix A.3. □

[14] By addition rule for mutually exclusive events, $P_{ht}^I + P_{lt}^I = 1$. Therefore $\rho_{ht} = log(\frac{P_{ht}^I}{1-P_{ht}^I}) = log(\frac{1-P_{lt}^I}{P_{lt}^I}) = -\rho_{lt}$.

The fact that P_i^A is a constant function also implies the current assets sizes have no impact on investors' belief. The action of being an acquirer itself is fully informative and makes the belief degenerate.

Lastly, I do not need to consider the equilibrium scenario where only one stand-alone IPO happens. Given that the low type has a dominant strategy of mimicking, all IPOs initiated by the high type are followed. The flip side implies the only possible stand-alone IPO, which is not followed, is started by the low type for certainty. But then investors will offer a fair price, which makes the low type worse-off than just accepting takeover.

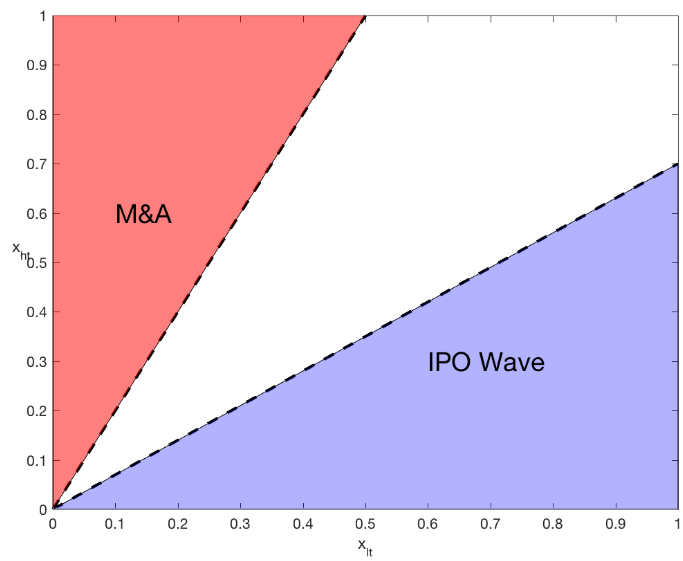
To summarize, I can focus on pooling IPOs and mergers. Therefore, ρ_{lt} is the key state variable and regulates the firm strategies. Because it essentially quantifies the mispricing of shares through equation (1.2). As in the standard real options model, I consider threshold strategies, which require that σ^i can be determined by verifying whether ρ_{lt} is above or below certain thresholds. Since firm returns are linear in its size and ρ_{lt} only changes the marginal rate of return, firm payoffs are perfectly separable. This implies that the thresholds must be constant, so the equilibrium strategy is stationary in belief and unrelated to the current stand-alone size of firm i. This is summarized in Lemma 2.

Lemma 2. *The equilibrium thresholds are constants, independent of x_{ht} and x_{it}.*

The procedure of solving the equilibrium is as follows. First, I guess that there are two constant thresholds β and η, $\eta > 0 > \beta$ such that if $\rho_{lt} \geq \eta$, both firms go IPO together and pooling IPO happens. If $\rho_{lt} \leq \beta$, the high type acquires the low type. Otherwise both firms delay expansions. Second, I solve this optimal stopping problem of the two firms using Bellman equations and boundary conditions. In this step, I show the existence and uniqueness of such threshold pair (β, η). Lastly, I verify that the guessed strategy is indeed an equilibrium by showing there exists no profitable deviations.

Figure 1.2: Characterizations of Equilibrium Outcomes in Three Regions

This figure plots the simulated equilibrium solution. Horizontal axis indicates assets size of firm l. Vertical axis indicates assets size of firm h. In top-left shaded area, the high type acquires the low type and goes IPO alone. In bottom-right shaded area, both firms go IPO without acquisition. In light area, both firms delay financial options.



1.3.2 Two-threshold Equilibrium

Figure 1.2 illustrates how the constructed two-threshold equilibrium works. The horizontal axis is the size of the low type x_{lt}, and the vertical axis is the size of the high type x_{ht}. Recall that ρ_{lt} is a function of $log(x_{lt}/x_{ht})$. Thus, a constant ρ-threshold maps to a straight line from the original point. The top-left shaded area maps to the case $\rho_{lt} \leq \beta$. When the more efficient firm has a sufficiently larger size relative to the less efficient one, investors almost perfectly identify the true types correctly. The low type has no incentives to wait as overpricing opportunities are low. So it accepts the takeover offer from firm h. On the contrary, $\rho_{lt} \geq \eta$ happens in the bottom-right area. Firm l has good lucks with large realized assets size, and investors mistakenly hold an optimistic belief about the truly less efficient firm. If so, the high type will file for IPO, aware of an upcoming underpricing cost. In the light area between them, belief is in the intermediate region, so both types postpone investments.

Given the equilibrium strategy, the next step is to quantify firm returns in different ending outcomes. The payoffs when firm h makes acquisitions are

$$
\begin{aligned}
R_h^m(x_{ht}, x_{lt}) &= (H - \alpha)\, x_{ht} + (H - L - \gamma_l)\, x_{lt}, \\
R_l^m(x_{ht}, x_{lt}) &= (L - \alpha + \gamma_l)\, x_{lt}.
\end{aligned}
$$

M&A is a "winning" scenario for the high type. By acquisition, it successfully signals its better quality to investors and receives a fair offer price. Besides, when the reservation value of low type is small, it enjoys additional net synergy through investing with larger total assets. The low type receives a fixed positive wedge above its NPV. However, this payoff may be smaller than being overpriced by optimistic investors. The payoffs in pooling IPOs are

$$R_h^I(x_{ht}, x_{lt}) = \left(H - \alpha - \frac{\alpha(H-L)\,exp(\rho_l)}{H + Lexp(\rho_l)}\right) x_{ht},$$
$$R_h^I(x_{ht}, x_{lt}) = \left(L - \alpha + \frac{\alpha(H-L)\,exp(\rho_l)}{Hexp(\rho_l) + L}\right) x_{lt}.$$

The pooling payoff consists of two parts. The first part is the stand-alone investment NPV, and the second part is the discount or premium due to mispricing. The high type is priced at a discount, and thus its total payoff is lower than NPV. The low type benefits and earns an additional premium. Therefore, pooling in IPOs is a "winning" scenario for the low type.

Each firm's trade-off is to choose between accepting its loser payoff right away and waiting for the winner payoff. Whenever one firm gives up, the other one wins automatically. This is how the game becomes a "race of unicorns".

1.3.3 Value Function

This section shows the existence and uniqueness of (η, β). The beginning step is to derive the Bellman equation of $V_i(x_{it}, x_{-it})$, and pin down the endogenous thresholds with boundary conditions. The technical difficulty is that, first, V_i has two state variables, and it requires solving a partial differential equation (PDE). Second, the threshold is characterized by belief ρ_{lt}, but V_i is a function of asset sizes. Thus the smooth pasting conditions are not clear.

However, the linearity of the model setup generates tractability, which implies V_i is homogeneous of degree one in assets size x_{it}. The value function satisfies $V_i(x_{it}, x_{-it}) = x_{it} V_i(1, x_{-it}/x_{it})$. Since ρ_{lt} is a function of $log(x_{lt}/x_{ht})$, $V_i(1, x_{-it}/x_{it})$ is essentially a function of belief. Denote it as $J_i(\rho_{lt})$. $J_i(\rho_{lt}) = V_i/x_{it}$ has concrete meanings in corporate finance. Recall that $V_i(x_{it}, x_{-it})$ is the valuation of firm i before

IPO. When x_{it} is referred as assets size, J_i mirrors the market-to-book ratio.

Since $V_i(x_{it}, x_{-it}) = x_{it} J_i(\rho_{lt})$, I can now generate the smooth pasting conditions through partial derivatives $\partial V_i / \partial \rho_{lt}$. I omit the Bellman equations of V_i and show them in (A.1) and (A.2) at Appendix A.2. There exist four value matching conditions by equations (A.3) to (A.6). There are also two smooth pasting conditions representing the optimality of thresholds. It is important to distinguish the "decision maker" in different cases. First, consider the pooling IPOs. If firm h moves first, it must be indifferent between waiting and initiating stand-alone IPO, knowing that the low type would follow. Firm l has a dominant strategy by mimicking and therefore has no indifferent conditions binding in this case. Thus the "decision maker" in pooling IPOs is the high type. Second, the low type optimally chooses to accept the offer when the high type acquires, at the cost of giving up potential pooling opportunities. Acquiring is a winning scenario for the high type. Therefore, making an acquisition offer is its dominant strategy. By applying the first-order condition for the respective decision makers, the two smooth pasting conditions follow as equations (A.7) and (A.8).

The next step is to characterize the thresholds with functions J_i. The problem reduces to solving a second-order ordinary differential equation system for $J_h(\rho_{lt})$ and $J(\rho_{ht})$:

$$(\mu_h - r) J_h(\rho_{lt}) - \left(\frac{\mu_h - \mu_l}{2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2} \right) J_h'(\rho_{lt}) + \frac{1}{2} \frac{(\mu_h - \mu_l)^2}{\sigma^2} J_h''(\rho_{lt}) = 0, \tag{1.4}$$

$$(\mu_l - r) J_l(\rho_{lt}) - \left(-\frac{\mu_h - \mu_l}{2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2} \right) J_l'(\rho_{lt}) + \frac{1}{2} \frac{(\mu_h - \mu_l)^2}{\sigma^2} J_l''(\rho_{lt}) = 0. \tag{1.5}$$

It is straightforward to show $J_h(\rho_{lt}) = C_{1h} exp(\theta_{1h}\rho_{lt}) + C_{2h} exp(\theta_{2h}\rho_{lt})$ and $J_l(\rho_{lt}) =$

$C_{1l}exp\left(\theta_{1l}\rho_{lt}\right)+C_{2l}exp\left(\theta_{2l}\rho_{lt}\right)$. θ_{ij}s are known constants as the roots of the characteristic functions in equations (1.4) and (1.5). The four constants C_{ij}s are the coefficients that will be pinned down together with the two boundaries η and β by the six boundary conditions.

The following assumptions are sufficient conditions for characterizing the solution as a two threshold equilibrium. Assumption 2 is a single-crossing condition. Both $\frac{L-\alpha+\frac{\alpha(H-L)}{H+L}}{L-\alpha+\gamma_l}$ and $\frac{H-\alpha+H-L-\gamma_l}{H-\alpha-\frac{\alpha(H-L)}{H+L}}$ represent the ratio of a player's winning payoff to its losing one if firm h's stopping strategy is $\eta \to 0$[15]. The assumption guarantees that firm l is sufficiently more resistant in waiting than firm h at the extreme case $\eta \to 0$. Without this assumption, the solved thresholds can possibly be degenerate as acquisition always happens right away. Assumption 2 restricts the learning precision to be smaller than 1, meaning that investors belief is slowly moving. This is a sufficient condition for monotonicity in proof.

Assumption 1. $\frac{L-\alpha+\frac{\alpha(H-L)}{H+L}}{L-\alpha+\gamma_l} \gg \frac{H-\alpha+H-L-\gamma_l}{H-\alpha-\frac{\alpha(H-L)}{H+L}} > 1.$

Assumption 2. $k = \left(\frac{\mu_h-\mu_l}{\sigma^2}\right)^{-1} > 1$

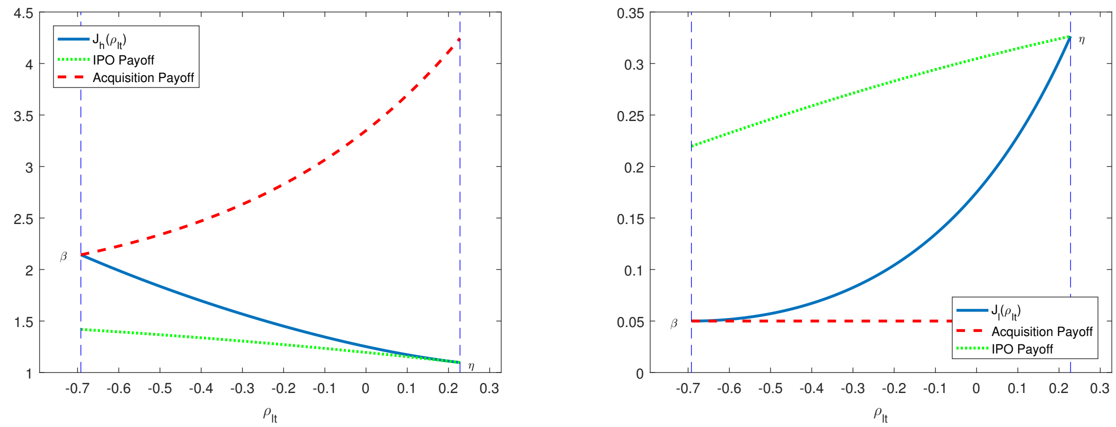
1.3.4 Existence and Uniqueness

Figure 1.3 illustrates the main idea of the two-threshold strategy. In both figures, the horizontal axis is the mistaken belief ρ_{lt} and the vertical axis is the payoff value. The dashed lines plot the payoffs if firm h acquires. The dotted lines plot the payoffs if both firms go public. The solid lines plot the waiting value J_i. The left picture represents the payoffs of the high type. The right picture represents the payoffs of the low type. For both players, their waiting value is lower than the payoff in their preferred scenarios. However, they cannot receive the winning payoffs unless their

[15]The boundary value 0 is mechanically selected as initial belief at $t = 0$ is $\rho_l = 0$. Technically a negative threshold η can exist with other assumptions. But then initial condition must be adjusted. Otherwise the game ends into pooling IPO immediately at $t = 0$.

Figure 1.3: Value Functions J_h and J_l

This figure plots the simulated value functions J_is. Left figure plots for h and right figure plots for l. Dashed lines are payoffs when h acquires l and dotted lines are payoffs in pooling IPO. Solid lines are equilibrium waiting function.



opponent give ups. This happens only if the mistaken belief is sufficiently large or small.

Consider acquisitions first. Suppose that the high type wants to acquire the low type before ρ_{lt} falls below β. Firm l finds it optimal to wait because its continuation value is higher than the payoff of being acquired. In this case, the low type will reject. The waiting value for firm l is strictly reducing as ρ_{lt} decreases, when investors become more pessimistic about the low type. This is because the low type loses its possibility of deceiving the investors. At the point when ρ_{lt} is equal to β, the discounted payoff of waiting for pooling equals to the current acquisition offer. As now firm l becomes indifferent, it accepts the takeover by firm h.

The logic of pooling IPOs follows similarly. Before belief ρ_{lt} crosses η, the high type's waiting value is strictly higher than its payoff in pooling IPOs. Thus it has no incentive to go stand-alone IPO first and let the low type imitate. So why does firm h becomes indifferent at η? Firm h is balancing between two types of costs. On the one hand, it could always go public without acquiring and suffer a short-term underpricing cost in pooling. On the other hand, it could choose to wait for the mistaken belief to adjust. This will take additional time, and waiting generates the costs of delaying profitable investments. If the current belief is extremely mistaken, adjustment takes so long that the cost in postponing investment outweighs the cost of underpricing. This explains why the high type will give up if ρ_{lt} is greater or equal to η. Theorem 3 lays out the uniqueness and existence of (β, η).

Theorem 3. *There exists a unique pair of $\eta > 0 > \beta$ that solves the optimal stopping problem. Define $P = \{(x_{ht}, x_{lt}) \,|\rho_{lt} \geq \eta\}$ and $M = \{(x_{ht}, x_{lt}) \,|\rho_{lt} \leq \beta\}$. Firm strategies are as follows:*

1. *When $\rho_{lt} \leq \beta$, the high type acquires the low type: If $(x_{ht}, x_{lt}) \in M$, $\sigma^A_{ht}(x_{ht}, x_{lt}) = \sigma^T_{lt}(x_{ht}, x_{lt}) = 1$. Otherwise $\sigma^A_{ht}(x_{ht}, x_{lt}) = \sigma^T_{lt}(x_{ht}, x_{lt}) = 0$.*

2. *When $\rho_{lt} \geq \eta$, both firms go public: If $(x_{ht}, x_{lt}) \in P$, $\sigma^I_{ht}(x_{ht}, x_{lt}) = \sigma^I_{lt}(x_{ht}, x_{lt}) = 1$. Otherwise $\sigma^I_{ht}(x_{ht}, x_{lt}) = \sigma^I_{lt}(x_{ht}, x_{lt}) = 0$.*

3. *When $\beta < \rho_{lt} < \eta$, both firms delay: For all (x_{ht}, x_{lt}), $\sigma^A_{lt}(x_{ht}, x_{lt}) = \sigma^T_{ht}(x_{ht}, x_{lt}) = 0$.*

To show this is an equilibrium, it remains to verify that no firms will deviate. At the boundaries, the decision-making type is indifferent by smooth pasting conditions, and the other type's optimality follows by dominance. I assume the following off-equilibrium belief:

1. If investors observe the deviation of a single firm going public, they will believe the deviating firm is bad type with certainty.

2. If investors observe the deviation of two firms going public, they will follow their prior belief.

3. If investors observe the deviation of a merged firm going public, they will believe the deviating firm is good type with certainty.

Theorem 4. *The strategies characterized in Theorem 3 is the unique Markov Perfect Bayesian Equilibrium given the above off-equilibrium belief.*

(i) is justified by the skimming property in this literature. In the model, whenever the high type is willing to go public, the low type has an incentive to imitate and extract information rent. The flip side is the following statement. Whenever one firm goes public but the other decides not to imitate, the former firm must be a low type. In other words, the only credible single IPO must be initiated by firm l. (iii) gives the largest incentive for firms to deviate by acquisition. It is the highest bar, and if I can prove no firms are willing to deviate given this belief, the equilibrium is robust to alternative assumptions.[16]

Verification First, no firm will deviate for a stand-alone IPO. Such deviating firm is believed to be the low type for certainty. Second, firm l could not benefit by deviating as an acquirer. The gross profit after firm l deviates is $(L-\alpha L/H)(x_{ht}+x_{lt})$. This is because investors regard the acquirer as a high type with probability one. Therefore the effective marginal cost reduces to L/H. A deviation is possible only if (i) firm l transfers sufficient payment to cover firm h's continuation value in the original equilibrium, and (ii) the net profit for firm l after payment is higher than its waiting value in equilibrium. In the proof of Theorem 4, I show that the gross profit is smaller than the total continuation payoffs:

$$V_h\left(x_{ht}, \rho_{lt}\right) + V_l\left(x_{lt}, \rho_{lt}\right) > \left(L - \alpha\frac{L}{H}\right)\left(x_{h\tau} + x_{l\tau}\right).$$

This implies it is impossible to find a transfer payment that satisfies both (i) and (ii)

[16](ii) rules one-threshold equilibrium where the high type never goes public and triggers IPO waves.

at the same time. The economics behind is that in expectation ρ_{lt} drifts down strictly. The prior belief corrects itself, and this is in favor of the high type. In other words, the high type has a strong motive to wait and $V_h(x_{ht}, \rho_{lt})$ is too large. Persuading firm h into deviation is so costly that the left profit is insufficient to motivate firm l's deviation.

Uniqueness First, Lemma 1 has already ruled out the equilibria where firm l acquires on path. Single IPO equilibria can also be ruled out with the skimming property. Thus, in the realm of two-threshold equilibria, the only two ending situations must be the high type acquiring and pooling IPOs. Since the threshold pair is unique, this is the unique two-threshold equilibrium. Second, one-threshold stopping strategies will have a two threshold deviation. For example, good firm only acquires but never triggers pooling IPOs. However, there is a profitable deviation where both firms go public when belief is sufficiently wrong.

Refinement Suppose I change (i) into the following: If investors observe the deviation of a single firm going public, they will believe the deviating firm is good type with certainty. Then the two-threshold strategy is not an equilibrium. The current belief can be refined by D1. D1 requires that after a single IPO, if there are more actions of investors that improve the equilibrium utility of type t' compared to t, then investors should believe they are facing type t' for certainty. In this model, investors' action is to give an offer price. This can be pinned down by a pseudo belief $\tilde{\rho}$ as there is a one-to-one mapping function from belief to price. Suppose the current belief is $\rho' \in (\beta, \eta)$. As shown in Figure 1.3, firm h's equilibrium waiting value is higher than the IPO offer defined with ρ' (dotted line). In other words, the lowest acceptable offer for it to deviate is pinned down by $\tilde{\rho}_h > \rho'$. On the contrary, firm l's continuation value is strictly smaller than the current IPO offer. Therefore the lowest deviating offer has a belief $\tilde{\rho}_l < \rho'$. This implies the prices that improve the high type's equilibrium payoff is a strict subset of those improving the low type's. Since the latter benefits more in this sense, investors will believe the deviating firm is the low type for sure.

Theorem 4 implies that acquiring is a good signal of quality in IPO. It is a selection process that only keeps the high type afterwards. On the contrary, when pooling IPOs happen, both firms are not acquiring and appear indistinguishable. In this case the high and low types are mixing. The average quality of non-acquiring IPOs is strictly lower than the acquiring ones. There are two testable implications of the model. First, conditional on observing private acquisitions, investors should be more confident and offer better share price. This reduces IPO underpricing. Second, acquiring IPOs should have better long-run performance since they are all high types. The average performance of non-acquiring IPOs is decreased by the existence of low types.

A sideline implication is that valuation of a single startup alone does not fully determine the timing of IPO. ρ_{lt} is a function of assets size ratio. Recently many startups with gigantic market valuations keep postponing their IPO timelines. The model argues that this is because these hefty valuations must be adjusted relatively to industry benchmarks. Investors are confident about a company's future growth only if it outperforms its competitors by a large degree. Today our economy has almost 400 unicorns[17], which implicitly raises the bar of valuation. Because these startups are hard to be differentiated from each other, they end up in the zone like $\rho_{lt} \in (\beta, \eta)$.

1.3.5 First Best Comparison

It is important to compare with the first-best solution. Without information asymmetry, investors observe the types of both firms. So if firm l goes public in a stand-alone IPO, its payoff will be the net benefit in expansion, $(L - \alpha)\, x_{lt}$. This is strictly lower

[17] The number was documented when the draft was written at August 2019. See the updated full list of startups valued at one billion or more: `https://www.cbinsights.com/research-unicorn-companies`

than the acquisition offer $(L - \alpha + \gamma_l)\, x_{lt}$. As a result, the low type will only exit in being taken over by the high type. Thus, this game becomes a standard real options model in which firm h decides when to acquire.

While firm h is balancing between waiting and making an acquisition offer right away, ρ_{lt} is still a relevant state variable. However, it now merely quantifies the ratio of asset sizes, which in turn affects firm h's acquisition payoff because larger x_{lt} generates a larger synergy. The following proposition shows that, information asymmetry is the only reason that deters efficiency reallocation as firm h acquires right away.

Proposition 5. *(No Delay) In the first-best solution, the high type will acquire the low type at the beginning of the game.*

Information imperfection in IPO markets creates an outside option for firm l as pooling becomes possible. The resistance of the low type generates delaying. On the contrary, when there is no information asymmetry, firm l will accept the acquisition offer regardlessly. In this case, firm h would only delay for generating a larger synergy. In other words, firm h would wait for a relatively larger size of firm l. However at any moment it waits, in expectation the synergy would decrease as the size ratio has a negative drift. Thus, it prefers to exercise acquisition immediately. This confirms that the efficiency loss in the private M&A market is a consequence of information imperfections.

1.4 Extensions

1.4.1 Endogenous Growth Rate

In the baseline model, the growth rates of firms are exogenously fixed. Suppose now at $t = 0^-$ (before the assets start to grow), both firms can make an one-time investment and increase the growth rate μ_i in a simultaneous subgame. In this section, I show that making these investments will hurt the individual firm itself. The result serves as a caveat for startups and their sponsoring VCs. It is a fallacy that speeding up growth by high burning rate will facilitate IPOs and receive better deals in exits.

The reason is that investors are rational and would adjust their belief process accordingly. In the model, growth rates chosen by firms are perfectly learned by the investors. Similarly, lavish burning rates and exorbitant investment speeds are observed and taken into account in reality. Investors would reasonably doubt that the large assets size is driven by capital injections rather than the underlying quality. Then they will update belief more conservatively given the same realization of sizes.

To see it in the model, recall that the belief process follows $d\rho_{lt} = -\frac{(\mu_h-\mu_l)^2}{\sigma^2}dt + \frac{\mu_h-\mu_l}{\sigma}dB_t$. Imagine that firm h increases its growth rate μ_h. This generates two effects on the belief process. First, the absolute value of drift increases (drift effect). This implies that in expectation, mistaken beliefs are corrected at a faster speed. Second, the volatility of belief is also increased (volatility effect). This implies the belief fluctuates at a larger degree. The effect of increasing μ_l is just the opposite.

Decreasing the mistaken belief ρ_{lt} at a faster speed will make firm l's resistance less valuable. Oppositely, as the volatility rises, firm l benefits in waiting in the sense of an option value. At any moment, the mistaken belief could bounce up by a larger degree, which legitimates rejecting acquisition. Therefore when μ_h increases, firm h benefits from the drift effect and but suffers from the volatility effect. The high type must judge whether the drift dominates the volatility effect. The problem for for firm l is similar. A larger μ_l decreases the drift (beneficial for firm l) and volatility (costly for firm l) simultaneously.

Proposition 6. *1. The two thresholds shift downward simultaneously as μ_h increases, i.e., $d\eta/d\mu_h$, $d\beta/d\mu_h < 0$. As a result, firm h is worse off initially, $\partial V_h\left(x_{h0}, x_{l0}\right)/\partial\mu_h < 0$.*

2. *The two thresholds shift upward simultaneously as μ_l increases, i.e., $d\eta/d\mu_l$, $d\beta/d\mu_l > 0$. As a result, firm l is worse off initially, $\partial V_l\left(x_{h0}, x_{l0}\right)/\partial\mu_l < 0$.*

The first statement indicates that as μ_h increases, the threshold for pooling IPOs, η, is closer to the initial belief but the threshold for acquisition, β, is further. This indicates the game is more likely to end in the pooling IPOs case. Therefore the expected payoff for firm h is lower. The second statement implies that as μ_l grows, changes in the thresholds are in the opposite direction. Thus firm l is more likely to be acquired. Even though both startups have the opportunities to boost their growth rates, they should choose to forsake the increments.

The volatility effect always dominates the drift effect. This can be illustrated by investigating firm l's decision at the boundary β. By equation (1.5), l's waiting value can be decomposed by the drift component related with $J_l'(\rho_{lt})$ and volatility component $J_l''(\rho_{lt})$. The smooth pasting condition indicates $J_l'(\beta)$ is 0. In other words, drift effect is minimal at the timing when the low type chooses to accept acquisition offer. The waiting value is solely determined by $J_l''(\beta)$. If the volatility of belief is larger, firm l's waiting value increases. Thus, it would choose to accept the takeover later, which explains why $d\beta/d\mu_h < 0$.

Notice the decision of firm h and l are mutually influenced. When firm l postpones its decision in accepting the takeover, firm h is anywhere worse off while waiting. It will take longer for acquisition to happen. This suggests the new waiting value for firm h decreases, which explains why $d\beta/d\mu_h < 0$.

One related question is about increasing growth rates of the industry. There is anecdotal evidence that recently all startups are inclined to premature scaling, which soon exhausts the innovation ability and burns the cash flow at an unsustainable speed.[18] The previous result highlights that individually scaling up too quickly is

[18] https://www.forbes.com/sites/nathanfurr/2011/09/02/1-cause-of-startup-death-premature-scaling/#571c23571fc9

detrimental to startups. Here a step further is taken by assuming that both firm h and l simultaneously increase their growth rates while keeping the wedge $\mu_h - \mu_l$ fixed. This mirrors the cash burning race where both firms are taking actions to enlarge their sizes, such as advertisements and price wars. Which type of firm benefits from such growth fights? How the social welfare changes in response?

Proposition 7. *Suppose $\mu_h = \mu_l + \delta$. Fixing δ and increasing μ_h and μ_l by the same degree will lower both η and β. The game ends more likely in the pooling IPOs case.*

The consequence of growth fights is that the threshold for pooling IPOs is closer to the initial belief but the threshold for acquisition is further. Therefore, pooling in IPOs is more likely to happen after both growth rates are increased. Efficient reallocation is blocked and this is a deadweight loss in social welfare. This analysis suggests that investment arms race will in the end let more poor startups become public. The funds are misallocated from the good companies to the bad ones.

If the wedge is fixed, increases in growth rates will not affect learning by investors. Therefore in Daley and Green (2012), such an increase will not affect equilibrium thresholds. This model is different. In equation (1.4) and (1.5), the effective discount rates are $r - \mu_h$ for firm h and $r - \mu_l$ for firm l. This is because as company grows, they gain in additional waiting value J_i at the rate of μ_i, which offsets discounting. Due to the fact that firm l is expected to wait longer for winning in equilibrium, the low type benefits relatively more with lower effective discount rates.

1.4.2 Nash Bargaining

In this section, I derive the acquisition offer in a Nash Bargaining way. The total surplus to divide is $(H - \alpha)(x_{ht} + x_{lt})$ when firm h acquires. The threat that both players can make is to go a stand-alone IPO while being regarded as the low type. Thus the disagreement payoff is $(H - \alpha H/L)\, x_{ht}$ for firm h and $(L - \alpha)\, x_{lt}$ for firm

l. Notice that firm h suffers maximal underpricing cost below its NPV. Denote Δ as the markup that firm l gets after bargaining, which is determined through

$$\begin{aligned} &\max_{\Delta} \quad \Delta^{1-\xi}\left((H-\alpha)(x_{ht}+x_{lt})-\Delta-(L-\alpha)x_{lt}-\left(H-\frac{H}{L}\alpha\right)x_{ht}\right)^{\xi} \\ =&\max_{\Delta} \quad \Delta^{1-\xi}\left((H-L)x_{lt}+\alpha\frac{H-L}{L}x_{ht}-\Delta\right)^{\xi}. \end{aligned} \tag{1.6}$$

$\xi \geq 0$ is the bargaining power of h. The Nash Bargaining solution $\Delta^{\star}$ that solves equation (1.6) is $(1-\xi)\left((H-L)x_{lt}+\alpha\frac{H-L}{L}x_{ht}\right)$. First, players share the synergy created by letting firm h invest. Second, firm l is in a stronger position to bargain as it suffers no informational cost from its NPV when making a threat. Unlike the baseline model, the fact that firm h will suffer from underpricing now endogenously transfers into the acquisition offer. The net payoffs for firms are

$$R_h^{NB}(x_{ht},x_{lt}) = \left(H-\alpha-(1-\xi)\alpha\frac{H-L}{L}\right)x_{ht}+\xi(H-L)x_{lt}, \tag{1.7}$$

$$R_l^{NB}(x_{ht},x_{lt}) = (L-\alpha+(1-\xi)(H-L))x_{lt}+(1-\xi)\alpha\frac{H-L}{L}x_{ht}. \tag{1.8}$$

In equation (1.7), the high type still suffers from "underpricing cost" due to the stronger threats of the low type. Firm h is balancing this cost with the distributed synergy $\xi(H-L)x_{lt}$. Its willingness to acquire depends on its bargaining power. In the extreme case, when $\xi = 0$, firm h earns no synergy and suffers a maximal underpricing cost. Therefore, it will never initiate a takeover on firm l. In effect, when firm h has small bargaining power, the only possible equilibrium outcome is pooling IPOs. The high type optimally chooses the stopping time τ to go public first and let the low type mimic.

Theorem 8. *There exists $\underline{\xi}$ such that if $\xi \leq \underline{\xi}$, firm h never acquires and goes for stand-alone IPO with belief $\rho_{lt} \leq \eta^{NB} \leq 0$. Then firm l imitates firm h's IPO. This* *is a one-threshold equilibrium that only pooling IPO happens.*

Intuitively, acquisition becomes too costly when firm l has huge bargaining power. Though firm h can perfectly signal itself, the value left on the table is even smaller than the net profit in pooling IPOs. On the contrary, when the high type dominates the negotiation and successfully restricts the markup that low type receives, the equilibrium strategies in the baseline model still holds.[19]

Theorem 9. *There exists $\bar{\xi}$ such that if $\xi \geq \bar{\xi}$ and $\frac{L-\alpha+g(1)}{L-\alpha} \gg \frac{H-\alpha+H-L}{H-\alpha-f(1)}$, the two-threshold equilibrium with a unique pair $\eta > 0 > \beta$ exists. Firm strategies are defined in the same way as Theorem 3.*

1.5 Empirical Implications

In this section, I test the empirical implications of the baseline model. Section 1.5.1 explains the construction of the sample and discusses model implications in line with the existing literature. Section 1.5.2 tests that IPOs with private acquisitions have lower first-day returns, and Section 1.5.3 shows they also have better long-term performance. Section 1.5.4 simulates a marketing timing framework where investors are rational.

1.5.1 Data

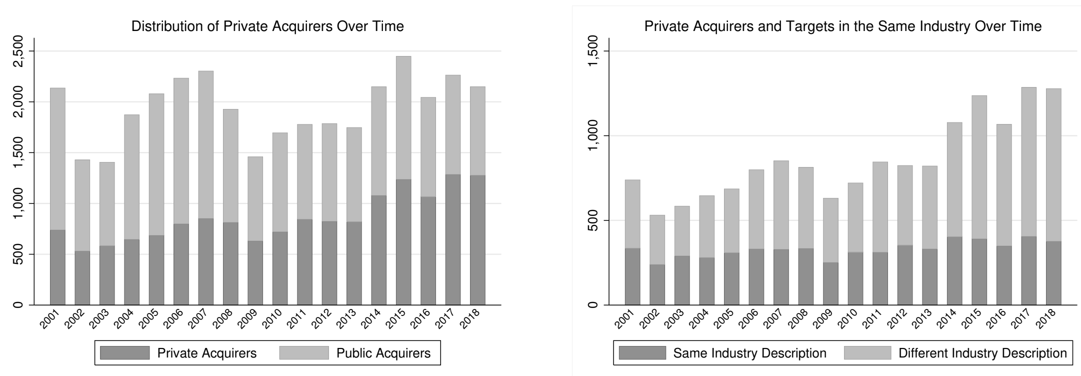
The stock issuance data are from SDC Platinum database. Following previous literature, I exclude ADRs, closed-end funds, REITs, financial companies (SIC code 6000—6799), firms not covered by CRSP within six months of offering and IPOs with offer price below $5.00 per share. I collect data on offer price, proceeds, total

[19]The requirement $\frac{L-\alpha+g(1)}{L-\alpha} \gg \frac{H-\alpha+H-L}{H-\alpha-f(1)}$ follows the same as **Assumption 2** to rule out acquisition threshold greater than 0.

Figure 1.4: Acquisitions Categorized by Private Acquirers and Industry Description

Data source is Thomson Reuters SDC Mergers and Acquisitions. All deals are selected if target company is private and belongs to high tech industry following definition of Loughran and Ritter (2004). In Panel (a), deals are categorized by whether acquirer company is public or private. In Panel (b), all deals with private acquirers are categorized by whether involved companies operate in the same industry. Industry definition follows SDC's mid-level industry definition.

(a) (b)



assets before issuance, number of bookrunners and whether company is backed by venture capitalists.

Private M&A data are from SDC Merger and Acquisition database (SDC M&A), covering deals from 2000—2017. Following Netter et al. (2011), I focus on U.S. acquirers and require the acquirer owned less than 50% of the target prior to the purchase and acquired 50% or more of the target. Acquirers not on CRSP are defined as private, and those on CRSP are public. There are many acquisitions made by private companies. For example, Figure 1.4 plots the yearly distribution of acquisitions of private targets in high tech industries categorized by whether the acquirer is public. The portion of private acquisitions is trending up in recent years. The average portion of private acquirers is 50.50%, which is comparable to both Maksimovic et al. (2013) and Netter et al. (2011).[20] Among these private deals, around one-third of them

[20]Maksimovic et al. (2013) shows 42% of asset buyers are public in a sample of U.S. manufacturing firms over the 1977 to 2004 period. Netter et al. (2011) shows only 52% of the U.S. acquirers have CRSP price data with a sample of completed mergers and acquisitions announced between 1992 and 2009.

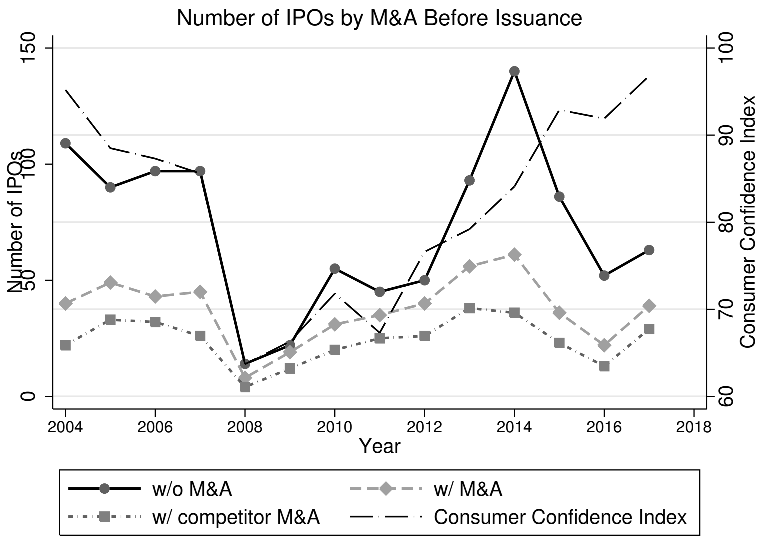
happens where the involved companies were competitors in the same industry. Given the M&A sample, I select the IPO sample period to be from 2004 to 2017 because companies that went public during 2000 to 2003 might have M&A deals before 2000 but are not covered. In total, I have $1,537$ IPOs.

The next step is to merge IPO data with M&A deals. Though both databases provide CUSIP for the issuers and acquirers, I cannot directly use it as the identifier. This is because the CUSIP and the company name documented in SDC M&A were historical upon the deal time. It is very likely that by the IPO time, the company had a new CUSIP and changed its firm name. For example, Snap Inc has CUSIP id "83304A" in the issuance database whereas it has the following name and CUSIP combinations in SDC M&A: "Snapchat Inc, 7A2488", "Snap Inc, 7A2488", "Snap Inc, 83304A", "Snap Inc, 9E4450".

I take the following procedures to match cases such as "Snapchat Inc, 7A2488" to "Snap Inc, 83304A". First I standardize the company names. Then I put each standardized company name in a separate name set NS_j indexed by a new "key" j and assign this key backwards to each firm. To illustrate, suppose now "Snap" is indexed by key "1" ($NS_1 = \{Snap\}$) and "Snapchat" is indexed by "2" ($NS_2 = \{Snapchat\}$). Secondly, for each key j, I document all CUSIPs belonged to j in a temporary set C_j. Using the previous example, this generates $C_1 = \{83304A, 7A2488, 9E4450\}$ and $C_2 = \{7A2488\}$. Thirdly, I update j in the following way. For any pair (j, j') such that $C_j \cap C_{j'} \neq \emptyset$ and $j' > j$, update NS_j to $NS_j \cup NS_{j'}$ and update all companies with key j' to key j. In other words, the higher key is dropped. In Snap Inc's example, both C_1 and C_2 contain "7A2488". All "Snapchat" companies will change their key to 1 and now $NS_1 = \{Snap, Snapchat\}$. Lastly I repeat the second and third steps until no keys are dropped. In each loop, C_j are recreated based on updated keys from previous step. The initial standardization step is very important so I manually verify it.

Figure 1.5: Number of IPOs Categorized by Private Acquisitions

This figure plots number of IPOs categorized private acquisitions. Solid line corresponds to IPOs without private acquisition. Short dashed line corresponds to IPOs with private acquisitions. Dotted line corresponds to IPOs with private acquisitions of competitors. Long dashed line corresponds to University of Michigan Consumer Sentiment Index. Left y-axis is for IPO number and the right is for consumer confidence index



Issuers and acquirers are matched based on the final converged key. Lastly, a deal is defined as *competitor M&A* if the acquirer and target operate in the same industry according to SDC's mid-level industry description. Stock return data are taken from CRSP, merged first through CUSIP and then manually supplemented. IPOs founding dates are downloaded from Jay Ritter's website or searched through Google.

Figure 1.5 plots the number of IPOs in different acquisition categories. In general, there are more IPOs without private M&As . The ratio between acquiring IPOs to the non-acquiring ones is approximately from 1:1 to 1:2. The number of acquiring IPOs is quite stable, around 50 per year, except for the recent depression in 2008 and 2009. On average, more than half of the acquiring IPOs have made acquisitions on their competitors in the same industry.

Table A.1 summarizes the statistics of main variables of the full sample as well as IPOs with private acquisitions. Acquiring IPOs tend to have higher offer price and generate more proceeds in offering. However, they are also generally larger in firm sizes and senior in firm ages. These variables also potentially help investors discern company qualities. It is possible that the lower underpricing is driven by such signals rather than previous acquisitions. So it is important to include them as control variables. The two groups are similar in other dimensions.

[Table A.1 Here]

The second row in Table A.2 shows the average first-day return for IPOs categorized by private acquisitions. In Panel A, I focus on IPOs in all industries. An IPO company has 1.646% lower first day return if it has ever made an acquisition before. Additionally, for companies that have taken over their competitors, there is 2.538% less underpricing. In Panel B, I focus on high tech IPOs. The first-day returns are consistently higher and only competitor acquisition group has less underpricing (0.683%).

Table A.2 also documents consistently better long-term operational performance measured by return on assets (ROA). Acquiring IPOs' ROAs are around 20% higher three or four years after their IPO time. Buy-and-hold returns are also higher for IPOs with private acquisitions. The concern of comparing raw cumulative returns is that they reflect different loadings on the underlying risk factors, and IPOs with private acquisitions are riskier. I address this concern through two methods. First, for each IPO, I sort it into one of 2×3 size and book-to-market portfolios using. I then use the corresponding portfolio's buy-and-hold return during the same period as a benchmark to adjust for risks. My results show that IPOs with private acquisitions still have less long-term underperformance.

[Table A.2 Here]

Before I move to the empirical tests, I discuss two implications of the model related to private acquisition patterns. First, the volume of IPOs without acquisitions fluctuates dramatically and is highly procyclical. For example, its correlation with consumer sentiment is huge. Acquiring IPOs are smoother. This is in line with the fact that private acquisitions respond less to business cycles and less wavelike. Consistently, my model predicts less private acquisitions in booming periods. This is because the less efficient firms are more likely to have optimistic beliefs then. First, when the economy grows strongly, startups can ride the tide with aggregate positive shocks, so their financial performance is usually good. Second, investors have more confidence in future growth and are optimistic about new technologies. Optimistic beliefs drive up the waiting value of bad companies. This forces the more efficient firms to withdraw their acquisition offer. In the end, we observe fewer private acquisitions and more non-acquiring IPOs.

The second implication is that information imperfections prevent efficient reallocations. A firm with higher productivity may not be able to acquire less efficient companies due to the existence of pooling IPOs. Recall that Proposition 5 implies the high type will acquire with probability one in the first-best case. In other words, firm-level productivity is a more powerful predictor of acquisition with complete information. Compared to private companies, public firms have more obligations in disclosure and therefore have less information asymmetry. This is why, in Maksimovic et al. (2013), the estimated marginal effect of productivity is 10 times larger in assets purchase decisions for public firms, compared to private companies.

