The Race of Unicorns:  
A Signaling Story of Private Acquisitions  

Xuelin Li*

Abstract  
Private companies frequently make acquisitions before IPOs. This paper explains these transactions through a signaling channel. I formalize a real option model where two heterogeneous private startups need IPO fundings when exercising and are able to acquire each other before IPO. Information imperfection exists as investors are unaware of the firm types. Consistent with the neoclassical view of M&As, assets flow from the less productive type to the firm with better technologies but not vice versa. Therefore, being an acquirer generates positive signals for IPO financing and solves the information asymmetry problem. However, efficiency improvement does not always happen. When investors have extremely mistaken priors on firm qualities, low type will resist takeover and force the high type to go public in pooling IPOs without acquiring. Activities that interfere with investor’s learning process, such as an arms race of cash burning, would intensify the pooling equilibrium and reduce allocation efficiencies. Empirically IPOs with previous private acquisitions have significantly less underpricing and better long-term performance, confirming the model mechanism.

Keywords:  Real Option, Startups, Merger and Acquisitions, Initial Public Offering, Information Asymmetry.

JEL Classification:  G14, G32, G34, D81, D82

*Li is at the University of Minnesota (Carlson). Email: lixx3811@umn.edu. I thank my advisors Andrew Winton and Martin Szydlowski and my committee members Raj Singh and Richard Thakor for numerous discussions, feedback and support. I would also like to thank Onur Bayar (Discussant), Frederico Belo, Arnoud W. A. Boot, Simon Gervais, Tong Liu, David Rahman, Aldo Rustichini, Merih Sevilir (Discussant), Fabian Trottner (Discussant), Zhe Wang, Ming Yang, Hongda Zhong and participants in the 2019 Financial Intermediation Research Society conference, 2019 Econometric Society Asia annual conference, 2019 Econometric Society North America Summer Conference, 2018 Northern Finance Association conference, 2018 Wharton Innovation Doctoral Symposium and UMN finance brownbag. Any errors are mine.
1. Introduction

Prior mergers and acquisitions (M&As) research focuses on a sample of large deals involving public firms. Much less is known about acquisitions between private companies, though quantitatively they account for a great portion of total transactions. From 2004 to 2018, 44.2% of private targets were acquired by other non-public companies in high technology industry.\(^1\) Such transactions play a pivotal role in shaping the landscape of product markets. Private acquisitions also exhibit very different patterns from public ones.\(^2\) Therefore, a theoretical framework which can explain such deals with some features of the private company universe is much in need.

In this paper, I focus on startups and a real option model with information imperfections is motivated by the following characteristics of them. First, startups usually have valuable growth options and investment opportunities. Second, they lack internal cash and rely on external fundings such as initial public offerings (IPOs). Lastly, their quality is hard to distinguish for investors and thus they face the adverse selection problem. In the model, there are two firms of heterogeneous types requiring IPO fundings when exercising the option, subject to a noisy learning process by investors. To explain private M&As, they are also allowed to acquire each other before going public. In equilibrium, acquisition decision is optimally determined together with the timing and possibility of stand-alone IPOs. Consistent with the neoclassical view of M&As,\(^3\) assets flow from the less productive type to the firm with better technologies but not vice versa. Since they indicate the better growth opportunities of acquirers than targets, private acquisitions generate positive signals during stock issuance. This prediction is supported by an empirical analysis of a sample of IPOs after 2004.

A more detailed preview of the model is as follows. Prior to IPO, each firm has assets under management stochastically evolving over time as their state variable. Their real option is a costly expansion technology that increases the current assets but they must raise funding externally to cover the cost. Firms differ in their ability of exercising the option so the high type generates higher return than the low type. Both firms know their own type as well as their opponent’s type but investors cannot distinguish their qualities. When acquisition happens, the target transfers its assets and the acquirer expands the combined assets using its own technology. Following Jovanovic and Rousseau (2002) and Maksimovic and Phillips

\(^1\)See Section 5 for empirical details of these observations.

\(^2\)For example, both Maksimovic et al. (2013) and Netter et al. (2011) document private M&A activities are much smoother and less wavelike. Besides, Maksimovic et al. (2013) find they are much less sensitive to firm productivity, credit market liquidity and industry-level misvaluation.

\(^3\)This idea can be illustrated by the “q-theory of mergers” (Jovanovic and Rousseau, 2002). If mergers are an alternative of investment, then high market-to-book (M/B) firms should buy the assets of firms with low opportunities.
When IPO market opens, investors observe whether acquisition happens and the sizes of IPO candidates’ assets.