1.5.2 IPO Underpricing

The main implication of the model is that private acquisitions send positive signals to investors and therefore the high type is fairly priced. If so, IPOs preceded by acquisitions will have a higher offer price and less underpricing. To test this hypothesis, I calculate the first-day returns using company's closing price in the first trading day. I estimate the following linear multivariate regression model:

$$Firstret_{i,t} = \beta MA + \gamma Controls + \eta_t + \mu_j + \epsilon_{i,t}. \tag{1.9}$$

In equation (1.9), $Firstret_{i,t}$ is the first-day return of IPO i at issuing year t. MA is an indicating dummy, equal 1 if i has made private acquisitions before IPO and 0 otherwise. Alternatively, I replace MA with MA^{comp}, a dummy variable indicating whether i has acquired a competitor before IPO. I include two types of specifications to show that the result is insensitive to the definition of competitors. In SDC, mid-level industry is classified into 85 markets such as E-commerce/B2B, Internet Software & Services and Software. These markets might be too broad or too granular in different cases, which makes MA^{comp} a noisy measure of competitor acquisition. On the contrary, the specification with MA provides a conservative estimates that are biased downwards by all non-competing acquisitions.

Following literature I include a few control variables. $ln(1 + TA)$ is logarithm of one plus total assets of IPO company before issuance. $ln(1 + age)$ is logarithm of one plus company age, defined as the years between IPO year and founding year. VC is an indicating dummy that takes 1 if company is VC-funded. $Hightech$ indicates whether the company belongs to a high tech industry following Loughran and Ritter (2004). $Bookrunners$ is the number of leading underwriters. $Nasdaq$ indicates whether the firm is listed at the Nasdaq Stock Market. I also consider including IPO year fixed effects η_t and industry fixed effects μ_j. If μ_j is specified, then $Hightech$ is dropped.

[Table A.3 Here]

Table A.3 provides the regression results. Having acquisitions significantly reduces underpricing in all specifications in Panel A. These results are robust to adding controls and different fixed effects. Notice that for each specification, replacing MA with MA^{comp} increases the magnitude of first-day return reduction. This confirms the earlier conjecture that MA is biased downwards by non-competitor M&As. In the full-fledged specification, IPOs with acquisitions have 3.136% lower first day return. The effect increases to 3.725% if competitor have been acquired.

I omit the estimates on control variables for saving spaces, but they largely conform with previous literature. As in Loughran and Ritter (2004), firm age significantly reduces first-day returns, whereas operation in the high tech industry increases them. The total assets variable is insignificant due to its common component with firm age and becomes negatively significant if the age variable is dropped. In line with recent literature (Loughran and McDonald, 2013, for eg.), I document a significant increase in underpricing if IPO is VC backed. Lastly, though insignificant, having more bookrunners slightly reduces underpricing, but listing in Nasdaq increases it.

1.5.3 Long-term Performance

The model highlights that private acquisitions are related with the better quality of startups. In this section I provide empirical support for this argument. The evidence distinguishes from the alternative mechanism that investors offer better price because they believe acquisitions reduce market competition.

In fact, investors should believe the opposite. Companies that have acquisitions along with or soon after IPO generally underperform compared to those that are not doing acquisitions (Brown et al., 2005; Brau et al., 2012). These results indicate that the consolidation of markets alone is not sufficient to sustain long term growth. A rational investor should also not reward private acquisitions if they merely help acquirers of random qualities with more monopoly power.

[Table A.4 Here]

In Panel A of Table A.4, IPO firms with private acquisitions have significantly better profitability three or four years after they go public. However, in Panel B, IPO firms that had acquired competitors do not seem to generate additional profits than firms with general private acquisitions. This is further ruling out the market competition channel.

1.5.4 Pseudo Market Timing

It is well documented that after IPO, new public firms underperform to their senior counterparts (Ritter, 1991). In an aggregate level, Baker and Wurgler (2000) find that IPOs concentrate before periods of low market performance. This evidence seems to imply that managers take advantage of investor's overoptimism and issue stocks when they know their performance in the subsequent periods is worse. In this model, firms evaluate market beliefs, which is formed rationally, when they decide whether or not to file for IPO. In a pooling IPO, less efficient firms raise fundings, and rational investors forecast this. The underpayment of the high types compensates for the expected loss in the low types. Investors are break-even given their belief as in Schultz (2003) with zero expected profit.

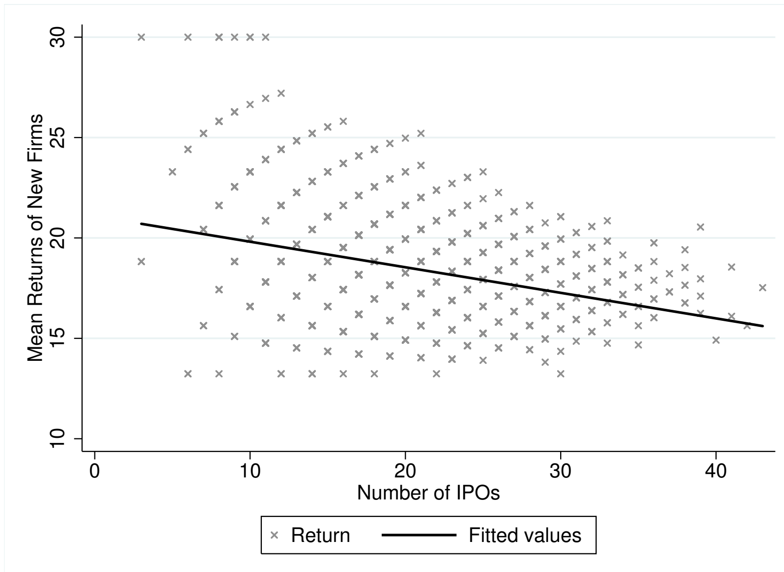
I argue that the realized performance of the new public firms varies with the number of new IPO firms. In an IPO wave, there are more pooling IPOs, containing more low-quality firms. IPO performance fluctuates due to the different characteristics of new firms on or off the wave. This composition effect coincides with the number of firms going public. So the pattern that IPO waves concentrate before periods of low market performance mechanically appears. However, this is not because the low types forecast any downturns and try to take advantage before that.

To illustrate this idea, I simulate the model in the following way.[21] Recall that

[21] This simulation is a qualitative exercise intended to illustrate the pattern of pseudo market timing.

Figure 1.6: Simulated IPO Numbers and Mean Returns

This figure plots the simulated data on number of IPOs and mean returns of new firms in each period. The horizontal axis is number of IPOs in a given period. The vertical axis is the mean return of all IPOs in that period. Solid line is the fitted linear regression.



the gross return rate of investing in the high type is $(H-\alpha)/\alpha$, and investing in the low type generates $(L-\alpha)/\alpha$. Consider amortizing the total returns in a perpetuity way. The periodic return for the high-quality firm is $r_h = r(H-\alpha)/\alpha$ and for the low-quality firm is $r_l = r(L-\alpha)/\alpha$. The simulation starts with 500 pairs of startups evolving independently, as in the baseline model. At any moment $t = k \cdot \Delta t$, (dB^n_{ht}, dB^n_{lt}) is drawn independently for the n^{th} pair of startups. If ρ^n_{lt} hits the thresholds β or η, then the pair act as in Theorem 3. Meanwhile, a new pair of startups will fill in the vacancy. At k^{th} step of the simulation, I document the number of newly public firms and the average returns of all IPOs in that step.

The parameters are $\mu_h = 0.05$, $\mu_l = -0.05$, $r = 0.15$, $\sigma^2 = 0.19$, $H = 2.55$, $L = 0.65$, $\alpha = 0.85$ and $\gamma_l = 0.25$. As a result, $r_h = 30\%$ and $r_l = -3.53\%$. The two boundaries η and β are solved to be 0.2280 and -0.6920. $\Delta t = 0.05$ and the maximal number of simulation steps is 5000. I select results from the steps 200^{th} to 5000^{th} for a stationary distribution. Figure 1.6 illustrates the result. The x-axis is the number of IPOs in a given period. The y-axis is the mean return of all IPOs in that period. For example, point $(30, 17.54)$ indicates in a period when 30 firms go public, the mean return of new firms is 17.54%. The decreasing pattern implies as more and more firms go public, the mean return afterwards is lower. I then sort the sample into deciles based on the number of IPOs in each period. The return difference between the top and bottom decile is -2.49% with a p-value less than 1%.

What if the belief of investors is biased for the less efficient types? The last part of this section revisits the overoptimism of investors by assuming that the belief process follows

$$d\rho_{lt} = -\left(\frac{(\mu_h - \mu_l)^2}{\sigma^2} - \delta\right) dt + \frac{\mu_h - \mu_l}{\sigma} dB_t.$$

Investors are no longer purely Bayesian learners. At any moment, they first update ρ_{lt} by observing size differences. However, they are mistakenly more confident about the quality of firm l. So they adjust the belief upwards by $\delta > 0$. The correction δ could be due to an inherent misconception of investors or firm l's marketing strategies. Given the new belief, the Bellman Equation of waiting functions J_i follows:

$$(\mu_h - r) J_h(\rho_{lt}) - \left(\frac{\mu_h - \mu_l}{2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2} - \delta\right) J_h'(\rho_{lt}) + \frac{1}{2}\frac{(\mu_h - \mu_l)^2}{\sigma^2} J_h''(\rho_{lt}) = 0 \tag{1.10}$$

$$(\mu_l - r) J_l(\rho_{lt}) - \left(-\frac{\mu_h - \mu_l}{2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2} - \delta\right) J_l'(\rho_{lt}) + \frac{1}{2}\frac{(\mu_h - \mu_l)^2}{\sigma^2} J_l''(\rho_{lt}) = 0 \tag{1.11}$$

A biased belief by δ only creates drift effect. The mistaken ρ_{lt} now decreases at a slower speed. Firm h expects to spend a longer time on reaching the acquisition threshold. Because delaying is costly, firm h's waiting value is anywhere strictly lower and it is now willing to give up for a pooling IPO earlier at $\eta' < \eta$. So the irrationality generates more opportunities for the less efficient types to join in IPO waves.

Proposition 10. *The two thresholds shift downward simultaneously as overoptimism δ is imposed, i.e., $d\eta/d\delta, d\beta/d\delta < 0$. Pooling IPOs become more likely. As a result, ex ante firm l is better off and firm h is worse off $\partial V_l(x_{l0}, \rho_{lo})/\partial\delta > 0$ and $\partial V_h(x_{l0}, \rho_{lo})/\partial\delta < 0$.*

1.6 Conclusion

Acquisitions by private companies are under-examined in the literature. This paper fills the gap by linking private acquisitions with signaling in IPOs. The mechanism is different from common arguments such as acquisitions reduce competition and creates monopoly power. Instead, the logic is rooted in the neoclassical view of M&As: Assets flow from the less productive firms to the ones with better technologies but not vice versa. I validate that, both theoretically and empirically, private acquisitions are determined by the quality of startups. In a real options model with information imperfections, the more efficient firm can initiate takeovers and therefore signal its quality. However, asset reallocations happen less frequently. Low types may resist selling for being overpriced in IPOs. Resistance level varies with the mistaken belief of outside investors. Especially when asset sizes are close, investors' belief can be extremely wrong. Then the high type will give up waiting because it takes prolonged expected time for the belief to self-correct.

Compared to IPOs with acquisitions simultaneously or soon afterward, I document that the issuing firms which have acquired their competitors perform significantly better after going public. This confirms the theoretical result that private acquisitions indicate better growth opportunities. Investors rationally take the deals as positive signals so that these companies have less underpricing and more proceeds in IPO.

Chapter 2

Venture Capital Contracts with Experimentation[1]

2.1 Introduction

A long line of literature in venture capital (VC) recognizes the entrepreneur's information advantage as an important source of frictions.[2] This information advantage, however, is not static. Running startups is inherently an act of experimentation, and as the startup progresses, both entrepreneurs and investors learn about its prospects.[3] Over time, this experimentation changes the entrepreneur's information advantage, and with it the optimal contract between entrepreneur and investor.

In this paper, we model startup contracts as a dynamic informed principal problem, and we characterize how experimentation changes the optimal contract over time. Common features of VC contracts emerge as equilibrium outcomes. The entrepreneur's equity share features vesting and dilution. Early payouts, pivots, and prestige projects act as signaling devices.

In our model, the entrepreneur (she) offers a contract to an investor (he) in each

[1]This chapter is co-authored with Martin Szydlowski.

[2]See Gompers (1995) for an early contribution. See also Admati and Pfleiderer (1994), Ueda (2004), and Piacentino (2019), among many others.

[3]See Kerr et al. (2014) for a summary of empirical evidence consistent with this view.

period. This contract consists of an equity share, which grants the investor a stake in the project if it succeeds, and an immediate payout. The entrepreneur is privately, but imperfectly, informed about the quality of her project, which is either good or bad. She is either a high type, who knows that her project is likely to be good, or a low type. The project requires costly experimentation by both the entrepreneur and the investor to succeed.[4] When both experiment, the good project generates a breakthrough with positive probability, while the bad project generates no breakthrough.[5] Over time, entrepreneur and investor learn from the absence of breakthroughs and revise their beliefs about the project downwards. The high and low type hold different beliefs at any point in time, but their beliefs converge as they learn about the project. Thus, experimentation gradually reduces adverse selection.

The optimal equity share features vesting and dilution. Since the investor becomes more pessimistic about the project over time, the entrepreneur pledges successively larger equity shares to prevent him from abandoning the project. Thus, the entrepreneur's share is diluted over time. Eventually, however, the low type entrepreneur starts liquidating, because she knows that her project is unlikely to succeed. Then, the investor updates his belief about the project upwards, because the likelihood that he faces the high type increases.[6] In response, the entrepreneur optimally lowers the investor's share and increases her own. This feature resembles a delayed vesting schedule: the entrepreneur's share initially decreases, but it starts to increase after sufficient time has passed. As Kaplan and Strömberg (2004) document, vesting and dilution feature prominently in VC contracts.

In equilibrium, the investor learns about the project from experimenting and from

[4]Thus, our model features double moral hazard. See e.g. Schmidt (2003), Casamatta (2003), Repullo and Suarez (2004), and Hellmann (2006). These papers do not feature experimentation.

[5]Thus, our model features an exponential bandit with good news, as in e.g. Bergemann and Hege (1998), Bergemann and Hege (2005), and Keller et al. (2005).

[6]Importantly, this is true even though the investor's belief that the project succeeds keeps decreasing. There are two conflicting effects here, and the "composition effect," i.e. the investor's changing belief about which type of entrepreneur he faces, dominates.

the entrepreneur's contract offers. This allows the entrepreneur to signal her type. We show that early payouts serve as signaling devices. Initially, the low and high type pool by offering the same contract, which consists of only an equity share. Although the high type can separate in any period, doing so is not optimal early on, because the cost of separating is too high. As time passes, however, experimentation reduces the amount of adverse selection, and therefore the cost of separating. Eventually, the high type separates by offering a payout. Then, the investor's belief about the project jumps upwards, because he learns that he is facing the high type. Again, the entrepreneur reduces the investor's share and increases her own. This resembles a management buyout (see again Kaplan and Strömberg (2004)).

We also show that pivots and prestige projects can act as signaling devices. Pivots are common among startups and many successful firms have emerged following a radical change in strategy.[7] The folklore explanation for pivots is simple: entrepreneurs realize that their idea is not working and radically change their approach. We instead show that by pivoting entrepreneurs can signal information.

Here is the intuition. Suppose that the entrepreneur can pay a fixed cost to start another project with the same ex-ante likelihood of success. The entrepreneur's type does not change once the new project is started, e.g. because it represents entrepreneurial skill.[8] Since the high type knows that her project is more likely to succeed, her value from pivoting is higher. Given the fixed cost, it may be optimal for

[7]Examples abound. Groupon initially started as a social network, Twitter emerged as a side project of an unsuccessful podcasting platform, and Instagram's founders initially developed an app for whiskey enthusiasts. See `https://www.inc.com/jeff-haden/21-side-projects-that-became-million-dollar-startups-and-how-yours-can-too.html` (last accessed 10/13/19) for other examples.

[8]This is a reasonable assumption. For many VC firms, the quality of the founding team determines whether they invest in a startup. VCs expect founders to change ideas, but they believe that founder ability is key to eventual success. See Gompers et al. (2019), who find that founder ability is the most important factor in VC financing decisions. Similarly, Gladstone and Gladstone (2002) note that "Most venture capitalists consider management to be the key to every successful venture capital investment. [...] You can have a good idea and poor management and lose every time; conversely, you can have a poor idea and a good management team and win every time."

the high type to pivot, but not for the low type. This renders separation feasible. We then characterize conditions such that optimal contract indeed features separation via a pivot.

Prestige projects are also common among early stage firms. Perhaps paradoxically, firms divert resources from their main project and use them to generate publicity and goodwill.[9] We show that prestige projects can serve as signaling devices, because they tempt low types to liquidate.

Suppose that the entrepreneur can divert resources towards a prestige project, which generates a higher outside option for her. For example, publicity may make it easier to obtain funding for another startup or to find outside employment.[10] This outside option is more appealing for the low type, who knows that her project is less likely to succeed. Once the high type implements the prestige project, the low type prefers to liquidate and takes the outside option.[11] Because of this, prestige projects can be used to signal, and we provide conditions such that the optimal contract indeed implements prestige projects.

Recent technological progress has dramatically transformed venture capital financing. As Kerr et al. (2014) report, the cost of starting internet companies has decreased radically,[12] which has prompted VC firms to adopt a "spray and pray" approach and

[9]For example, WeWork, a co-working platform, founded an elementary school (see `https://www.reuters.com/article/us-wework-wegrow/wework-to-close-its-wegrow-elementary-school-in-new-york-next-year-idUSKBN1WQ28V`, last accessed 10/13/19), Uber offered helicopter rides (see `https://www.theverge.com/2019/10/3/20897427/uber-helicopter-trips-manhattan-jfk-airport-price`, last accessed 10/13/19), and Tesla's Elon Musk has sold a device which closely resembles, but is not, a flamethrower (see `https://www.boringcompany.com/not-a-flamethrower`, last accessed 10/13/19).

[10]This is consistent with Gompers et al. (2010), who document that and entrepreneur's past performance is an important consideration in funding decisions.

[11]Importantly, we show that the low type prefers to take the outside option even though the alternative is imitating the high type.

[12]They write "firms in these sectors that would have cost $5 million to set up a decade ago can be done for under $50,000 today. For example, open-source software lowers the costs associated with hiring programmers. In addition, fixed investments in high-quality infrastructure, servers, and other hardware are no longer necessary [...] because they can be rented in tiny increments from cloud computing providers"

to fund a large number of startups with limited vetting and support. Simultaneously, cohort-based accelerators (such as Y-Combinator) have increased entry by relatively inexperienced founders. Arguably, experimentation about startups has sped up, and investors discover more quickly whether a startup is going to be successful. An important question is how these developments affect the adverse selection friction between the entrepreneur and the VC. Does technological progress alleviate adverse selection? Or does it cause adverse selection to persist longer?

Surprisingly, technological progress delays separation in our model. Adverse selection persists longer and the venture capital market is, in this sense, less efficient. We cast these results in terms of comparative statics. First, as the cost of running the startup decreases for the entrepreneur, the high type separates later. Intuitively, the lower cost makes it more appealing for the low type to imitate, which increases the cost of separating and delays separation. Second, as VCs become less involved in a startup, their cost of experimentation decreases, which also delays separation by making imitation more appealing for the low type. Third, if breakthroughs for the good project arrive more quickly, i.e. learning speeds up, adverse selection may also persist longer. Finally, we consider the effect of accelerators. When less experienced entrepreneurs found startups, the likelihood of facing the high type decreases. Then, pooling becomes costlier for the high type, and she separates earlier.

Liquidation rights are commonly used to protect investors from the entrepreneur's information advantage (see again Kaplan and Strömberg (2004)). However, increasing investors' liquidation rights may backfire, because it delays separation. The effect is similar to the one of our example on prestige projects. When investors get favorable liquidation rights, the entrepreneur's value from abandoning the project is lower. Then, the low type is willing to continue longer, which makes it costlier to separate. This result is broadly consistent with Ewens et al. (2019), who estimate that investor liquidation preference is detrimental to firm value.

Technical Contribution Our model is a dynamic informed-principal problem with experimentation. In equilibrium, each entrepreneur type and the investor learn *differently* from the absence of breakthroughs and their beliefs follow different laws of motion. In addition, signaling occurs on the equilibrium path either by the low type liquidating or by offering a different contract from the high type. As time passes, learning diminishes the amount of adverse selection, which introduces subtle dynamic incentives. The optimal contract trades off separating today against pooling and separating at a lower cost tomorrow. Despite these apparent complexities, we characterize the ex-ante optimal contract. In fact, we characterize the entire set of optimal pooling and separating Perfect Bayesian Equilibria (PBE). We do this with only minimal restrictions: the entrepreneur lacks commitment and offers a contract in each period[13] and once beliefs are degenerate, they stay that way. Limited commitment is reasonable in our setting. Startups face rapid changes and significant uncertainty, which often render contractual commitments moot.[14] The latter assumption is common in dynamic adverse selection models. It serves to avoid situations in which the investor is offered a contract which she believes will be offered with probability zero.[15]

2.2 Literature

Our paper contributes to the literature on experimentation in venture capital financing. In seminal work, Bergemann and Hege (1998) and Bergemann and Hege (2005) study optimal contracts when the entrepreneur can steal in an experimentation setting. We extend this literature by considering private information on the

[13]We share this feature with the literature on relational contracts with adverse selection, i.e. Halac (2012), Malcomson (2016), Fahn and Klein (2017), and Kartal (2018).

[14]See Kaplan and Strömberg (2001), Kaplan and Strömberg (2003), Kaplan and Strömberg (2004) and Kerr et al. (2014).

[15]See Osborne and Rubinstein (1990)'s "Never Dissuaded Once Convinced" condition. As is well-known, Bayes' rule does not apply to such situations.

entrepreneur's side and by characterizing how adverse selection changes as information arrives over time. In Bergemann and Hege's papers, the questions about signaling and separation do not arise, because information is symmetric. Hence, our results on early payouts, pivots, and prestige projects cannot be obtained in their frameworks.

Methodologically, the closest paper is Kaniel and Orlov (2018), which studies the relationship between a mutual fund family and a manager. As in our paper, there is experimentation about the manager's skill and information is revealed by both news arrival and retention/continuation decisions. In their paper, however, retention is the only signaling device, whereas in our paper, the terms of the contract also can be used to signal. Our results on vesting and dilution, early payouts, pivots, and prestige project do not appear in Kaniel and Orlov (2018).

Also closely related is Azarmsa and Cong (2018). They study a model of venture capital financing in which the entrepreneur reveals information at an interim date, at which an additional investment is required. They characterize how information is revealed and how information disclosure leads to a hold-up problem. Our paper shares a similar spirit, since in each period the investor must make payments to continue with the project. However, information in our setting is revealed both exogenously, via the arrival of breakthroughs, and indirectly, via the choice of contract and the liquidation decisions. While Azarmsa and Cong (2018) characterize the optimal security design and its interaction with optimal disclosure, we focus on how signaling reveals information over time.

Bouvard (2012) studies a real options model. In his paper, information arrives via perfect bad news, the entrepreneur has private information about the project, and she commits to a contract which specifies the duration of experimentation together with performance-contingent payments. In Bouvard's paper, signaling occurs through excessively delaying investment and through cash-flow rights. Our paper features fundamentally different economic forces, which lead to different predictions. In our setting, the project cannot be delayed, because we do not have a real option. Instead, the project has a chance to succeed in each period as long as the entrepreneur and the investor do not liquidate. As we show, signaling though cash-flow rights is possible, but never optimal. Instead, signaling occurs via payouts, pivots, or prestige projects, which play no role in Bouvard's paper.

Similar to Bouvard (2012), a number of papers study signaling via the length of experimentation, i.e. Grenadier et al. (2014), Dong (2016), and Thomas (2019). In these papers, the experimenter has private information and can choose to continue experimenting for an excessive amount of time to signal his type. In our setting, the entrepreneur chooses both how long to experiment and which contract to offer to investors. We characterize how signaling shapes the optimal contract, while there is no contracting in the above papers.

To render our analysis tractable, we borrow from the literature on relational contracts with adverse selection (i.e. Halac (2012), Fahn and Klein (2017), and Kartal (2018)). Just as these papers, we assume that the entrepreneur, who acts as the principal in our setting, does not have commitment and optimizes period-by-period. The key difference is that in our model, all parties learn about the project by observing whether a success arrived, so that the degree of adverse selection changes over time. By contrast, in Halac (2012), Fahn and Klein (2017), and Kartal (2018), there is no exogenous information about the principal's type and agents can learn only by observing the principal's choices. Indeed, the arrival of information is crucial for our results. Without it, the high type would separate either immediately or never and there would be no dynamics.

2.3 Model

Environment An entrepreneur (she) needs to contract with an investor (he) to start a project. The project is either good or bad. It requires costly experimentation by both the entrepreneur and the investor. When both experiment, the good project generates a single payoff V, which realizes with probability $\lambda \in (0,1)$ in each period $t \in \{1,2,...\}$. The bad project never generates a payoff.[16] Once either the entrepreneur or the investor stops experimenting, the project is irreversibly liquidated.

The entrepreneur is privately, but imperfectly, informed about the quality of the project. We denote the entrepreneur's type with $\theta \in \{l,h\}$. A high type entrepreneur knows that the project is good with ex-ante probability p_1^h, while a low type knows that the project is good with probability p_1^l, where $0 < p_1^l < p_1^h < 1$. The ex-ante likelihood of a high type is $q_0 \in (0,1)$.[17] Both parties are risk-neutral and have a common discount factor $\delta \in (0,1)$.

Contracts At the beginning of each period t, the entrepreneur chooses a liquidation probability $l_t^\theta \in [0,1]$. If she continues, the entrepreneur pays a cost $k > 0$ and offers the investor a contract $C_t^\theta = \left(d_t^\theta, \alpha_t^\theta\right)$. This contract consists of a payout[18] $d_t^\theta \geq 0$ and an equity share $\alpha_t^\theta \in [0,1]$. The equity share is contingent on the project succeeding, but d_t^θ is not contingent and paid immediately instead.[19] Given the contract, the investor chooses whether to continue $(e_t = 1)$ or whether to abandon the project $(e_t = 0)$.[20] Continuing has cost $c > 0$ for the investor. This cost can represent a cash

[16] Thus, we have an exponential bandit with good news, as in Bergemann and Hege (1998), Bergemann and Hege (2005), and Keller et al. (2005).

[17] In the model, q_t is updated at the beginning of each period, while p_t is updated at the end. This is why our notation for the ex-ante probabilities, i.e., p_1^θ vs. q_0, differs.

[18] Or, equivalently, any costly action which does not affect the project value or likelihood of success.

[19] Once the equity is pledged, it is enforceable. Thus, the entrepreneur cannot renege once the project succeeds, unlike in e.g. Halac (2012).

[20] Thus, offering a contract for which the investor is not willing to experiment is the same as liquidating the project. We retain l_t^θ for notational clarity.

investment which is made each period, the opportunity cost of already committed funds, or advising effort. If the project is liquidated, the entrepreneur and investor each receive an outside option of zero.[21]

Beliefs The high type entrepreneur, the low type entrepreneur, and the investor each have different beliefs about the likelihood that the project is good. Figure 2.1 shows how beliefs are updated.

Each entrepreneur type learns from the absence of successes. Type θ enters period t with belief p_t^θ. Without a success, she updates her belief to

$$p_{t+1}^\theta = \frac{(1-\lambda)\, p_t^\theta}{1-\lambda p_t^\theta}, \tag{2.1}$$

which is strictly decreasing over time. Equation (2.1) implies that $p_t^h > p_t^l$ for all t. That is, the high type believes that her project is more likely to succeed at all times.

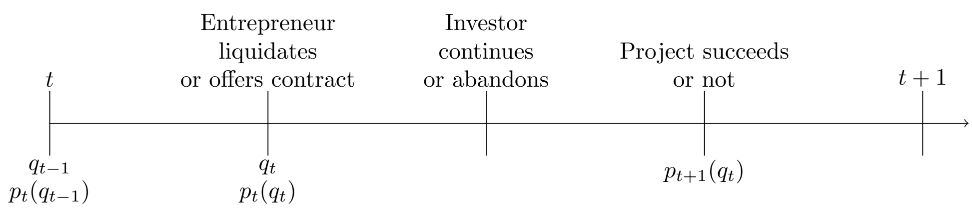
The investor learns from two sources. First, he may learn about entrepreneur's type from the contract offered. At the beginning of period t, he believes he is facing the high type with probability q_{t-1}. Upon observing the contract, he updates this belief to q_t. From the investor's perspective, the likelihood that the project is good is then

$$p_t\,(q_t) = q_t p_t^h + (1-q_t)\, p_t^l. \tag{2.2}$$

Second, if the project does not succeed, he updates this belief to $p_{t+1}\,(q_t)$, using Bayes' rule in Equation (2.1).

[21] The value of the outside option does not affect the qualitative properties of the contract. We set it to zero to simplify notation. We study the case when the entrepreneur can choose different projects, which have different outside options, in Section 2.7.2.

Figure 2.1: Timeline and Beliefs



Payoffs Denote with τ the period in which the project ends, either because it succeeds or because it is liquidated.[22] Denote with $\mathbf{1}_s$ the indicator function which is one if and only if the project succeeds in period s.

In period t, the payoffs for the type-θ entrepreneur and the investor are

$$\Pi_t^\theta = E_t^\theta \left[\sum_{s=t}^{\tau} \delta^{s-t} \left(\mathbf{1}_s \left(1-\alpha_s\right) V - k - d_s \right) \right] \tag{2.3}$$

and

$$U_t = E_t \left[\sum_{s=t}^{\tau} \delta^{s-t} \left(\mathbf{1}_s \alpha_s V - c + d_s \right) \right]. \tag{2.4}$$

Here, E_t^θ is the expectation given type θ's belief p_t^θ and E_t is the expectation given the investor's beliefs $p_t\left(q_t\right)$ and q_t.

For $t < \tau$, the entrepreneur's and investor's values can be written recursively as

$$\Pi_t^\theta = \left(1 - l_t^\theta\right) \left(\lambda p_t^\theta \left(1-\alpha_t\right) V - k - d_t^\theta + \delta \left(1 - \lambda p_t^\theta\right) \Pi_{t+1}^\theta \right) \tag{2.5}$$

[22]Since liquidation is irreversible, $e_t = 1$ for all $t < \tau$.

and

$$U_t = (1 - l_t(q_{t-1}))\left(\lambda p_t(q_t)\,\alpha_t V - c + q_t d_t^h + (1-q_t)d_t^l + \delta(1 - \lambda p_t(q_t))\,U_{t+1}\right), \quad (2.6)$$

where

$$1 - l_t(q_{t-1}) = q_{t-1}\left(1 - l_t^h\right) + (1 - q_{t-1})\left(1 - l_t^l\right)$$

is the investor's expectation about the entrepreneur's liquidation probability.

Equilibrium Concept We focus on Perfect Bayesian Equilibria. A Perfect Bayesian Equilibrium is a set of strategies and posterior beliefs, such that the strategies are sequentially rational at each history given the beliefs, and the beliefs are updated according to Bayes' rule whenever possible. We provide a formal equilibrium definition in Appendix Appendix B.1. Following Halac (2012), we require Bayesian updating both on and off the equilibrium path. Bayes' rule does not apply at histories at which the investor's belief about the entrepreneur is degenerate.[23] We follow the literature in making the following assumption.[24]

Assumption 3. *If, at any history, the investor believes he is facing type θ with certainty, he will continue to believe so no matter which contract is offered.*

Throughout the paper, we refer to a Perfect Bayesian Equilibrium as *equilibrium.* We consider pooling and separating equilibria. In a pooling equilibrium, both types offer the same contract each period, but the low type may liquidate the project earlier and thereby reveal her type. In a separating equilibrium, types *separate in period t* if

[23]That is, $q_t \in \{0, 1\}$. These histories arise after the high type successfully separates from the low type. See e.g. Section 2.5.2.

[24]This, or similar assumptions, are common in dynamic adverse selection models. See e.g. Osborne and Rubinstein (1990)'s "Never Dissuaded Once Convinced" condition.

they pool until period $t-1$ and offer different contracts in period t. An equilibrium contract is *optimal* if it maximizes a weighted average of the low and high type's ex-ante values, where $\gamma \in [0,1]$ is the weight on the high type.[25]

Parametric Assumptions To avoid uninteresting cases, we maintain the following assumptions throughout the paper.

Assumption 4. *In the first best, the good project is never liquidated, i.e.,*

$$\lambda V > k + c, \tag{2.7}$$

and in the pooling equilibrium, the low type does not immediately liquidate, i.e.,

$$\lambda p_1^l \left(1 - \frac{c}{\lambda p_1\left(q_0\right) V}\right) V > k. \tag{2.8}$$

Without Equation (2.7), the entrepreneur would immediately liquidate the project in any equilibrium. Without Equation (2.8), there may exist a pooling equilibrium in which either the low type or both types immediately liquidate the project. Then, the investor's belief evolution is trivial. He either learns nothing (if both liquidate) or immediately learns he is facing the high type (if only the low type liquidates).

2.4 Discussion

We now discuss how our modeling assumptions map to observed patterns in venture capital and how they relate to existing literature.

[25]We focus on optimal PBE throughout the paper. As a robustness check, we adapt the D1 criterion of Cho and Kreps (1987) to our setting and show that our results are qualitatively unchanged. The extension is in Section 2.8.

Venture Capital In our model, both the entrepreneur and investor exert effort. This modeling choice is consistent with a long line of literature on "double moral hazard" in venture capital (see Schmidt (2003), Casamatta (2003), Repullo and Suarez (2004), and Hellmann (2006)). It is also consistent with a substantial empirical literature, which documents that venture capital investors provide valuable services to entrepreneurs (see Sahlman (1990), Gorman and Sahlman (1989), Lerner (1995), and Hellmann and Puri (2000)). These services, which include providing advice, helping determine strategy, or helping recruit talent, are important for a firm's success (see e.g. Kortum and Lerner (2000) and Bernstein et al. (2016)). In addition to effort, we can interpret the investor's cost c as an opportunity cost of already committed funds or as investments under a given staging structure. We can interpret the investor's exit as the firm shutting down or being bought out, or as the founder being replaced (see Wasserman (2003)). We can interpret the arrival of a success as an IPO.

Startup firms are subject to rapid changes and significant uncertainty, which often render contractual commitments moot (see Kaplan and Strömberg (2001), Kaplan and Strömberg (2003), Kaplan and Strömberg (2004) and Kerr et al. (2014)). Our modeling of contracts is consistent with this view. Instead of committing to a long-term contract at the beginning, the entrepreneur offers a sequence of contracts to the investor.[26]

Our model features binary outcomes (either success or no success). Thus, every payment contingent on success is equivalent to an equity share. While this is stylized, equity indeed makes up a majority of venture capitalists' compensation (see e.g. Kaplan and Strömberg (2004)).[27]

[26]Similar commitment issues underlie the literature on incomplete contracts and control rights (see Aghion and Tirole (1994), Aghion et al. (2004), and Dessein (2005)) and the literature on holdup problems (see Rajan (1992), Admati and Pfleiderer (1994), Burkart et al. (1997), Gompers and Lerner (2004), Azarmsa and Cong (2018), and Inderst and Vladimirov (2019)). The particular allocation of control rights is irrelevant in our model. If, as in the literature on hold-up, we assume that the investor makes all liquidation decisions, all results remain unchanged.

[27]Specifically, in their Table 1, the vast majority of financing is in the form of convertible preferred

Experimentation and Adverse Selection We assume that entrepreneur and investor learn about the firm over time. This is consistent with Kerr et al. (2014), who document that even conditional on investing, VCs face significant residual uncertainty, and with Ewens et al. (2018), who document that investors adjust their contracts as information becomes available. Our modeling of learning follows Bergemann and Hege (1998) and Bergemann and Hege (2005), who also assume that information arrives in the form of successes or their absence. Given the substantial skewness of returns in the venture capital industry, which features few startups with high profits and many startups with profits close to zero, this assumption is reasonable (see Scherer and Harhoff (2000) and Hall and Woodward (2010)).

An overwhelming part of the venture capital literature highlights the entrepreneur's information advantage as a source of frictions, going back at least to Gompers (1995). As investors learn about the firm, however, the information advantage disappears and the optimal contract changes (see Kaplan and Strömberg (2003) and Ewens et al. (2018) for evidence). This is exactly what happens in our model. As time passes, the projects of the good and bad type become indistinguishable. This evolution of the adverse selection friction is a key driver for our results.

Alternative Formulations In our model, the entrepreneur has private information, chooses the contract, and has all bargaining power. This is a common modeling choice (see e.g. Gale and Hellwig (1985), Innes (1990), and Bouvard (2012)), which can easily be relaxed. If, as in Axelson (2007), we assume that the investor has private information and chooses the contract, our entire analysis goes through, except that the roles of the entrepreneur's and investor's share are reversed.

Alternatively, we could interpret the model as contracting between a founder

---

stock. This type of equity is equivalent to straight equity in our model. We keep the security design simple and instead focus on the role of learning and adverse selection. See Schmidt (2003), Casamatta (2003), and Hellmann (2006) for models in which convertible and straight equity differ.

and an early employee, who is compensated by a significant equity stake. Such arrangements are common in startups (see Hand (2008)) and other industries (see Eisfeldt et al. (2018)). In this interpretation, the employee learns about the firm's prospects the longer he is employed and prefers to leave if the prospects become sufficiently unfavorable. None of our results change.

Recently, Robb and Robinson (2014) and Mann (2018) have documented that bank finance is a significant source of startup capital. Although we prefer the interpretation with equity and venture capitalists, our model is consistent with this alternative. If the investor is a bank, the cost c can represent monitoring effort. Since our outcome is binary, debt and equity contracts are equivalent. Thus, we can define a face value of debt F_t, which is to be repaid once the project succeeds, so that the entrepreneur's payout is $V - F_t$. This face value is renegotiated continuously, as in Hart and Moore (1998).

2.5 Analysis

We start with some notation. We continue denoting a generic payoff for type θ with Π_t^θ. We denote the payoff given belief q_t and contract C_t as $\Pi_t^\theta\left(q_t, C_t\right)$, irrespectively of whether this is on or off the equilibrium path. Finally, we denote with $\Pi_t^\theta\left(q_t\right)$ the *equilibrium* payoff given belief q_t. All proofs are in the Appendix.

2.5.1 Symmetric Information Benchmark

Suppose that the entrepreneur's type is public and that she offers a contract $\bar{C}_t^\theta = \left(\bar{d}_t^\theta, \bar{\alpha}_t^\theta\right)$. The investor's belief about the project is the same as the entrepreneur's, i.e., $p_t\left(q_t\right) = p_t^\theta$. The investor is willing to continue the project whenever

$$\lambda p_t^\theta \alpha_t^\theta V - c + \delta\left(1 - \lambda p_t^\theta\right) U_{t+1} \geq 0. \tag{2.9}$$

The left-hand side (LHS) is the investor's payoff from continuing,[28] which must exceed his outside option. The optimal contract leaves the investor indifferent between continuing or abandoning the project. That is, $U_t = 0$ for all t, the optimal share is

$$\bar{\alpha}_t^\theta = \frac{c}{\lambda p_t^\theta V}, \tag{2.10}$$

and $\bar{d}_t^\theta = 0$. Any other contract can be improved upon by the entrepreneur. If $U_t > 0$ for some t, lowering either d_t^θ or α_t^θ increases the entrepreneur's payoff without violating Equation (2.9).

The optimal equity share is increasing in time. As time passes without a success, the investor becomes more pessimistic about the project, i.e. p_t^θ decreases. Then, the entrepreneur must pledge a larger share to ensure that the investor continues. Moreover, the low type pledges a larger share than the high type, i.e. $\bar{\alpha}_t^l > \bar{\alpha}^h$, because the likelihood that the low type's project succeeds is lower.

Given the optimal share $\bar{\alpha}_t^\theta$, the entrepreneur's payoff in period t is

$$\Pi_t^\theta = \left(1 - l_t^\theta\right)\left(\lambda p_t^\theta V - c - k + \delta\left(1 - \lambda p_t^\theta\right)\Pi_{t+1}^\theta\right). \tag{2.11}$$

When t becomes large, the entrepreneur's value becomes negative, because the project is unlikely to succeed. Then, she liquidates. Since the low type's project is less likely to succeed than the high type's, the low type liquidates earlier. We summarize these results in the following Lemma.

Lemma 11. *The type-θ entrepreneur offers share*

$$\bar{\alpha}_t^\theta = \frac{c}{\lambda p_t^\theta V}$$

[28]Recall that at the time when the investor decides whether to continue, d_t^θ has already been paid and is therefore sunk.

and liquidates the project whenever $\lambda p_t^\theta V - c - k \leq 0$. Let τ^θ be the period in which liquidation occurs under symmetric information. We have $\tau^l \leq \tau^h$ and $\bar{\alpha}_t^l > \bar{\alpha}_t^h$ for all $t < \tau^l$.

In the following, we denote the high and low type's symmetric information payoffs as $\Pi_t^h(1)$ and $\Pi_t^l(0)$.[29]

2.5.2 Cashless Entrepreneur

With private information, offering the symmetric information contracts is not incentive compatible. Since $\bar{\alpha}_t^h < \bar{\alpha}_t^l$, the low type prefers to imitate the high type, because then she can offer a lower equity share. This is the source of adverse selection in our model.

We first consider a cashless entrepreneur who cannot provide payouts to the investor unless the project succeeds. That is, we set $d_t^\theta = 0$ for all t and θ. Then, the optimal equity contract is pooling. Both types offer the same equity share, but the low type liquidates earlier than the high type and thereby reveals her type. The entrepreneur's share first decreases and then increases. This resembles dilution, i.e. as the project continues the entrepreneur's share becomes increasingly diluted, and vesting, i.e. after enough time has passed, the entrepreneur's shares vest and her stake in the firm increases. Both of these features are prominent in VC contracts (see Kaplan and Strömberg (2001)).

Proposition 12. *The following pooling contract is optimal. There exist two periods $\underline{\tau}^l \leq \bar{\tau}^l$, such that both types continue for $t < \underline{\tau}^l$. The low type liquidates with positive probability for $t \geq \underline{\tau}^l$ and liquidates with certainty in period $\bar{\tau}^l$. For all $t < \bar{\tau}^l$, both*

[29] That is, if $\bar{C}_t^\theta$ is the optimal symmetric information contract, then $\Pi_t^h(1) = \Pi_t^h\left(1, \bar{C}_t^h\right)$ and $\Pi_t^l(0) = \Pi_t^l\left(0, \bar{C}_t^l\right)$.

types offer equity share

$$\alpha_t^P = \frac{c}{\lambda p_t\left(q_t\right) V}. \tag{2.12}$$

This share is increasing in time for $t < \underline{\tau}^l$ and decreasing for $t \geq \underline{\tau}^l$. After the low type liquidates, the high type offers $\bar{\alpha}_t^h$ and continues until period τ^h.

We now informally derive the main results of the proposition.[30] The low type knows that her project is less likely to succeed. Thus, when offering the same contract as the high type, her value from continuing is lower. After enough time without a success, the low type starts liquidating with positive probability, while the high type continues. Thus, even though both types offer the same contract, the investor learns the entrepreneur's type over time. Based on the low type's liquidation strategy, the investor updates his belief according to

$$q_t = \frac{q_{t-1}}{q_{t-1} + \left(1 - q_{t-1}\right)\left(1 - l_t^l\right)}. \tag{2.13}$$

The belief is constant when the low type does not liquidate ($l_t^l = 0$) and increasing when she does ($l_t^l > 0$). Since the low type never liquidates before period $\underline{\tau}_l$, we have $q_t = q_0$ for any $t < \underline{\tau}^l$.

The investor continues the project whenever

$$\lambda p_t\left(q_t\right) \alpha_t^P V - c + \delta\left(1 - \lambda p_t\left(q_t\right)\right) U_{t+1} \geq 0. \tag{IC_I}$$

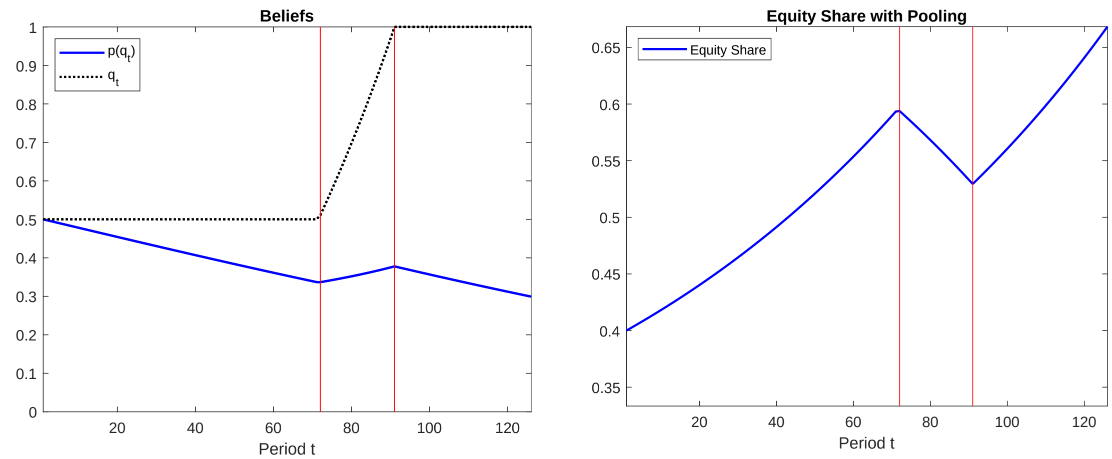
The optimal pooling equity share (in Equation (2.12)) leaves the investor indifferent between continuing and abandoning the project. Any higher share is suboptimal, because both types can lower the share until the investor's incentive compatibility (IC) condition (IC_I) binds.

[30] The formal proof is in Appendix Appendix B.2.

Figure 2.2: Pooling Equilibrium

On each panel, the left vertical line indicates $\underline{\tau}^l$ and the right vertical line indicates $\bar{\tau}^l$. Before $\underline{\tau}^l$, the investordoes not change his beliefs about the entrepreneur's type, because neithertype liquidates. Between $\underline{\tau}^l$ and $\bar{\tau}^l$, the low type liquidates with positive probability, sothat q_t increases. This causes $p_t\left(q_t\right)$, theinvestor's belief about the project, to increase (left panel).Between $\underline{\tau}^l$ and $\bar{\tau}^l$, the optimal pooling share decreases (right panel), because theinvestor becomes more optimistic about the project, which resembles avesting schedule for the entrepreneur.

(a) Investor's Beliefs (b) Investor's Equity Share



The optimal equity share is increasing in time when the low type does not liquidate, i.e. before period $\underline{\tau}^l$, because the investor's belief $p_t\left(q_0\right)$ is decreasing.[31] To keep the investor indifferent, his share must increase. However, starting from period $\underline{\tau}^l$, the low type liquidates with positive probability and the equity share is decreasing. Intuitively, the low type must be indifferent between liquidating and continuing, i.e. $\Pi_t^l = \Pi_{t+1}^l = 0$, and Equation (2.5) reduces to

$$\lambda p_t^l\left(1-\alpha_t^P\right)V = k.$$

The belief p_t^l decreases over time, so the equity share must also decrease to keep the low type indifferent. The equilibrium liquidation probability l_t^l, together with Bayes rule in Equation (2.13), ensure that the investor continues the project in any

[31] Recall that both p_t^h and p_t^l decrease without a success. Keeping q_t at q_0, $p_t(q_t)$ decreases as well.

such period.[32] Intuitively, when the low type liquidates with positive probability, q_t increases, and thus the investor is willing to continue despite receiving a lower share.

In period $\bar{\tau}^l$, the low type liquidates with certainty and the investor learns whether he is facing the high type. The high type then offers the symmetric information contract $\bar{\alpha}_t^h$. Figure 2.2 illustrates the dynamics of the investor's equity share and beliefs.[33]

In addition to pooling equilibria, there exist equilibria in which the high type separates in period t. That is, both types offer a pooling contract before period t, and the high type separates in period t by offering an inefficiently large equity share. However, as we show next, all separating equilibria are suboptimal.

Proposition 13. *For any $t < \bar{\tau}^l$, there exists an equilibrium in which the high type separates in period t by offering a share $\alpha_t^h > \alpha_t^P$ and both types pool before period t. Any such equilibrium is suboptimal.*

In a separating equilibrium, the following IC conditions must hold. The low type prefers to reveal her type instead of offering the high type's equity share α_t^h, i.e.

$$\Pi_t^l(0) \geq \lambda p_t^l \left(1 - \alpha_t^h\right) V - k + \delta \left(1 - \lambda p_t^l\right) \Pi_{t+1}^l (1) . \qquad (IC_l)$$

Similarly, the high type prefers to offer α_t^h, so that

$$\Pi_t^h(0) \leq \lambda p_t^h \left(1 - \alpha_t^h\right) V - k + \delta \left(1 - \lambda p_t^h\right) \Pi_{t+1}^h (1) . \qquad (IC_h)$$

The continuation values following separation are the symmetric information values

[32] That is, we have for $\underline{\tau}^l \leq t < \bar{\tau}^l$,

$$\alpha_t^P = \frac{\lambda p_t^l V - k}{\lambda p_t^l V} = \frac{c}{\lambda p_t\left(q_t\right)} V.$$

[33] The investor's share is α_t, which is depicted in Figure 2.2, and the entrepreneur's share is $1 - \alpha_t$.

$\Pi^h_{t+1}(1)$ and $\Pi^l_t(0)$. If the low type imitates the high type, she optimally offers the high type's symmetric information share $\bar{\alpha}^h_{t+1}$ the next period, and we denote her value with $\Pi^l_{t+1}(1)$, while if the high type imitates the low type, she optimally offers $\bar{\alpha}^l_t$ and receives $\Pi^h_t(0)$.[34]

Intuitively, pledging a large share dissuades the low type from imitating, because she has to give up a larger portion of the project's value if it succeeds. Then, the low type prefers instead to be discovered. The high and low type's values satisfy a variant of single crossing in any period $t < \bar{\tau}^l$,[35]

$$\frac{\Pi^h_t(1) - \Pi^h_t(0)}{p^h_t} \geq \frac{\Pi^l_t(1) - \Pi^l_t(0)}{p^l_t}. \tag{2.14}$$

That is, the value of being perceived as the high type is larger for the high type than for the low type.[36] Because single crossing holds, the high type can separate in any period.

However, separating via a higher share is relatively costly. The high type's project is more likely to succeed, and thus she is more likely to pay the investor. If she increases the share by ε, she reduces the low type's value from imitating by $\lambda p^l_t V\varepsilon$ and her own value by $\lambda p^h_t V\varepsilon$. Thus, to reduce the low type's value by one, the high type has to give up value $p^h_t/p^l_t > 1$. Because of this, the high type's cost of deterring the low type exceeds her benefit from separating. Thus, the optimal equilibrium is pooling.

[34] This is because of Assumption 3. After types separate, the investor never changes his belief, and he will accept any contract which promises a share of at least $\bar{\alpha}^h_t$ (if $q = 1$) or $\bar{\alpha}^l_t$ (if $q = 0$). The optimal contract for type θ is then the symmetric information contract of Section 2.5.1.

[35] See Lemma 34 in Appendix Appendix B.2.3.

[36] Intuitively, if type θ is being perceived as the high type in a future period $s > t$, she can offer a share $\bar{\alpha}^h_s$, while when she is perceived as the low type, she offers share $\bar{\alpha}^l_s$. Type θ's value of being perceived as the high type is thus $\lambda p^\theta_s(\bar{\alpha}^l_s - \bar{\alpha}^h_s)$. This value is larger for the high type, whose project is more likely to succeed. After discounting and considering the high and low type's liquidation decisions, this leads to Equation (2.14).

2.5.3 Cash Payments

We now consider the case with payouts and show that early payouts can be used to signal. For tractability, we impose the following parametric assumption.

Assumption 5. *We have* $\left(1-\lambda p_1^h\right)\left(1-\lambda p_1^h\right) > 1-\lambda$.

Assumption 5 implies that the degree of adverse selection, as measured by the difference in the high and low type's beliefs, $p_t^h - p_t^l$, is decreasing over time.[37] It holds whenever the initial probabilities p_1^l and p_1^h are sufficiently small.

In the optimal contract, the high type separates by offering a payout in period τ_S, but separation is inefficiently delayed. Intuitively, the high type prefers to wait until the degree of adverse selection has decreased, since separating earlier is too costly.

Proposition 14. *If q_0 is sufficiently small and γ is sufficiently large, the optimal contract is separating in period τ_S. It consists of a payment $d_{\tau_S}^h$ and the high type's symmetric information share $\bar{\alpha}_{\tau_S}^h$. Before period τ_S, both types offer the pooling contract of Proposition 12. If either q_0 is sufficiently large, or if γ is sufficiently small, pooling is optimal.*

With payouts, the pooling equilibrium of Proposition 12 still exists. Moreover, it is optimal among all pooling equilibria.[38] Suppose that the high type separates in period t by offering a contract $C_t^h = \left(d_t^h, \alpha_t^h\right)$. The relevant IC conditions are

$$\Pi_t^l(0) \geq \lambda p_t^l \left(1-\alpha_t^h\right) V - d_t^h - k + \delta\left(1-\lambda p_t^l\right) \Pi_{t+1}^l(1) \qquad (\tilde{IC}_l)$$

[37] This is because

$$p_{t+1}^h - p_{t+1}^l = \frac{1-\lambda}{\left(1-\lambda p_t^h\right)\left(1-\lambda p_t^h\right)}\left(p_t^h - p_t^l\right).$$

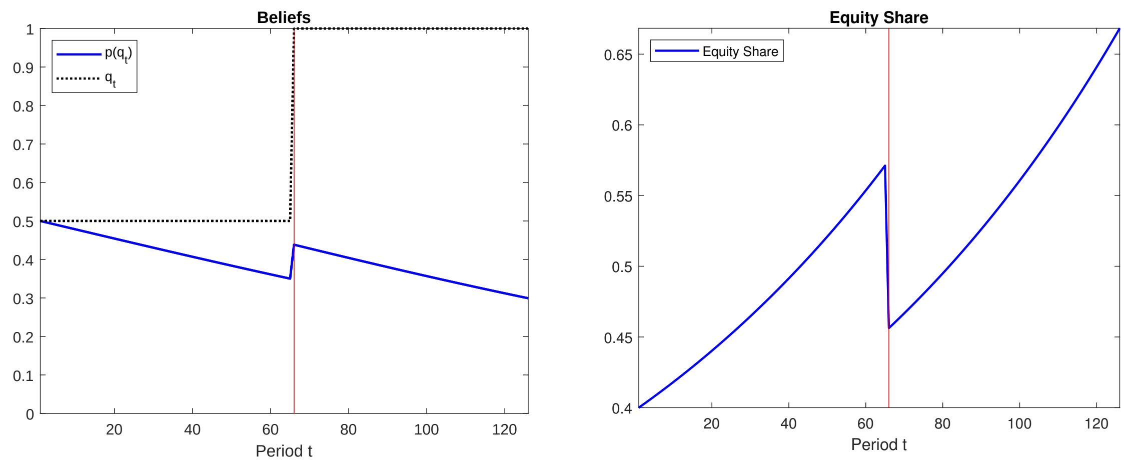
Since $p_t^\theta < p_1^\theta$, it is sufficient to show this condition holds at $t=1$. We use Assumption 5 to establish a single-crossing condition in Lemma 39. Note that $p_t^h - p_t^l$ is decreasing at time t whenever p_t^h and p_t^l are sufficiently small. Thus, even without the assumption, $p_t^h - p_t^l$ is decreasing when t is sufficiently large. Assumption 5 merely ensures that this is true for all t.

[38] Intuitively, any pooling equilibrium in which there are positive payouts $d_t^P > 0$ leaves rents to the investor and can be improved upon by setting the payouts to zero.

Figure 2.3: Separating Equilibrium

The vertical line indicates the time at which the high typeseparates. Before, the investor's belief about the project is decreasing.Once the investor learns that he is facing the high type, the beliefs $q_t$and $p_t\,(q_t)$ both jump upwards (left panel). Beforeseparation, the high type successively pledges higher shares to theinvestor. Once the high type separates, the investor's share drops and the entrepreneur's share increases (right panel).

(a) Investor's Beliefs (b) Investor's Equity Share



for the low type and

$$\Pi^h_t(0) \leq \lambda p^h_t \left(1 - \alpha^h_t\right) V - d^h_t - k + \delta \left(1 - \lambda p^h_t\right) \Pi^h_{t+1}\,(1) \qquad (\tilde{IC}_h)$$

for the high type. The first condition states that the low type's value from offering C^h_t and imitating the high type must be lower than the value from revealing her type. Specifically, once her type is discovered, the low type offers her symmetric information contract $\bar{C}^l_t = (0, \bar{\alpha}^l_t)$.[39] The second condition states that offering C^h_t must indeed be optimal for the high type. As before, once separated, the continuation values are the symmetric information values $\Pi^h_{t+1}\,(1)$ and $\Pi^l_t\,(0)$.

Separating via the equity share is costly for the high type. As we described in

[39]This is straightforward. If the low type were to offer any other contract with $\alpha^l_t > \bar{\alpha}^l_t$ or $d^l_t > 0$ which reveals her type, that contract would be suboptimal.

Section 2.5.2, to reduce the low type's payoff by one, the high type gives up a payoff of $p_t^h/p_t^l > 1$. By contrast, if she separates via the payout d_t^h, her cost of reducing the low type's payoff is one. Thus, separating via a payout is cheaper and the equity share is not distorted, i.e. $\alpha_t^h = \bar{\alpha}_t^h$. The low type's IC constraint binds in an optimal contract and the payout reduces to[40]

$$d_t^h = \Pi_t^l(1) - \Pi_t^l(0). \tag{2.15}$$

That is, the payment equals the value of imitating for the low type, which is given by the difference in continuation values at beliefs $q_t = 1$ and $q_t = 0$.

Because the cost of separating is smaller, the high type may prefer to separate rather than pooling forever. This is true when pooling is relatively costly, i.e. when q_0 is small. Intuitively, when the investor believes he is unlikely to be facing the high type, i.e. q_0 is small, the pooling equity share α_t^P is large. Then, the high type must give up a large portion of the project when she continues pooling and her value of separating is relatively large. For sufficiently small q_0, there exists periods in which the value of separating value outweighs the cost.

In equilibrium, separation is inefficiently delayed. The high type prefers to separate whenever the loss from pooling exceeds the cost of separating d_t^h. As time passes, the pooling equilibrium becomes progressively worse, and, compared to separating, the high type must pledge successively larger shares.[41] Simultaneously, the cost of separating d_t^h decreases, because the low type's project becomes less likely to succeed. After sufficient time has passed, the high type prefers to separate.

The low type, by contrast, always prefers to pool until the project is liquidated. Whenever γ, the weight on the high type's payoff, is small, the optimal contract is

[40] This follows from plugging $\alpha_t^h = \bar{\alpha}_t^h$ into Equation ($\tilde{IC}_l$).

[41] The ratio $\alpha_t^P/\bar{\alpha}_t^h$ is monotonically increasing, which implies that the high type's "adverse selection discount" becomes progressively worse.

pooling, while when the weight is large, it is separating. Figure 2.3 illustrates the dynamics of the equity share and the investor's beliefs when separation is optimal.

Interpretation When the high type separates, she pays the investor, and simultaneously reduces the investor's equity share and increases her own. This resembles performance-contingent vesting: in period τ_S, the entrepreneur takes a costly action and is rewarded with immediate vesting of shares. As Kaplan and Strömberg (2004) document, such contingent vesting schemes are common in VC contracts. Alternatively, we can interpret this early payout as a buyback. That is, the entrepreneur pays the investor to repurchase a fraction of his shares, so that her own share increases and the investor's share declines. In reality, such buybacks occur most often among later stage startups, as is the case in our model.[42]

2.6 Applied Results

Our results explain important features of venture capital contracts. First, the entrepreneur's share decreases over time, which is consistent with e.g. Kaplan and Strömberg (2003) and Kaplan and Strömberg (2004). In our model, the investor becomes more pessimistic as time passes and the entrepreneur must pledge successively larger shares to prevent the investor from leaving. Second, the optimal contract features vesting. After sufficient time has passed, the investor expects the low type to liquidate, and he becomes more optimistic over time even without observing a breakthrough. Then, the entrepreneur's share increases, which can be implemented by a delayed vesting schedule. Finally, the separating equilibrium of Proposition 14 can be implemented via a milestone with a contingent vesting clause. In period τ_S, the entrepreneur takes a costly action, and her share in the company increases. As

[42] See e.g. https://www.wsj.com/articles/ditch-the-venture-model-say-founders-who-buy-out-early-investors-to-make-a-clear-break-1531827001 for a prominent case.

Kaplan and Strömberg (2004) document, contingent vesting provisions are prominent in VC contracts.

Technological advances in recent decades have dramatically changed how VCs finance startups (see Kerr et al. (2014) and Ewens et al. (2018)). Three aspects are particularly relevant to our model. First, the cost of experimentation has declined for both entrepreneurs and investors. For example, could computing services have lowered the cost of operating IT startups by orders of magnitude. As a result, venture capital firms have adopted a “spray-and-pray” approach, now funding a large number of startups with limited vetting and oversight. Second, cohort-based accelerators (such as Y-Combinator) have increased entry by relatively inexperienced founders. Third, existing advances have made follow-up innovations easier, so that learning about startups has sped up. We now investigate how these aspects affect adverse selection in our model.

Cost of Experimenting When the entrepreneur’s cost of experimenting k decreases, both the low and high type are willing to continue longer. This increases the information rent that the high type must give up to dissuade the low type from imitating. As a result, separation becomes costlier and the high type separates later. Thus, adverse selection persists longer. We illustrate this result in Figure 2.4a, which is obtained by solving the model numerically.

To evaluate how VCs ”spray-and-pray” approach affects adverse selection, consider a reduction in the investor’s cost to experiment c. For example, the investor may exert less effort in advising the entrepreneur, making experimentation less costly for him. Again, τ_S increases and adverse selection persists longer. Intuitively, as c decreases, the optimal equity share decreases as well. This is true both when the two types pool (see Equation (2.12)) and when they separate (see Equation (2.10)). This decrease, in turn, makes both types of entrepreneur more willing to continue. Just as in the previous case, separation becomes costlier for the high type. Figure 2.4b illustrates the result.

Entrepreneur Quality Accelerators have allowed relatively inexperienced founders to start firms and to receive funding. In our model, this corresponds to a decrease in q_0, the ex-ante quality of the entrepreneur. One might expect that funding lower quality startups leads to more adverse selection. However, the effect is more subtle: as q_0 decreases, the high type separates earlier. Intuitively, as q_0 decreases, the high type's "adverse selection discount," which she suffers in the pooling equilibrium, becomes worse. The cost of separating from the low type, however, is independent of q_0. Thus, the high type separates earlier. We illustrate this result in Figure 2.4c.

Faster Learning As technology increases λ, the likelihood of breakthroughs, learning about the entrepreneur's type speeds up. However, the high type does not necessarily separate earlier. The effect is subtle, because the cost of separating d_t (in Equation (2.15)) is non-monotone in λ. The low type's expected value from imitating the high type, $\lambda p_t^l \left(\bar{\alpha}_t^l - \bar{\alpha}_t^h\right) V$, increases in both λ and t. As λ increases, this expression increases for any fixed t,[43] which increases the cost of separation. However, as λ increases, breakthroughs arrive earlier on average, which lowers the low type's expected value of imitating and therefore the cost of separating. Because of these conflicting effects, the cost of separating may increase or decrease in as λ increases. When the cost increases, the high type prefers to separate later, to reduce the cost of doing so. Figure 2.4d shows that the optimal separation time is indeed non-monotone in λ.

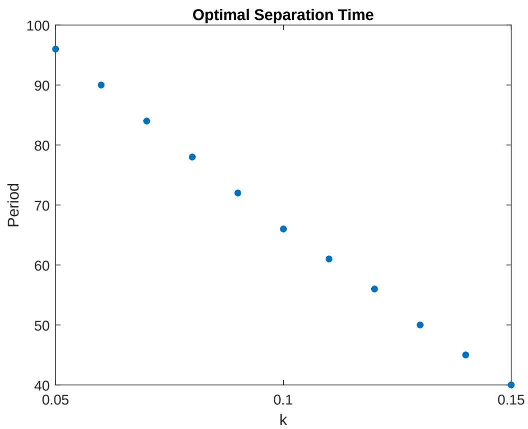
[43] Specifically, we have

$$\lambda p_t^l \left(\bar{\alpha}_t^l - \bar{\alpha}_t^h\right) V = c\left(1 - \frac{p_t^l}{p_t^h}\right).$$

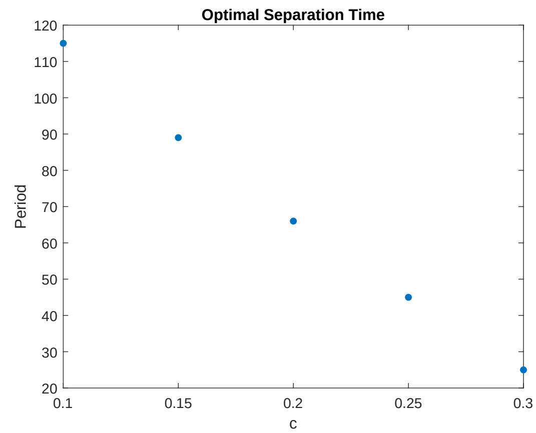
As λ increases, p_t^l/p_t^h decreases for any fixed t, because learning speeds up.

Figure 2.4: Comparative Statics

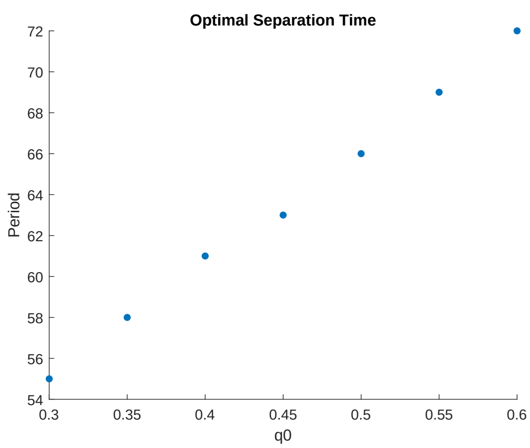
(a) As the cost of experimenting $k$decreases, the high type separates later from the low type.



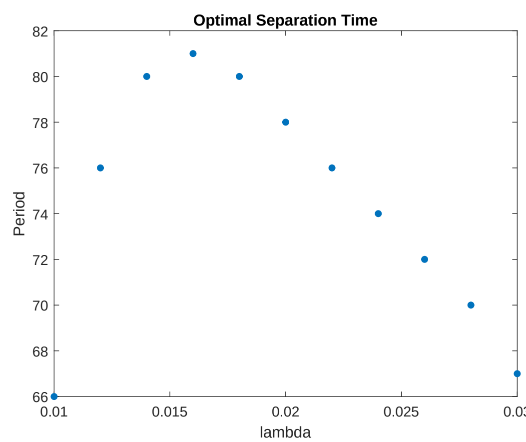
(b) As experimentation becomescheaper for the investor, the high type separates later.



(c) As the expected quality of theentrepreneur decreases, the high type separates earlier.



(d) As the likelihood ofbreakthrough increases, the high type does not necessarily sparate earlier.



Liquidation Rights To protect investors from the entrepreneur's information advantage, many VC contracts grant them favorable liquidation rights, which allow them to capture a larger part of the firm's value in bankruptcy (see Kaplan and Strömberg (2003)). However, as we show next, liquidation rights come with a downside because they delay separation.

Suppose that the firm has value $V_L < V$ if it is liquidated. Let $\alpha_L \in [0, 1]$ index the investor's liquidation rights, so that he captures value $\alpha_L V_L$, while the entrepreneur captures the remainder $(1 - \alpha_L) V_L$. Stronger liquidation rights make the investor less willing to continue, because the outside option of forcing liquidation becomes more appealing. To ensure that the investor continues, the entrepreneur must pledge a larger share, which is now given by

$$\alpha_t^P = \frac{c + \alpha_L V_L}{\lambda p_t (q_t)}$$

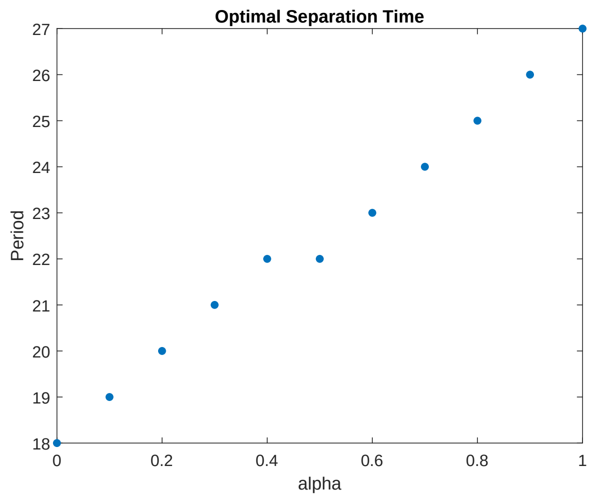
in the pooling contract and

$$\bar{\alpha}_t^\theta = \frac{c + \alpha_L V_L}{\lambda p_t^\theta}$$

once her type is revealed. This makes the low type less willing to continue, which should make separation easier to achieve. At the same time, however, increasing the investor's liquidation rights decreases the entrepreneur's value of liquidating the project, which is given by $(1 - \alpha_L) V_L$. Intuitively, as α_L increases, low type's value from imitating the high type increases, because she expects to continue the project longer. As Figure 2.5 shows, the second effect dominates. Then, separation becomes more costly as α_L increases and the high type separates later. This result is broadly consistent with Ewens et al. (2019), who estimate a matching model between VCs and entrepreneurs and find that the investor's liquidation preference reduces firm value.

Figure 2.5: Liquidation Rights

As the investor's liquidation preference increases, the high type separates later.



Overall, our comparative statics suggest that common wisdom about technology and contractual provisions may be misleading. Recent technological progress, which has speed up learning and reduced the cost of experimentation, may actually have worsened adverse selection between founders and VCs, by delaying the time at which low types are revealed. Liquidation preferences for investors may have a similar downside.

We have carried out these comparative statics in the context of Proposition 14. Suitably modified, all the results in this section carry over to the pooling equilibrium of Section 2.5.2, when the entrepreneur is cashless.[44] They do not rely on the high type being able to separate.

[44]We omit describing them in detail to save space. A description and accompanying figures are available from the authors upon request.

2.7 Signaling with Strategy

2.7.1 Pivots

Pivots are common among startups.[45] When entrepreneurs realize that their project is not succeeding, they may abandon it and pivot to a new one. In this section, we show that pivots can also signal information. For the sake of clarity, we focus on pivots as the only separating device. Thus, we assume that the entrepreneur cannot pay investors, as in Section 2.5.2.

The entrepreneur now has a real option to pivot. That is, she can abandon the current project, which is visible to investors, and start a new one at fixed cost $F > 0$. The entrepreneur's type is fixed across projects, e.g. because it represents entrepreneurial ability.[46] Thus, the new project of type θ has likelihood p_1^θ of being good. Since the high type's project is more likely to succeed, i.e. $p_1^h > p_1^l$, her value from pivoting is higher than the low type's. Thus, the fixed cost may dissuade the low type from pivoting, but not the high type, which allows the high type to separate.

In a separating equilibrium, the following IC conditions must hold. First, the low type prefers not to imitate the high type by pivoting as well, i.e.

$$\Pi_1^l(1) - F \leq \Pi_t^l(0). \tag{IC_l^{Piv}}$$

The LHS is the low type's value from pivoting. She starts a new project, so that her belief is p_1^l, and the investor believes he is facing the high type, i.e. $q = 1$. When the

[45] See the examples in the introduction.

[46] This assumption is consistent with reality. Many VC firms prioritize the quality of the founding team in their financing decisions over the particular business idea. They anticipate that the founders may change direction, but believe that the founders' quality is most important for the eventual outcomes. See Gompers et al. (2019), who find that "the management team is the most important factor VCs consider in choosing portfolio company investments" and that "ability is the most mentioned factor." See also the broader discussion in Gompers and Lerner (2001) and Kaplan et al. (2009).

low type does not pivot, she continues her initial project, which has a likelihood of p_t^l of being good, but the investor knows he is facing the low type.

The high type's IC constraint is similarly given by

$$\Pi_1^h(1) - F \geq \Pi_t^h(0), \tag{IC_h^{Piv}}$$

i.e. the high type prefers to pivot rather than continuing her initial project and being perceived as the low type.

Combining the two IC constraints, we see that when the fixed cost is too low, the low type imitates. Conversely, when the cost is too high, pivoting is not incentive compatible for high type. Overall, separating via a pivot is feasible whenever

$$F \in \left[\Pi_1^l(1) - \Pi_t^l(0), \Pi_1^h(1) - \Pi_t^h(0)\right]. \tag{2.16}$$

The interval on the RHS is non-empty for all t, because the high type's value from pivoting is higher.[47] To rule out uninteresting cases and to ensure tractability, make the following assumptions, in addition to Assumptions 3-5.

Assumption 6. *Both types pivot rather than liquidate, i.e.*

$$\Pi_1^l(0) - F \geq 0, \tag{2.17}$$

and separation is feasible, i.e.

$$F \geq \Pi_1^l(1) - \Pi_1^l(0). \tag{2.18}$$

[47] See Corollary 50 in Appendix Appendix B.2.5. This is another variant of single crossing. The value of pivoting vs. continuing and being perceived as the low type is higher for the high type, i.e.,

$$\Pi_1^l(1) - \Pi_t^l(0) \leq \Pi_1^h(1) - \Pi_t^h(0).$$

Equation (2.17) is not crucial, but allows us to reduce the number of cases we have to consider.[48] Without Equation (2.18), the low type always imitates the high type by pivoting at the same time, so that separating is not feasible.[49]

Proposition 15. *Suppose that δ and F are sufficiently small and that γ is sufficiently large. Then, the optimal contract features pooling in all periods $t < \tau_S$, and in period τ_S, the high type separates via a pivot.*

Here is the intuition. Early on, pivoting is not incentive compatible for the high type, because the value $\Pi_t^h(0)$ is relatively large. Intuitively, pivoting is not worth it when the high type is still optimistic about the project. As time passes, $\Pi_t^h(0)$ decreases and the high type eventually prefers to pivot. As long as t is not too large, the low type's IC condition holds as well, because her value from pivoting is lower.[50] As more time passes, however, the low type prefers to pay the fixed cost and imitate the high type, because her value from separating, $\Pi_t^l(0)$, becomes too small. As long as F is sufficiently small, the high type prefers to pivot before this happens.[51] Then, separating is optimal whenever γ, the weight on the high type's value, is sufficiently large.[52]

2.7.2 Prestige Projects

Many early stage firms divert resources towards prestige projects in order to generate publicity or goodwill.[53] As we show next, such prestige projects can act as signaling

[48]Without the assumption, the high and/or the low type may prefer to liquidate rather than to pivot in the first best, which yields different value functions for the equilibrium.

[49]For separating to be feasible, there also has to exist a period t in which $F \leq \Pi_1^h(1) - \Pi_t^h(0)$. Since $\Pi_t^h(0)$ is decreasing in t and eventually reaches zero, a sufficient condition is $F \leq \Pi_1^h(1)$. This condition holds because of Equation (2.17), since we have $\Pi_1^h(1) \geq \Pi_1^l(1) \geq \Pi_1^l(0)$.

[50]We can see this from the IC conditions (IC_l^{Piv}) and (IC_h^{Piv}), which satisfy a single crossing type condition.

[51]Importantly, it is possible to pick F sufficiently small without violating Equation (2.18).

[52]The assumption that δ is small allows us to simplify the derivations.

[53]For example, DoorDash, a food delivery platform, has started donating drivers' time to deliver surplus food from restaurants to nonprofits (see https://thespoon.tech/with-project-dash-doordash-

devices.

The entrepreneur can now implement a publicly observable prestige project in each period. Doing so reduces the payoff of the original project and generates a higher outside option for the entrepreneur. Specifically, when implementing the prestige project, a breakthrough yields value $V - V_0$ and the outside option is $\pi > 0$. For example, generating prestige makes it more likely that the entrepreneur can fund another startup[54] or obtain outside employment, but to do so the entrepreneur has to divert resources from her main project. The entrepreneur decides whether to implement the prestige project at the same time she offers the contract to the investor.[55] For the sake of clarity, we assume that the prestige project is the only signaling device and that the entrepreneur cannot pay the investor, as in Section 2.5.2.

The high type can use the prestige project to separate. After enough time has passed, the low type is pessimistic about the likelihood of success and her value from continuing to experiment is small. Then, she liquidates immediately upon implementing the prestige project, since taking the higher outside option is more valuable than continuing. Because of this, she cannot mimic the high type. Since the prestige project has a lower value, separating is costly and can be suboptimal. Intuitively, if both V_0 and π are very large, both types prefer to liquidate instead of continuing with the prestige project. As we show in the proposition below, for certain parameter values, separating via a prestige project is optimal.[56]

---

uses-logistics-to-rescue-over-1-million-pounds-of-surplus-food/, last accessed 10/13/19). SpaceX, an aerospace manufacturer, has a stated, if lofty, goal to colonize Mars with one million inhabitants (see `https://www.businessinsider.com/elon-musk-mars-iac-2017-transcript-slides-2017-10`, last accessed 10/13/19).

[54]This is broadly consistent with Gompers et al. (2010), who find that an entrepreneur's past success is an important factor for VC financing decisions.

[55]The particular timing is irrelevant, as long as the decision to implement the prestige project occurs before the investor's continuation decision.

[56]Broadly, V_0 cannot be too large, because then the high type never prefers to separate, but it also cannot be too small, because otherwise the low type is never dissuaded from mimicking. The range of feasible V_0 is affected by the outside option π. It is larger whenever π is smaller.

Proposition 16. *Suppose that γ is sufficiently large. Then, there exists a pair (π, V_0) such that the optimal contract features pooling in all periods $t < \tau_S$, and in period τ_S, the high type separates by implementing a prestige project and the low type liquidates.*

Formally, the following IC constraints hold when the high type separates at time t. First, the low type prefers not to implement the prestige project, i.e.,

$$\Pi_t^l(1) - \lambda p_t^l V_0 \leq \Pi_t^l(0). \qquad (IC_l^{Pres})$$

The LHS is the low type's value from imitating the high type, which consists of her continuation value $\Pi_t^l(1)$ and the loss in value if the project succeeds, while the RHS is the low type's value from separating. The high type's IC constraint is similarly given by

$$\Pi_t^h(1) - \lambda p_t^h V_0 \geq \Pi_t^h(0). \qquad (IC_h^{Pres})$$

Separating is feasible whenever

$$\frac{\Pi_t^h(1) - \Pi_t^h(0)}{p_t^h} \geq \lambda V_0 \geq \frac{\Pi_t^l(1) - \Pi_t^l(0)}{p_t^l}. \qquad (2.19)$$

A variant of the single-crossing condition in Equation (2.14) holds in this extension. Thus, the interval in Equation (2.19) is nonempty, and for each period, there exists a V_0 that separates types. When the low type liquidates upon implementing the prestige project, her IC constraint becomes

$$\Pi_t^l(1) - \lambda p_t^l V_0 \leq \pi.$$

Whenever the outside option π is sufficiently close to $\Pi_t^l(1)$, separating can be achieved relatively cheaply. That is, the loss in value V_0 which is necessary for the low type to separate is relatively small. This shows, intuitively, that there exists a pair (π, V_0) which makes separation optimal for the high type. As in the main model, the low type prefers to never separate. Thus, separating is ex-ante optimal whenever γ, the weight on the high type's value, is sufficiently large.

2.8 Equilibrium Refinements

Multiple equilibria exist in this game. The analysis in Section 2.5 focuses on Pareto optimality. Instead, in this section we adopt the D1 criterion of Cho and Kreps (1987). It specifies that the belief of the VC upon observing an off-equilibrium contract should place no weight on type-θ entrepreneur if the other type θ' has a strict incentive to deviate whenever the first type has a strict or weak incentive to deviate. We first need to modify the definition of the D1 criterion as we are working in a dynamic setup different from Cho and Kreps (1987). This requires specifying the continuation game after deviating.

To be precise, consider the entrepreneur offers a deviated contract $C_t' = (d_t', \alpha_t')$ and the VC's off-equilibrium belief is q_t'. We assume that following the deviation at t, the entrepreneur will offer an optimal contract that maximizes her continuation value given q_t'. The low type always prefers to offer the pooling contract in Proposition 12, and the high type will either offer a pooling contract or a separating contract as in Proposition 14. The choice of continuation contracts therefore is decided by q_t'.[57] Define $M^\theta(C_t') \in [0,1]$ as the set of off-equilibrium beliefs that will make the θ type weakly better off after paying C_t'. If $M^{\theta'}(C_t') \subset M^\theta(C_t')$, then the D1 criterion assumes that the VC believes facing θ with certainty upon seeing C_t'. We further define a tie-breaking rule that if $M^h(C_t') = M^l(C_t')$ then $q_t' = q_t$. In this definition, we focus on optimal contracts in continuation.[58] Doing so is sufficient for characterizing the

[57]The entrepreneur will offer the symmetric information contract if belief is degenerate.
[58]This is feasible because we require Bayesian updating both on and off the equilibrium path.

largest set of beliefs that make deviations weakly profitable.[59]

Proposition 17. *The unique equilibrium satisfying the D1 criterion is a separating contract if payouts d_t are feasible. Otherwise, the pooling contract in Proposition 12 is the unique equilibrium.*

If d_t is feasible, then the unique equilibrium features separation at a weakly earlier time than τ_S defined in the optimal separating contract of Proposition 14. Any equilibrium for which the following equation holds at some t will be pruned by the D1 criterion:

$$\Pi_t^l(1) - \Pi_t^l(q_t) < \Pi_t^h(1) - \Pi_t^h(q_t). \tag{2.20}$$

Equation (2.20) implies that relative to the equilibrium payoff, the incentive for the high type to separate is strictly larger than the incentive for the low type to imitate. This guarantees the existence of a deviating payout d_t' such that $M^l\left(d_t'\right) \subset M^h\left(d_t'\right)$ and the off-equilibrium belief is $q_t' = 1$ under the D1 criterion. Then at least the high type will deviate. In the proof, we show that there exists a $\tau_S^\star$ such that Equation (2.20) always holds for any $t > \tau_S^\star$, in all equilibria featuring separation later than $\tau_S^\star$. This is because when the low type becomes pessimistic at later periods, she is close to liquidate even with offering $\bar{\alpha}_t^h$. Therefore her imitation incentive is very low. But at the same period, the high type still has gains from separation because he persists strictly longer when offering the symmetric information contract.

Recall that in Proposition 14, separation is inefficiently delayed. For the high type, the cost of separating becomes relatively cheaper compared to the loss of pooling at

[59]If the entrepreneur is weakly better off with some suboptimal continuation contract, then she must be strictly better with an optimal contract given the same q_t'. However, if she is just indifferent with an optimal continuation contract given q_t', then she will become strictly worse off with suboptimal contacts.

later periods. In other words, she prefers to separate only if

$$\Pi_t^l(1) - \Pi_t^l(0) < \Pi_t^h(1) - \Pi_t^h(q_t). \tag{2.21}$$

Since $\Pi_t^l(q_t) \geq \Pi_t^l(0)$,[60] the optimal separating contract may be pruned. As a result the D1-refined equilibrium separates types weakly earlier at $\tau_S^\star \leq \tau_S$.

If entrepreneurs are cashless and only equity payments are feasible, we need to normalize the incentives by the probability of paying the deviating contract, which leads to a variant of Equation (2.20):

$$\frac{\Pi_t^l(1) - \Pi_t^l(q_t)}{p_t^l} < \frac{\Pi_t^h(1) - \Pi_t^h(q_t)}{p_t^h}. \tag{2.22}$$

This never holds for the optimal pooling contract since the marginal cost of separation increases to $p_t^h/p_t^l > 1$. Therefore it survives the D1 criterion. For the same reason, the separating contract in Proposition 13 is pruned. At the separating period, the entrepreneur will deviate to a smaller equity contract. Since the high type has a strictly larger expected payment reduction, the off-equilibrium belief is $q_t' = 1$, which makes the deviation profitable for both types.

Lastly, we discuss the necessity of using the D1 criterion instead of the intuitive criterion in Cho and Kreps (1987). The definition of the intuitive criterion in our setting is as follows. Suppose there exists a deviating contract C_t' such that (i) the low is strictly worse off even if she is regarded as the high type, i.e. $M^l(C_t') = \emptyset$, and (ii) the high type is strictly better off if the VC knows she is not the low type, i.e $\{1\} \subset M^h(C_t')$, then the VC believes that the deviating entrepreneur is the high type.[61] Proving Proposition 17 takes two steps. First, we need to prune all other equilibria for uniqueness. Following the intuitive criterion is equivalent to using the

[60] The inequality is strict for all pooling equilibria at $t \leq \underline{\tau}_l$ and all separating equilibria at $t \leq \tau_S$.

[61] The case when the VC believes facing the low type is interchangeably defined by switching the above types.

D1 criterion in this step. This is because we can find a deviating contract satisfying both (i) and (ii) if Equation (2.20) holds.[62] Second, we need to verify that no deviation eliminates the unique equilibrium for existence. Since Equation (2.20) does not hold at any t for this equilibrium, the following deviating payout exists

$$d'_t < \Pi^h_t(1) - \Pi^h_t(q_t) < \Pi^l_t(1) - \Pi^l_t(q_t).$$

Given this deviation, the D1 criterion can pin down the off-equilibrium to be $q'_t = 0$, asserting that the unique equilibrium is not pruned. However, the low type will be strictly better off if she is regarded as the high type after the same deviation. This violates (i) and the off-equilibrium belief is not defined under the intuitive criterion. In this case, we cannot rigorously check whether a profitable deviation exists for the unique equilibrium. This is why we have to rely on the D1 criterion.

2.9 Conclusion

In this paper, we model startup financing a contracting problem with private information, in which both the entrepreneur and investor learn about the startup over time. We recover common features of VC contracts, such as dilution, vesting, and buyouts, as equilibrium outcomes.