Investors then use these two pieces of information to infer firm types. When acquisition does not happen, low type will always choose to mimic the high type’s IPO. In this case, assets sizes are the sufficient statistic. High type’s asset has a higher expected growth rate and investors learn based on the realized sizes using Bayes’ rule. Then they will suggest a pooling offer price, which makes the de facto better company suffer an underpricing cost. But if acquisition happens, investors will evaluate the equilibrium type of the acquirer. If the neoclassical prediction of M&As is still valid, investors believe the acquiring company is the high type with probability one. Then they will offer a fair price.

However, the fact that the high type is a more efficient player is not a priori justification that acquisitions will not happen in the other way around given information asymmetry. Theoretically a deviation where the low type acquires the high type is still possible. The loss of efficiency could be compensated by the information rent extracted through deceiving investors. I show that such a deviation is not profitable by comparing the maximal IPO proceeds when the low type acquires to the total continuation payoffs of the two players. The latter is strictly larger because the high type prefers to wait due to favorable information process. Even without acquisitions, investors’ belief will converge to the truth just by observing firm sizes in limits. Therefore the low type must give up the majority of the profits to persuade the high type into deviations, which leaves insufficient portion for itself.

That said, information asymmetry does impact efficient reallocation in a different way: M&As do not happen with certainty and will be delayed. Asset sizes are noisy signals of firm types. In realization, investors may have a highly optimistic belief on the low type, which makes it reject the takeover offer. Then the high type will give up acquisitions and go public, knowing that the low type would mimic in a pooling IPO. The reason is that a mistaken belief takes prolonged periods to revert. For the high type, waiting for additional news incurs the cost of delaying expansion. This cost finally outweighs the loss in a pooling IPO. This explains why there exists a fair amount of non-acquiring IPOs. On average, such IPOs

---

4 The notion of ability is interpreted as managerial and organizational talent, dating at least back to Lucas (1978). The key assumption is that the total profit of acquisition strictly increases in the acquirer’s ability and efficiency gain is realized only when the high type acquires. However, the opposite assumption of technology complementarity as in Rhodes-Kropf and Robinson (2008) can also be accommodated and discussed in the model section.

5 This paper tells a story of firm size difference but the readers could interchangeably consider it as any source of information that is available in the firm’s financial reports or news coverage. More generally, the difference could be due to market share, earnings or even consumer reviews. The emphasis here is what matters for investors is not the absolute value of these measures but rather the relative performance.

6 As in common continuous time model, the continuation payoff is the expected value while players choose to wait in equilibrium.
have poorer long-term performance because they also include less efficient firms, whereas after acquisitions, only high types survive. Lastly, investors’ belief is intermediate when the difference between asset sizes is moderate. Since no type is willing to give up waiting, the game will fall into a region where both startups delay expansion opportunities. The equilibrium outcomes can thus be characterized by three regions, depending on how accurate the investors’ belief is, based on asset sizes.

The existence of delaying expansion and M&A failure is a consequence of information asymmetry. The less efficient firm relies on the outside option of pooling to reject acquisition offers. If investors are perfectly informed about firm ability, low type cannot be overpriced when going public alone. It is Pareto-optimal for the high type to acquire and transfer part of the profit to the low type. Therefore, acquisition will happen right away in the first-best case.

I then extend the model in several ways. I first assume startups can make one-time investments to boost asset growth rates. Though realized greater firm size indicates better quality, firms should not overinvest in equilibrium. The economics is that investors are rational and update belief with respect to the true data generating process. For example, if the less efficient firm increases its growth rate, investors would think realized good performance is manipulated by investments rather than due to better qualities. So they update belief more conservatively. The reduced belief volatility will make the low type worse off while waiting. Additionally, I consider a cash-burning case where both firms increase the growth rates by the same amount. This will make the game more likely to end in a pooling IPO and reduce allocation efficiency in startups. Finally, I include a Nash Bargaining process to pin down an endogenous acquisition offer and show the robustness of equilibrium to alternative bargaining assumption.

Empirically this model has implications on the different patterns between private and public acquisitions. First, information asymmetry prevents the more efficient private firms from acquiring. Thereby as in Maksimovic et al. (2013), firm-level productivity has less predicting power on being a private acquirer compared to public firms. Secondly, private acquisitions should not be procyclical. This is because in booming periods, less efficient firms will have better asset performance and investors are more optimistic about them. As a result, they are reluctant to accept the takeover offer and IPOs without acquiring will increase. This explains why M&A activities become less wavelike once private acquirers are included (Netter et al., 2011).