In the literature, the dominant explanation for pivots is that entrepreneurs realize that their project is unlikely to succeed and therefore start a new one. Our model nests this explanation and provides a new one: pivots can act as signaling devices. This, to our knowledge, has not been recognized in the literature on venture capital.

Our comparative statics reveal that the interaction between technological progress and adverse selection is a subtle one. Technological progress is commonly believed to

[62] The deviating payout is $d'_t = \Pi^l_t(1) - \Pi^l_t(q_t) + \varepsilon$, where ε is a sufficiently small positive number.

improve learning by entrepreneurs and investors alike and therefore to reduce adverse selection. As we show, this is not necessarily the case. Instead, recent developments in VC financing may have allowed adverse selection to persist longer. These predictions would be interesting to test. Our model provides an empirical link for such a test, since the rate of low types exiting is a measure for the persistence of adverse selection.

References

Admati, A. R. and Pfleiderer, P. (1994). Robust financial contracting and the role of venture capitalists. *Journal of Finance*, 49(2):371–402.

Aghion, P., Bolton, P., and Tirole, J. (2004). Exit options in corporate finance: Liquidity versus incentives. *Review of Finance*, 8(3):327–353.

Aghion, P. and Tirole, J. (1994). The management of innovation. *Quarterly Journal of Economics*, 109(4):1185–1209.

Asker, J., Farre-Mensa, J., and Ljungqvist, A. (2011). Comparing the investment behavior of public and private firms. Technical report, National Bureau of Economic Research.

Axelson, U. (2007). Security design with investor private information. *Journal of Finance*, 62(6):2587–2632.

Azarmsa, E. and Cong, L. W. (2018). Persuasion in relationship finance. *Becker Friedman Institute for Research in Economics Working Paper*, (2017-16).

Baker, M. and Wurgler, J. (2000). The Equity Share in New Issues and Aggregate Stock Returns. *The Journal of Finance*, 55(5):2219–2257.

Ball, E., Chiu, H. H., and Smith, R. (2011). Can VCs Time the Market? An Analysis of Exit Choice for Venture-backed Firms. *Review of Financial Studies*, 24(9):3105–3138.

Barrot, J.-N. (2016). Investor Horizon and the Life Cycle of Innovative Firms: Evidence from Venture Capital. *Management Science*, 63(9):3021–3043.

Bayar, O. and Chemmanur, T. J. (2012). IPOs versus Acquisitions and the Valuation Premium Puzzle: A Theory of Exit Choice by Entrepreneurs and Venture Capitalists. *Journal of Financial and Quantitative Analysis*, 46(06):1755–1793.

Bena, J. and Li, K. (2014). Corporate Innovations and Mergers and Acquisitions. *The Journal of Finance*, 69(5):1923–1960.

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

Bergemann, D. and Hege, U. (2005). The financing of innovation: Learning and stopping. *RAND Journal of Economics*, 36(4):719–752.

Bernstein, S. (2015). Does going public affect innovation? *The Journal of Finance*, 70(4):1365–1403.

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

Bouvard, M. (2012). Real option financing under asymmetric information. *Review of Financial Studies*, 27(1):180–210.

Brau, J. C., Couch, R. B., and Sutton, N. K. (2012). The Desire to Acquire and IPO Long-Run Underperformance. *Journal of Financial and Quantitative Analysis*, 47(3):493–510.

Brau, J. C., Francis, B., and Kohers, N. (2003). The Choice of IPO versus Takeover: Empirical Evidence. *The Journal of Business*, 76(4):583–612.

Brown, K. C., Dittmar, A., and Servaes, H. (2005). Corporate Governance, Incentives, and Industry Consolidations. *Review of Financial Studies*, 18(1).

Budish, E., Roin, B. N., and Williams, H. (2015). Do firms underinvest in long-term research? evidence from cancer clinical trials. *American Economic Review*, 105(7):2044–85.

Burkart, M., Gromb, D., and Panunzi, F. (1997). Large shareholders, monitoring, and the value of the firm. *Quarterly Journal of Economics*, 112(3):693–728.

Casamatta, C. (2003). Financing and advising: Optimal financial contracts with venture capitalists. *Journal of Finance*, 58(5):2059–2085.

Celikyurt, U., Sevilir, M., and Shivdasani, A. (2010). Going Public to Acquire? the Acquisition Motive in IPOs. *Journal of Financial Economics*, 96(3):345–363.

Chiu, J. and Koeppl, T. V. (2016). Trading Dynamics with Adverse Selection and Search: Market Freeze, Intervention and Recovery. *The Review of Economic Studies*, 83(3):969–1000.

Cho, I.-K. and Kreps, D. M. (1987). Signaling games and stable equilibria. *Quarterly Journal of Economics*, 102(2):179–221.

Cho, I.-K. and Sobel, J. (1990). Strategic Stability and Uniqueness in Signaling Games. *Journal of Economic Theory*, 50(2):381–413.

Cunningham, C., Ederer, F., and Ma, S. (2018). Killer Acquisitions. *Available at SSRN*.

Daley, B. and Green, B. (2012). Waiting for News in the Market for Lemons. *Econometrica*, 80(4):1433–1504.

Daley, B. and Green, B. (2016). An Information-based Theory of Time-Varying Liquidity. *The Journal of Finance*, 71(2):809–870.

Dessein, W. (2005). Information and control in ventures and alliances. *Journal of Finance*, 60(5):2513–2549.

Devos, E., Kadapakkam, P.-R., and Krishnamurthy, S. (2008). How Do Mergers Create Value? A Comparison of Taxes, Market Power, and Efficiency Improvements as Explanations for Synergies. *The Review of Financial Studies*, 22(3):1179–1211.

Dixit, A. and Pindyck, R. (1994). *Investment under Uncertainty*. Princeton University Press.

Doidge, C., Karolyi, G. A., and Stulz, R. M. (2017). The us Listing Gap. *Journal of Financial Economics*, 123(3):464–487.

Dong, M. (2016). Strategic Experimentation with Asymmetric Information. *Working Paper*.

Eisfeldt, A. L., Falato, A., and Xiaolan, M. Z. (2018). Human Capitalists. *Working Paper*.

Ewens, M., Gorbenko, A. S., and Korteweg, A. (2019). Venture capital contracts. Technical report, National Bureau of Economic Research.

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

Fahn, M. and Klein, N. A. (2017). Relational contracts with private information on the future value of the relationship: The upside of implicit downsizing costs. *Working Paper*.

Fishman, M. J. and Parker, J. A. (2015). Valuation, Adverse Selection, and Market Collapses. *The Review of Financial Studies*, 28(9):2575–2607.

Fuchs, W., Green, B., and Papanikolaou, D. (2016). Adverse Selection, Slow-Moving Capital, and Misallocation. *Journal of Financial Economics*, 120(2):286–308.

Gale, D. and Hellwig, M. (1985). Incentive-compatible debt contracts: The one-period problem. *Review of Economic Studies*, 52(4):647–663.

Gao, H., Harford, J., and Li, K. (2013). Determinants of corporate cash policy: Insights from private firms. *Journal of Financial Economics*, 109(3):623–639.

Gladstone, D. and Gladstone, L. (2002). *Venture capital handbook: an entrepreneur's guide to raising venture capital*. Prentice Hall, NJ.

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

Gompers, P. A. (1996). Grandstanding in the Venture Capital Industry. *Journal of Financial Economics*, 42(1):133–156.

Gompers, P. A., Gornall, W., Kaplan, S. N., and Strebulaev, I. A. (2019). How do venture capitalists make decisions? *Journal of Financial Economics*.

Gompers, P. A., Kovner, A., Lerner, J., and Scharfstein, D. (2010). Performance persistence in entrepreneurship. *Journal of Financial Economics*, 96(1):18–32.

Gompers, P. A. and Lerner, J. (2001). *The money of invention: How venture capital creates new wealth*. Harvard Business School Press, Boston, MA.

Gompers, P. A. and Lerner, J. (2004). *The venture capital cycle*. MIT press, Cambridge, Massachusets.

Gorbenko, A. S. and Malenko, A. (2017). The Timing and Method of Payment in Mergers when Acquirers are Financially Constrained. *The Review of Financial Studies*, 31(10):3937–3978.

Gorman, M. and Sahlman, W. A. (1989). What do venture capitalists do? *Journal of Business Venturing*, 4(4):231–248.

Grenadier, S. R. and Malenko, A. (2011). Real Options Signaling Games with Applications to Corporate Finance. *The Review of Financial Studies*, 24(12):3993–4036.

Grenadier, S. R., Malenko, A., and Strebulaev, I. A. (2014). Investment Busts, Reputation, and the Temptation to Blend in with the Crowd. *Journal of Financial Economics*, 111(1):137–157.

Gul, F. and Pesendorfer, W. (2012). The War of Information. *The Review of Economic Studies*, 79(2):707.

Halac, M. (2012). Relational contracts and the value of relationships. *American Economic Review*, 102(2):750–79.

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

Hand, J. R. (2008). Give everyone a prize? Employee stock options in private venture-backed firms. *Journal of Business Venturing*, 23(4):385–404.

Hart, O. and Moore, J. (1998). Default and renegotiation: A dynamic model of debt. *Quarterly Journal of Economics*, 113(1):1–41.

Hege, U. and Hennessy, C. (2010). Acquisition Values and Optimal Financial (in) Flexibility. *The Review of Financial Studies*, 23(7):2865–2899.

Hellmann, T. (2006). IPOs, acquisitions, and the use of convertible securities in venture capital. *Journal of Financial Economics*, 81(3):649–679.

Hellmann, T. and Puri, M. (2000). The interaction between product market and financing strategy: The role of venture capital. *Review of Financial Studies*, 13(4):959–984.

Higgins, M. J. and Rodriguez, D. (2006). The Outsourcing of R&D Through Acquisitions in the Pharmaceutical Industry. *Journal of Financial Economics*, 80(2):351–383.

Inderst, R. and Vladimirov, V. (2019). Growth firms and relationship finance: A capital structure perspective. *Management Science*.

Innes, R. D. (1990). Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory*, 52(1):45–67.

Jensen, M. C. (1986). Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. *The American Economic Review*, 76(2):323–329.

Jovanovic, B. and Rousseau, P. L. (2002). The Q-Theory of Mergers. *American Economic Review*, 92(2):198–204.

Kaniel, R. and Orlov, D. (2018). Family knows best: Fund advisors as talent-rating agencies. *Working Paper*.

Kaplan, S. N., Sensoy, B. A., and Strömberg, P. (2009). Should investors bet on the jockey or the horse? Evidence from the evolution of firms from early business plans to public companies. *The Journal of Finance*, 64(1):75–115.

Kaplan, S. N. and Strömberg, P. (2001). Venture capitals as principals: Contracting, screening, and monitoring. *American Economic Review*, 91(2):426–430.

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

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

Kartal, M. (2018). Honest equilibria in reputation games: The role of time preferences. *American Economic Journal: Microeconomics*, 10(1):278–314.

Keller, G., Rady, S., and Cripps, M. (2005). Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68.

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

Kim, E. H. and Singal, V. (1993). Mergers and Market Power: Evidence from the Airline Industry. *The American Economic Review*, pages 549–569.

Kortum, S. S. and Lerner, J. (2000). Assessing the impact of venture capital on innovation. *RAND Journal of Economics*, 31(4):674–692.

Krieger, J., Li, X., and Thakor, R. T. (2018). Find and Replace: R&D Investment Following the Erosion of Existing Products. *Available at SSRN.*

Lambrecht, B. M. (2004). The Timing and Terms of Mergers Motivated by Economies of Scale. *Journal of Financial Economics*, 72(1):41–62.

Lerner, J. (1995). Venture capitalists and the oversight of private firms. *Journal of Finance*, 50(1):301–318.

Levine, O. (2017). Acquiring Growth. *Journal of Financial Economics*, 126(2):300–319.

Li, X. (2013). Productivity, Restructuring, and the Gains from Takeovers. *Journal of Financial Economics*, 109(1):250–271.

Loughran, T. and McDonald, B. (2013). IPO First-day Returns, Offer Price Revisions, Volatility, and Form S-1 Language. *Journal of Financial Economics*, 109(2):307–326.

Loughran, T. and Ritter, J. (2004). Why Has IPO Underpricing Changed over Time? *Financial Management*, 33(3):5–37.

Lucas, R. E. (1978). On the Size Distribution of Business Firms. *The Bell Journal of Economics*, pages 508–523.

Maksimovic, V. and Phillips, G. (2001). The Market for Corporate Assets: Who Engages in Mergers and Asset Sales and Are There Efficiency Gains? *The Journal of Finance*, 56(6):2019–2065.

Maksimovic, V. and Phillips, G. (2002). Do Conglomerate Firms Allocate Resources inefficiently Across Industries? Theory and Evidence. *The Journal of Finance*, 57(2):721–767.

Maksimovic, V., Phillips, G., and Prabhala, N. R. (2011). Post-Merger Restructuring and the Boundaries of the Firm. *Journal of Financial Economics*, 102(2):317–343.

Maksimovic, V., Phillips, G., and Yang, L. (2013). Private and Public Merger Waves. *The Journal of Finance*, 68(5):2177–2217.

Malcomson, J. M. (2016). Relational incentive contracts with persistent private information. *Econometrica*, 84(1):317–346.

Mann, W. (2018). Creditor rights and innovation: Evidence from patent collateral. *Journal of Financial Economics*, 130(1):25–47.

Netter, J., Stegemoller, M., and Wintoki, M. B. (2011). Implications of Data Screens on Merger and Acquisition Analysis: A Large Sample Study of Mergers and Acquisitions from 1992 to 2009. *The Review of Financial Studies*, 24(7):2316–2357.

Osborne, M. J. and Rubinstein, A. (1990). *Bargaining and markets*. Academic press.

Piacentino, G. (2019). Venture capital and capital allocation. *Journal of Finance*, 74(3):1261–1314.

Poulsen, A. B. and Stegemoller, M. (2008). Moving from Private to Public Ownership: Selling out to Public Firms versus Initial Public Offerings. *Financial Management*, 37(1):81–101.

Rajan, R. G. (1992). Insiders and outsiders: The choice between informed and arm's-length debt. *Journal of Finance*, 47(4):1367–1400.

Repullo, R. and Suarez, J. (2004). Venture capital finance: A security design approach. *Review of Finance*, 8(1):75–108.

Rhodes-Kropf, M. and Robinson, D. T. (2008). The Market for Mergers and the Boundaries of the Firm. *The Journal of Finance*, 63(3):1169–1211.

Ritter, J. R. (1991). The Long-run Performance of Initial Public Offerings. *The Journal of Finance*, 46(1):3–27.

Ritter, J. R. (2015). Growth Capital-Backed IPOs. *Financial Review*, 50(4):481–515.

Robb, A. M. and Robinson, D. T. (2014). The capital structure decisions of new firms. *Review of Financial Studies*, 27(1):153–179.

Roll, R. (1986). The Hubris Hypothesis of Corporate Takeovers. *Journal of Business*, pages 197–216.

Sahlman, W. A. (1990). The structure and governance of venture-capital organizations. *Journal of Financial Economics*, 27(2):473–521.

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

Schmidt, K. M. (2003). Convertible securities and venture capital finance. *Journal of Finance*, 58(3):1139–1166.

Schultz, P. (2003). Pseudo Market Timing and the Long-run Underperformance of IPOs. *The Journal of Finance*, 58(2):483–518.

Sheen, A. (2014). The Real Product Market Impact of Mergers. *The Journal of Finance*, 69(6):2651–2688.

Shleifer, A. and Vishny, R. W. (2003). Stock Market Driven Acquisitions. *Journal of Financial Economics*, 70(3):295–311.

Strebulaev, I. A., Zhu, H., and Zryumov, P. (2016). Optimal Issuance under Information Asymmetry and Accumulation of Cash Flows. *Rock Center for Corporate Governance at Stanford University Working Paper No. 164*.

Thomas, C. (2019). Experimentation with reputation concerns – Dynamic signalling with changing types. *Journal of Economic Theory*, 179:366–415.

Ueda, M. (2004). Banks versus venture capital: Project evaluation, screening, and expropriation. *Journal of Finance*, 59(2):601–621.

Wasserman, N. (2003). Founder-CEO succession and the paradox of entrepreneurial success. *Organization Science*, 14(2):149–172.

Zryumov, P. (2015). Dynamic Adverse Selection: Time-Varying Market Conditions and Endogenous Entry. *Available at SSRN*.

Appendix A

Appendix for Chapter 1

A.1 Tables

Table A.1: **Summary Statistics**

This table provides a summary statistics of main variables. *Offer price* is the dollar price per share at IPO. *Proceeds* is the total proceeds from IPO. *Total Assets* is total assets of IPO company before issuance. *Age* is the year time between IPO year and founding year. VC is an indicating dummy that takes 1 if company is VC-funded. $Hightech$ indicates whether the company belongs to high tech industry following Loughran and Ritter (2004). $Bookrunners$ is the number of lead managers. $Nasdaq$ indicates whether the firm is listed at the Nasdaq Stock Market.

	(1) **Mean**	**(2)** **Std**	**(3)** **Min**	**(4)** **Max**	**(5)** **N**
Panel A: Full Sample					
Offer Price	14.11	6.02	5.00	85.00	1537.00
Proceeds (Mil.)	207.05	634.08	3.50	16006.90	1537.00
Total Assets (Mil.)	854.04	4744.24	0.20	137238.00	1475.00
Age	19.44	23.41	1.00	166.00	1530.00
VC	0.48	0.50	0.00	1.00	1537.00
High Tech (dummy)	0.31	0.46	0.00	1.00	1537.00
Bookrunners	2.68	1.73	1.00	13.00	1537.00
Nasdaq (dummy)	0.66	0.47	0.00	1.00	1537.00
Panel B: $MA = 1$					
Offer Price	14.93	6.93	5.00	85.00	524.00
Proceeds (Mil.)	297.28	1033.66	7.00	16006.90	524.00
Total Assets (Mil.)	1508.97	7803.03	0.20	137238.00	503.00
Age	22.27	25.47	1.00	166.00	522.00
VC	0.45	0.50	0.00	1.00	524.00
High Tech (dummy)	0.38	0.49	0.00	1.00	524.00
Bookrunners	2.97	1.87	1.00	13.00	524.00
Nasdaq (dummy)	0.54	0.50	0.00	1.00	524.00

Table A.2: **First-day and Long-run Returns on IPOs Categorized by M&A Before**

This table provides results on average first-day return, buy-and-hold returns and return of assets categorized by whether the IPO firm has acquisitions before. *First-Day Return* is defined as the return of closing price in the first trading day over offer price. *3yr Return* is the cumulative return of holding this IPO from its first trading day for three years, excluding first-day return. *3yr Adj. Return* adjusts *3yr Return* by subtracting the cumulative return, during the same holding period, of a corresponding Fama–French Size and Book-to-Market (2 × 3) portfolios from it. *3yr ROA* is the ROA three years after the IPO time. *4yr Return*, *4yr Adj. Return* and *4yr ROA* are defined similarly with a 4-year window. The first three columns are results of the full sample and the last three columns are results of the IPOs in high technology industries. Except *IPO Num*, all numbers are percentage points. Columns (1) and (4) show the result of IPOs without private acquisitions. Columns (2) and (5) show the results of IPOs with private acquisitions. Columns (3) and (6) show the results of IPOs with private acquisitions of a competitor. t statistics are displayed in the parentheses.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	All Industry			**High Tech**		
	w/o Acq	**w/ acq**	**w/ acqcom**	**w/o Acq**	**w/ acq**	**w/ acqcom**
IPO Num	1,012	524	339	282	198	118
First-Day Return	14.724***	13.078***	12.186***	18.508***	18.773***	17.825***
	(17.172)	(13.394)	(10.900)	(10.641)	(10.579)	(8.346)
3yr ROA	-19.600***	-1.489	-0.289	-15.621***	-1.977	-2.989
	(-9.338)	(-0.927)	(-0.127)	(-3.818)	(-1.068)	(-1.039)
4yr ROA	-25.604***	-3.373	-0.769	-16.580***	-2.238	-1.002
	(-6.576)	(-1.611)	(-0.280)	(-3.867)	(-0.945)	(-0.498)
3yr Return	16.991***	25.061***	28.707***	17.942***	31.938***	30.888***
	(3.971)	(5.286)	(4.956)	(2.833)	(4.057)	(3.120)
3yr Adj. Return	-5.208	-0.020	2.750	-0.988	5.877	3.098
	(-1.243)	(-0.004)	(0.498)	(-0.161)	(0.773)	(0.321)
4yr Return	24.101***	32.749***	35.183***	32.336***	43.832***	43.156***
	(4.766)	(5.775)	(5.231)	(4.162)	(4.522)	(3.550)
4yr Adj. Return	-5.419	0.284	1.345	6.602	10.056	6.987
	(-1.083)	(0.052)	(0.208)	(0.869)	(1.065)	(0.587)

Table A.3: **Regressions of First-day Returns on Private Acquisition Indicators**

This table provides estimations on the following equation:

$$Firstret_{i,t} = \beta MA + \gamma Controls + \eta_t + \mu_j + \epsilon_i, t$$

$Firstret_{i,t}$ is defined as the return of closing price in the first trading day over offer price. MA is an indicating dummy, equal 1 if i has made private acquisitions before IPO and 0 otherwise. MA^{comp} is a dummy variable indicating whether i has acquired a competitor before IPO. Control variables include $ln(1+TA)$, $ln(1+age)$, VC, $Hightech$, $Bookrunners$ and $Nasdaq$. Robust standard errors are in parentheses. A constant term is included in all regressions (not reported). *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	(1) *Firstret*	**(2)** *Firstret*	**(3)** *Firstret*
Panel A: Private Acquisitions			
MA	-2.945*	-2.661**	-3.136*
	(-1.824)	(-2.060)	(-1.945)
Controls	N	Y	Y
Year Fixed Effects	Y	Y	Y
Industry Fixed Effects	Y	N	Y
#Obs	1391	1467	1323
Adj. R^2	0.06	0.04	0.09
Panel B: Private Acquisitions of Competitors			
MA^{comp}	-3.554**	-3.017**	-3.725**
	(-2.147)	(-2.238)	(-2.244)
Controls	N	Y	Y
Year Fixed Effects	Y	Y	Y
Industry Fixed Effects	Y	N	Y
#Obs	1391	1467	1323
Adj. R^2	0.06	0.04	0.09

Table A.4: **Regressions of First-day Returns on Private Acquisition Indicators**

This table provides estimations on the following equation:

$$ROA_{i,t} = \beta MA + \gamma Controls + \eta_t + \mu_j + \epsilon_i, t$$

$ROA_{i,t}$ is defined as the return on assets 3 years (Columns 1–3) or 4 years (Columns 4–6) after the firm go public. MA is an indicating dummy, equal 1 if i has made private acquisitions before IPO and 0 otherwise. MA^{comp} is a dummy variable indicating whether i has acquired a competitor before IPO. Control variables include $ln(1+TA)$, $ln(1+age)$, VC, $Hightech$, $Bookrunners$ and $Nasdaq$. Robust standard errors are in parentheses. A constant term is included in all regressions (not reported). *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	(1) *3yr ROA*	**(2)** *3yr ROA*	**(3)** *3yr ROA*	**(4)** *4yr ROA*	**(5)** *4yr ROA*	**(6)** *4yr ROA*
Panel A: Private Acquisitions						
MA	11.884***	5.022***	5.503**	20.850**	4.526*	8.476**
	(3.609)	(2.701)	(2.055)	(2.532)	(1.646)	(2.296)
Controls	N	Y	Y	N	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Industry Fixed Effects	Y	N	Y	Y	N	Y
#Obs	1373	1451	1306	1372	1450	1305
Adj. R^2	0.14	0.22	0.23	-0.00	0.11	0.08
Panel B: Private Acquisitions of Competitors						
MA^{comp}	10.769***	1.141	1.914	22.843***	2.457	7.867**
	(3.439)	(0.557)	(0.738)	(2.598)	(0.832)	(2.011)
Controls	N	Y	Y	N	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Industry Fixed Effects	Y	N	Y	Y	N	Y
#Obs	1373	1451	1306	1372	1450	1305
Adj. R^2	0.14	0.22	0.23	0.00	0.11	0.08

A.2 Omitted Equations in Section 1.3.3

Using Ito's lemma, the Bellman Equation for the two companies are:

$$
\begin{aligned}
rV_h(x_{ht}, \rho_{lt}) = & \underbrace{\mu_h x_{ht} \frac{\partial V_h(x_{ht}, \rho_{lt})}{\partial x} + \frac{1}{2}\frac{\sigma^2}{2} x_{ht}^2 \frac{\partial V_h^2(x_{ht}, \rho_{lt})}{\partial x^2}}_{Size\ Effect} \\
& + \underbrace{\frac{-(\mu_h - \mu_l)^2}{\sigma^2} \frac{\partial V_h(x_{ht}, \rho_{lt})}{\partial \rho} + \frac{1}{2}\frac{(\mu_h - \mu_l)^2}{\sigma^2} \frac{\partial V_h^2(x_{ht}, \rho_{lt})}{\partial \rho^2}}_{Belief\ Effect} \\
& \underbrace{-\frac{1}{2}(\mu_h - \mu_l) x_{ht} \frac{\partial V_h^2(x_{ht}, \rho_{lt})}{\partial x \partial \rho}}_{Cross\ Effect}
\end{aligned} \tag{A.1}
$$

$$
\begin{aligned}
rV_l(x_{lt}, \rho_{lt}) = & \underbrace{\mu_l x_{lt} \frac{\partial V_l(x_{lt}, \rho_{lt})}{\partial x} + \frac{1}{2}\frac{\sigma^2}{2} x_{lt}^2 \frac{\partial V_l^2(x_{lt}, \rho_{lt})}{\partial x^2})}_{Size\ Effect} \\
& + \underbrace{\frac{-(\mu_h - \mu_l)^2}{\sigma^2} \frac{\partial V_l(x_{lt}, \rho_{lt})}{\partial \rho} + \frac{1}{2}\frac{(\mu_h - \mu_l)^2}{\sigma^2} \frac{\partial V_l^2(x_{lt}, \rho_{lt})}{\partial \rho^2}}_{Belief\ Effect} \\
& + \underbrace{\frac{1}{2}(\mu_h - \mu_l) x_{lt} \frac{\partial V_l^2(x_{lt}, \rho_{lt})}{\partial x \partial \rho}}_{Cross\ Effect}.
\end{aligned} \tag{A.2}
$$

When game ends, firms receive their payoffs correspondingly:

$$V_h(x_{h\tau}, \eta) = (H - \alpha - \frac{\alpha(H-L)exp(\eta)}{(H + Lexp(\eta))})x_{h\tau} \tag{A.3}$$

$$V_l(x_{l\tau}, \eta) = (L - \alpha + \frac{\alpha(H-L)exp(\eta)}{(Hexp(\eta) + L)})x_{l\tau} \tag{A.4}$$

$$V_h(x_{h\tau}, \beta) = (H - a)x_{h\tau} + (H - L - \gamma_l)x_{l\tau} \tag{A.5}$$

$$V_l(x_{l\tau}, \beta) = (L - a + \gamma_l)x_{l\tau}. \tag{A.6}$$

The two smooth pasting conditions are

$$\frac{\partial V_l(x_{ht},\beta)}{\partial \beta} = 0 \tag{A.7}$$

$$\frac{\partial V_h(x_{ht},\eta)}{\partial \eta} = \frac{-\alpha(H-L)Hexp(\eta)}{(H+Lexp(\eta))^2}x_{h\tau}. \tag{A.8}$$

The last step is to transform all the boundary conditions in forms of J_i:

$$J_h(\eta) = H - \alpha - \frac{\alpha(H-L)exp(\eta)}{(H+Lexp(\eta))}$$

$$J_l(\eta) = L - \alpha + \frac{\alpha(H-L)exp(\eta)}{(Hexp(\eta)+L)}$$

$$J_h(\beta) = H - a + (H - L - \gamma_l)e^{\beta(\frac{\mu_h-\mu_l}{\sigma^2})^{-1}}$$

$$J_l(\beta) = L - a + \gamma_l$$

$$J_h'(\eta) = -\frac{\alpha H(H-L)exp(\eta)}{(H+Lexp(\eta))^2}$$

$$J_l'(\beta) = 0.$$

Combine the above boundary conditions with equation (1.4) and (1.4) and solve (β,η).

A.3 Proofs

Proof of Lemma 1

Proof. Fix a pair (x_1, x_2). First, there exists no equilibria in which

$$\sigma_{ht}^{A}\left(x_1, x_2\right)\sigma_{lt}^{T}\left(x_1, x_2\right) > 0,$$

and

$$\sigma_{lt}^{A}\left(x_1, x_2\right)\sigma_{ht}^{T}\left(x_1, x_2\right) > 0.$$

If so, both types must be indifferent between being an acquirer or a target, which implies their total payoffs are the same with different acquiring types. This implies,

$$\underbrace{H(x_1 + x_2)(1 - s_t)}_{h\text{-acquirer total net value}} = \underbrace{L(x_1 + x_2)(1 - s_t)}_{l\text{-acquirer total net value}}.$$

Since H¿L, the above equation is only valid when $s_t = 1$. But then both players have 0 payoffs, which is contradictory.

Second, the following type of equilibria cannot exist:

$$\sigma_{ht}^{A}\left(x_1, x_2\right)\sigma_{lt}^{T}\left(x_1, x_2\right) > 0, \quad \sigma_{lt}^{A}\left(x_1, x_2\right)\sigma_{ht}^{T}\left(x_1, x_2\right) = 0,$$
$$\sigma_{ht}^{A}\left(x_2, x_1\right)\sigma_{lt}^{T}\left(x_2, x_1\right) = 0, \quad \sigma_{lt}^{A}\left(x_2, x_1\right)\sigma_{ht}^{T}\left(x_2, x_1\right) > 0.$$

In other words, if the high type's size is x_1 and low type's size is x_2, then the high type acquires. On the contrary, if the high type's size is x_2 and low type's size is x_1, then the low type acquires. Firm l's payoff at (x_2, x_1) is $(1 - s_t)\, L\,(x_1 + x_2) - (H - \alpha + \gamma_h)\, x_2$.

The following deviation is strictly profitable:

$$\tilde{\sigma}_{ht}^{A}\left(x_2, x_1\right)\tilde{\sigma}_{lt}^{T}\left(x_2, x_1\right) = 1, \quad \tilde{\sigma}_{lt}^{A}\left(x_2, x_1\right)\tilde{\sigma}_{ht}^{T}\left(x_2, x_1\right) = 0,$$

while the high type still gets $\left(H - \alpha + \gamma_h\right) x_1'$.[1] Notice this deviation will not affect investors belief and therefore s_t. This deviation generates $\left(1 - s_t\right) H \left(x_1 + x_2\right) - \left(H - \alpha + \gamma_h\right) x_2$ for firm l.

Lastly, it is intuitive that there exits no equilibria in which only the low-quality firm acquires. □

Proof of Lemma 2

Proof. Here I only illustrate for firm h's problem since the same logic applies to firm l's. Suppose the initial sizes are x_{ht} and x_{lt}. Now suppose we change firm h's assets size without affecting the investors belief upon pooling. This is done as follows. Given any non-zero $\alpha > 0$, let the new firm sizes to be $x_{ht}' = \alpha x_{ht}$. By equation (1.3), firm l must be adjusted to $x_{lt}' = \alpha x_{lt}$. Consider the new firm h's problem:

$$\begin{aligned}
&\sup_{\sigma^i} \quad \mathbb{E}^h \left(\int_t^\tau e^{-r\tau} \left(R_i^I \left(x_{h\tau}', x_{l\tau}'\right) \mathbb{1}_\tau^I + R_i^m \left(x_{h\tau}', x_{l\tau}'\right) \mathbb{1}_\tau^m \right) | x_{ht}', x_{lt}' \right) \\
&= \sup_{\sigma^i} \quad \mathbb{E}^h \left(\int_t^\tau e^{-r\tau} \left(R_i^I \left(x_{ht}' e^{\left(\mu_h - \frac{\sigma^2}{4}\right)(\tau - t) + \frac{1}{\sqrt{2}}\sigma B_{h\tau}}, x_{lt}' e^{\left(\mu_l - \frac{\sigma^2}{4}\right)(\tau - t) + \frac{1}{\sqrt{2}}\sigma B_{l\tau}} \right) \mathbb{1}_\tau^I \right.\right. \\
&\qquad \left.\left. + R_i^m \left(x_{ht}' e^{\left(\mu_h - \frac{\sigma^2}{4}\right)(\tau - t) + \frac{1}{\sqrt{2}}\sigma B_{h\tau}}, x_{lt}' e^{\left(\mu_l - \frac{\sigma^2}{4}\right)(\tau - t) + \frac{1}{\sqrt{2}}\sigma B_{l\tau}} \right) \mathbb{1}_\tau^m \right) | x_{ht}', x_{lt}' \right) \\
&= \sup_{\sigma^i} \quad \mathbb{E}^h \left(\int_t^\tau e^{-r\tau} \left(\alpha R_i^I \left(x_{ht} e^{\left(\mu_h - \frac{\sigma^2}{4}\right)(\tau - t) + \frac{1}{\sqrt{2}}\sigma B_{h\tau}}, x_{lt} e^{\left(\mu_l - \frac{\sigma^2}{4}\right)(\tau - t) + \frac{1}{\sqrt{2}}\sigma B_{l\tau}} \right) \mathbb{1}_\tau^I \right.\right. \\
&\qquad \left.\left. + \alpha R_i^m \left(x_{ht} e^{\left(\mu_h - \frac{\sigma^2}{4}\right)(\tau - t) + \frac{1}{\sqrt{2}}\sigma B_{h\tau}}, x_{lt} e^{\left(\mu_l - \frac{\sigma^2}{4}\right)(\tau - t) + \frac{1}{\sqrt{2}}\sigma B_{l\tau}} \right) \mathbb{1}_\tau^m \right) | x_{ht}', x_{lt}' \right) \\
&= \sup_{\sigma^i} \quad \mathbb{E}^h \alpha \left(\int_t^\tau e^{-r\tau} \left(R_i^I \left(x_{h\tau}, x_{l\tau}\right) \mathbb{1}_\tau^I + R_i^m \left(x_{h\tau}, x_{l\tau}\right) \mathbb{1}_\tau^m \right) | x_{ht}', x_{lt}' \right) \\
&= \sup_{\sigma^i} \quad \mathbb{E}^h \alpha \left(\int_t^\tau e^{-r\tau} \left(R_i^I \left(x_{h\tau}, x_{l\tau}\right) \mathbb{1}_\tau^I + R_i^m \left(x_{h\tau}, x_{l\tau}\right) \mathbb{1}_\tau^m \right) | \alpha x_{ht}, \alpha x_{lt} \right) \\
&= \sup_{\sigma^i} \quad \mathbb{E}^h \alpha \left(\int_t^\tau e^{-r\tau} \left(R_i^I \left(x_{h\tau}, x_{l\tau}\right) \mathbb{1}_\tau^I + R_i^m \left(x_{h\tau}, x_{l\tau}\right) \mathbb{1}_\tau^m \right) | x_{ht}, x_{lt} \right)
\end{aligned}$$

[1]Off equilibrium acquisition offer does not have to equal exogenous reservation value.

The second equality follows from the fact that returns are linear in asset sizes. The last equality follows from the fact that investors belief is decided by assets size ratios (pooling IPOs case) or independent of sizes (acquisition case). Therefore the stopping time τ will not be affected. $\square$

Proof of Theorem 3

We begin the proof by characterizing certain relationships of θ_{1l}, θ_{2l}, θ_{1h} and θ_{2h}.

Lemma 18. *(i)* $\theta_{1h} < 0 < 1 < \theta_{2h}$, $\theta_{1l} < 0 < \theta_{2l}$; *(ii)* $\theta_{1h} + \theta_{2h} = 2 + (\frac{\mu_h-\mu_l}{\sigma^2})^{-1}$, $\theta_{1l} + \theta_{2l} = 2 - (\frac{\mu_h-\mu_l}{\sigma^2})^{-1}$; *(iii)* $\theta_{2h} - \theta_{2l} = \theta_{1h} - \theta_{1l} = (\frac{\mu_h-\mu_l}{\sigma^2})^{-1}$.

Proof. By solving the characteristic function, we have

$$\theta_{1h} = \frac{\frac{\mu_h-\mu_l}{2} + \frac{(\mu_h-\mu_l)^2}{\sigma^2} - \sqrt{(\frac{\mu_h-\mu_l}{2} + \frac{(\mu_h-\mu_l)^2}{\sigma^2})^2 - 2\frac{(\mu_h-\mu_l)^2}{\sigma^2}(\mu_h - r)}}{\frac{(\mu_h-\mu_l)^2}{\sigma^2}}$$

$$\theta_{2h} = \frac{\frac{\mu_h-\mu_l}{2} + \frac{(\mu_h-\mu_l)^2}{\sigma^2} + \sqrt{(\frac{\mu_h-\mu_l}{2} + \frac{(\mu_h-\mu_l)^2}{\sigma^2})^2 - 2\frac{(\mu_h-\mu_l)^2}{\sigma^2}(\mu_h - r)}}{\frac{(\mu_h-\mu_l)^2}{\sigma^2}}$$

$$\theta_{1l} = \frac{-\frac{\mu_h-\mu_l}{2} + \frac{(\mu_h-\mu_l)^2}{\sigma^2} - \sqrt{(-\frac{\mu_h-\mu_l}{2} + \frac{(\mu_h-\mu_l)^2}{\sigma^2})^2 - 2\frac{(\mu_h-\mu_l)^2}{\sigma^2}(\mu_l - r)}}{\frac{(\mu_h-\mu_l)^2}{\sigma^2}}$$

$$\theta_{2l} = \frac{-\frac{\mu_h-\mu_l}{2} + \frac{(\mu_h-\mu_l)^2}{\sigma^2} + \sqrt{(-\frac{\mu_h-\mu_l}{2} + \frac{(\mu_h-\mu_l)^2}{\sigma^2})^2 - 2\frac{(\mu_h-\mu_l)^2}{\sigma^2}(\mu_l - r)}}{\frac{(\mu_h-\mu_l)^2}{\sigma^2}}$$

Notice

$$
\begin{aligned}
&(\frac{\mu_h-\mu_l}{2}+\frac{(\mu_h-\mu_l)^2}{\sigma^2})^2-2\frac{(\mu_h-\mu_l)^2}{\sigma^2}(\mu_h-r)\\
=&(\frac{\mu_h-\mu_l}{2})^2+(\frac{(\mu_h-\mu_l)^2}{\sigma^2})^2+2\frac{\mu_h-\mu_l}{2}\frac{(\mu_h-\mu_l)^2}{\sigma^2}-2\frac{(\mu_h-\mu_l)^2}{\sigma^2}(\mu_h-r)\\
=&(\frac{\mu_h-\mu_l}{2})^2+(\frac{(\mu_h-\mu_l)^2}{\sigma^2})^2+2\frac{(\mu_h-\mu_l)^2}{\sigma^2}(-\frac{\mu_h-\mu_l}{2}-(\mu_l-r))\\
=&(\frac{\mu_h-\mu_l}{2})^2+(\frac{(\mu_h-\mu_l)^2}{\sigma^2})^2-2\frac{\mu_h-\mu_l}{2}\frac{(\mu_h-\mu_l)^2}{\sigma^2}-2\frac{(\mu_h-\mu_l)^2}{\sigma^2}(\mu_h-r)\\
=&(-\frac{\mu_h-\mu_l}{2}+\frac{(\mu_h-\mu_l)^2}{\sigma^2})^2-2\frac{(\mu_h-\mu_l)^2}{\sigma^2}(\mu_l-r)
\end{aligned}
$$

Thus, the part under square root in the above equations are the same. To save space, we define

$$(\frac{\mu_h-\mu_l}{\sigma^2})^{-1}=k$$

$$\frac{\sqrt{(\frac{\mu_h-\mu_l}{2}+\frac{(\mu_h-\mu_l)^2}{\sigma^2})^2-2\frac{(\mu_h-\mu_l)^2}{\sigma^2}(\mu_h-r)}}{\frac{(\mu_h-\mu_l)^2}{\sigma^2}}=\Delta$$

and use those notations hereafter.

Thus, $\theta_{1h}=1+\frac{k}{2}-\Delta$, $\theta_{2h}=1+\frac{k}{2}+\Delta$, $\theta_{1l}=1-\frac{k}{2}-\Delta$, $\theta_{2l}=1-\frac{k}{2}+\Delta$. It's easy to check (ii), (iii) and $\theta_{2h}>1$. Lastly $\theta_{1l}\theta_{2l}=2(\mu_l-r)<0$, $\theta_{1h}\theta_{2h}=2(\mu_h-r)<0$. □

Proof. To show the main results, we first replace the boundary conditions with explicit expression of J_h and J_l:

$$C_{1h}exp(\theta_{1h}\eta) + C_{2h}exp(\theta_{2h}\eta) = H - \alpha - \frac{\alpha(H-L)exp(\eta)}{H + Lexp(\eta)} \tag{A.1}$$

$$C_{1l}exp(\theta_{1l}\eta) + C_{2l}exp(\theta_{2l}\eta) = L - \alpha + \frac{\alpha(H-L)exp(\eta)}{Hexp(\eta) + L} \tag{A.2}$$

$$C_{1h}exp(\theta_{1h}\beta) + C_{2h}exp(\theta_{2h}\beta) = H - \alpha + (H - L - \gamma_l)exp(k\beta) \tag{A.3}$$

$$C_{1l}exp(\theta_{1l}\beta) + C_{2l}exp(\theta_{2l}\beta) = L - \alpha + \gamma_l \tag{A.4}$$

$$C_{1h}\theta_{1h}exp(\theta_{1h}\eta) + C_{2h}\theta_{2h}exp(\theta_{2h}\eta) = -\frac{\alpha H(H-L)exp(\eta)}{(H + Lexp(\eta))^2} \tag{A.5}$$

$$C_{1l}\theta_{1l}exp(\theta_{1l}\beta) + C_{2l}\theta_{2l}exp(\theta_{2l}\beta) = 0 \tag{A.6}$$

Define $x = exp(\eta)$ and $y = exp(\beta)$. Notice since $\beta \leq 0 \leq \eta$, $0 \leq y \leq 1 \leq x$. Replace $exp(\eta)$ and $exp(\beta)$ with x and y and use **Lemma 18** (iii) to change θ_{1h} and θ_{2h} into θ_{1l} and θ_{2l} into (A.1) to (A.6):

$$C_{1h}x^{\theta_{1l}} + C_{2h}x^{\theta_{2l}} = (H - \alpha - f(x))x^{-k} \tag{A.7}$$

$$C_{1l}x^{\theta_{1l}} + C_{2l}x^{\theta_{2l}} = L - \alpha + g(x) \tag{A.8}$$

$$C_{1h}y^{\theta_{1l}} + C_{2h}y^{\theta_{2l}} = (H - \alpha + (H - L - \gamma_l)y^k)y^{-k} \tag{A.9}$$

$$C_{1l}y^{\theta_{1l}} + C_{2l}y^{\theta_{2l}} = L - \alpha + \gamma_l \tag{A.10}$$

$$C_{1h}\theta_{1h}x^{\theta_{1l}} + C_{2h}\theta_{2h}x^{\theta_{2l}} = -f'(x)x^{-k} \tag{A.11}$$

$$C_{1l}\theta_{1l}y^{\theta_{1l}} + C_{2l}\theta_{2l}y^{\theta_{2l}} = 0 \tag{A.12}$$

where $f(x) = \frac{\alpha(H-L)x}{H+Lx}$, $f'(x) = \frac{\alpha H(H-L)x}{(H+Lx)^2}$ and $g(x) = \frac{\alpha(H-L)x}{Hx+L}$.