This model also generates a unique implication on signaling during IPOs. Because private acquisitions indicate the acquirer’s better quality, it is expected to see acquiring IPOs have less underpricing. To validate the efficiency channel, acquiring IPOs will also have
better long-term growth. To test these predictions, I merge stock issuance data from SDC Platinum with private M&A data in SDC Merger and Acquisition database. Companies that have acquired their competitors on average suffer less underpricing in IPO and significantly outperform afterwards. I document 2.538% lower first-day return and 8.965% higher market-adjusted three-year buy-and-hold return for firms with private acquisitions than non-acquiring companies. The better performance cannot be due to horizontal M&As reducing competitions alone. If it were the case, IPOs that make acquisitions simultaneously or soon afterwards should have similarly better performance. However, after private companies merge into a large public one in roll-up IPOs, these firms deliver extremely poor stock returns in the following years (Brown et al., 2005, Ritter, 2015). Additionally, Brau et al. (2012) find IPOs that acquire within a year of going public significantly underperform in the long term. Therefore, market consolidation is not sufficient to justify long-term growth of IPOs. Rational investors should also not reward acquirers of random qualities in IPO offers if acquisitions are not correlated with firm qualities.

Finally, I simulate the model and show that average performance of new public companies decreases with the number of IPOs. This corresponds to the empirical observation that IPO waves usually occur right before market downturns. This model implies the composition of firm types matters. When private acquisition happens, fewer but more efficient companies go public whereas low quality companies also join in the IPO waves. The different characteristics of companies drive the different market performance afterwards. However, investors rationally expect that and adjust the offer price. Therefore, such a pattern could occur without behavioral explanation such as low-quality startups taking advantage of overoptimism, but instead due to a composition effect.

Literature This paper adds to the literature on mergers and acquisitions in the four distinct ways. First, I offer a theoretical framework that generates implications closely related with the patterns of private acquisitions. Empirically, both Maksimovic et al. (2013) and Netter et al. (2011) show the commonly-used sample of large deals with public companies is an underrepresentation of the whole M&A space and many findings cannot be generalized when private transactions are included. I argue that adverse selection and the need of external financing are the key factors to drive the difference. Second, I extend the neoclassical model of M&As to an environment with information asymmetry. I assert that acquisitions still represent the efficient reallocation of assets, which is consistent with models under perfect information (Jovanovic and Rousseau, 2002, Maksimovic and Phillips, 2002) and broadly with the empirical evidence. Yet the result is not simply a reassurance. Though the direction

7 The evidence 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.,
of asset sales remains the same, the occurrence is delayed or even blocked by pooling IPOs. Therefore information imperfection generates efficiency loss in the market of assets.

Thirdly, this paper adds to the literature on the role of growth options or generally intangible assets in M&A. Closely related is Lambrecht (2004), which analyzes how the timing and profit-sharing of mergers are determined in a two-player real option model. The merger there is motivated by economies of scale and firms have no ability differences. Therefore, there exists no possible implications on signaling in Lambrecht (2004). Instead, the author focuses on how the bargaining process of merger endogenously impacts its timing. The distinction is that in this paper, the timing and even the occurrence of M&A endogenously hinges on the belief accuracy of investors. The bargaining protocol is exogenously given. Therefore the two papers complement each other. Alternatively, Levine (2017) models growth opportunities as “seeds” to constrain capital investment.\(^8\)

Lastly, I add 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. Caveats are in order, however. Signaling is not the only motivation of private acquisitions and reversely they are not the only signal to IPO investors. Acquisition and public offerings are both important corporate decisions made out of multiple factors. In fact, private acquisition is not even a perfect signaling tool in the sense that less efficient firms may reject the offer and force pooling IPOs. In this case, other channels of signaling, e.g., coupling with prestigious underwriters (Carter et al., 1998) and all-star analyst coverage (Cliff and Denis, 2004), are still relevant and cannot be substituted.

I add to the literature of the relationships between IPOs and M&As in the following two aspects. First, going public and being acquired are used to be assumed as substituting exit choices for startups. For example, in Bayar and Chemmanur (2012), a startup being acquired is more resilient to future competitions and has less information asymmetry with the bidder, but its owner loses the bargaining advantage and private benefit of control compared to IPO. In this paper, I do not assume they are mutually exclusive choices. In fact, the high type will acquire and then go public with some probability. It is only an equilibrium outcome that the low type will choose between them. The implications are also consistent with the empirical evidence on the timing when IPO exits are relatively more (Ball et al., 2011), and

---

\(^8\)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).
firm-level characteristics that predict higher chances of IPO compared to being acquired (Brau et al., 2003, Poulsen and Stegemoller, 2008). Secondly, 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). I focus on the acquisitions before IPOs and show these acquiring firms have substantially better long-term performance after being public, which differentiate them from IPOs seeking to acquire. Thereby I argue that firm qualities have predicting power on being a private acquirer.

Theoretically this paper extends the literature of real option models with information imperfection. As in Grenadier and Malenko (2011), the exercising return has two parts: a direct project payoff and an indirect belief component depending on a third party’s, which is the IPO investors’ in this model, assessment of types. Therefore players have incentives to manipulate the timing to signal. Extending this framework, other choices, such as debt or equity financing, can be informative as well (Morellec and Schürhoff, 2011). However, these papers are all a single-player optimization problem. Closer related is Gorbenko and Malenko (2017). They consider a model where two bidders with different marginal cost of cash, privately knowing their synergy, decide when to approach the target and the method of payment. Signaling is achieved by paying costly cash. Though Gorbenko and Malenko (2017) has two strategic players, my setup is different. First, in their model players have one common state variable, the stand-alone value, orthogonal to the target’s prior belief. In my model, both startups have their own asset sizes evolving with different drifts. These two state variables are informative and exogenously move the belief of the investors. Second, in my model I consider how signaling is achieved through interactions between players. The idea is motivated by the efficiency reallocation view, which predicts the high type is more likely to acquire. Whereas in Gorbenko and Malenko (2017), the question is not to consider whether one bidder can pay the other to give up acquiring and signal its high valuation of M&As.

This paper is also related with the dynamic adverse selection models (Daley and Green, 2012) in the sense that, the investors’ belief when pooling IPO happens stochastically moves and serves as the key state variable to influence 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 freeze (Fishman and Parker, 2015) and the recovery of it (Chiu and Koeplpl, 2016). In Strebulaev et al. (2016), cash flow plays a dual role of signaling and loosening financial constraint, which both benefit the high type. In my model, asset 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 more synergy but investors would be more optimistic about the less efficient firm, which makes it reluctant to accept takeover offers.
Though the model generates a two-threshold equilibrium as in Daley and Green (2012), the fact that it has two strategic players changes the equilibrium outcome when separation happens. 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 the high type will initiate acquisition with certainty below the boundary since this is its dominant strategy motivated by the efficiency gain, and low type will accept. Therefore I have a full separation 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 shows the equilibrium is characterized by thresholds of belief. Section 4 extends the model in several ways. Section 5 tests and discusses the model’s implications empirically. Section 6 concludes.

2. Model

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,

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

\( B_{ht} \) and \( B_{lt} \) are two independent standard Brownian motions on the canonical probability space \( \{\Omega, F, Q\} \). \( \mu_i \) 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 implies delaying is costly for both types and regulates finite solutions.\(^9\)

Asset sizes are included for two reasons. First, they serve as the base value for the firm’s real option. Second, size growth provides an exogenous channel of learning. The idea follows real life examples when an investor wants to identify the better stock out of two similar ones. The beginning step is to download financial data and calculate realized growth rates. A firm growing rapidly in realization is considered as the high type with higher probability.

Real Option. Both firms have a costly real option to speed up growth, modeled as a linear expansion technology. The stand-alone NPV when exercising is \((H - \alpha) x_{ht}\) for firm \( h \) and

\(^9\)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.
$(L - \alpha) x_l$ for firm $l$. The return satisfies $(i) \ H - \alpha > 0$ and $(ii) \ H > L$. The first restriction implies after exercising, NPV for the high type strictly increases. The second implies the high type generates strictly higher return than the low type. I interchangeably call firm $h$ as the more efficient type, motivated by the fact that it has better expansion technology and asset growth.

$\alpha x_{it}$ is the upfront cost of expansion. Since startups in general lack internal cash and rely heavily on external financing, firms will raise fundings through IPOs to cover the cost. They endogenously choose the timing of IPO to exercise the real option.

To avoid technical complexities, if a single player exercises option at $\tau$, firm types become public information at $\tau^+$ and the remaining firm exercises the option instantaneously soon. Thus, game effectively ends after one execution.