We use equation (A.10) and (A.12) to solve C_{1l} and C_{2l} and (A.7) and (A.11) to solve C_{1h} and C_{2h}. Then replace the solved constants in (A.8) and (A.9) respectively.

$$[(H-\alpha-f(x))\theta_{2h}+f'(x)](\frac{x}{y})^{-\theta_{1l}}-[(H-\alpha-f(x))\theta_{1h}+f'(x)](\frac{x}{y})^{-\theta_{2l}}$$
$$=(\theta_{2l}-\theta_{1l})[(H-\alpha)+(H-L-\gamma_l)y^k](\frac{x}{y})^k \tag{A.13}$$

$$(L-\alpha+\gamma_l)(\theta_{2l}(\frac{x}{y})^{\theta_{1l}}-\theta_{1l}(\frac{x}{y})^{\theta_{2l}})=(\theta_{2l}-\theta_{1l})(L-\alpha+g(x)) \tag{A.14}$$

Instead of solving y directly, we solve $m=\frac{x}{y}\geq 1$ instead as it's easier to deal with.

$$[(H-\alpha-f(x))\theta_{2h}+f'(x)]m^{-\theta_{1l}}-[(H-\alpha-f(x))\theta_{1h}+f'(x)]m^{-\theta_{2l}}$$
$$-(\theta_{2h}-\theta_{1h})(H-\alpha)m^k=(\theta_{2l}-\theta_{1l})(H-L-\gamma_l)x^k \tag{A.15}$$

$$(L-\alpha+\gamma_l)(\theta_{2l}m^{\theta_{1l}}-\theta_{1l}m^{\theta_{2l}})=(\theta_{2l}-\theta_{1l})(L-\alpha+g(x)) \tag{A.16}$$

Step 1. For any $x\geq 1$, there exists a unique $m_h(x)\geq 1$ and $m_l(x)\geq 1$ that solves the equations (A.15) and (A.16) correspondingly.

To show the part of $m_h(x)$, the LHS of (A.15) is

$$\underbrace{\{[(H-\alpha-f(x))\theta_{2h}+f'(x)]m^{-\theta_{1h}}-[(H-\alpha-f(x))\theta_{1h}+f'(x)]m^{-\theta_{2h}}-(\theta_{2h}-\theta_{1h})(H-\alpha)\}}_{A}m^k$$

Notice $\frac{\partial A}{\partial m}=-\theta_{1h}[(H-\alpha-f(x))\theta_{2h}+f'(x)]m^{-\theta_{1h}-1}+\theta_{2h}[(H-\alpha-f(x))\theta_{1h}+f'(x)]m^{-\theta_{2h}-1}>0$. When $m\to\infty$, $A\to\infty$. When $m=1$, $A=-(\theta_{2h}-\theta_{1h})f(x)<0$. Thus, by intermediate value theorem, there exists $m_h(x)$ that solves the equation. To show uniqueness, notice for any x, the solution exists only when $A>0$. The derivative of the LHS is $\frac{\partial A}{\partial m}m^k+kAm^{k-1}>0$ then, which implies uniqueness.

To show the part of $m_l(x)$, The derivative of the LHS of (A.16) is

$$(L-\alpha+\gamma_l)\theta_{1l}\theta_{2l}(m^{\theta_{1l}-1}-m^{\theta_{2l}-1})$$

As $\theta_{1l}\theta_{2l}<0$ and $\theta_{1l}<\theta_{2l}$, the LHS is monotonically increasing by m. It's easy to check when $m=1$, the LHS is $L-\alpha+\gamma<L-\alpha+g(1)$ by **Assumption 2**. The latter is the minimum of RHS. When $m\to\infty$, $LHS\to\infty$. Thus there is a unique $m_l(x)$ solves the equation.

Step 2. $m_h'(x)>0$ and $m_l'(x)>0$

To show $m_h'(x)>0$, by (A.15)

$$(\frac{\partial A}{\partial m}m^k+kAm^{k-1})dm=(-\frac{\partial A}{\partial x}m^k+(\theta_{2h}-\theta_{1h})(H-L-\gamma_l)kx^{k-1})dx$$

It remains to check $\frac{\partial A}{\partial x}$,

$$\begin{aligned}\frac{\partial A}{\partial x}=&(-f'(x)\theta_{2h}+f''(x))m^{-\theta_{1h}}-(-f'(x)\theta_{1h}+f''(x))m^{-\theta_{2h}}\\=&(-f'(x)(\frac{k}{2}+\Delta)+f''(x)-f'(x))m^{-\theta_{1h}}-(-f'(x)(\frac{k}{2}-\Delta)+f''(x)-f'(x))m^{-\theta_{2h}}\\<&(-f'(x)(\frac{k}{2}+\Delta)+f''(x)-f'(x))m^{-\theta_{1h}}-(-f'(x)(\frac{k}{2}+\Delta)+f''(x)-f'(x))m^{-\theta_{2h}}\\=&(-f'(x)(\frac{k}{2}+\Delta)+f''(x)-f'(x))(m^{-\theta_{1h}}-m^{-\theta_{2h}})\\<&0\end{aligned}$$

The second line comes from **Lemma 18.** The last line comes from $f''(x)<f'(x)$ for any given $x\geq 1$ and $-\theta_{1h}>-\theta_{2h}$. Thus

$$\frac{dm_h}{dx}=\frac{-\frac{\partial A}{\partial x}m^k+(\theta_{2h}-\theta_{1h})(H-L-\gamma_l)kx^{k-1}}{\frac{\partial A}{\partial m}m^k+kAm^{k-1}}>0 \tag{A.17}$$

Similarly,

$$\frac{dm_l}{dx} = \frac{g'(x)(\theta_{2l} - \theta_{1l})}{(L - \alpha + \gamma_l)\theta_{1l}\theta_{2l}(m^{\theta_{1l}-1} - m^{\theta_{2l}-1})} > 0$$

Step 3. $m_h(1) < m_l(1)$ and there exists $\bar{x}$ such that $m_h(x) > m_l(x)$ whenever $x > \bar{x}$.

The second part is easy to check. This is because as $x \to \infty$, RHS of (A.16) is bounded. Thus, m_l is bounded as $x \to \infty$. However, RHS of (A.15) is unbounded as $x \to \infty$. Therefore it must be the case that $m_h \to \infty$.

To show the first part, first notice

$$\begin{aligned}
&\{[(H - \alpha - f(1))\theta_{2h} + f'(1)]m_l(1)^{-\theta_{1h}} - [(H - \alpha - f(1))\theta_{1h} + \\
&f'(1)]m_l(1)^{-\theta_{2h}} - (\theta_{2h} - \theta_{1h})(H - \alpha)\}m_l(1)^k \\
>&(H - \alpha - f(1))\theta_{2h}m_l(1)^{-\theta_{1h}} - (H - \alpha - f(1))\theta_{1h}m_l(1)^{-\theta_{2h}} - (\theta_{2h} - \theta_{1h})(H - \alpha) \\
=&(H - \alpha - f(1))(\theta_{2h}m_l(1)^{-\theta_{1h}} - \theta_{1h}m_l(1)^{-\theta_{2h}}) - (\theta_{2h} - \theta_{1h})(H - \alpha)
\end{aligned}$$

Thus, if we can show $(H - \alpha - f(1))(\theta_{2h}m_l(1)^{-\theta_{1h}} - \theta_{1h}m_l(1)^{-\theta_{2h}}) > (\theta_{2h} - \theta_{1h})(H - \alpha + H - L - \gamma_l)$, by monotonicity we prove $m_h(1) < m_l(1)$. This is equivalent to

$$\frac{L - \alpha + g(1)}{L - \alpha + \gamma_l}\frac{\theta_{2h}m_l(1)^{-\theta_{1h}} - \theta_{1h}m_l(1)^{-\theta_{2h}}}{\theta_{2l}m_l(1)^{\theta_{1l}} - \theta_{1l}m_l(1)^{\theta_{2l}}} > \frac{H - \alpha + H - L - \gamma_l}{H - \alpha - f(1)}$$

which is **Assumption 2**.

Thus, there exists $x^\star$ such that $m_h(x^\star) = m_l(x^\star)$.

Step 4. $x^\star$ is unique.

We prove by showing that $\frac{d^2m_h}{dx^2} > 0$. The result in **Step 3** indicates that there must be $2k+1$ intersections between $m_h(1)$ and $m_l(1)$, where k is an integer. Suppose $k \neq 0$. Then one could find two consecutive intersections where at the first one, m_h crosses m_l from below but crosses m_l from above at the second. It's easy to verify

$g'(x) < 0$ and $\frac{\partial \theta_{1l}\theta_{2l}(m^{\theta_{1l}-1} - m^{\theta_{2l}-1})}{\partial m} > 0$. Thus $\frac{d^2 m_l}{dx^2} < 0$. This implies the $\frac{dm_h}{dx}$ must be decreasing from the first intersection to the second. Thus there $k = 0$.

To see $\frac{d^2 m_h}{dx^2} > 0$. First, $\frac{\partial^2 A}{\partial x \partial m} = (-f'(x)(\frac{k}{2} + \Delta) + f''(x) - f'(x))(-\theta_{1h} m^{-\theta_{1h}-1} + \theta_{2h} m^{-\theta_{2h}-1}) < 0$. Secondly, applying **Lemma 18**

$$\begin{aligned}\frac{\partial A}{\partial m} m^k + kAm^{k-1} &= \theta_{2l}[(H - \alpha - f(x)\theta_{1h} + f'(x)]m^{-\theta_{2l}-1} \\ &- \theta_{1l}[(H - \alpha - f(x)\theta_{2h} + f'(x)]m^{-\theta_{1l}-1} - km^{k-1}(H - \alpha)(\theta_{2h} - \theta_{1h})\end{aligned} \tag{A.18}$$

Taking derivative of (A.18) yields $\frac{\partial \frac{\partial A}{\partial m} m^k + kAm^{k-1}}{\partial m} < 0$. In **Step 2** we've shown $\frac{\partial m_h}{\partial x} > 0$, thus

$$\frac{\partial(\frac{\partial A}{\partial m} m^k + kAm^{k-1})}{\partial x} = \frac{\partial \frac{\partial A}{\partial m} m^k + kAm^{k-1}}{\partial m} \frac{\partial m}{\partial x} + \frac{\partial^2 A}{\partial x \partial m} m^k + k \frac{\partial A}{\partial x} m^{k-1} < 0$$

This implies the denominator is decreasing in x. Similarly, one could show $\frac{\partial^2 A}{\partial x^2} < 0$ and

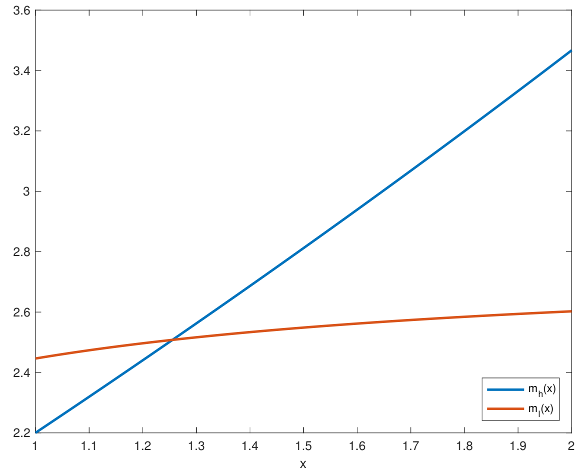
$$\begin{aligned}\frac{\partial(-\frac{\partial A}{\partial x} m^k + (\theta_{2h} - \theta_{1h})(H - L - \gamma_l)kx^{k-1})}{\partial x} &= \underbrace{(-\frac{\partial^2 A}{\partial x \partial m} m^k - k \frac{\partial A}{\partial x} m^{k-1})}_{>0} \frac{\partial m}{\partial x} - \frac{\partial^2 A}{\partial x^2} m^k \\ &+ (\theta_{2h} - \theta_{1h})(H - L - \gamma_l)k(k-1)x^{k-2} > 0\end{aligned}$$

This implies the numerator is increasing in x. This proves $\frac{d^2 m_h}{dx^2} > 0$.

To summarize, we have shown that $m_h(x)$ and $m_l(x)$ are both strictly increasing in x. $m_h(x)$ is below $m_l(x)$ when x is small but above it when x is large. There is only one crossing because $\frac{d^2 m_h}{dx^2} > 0$ and $\frac{d^2 m_l}{dx^2} < 0$. See the figure below for simulated solution of $m_h(x)$ and $m_l(x)$.

□

Figure A.1: Simulated $m_h(x)$ and $m_l(x)$



Proof of Theorem 4

Proof. It remains to show l could not benefit by deviating to act as an acquirer. First notice given the two thresholds (β, η) and the proposed strategy, the sum of waiting value $V_i(x_{it}, \rho_{lt})$ at ρ_{lt} can be expressed as

$$\begin{aligned} SW(x_{ht}, x_{lt}, \rho_{lt}) =& [(H-\alpha)(x_{ht}+x_{lt})]E(e^{-rT(\beta)}|\rho_{lt})P(\rho_{lT}=\beta) \\ & + [(H-\alpha-\frac{\alpha(H-L)exp(\eta)}{(H+Lexp(\eta))})x_{ht} + (L-\alpha+\frac{\alpha(H-L)exp(\eta)}{(Hexp(\eta)+L)})x_{lt}] \\ & E(e^{-rT(\eta)}|\rho_{lt})P(\rho_{lT}=\eta) \end{aligned}$$

In the above equation, $T(\eta)$ and $T(\beta)$ is the first hitting time of ρ_{lt} to η and β. Given that $d\rho_{lt} = -\frac{(\mu_h-\mu_l)^2}{\sigma^2}dt + \frac{\mu_h-\mu_l}{\sigma}dB_t$ follows a Brownian Motion, and define $\theta_1 < \theta_2$ as the two roots for function $\frac{1}{2}(\frac{\mu_h-\mu_l}{\sigma})^2\theta^2 - \frac{(\mu_h-\mu_l)^2}{\sigma^2}\theta - r = 0$. Then following standard methods we could generate the analytical form of $E(e^{-rT(\beta)}|\rho_{lt})$ and

$E(e^{-rT(\eta)}|\rho_{lt})$ to be

$$\psi(\rho_{lt}) = E(e^{-rT(\beta)}|\rho_{lt})P(\rho_{lT} = \beta) = \frac{e^{\theta_1\rho_{lt}}e^{\theta_2\eta} - e^{\theta_2\rho_{lt}}e^{\theta_1\eta}}{e^{\theta_1\beta}e^{\theta_2\eta} - e^{\theta_2\beta}e^{\theta_1\eta}}$$
$$\Psi(\rho_{lt}) = E(e^{-rT(\eta)}|\rho_{lt})P(\rho_{lT} = \eta) = \frac{e^{\theta_1\beta}e^{\theta_2\rho_{lt}} - e^{\theta_2\beta}e^{\theta_1\rho_{lt}}}{e^{\theta_1\beta}e^{\theta_2\eta} - e^{\theta_2\beta}e^{\theta_1\eta}}$$

Also it is easy to verify that $\frac{\partial\psi(\rho_{lt})}{\partial\rho_{lt}} < 0$ and $\frac{\partial\Psi(\rho_{lt})}{\partial\rho_{lt}} > 0$. Together with the fact that $(H-\alpha)(x_{ht}+x_{lt}) > (H-\alpha-\frac{\alpha(H-L)exp(\eta)}{(H+Lexp(\eta))})x_{ht} + (L-\alpha+\frac{\alpha(H-L)exp(\eta)}{(Hexp(\eta)+L)})x_{lt}$, it's easy to verify that $SW(\rho_{lt})$ is decreasing in ρ_{lt}.

Now suppose a deviation to l as the acquirer is possible at $\rho'_{lt} < \eta$, then h must be willing to accept the offer

$$T_{ht} > V_h(x_{ht}, \rho'_{lt})$$

Similarly, l must be willing to deviate

$$(L - \alpha\frac{L}{H})(x_{ht} + x_{lt}) - T_{ht} > V_l(x_{lt}, \rho'_{lt})$$

This implies in such deviation, the total profit is

$$D(\rho'_{lt}) = (L - \alpha\frac{L}{H})(x_{h\tau} + x_{l\tau}) > SW(\rho'_{lt})$$

and $D'(\rho'_{lt}) > 0$. Therefore $D(\eta) > SW(\eta)$, in other words

$$\begin{aligned}
(L-\alpha\frac{L}{H})(x_{ht}+x_{lt}) > SW(\rho'_{lt}) > (H-\alpha-\frac{\alpha(H-L)exp(\eta)}{(H+Lexp(\eta))})x_{ht} \\
+ (L-\alpha+\frac{\alpha(H-L)exp(\eta)}{(Hexp(\eta)+L)})x_{lt} \\
\iff \quad \frac{\alpha(H-L)}{H} > (H-L-\frac{H-L}{H}\alpha - \frac{\alpha(H-L)exp(\eta)}{(Hexp(\eta)+L)})\frac{x_{ht}}{x_{lt}} + \frac{\alpha(H-L)exp(\eta)}{(Hexp(\eta)+L)} \\
\iff \quad \frac{\alpha(H-L)}{H} > (H-L-\frac{H-L}{H}\alpha - \frac{\alpha(H-L)exp(\eta)}{(H+Lexp(\eta))})e^{-\frac{\eta\sigma^2}{\mu_h-\mu_l}} + \frac{\alpha(H-L)exp(\eta)}{(Hexp(\eta)+L)}.
\end{aligned}$$

Notice the RHS is decreasing in η. As $\eta \to +\infty$, RHS is $\frac{\alpha(H-L)}{H}$, which is equal to LHS. Since in equilibrium η is finite, it implies the above inequality cannot hold. □

Proof of Proposition 5

Proof. As we have argued in the paper, the only outcome in the game is that h acquires l. Thus the problem is a simple optimal stopping problem for h. The value function of l is pinned down once the threshold is solved in h's problem as l now is totally passive. Define $V_h^\star(x_{ht}, \rho_{lt}) = x_{ht} J_h^\star(\rho_{lt})$ and $V_l^\star(x_{lt}, \rho_{lt}) = x_{lt} J_l^\star(\rho_{lt})$ as the value functions of h and l in the first-best solution. The Bellman Equations of $J_i^\star(\rho_{lt})$ are the same as $J_i(\rho_{lt})$ so we still have $J_i^\star(\rho_{lt}) = C_{1i}^\star exp(\theta_{1i}\rho_{lt}) + C_{2i}^\star exp(\theta_{2i}\rho_{lt})$.

If in equilibrium h delays its financing decision, there will be two boundaries (potentially infinite) $\beta_1^\star < 0 < \beta_2^\star$ where h is waiting in between. The boundary conditions are:

$$
\begin{aligned}
J_h^\star(\beta_1^\star) &= (H-a) + (H-L-\gamma_l)exp(k\beta_1^\star) \\
J_h^{\star\prime}(\beta_1^\star) &= k(H-L-\gamma_l)exp(k\beta_1^\star) \\
J_h^\star(\beta_2^\star) &= (H-a) + (H-L-\gamma_l)exp(k\beta_2^\star) \\
J_h^{\star\prime}(\beta_2^\star) &= k(H-L-\gamma_l)exp(k\beta_2^\star)
\end{aligned}
$$

Replace $exp(\beta_1^\star)$ with y_1 and $exp(\beta_2^\star)$ with y_2. We solve $C_{1h}^\star$ and $C_{2i}^\star$ from the first two equations

$$
\begin{aligned}
(\theta_{2h}-\theta_{1h})C_{1h}^\star &= \theta_{2h}(H-\alpha)y_1^{-\theta_{1h}} + (\theta_{2h}-k)(H-L-\gamma_l)y_1^{-\theta_{1h}+k} \\
(\theta_{1h}-\theta_{2h})C_{2h}^\star &= \theta_{1h}(H-\alpha)y_1^{-\theta_{2h}} + (\theta_{1h}-k)(H-L-\gamma_l)y_1^{-\theta_{2h}+k}
\end{aligned}
$$

and from the last two equations

$$(\theta_{2h}-\theta_{1h})C^\star_{1h} = \theta_{2h}(H-\alpha)y_2^{-\theta_{1h}} + (\theta_{2h}-k)(H-L-\gamma_l)y_2^{-\theta_{1h}+k}$$
$$(\theta_{1h}-\theta_{2h})C^\star_{2h} = \theta_{1h}(H-\alpha)y_2^{-\theta_{2h}} + (\theta_{1h}-k)(H-L-\gamma_l)y_2^{-\theta_{2h}+k}$$

Define $f_1(x) = \theta_{2h}(H-\alpha)x^{-\theta_{1h}} + (\theta_{2h}-k)(H-L-\gamma_l)x^{-\theta_{1h}+k}$ and $f_2(x) = \theta_{1h}(H-\alpha)x^{-\theta_{2h}} + (\theta_{1h}-k)(H-L-\gamma_l)x^{-\theta_{2h}+k}$. From the above equations, we have $f_1(y_1) = f_1(y_2)$ and $f_2(y_2) = f_2(y_2)$. But for any $x \in [-\infty, +\infty]$, both $f_1'(x)$ and $f_2'(x)$ are strictly positive. Thus, it's only possible that $\beta_1^\star = \beta_2^\star = \beta^\star$, which is contradictory to the assumption of two-threshold equilibrium.

As there is only one threshold, when $\rho_{lt} \to -\infty$ ($\frac{x_{lt}}{x_{ht}} \to 0$), the value function must be bounded so $C^\star_{1i} = 0$. Suppose $exp(k\beta^\star) \neq 0$. Then $exp(k\beta^\star) = \frac{\theta_{2h}(H-\alpha)}{(\frac{k}{\theta_{2h}}-1)(H-L-\gamma_l)}$. If $\frac{k}{\theta_{2h}} - 1 > 0$, then there exists a $\beta^\star$ that solves the problem. However, $\theta_{2h} = 1 + \frac{k}{2} + k\sqrt{(\frac{1}{2}+\frac{1}{k})^2 - 2\frac{\mu_h - r}{\sigma^2}} > 1 + \frac{k}{2} + \frac{k}{2} > k$. This indicates $exp(k\beta^\star) = 0$. In this case acquisition always happen regardless of the value of ρ_{lt}. □

Proof of Proposition 6

We start with the following lemma.

Lemma 19. *i)* $\frac{d\theta_{1h}}{d\mu_h} > 0$, $\frac{d\theta_{1l}}{d\mu_h} > 0$, $\frac{d\theta_{2h}}{d\mu_h} < 0$, $\frac{d\theta_{2l}}{d\mu_h} < 0$.

ii) $\frac{d\theta_{1h}}{d\mu_l} < 0$, $\frac{d\theta_{1l}}{d\mu_l} < 0$, $\frac{d\theta_{2h}}{d\mu_l} > 0$, $\frac{d\theta_{2l}}{d\mu_l} > 0$.

iii) Suppose $\mu_h = \mu_l + \delta$. Fixing δ and increasing μ_h and μ_l simultaneously will increase θ_{1h} and θ_{1l} but decrease θ_{2h} and θ_{2l}.

Proof. i) Let $A = (\frac{1}{2}+\frac{\mu_h-\mu_l}{\sigma^2})^2 - 2\frac{(\mu_h-r)}{\sigma^2}$ and $k = (\frac{\mu_h-\mu_l}{\sigma^2})^{-1}$. Then $\theta_{1h} = 1+(\frac{1}{2}-\sqrt{A})k$, $\theta_{1l} = 1-(\frac{1}{2}+\sqrt{A})k$, $\theta_{2h} = 1+(\frac{1}{2}+\sqrt{A})k$, $\theta_{2l} = 1-(\frac{1}{2}-\sqrt{A})k$. Due to symmetry, we only need to check $\frac{d(\frac{1}{2}-\sqrt{A})k}{d\mu_h}$ and $\frac{d(\frac{1}{2}+\sqrt{A})k}{d\mu_h}$.

$$
\begin{aligned}
&sgn(\frac{d(\frac{1}{2}-\sqrt{A})k}{d\mu_h}) \\
=&sgn(\frac{1}{2\sqrt{A}}-\frac{k}{2}+\sqrt{A}k-\frac{1}{\sqrt{A}k}) \\
=&sgn((\sqrt{A}k-1)(1+\frac{1}{\sqrt{A}k}-\frac{1}{2\sqrt{A}}))>0
\end{aligned}
$$

Since $\sqrt{A}k > \sqrt{(\frac{1}{2}+\frac{1}{k})^2}k = \frac{k}{2}+1 > 1$ and $\frac{1}{2\sqrt{A}} < \frac{1}{2\sqrt{(\frac{1}{2}+\frac{1}{k})^2}} < 1$.

$$
\begin{aligned}
&sgn(\frac{d(\frac{1}{2}+\sqrt{A})k}{d\mu_h}) \\
&= sgn(\frac{1}{\sqrt{A}k}-\frac{1}{2\sqrt{A}}-\frac{k}{2}-\sqrt{A}k)<0
\end{aligned}
$$

Since $\sqrt{A}k > 1$, $\frac{1}{\sqrt{A}k}-\frac{1}{2\sqrt{A}}-\frac{k}{2}-\sqrt{A}k<0$.

ii)

$$
\begin{aligned}
&sgn(\frac{d(\frac{1}{2}-\sqrt{A})k}{d\mu_l}) \\
=&sgn(\frac{1}{2\sqrt{A}}+\frac{k}{2}-\sqrt{A}k+\frac{1}{\sqrt{A}k}) \\
&= sgn((\sqrt{A}k+1)(1+\frac{1}{2}k-\sqrt{A}k))<0
\end{aligned}
$$

Since $\sqrt{A}k > 1+\frac{1}{2}k$.

$$
\begin{aligned}
&sgn(\frac{d(\frac{1}{2}+\sqrt{A})k}{d\mu_l})\\
=&sgn(\frac{k}{2}+\sqrt{A}k-\frac{1}{\sqrt{A}k}-\frac{1}{2\sqrt{A}})\\
&= sgn(\frac{1}{2}\sqrt{A}(\sqrt{A}k-1)+\sqrt{A}k-\frac{1}{\sqrt{A}k}) < 0
\end{aligned}
$$

Since $\sqrt{A}k > 1$.

iii) Since δ is fixed, thus k is unchanged but A is smaller. Thus θ_{1h} and θ_{1l} are larger but decrease θ_{2h} and θ_{2l} are smaller. □

Proof. Rearrange equations (A.15) and (A.16) to be

$$M_1 m^k = (H - L - \gamma_l)x^k,$$

and

$$(L - \alpha + \gamma_l)M_2 = L - \alpha + g(x).$$

Take the derivative of M_1 w.r.t μ_h

$$
\begin{aligned}
\frac{\partial M_1}{\partial \mu_h} =& \frac{d\theta_{1h}}{d\mu_h}\underbrace{[(H-\alpha-f(x))\theta_{2h}+f'(x)]}_{>0}(\frac{m^{-\theta_{1h}}-m^{-\theta_{2h}}}{\theta_{2h}-\theta_{1h}}-m^{-\theta_{1h}}log(m))\\
&+\frac{d\theta_{2h}}{d\mu_h}[(H-\alpha-\underset{<0}{f(x))\theta_{1h}}+f'(x)](\frac{m^{-\theta_{2h}}-m^{-\theta_{1h}}}{\theta_{2h}-\theta_{1h}}+m^{-\theta_{2h}}log(m))
\end{aligned}
$$

We can show both $\frac{m^{-\theta_{1h}}-m^{-\theta_{2h}}}{\theta_{2h}-\theta_{1h}} - m^{-\theta_{1h}}log(m)$ and $\frac{m^{-\theta_{2h}}-m^{-\theta_{1h}}}{\theta_{2h}-\theta_{1h}} + m^{-\theta_{2h}}log(m)$ are negative if $m > 1$. We only list the proof of $\frac{m^{-\theta_{1h}}-m^{-\theta_{2h}}}{\theta_{2h}-\theta_{1h}} - m^{-\theta_{1h}}log(m)$ as the other is identical. First notice when $m = 1$, $\frac{m^{-\theta_{1h}}-m^{-\theta_{2h}}}{\theta_{2h}-\theta_{1h}} - m^{-\theta_{1h}}log(m) = 0$. Taking derivative of the function yields

$$\begin{aligned}
&\frac{-\theta_{1h}m^{-\theta_{1h}-1}+\theta_{2h}m^{-\theta_{2h}-1}}{\theta_{2h}-\theta_{1h}} - m^{-\theta_{1h}-1} + \theta_{1h}m^{-\theta_{1h}-1}log(m) \\
<&\frac{-\theta_{1h}m^{-\theta_{1h}-1}+\theta_{2h}m^{-\theta_{2h}-1}}{\theta_{2h}-\theta_{1h}} - m^{-\theta_{1h}-1} \\
=&\frac{\theta_{2h}(m^{-\theta_{2h}-1}-m^{-\theta_{1h}-1})}{\theta_{2h}-\theta_{1h}} < 0
\end{aligned}$$

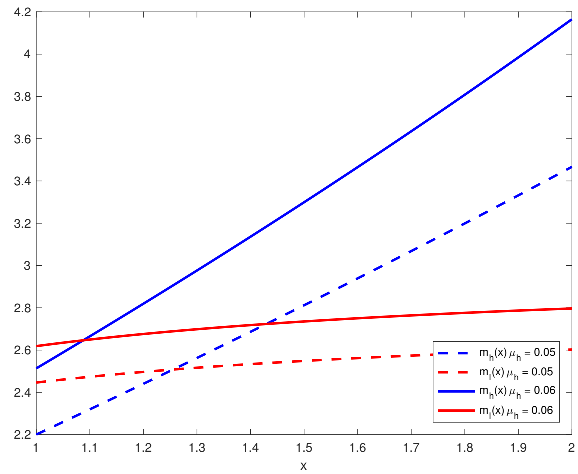
The second line is due to $\theta_{1h} < 0$ and $log(m) > 0$. By **Lemma 19**, $\frac{d\theta_{1h}}{d\mu_h} > 0$ and $\frac{d\theta_{2h}}{d\mu_h} < 0$. Lastly we have $\frac{\partial M_1}{\partial \mu_h} < 0$. One can also easily check $\frac{dm^k}{d\mu_h} < 0$. For any given value of m, the LHS of equation (**??**) is strictly decreasing in μ_h. Therefore $\frac{dm_h(x;\mu_h)}{d\mu_h} > 0$. In other words, $m_h(x)$ shifts towards northwest as μ_h increases. Applying the same kind of trick, we can show

$$\begin{aligned}
\frac{\partial M_2}{\partial \mu_h} = &\frac{\theta_{2l}}{\theta_{2l}-\theta_{1l}}\frac{d\theta_{1l}}{d\mu_h}(\underbrace{\frac{m^{\theta_{1l}}-m^{\theta_{2l}}}{\theta_{2l}-\theta_{1l}} + log(m)m^{\theta_{1l}})}_{<0} \\
&+ \frac{\theta_{1l}}{\theta_{2l}-\theta_{1l}}\frac{d\theta_{2l}}{d\mu_h}(\underbrace{\frac{m^{\theta_{2l}}-m^{\theta_{1l}}}{\theta_{2l}-\theta_{1l}} - log(m)m^{\theta_{2l}})}_{<0} < 0
\end{aligned}$$

Thus, both $m_h(x)$ and $m_l(x)$ shift towards northwest as μ_h increases. Denote $m^\star(\mu_h)$ and $x^\star(\mu_h)$ as solution of the equations. By comparing $\frac{\partial M_1}{\partial \mu_h}$ over $\frac{\partial M_1}{\partial m}$ and $\frac{\partial M_2}{\partial \mu_h}$ over $\frac{\partial M_2}{\partial m}$, one could show the distance the $m_h(x)$ moves upwards with is larger than $m_l(x)$ does. This regulates that $\frac{dx^\star(\mu_h)}{d\mu_h} < 0$. Notice $y^\star = \frac{x^\star}{m^\star} = cot(\theta_x)$, where θ_x is the angle between the line $(x^\star, m^\star) \to (0, 0)$ and $(x^\star, 0) \to (0, 0)$. As $(x^\star, m^\star)$ shifts northwest, θ_x increases and $cot(\theta_x)$ decreases. Hence $\frac{dy^\star(\mu_h)}{d\mu_h} < 0$.

The proof for ii) is similar. Replace all $\frac{d\theta}{d\mu_h}$ with $\frac{d\theta}{d\mu_l}$ in the above derivatives. It's easy to verify all signs of inequality are reversed. $\square$

Figure A.2: Simulated $m_h(x)$ and $m_l(x)$ as μ_h Increases



Proof of Proposition 7

Proof. The proof is similar to Proposition 6. Denote θ'_{ij} as the changes of θ when μ_h and μ_l are simultaneously increased without changing the wedge. By **Lemma 19**, $\theta'_{1h} > 0$, $\theta'_{1l} > 0$, $\theta'_{2h} < 0$, and $\theta'_{2l} < 0$. Using the same notation, $M'_1 < 0$ and $M'_2 < 0$. Thus, both $m_h(x)$ and $m_l(x)$ shift towards northwest as μ_h and μ_l increase. The rest follows argument in Proposition 6. □

Proof of Theorem 8

Proof. The statement is true if the underpricing cost in pooling IPO is larger than h's net benefit in acquisition $\forall \rho \leq 0$. It is equivalent to:

$$
\begin{aligned}
& -\frac{\alpha(H-L)exp(\rho)}{H+Lexp(\rho)} > -(1-\xi)\alpha\frac{H-L}{L} + \xi(H-L)\frac{x_{lt}}{x_{ht}} \\
\iff \quad & -\frac{\alpha(H-L)exp(\rho)}{H+Lexp(\rho)} > -(1-\xi)\alpha\frac{H-L}{L} + \xi(H-L)exp(k\rho) \\
\iff \quad & (1-\xi)\alpha\frac{H-L}{L} > \frac{\alpha(H-L)exp(\rho)}{H+Lexp(\rho)} + \xi(H-L)exp(k\rho)
\end{aligned}
$$

The RHS of last line is increasing in ρ. Therefore it is sufficient to verify $(1-\xi)\alpha\frac{H-L}{L} > \frac{\alpha(H-L)}{H+L} + \xi(H-L)$, which is true if and only if $\xi \leq \underline{\xi} = \frac{\frac{\alpha}{L}-\frac{\alpha}{H+L}}{1+\frac{\alpha}{L}} < 1$. The remaining statement comes from solving a standard optimal problem for h. □

Proof of Theorem 9

Proof. Following the same procedure of proving Theorem 3, I first write down the new boundary conditions under Nash Bargaining:

$$C_{1h}exp(\theta_{1h}\eta) + C_{2h}exp(\theta_{2h}\eta) = H - \alpha - \frac{\alpha(H-L)exp(\eta)}{H + Lexp(\eta)} \tag{A.19}$$

$$C_{1l}exp(\theta_{1l}\eta) + C_{2l}exp(\theta_{2l}\eta) = L - \alpha + \frac{\alpha(H-L)exp(\eta)}{Hexp(\eta) + L} \tag{A.20}$$

$$C_{1h}exp(\theta_{1h}\beta) + C_{2h}exp(\theta_{2h}\beta) = H - \alpha - (1-\xi)\alpha\frac{H-L}{L} + \xi(H-L)exp(k\beta) \tag{A.21}$$

$$C_{1l}exp(\theta_{1l}\beta) + C_{2l}exp(\theta_{2l}\beta) = L - \alpha + (1-\xi)(H-L) + (1-\xi)\alpha\frac{H-L}{L}exp(-k\beta) \tag{A.22}$$

$$C_{1h}\theta_{1h}exp(\theta_{1h}\eta) + C_{2h}\theta_{2h}exp(\theta_{2h}\eta) = -\frac{\alpha H(H-L)exp(\eta)}{(H + Lexp(\eta))^2} \tag{A.23}$$

$$C_{1l}\theta_{1l}exp(\theta_{1l}\beta) + C_{2l}\theta_{2l}exp(\theta_{2l}\beta) = -k(1-\xi)\alpha\frac{H-L}{L}exp(-k\beta) \tag{A.24}$$

Define $x = exp(\eta)$ and $y = exp(\beta)$. Notice since $\beta \leq 0 \leq \eta$, $0 \leq y \leq 1 \leq x$. Replace $exp(\eta)$ and $exp(\beta)$ with x and y and use **Lemma 18** (iii) to change θ_{1h} and θ_{2h} into θ_{1l} and θ_{2l} into (A.19) to (A.24):

$$C_{1h}x^{\theta_{1l}} + C_{2h}x^{\theta_{2l}} = (H - \alpha - f(x))x^{-k} \tag{A.25}$$

$$C_{1l}x^{\theta_{1l}} + C_{2l}x^{\theta_{2l}} = L - \alpha + g(x) \tag{A.26}$$

$$C_{1h}y^{\theta_{1l}} + C_{2h}y^{\theta_{2l}} = (H - \alpha - (1-\xi)\alpha\frac{H-L}{L} + \xi(H-L)y^k)y^{-k} \tag{A.27}$$

$$C_{1l}y^{\theta_{1l}} + C_{2l}y^{\theta_{2l}} = L - \alpha + (1-\xi)(H-L) + (1-\xi)\alpha\frac{H-L}{L}y^{-k} \tag{A.28}$$

$$C_{1h}\theta_{1h}x^{\theta_{1l}} + C_{2h}\theta_{2h}x^{\theta_{2l}} = -f'(x)x^{-k} \tag{A.29}$$

$$C_{1l}\theta_{1l}y^{\theta_{1l}} + C_{2l}\theta_{2l}y^{\theta_{2l}} = -k(1-\xi)\alpha\frac{H-L}{L}y^{-k} \tag{A.30}$$

where $f(x) = \frac{\alpha(H-L)x}{H+Lx}$, $f'(x) = \frac{\alpha H(H-L)x}{(H+Lx)^2}$ and $g(x) = \frac{\alpha(H-L)x}{Hx+L}$.

We use equation (A.28) and (A.30) to solve C_{1l} and C_{2l} and (A.25) and (A.29) to solve C_{1h} and C_{2h}. Then replace the solved constants in (A.26) and (A.27) respectively.

Using the same trick, we solve $m = \frac{x}{y} \geq 1$ and x:

$$\begin{aligned}&[(H-\alpha-f(x))\theta_{2h} + f'(x)]m^{-\theta_{1l}} - [(H-\alpha-f(x))\theta_{1h} + f'(x)]m^{-\theta_{2l}} \\ &-(\theta_{2h}-\theta_{1h})(H-\alpha-(1-\xi)\alpha\frac{H-L}{L})m^k = (\theta_{2l}-\theta_{1l})\xi(H-L)x^k\end{aligned} \tag{A.31}$$

$$\begin{aligned}&(L-\alpha+(1-\xi)(H-L))(\theta_{2l}m^{\theta_{1l}} - \theta_{1l}m^{\theta_{2l}}) \\ &+(1-\xi)\alpha\frac{H-L}{L}x^{-k}(\theta_{2h}m^{\theta_{1h}} - \theta_{1h}m^{\theta_{2h}}) = (\theta_{2l}-\theta_{1l})(L-\alpha+g(x))\end{aligned} \tag{A.32}$$

I go through the same steps of proofs in Theorem 3. Since most derivations are essentially the same, I only highlight the different parts to avoid repetition.

Step 1. If $\xi > \frac{1+\frac{\alpha}{L}-\frac{\alpha}{H+L}}{1+\frac{\alpha}{L}}$, then $\forall x \geq 1$ there exists a unique $m_h(x) \geq 1$ and $m_l(x) \geq 1$ that solves the equations (A.31) and (A.32) correspondingly.

The statement on $m_h(x)$ can be shown in the same way as before. For $m_l(x)$, the

LHS of (A.32) is monotonically still increasing by m. When $m \to \infty$, $LHS \to \infty$. When $m = 1$, the LHS is $(L - \alpha + (1-\xi)(H-L)) + (1-\xi)\alpha\frac{H-L}{L}$. It is smaller than $L - \alpha + \gamma < L - \alpha + g(1)$ if and only if $\xi > \frac{1+\frac{\alpha}{L}-\frac{\alpha}{H+L}}{1+\frac{\alpha}{L}}$. $L - \alpha + g(1)$ is the minimum of RHS. This proves the statement.

Step 2. $m_h'(x) > 0$ and $m_l'(x) > 0$.

$\frac{dm_h}{dx}$ follows the same as before.

$$\frac{dm_l}{dx} = \frac{A(x)}{B(m)} > 0$$

Where $A(x) = g'(x)(\theta_{2l} - \theta_{1l}) + k(1-\xi)\alpha\frac{H-L}{L}x^{-k-1}(\theta_{2h}m^{\theta_{1h}} - \theta_{1h}m^{\theta_{2h}}) > 0$ and $B(m) = (L-\alpha+\gamma_l)\theta_{1l}\theta_{2l}(m^{\theta_{1l}-1} - m^{\theta_{2l}-1}) + (1-\xi)\alpha\frac{H-L}{L}x^{-k}\theta_{1h}\theta_{2h}(m^{\theta_{1h}-1} - m^{\theta_{2h}-1}) > 0$.

Step 3. $m_h(1) < m_l(1)$ and there exists $\bar{x}$ such that $m_h(x) > m_h(x)$ whenever $x > \bar{x}$.