**Acquisition.** Prior to 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.\(^\text{10}\) Specifically, 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 perfect information, only the more efficient type will acquire as in the neoclassical M&A literature. Yet this is not necessarily true with information imperfection as firms may collude to deceive investors and extract information rents to offset the loss.

The acquirer transfers part of the expansion profit to target as acquisition offers. This can be realized in a stock-exchange transaction such that the target receives shares of the merged company.\(^\text{11}\) 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 larger than $(H - \alpha + \gamma_h) x_{ht}$ for firm $h$ as the target and $(L - \alpha + \gamma_l) x_{lt}$ for firm $l$ as the target.\(^\text{12}\)

Firms endogenously choose the timing of acquisition, jointly with the decision of a stand-alone IPO, knowing that: $(i)$ an acquisition deal is successful only if both parties are willing to, and $(ii)$ acquisition is observable to investors. $(i)$ implies that the target would reject

---

10I assume 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 size $x_{mt}$ 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.

11This simplifies the acquisition process and avoid alternative signaling concerns through methods of payments. See Hege and Hennessy (2010), Lambrecht (2004) for a similar assumption.

12In Section 4.3, I endogenize offer value through Nash bargaining where the threat point is to go public alone while being regarded as the low type.
being acquired if the offer value is below its equilibrium continuation value or the exogenous reservation value. \((ii)\) assumes when there is only one firm going public, the investors could distinguish the merged firm from a stand-alone IPO.\(^{13}\)

**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 \(i\) decide 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 \(\alpha x_{it}\) 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 decision of the first one. The order is indistinguishable in investor’s perspective in a pooling IPO.

Following the real option literature, I consider Markov stopping strategies, which is a 3-tuple \(\sigma_{it} = (\sigma^A_{it}(x_{ht}, x_{lt}), \sigma^T_{it}(x_{ht}, x_{lt}), \sigma^I_{it}(x_{ht}, x_{lt}))\). \(\sigma^A_{it} : X^h_t \times X^l_t \rightarrow [0, 1]\) is the probability of firm \(i\) making an acquisition offer at Step 3. \(\sigma^T_{it} : X^h_t \times X^l_t \rightarrow [0, 1]\) is the probability of firm \(i\) accepting an offer, conditional on it receives one. \(\sigma^I_{ht} : X^h_t \times X^l_t \rightarrow [0, 1]\) is the probability of firm \(h\)’s stand-alone IPO when it moves first in the IPO stage. \(\sigma^I_{lt} : X^h_t \times X^l_t \rightarrow [0, 1]\) 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 \(\sigma^I_{lt}\) as a mapping from the size space, rather than make it a function of the realized action of firm \(h\), with the restriction that \(\sigma^I_{ht}(x_{ht}, x_{lt}) = \sigma^I_{lt}(x_{ht}, x_{lt})\). Lastly, I will focus on the equilibria that acquisition offer equals to the reservation value of the target in equilibrium.\(^{14}\) Hence offer value is not in the strategy space.

---

\(^{13}\)In equilibrium, a single stand-alone IPO will never happen. This assumption is required for off-equilibrium deviations.

\(^{14}\)It does not rule out off-equilibrium deviations with offers higher than the reservation value. A full-fledged
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 investors, who initially believe both firms have equal chance to be the high type. When IPO market opens, they observe all IPO candidates and their assets sizes. The information set is therefore $I_t = \{x_{kt}\}_{k \in C_t}$. $C_t$ is the index set IPO candidates. $C_t = \{h, l\}$ in a pooling IPO and $C_t = \{m\}$ if merger happens. Besides, they are aware of the data generating process of sizes and are able to apply Bayes’ rule when updating beliefs. When IPO happens, investors make pricing decisions based on

$$s_{kt} (HP_{kt} + L (1 - P_{kt})) = \alpha.$$  

$s_{kt}$ are the shares issued to investors. $1 - s_{kt}$ remains for original owners of IPO candidate $k$. The price of issuing shares is a function of $P_{kt}$, which is candidate $k$’s probability of being a high type conditional on $I_t$. $(HP_{kt} + L (1 - P_{kt}))$ is also the expected gross return of expansion in investors’ belief. Equation (2) is a zero-profit condition. To motivate this, One could assume that there are at least two investors in the IPO market who compete in a Bertrand fashion. Let $Z_i \in \{H, L\}$ denote the return rate of firm $i$. A stand-alone IPO has payoff functions for firm $i$

$$R^I_i (x_{it}, x_{-it}) = (1 - s_{it}) Z_i x_{it}.$$ 

A merged IPO, when $i$ is the acquirer has payoff functions

$$R^m_i (x_{it}, x_{-it}) = (1 - s_{mt}) Z_i x_{mt} - (Z_{-i} - \alpha + \gamma_{-i}) x_{-it},$$

$$R^{m,-i}_i (x_{it}, x_{-it}) = (Z_{-i} - \alpha + \gamma_{-i}) x_{-it},$$

for the acquirer and target respectively.