It only remains to check the first part. Consider the solution of $m_l(1)$ in (A.32), satisfying

$$\begin{aligned}(L - \alpha + (1-\xi)(H-L))(\theta_{2l}m_l(1)^{\theta_{1l}} - \theta_{1l}m_l(1)^{\theta_{2l}})& \\ +(1-\xi)\alpha\frac{H-L}{L}(\theta_{2h}m_l(1)^{\theta_{1h}} - \theta_{1h}m_l(1)^{\theta_{2h}}) =& (\theta_{2l} - \theta_{1l})(L - \alpha + g(1))\end{aligned} \tag{A.33}$$

Notice $\theta_{2h}m_l(1)^{\theta_{1h}} - \theta_{1h}m_l(1)^{\theta_{2h}} < \theta_{2l}m_l(1)^{\theta_{1l}} - \theta_{1l}m_l(1)^{\theta_{2l}}$. This implies

$$\theta_{2l}m_l(1)^{\theta_{1l}} - \theta_{1l}m_l(1)^{\theta_{2l}} > \frac{(\theta_{2l} - \theta_{1l})(L - \alpha + g(1))}{[L - \alpha + (1-\xi)(H-L)] + (1-\xi)\alpha\frac{H-L}{L}} \tag{A.34}$$

Now consider the following equation:

$$
\begin{aligned}
&\{[(H-\alpha-f(1))\theta_{2h}+f'(1)]m_l(1)^{-\theta_{1h}}-[(H-\alpha-f(1))\theta_{1h}+f'(1)]m_l(1)^{-\theta_{2h}}\\
&-(\theta_{2h}-\theta_{1h})(H-\alpha-(1-\xi)\alpha\frac{H-L}{L})\}m_l(1)^k\\
>&(H-\alpha-f(1))\theta_{2h}m_l(1)^{-\theta_{1h}}-(H-\alpha-f(1))\theta_{1h}m_l(1)^{-\theta_{2h}}\\
&-(\theta_{2h}-\theta_{1h})(H-\alpha-(1-\xi)\alpha\frac{H-L}{L})\\
=&(H-\alpha-f(1))(\theta_{2h}m_l(1)^{-\theta_{1h}}-\theta_{1h}m_l(1)^{-\theta_{2h}})\\
&-(\theta_{2h}-\theta_{1h})(H-\alpha-(1-\xi)\alpha\frac{H-L}{L})
\end{aligned}
$$

Thus, if we can show $(H-\alpha-f(1))(\theta_{2h}m_l(1)^{-\theta_{1h}}-\theta_{1h}m_l(1)^{-\theta_{2h}})>(\theta_{2h}-\theta_{1h})(H-\alpha-(1-\xi)\alpha\frac{H-L}{L}+\xi(H-L))$, by monotonicity we prove $m_h(1)<m_l(1)$. By equation (A.34), a sufficient condition is

$$
\frac{L-\alpha+g(1)}{[L-\alpha+(1-\xi)(H-L)]+(1-\xi)\alpha\frac{H-L}{L}}\frac{\theta_{2h}m_l(1)^{-\theta_{1h}}-\theta_{1h}m_l(1)^{-\theta_{2h}}}{\theta_{2l}m_l(1)^{\theta_{1l}}-\theta_{1l}m_l(1)^{\theta_{2l}}} \tag{A.35}
$$

$$
>\frac{H-\alpha-(1-\xi)\alpha\frac{H-L}{L}+\xi(H-L)}{H-\alpha-f(1)} \tag{A.36}
$$

When $\xi=1$, this equation reduces to

$$
\frac{L-\alpha+g(1)}{L-\alpha}\gg\frac{H-\alpha+H-L}{H-\alpha-f(1)}. \tag{A.37}
$$

By continuity, if this is true, there exists a threshold $\xi^\star$ such that (A.36) is true whenever $\xi\geq\xi^\star$. Denote $\bar{\xi}=\max\{\xi^\star,\frac{1+\frac{\alpha}{L}-\frac{\alpha}{H+L}}{1+\frac{\alpha}{L}}\}$. The theorem is proved. □

Proof of Proposition 10

Proof. Rewrite A as $(\frac{1}{2}+\frac{\mu_h-\mu_l}{\sigma^2}-\delta)^2-2\frac{(\mu_h-r)}{\sigma^2}$. Obviously $\frac{dA}{d\delta}<0$. Therefore $\frac{d\theta_{1h}}{d\delta}>0$, $\frac{d\theta_{1l}}{d\delta}>0$, $\frac{d\theta_{2h}}{d\delta}<0$, $\frac{d\theta_{2l}}{d\delta}<0$. The rest follows proof of Proposition 6. □

Appendix B

Appendix for Chapter 2

B.1 Equilibrium Definition

Since the game ends as soon as $e_t = 0$, the only relevant history is the one in which the investor has chosen $e_s = 1$ for all $s < t$. We thus do not need to keep track of e_t. Let $\mathcal{C} = [0, \infty) \times [0, 1]$ be the space of possible contracts in any given period, where each contract is given by a pair $C_t = (d_t, \alpha_t)$.

A history for entrepreneur type θ in period t is given by $h^{\theta t} = \{\theta, C_1, C_2, ..., C_{t-1}\} \in H^{\theta t}$. A history for the investor is given by $h^t = \{C_1, C_2, ..., C_{t-1}, C_t\} \in H^t$. A strategy for a type θ entrepreneur is pair $\sigma_t^\theta = \left(g_t^\theta\left(h^t, C\right), l_t^\theta\left(h^t\right)\right)$, where $g_t^\theta : H^{\theta t} \to \Delta\left(\mathcal{C}\right)$ is the probability she offers contract C given history h^t and $l_t^\theta : H^{\theta t} \to [0, 1]$ is the probability she liquidates the project. A strategy for the investor is $\sigma_t : H^t \to [0, 1]$, which maps (h^t, C_t) into a distribution over $e_t \in \{0, 1\}$.

A Perfect Bayesian Equilibrium consists of strategies σ^θ and σ, and beliefs p_t^θ, $p_t\left(q_t\right)$, and q_t, such that for all $t \leq \tau$, σ_t^h, σ_t^l, and σ_t are sequentially rational at all histories and beliefs satisfy Bayes' rule whenever possible.

This definition is consistent with the following extensive form stage game in each period t:

- Stage 1: The entrepreneur chooses l_t^θ.

- Stage 2: If the entrepreneur has not liquidated, she chooses C_t conditional on $h^{\theta t}$.
- Stage 3: The investor observes C_t (and the fact that the game has not ended yet) and chooses e_t conditional on h^t (which includes C_t).
- Stage 4: The project succeeds or not. If not, the game proceeds to period $t+1$.

B.2 Proofs

B.2.1 Proof of Lemma 11

To establish that liquidation is optimal whenever $\lambda p_t^\theta V - c - k \leq 0$, suppose by way of contradiction that the period payoff is strictly positive when the entrepreneur liquidates. Then $\bar{\Pi}_t^\theta > 0$, because the entrepreneur always has the option of liquidating in the next period, so that $\bar{\Pi}_{t+1}^\theta \geq 0$. Thus, liquidating in period t cannot be optimal. Conversely, if the period payoff is non-positive, it will be strictly negative in all future periods, because the belief p_t^θ is strictly decreasing. Thus, it must be the case that $\bar{\Pi}_{t+1}^\theta < 0$. If the entrepreneur continues the project, she earns a negative value. Thus, liquidating is optimal.

Period τ^θ is defined as the first period in which $\lambda p_t^\theta V - c - k$ becomes negative, or, equivalently, the first period for which

$$p_t^\theta \leq \frac{c+k}{\lambda V}.$$

Since $p_t^l < p_t^h$ for all t, we have $\tau^l < \tau^h$.

That the optimal equity share is given $\bar{\alpha}_t^\theta$ has been established in the text.

B.2.2 Proof of Proposition 12

We start with some preliminaries. First, no pooling equilibrium with inefficient liquidation can be optimal.

Lemma 20. *There exists no optimal pooling equilibrium in which* $l_t^l = l_t^h = 1$*, but* $\Pi_t^h(1) > 0$.

Proof. To characterize alternative equilibria with higher payoffs, we must consider a number of cases. Let $\Pi_t^l(q_{t-1})$ be the low type's value in period t under the strategies $l_t^{l\prime} = 0$, $l_t^{h\prime} = 0$, $\alpha_t' = \alpha_t^P(q_{t-1})$, $l_{t+1}^{l\prime} = l_t^l$, etc. Note that we have $\Pi_t^h(1) \geq \Pi_t^l(1) \geq \Pi_t^l(q_{t-1})$. Thus, if $\Pi_t^l(q_{t-1}) \geq 0$, both types simply continue and offer contract $\alpha_t' = \alpha_{t-1}^P$. Then, the belief in the alternative equilibrium is $q' = q_{t-1}$, which ensures that the investor's IC condition holds. The payoffs are then larger: $\Pi_t^{h\prime} > \Pi_t^{l\prime} = \Pi_t^l(q_{t-1}) \geq 0$.[1]

Suppose instead that $\Pi_t^l(q_{t-1}) < 0 \leq \Pi_t^l(1) \leq \Pi_t^h(1)$. Then, by continuity, there exists a belief q' such that $\Pi_t^l(q') = 0$. Consider the following alternative equilibrium: $l_t^{l\prime}, l_t^{h\prime} \in (0,1)$ such that

$$1 - l_t^{l\prime} = \left(1 - l_t^{h\prime}\right) q_{t-1} \left(\frac{1-q'}{q'}\right)$$

and

$$\alpha_t' = \frac{c}{\lambda p_t(q') V}.$$

The liquidation probabilities induce the belief q'. This yields the same payoff as in equilibrium for the low type, but a strictly larger payoff for the high type. Finally, suppose that $\Pi_t^h(1) > 0 = \Pi_t^l(1)$. Then, picking $l_t^{h\prime} = 0$ and $l_t^{l\prime} = 1$ is a Pareto improvement. $\square$

[1] Recall that in any pooling equilibrium $\Pi_t^h > \Pi_t^l$, by Lemma 21.

In equilibrium, the high type receives a larger payoff.

Lemma 21. *In any pooling equilibrium, we have* $\Pi_t^h \geq \Pi_t^l$*. If* $l_t^l < 1$*, then* $\Pi_t^h > \Pi_t^l$*.*

Proof. Since choosing l_t^l is (weakly) suboptimal for the high type, his value satisfies

$$\Pi_t^h \geq \sum_{s=t}^{\infty} \delta^{s-t} \left[\prod_{t \leq u \leq s-1} \left(1 - \lambda p_u^h\right) \left(1 - l_u^l\right) \right] \left(1 - l_s^l\right) \left(\lambda p_s^h \left(1 - \alpha_s^P\right) V - k\right). \quad \text{(B.1)}$$

Using the updating rule in Equation (2.1) repeatedly yields

$$\prod_{t \leq u \leq s-1} \left(1 - \lambda p_u^h\right) \lambda p_s^h = \lambda p_t^h \left(1 - \lambda\right)^{s-t},$$

so that the RHS of Equation (B.1) equals

$$\begin{aligned} & p_t^h \sum_{s=t}^{\infty} \left(\delta \left(1 - \lambda\right)\right)^{s-t} \left[\prod_{t \leq u \leq s} \left(1 - l_u^l\right) \right] \lambda \left(1 - \alpha_s^P\right) V \\ & - \sum_{s=t}^{\infty} \delta^{s-t} \left[\prod_{t \leq u \leq s-1} \left(1 - \lambda p_u^h\right) \left(1 - l_u^l\right) \right] \left(1 - l_s^l\right) k. \end{aligned}$$

We similarly obtain for the low type

$$\begin{aligned} \Pi_t^l \;=\; & p_t^l \sum_{s=t}^{\infty} \left(\delta \left(1 - \lambda\right)\right)^{s-t} \left[\prod_{t \leq u \leq s} \left(1 - l_u^l\right) \right] \lambda \left(1 - \alpha_s^P\right) V \\ & - \sum_{s=t}^{\infty} \delta^{s-t} \left[\prod_{t \leq u \leq s-1} \left(1 - \lambda p_u^l\right) \left(1 - l_u^l\right) \right] \left(1 - l_s^l\right) k. \end{aligned}$$

Since $p_t^h > p_t^l$ for all t, combining the two expressions yields $\Pi_t^h \geq \Pi_t^l$ and the inequality is strict if $l_t^l < 1$. The lemma implies that if the low type does not liquidate, the high type will not liquidate either. We will exploit this fact throughout. □

Since the high type receives a higher payoff, she liquidates later.

Corollary 22. *Whenever $l_t^l = 0$, we have $l_t^h = 0$. Whenever $l_t^h > 0$, we have $l_t^l = 1$. There exists no equilibrium in which $l_t^l, l_t^h \in (0,1)$.*

Proof. Liquidating with probability $l_t^l \in (0,1)$ is optimal for the low type if and only if

$$\Pi_t^l = 0,$$

where Π_t^l is the equilibrium value of the low type. Similarly, liquidating with probability $l_t^h \in (0,1)$ is optimal for the high type if and only if

$$\Pi_t^h = 0.$$

Lemma 21 then implies the results. □

Moreover, the constraint $\Pi_t^\theta \geq \Pi_t^\theta(0)$ does not bind in equilibrium whenever both types continue.

Lemma 23. *For all $t < \bar{\tau}^l$, we have $\Pi_t^\theta > \Pi_t^\theta(0)$.*

Proof. We have, using a similar calculation as in Lemma 21,

$$\begin{aligned}\Pi_t^h - \Pi_t^h(0) =& \lambda p_t^h V \sum_{s=t}^{\bar{\tau}^l - 1} (\delta(1-\lambda))^{s-t} \left(\bar{\alpha}_t^l - \alpha_s^P\right) \\ &+ \delta^{\bar{\tau}^l - t} \Pi_{t \leq u < \bar{\tau}^l - 1} \left(1 - \lambda p_u^h\right) \left(\Pi_{\bar{\tau}^l}^h(1) - \Pi_{\bar{\tau}^l}^h(0)\right)\end{aligned}$$

and[2]

$$\Pi_t^l - \Pi_t^l(0) \geq \lambda p_t^l V \sum_{s=t}^{\bar{\tau}^l} (\delta(1-\lambda))^{s-t} \Pi_{t \leq u < s-1} \left(1 - l_u^l\right) \left(\bar{\alpha}_s^l - \alpha_s^P\right).$$

[2]Note that by construction, $\Pi_{\bar{\tau}^l}^l = \Pi_{\bar{\tau}^l}^l(0) = 0$.

For all $t < \bar{\tau}^l$, we have $q_0 \leq q_t$ and therefore $\alpha_t^P < \bar{\alpha}_t^l$. We also have $l_t^l < 1$ and $\Pi_t^h(1) \geq \Pi_t^h(0)$ for all t. Thus, both expressions are strictly positive. □

We are now done with preliminaries and ready to prove the proposition. We first show that the contract in Proposition 12 is optimal among all pooling contracts.

Proposition 24. *Any other pooling equilibrium yields weakly lower payoffs for both types than the equilibrium in Proposition 12.*

Specifically, the next series of Lemmas establishes that $U_t = 0$, that

$$\alpha_t^P = \frac{c}{\lambda p_t(q_t) V}$$

for all $t < \bar{\tau}^l$, and that no equilibrium with different payoffs can be optimal.

Lemma 25. *Suppose that $1 = l_t^l \geq l_t^h$. Then, $U_t = 0$ and if $l_t^h < 1$, the high type offers the optimal contract $\bar{\alpha}_s^h$ for all $s \geq t$.*

Proof. That $U_t = 0$ follows directly from Equation (B.2). Since $l_t^l = 1$, we have $q_t = 1$. Then, any equilibrium must have the high type offer $\bar{\alpha}_s^h$ for $s \geq t$. Overpaying the investor, i.e. $\alpha_s' > \bar{\alpha}_s^h$, is not optimal for the high type and liquidating for $t < s < \tau^h$ cannot be optimal either. □

Thus, if $l_t^l = 1$, the contract is uniquely pinned down for $s \geq t$. We can therefore restrict attention to periods in which $l_t^l < 1$ and, by Corollary 22, $l_t^h = 0$. We keep this restriction throughout the remainder of the section.

Lemma 26. *Any optimal pooling contract features $U_t = 0$ and*

$$\alpha_t^P = \frac{c - \delta(1 - \lambda p_t(q_t)) U_{t+1}}{\lambda p_t(q_t) V}$$

whenever $\Pi_t^l > 0$.

Proof. The investor experiments whenever

$$U_t = (1 - l_t(q_{t-1}))\left(\lambda p_t(q_t)\,\alpha_t^P V - c + \delta(1 - \lambda p_t(q_t))\,U_{t+1}\right) \geq 0. \tag{B.2}$$

Suppose that $U_t > 0$, $\Pi_t^l > 0$, and $l_t^l = l_t^h = 0$ for some t.[3] We can generate an improvement for the entrepreneur by picking equity share $\alpha_t' = \alpha_t^P - \varepsilon$, where, ε is chosen sufficiently small to ensure that the investor's value remains positive. This clearly increases both types' payoffs. □

We next show that for any pooling equilibrium in which $\Pi_t^l = 0$ and $U_t > 0$, there exists another equilibrium in which $U_t = 0$ and which yields at least weakly increases the payoffs to both types in period t. We distinguish two cases, when $\Pi_{t+1}^l > 0$ and when $\Pi_{t+1}^l = 0$.

Lemma 27. *Suppose that $\Pi_t^l = \Pi_{t+1}^l = 0$. Then, any pooling equilibrium in which*

$$\alpha_t^P > \frac{c}{\lambda p_t(q_{t-1})\,V}$$

is not optimal.

[3]Recall that $l_t^l = 0$ implies $l_t^h = 0$ by Corollary 22.

Proof. Consider the following alternative equilibrium

$$\begin{aligned}
\alpha_t' &= \frac{c}{\lambda p_t\left(q_{t-1}\right)V} \\
l_t^{l\prime} &= l_t^{h\prime} = 0 \\
q_t' &= q_{t-1} \\
\alpha_{t+1}' &= \alpha_t^P \\
l_{t+1}^{\theta\prime} &= l_t^\theta \text{ for } \theta = l, h \\
q_{t+1}' &= q_t \\
&\ldots
\end{aligned}$$

We now verify that the alternative contract is indeed an equilibrium and improves the entrepreneur's payoffs. First, in any equilibrium, we have $U_{t+1} \geq 0$. Thus, the investor experiments whenever his share exceeds $c/\left(\lambda p_t\left(q_t'\right)V\right)$. In particular, he experiments at α_t' given belief $q_t' = q_{t-1}$. Second, since $\Pi_{t+1}^{l\prime} = \Pi_t^l = 0$ and $\alpha_t' < \alpha_t^P$, it must be the case that $\Pi_t^{l\prime} > \Pi_t^l = 0$ and therefore $l_t^{l\prime} = 0$ is optimal. A similar argument holds for type h, which implies that $\Pi_t^{h\prime} > \Pi_t^h \geq 0$ and that $l_t^{h\prime} = 0$ is optimal.[4] Third, we have $l_t^{l\prime} = l_t^{h\prime} = 0$ and Bayesian updating implies that $q_t' = q_{t-1}$. Finally, since the continuation game in period $t+1$ in the alternative equilibrium is the same as the continuation game in period t under the original equilibrium, all conditions are satisfied from period $t+1$ onward. Thus, we have constructed an equilibrium which improves the entrepreneur's payoffs. □

Lemma 28. *If* $\Pi_t^l = \Pi_{t+1}^l = 0$ *and*

$$\alpha_t^P \leq \frac{c}{\lambda p_t\left(q_{t-1}\right)V},$$

then there exists another pooling equilibrium which yields the same payoffs to both

[4]Of course, it is possible that $l_t^{h\prime} = l_t^h = 0$, which happens whenever $\Pi_t^h > 0$.

types in period t and which satisfies

$$\alpha_t^P = \frac{c}{\lambda p_t\left(q_t\right)V}.$$

Proof. Consider the following alternative equilibrium. We pick

$$\begin{aligned}
\alpha_t' &= \alpha_t^P \\
q_t' &: \alpha_t^P = \frac{c}{\lambda p_t\left(q_t'\right)V} \\
l_t^{h\prime} &= l_t^h \\
1 - l_t^{l\prime} &= \left(1 - l_t^h\right) q_{t-1}\left(\frac{1-q_t'}{q_t'}\right) \\
\alpha_{t+1}' &= \alpha_{t+1}^P \\
q_{t+1}' &= q_{t+1} \\
l_{t+1}^{h\prime} &= l_{t+1}^h \\
1 - l_{t+1}^{l\prime} &= \left(1 - l_{t+1}^h\right) q_t'\left(\frac{1-q_{t+1}'}{q_{t+1}'}\right).
\end{aligned}$$

That is, we keep the equity share the same, i.e. $\alpha_t' = \alpha_t^P$. However, we change the likelihood of termination $l_t^{l\prime}$ so that the belief q_t' satisfies

$$\alpha_t^P = \frac{c}{\lambda p_t\left(q_t'\right)V}$$

under Bayes' rule.[5] Note that this implies $q_t' \geq q_{t-1}$. In period $t+1$, we keep the equity share and beliefs the same as in the original equilibrium, but we again adjust type l's likelihood of liquidation so that $q_{t+1}' = q_{t+1}$.[6] From period $t+2$ onward, the strategies and beliefs in the alternative equilibrium are the same as in the original one.

[5]This is always possible, since we are considering the case when $l_t^h < 1$.

[6]That is, the investor's beliefs in the alternative and original equilibrium coincide.

Let us confirm that the alternative equilibrium exists. First, since $q_t' \geq q_{t-1}$, the investor's IC condition in Equation (B.2) holds in period t given equity share α_t' and belief q_t'. Similarly, his IC condition in period $t+1$ holds because $q_{t+1}' = q_{t+1}$ and $\alpha_{t+1}' = \alpha_{t+1}^P$. Second, we have $\Pi_{t+1}^{l\prime} = \Pi_{t+1}^l = 0$, which holds because $\alpha_{t+1}' = \alpha_{t+1}^P$ and because the continuation strategies after time $t+1$ are the same as in the original equilibrium. We also have $\Pi_t^{l\prime} = \Pi_t^l = 0$, because $\alpha_t' = \alpha_t^P$ and $\Pi_{t+1}^{l\prime} = \Pi_{t+1}^l$.[7] Since the low type is indifferent in period t, we can freely pick $l_t^{l\prime}$ to ensure that the investor's belief is indeed q_t'. Similarly, we can pick $l_{t+1}^{l\prime}$ such that $q_{t+1}' = q_{t+1}$. □

Now, we consider the case when $\Pi_t^l = 0$ and $\Pi_{t+1}^l > 0$. We will show that either (i) this case is equivalent to the previous one, where $\Pi_t^l = \Pi_{t+1}^l = 0$, or (ii) we can pick an alternative equilibrium in which $U_t = 0$.

Lemma 29. *Suppose that $\Pi_t^l = 0$ and $\Pi_{t+1}^l > 0$. Then, there exists another pooling equilibrium which yields the same payoffs to both types and in which either $\Pi_{t+1}^l = 0$ or $U_t = 0$.*

Proof. Consider the following alternative equilibrium

$$
\begin{aligned}
\alpha_t' &= \alpha_t^P - \varepsilon \\
\alpha_{t+1}' &= \alpha_{t+1}^P + \varepsilon / \left(\delta \left(1 - \lambda\right)\right) \\
q_t' &= q_t \\
q_{t+1}' &= q_{t+1} \\
l_s^{\theta\prime} &= l_s^\theta \text{ for } \theta = l, h \text{ and } s = t, t+1.
\end{aligned}
$$

Let us confirm that this is indeed an equilibrium. Type l's payoff in the alternative

[7] A similar argument for the high type yields $\Pi_t^{h\prime} = \Pi_t^h$ and $\Pi_{t+1}^{h\prime} = \Pi_{t+1}^h$. Thus, both types' payoffs are unchanged.

equilibrium is

$$
\begin{aligned}
\Pi_t^l\left(q_t, \alpha_t'\right) &= \left(1-l_t^{l\prime}\right)\left(\lambda p_t^l\left(1-\alpha_t^p+\varepsilon\right)V-k+\delta\left(1-\lambda p_t^l\right)\Pi_{t+1}^l\left(q_{t+1}, \alpha_{t+1}'\right)\right) \\
&= \left(1-l_t^{l\prime}\right)\left(\lambda p_t^l\left(1-\alpha_t^p+\varepsilon\right)V-k+\delta\left(1-\lambda p_t^l\right)\Pi_{t+1}^l\left(q_{t+1}, \alpha_{t+1}^P\right)\right. \\
&\quad \left. -\delta\left(1-\lambda p_t^l\right)\lambda p_{t+1}^l\frac{\varepsilon V}{\delta\left(1-\lambda\right)}\right).
\end{aligned}
$$

Using the updating rule in Equation (2.1) yields

$$
\begin{aligned}
\Pi_t^l\left(q_t, \alpha_t'\right) &= \left(1-l_t^{l}\right)\left(\lambda p_t^l\left(1-\alpha_t^p+\varepsilon\right)V-k+\delta\left(1-\lambda p_t^l\right)\Pi_{t+1}^l\left(q_{t+1}, \alpha_{t+1}^P\right)\right. \\
&\quad \left. -\delta\left(1-\lambda\right)\lambda p_t^l\frac{\varepsilon V}{\delta\left(1-\lambda\right)}\right) \\
&= \Pi_t^l.
\end{aligned}
$$

Thus, type l receives the same payoff as in equilibrium. A similar calculation for type h yields

$$\Pi_t^h\left(q_t, \alpha_t'\right) = \Pi_t^h.$$

Thus, $l_t^{\theta\prime} = l_t^\theta$ for $\theta = l, h$ is optimal. Since the liquidation probabilities are the same, the beliefs are the same as well, i.e. $q_t' = q_t$. Notice that $\Pi_{t+1}^{\theta\prime}$ is decreasing in ε. Thus, if we pick ε sufficiently large, we have $\Pi_{t+1}^{l\prime} = 0$ and we can then pick $l_{t+1}^{l\prime}$ to ensure that $q_{t+1}' = q_{t+1}$.

Finally, it remains to check whether the investor's incentive compatibility constraint holds in period t in the alternative equilibrium. If this is not true, i.e. for the ε for which $\Pi_{t+1}^l = 0$, we have $U_t < 0$, then, since U_t is continuous in ε, there exists another ε' such that $U_t = 0$. We have thus established the result in the statement of the lemma. □

So far, we have shown that for each t, any optimal pooling contract must feature either $U_t = 0$ or $\alpha_t^p = \frac{c}{\lambda p_t(q_t)V}$. The following Lemma shows concludes this part of our argument by showing that $U_t = 0$ for all t.

Lemma 30. *Suppose that either $U_t = 0$ or $\alpha_t^p = c/(\lambda p_t(q_t)V)$ for all t. Then, for all t, we have $U_t = 0$ and*

$$\alpha_t^p = \frac{c}{\lambda p_t(q_t)V}.$$

Proof. Suppose $U_t > 0$. Then, we have

$$U_t = \delta(1 - \lambda p_t(q_t))U_{t+1}$$

and thus $U_{t+1} > 0$. Proceeding inductively, we must have $U_s > 0$ for all $s \geq t$. But this is impossible. Under the contract α_t^P, the low type will eventually liquidate with probability one, which leaves the investor with a continuation value of zero, since either the game ends or the high type offers his optimal symmetric information contract. Therefore, we must have $U_t = 0$ for all t. But this immediately implies that

$$\alpha_t^p = \frac{c}{\lambda p_t(q_t)V}$$

for all t. □

We have now shown that there is no pooling equilibrium which yields a strictly higher payoff to any type than the equilibrium of Proposition 12. We next show that the equilibrium of Proposition 12 exists. A necessary condition is that given

$$\alpha_t^P = \frac{c}{\lambda p_t(q_t)V},$$

the following conditions are satisfied. For all t and $\theta \in \{l, h\}$,

$$\begin{aligned} \Pi_t^\theta &\geq \Pi_t^\theta(0) \\ \Pi_t^l &> 0 \Rightarrow l_t^h = l_t^l = 0 \\ l_t^\theta &> 0 \Rightarrow \Pi_t^\theta = 0 \end{aligned} \tag{B.3}$$

and q_t satisfies Equation (2.13).[8] The first equation says that deviating to any other contract (in which case we can set the off-path belief to zero) makes each type worse off than staying in equilibrium. The two following equations ensure that the liquidation decisions are optimal for both types. Note that we do not have to consider the investor's incentives, since in this equilibrium, we have $U_t = 0$ for all t.

Lemma 31. *Let $\underline{\tau}^l$ be the first period for which*

$$\lambda p_t^l \left(1 - \frac{c}{\lambda p_t(q_0) V}\right) V - k \leq 0$$

and let $\bar{\tau}^l$ be the first period for which

$$\lambda p_t^l \left(1 - \bar{\alpha}_t^h\right) V - k \leq 0.$$

We have $1 < \underline{\tau}^l \leq \bar{\tau}^l$. Consider the following strategies and beliefs. For any $t < \underline{\tau}^l$, we have $q_t = q_0$ and $l_t^l = l_t^h = 0$. If $\underline{\tau}^l < \bar{\tau}^l$, then for any $\underline{\tau}^l \leq t < \bar{\tau}^l$, we have $l_t^h = 0$, l_t^l satisfies

$$q_t = \frac{q_{t-1}}{q_{t-1} + (1 - q_{t-1})\left(1 - l_t^l\right)},$$

[8] Recall that we are restricting attention to times at which $l_t^l < 1$, since by Lemma 25, the contract is pinned down if $l_t^l = 1$. Thus, Bayes' rule applies.

and q_t satisfies

$$\lambda p_t^l \left(1 - \frac{c}{\lambda p_t\left(q_t\right) V}\right) V = k. \tag{B.4}$$

Finally, we have $l^l_{\bar{\tau}^l} = 1$. Then, the belief q_t is strictly increasing for all $\underline{\tau}^l \leq t < \bar{\tau}^l$ and the conditions in Equation (B.3) *are satisfied.*

Intuitively, we adjust liquidation probabilities so that the low type remains indifferent between continuing and liquidating, given that q_t is updated using Bayes rule. As we show, this is possible and satisfies all relevant incentive constraints.

Proof. Condition (2.8) implies that

$$\lambda p_1^l \left(1 - \alpha_1^P\right) V - k > 0.$$

Thus, $l_1^l = l_1^h = 0$, $q_1 = q_0$, and $\Pi_1^l > 0$. For any t such that $l_s^l = 0$ for all $s \leq t$, we have $q_t = q_0$ and the low type's period payoff is

$$\lambda p_t^l \left(1 - \frac{c}{\lambda p_t\left(q_0\right) V}\right) V - k.$$

This expression crosses zero exactly once from above, since $p_t\left(q_0\right)$ is strictly decreasing in t and vanishes as t becomes large. Thus, $\underline{\tau}^l$ exists and we have $\underline{\tau}^l > 1$. We have $\Pi_t^l > 0$ for $t < \underline{\tau}^l$ and thus $l_t^l = l_t^h = 0$ is optimal for any such t.

A similar argument implies that $\bar{\tau}^l$ exists. That $\underline{\tau}^l \leq \bar{\tau}^l$ is straightforward, because $\alpha_t^P \geq \alpha_t^h$ for all t.

In period $\underline{\tau}^l$, the low type liquidates with strictly positive probability. Here is the argument. If she continues with certainty, then the equilibrium must feature a belief $q_{t+1} = q_0$ and a contract $\alpha_{t+1}^P = c / \left(\lambda p_{t+1}\left(q_0\right) V\right)$, etc. But then, the low type's period payoff for any period $s > t$ in which she continues is strictly negative, so that

$\Pi^l_{t+1} \leq 0$. This, in turn implies that

$$\Pi^l_t = \left(1 - l^l_t\right) \left(\lambda p^l_t \left(1 - \alpha^P_t\right) V - k + \delta \left(1 - \lambda p^l_t\right) \Pi^l_{t+1}\right) < 0,$$

so that the low type's decision to continue in period t must be suboptimal. Thus, no such equilibrium can exist, and we have $\Pi^l_{\underline{\tau}^l} = 0$.

Now, consider the case when $\bar{\tau}^l > \underline{\tau}^l$[9] and recall that by construction of $\bar{\tau}^l$, we have

$$\lambda p^l_{\bar{\tau}^l} \left(1 - \alpha^h_{\bar{\tau}^l}\right) V \leq k. \tag{B.5}$$

Then, for any $\underline{\tau}^l \leq t < \bar{\tau}^l$, there exists a unique belief $q_t \in (q_0, 1)$ such that

$$\lambda p^l_t \left(1 - \frac{c}{\lambda p_t\left(q_t\right) V}\right) V = k. \tag{B.6}$$

Using the updating rule in Equation (2.13), we can find a unique l^l_t which induces belief q_t given q_{t-1}. We can now inductively construct a sequence $\left\{l^l_t, q_t\right\}$ such that Equation (B.6) holds in each period $\underline{\tau}^l \leq t < \bar{\tau}^l$. In period $\bar{\tau}^l$, we pick $l^l_{\bar{\tau}^l} = 1$. If $\underline{\tau}^l = \bar{\tau}^l$, we also pick $l^l_{\bar{\tau}^l} = 1$.

In periods $\underline{\tau}^l \leq t < \bar{\tau}^l$, we have $\Pi^h_t > 0$ and therefore it is optimal for the high type to continue. For the low type, the indifference condition $\Pi^l_t = 0$ must hold. This is true. In period $\bar{\tau}^l$, the low type receives zero value, i.e. $\Pi^l_{\bar{\tau}^l} = 0$, and in any period $\underline{\tau}^l \leq t < \bar{\tau}^l$, Equation (B.6) implies that her period payoff is zero. Backwards induction then implies that $\Pi^l_t = 0$.

In period $\bar{\tau}^l$, it is optimal for the low type to liquidate with certainty. We must distinguish two cases. Suppose that $l^h_{\bar{\tau}^l} < 1$. Then, Equation (2.13) implies that

[9] This implies that $\lambda p^l_t \left(1 - \bar{\alpha}^h_t\right) V > k$ at $t = \underline{\tau}^l$.

$q_{\bar{\tau}^l} = 1$. If the low type continues instead, she receives a payoff of

$$\lambda p^l_{\bar{\tau}^l}\left(1 - \alpha^h_{\bar{\tau}^l}\right) V - k + \delta\left(1 - \lambda p^l_{\bar{\tau}^l}\right)\Pi^l_{\bar{\tau}^l+1}.$$

Equation (B.5) implies that the deviation payoff is negative. Thus, the low type indeed prefers to liquidate. Now, suppose that $l^h_{\bar{\tau}^l} = 1$, so that the game ends with certainty in period $\bar{\tau}^l$. In that case, Equation (B.5) guarantees that the low type's deviation payoff is negative for any off-path belief. $\square$

We have now established that the proposed strategies satisfy the necessary conditions in Equation (B.3).[10] It remains to show that the pooling equilibrium yields at least a weakly higher payoff to both types than any separating equilibrium. To prove the result, we must first characterize separating equilibria. Therefore, we defer this proof. It can be found in Corollary 37 below.

We conclude by showing that q_t is strictly increasing for $\underline{\tau}^l \leq t < \bar{\tau}^l$. This follows, because $p_t(q)$ is strictly decreasing in t for any fixed q and strictly increasing in q for any fixed t. Thus, to satisfy the indifference condition in Equation (B.6) in consecutive periods, q_t must be strictly increasing.

Finally, the equilibrium in Lemma 31 is unique, provided we fix the equity share offered. To establish this, we need the following auxiliary result.

Lemma 32. *For $\underline{\tau}^l \leq t < \bar{\tau}^l$, we have*

$$\alpha^P_t = \frac{\lambda p^l_t V - k}{\lambda p^l_t V}$$

[10] Recall that $\alpha^P_t = \frac{c}{\lambda p_t(q_t)V}$ guarantees that the investor is always willing to experiment. Thus, we do not need to consider the investor's incentive compatibility constraints. Similarly, we do not need to consider deviations in the contract offered, i.e. $\alpha'_t \neq \alpha^P_t$, since the off-path belief $q' = 0$ renders such deviations unprofitable for either type.

and $\Pi_t^l = 0$. *For* $t < \underline{\tau}^l$, *we have*

$$\alpha_t^P < \frac{\lambda p_t^l V - k}{\lambda p_t^l V}.$$

Proof. For $t < \underline{\tau}^l$, the low type continues with certainty and $q_t = q_0$, so that

$$\alpha_t^P = \frac{c}{\lambda p_t\left(q_0\right) V}.$$

The inequality

$$\frac{c}{\lambda p_t\left(q_0\right) V} \geq \frac{\lambda p_t^l V - k}{\lambda p_t^l V},$$

is equivalent to

$$c \geq \lambda p_t\left(q_0\right) V - k\left(q_0 \frac{p_t^h}{p_t^l} + \left(1 - q_0\right)\right). \tag{B.7}$$

Since $p_t\left(q_0\right)$ is decreasing in t and p_t^h / p_t^l is increasing, the RHS is decreasing. We will exploit this fact throughout the proof.

First, we show that there exists no $\underline{\tau}^l \leq t < \bar{\tau}^l$ for which $\Pi_t^l > 0$. Assume towards a contradiction that there exists such a period and let $\hat{t}$ be the largest one. This implies that $\Pi_{\hat{t}+1}^l = 0$, $l_{\hat{t}}^l = 0$, $q_{\hat{t}} = q_{\hat{t}-1}$, and

$$\Pi_{\hat{t}}^l = \lambda p_{\hat{t}}^l\left(1 - \frac{c}{\lambda p_{\hat{t}}\left(q_{\hat{t}-1}\right) V}\right) V - k > 0.$$

By construction of $\underline{\tau}^l$, we have

$$\frac{c}{\lambda p_{\underline{\tau}^l}^l\left(q_0\right) V} \geq \frac{\lambda p_{\underline{\tau}^l}^l V - k}{\lambda p_{\underline{\tau}^l}^l V}.$$

If $\Pi_t^l > 0$ for all $\underline{\tau}^l \leq t < \hat{t}$, then we have $q_{\hat{t}} = q_0$. But since the RHS in Equation (B.7) is decreasing in time, this implies that

$$\frac{c}{\lambda p_{\hat{t}}\left(q_0\right) V} > \frac{\lambda p_{\hat{t}} V - k}{\lambda p_{\hat{t}} V} \tag{B.8}$$

and therefore $\Pi_{\hat{t}}^l < 0$, a contradiction.

Otherwise, there exists a $\underline{\tau}^l \leq \tilde{t} < \hat{t}$ such that $\Pi_{\tilde{t}}^l = 0$ and $\Pi_t^l > 0$ for all $\tilde{t} < t \leq \hat{t}$. Then, we have $l_t^l = 0$ for any such t and $q_{\hat{t}} = q_{\tilde{t}}$. Moreover,

$$\Pi_{\tilde{t}}^l = \lambda p_{\tilde{t}}^l \left(1 - \alpha_{\tilde{t}}^P\right) V - k + \delta \left(1 - \lambda p_{\tilde{t}}^l\right) \Pi_{\tilde{t}+1}^l = 0$$

implies that

$$\frac{c}{\lambda p_{\tilde{t}}^l \left(q_{\tilde{t}}\right) V} \geq \frac{\lambda p_{\tilde{t}}^l V - k}{\lambda p_{\tilde{t}}^l V},$$

since $\Pi_{\tilde{t}+1}^l \geq 0$. Because the belief does not change between $\tilde{t}$ and $\hat{t}$, a variant of Equation (B.7) implies that Inequality (B.8) holds again, and we get a contradiction.

That

$$\alpha_t^P = \frac{\lambda p_t^l V - k}{\lambda p_t^l V}$$

for $\underline{\tau}^l \leq t < \bar{\tau}^l$ follows because $\Pi_t^l = \Pi_{t+1}^l = 0$ for any such t, so that

$$\lambda p_t^l \left(1 - \alpha_t^P\right) V - k + \delta \left(1 - \lambda p_t^l\right) \cdot 0 = 0.$$

To show the second part of the lemma, note simply that by construction, $\underline{\tau}^l$ is the first period in which Inequality (B.7) holds. $\square$

Corollary 33. *Suppose that*

$$\alpha_t^P = \frac{c}{\lambda p_t(q_t)V}.$$

Then, the equilibrium of Lemma 31 is unique, i.e. there does not exist another pooling equilibrium with the same equity share but different liquidation probabilities.

Proof. Any equilibrium which features liquidation before period $\underline{\tau}^l$ violates the entrepreneur's incentive constraints. By Lemma 32, we have for any $\underline{\tau}^l \leq t < \bar{\tau}^l$,

$$\frac{c}{\lambda p_t(q_t)} = \frac{\lambda p_t^l V - k}{\lambda p_t^l V}.$$

Thus, the sequence of beliefs $\{q_t\}$ is unique and so is the sequence of liquidation probabilities $\{l_t^l\}$. Any other choice of liquidation probabilities will either violate the low type's incentive constraint for some t or violate Bayesian updating in Equation (2.13). Finally, there is no equilibrium in which the low type continues past $\bar{\tau}^l$, because even if $q_t = 1$, her value from continuing is negative. □

B.2.3 Proof of Proposition 13

We now construct separating equilibria and show that any separating equilibrium is suboptimal. To show existence, we must ensure that the low type does not mimic the high type and vice versa. For this, we need to consider the continuation payoff of the high type when $q = 0$, i.e. the investor believes he is facing the low type, and the low type's continuation payoff when $q = 1$. If $q = 0$, the high type's continuation contract is $\bar{\alpha}_{t+1}^l$. Any lower share leads to the investor abandoning the project while any higher share is suboptimal.[11] We denote the high type's continuation value in that case as

[11] If $q = 0$, then the investor experiments whenever $\lambda p_t^l(\alpha_t V_t - c) + \delta(1 - \lambda p_t^l)U_{t+1} \geq 0$. The optimal contract for type h induces $U_t = 0$ for all t, just as in the symmetric information benchmark.

$\Pi_{t+1}^h(0) = \Pi_{t+1}^h\left(0, \bar{\alpha}_{t+1}^l\right)$. Similarly, if the low type succeeds in mimicking the high type, he optimally offers $\bar{\alpha}_{t+1}^h$ and receives a value of $\Pi_{t+1}^l(1) = \Pi_{t+1}^l\left(1, \bar{\alpha}_{t+1}^h\right)$. When the investor's beliefs are degenerate, the project is liquidated at a deterministic time. We denote with $\tau^{\theta\prime}$ the liquidation times after a deviation. That is $\tau^{l\prime}$ is the low type's liquidation time if $q' = 1$ and $\tau^{h\prime}$ is the high type's liquidation time when $q' = 0$. We have $\tau^{l\prime} \geq \tau^l$ and $\tau^{h\prime} \leq \tau^h$.

If $t < \tau^l$, combining the two incentive constraints yields the necessary condition

$$\alpha_t^h \in \left[\bar{\alpha}_t^l + \delta(1-\lambda)\frac{\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0)}{\lambda p_{t+1}^l V}, \bar{\alpha}_t^l + \delta(1-\lambda)\frac{\Pi_{t+1}^h(1) - \Pi_{t+1}^h(0)}{\lambda p_{t+1}^h V}\right], \quad \text{(B.9)}$$

while if $\tau^l \leq t < \tau^{l\prime}$, we have

$$\alpha_t^h \in \left[\frac{\lambda p_t^l V - k}{\lambda p_t^l V} + \delta(1-\lambda)\frac{\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0)}{\lambda p_{t+1}^l V}, \bar{\alpha}_t^l + \delta(1-\lambda)\frac{\Pi_{t+1}^h(1) - \Pi_{t+1}^h(0)}{\lambda p_{t+1}^h V}\right], \quad \text{(B.10)}$$

because the low type liquidates if his type is revealed. If $t = \tau^{l\prime}$, the low type liquidates even if she successfully imitates the high type and therefore the high type simply offers the symmetric information contract, i.e. $\alpha_t^h = \bar{\alpha}_t^h$.

Finally, there is no equilibrium in which the high type separates in period $t > \tau^{l\prime}$. Any such equilibrium requires pooling in period $\tau^{l\prime}$. But even if the belief under pooling were $q_{\tau^{l\prime}} = 1$, the low type would liquidate with certainty. Thus, period $t > \tau^{l\prime}$ cannot be reached.

The intervals in Equation (B.9) and (B.10) are nonempty, because the high and low type's values satisfy a variant of single crossing. We prove this in Lemma 34 below.[12]

[12] Specifically, in Equation (B.10), we have $\lambda p_t^l\left(1 - \bar{\alpha}_t^l\right) V \leq k$, which implies that

$$\frac{\lambda p_t^l V - k}{\lambda p_t^l V} \leq \bar{\alpha}_t^l.$$

Lemma 34. *We have*

$$\frac{\Pi_t^h(1) - \Pi_t^h(0)}{p_t^h} \geq \frac{\Pi_t^l(1) - \Pi_t^l(0)}{p_t^l}$$

for all $t < \tau^h$.

Proof. The low type's gain from imitating the high type vs. revealing her type in a given period is

$$\Delta_t^l = \begin{cases} \lambda p_t^l(\bar{\alpha}_t^l - \bar{\alpha}_t^h)V & \text{if } t < \tau^l, \\ \lambda p_t^l\left(1 - \bar{\alpha}_t^h\right)V - k & \text{if } \tau^l \leq t < \tau^{l\prime}, \\ 0 & \text{if } \tau^{l\prime} \leq t. \end{cases} \tag{B.11}$$

If the project succeeds before τ_l, the low type pays the investor $\bar{\alpha}_t^h V$ instead of $\bar{\alpha}_t^l V$. If the project succeeds after τ_l, she receives an additional continuation value since she would have liquidated the project otherwise.

Similarly, the gain for the high type from being indeed perceived as the high type is

$$\Delta_t^h = \begin{cases} \lambda p_t^h(\bar{\alpha}_t^l - \bar{\alpha}_t^h)V & \text{if } t < \tau^{h\prime}, \\ \lambda p_t^h\left(1 - \bar{\alpha}_t^h\right)V - k & \text{if } \tau^{h\prime} \leq t < \tau^h, \\ 0 & \text{if } \tau^h \leq t. \end{cases} \tag{B.12}$$

Thus, we have

$$\Pi_t^l(1) - \Pi_t^l(0) = E_t^l\left[\sum_{s=t}^{\tau^{l\prime}-1} \delta^{s-t}\Delta_s^l\right]$$

Together with Lemma 34, this ensures that the interval in Equation (B.10) is nonempty.

and

$$\Pi_t^h(1) - \Pi_t^h(0) = E_t^h \left[\sum_{s=t}^{\tau^h - 1} \delta^{s-t} \Delta_s^h \right].$$

To prove the result, we distinguish two cases. Suppose first that $\tau^l \leq \tau^{l\prime} \leq \tau^{h\prime} \leq \tau^h$. Then, we have

$$\begin{aligned} \Pi_t^h(1) - \Pi_t^h(0) &\geq E_t^h \left[\sum_{s=t}^{\tau^{h\prime}-1} \delta^{s-t} \Delta_s^h \right] \\ &= \lambda p_t^h \sum_{s=t}^{\tau^{h\prime}-1} \left(\delta\left(1-\lambda\right)\right)^{s-t} (\bar{\alpha}_s^l - \bar{\alpha}_s^h) V, \end{aligned}$$

and

$$\begin{aligned} \Pi_t^l(1) - \Pi_t^l(0) &\leq E_t^l \left[\sum_{s=t}^{\tau^{l\prime}-1} \delta^{s-t} \lambda p_s^l \left(\bar{\alpha}_s^l - \bar{\alpha}_s^h\right) V \right] \\ &= \lambda p_t^l \sum_{s=t}^{\tau^{l\prime}-1} \left(\delta\left(1-\lambda\right)\right)^{s-t} \left(\bar{\alpha}_s^l - \bar{\alpha}_s^h\right) V. \end{aligned}$$

Since $\tau_h' \geq \tau_l'$ and $\bar{\alpha}_t^l > \bar{\alpha}_t^h$ for all t, the above expressions imply

$$\frac{\Pi_t^h(1) - \Pi_t^l(0)}{p_t^h} \geq \frac{\Pi_t^l(1) - \Pi_t^l(0)}{p_t^l},$$

which is what we set out to prove.

Suppose now that $\tau^l \leq \tau^{h\prime} < \tau^{l\prime} \leq \tau^h$. We have

$$\begin{aligned}\Pi_t^h(1) - \Pi_t^h(0) &\geq E_t^h\left[\sum_{s=t}^{\tau^{l\prime}-1} \delta^{s-t}\Delta_s^h\right] \\ &= E_t^h\left[\sum_{s=t}^{\tau^{h\prime}-1} \delta^{s-t}\lambda p_s^h\left(\bar{\alpha}_s^l - \bar{\alpha}_s^h\right)V + \sum_{s=\tau^{h\prime}}^{\tau^{l\prime}-1} \delta^{s-t}\left(\lambda p_s^h\left(1-\bar{\alpha}_s^h\right)V - k\right)\right]\end{aligned}$$

and

$$\begin{aligned}\Pi_t^l(1) - \Pi_t^l(0) &= E_t^l\left[\sum_{s=t}^{\tau^{l\prime}-1} \delta^{s-t}\Delta_s^l\right] \\ &= E_t^l\left[\sum_{s=t}^{\tau^l-1} \delta^{s-t}\lambda p_s^h\left(\bar{\alpha}_s^l - \bar{\alpha}_s^h\right)V + \sum_{s=\tau^l}^{\tau^{l\prime}-1} \delta^{s-t}\left(\lambda p_s^h\left(1-\bar{\alpha}_s^h\right)V - k\right)\right].\end{aligned}$$

If $t \leq s < \tau^l$, both types continue. Then, we have

$$\frac{E_t^h\left[\Delta_s^h\right]}{p_t^h} = (1-\lambda)^{s-t}\lambda\left(\bar{\alpha}_s^l - \bar{\alpha}_s^h\right)V = \frac{E_t^l\left[\Delta_s^l\right]}{p_t^l}.$$

If $\tau^l \leq s < \tau^{h'}$, then the high type always continues, while the low type liquidates if her type is known. Therefore, we have

$$\lambda p_s^l\left(1-\bar{\alpha}_s^l\right)V \leq k$$

and

$$\lambda p_s^l\left(1-\bar{\alpha}_s^h\right)V \geq k,$$

which together imply

$$\lambda p_s^l \left(\bar{\alpha}_s^l - \bar{\alpha}_s^h\right) V \geq \lambda p_s^l \left(1 - \bar{\alpha}_s^h\right) V - k.$$

This inequality, in turn, implies that

$$\begin{aligned}
\frac{E_t^h \left[\Delta_s^h\right]}{p_t^h} &= \frac{1}{p_t^h} \left[\prod_{t \leq u < s-1} \left(1 - \lambda p_u^h\right)\right] \lambda p_s^h \left(\bar{\alpha}_s^l - \bar{\alpha}_s^h\right) V \\
&= (1-\lambda)^{s-t} \lambda \left(\bar{\alpha}_s^l - \bar{\alpha}_s^h\right) V \\
&= \frac{1}{p_t^l} \left[\prod_{t \leq u < s-1} \left(1 - \lambda p_u^l\right)\right] \lambda p_s^l \left(\bar{\alpha}_s^l - \bar{\alpha}_s^h\right) V \\
&\geq \frac{1}{p_t^l} E_t^l \left[\lambda p_s^l \left(1 - \bar{\alpha}_s^h\right) V - k\right] \\
&= \frac{E_t^l \left[\Delta_s^l\right]}{p_t^l}.
\end{aligned}$$

If $\tau^{h'} \leq s < \tau^{l'}$, then both types liquidate if $q_s = 0$ and continue if $q_s = 1$. We have

$$\begin{aligned}
\frac{E_t^h \left[\Delta_s^h\right]}{p_t^h} &= \frac{1}{p_t^h} \left[\prod_{t \leq u < s-1} \left(1 - \lambda p_u^h\right)\right] \left(\lambda p_s^h \left(1 - \bar{\alpha}_s^h\right) V - k\right) \\
&= (1-\lambda)^{s-t} \lambda \left(1 - \bar{\alpha}_s^h\right) V - \frac{1}{p_t^h} \left[\prod_{t \leq u < s-1} \left(1 - \lambda p_u^h\right)\right] k.
\end{aligned}$$

An analog expression holds for the low type. Since $p_t^h > p_t^l$ for all t, we have

$$\frac{1}{p_t^h} \prod_{t \leq u < s-1} \left(1 - \lambda p_u^h\right) < \frac{1}{p_t^l} \prod_{t \leq u < s-1} \left(1 - \lambda p_u^l\right)$$

and therefore

$$\frac{E_t^h \left[\Delta_s^h\right]}{p_t^h} \geq \frac{E_t^l \left[\Delta_s^l\right]}{p_t^l}.$$

Combining the three cases yields the result.

Finally, for $\tau^{l\prime} \leq t < \tau^h$, the result is obvious. The low type always liquidates and receives zero, while the high type continues if his type is known and receives a strictly positive payoff. $\square$

The above Lemma establishes that for each $t < \tau^h$, there is an equilibrium in which the high type separates in period t. In the optimal separating equilibrium, the low type's IC constraint binds, i.e.,

$$\alpha_t^h = \bar{\alpha}_t^l + \delta \frac{1-\lambda}{\lambda V} \frac{\Pi_{t+1}^l (1) - \Pi_{t+1}^l (0)}{p_{t+1}^l} \tag{B.13}$$

if $t < \tau^l$ and[13]

$$\alpha_t^h = \frac{\lambda p_t^l V - k}{\lambda p_t^l V} + \delta \frac{1-\lambda}{\lambda V} \frac{\Pi_{t+1}^l (1)}{p_{t+1}^l} \tag{B.14}$$

if $\tau^l \leq t < \tau^h$.

We now show that in any separating equilibrium, both types would at least weakly prefer to offer the pooling contract. We split the argument into two cases, depending on whether the low type continues once her type is revealed. The low type is at least weakly better off in the pooling equilibrium compared to the separating equilibrium. Thus, we only need to show that the high type is better off.

Lemma 35. *Any equilibrium in which the high type separates in period* $t < \min\{\tau^l, \tau^h - 1\}$ *is suboptimal. The entrepreneur can strictly improve by pooling in period* t *and separating in period* $t+1$.

Proof. The high type's payoff from separating in period t is

$$\Pi_t^h = \lambda p_t^h \left(1 - \alpha_t^h\right) V - k + \delta \left(1 - \lambda p_t^h\right) \Pi_{t+1}^h (1),$$

[13]Note that $\Pi_{t+1}^l (0) = 0$ in this case.

where $\Pi_{t+1}^{h}(1)$ is the symmetric information continuation payoff, which we defined in Section 2.5.1, while her payoff from pooling in period t and separating in period $t+1$ is

$$\Pi_t^{h\prime} = \lambda p_t^h \left(1-\alpha_t^P\right) V - k + \delta \left(1 - \lambda p_t^h\right) \Pi_{t+1}^{h\prime},$$

where $\Pi_{t+1}^{h\prime}$ is her payoff from offering the separating contract in period $t+1$.[14]

Suppose first that $t < \tau^l - 1 \leq \tau^{l\prime}$. Then, the low type continues after her type is revealed, both in the initial separating contract and in the alternative one. Using Equation (B.13), we can write

$$\Pi_t^h = \lambda p_t^h \left(V - \frac{c}{\lambda p_t^l}\right) - k - \delta (1-\lambda) p_t^h \frac{\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0)}{p_{t+1}^l} + \delta \left(1 - \lambda p_t^h\right) \Pi_{t+1}^h(1)$$

and

$$\begin{aligned}
\Pi_t^{h\prime} \;=\;& \lambda p_t^h \left(V - \frac{c}{\lambda p_t(q_t)}\right) - k \\
&+\delta \left(1-\lambda p_t^h\right) \left(\lambda p_{t+1}^h \left(V - \frac{c}{\lambda p_{t+1}^l}\right) - k\right. \\
&\left. -\delta (1-\lambda) p_{t+1}^h \frac{\Pi_{t+2}^l(1) - \Pi_{t+2}^l(0)}{p_{t+2}^l} + \delta \left(1 - \lambda p_{t+1}^h\right) \Pi_{t+2}^h(1)\right).
\end{aligned}$$

We can now plug in the expression

$$\frac{\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0)}{p_{t+1}^l} = \frac{1}{p_{t+1}^l} \left(\lambda p_{t+1}^l \left(\frac{c}{\lambda p_{t+1}^l} - \frac{c}{\lambda p_{t+1}^h}\right) + \delta \left(1 - \lambda p_{t+1}^l\right) \left(\Pi_{t+2}^l(1) - \Pi_{t+2}^l(0)\right)\right),$$

plug in the symmetric information value

$$\Pi_{t+1}^h(1) = \lambda p_{t+1}^h V - c - k + \delta \left(1 - \lambda p_{t+2}^h \Pi_{t+2}^h(1)\right),$$

[14]For $t < \tau^{l\prime}$, the high type does not liquidate when offering the pooling contract. See Proposition 12.

and use the Bayesian updating rule in Equation (2.1). This yields, after some algebra,

$$\Pi_t^{h\prime} - \Pi_t^h = \lambda p_t^h \left(\frac{c}{\lambda p_t\left(q_t\right)} - \frac{c}{\lambda p_t^l} \right) > 0.$$

Thus, pooling in period t and separating in period $t+1$ yields a strictly larger payoff for the high type.

If $t = \tau^l - 1 < \tau^{l\prime}$, the low type liquidates in period $t+1$ if her type is revealed. This changes the high type's payoff from separating later. The separating contract in period $t+1$ is given by Equation (B.14) and we have

$$\frac{\Pi_{t+1}^l\left(1\right) - \Pi_{t+1}^l\left(0\right)}{p_{t+1}^l} = \frac{1}{p_{t+1}^l} \left(\lambda p_{t+1}^l \left(V - \frac{c}{\lambda p_{t+1}^h} \right) - k + \delta \left(1 - \lambda p_{t+1}^l\right) \Pi_{t+2}^l(1) \right).$$

A similar argument as in the previous case yields

$$\Pi_t^{h\prime} - \Pi_t^h = \lambda p_t^h \left(\frac{c}{\lambda p_t^l} - \frac{c}{\lambda p_t(q_t)} \right) > 0.$$

□

Note that the alternative equilibrium we construct is only meaningful if the high type does not liquidate in period $t+1$. This complication occurs when $t = \tau^h - 1$. Then, the pooling and separating contracts coincide, i.e.

$$\alpha_t^P = \alpha^h = \frac{\lambda p_t^l V - k}{\lambda p_t^l V}.$$

This is because we have $\tau^h \geq \tau^{l\prime}$, so the low type will liquidate with certainty in period $t+1$ under both the pooling and separating contracts. In this case, pooling and separating contracts yields the same payoffs to both types, and they both induce liquidation. The distinction in that case is thus purely notational.

Now, we consider the case when the low type liquidates if her type is known and the separating contract is given by Equation (B.14).

Lemma 36. *Any equilibrium in which the high type separates in period $\tau^l \leq t < \tau^{l'}$ is suboptimal. If instead the entrepreneur pools in period t and separates in period $t+1$, her payoff is at least weakly larger.*

Proof. Suppose first that $t < \tau^{l'} - 1$. Using Equation (B.14), the high type's payoff from separating is

$$\Pi_t^h = \lambda p_t^h \left(1 - \frac{\lambda p_t^l V - k}{\lambda p_t^l V}\right) V - k - \delta \left(1 - \lambda\right) p_t^h \frac{\Pi_{t+1}^l \left(1\right)}{p_{t+1}^l} + \delta \left(1 - \lambda p_t^h\right) \Pi_{t+1}^h \left(1\right),$$

and her payoff from pooling in period t and separating in period $t+1$ is

$$\begin{aligned}
\Pi_t^{h\prime} &= \lambda p_t^h \left(1 - \alpha_t^P\right) V - k + \delta \left(1 - \lambda p_t^h\right) \Pi_{t+1}^{h\prime} \\
&= \lambda p_t^h \left(1 - \alpha_t^P\right) V - k \\
&\quad + \delta \left(1 - \lambda p_t^h\right) \left(k \left(\frac{p_{t+1}^h}{p_{t+1}^l} - 1 \right) \right. \\
&\quad \left. - \delta \left(1 - \lambda\right) p_{t+1}^h \frac{\Pi_{t+2}^l \left(1\right)}{p_{t+2}^l} + \delta \left(1 - \lambda p_{t+1}^h\right) \Pi_{t+2}^h \left(1\right) \right).
\end{aligned}$$

Using

$$\Pi_{t+1}^l \left(1\right) = \lambda p_{t+1}^l \left(V - \frac{c}{\lambda p_{t+1}^h}\right) - k + \delta \left(1 - \lambda p_{t+1}^l\right) \Pi_{t+2}^l \left(1\right)$$

and substituting the high type's symmetric information value $\Pi_{t+1}^h \left(1\right)$, we can write

$$\begin{aligned}
\Pi_t^h &= \lambda p_t^h \left(1 - \frac{\lambda p_t^l V - k}{\lambda p_t^l V}\right) V - k \\
&\quad - \delta \left(1 - \lambda\right) p_t^h \left(\lambda V - \frac{c}{p_{t+1}^h} - \frac{k}{p_{t+1}^l} + \delta \left(1 - \lambda\right) \frac{\Pi_{t+2}^l \left(1\right)}{p_{t+2}^l} \right) \\
&\quad + \delta \left(1 - \lambda p_t^h\right) \left(\lambda p_{t+1}^h V - c - k + \delta \left(1 - \lambda p_{t+1}^h\right) \Pi_{t+2}^h \left(1\right) \right),
\end{aligned}$$

which yields

$$\Pi_t^{h\prime} - \Pi_t^h = \lambda p_t^h \left(\frac{\lambda p_t^l V - k}{\lambda p_t^l V} - \alpha_t^P \right) V$$

after some algebra. Lemma 32 then implies that[15] $\Pi_t^{h\prime} \geq \Pi_t^h$.

Finally, if $t = \tau^{l'} - 1$, the low type liquidates in period $t+1$ even after mimicking the high type. In that case, we have $\Pi_{t+1}^l(1) = 0$. A similar calculation as above yields

$$\Pi_t^{h\prime} - \Pi_t^h = \lambda p_t^h \left(\frac{\lambda p_t^l V - k}{\lambda p_t^l V} - \alpha_t^P \right) V - \delta (1-\lambda) \frac{p_t^h}{p_{t+1}^l} \left(\lambda p_{t+1}^l \left(V - \frac{c}{\lambda p_{t+1}^h} \right) - k \right).$$

The first term is positive because of Lemma 32. The second term is positive because the low type prefers to liquidate even if $q_{t+1} = 1$, which implies that

$$\lambda p_{t+1}^l \left(V - \frac{c}{\lambda p_{t+1}^h} \right) \leq k.$$

Thus, we again have $\Pi_t^{h\prime} - \Pi_t^h \geq 0$. □

Corollary 37. *The pooling equilibrium yields an at least weakly higher payoff for both types than any separating equilibrium.*

Proof. We can apply the two previous Lemmas inductively. Separating in period t is Pareto dominated by separating in period $t+1$, which is Pareto dominated by separating in period $t+2$, etc. Thus, separating in period t is Pareto dominated by pooling in period $\tau^{l\prime} - 1$. In period $\tau^{l\prime}$, the low type liquidates with certainty in the pooling equilibrium, so pooling and separating contracts are identical. □

This concludes our proof of Proposition 13.

[15]The inequality binds for $t > \underline{\tau}^l$.

B.2.4 Proof of Proposition 14

Combining the incentive constraints above yields the necessary condition

$$d_t^h \in \left[\Pi_t^l(1) - \Pi_t^l(0), \Pi_t^h(1) - \Pi_t^h(0)\right]. \tag{B.15}$$

In Lemma 34, we have shown that the entrepreneur's values satisfy the single-crossing condition

$$\frac{\Pi_t^h(1) - \Pi_t^h(0)}{p_t^h} \geq \frac{\Pi_t^l(1) - \Pi_t^l(0)}{p_t^l}.$$

Since $p_t^h > p_t^l$, Equation (2.14) implies the set in Equation (B.15) is non-empty. Thus, for any $t < \underline{\tau}^l$, there exists an equilibrium in which the high type separates in period t.

As mentioned in the text, the optimal separating contract must be the one with the lowest cost, i.e.,

$$d_t^h = \Pi_t^l(1) - \Pi_t^l(0).$$

Next, we establish that the pooling equilibrium constructed in Proposition 12 is suboptimal compared to separating in some period τ_S. For $t < \underline{\tau}^l$, any optimal pooling equilibrium must feature $d_t^P = 0$, otherwise, both types could improve by offering no payouts.[16] Thus, the equilibrium of Proposition 12 remains the optimal pooling equilibria.

We first show that the high type prefers to separate rather than continue pooling for any $t \geq \tau_S$.

[16]Recall that the equity share α_t^P provides all incentives to the investor while d_t^P simply serves as a transfer.

Lemma 38. *For any $t < \underline{\tau}^l$, type h prefers to separate with payouts in period t instead of pooling in period t and separating in period $t+1$ if and only if*

$$f_t(q_0) = c\left(\frac{(1-q_0)\,p_t^h}{q_0 p_t^h + (1-q_0)\,p_t^l} - 1\right) - \delta\lambda p_t^h\left(\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0)\right) > 0.$$

Proof. Using Equation (2.15) and the fact that $\alpha_t^h = \bar{\alpha}_t^h$, if the high type separates in period t, her value is

$$\Pi_t^h = \lambda p_t^h\left(V - \frac{c}{\lambda p_t^h}\right) - k + \delta\left(1-\lambda p_t^h\right)\Pi_{t+1}^h(1) - \left(\Pi_t^l(1) - \Pi_t^l(0)\right),$$

while if she pools in period t and separates in period $t+1$, her value is

$$\Pi_t^{h\prime} = \lambda p_t^h\left(V - \frac{c}{\lambda p_t(q_0)}\right) - k + \delta\left(1-\lambda p_t^h\right)\left(\Pi_{t+1}^h(1) - \left(\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0)\right)\right).$$

In the second case, type h offers the optimal pooling contract a_t^p before separation. By construction, no liquidation occurs before τ^l in the pooling equilibrium and therefore $q_t = q_0$. She prefers to separate earlier if and only if

$$\Pi_t^h - \Pi_t^{h\prime} = \lambda p_t^h\left(\frac{c}{\lambda p_t(q_0)} - \frac{c}{\lambda p_t^h}\right) - \left(\Pi_t^l(1) - \Pi_t^l(0)\right) + \delta\left(1-\lambda p_t^h\right)\left(\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0)\right)$$

is positive. Using the fact that

$$\Pi_t^l(1) - \Pi_t^l(0) = \lambda p_t^l\left(\frac{c}{\lambda p_t^l} - \frac{c}{\lambda p_t^h}\right) + \delta\left(1-\lambda p_t^l\right)\left(\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0)\right),$$

we have, after some algebra,

$$\frac{p_t^h}{p_t^h - p_t^l}\left(\Pi_t^h - \Pi_t^{h\prime}\right) = c\left(\frac{(1-q_0)\,p_t^h}{q_0 p_t^h + (1-q_0)\,p_t^l} - 1\right) - \delta\lambda p_t^h\left(\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0)\right) = f_t(q_0). \tag{B.16}$$

Since $p_t^h > p_t^l$, $\Pi_t^h - \Pi_t^{h\prime}$ is positive if and only if $f_t(q_0)$ is positive. $\square$

The next lemma establishes the monotonicity of $f_t(\cdot)$ given the initial belief q_0.

Lemma 39. *Given q_0, f_t strictly increases in $1 \le t < \underline{\tau}^l$, i.e.,*

$$f_1(q_0) < f_2(q_0) < ... < f_{\underline{\tau}^l - 1}(q_0).$$

Proof. We first show that p_t^l / p_t^h decreases in t. This is because

$$\frac{p_{t+1}^l}{p_{t+1}^h} = \frac{p_t^l}{p_t^h} \frac{1 - \lambda p_t^h}{1 - \lambda p_t^l} < \frac{p_t^l}{p_t^h}.$$

This implies

$$\frac{1 - q_0}{q_0 + (1 - q_0) \frac{p_t^l}{p_t^h}} < \frac{1 - q_0}{q_0 + (1 - q_0) \frac{p_{t+1}^l}{p_{t+1}^h}},$$

which is equivalent to

$$\frac{(1 - q_0) p_t^h}{q_0 p_t^h + (1 - q_0) p_t^l} < \frac{(1 - q_0) p_{t+1}^h}{q_0 p_{t+1}^h + (1 - q_0) p_{t+1}^l}.$$

For the second part, we first generate an upper bound of $\delta \left(\Pi_t^l(1) - \Pi_t^l(0) \right)$. To start, notice that for $t < \underline{\tau}^l$,

$$\frac{c}{\lambda p_{t+1}^l} - \frac{c}{\lambda p_{t+1}^h} = \frac{c \left(1 - \lambda p_t^h\right)}{(1 - \lambda) p_t^h} - \frac{c \left(1 - \lambda p_t^l\right)}{(1 - \lambda) p_t^l} = \frac{1}{1 - \lambda} \left(\frac{c}{\lambda p_t^l} - \frac{c}{\lambda p_t^h} \right),$$

and for $t \ge \underline{\tau}^l$, we have

$$\lambda p_t^l \left(V - \frac{c}{\lambda p_t^l} \right) - k \le 0 = \lambda p_t^l \left(\frac{c}{\lambda p_t^h} - \frac{c}{\lambda p_t^h} \right),$$

since type l liquidates when she reveals her type, so that

$$V - \frac{c}{\lambda p_t^h} - \frac{k}{\lambda p_t^l} \leq \frac{c}{\lambda p_t^l} - \frac{c}{\lambda p_t^h}.$$

Therefore,[17]

$$\begin{aligned}
& \delta\left(\Pi_t^l(1) - \Pi_t^l(0)\right) \\
= \; & \delta\lambda p_t^l \left(\sum_{s=t}^{\underline{\tau^l}-1} (\delta(1-\lambda))^{s-t} \left(\frac{c}{\lambda p_s^l} - \frac{c}{\lambda p_s^h} \right) \right. \\
& \left. + \sum_{s=\underline{\tau^l}}^{\tau^{l\prime}-1} (\delta(1-\lambda))^{s-t} \left(V - \frac{c}{\lambda p_s^h} - \frac{k}{\lambda p_s^l} \right) \right) \\
\leq \; & \delta\lambda p_t^l \left(\frac{c}{\lambda p_t^l} - \frac{c}{\lambda p_t^h} \right) \left(\sum_{s=t}^{\tau^{l\prime}-1} \delta^{s-t} \right) \\
\leq \; & \frac{\delta}{1-\delta} \lambda p_t^l \left(\frac{c}{\lambda p_t^l} - \frac{c}{\lambda p_t^h} \right).
\end{aligned}$$

Second, using a similar derivation, if $\underline{\tau^l} < \tau^{l\prime}$ or if $t < \underline{\tau^l} - 1$, we have

$$\begin{aligned}
& \delta p_t^h \left(\Pi_t^l(1) - \Pi_t^l(0)\right) - \delta p_{t+1}^h \left(\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0)\right) \\
= \; & \delta\lambda p_t^h p_t^l \left(\frac{c}{\lambda p_t^l} - \frac{c}{\lambda p_t^h} \right) - \left(1 - \delta \frac{\left(1-\lambda p_t^h\right)\left(1-\lambda p_t^l\right)}{1-\lambda} \right) \delta p_{t+1}^h \left(\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0)\right) \\
\geq \; & \delta\lambda p_t^h p_t^l \left(\frac{c}{\lambda p_t^l} - \frac{c}{\lambda p_t^h} \right) - \left(1 - \delta \frac{\left(1-\lambda p_t^h\right)\left(1-\lambda p_t^l\right)}{1-\lambda} \right) \delta\lambda p_{t+1}^h p_{t+1}^l \left(\frac{c}{\lambda p_{t+1}^l} - \frac{c}{\lambda p_{t+1}^h} \right) \\
= \; & \frac{\delta}{1-\delta} \lambda p_t^h p_t^l \left(\frac{c}{\lambda p_t^l} - \frac{c}{\lambda p_t^h} \right) \frac{1-\lambda}{\left(1-\lambda p_t^h\right)\left(1-\lambda p_t^l\right)} \left(\frac{\left(1-\lambda p_t^h\right)\left(1-\lambda p_t^l\right)}{1-\lambda} - 1 \right) \\
> \; & 0.
\end{aligned}$$

The inequality comes from Assumption 5. If $\underline{\tau^l} = \tau^{l\prime}$ and $t = \underline{\tau^l} - 1$, then $\Pi_{t+1}^l(1) -$

[17] If $\underline{\tau^l} > \tau^{l\prime} - 1$, we use the convention $\sum_{s=\underline{\tau^l}}^{\tau^{l\prime}-1}(...) = 0$.

$\Pi^l_{t+1}(0) = 0$ and the expression is still positive. Together with the previous result, this generates the Lemma statement. $\square$

Next, we will characterize τ_S for two cases: $\underline{\tau}^l = \tau^{l\prime}$ and $\underline{\tau}^l < \tau^{l\prime}$.

Lemma 40. *If $\underline{\tau}^l = \tau^{l\prime}$, there exists a threshold $\bar{q}$ such that the high type always prefers to pool if and only if $q_0 \geq \bar{q}$.*

Proof. If $\underline{\tau}^l = \tau^{l\prime}$, then

$$f_{\underline{\tau}^l-1}(q_0) = c\left(\frac{(1-q_0)\,p^h_{\underline{\tau}^l-1}}{q_0 p^h_{\underline{\tau}^l-1} + (1-q_0)\,p^l_{\underline{\tau}^l-1}} - 1\right).$$

If $q_0 > 0.5$, the above expression is strictly negative. By continuity, there exists a $\bar{q} < 0.5$ such that $f_{\underline{\tau}^l-1}(q_0) \leq 0$ if and only if $q_0 \geq \bar{q}$. Since $f_t(q_0)$ is strictly increasing in t, we have $f_t(q_0) < 0$ for all $t < \underline{\tau}^l - 1$. Thus, for any equilibrium with separation in period $t \leq \underline{\tau}^l - 1$, we can increase the high type's payoff by separating in period $t+1$ instead. Applying the argument inductively implies that the high type prefers to never separate. $\square$

Corollary 41. *If $\underline{\tau}^l = \tau^{l\prime}$ and $q_0 \geq \bar{q}$, the optimal contract is pooling.*

Lemma 42. *If $\underline{\tau}^l = \tau^{l\prime}$ and $q_0 < \bar{q}$, we have $1 \leq \tau_S \leq \underline{\tau}^l - 1$.*

Proof. Define $\tau_S = \min\{t | f_t(q_0) \geq 0\}$. Since $q_0 < \bar{q}$, this set is non-empty, and we have $\tau_S \leq \underline{\tau}^l - 1$. $\square$

Now, consider the case $\underline{\tau}^l < \tau^{l\prime}$. In Section Appendix B.2.2, we have shown that separating through α^h_t has the same payoff as the pooling equilibrium for the high type when $\underline{\tau}^l \leq t < \tau^{l\prime}$. Since separating through payouts is cheaper, it is now preferred by the high type.

Lemma 43. *If $\underline{\tau}^l < \tau^{l\prime}$, the high type strictly prefers separating in period $\underline{\tau}^l$ and we have $\tau_S \leq \underline{\tau}^l$.*

Proof. For all $t \geq \underline{\tau}^l$, we can replicate the argument of Lemma 38. The difference is that separating in period $t+1$ has the following payoff structure, due to the change of α_t^p:

$$\lambda p_t^h \left(V - \frac{\lambda p_t^l V - k}{\lambda p_t^l}\right) - k + \delta \left(1 - \lambda p_t^h\right) \left(\Pi_{t+1}^h (1) - \left(\Pi_{t+1}^l (1) - \Pi_{t+1}^l (0)\right)\right).$$

Besides,

$$\Pi_t^l (1) - \Pi_t^l (0) = \lambda p_t^l \left(V - \frac{c}{\lambda p_t^h}\right) - k + \delta \left(1 - \lambda p_t^l\right) \left(\Pi_{t+1}^l (1) - \Pi_{t+1}^l (0)\right).$$

The difference in value for type h from separating in period t vs. period $t+1$ is

$$\Pi_t^{h\prime} - \Pi_t^h = \frac{\lambda \left(p_t^h - p_t^l\right)}{\lambda p_t^l} \left(\lambda p_t^l \left(V - \frac{c}{\lambda p_t^h}\right) - k\right) - \delta\lambda \left(p_t^h - p_t^l\right) \left(\Pi_{t+1}^l (1) - \Pi_{t+1}^l (0)\right).$$

Using the same argument as in Lemma 39, we can show

$$\Pi_{t+1}^l (1) - \Pi_{t+1}^l (0) < \frac{1}{\lambda} \left(\lambda p_{t+1}^l \left(V - \frac{c}{\lambda p_{t+1}^h}\right) - k\right) < \frac{1}{\lambda p_t^l} \left(\lambda p_t^l \left(V - \frac{c}{\lambda p_t^h}\right) - k\right).$$

Therefore, $\Pi_t^{h\prime} - \Pi_t^h$ is strictly positive for all $t \geq \underline{\tau}^l$. Define $\tau_S = \min\{t | f_t(q_0) \geq 0, t \leq \underline{\tau}^l - 1\} \cup \{\underline{\tau}^l\}$, then the high type will optimal separate at τ_S. □

Next, we provide a sufficient condition for when $\underline{\tau}^l = \tau^{l\prime}$.

Lemma 44. *We have $\underline{\tau}^l = \tau^{l\prime}$ if and only if q_0 exceeds a threshold $\bar{q}'$.*

Proof. Given p_1^l and p_1^h, by Proposition 12, $\underline{\tau}^l$ is the first period when

$$\lambda p_t^l \left(V - \frac{c}{\lambda p_t(q_0)}\right) - k \leq 0.$$

Since $p_t(q_0)$ increases in q_0 and $\lim_{q_0 \to 1} p_t(q_0) = p_t^h$, $\underline{\tau}^l$ increases in q_0 and converges to $\tau^{l\prime}$. By continuity, there exists $\bar{q}'$ such that $\underline{\tau}^l = \tau^{l\prime}$ for all $q_0 > \bar{q}'$. □

Combining Lemma 40 and 44, if $q_0 > \max\{\bar{q}, \bar{q}'\}$, the high type optimally chooses to pool. The low type, of course, prefers to never separate. The optimal equilibrium is pooling regardless of γ. This equilibrium can also be interpreted as "separating" at $\bar{\tau}^l$. In that period, the low type liquidates with certainty in the pooling equilibrium, so pooling and separating are equivalent. If $q_0 < \max\{\bar{q}, \bar{q}'\}$, there exists an optimal timing τ_S when the high type prefers to separate. Now we are optimizing over the finite set $\left\{\tau_S, \bar{\tau}^l\right\}$, an optimal separating time exists. Clearly, whenever γ is sufficiently small, pooling is optimal while whenever γ is sufficiently large, the optimal separating time satisfies either $f_{\tau_S}(q_0) > 0$ or $\tau_S = \underline{\tau}^l$.

Corollary 45. *The separation time in equilibrium weakly increases in q_0.*

Proof. Denote $\underline{\tau}^l$ as a function of q_0 by $\underline{\tau}^l(q_0)$. Then $\underline{\tau}^l(q_0)$ is a weakly increasing step function. We start from $q_0 < \min\{\bar{q}, \bar{q}'\}$. Notice $f_t(q_0)$ strictly decreases in q_0 for all $t \leq \tau^{l\prime} - 1$, therefore the first time when $f_t(q_0)$ crosses 0 from below weakly delays. If $f_t(q_0) < 0$ for all $t \leq \underline{\tau}^l(q_0) - 1$, then the separating time is $\underline{\tau}^l(q_0)$, which still weakly increases in q_0. This process continues until $q_0 = \bar{q}'$ and $\underline{\tau}^l(q_0) = \tau^{l\prime}$, then the separation timing weakly increases from $\tau^{l\prime} - 1$ to $\tau^{l\prime}$, depending on the relationship between $\bar{q}$ and $\bar{q}'$. $\square$

B.2.5 Proof of Proposition 15

Since the proof is similar to the proof of Proposition 14, we only provide a sketch.

Lemma 46. *With observable types, the high type pivots in period τ^h, which is the first period for which*

$$\Pi_1^h(1) - F \geq \Pi_t^h(1) \tag{B.17}$$

while the low type pivots in period τ^l,which is the first period for which

$$\Pi_1^l(0) - F \geq \Pi_t^l(0). \tag{B.18}$$

That τ^h and τ^l satisfy Equations (B.17) and (B.18) is straightforward, and we omit the proof.

Lemma 47. *In the optimal pooling equilibrium, the low type liquidates with positive probability l_t^l whenever $\underline{\tau}^l \leq t \leq \bar{\tau}^l$. The equity share is given by*

$$\alpha_t^P = \frac{c}{\lambda p_t(q_t) V}.$$

The period $\underline{\tau}^l$ is the first period in which

$$\Pi_t^l = \Pi_1^l(0) - F.$$

The proof is analogous to the proof of Proposition 12 and hence omitted. The only difference is that when the low type liquidates with positive probability, her value satisfies

$$\lambda p_t^l \left(1 - \alpha_t^P\right) V - k + \delta\left(1 - \lambda p_t^l\right)\left(\Pi_1^l(0) - F\right) = \Pi_1^l(0) - F$$

and the pooling share α_t^P is now given by

$$\alpha_t^P = \frac{\lambda p_t^l V - k}{\lambda p_t^l V} - \frac{\Pi_1^l(0) - F}{\lambda p_t^l V} \frac{1 - \delta\left(1 - \lambda p_t^l\right)}{1 - \delta}$$

for $\underline{\tau}^l \leq t \leq \bar{\tau}^l$.

The following Lemma allows us to restrict attention to contracts which separate before period τ^l.

Lemma 48. *We have $\tau^l \leq \underline{\tau}^l$. Separating via pivots is not incentive compatible in any period $t \geq \tau^l$.*

Proof. That $\tau^l \leq \underline{\tau}^l$ follows from the fact that $\Pi_t^l \geq \Pi_t^l(0)$ for all t. We have for $t \geq \tau^l$, $\Pi_1^l(1) - F \geq \Pi_1^l(0) - F \geq \Pi_t^l(0)$, which implies that separating in period $t \geq \tau^l$ is not incentive compatible. □

We next show that the set $\left[\Pi_1^l(1) - \Pi_t^l(0), \Pi_1^h(1) - \Pi_t^h(0)\right]$ in Equation (2.16) is nonempty and characterize how this set changes over time.

Lemma 49. *For δ sufficiently small, $\Pi_t^h(1) - \Pi_t^l(1)$ is decreasing in t and for all t, we have*

$$\Pi_1^l(1) - \Pi_t^l(0) \leq \Pi_1^h(1) - \Pi_t^h(0)$$

Proof. We have[18]

$$\begin{aligned}\Pi_t^h(1) - \Pi_t^l(1) \quad = \quad & \sum_{s=t}^{\tau_h} \delta^{(s-t)} \left[\prod_{t \leq u < s} \left(1 - \lambda p_u^h\right)\right] \left(\lambda p_s^h \left(1 - \bar{\alpha}_s^h\right) V - k\right) \\ & - \sum_{s=t}^{\tau_l'} \delta^{(s-t)} \left[\prod_{t \leq u < s} \left(1 - \lambda p_u^l\right)\right] \left(\lambda p_s^l \left(1 - \bar{\alpha}_s^h\right) V - k\right).\end{aligned}$$

Since $\tau_l' \leq \tau_h$, we have for any $t \leq \tau_l'$

$$\Pi_t^h(1) - \Pi_t^l(1) = \lambda \left(p_t^h - p_t^l\right)\left(1 - \bar{\alpha}_t^h\right) V + \delta \left(\left(1 - \lambda p_t^h\right) \Pi_{t+1}^h(1) - \left(1 - \lambda p_{t+1}^l\right) \Pi_{t+1}^l\right).$$

As δ becomes small, this expression converges to

$$\Pi_t^h(1) - \Pi_t^l(1) = \lambda \left(p_t^h - p_t^l\right)\left(1 - \bar{\alpha}_t^h\right) V.$$

[18] Recall that τ_l' is defined as the time at which the low type liquidates when the investor's belief is $q_t = 1$.

Then,

$$\Pi_t^h\left(1\right)-\Pi_t^l\left(1\right)-\left(\Pi_t^h\left(1\right)-\Pi_{t+1}^l\left(1\right)\right)=\lambda V\left(p_t^h-p_t^l-\left(p_{t+1}^h-p_{t+1}^l\right)\right)+c\left(\frac{p_t^l}{p_t^h}-\frac{p_{t+1}^l}{p_{t+1}^h}\right).$$

This expression is positive, because by Assumption 5, $p_t^h - p_t^l$ is decreasing in t and because p_t^l/p_t^h is also decreasing in t, which can be seen from Bayes' rule.

To show that

$$\Pi_1^l(1) - \Pi_t^l(0) \leq \Pi_1^h(1) - \Pi_t^h(0),$$

we can rearrange this expression as

$$\Pi_t^h\left(0\right) - \Pi_t^l\left(0\right) \leq \Pi_1^h\left(1\right) - \Pi_1^l\left(1\right).$$

We have $\Pi_t^h\left(0\right) - \Pi_t^l\left(0\right) \leq \Pi_t^h\left(1\right) - \Pi_t^l\left(1\right)$, which follows from a similar argument as in Lemma 34. Thus, a sufficient condition for the inequality to hold is that

$$\Pi_t^h\left(1\right) - \Pi_t^l\left(1\right) \leq \Pi_1^h\left(1\right) - \Pi_1^l\left(1\right),$$

which in turn holds if $\Pi_t^h\left(1\right) - \Pi_t^l\left(1\right)$ is decreasing in t, which we just established. □

Corollary 50. *For F sufficiently small, there exists a period t, such that*

$$F \in \left[\Pi_1^l(1) - \Pi_t^l(0), \Pi_1^h(1) - \Pi_t^h(0)\right].$$

This result implies that there are two periods $1 \leq \underline{\tau}_{Piv} \leq \bar{\tau}_{Piv}$, such that separating via a pivot is feasible whenever $\underline{\tau}_{Piv} \leq t \leq \bar{\tau}_{Piv}$. Intuitively, for $t < \underline{\tau}_{Piv}$, separation is not feasible, because it is too costly for the high type. As t increases beyond $\bar{\tau}_{Piv}$, we eventually have $F < \Pi_1^l(1) - \Pi_t^l(0)$, so that separation is not incentive-compatible, because the low type would pivot together with the high type.

Lemma 51. *For δ and F sufficiently small, there exists a period $\underline{\tau}_{Piv} \leq \tau_S^h \leq \bar{\tau}_{Piv}$, such that the high type prefers to separate by pivoting in period τ_S^h and prefers to continue pooling for all $t < \tau_S^h$.*

Proof. Lemma 48 implies that separation is not incentive compatible after period τ^l. Thus, we can wlog restrict attention to the case when $\bar{\tau}_{Piv} \leq \tau^l \leq \underline{\tau}^l$. This implies that at any time at which we consider separation, we have $q_t = q_0$.

If the high type separates in period t,her payoff is $\Pi_1^h(1) - F$, while if she pools in period t and separates in period $t+1$, her payoff is

$$\lambda p_t^h \left(1 - \alpha_t^P\right) V - k + \delta \left(1 - \lambda p_t^h\right) \left(\Pi_1^h(1) - F\right).$$

The high type prefers separating in period t rather than in period $t+1$ whenever

$$f_t := \left(1 - \delta\left(1 - \lambda p_t^h\right)\right) \left(\Pi_1^h(1) - F\right) - \left(\lambda p_t^h \left(1 - \alpha_t^P\right) V - k\right) \geq 0.$$

Since we have $t \leq \underline{\tau}^l$, we can write

$$\begin{aligned} \lambda p_t^h \left(1 - \alpha_t^P\right) V &= \lambda p_t^h V - c \frac{p_t^h}{p_t(q_0)} \\ &= \lambda p_t^h V - c \frac{1}{q_0 + (1 - q_0) \frac{p_t^l}{p_t^h}}, \end{aligned}$$

which is strictly decreasing in t, because both p_t^h and p_t^l / p_t^h are strictly decreasing. For δ sufficiently small, f_t is then strictly increasing in t and crosses zero at most once. Let τ_S^h denote the crossing time. If τ_S^h does not exist, then the high type prefers to never separate via a pivot.

We now show that τ_S^h exists. Pick a period t such that both Equation (IC_h^{Piv}) and Equation (IC_l^{Piv}) hold. By Corollary 50, such a period exists whenever F is sufficiently small. Using Equation (IC_l^{Piv}), we have

$$\begin{aligned} f_t \geq & \left(1-\delta\left(1-\lambda p_t^h\right)\right)\left(\Pi_1^h(1)-\left(\Pi_1^l(1)-\Pi_t^l(0)\right)\right) \\ & -\left(\lambda p_t^h\left(1-\alpha_t^P\right) V-k\right). \end{aligned}$$

As δ becomes small, the RHS approaches

$$\lambda\left(p_1^h-p_1^l\right)\left(1-\bar{\alpha}_1^h\right) V-\lambda\left(p_t^h-p_t^l\right) V+c\left(\frac{p_t^h}{p_t\left(q_0\right)}-1\right).$$

By Assumption 5, $p_t^h - p_t^l$ is decreasing. Thus, the expression is positive, which implies that $f_t \geq 0$ and that τ_S^h exists.

That τ_S^h is the optimal time to separate for the high type follows from a similar argument as in the proof of Proposition 14. Specifically, for $t < \tau_S^h$, the high type prefers to separate later, while for $t > \tau_S^h$, the prefers to separate earlier. Thus, separating in period τ_S^h maximizes her ex-ante value. □

The following Lemma concludes our argument.

Lemma 52. *For γ sufficiently large, the optimal contract features separation via pivots.*

The proof follows the same lines as the proof of Proposition 14. The low type prefers to never separate via a pivot. Instead, she prefers to only pivot in period τ^l. The high type prefers separating via a pivot in period τ_S^h. For γ large, there exists a period τ_S such that $\underline{\tau}_{Piv} \leq \tau_S \leq \bar{\tau}_{Piv}$, so that pivoting in period τ_S is optimal.

B.2.6 Proof of Proposition 16

The proof of Proposition 16 is similar to the proof of Proposition 14. We hence only provide a sketch.

Under symmetric information, each type liquidates at time τ^{θ}. We have $\tau^{l} \leq \tau^{h}$. Each type implements the prestige project at the time of liquidation, since by doing so she receives a higher outside option and does not need to suffer the decrease in project value. Thus, each type liquidates at the first time at which

$$\Pi_t^{\theta} \leq \pi.$$

A pooling equilibrium exists and has the same features as the equilibrium in Proposition 12. Specifically, the equity share is still given by

$$\alpha_t^P = \frac{c}{\lambda p_t(q_t) V},$$

and there exist two times $\underline{\tau}^l \leq \bar{\tau}^l$, such that in any period between $\underline{\tau}^l$ and $\bar{\tau}^l$, the low type randomizes between implementing the prestige project and liquidating or continuing. In any such period, the following indifference condition holds:

$$\lambda p_t^l \left(1 - \alpha_t^P\right) V - k + \delta \left(1 - \lambda p_t^l\right) \pi = \pi. \tag{B.19}$$

This equation is the analog of Equation (B.4) in the baseline model. It implies that the pooling equity share satisfies

$$\alpha_t^P = \frac{\lambda p_t^l V - k}{\lambda p_t^l V} - \pi \frac{1 - \delta \left(1 - \lambda p_t^l\right)}{\lambda p_t^l}. \tag{B.20}$$

We now consider a separating equilibrium in period t. Using the IC conditions for the low and high type, Equations (IC_l^{Pres}) and (IC_h^{Pres}), we can see that separating is incentive compatible whenever

$$\frac{\Pi_t^l(1) - \Pi_t^l(0)}{p_t^l} \leq \lambda V_0 \leq \frac{\Pi_t^h(1) - \Pi_t^h(0)}{p_t^h}.$$

A similar argument as in Lemma 34 implies that

$$\frac{\Pi_t^l(1)-\Pi_t^l(0)}{p_t^l}\leq\frac{\Pi_t^h(1)-\Pi_t^h(0)}{p_t^h}.$$

Thus, for any period t, there exists a V_0, such that we can achieve separation in that period. It only remains to find conditions such that separation is optimal.

Following a similar argument as in Lemma 38, consider the high type's value from separating in period t versus separating in period $t+1$. If the high type separates in period $t<\bar{\tau}^l-1$, her value is

$$\Pi_t^h(1)-\lambda p_t^h V_0,$$

while if she separates in period $t+1$, her value in period t is

$$\lambda p_t^h\left(1-\alpha_t^P\right)V-k+\delta\left(1-\lambda p_t^h\right)\left(\Pi_{t+1}^h(1)-(1-\delta)\,\lambda p_{t+1}^h V_0\right).$$

After some algebra, the difference between these two values is

$$f_t=c\left(\frac{p_t^h}{p_t\left(q_t\right)}-1\right)-\lambda p_t^h V_0\left(1-\delta\left(1-\lambda\right)\right).$$

Whenever this expression is positive, the high type prefers separating in period t rather than in period $t+1$.

For $t=\bar{\tau}^l-1$, the high type knows that the low type will liquidate in period $\bar{\tau}^l$. Thus, separating in period $\bar{\tau}^l$ does not require costly signaling, and we have

$$f_{\bar{\tau}^l-1}=c\left(\frac{p_{\bar{\tau}^l-1}^h}{p_{\bar{\tau}^l-1}\left(q_{\bar{\tau}^l-1}\right)}-1\right)-\lambda p_{\bar{\tau}^l-1}^h V_0$$

Next, we find a sufficient condition, such that $f_t > 0$. Consider the region $\underline{\tau}^l \leq t \leq \bar{\tau}^l$. On this region, the low type will liquidate if her type is discovered. Thus, the low type's IC condition becomes

$$\Pi_t^l(1) - \lambda p_t^l V_0 \leq \pi,$$

since we have $\Pi_t^l(0) = \pi$. Let us pick V_0 such that the above inequality binds. Then, the high type prefers separating in period t over separating in the next period whenever

$$\hat{f}_t = c\left(\frac{p_t^h}{p_t(q_t)} - 1\right) - \frac{p_t^h}{p_t^l}\left(\Pi_t^l(1) - \pi\right)\left(1 - \delta(1-\lambda)\right) \geq 0$$

for $t < \bar{\tau}^l - 1$ or

$$\hat{f}_{\bar{\tau}^l-1} = c\left(\frac{p_{\bar{\tau}^l-1}^h}{p_{\bar{\tau}^l-1}(q_{\bar{\tau}^l-1})} - 1\right) - \frac{p_{\bar{\tau}^l-1}^h}{p_{\bar{\tau}^l-1}^l}\left(\Pi_{\bar{\tau}^l-1}^l(1) - \pi\right) \geq 0.$$

Here is the significance of $\hat{f}_t$. Whenever $\hat{f}_t$ is positive, there exists a V_0, such that f_t is positive. Thus, $\hat{f}_t > 0$ is a necessary and sufficient condition for the existence of a V_0 such that the high type prefers to separate in period t rather than in period $t+1$.

Pick period $t = \bar{\tau}^l - 1$. The low type liquidates in period $t+1 = \bar{\tau}^l$, even if the investor's belief is $q_{t+1} = 1$. Thus, we have

$$\Pi_{\bar{\tau}^l-1}^l(1) = \lambda p_{\bar{\tau}^l-1}^l V - c\frac{p_{\bar{\tau}^l-1}^l}{p_{\bar{\tau}^l-1}^h} - k + \delta\left(1 - \lambda p_{\bar{\tau}^l-1}^l\right)\pi.$$

Plugging this expression into $\hat{f}_t$, we get,

$$\begin{aligned}
\hat{f}_{\bar{\tau}^l-1} &= -\frac{p^h_{\bar{\tau}^l-1}}{p^l_{\bar{\tau}^l-1}}\left(\lambda p^l_{\bar{\tau}^l-1}V - c\frac{p^l_{\bar{\tau}^l-1}}{p_{\bar{\tau}^l-1}\left(q_{\bar{\tau}^l-1}\right)} - k + \delta\left(1-\lambda p^l_{\bar{\tau}^l-1}\right)\pi - \pi\right)\\
&= -\frac{p^h_{\bar{\tau}^l-1}}{p^l_{\bar{\tau}^l-1}}\left(\Pi^l_{\bar{\tau}^l-1} - \pi\right) = 0,
\end{aligned}$$

where $\Pi^l_{\bar{\tau}^l-1}$ is the low type's value in period $\bar{\tau}^l-1$ if there is pooling. Since $\bar{\tau}^l-1 \geq \underline{\tau}^l$, we have $\Pi^l_{\bar{\tau}^l-1} = \pi$.

For any $\underline{\tau}^l \leq t < \bar{\tau}^l - 1$, we can write

$$\begin{aligned}
\hat{f}_t &= c\left(\frac{p^h_t}{p_t\left(q_t\right)} - 1\right) - \frac{p^h_t}{p^l_t}\left(\Pi^l_t\left(1\right) - \pi\right)\left(1-\delta\left(1-\lambda\right)\right)\\
&= c\left(\frac{p^h_t}{p_t\left(q_t\right)} - 1\right) - \frac{p^h_t}{p^l_t}\left(\lambda\left(p^l_tV - c\frac{p^l_t}{p^h_t} - k\right) + \delta\left(1-\lambda p^l_t\right)\Pi^l_{t+1}\left(1\right) - \pi\right)\\
&\quad +\delta\left(1-\lambda\right)\frac{p^h_t}{p^l_t}\left(\Pi^l_t\left(1\right) - \pi\right)\\
&= -\frac{p^h_t}{p^l_t}\left(\lambda p^l_tV - c\frac{p^l_t}{p_t\left(q_t\right)} - k + \delta\left(1-\lambda p^l_t\right)\Pi^l_{t+1}\left(1\right) - \pi\right)\\
&\quad +\delta\left(1-\lambda\right)\frac{p^h_t}{p^l_t}\left(\Pi^l_t\left(1\right) - \pi\right)\\
&= \delta\frac{p^h_t}{p^l_t}\left(\left(1-\lambda\right)\left(\Pi^l_t\left(1\right) - \pi\right) - \left(1-\lambda p^l_t\right)\left(\Pi^l_{t+1}\left(1\right) - \pi\right)\right),
\end{aligned}$$

since Equation (B.19) implies that

$$\lambda p^l_tV - c\frac{p^l_t}{p_t\left(q_t\right)} - k - \pi = -\delta\left(1-\lambda p^l_t\right)\pi.$$

Now, we pick $t = \bar{\tau}^l - 2$ and we pick $\pi \leq \Pi^l_{\bar{\tau}^l-1}(1)$, arbitrarily close to $\Pi^l_{\bar{\tau}^l-1}(1)$. Then, we have $\hat{f}_{\bar{\tau}^l-2} > 0$ and $\hat{f}_{\bar{\tau}^l-1} = 0$. Thus, there exists a V_0 such that the two IC conditions in Equation (IC^{Pres}_l) and Equation (IC^{Pres}_h) hold in period $\bar{\tau}^l-2$ and such that $f_{\bar{\tau}^l-2} > 0$. This implies that the high type prefers separating in period $\bar{\tau}^l - 2$

over separating in any later period. This, in turn, implies that there exists a period $\tau_S^h \leq \bar{\tau}^l - 2$ at which the high type prefers to separate.

As before, the low type strictly prefers to pool until period $\bar{\tau}^l$. Thus, for γ sufficiently large, there exists an optimal separation period $\tau_S < \bar{\tau}^l - 1$ which maximizes the entrepreneur's ex-ante value.

B.2.7 Proof of Proposition 17

We start with a monotonicity lemma so that we can characterize $M^\theta\left(C_t'\right)$ by a lower bound belief.

Lemma 53. *In the optimal pooling equilibrium, if* $1 > q_1 > q_2 > 0$*, then* $\Pi_t^\theta\left(q_1\right) > \Pi_t^\theta\left(q_2\right)$ *holds.*

Proof. We start with the low type,

$$
\begin{aligned}
& \Pi_t^l\left(q_1\right) - \Pi_t^l\left(q_2\right) \\
= \;& \lambda p_t^l \left(\sum_{s=t}^{\underline{\tau}_2^l - 1} \left(\delta\left(1-\lambda\right)\right)^{s-t} \left(\frac{c}{\lambda p_s\left(q_2\right)} - \frac{c}{\lambda p_s(q_1)} \right) \right. \\
& \left. + \sum_{s=\underline{\tau}_2^l}^{\underline{\tau}_1^l - 1} \left(\delta(1-\lambda)\right)^{s-t} \left(V - \frac{c}{\lambda p_s(q_1)} - \frac{k}{\lambda p_s^l} \right) \right) \\
> \;& 0.
\end{aligned}
$$

Recall that $\underline{\tau}_i^l$ is the first period when the low type starts to drop out with positive probability if $q_t = q_i$. Notice for $t \geq \underline{\tau}_1^l$, both contracts offer the same expected payoff 0 for the low type. In other words, these two contracts coincide in such periods. The definition of $\underline{\tau}_i^l$ implies $\underline{\tau}_1^l \geq \underline{\tau}_2^l$.[19] Similarly, for the high type,

$$\begin{aligned}
& \Pi_t^h(q_1) - \Pi_t^h(q_2) \\
= \; & \lambda p_t^h \left(\sum_{s=t}^{\underline{\tau}_2^l - 1} (\delta(1-\lambda))^{s-t} \left(\frac{c}{\lambda p_s(q_2)} - \frac{c}{\lambda p_s(q_1)} \right) \right. \\
& \left. + \sum_{s=\underline{\tau}_2^l}^{\underline{\tau}_1^l - 1} (\delta(1-\lambda))^{s-t} \left(\alpha_s^p - \frac{c}{\lambda p_s(q_1)} \right) \right) \\
> \; & 0,
\end{aligned}$$

where $\alpha_s^p = \left(\lambda p_s^l V - k\right)/\lambda p_s^l$ is the pooling contract that generates 0 payoff for the low type in expectation. □

Corollary 54. *In the optimal pooling equilibrium,* $\Pi_t^l(q) > \Pi_t^l(0)$ *holds for any* $0 < q \leq 1$.

Proof. After some algebra, we have

$$\begin{aligned}
& \Pi_t^l(q) - \Pi_t^l(0) \\
= \; & \lambda p_t^l \left(\sum_{s=t}^{\underline{\tau_l} - 1} (\delta(1-\lambda))^{s-t} \left(\frac{c}{\lambda p_s^l} - \frac{c}{\lambda p_s(q)} \right) \right. \\
& \left. + \sum_{s=\underline{\tau_l}}^{\tau_l - 1} (\delta(1-\lambda))^{s-t} \left(V - \frac{c}{\lambda p_s(q)} - \frac{k}{\lambda p_s^l} \right) \right) \\
> \; & 0,
\end{aligned}$$

where $\underline{\tau_l}$ is the first period when the low type starts to drop out with positive probability with the optimal pooling contract.[20] □

Lemma 53 implies $M^\theta(C_t')$ must be a closed interval $\left[q^\theta(C_t'), 1\right]$. It is sufficient to

[19] If $\underline{\tau}_1^l = \underline{\tau}_2^l$, then we follow the convention that $\sum_{s=\underline{\tau}_2^l}^{\underline{\tau}_1^l - 1} x_s = 0$.

[20] In the case of $q = 1$, this is the first period when the low type drops out if she is regarded as the high type.

find the off-equilibrium belief $q^{\theta}(C_t')$ such that the θ type is indifferent with deviating and offering the optimal contract in continuation. Therefore, if $q^{\theta}(C_t') < q^{\theta'}(C_t')$, then $M^{\theta'}(C_t') \subset M^{\theta}(C_t')$ and the VC should believe facing θ with certainty.

We first focus on the case when payouts d_t^h are feasible. Lemma 55 shows if both the optimal pooling equilibrium and separating equilibrium exist, the D1 criterion will pick the separating equilibrium.

Lemma 55. *The optimal pooling equilibrium does not satisfy the D1 criterion if there exists an optimal separating equilibrium that pays out $d_{\tau_s}^h$ at τ_s.*

Proof. Consider period τ_s, by the construction and definition of τ_s and $d_{\tau_s}^h$, the following equation must be true according to the proof of Proposition 3:

$$\Pi_t^h(q_t) \leq \Pi_t^h(1) - d_{\tau_s}^h.$$

Also notice

$$\Pi_t^l(q_t) > \Pi_t^l(0) = \Pi_t^l(1) - d_{\tau_s}^h,$$

The first inequality follow from Corollary 54 and the second equality follows the fact that $d_{\tau_s}^h = \Pi_t^l(1) - \Pi_t^l(0)$. This implies that there exists an $\varepsilon > 0$ such that $d' = d_{\tau_s}^h - \varepsilon$ and $M^l(d') = \emptyset$ and $\{1\} \subset M^h(d')$. □

Next we lay out a necessary condition that all separating equilibria have to satisfy in order to survive the D1 criterion.

Lemma 56. *For any equilibrium separating with $d_{\tau_S}^h = \Pi_{\tau_S}^l(1) - \Pi_{\tau_S}^l(0)$ at τ_S, the equilirbium does not satisfy the D1 criterion if there exists $t < \tau_S$ such that Equation (2.20) is true.*

Proof. We prove by constructing a profitable deviation d' for the high type. Consider a deviation d' satisfying

$$\Pi_t^l(1) - \Pi_t^l(q_t) < d' < \Pi_t^l(1) - \Pi_t^l(0).$$

The existence of d' is guaranteed by $\Pi_t^l(q_t) > \Pi_t^l(0)$. The above equation implies that the low type is strictly worse off by paying any d' even though it was regarded as the high type. Therefore $M^l(d') = \emptyset$. We then show that there exist some d' such that $M^h(d')$ is not empty if Equation (2.20) is true. To see this, notice

$$\begin{aligned} & \Pi_t^l(1) - \Pi_t^l(q_t) < \Pi_t^h(1) - \Pi_t^h(q_t) \\ \Leftrightarrow \quad & \Pi_t^h(q_t) < \Pi_t^h(1) - \left(\Pi_t^l(1) - \Pi_t^l(q_t)\right). \end{aligned}$$

By continuity, there exists an $\varepsilon > 0$, such that $d' = \Pi_t^l(1) - \Pi_t^l(q_t) + \varepsilon$, satisfying $1 \subset M^h(d')$ and $M^l(d') = \emptyset$. The refined off-equilibrium belief is $q' = 1$ upon observing d'. But then the high type will deviate. $\square$

To characterize the unique equilibrium under the D1 criterion, define

$$g_t(q_0) = c\left(\frac{p_t^h}{q_0 p_t^h + (1-q_0)\, p_t^l} - 1\right) - \delta\lambda p_t^h \left(\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0)\right) > f_t(q_0),$$

where $f_t(q_0)$ is defined in Lemma 38. Following the same proof of Lemma 39, we can show

$$g_1(q_0) < g_2(q_0) < \ldots < g_{\underline{\tau_l}-1}(q_0) < g_{\underline{\tau_l}}(q_0).$$

Denote $\tau_s^\star$ as the first period at which $g_t(q_0) > 0$. $\tau_s^\star$ always exists. This is because in Lemma 43, we have essentially proved if $\underline{\tau_l} < \tau^{l\prime}$, then $f_{\underline{\tau_l}}(q_0) > 0$ and hence

$g_{\underline{\tau_l}}(q_0) > f_{\underline{\tau_l}}(q_0) > 0$. If $\underline{\tau_l} = \tau^{l\prime}$, then

$$g_{\underline{\tau_l}-1}(q_0) = c\left(\frac{p^h_{\underline{\tau_l}-1}}{q_0 p^h_{\underline{\tau_l}-1} + (1-q_0)\, p^l_{\underline{\tau_l}-1}} - 1\right) > 0$$

since $p^h_{\underline{\tau_l}-1} > p^l_{\underline{\tau_l}-1}$. The following Lemma is useful for comparing the efficiency across different separating equilibria.

Lemma 57. *For $t \leq \underline{\tau_l} - 1$, suppose $\Pi^\theta_t(q_0)$ is the equilibrium payoff if the high type optimally separates at t. Suppose $\Pi'^\theta_t(q_0)$ is the equilibrium payoff if the high type optimally separates at $t+1$. Then both*

$$\Pi^h_t(q_t) - \Pi'^h_t(q_t) > \Pi^l_t(q_t) - \Pi'^l_t(q_t) \tag{B.21}$$

and

$$\Pi^h_t(1) - \Pi'^h_t(q_t) > \Pi^l_t(1) - \Pi'^l_t(q_t) \tag{B.22}$$

are equivalent to $g_t(q_0) > 0$.

Proof. We start with the following equivalent result:

$$\begin{aligned} & \Pi^h_t(q_t) - \Pi'^h_t(q_t) > \Pi^l_t(q_t) - \Pi'^l_t(q_t) \\ \Leftrightarrow\ & \Pi^h_t(1) - \left(\Pi^l_t(1) - \Pi^l_t(0)\right) - \Pi'^h_t(q_t) > \Pi^l_t(0) - \Pi'^l_t(q_t) \\ \Leftrightarrow\ & \Pi^h_t(1) - \Pi'^h_t(q_t) > \Pi^l_t(1) - \Pi'^l_t(q_t). \end{aligned}$$

Notice

$$\begin{aligned} & \Pi^l_t(1) - \Pi'^l_t(q_t) \\ =\ & \lambda p^l_t\left(\frac{c}{\lambda p_t(q_0)} - \frac{c}{\lambda p^h_t}\right) + \delta\left(1 - \lambda p^l_t\right)\left(\Pi^l_{t+1}(1) - \Pi^l_{t+1}(0)\right), \end{aligned}$$

and

$$
\begin{aligned}
& \Pi_t^h(1) - \Pi_t'^h(q_t) \\
= \; & \lambda p_t^h \left(\frac{c}{\lambda p_t(q_0)} - \frac{c}{\lambda p_t^h} \right) + \delta \left(1 - \lambda p_t^h\right) \left(\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0) \right).
\end{aligned}
$$

This implies

$$
\begin{aligned}
& \left(\Pi_t^h(1) - \Pi_t'^h(q_t)\right) - \left(\Pi_t^l(1) - \Pi_t'^l(q_t)\right) \\
= \; & \lambda \left(p_t^h - p_t^l\right) \left(\left(\frac{c}{\lambda p_t(q_0)} - \frac{c}{\lambda p_t^h} \right) - \delta \left(\Pi_{t+1}^l(1) - \Pi_{t+1}^l(0) \right) \right) \\
= \; & \frac{\lambda \left(p_t^h - p_t^l\right)}{\lambda p_t^h} g_t(q_0).
\end{aligned}
$$

Since $\lambda \left(p_t^h - p_t^l\right) > 0$ is true for any t, the sign of $g_t(q_0)$ determines the relationship between $\Pi_t^h(1) - \Pi_t'^h(q_t)$ and $\Pi_t^l(1) - \Pi_t'^l(q_t)$. □

Now we prove the unique equilibrium is the high type separating with $d_{\tau_s^\star}^h = \Pi_{\tau_s^\star}^l(1) - \Pi_{\tau_s^\star}^l(0)$ at $\tau_s^\star$.

Lemma 58. *If $\tau_s^\star \neq \tau_s$, then the equilibrium separating with $d_{\tau_S}^h = \Pi_{\tau_S}^l(1) - \Pi_{\tau_S}^l(0)$ at τ_S does not satisfy D1 criterion.*

Proof. First if $\tau_s > \tau_s^\star$, consider $t = \tau_s - 1$ and by definition $g_t(q_0) > 0$. Following the proof of Lemma 57, this implies Equation (2.20) is true so that a profitable deviation exists for the high type. Next if $\tau_s < \tau_s^\star$, consider the following deviation: Offering no payouts at t, α_t^P at t, and $d_{t+1}^h = \Pi_{t+1}^l(1) - \Pi_{t+1}^l(0)$. In other words, the entrepreneur deviates to separate at $t+1$. Notice the payoff of this deviation does not depend on the off equilibrium belief at t. This is because the payment at t is fixed and the belief at $t+1$ will jump to either 1 or 0, decided by whether d_{t+1}^h is made. Since separation is delayed, the low type is strictly better off. Moreover, $f_t(q_0) < g_t(q_0) < 0$

implies the high type is strictly better off by delaying separation as well. Therefore, regardless of the belief of the VC, both type will deviate. □

Lemma 59. *The following equilibrium survives the D1 criterion: The high type pays* $d^h_{\tau_s^\star} = \Pi^l_{\tau_s^\star}(1) - \Pi^l_{\tau_s^\star}(0)$ *at* $\tau_s^\star$ *to separate.*

Proof. First, for period $\tau_s^\star$, a deviation that pays d' strictly larger than $\Pi^l_{\tau_s^\star}(1) - \Pi^l_{\tau_s^\star}(0)$ will make both types strictly worse off, even if the off-equilibrium belief is $q_t' = 1$. Therefore, it is only worth checking $d' < \Pi^l_{\tau_s^\star}(1) - \Pi^l_{\tau_s^\star}(0)$. By continuity, there exists a q such that

$$d' = \Pi^l_{\tau_s^\star}(q) - \Pi^l_{\tau_s^\star}(0).$$

In other words, if the off-equilibrium belief is q, then the low type is indifferent between deviating or not, i.e. $M^l(d') = [q, 1]$. We then show that the high type is strictly worse off after paying d' and induces an off-equilibrium belief q. First, consider the case when q is large enough so that the high type offers optimal pooling contract in continuation. We want to show

$$\begin{aligned} & \Pi^h_{\tau_s^\star}(1) - d^h_{\tau_s^\star} > \Pi^h_{\tau_s^\star}(q) - d' \\ \Leftrightarrow\ & \Pi^h_{\tau_s^\star}(1) - \Pi^l_{\tau_s^\star}(1) > \Pi^h_{\tau_s^\star}(q) - \Pi^l_{\tau_s^\star}(q) \\ \Leftrightarrow\ & \Pi^h_{\tau_s^\star}(1) - \Pi^h_{\tau_s^\star}(q) > \Pi^l_{\tau_s^\star}(1) - \Pi^l_{\tau_s^\star}(q). \end{aligned} \tag{B.23}$$

The RHS equals to,

$$\begin{aligned}
& \Pi^{l}_{\tau^{\star}_{s}}(1) - \Pi^{l}_{\tau^{\star}_{s}}(q) \\
= \ & \lambda p^{l}_{\tau^{\star}_{s}} \left(\sum_{s=\tau^{\star}_{s}}^{\underline{\tau}^{l}-1} \left(\delta\left(1-\lambda\right)\right)^{s-t} \left(\frac{c}{\lambda p_{s}\left(q\right)} - \frac{c}{\lambda p^{h}_{s}} \right) \right. \\
& \left. + \sum_{s=\underline{\tau}^{l}}^{\tau^{l\prime}-1} \left(\delta(1-\lambda)\right)^{s-t} \left(V - \frac{c}{\lambda p^{h}_{s}} - \frac{k}{\lambda p^{l}_{s}} \right) \right).
\end{aligned}$$

The LHS is equal to

$$\begin{aligned}
& \Pi^{h}_{\tau^{\star}_{s}}(1) - \Pi^{h}_{\tau^{\star}_{s}}(q) \\
= \ & \lambda p^{h}_{\tau^{\star}_{s}} \left(\sum_{s=\tau^{\star}_{s}}^{\underline{\tau}^{l}-1} \left(\delta\left(1-\lambda\right)\right)^{s-t} \left(\frac{c}{\lambda p_{s}\left(q\right)} - \frac{c}{\lambda p^{h}_{s}} \right) \right. \\
& \left. + \sum_{s=\underline{\tau}^{l}}^{\tau^{l\prime}-1} \left(\delta(1-\lambda)\right)^{s-t} \left(c^{p}_{s} - \frac{c}{\lambda p^{h}_{s}} \right) \right).
\end{aligned}$$

Recall that $c^{p}_{s} = \left(\lambda p^{l}_{s} V - k\right) / \lambda p^{l}_{s}$ is the optimal pooling contract for $s \geq \underline{\tau}^{l}$. It can be shown that

$$\alpha^{p}_{s} - \frac{c}{\lambda p^{h}_{s}} = V - \frac{c}{\lambda p^{h}_{s}} - \frac{k}{\lambda p^{l}_{s}}.$$

Therefore, $\Pi^{h}_{\tau^{\star}_{s}}(1) - \Pi^{h}_{\tau^{\star}_{s}}(q) > \Pi^{l}_{\tau^{\star}_{s}}(1) - \Pi^{l}_{\tau^{\star}_{s}}(q)$ holds since $p^{h}_{\tau^{\star}_{s}} > p^{l}_{\tau^{\star}_{s}}$. Next, consider the case when q is small so the high type prefers to separate again some future time

τ'_S. The payoff for her in this case is

$$\begin{aligned}
& \Pi^h_{\tau^\star_s}(q) - d' \\
= \; & \lambda p^h_{\tau^\star_s}\left(\sum_{s=\tau^\star_s}^{\tau'_S-1}(\delta(1-\lambda))^{s-t}\left(V-\frac{c}{\lambda p_s(q)}-\frac{k}{\lambda p^h_s}\right)\right) \\
& + \left(\prod_{s=\tau^\star_s}^{\tau'_S-1}\delta(1-\lambda p^h_s)\right)\left(\Pi^h_{\tau'_S}(1)-\left(\Pi^l_{\tau'_S}(1)-\Pi^l_{\tau'_S}(0)\right)\right) \\
& - \left(\Pi^l_{\tau^\star_s}(q)-\Pi^l_{\tau^\star_s}(0)\right) \\
= \; & C-\lambda\left(p^h_{\tau^\star_s}-p^l_{\tau^\star_s}\right)\left(\sum_{s=\tau^\star_s}^{\tau'_S-1}(\delta(1-\lambda))^{s-t}\frac{c}{\lambda p_s(q)}\right),
\end{aligned}$$

where C is a term not containing q. Imagine that we first increase q while keeping τ'_S fixed. The above term strictly increases since $dp_s(q)/dq > 0$. The subtlety is, at some point, the high type will prefer delaying separation to $\tau'_S + 1$. Applying $f_{\tau'_S}(q) < 0$ we could still prove the above term keeps increasing when τ'_S jumps. Intuitively, the high type will delay separation if doing so is not profitable. This implies $\Pi^h_{\tau^\star_s}(q) - d'$ is strictly increasing in q. But for a sufficient large q, the high type will prefer to pool in continuation, which we have already shown to be dominated by the equilibrium payoff. Therefore, the high is worse off for any off-equilibrium belief q such that she will separate in the continuation game.

To sum, for any d', if the low type breaks even with some off equilibrium belief q, then the high type will become strictly worse off for the same belief, i.e. $M^h(d') \subset M^l(d')$. Then the refined off-equilibrium belief is to believe facing the low type with certainty after observing b' and hence no one will deviate.

For all $t > \tau^\star_s$, belief is degenerate. It remains to prove that no types will deviate at $t < \tau^\star_s$ by offering a payout earlier. First, consider the deviating payout d'_t such that

$$\Pi_t^l(1) - \Pi_t^l(q_0) < d' < \Pi_t^l(1) - \Pi_t^l(0).$$

This is a payout that will make the low type strictly worse off even if she is regarded as the high type with certainty, i.e. $M^l(d') = \emptyset$. Next we can show that the high type is not willing to deviate as well. Denote $\Pi_t'^{\theta}(q_0)$ as the equilibrium payoff if the high type optimally separates at $t < \tau_s^\star$. Notice since $g_s(q_0) < 0$ for all $t \leq s < \tau_s^\star$, using Lemma 57 repeatedly implies $\Pi_t'^h(q_0) - \Pi_t'^l(q_0) < \Pi_t^h(q_0) - \Pi_t^l(q_0)$. Then we have

$$\begin{aligned}
& \Pi_t^h(q_0) - \left(\Pi_t^h(1) - d'\right) \\
> \ & \Pi_t^h(q_0) - \left(\Pi_t^h(1) - \left(\Pi_t^l(1) - \Pi_t^l(q_0)\right)\right) \\
= \ & \left(\Pi_t^h(q_0) - \Pi_t^l(q_0)\right) - \left(\Pi_t^h(1) - \Pi_t^l(1)\right) \\
> \ & \left(\Pi_t'^h(q_0) - \Pi_t'^l(q_0)\right) - \left(\Pi_t^h(1) - \Pi_t^l(1)\right) \\
= \ & -\frac{\lambda\left(p_t^h - p_t^l\right)}{\lambda p_t^h} g_t(q_0).
\end{aligned}$$

The last equality follows the same algebra as the proof in Lemma 58. Since $g_t(q_0) < 0$, this implies $\Pi_t^h(q_0) > \Pi_t^h(1) - d'$ and the high type will not deviate with the highest belief so $M^h(d') = \emptyset$ as well.

Next consider

$$d' \leq \Pi_t^l(1) - \Pi_t^l(q_0).$$

Therefore, there exists q such that $d' = \Pi_t^l(q) - \Pi_t^l(q_0)$ and $M^l(d') = [q, 1]$. Since $g_s(q_0) < 0$ for all $t \leq s \leq \tau_s^\star - 1$, using Lemma 57 repeatedly implies

$$\Pi_t^h(q_0) - \Pi_t^l(q_0) > \Pi_t^h(1) - \Pi_t^l(1).$$

The above equation is equivalent to

$$\begin{aligned} & \Pi_t^h(q_0) \\ > \; & \Pi_t^h(1) - \left(\Pi_t^l(1) - \Pi_t^l(q_0)\right) \\ > \; & \Pi_t^h(q) - \left(\Pi_t^l(q) - \Pi_t^l(q_0)\right). \end{aligned}$$

The last line follows from $\Pi_t^h(1) - \Pi_t^h(q) > \Pi_t^l(1) - \Pi_t^l(q)$, validated with the same algebra of proving Equation (B.23). This implies the high type is strictly worse off if the off-equilibrium belief is q. In other words, the high type requires a strictly higher belief to break even. Hence, $M^h(d') \subset M^l(d')$. Then the refined off-equilibrium belief is to believe facing the low type with certainty after observing b' and hence no one will deviate. $\square$

So far we have shown if payouts are feasible, there exists a separating equilibrium surviving the D1 criterion. The optimal pooling equilibrium and all other separating equilibria do not satisfy it. To prove the uniqueness, it remains to show all suboptimal pooling equilbria are also pruned. We delay this after Corollary 62. We now prove the second statement of the Proposition, which applies to the cashless entrepreneurs. The following Lemma is useful as an extension of Lemma 34.

Lemma 60. *For any $0 < q < 1$, the following equation holds for all t*

$$\frac{\Pi_t^h(q) - \Pi_t^h(0)}{\lambda p_t^h} > \frac{\Pi_t^l(q) - \Pi_t^l(0)}{\lambda p_t^l}.$$

Proof.

$$
\begin{aligned}
& \Pi_t^h(q) - \Pi_t^h(0) \\
= \ & \lambda p_t^h \left(\sum_{s=t}^{\tau_h'-1} (\delta(1-\lambda))^{s-t} \left(\frac{c}{\lambda p_s^l} - \frac{c}{\lambda p_s(q_1)} \right) \right. \\
& + \sum_{s=\tau_h'}^{\underline{\tau_l}-1} (\delta(1-\lambda))^{s-t} \left(V - \frac{c}{\lambda p_s(q_1)} - \frac{k}{\lambda p_s^h} \right) \\
& \left. + \sum_{s=\max\{\tau_h',\underline{\tau_l}\}}^{\tau^h-1} (\delta(1-\lambda))^{s-t} \left(V - \alpha_s^p - \frac{k}{\lambda p_s^h} \right) \right).
\end{aligned}
$$

Recall that τ_h' is the optimal stopping time if the high type is regarded as the low type. We use $\max\{\tau_h', \underline{\tau_l}\}$ in the last line because τ_h' could be strictly larger than $\underline{\tau_l}$, in which case we set the second term 0. Also notice that,

$$
\begin{aligned}
& V - \frac{c}{\lambda p_s(q_1)} - \frac{k}{\lambda p_s^l} \geq \frac{c}{\lambda p_s^l} - \frac{c}{\lambda p_s(q_1)} \\
\Leftrightarrow \ & \lambda p_s^l \left(V - \frac{c}{\lambda p_s^l} \right) - k \geq 0 \\
\Leftrightarrow \ & s \leq \tau_l.
\end{aligned}
$$

In the last line we use weak inequality because we follow the tie-breaker that players stop with certainty if they are indifferent. This implies

$$
\begin{aligned}
& \frac{\Pi_t^h(q) - \Pi_t^h(0)}{\lambda p_t^h} \\
> \; & \sum_{s=t}^{\tau_h'-1} (\delta(1-\lambda))^{s-t} \left(\frac{c}{\lambda p_s^l} - \frac{c}{\lambda p_s(q_1)} \right) + \sum_{s=\tau_h'}^{\tau_l-1} (\delta(1-\lambda))^{s-t} \left(V - \frac{c}{\lambda p_s(q_1)} - \frac{k}{\lambda p_s^l} \right) \\
\geq \; & \sum_{s=t}^{\tau^l-1} (\delta(1-\lambda))^{s-t} \left(\frac{c}{\lambda p_s^l} - \frac{c}{\lambda p_s(q_1)} \right) + \sum_{s=\tau_l}^{\tau_l-1} (\delta(1-\lambda))^{s-t} \left(V - \frac{c}{\lambda p_s(q_1)} - \frac{k}{\lambda p_s^l} \right) \\
= \; & \frac{\Pi_t^l(q) - \Pi_t^l(0)}{\lambda p_t^l}.
\end{aligned}
$$

The first inequality follows from $k/p_s^h < k/p_s^l$. The second inequality is strict unless $\tau_h' = \tau_l$. □

We start with eliminating all suboptimal equilibria.

Lemma 61. *No suboptimal separating equilibrium survives the D1 criterion for cashless entrepreneurs.*

Proof. Rewrite the separating contract α_t^h as $c/\left(\lambda p_t^h\right) + \Delta$, $\Delta > 0$. By definition, the following equation is true, i.e. the low type is indifferent between mimicking or not,

$$\Pi_t^l(0) = \Pi_t^l(1) - \lambda p_t^l \Delta.$$

By Lemma 60 and the fact this is suboptimal,

$$\Pi_t^h(0) < \Pi_t^h(1) - \lambda p_t^h \Delta < \Pi_t^h(q_t).$$

By continuity and Lemma 53, there exists $0 < q' < q_t$ such that $\Pi_t^h(1) - \lambda p_t^h \Delta =$

$\Pi_t^h(q')$. Define $\Delta' < \Delta$ such that [21]

$$\Pi_t^l(1) - \lambda p_t^l \Delta = \lambda p_t^l \left(V - \frac{c}{\lambda p_t^h} - \Delta' \right) - k + \delta \left(1 - \lambda p_t^l\right) \Pi_{t+1}^l(q').$$

In other words, if the deviating contract is $\alpha_t' = c/\left(\lambda p_t^h\right) + \Delta'$, then $M_t^l\left(\alpha_t'\right) = [q', 1]$. However, the high type is strictly better off after paying additional Δ' instead of Δ. Therefore, $[q', 1] \subset M^h\left(\alpha_t'\right)$. This implies the off-equilibrium belief is $q' = 1$ after observing α_t', then both types will deviate. □

Corollary 62. *No suboptimal pooling contracts survive D1 criterion.*

Proof. Rewrite the suboptimal pooling contract as $\alpha_t^p + \Delta$, $\Delta > 0$. By definition, the following equation is true:

$$\lambda p_t^l \left(V - \alpha_t^p - \Delta\right) - k + \delta \left(1 - \lambda p_t^l\right) \Pi_{t+1}^l(q_t') = 0.$$

This is because for $q_t' > q_t$ to be true, the low type drops out with non-zero probability. If

$$\Pi_t^h\left(0\right) < \lambda p_t^h \left(V - \alpha_t^p - \Delta\right) - k + \delta \left(1 - \lambda p_t^h\right) \Pi_{t+1}^h(q_t') < \Pi_t^h\left(q_t\right),$$

then define q' and Δ' similarly as the proof of Lemma 61, and the rest of proof is essentially the same. Alternatively if,

$$\lambda p_t^h \left(V - \alpha_t^p - \Delta\right) - k + \delta \left(1 - \lambda p_t^h\right) \Pi_{t+1}^h(q_t') < \Pi_t^h\left(0\right),$$

[21]

$$\frac{\lambda}{1-\lambda}\left(\Delta - \Delta'\right) = \delta \frac{\Pi_{t+1}^l(1) - \Pi_{t+1}^l(q')}{p_{t+1}^l}.$$

The difference is not 0 because suboptimal separation happens before $\underline{\tau_l}$. See Lemma 34.

then both types strictly want to deviate by offering a_t^l even though they are regarded as the low type. Lastly, if

$$\lambda p_t^h \left(V - \alpha_t^p - \Delta\right) - k + \delta \left(1 - \lambda p_t^h\right) \Pi_{t+1}^h(q_t') = \Pi_t^h \left(0\right),$$

then players still want to deviate unless $\Pi_t^l(0) = 0$. In the latter case, we have $M^h\left(\alpha_t^l\right) = M^l\left(\alpha_t^l\right) = [0,1]$. Our tie-breaking rule stipulates the off-equilibrium belief is $q_t' = q_t$ so both types will deviate. □

Notice the proof of Corollary 62 shows deviating with equity payment exists for suboptimal pooling equilibria. So we can eliminate them regardless of whether payouts are feasible. This concludes that the first statement of Proposition 17. For the second statement, we next show the optimal pooling equilibrium survives the D1 criterion. We express any deviation in the format as $\alpha_t' = a_t^p + \Delta$, satisfying

$$\begin{aligned} \lambda p_t^l \Delta &= \delta\left(1 - \lambda p_t^l\right)\left(\Pi_{t+1}^l\left(q\right) - \Pi_{t+1}^l\left(q_{t+1}\right)\right) \\ \Longrightarrow \quad \frac{\lambda}{1-\lambda}\Delta &= \delta \frac{\Pi_{t+1}^l\left(q\right) - \Pi_{t+1}^l\left(q_{t+1}\right)}{p_{t+1}^l}. \end{aligned}$$

To understand the first line, the LHS is the expected additional payment, and the RHS is the jump in continuation value if the off-equilibrium belief is $q > q_{t+1}$. In other words, q is the off-equilibrium belief that makes the low type just indifferent and $M^l(\alpha_t') = [q, 1]$. With the same algebra of Lemma 53, for any $q > q_{t+1}$, we have

$$\delta \frac{\Pi_{t+1}^l\left(q\right) - \Pi_{t+1}^l\left(q_{t+1}\right)}{p_{t+1}^l} = \delta \frac{\Pi_{t+1}^h\left(q\right) - \Pi_{t+1}^h\left(q_{t+1}\right)}{p_{t+1}^h}.$$

This implies $M^l(\alpha_t') = M^h(\alpha_t')$ and out tie-breaking rule implies the off-equilibrium belief stays q_t, which is weakly lower than the q_{t+1}. But then both types will become strictly worse off by deviating.