Equilibrium Concept. Firm $i$’s strategy $\sigma^i$ maximizes its continuation $V_i (x_{it}, x_{-it})$ given its opponent strategy $\sigma^{-i}$. Formally, it solves (FP) given $\sigma^{-i}$:

$$V_i (x_{it}, x_{-it}) = \sup_{\sigma^i} \mathbb{E}^i \left( \int_0^\tau e^{-r\tau} \left( R^I_i (x_{i\tau}, x_{-i\tau}) 1^I_\tau + R^m_i (x_{i\tau}, x_{-i\tau}) 1^m_\tau \right) | x_{it}, x_{-it} \right).$$  

(FP)

In (FP), $\tau$ is the random stopping time. $1^I_\tau$ is the indicating function of a stand-alone IPO and $1^m_\tau$ is the indicating function of mergers. $\sigma^i$ and $\sigma^{-i}$ together decide the expected probability of these events and whether the player is an acquirer or a target in acquisitions. A Markov model with endogenous offer value can be pinned down by a commonly preferred stopping threshold of both types. See page 50 in Lambrecht (2004).

The investors know there is exactly one high type and one low type, but they do not know “who is who”.  

15
Perfect Bayesian Equilibrium is a pair of \( (\sigma^h, \sigma^l) \) 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 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). Especially, 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 when pooling IPO is highly likely. This alternative case is also easier to solve, since the merged firm’s efficiency is not random, which automatically “signals”, and the belief of investors when seeing M&As is degenerate.

Lastly, I assume sequential moves in each stage at time \( t \). This is because a pooling IPO is 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 IPO is 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.
3. Equilibrium

3.1. Belief

Consider the first case when merger happens. Investors only observe one firm $m$ with size $x_{mt}$. The probability that the acquirer is firm $i$ follows

$$q_i(x_{mt}) = \int_{x_1 + x_2 = x_{mt}} \sigma^A_{it}(x_1, x_2) \sigma^T_{-it}(x_1, x_2) f^i_t(x_1) f^t_{-i}(x_2) \, dx_1,$$

where $f^i_t$ is the size distribution of firm $i$ at time $t$, derived from Equation (1). To understand this equation, given that the acquiring firm $i$ has size $x_1$, the target must have size $x_2$ equal to $x_{mt} - x_1$. Conditional on the realization of $(x_1, x_2)$, the probability of a successful deal is the probability that firm $i$ will make the offer ($\sigma^A_{it}(x_1, x_2)$) times the probability that firm $-i$ will accept the offer ($\sigma^T_{-it}(x_1, x_2)$). And lastly the expectation is taken over the size distributions. Using Bayes’ rule, the probability that the acquirer is the high type is

$$P^m(x_{mt}) = \frac{q_h(x_{mt})}{q_h(x_{mt}) + q_l(x_{mt})}.$$

Lemma 1 implies I can focus on equilibria in which only firm $h$ acquires. In other words, it is sufficient to consider equilibria with $P^m(x_{mt}) \equiv 1$. The intuition is that given a pair of $(x_1, x_2)$, there exist 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 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 identify of acquirer changes. This is contradictory.

**Lemma 1** There exist no equilibria in which firm $l$ is acquirer with strictly positive probability, i.e., $q_l(x_{mt}) = 0$ for all $x_{mt}$.

**Proof.** All proofs are omitted and shown separately in Appendix C.

The fact that $P^m(x_{mt})$ is a constant function also implies the combined assets size $x_{mt}$ has no impact on investors’ belief. The action of being an acquirer itself is fully informative and makes belief degenerate.

Next we consider the case of pooling IPOs. In this case investors meet both firms and observe their sizes $x_{ht}$ and $x_{lt}$. Since in equilibrium the low type will adopt a mimicking
strategy, \( \sigma^I_{ht}(x_{ht}, x_{lt}) = \sigma^I_{lt}(x_{ht}, x_{lt}) \) and thus stand-alone IPO strategies have no impact on belief updating. In fact, asset sizes are the sufficient statistics. Firm \( i \)'s probability of being the high type is

\[
P_i^I(x_{it}, x_{-it}) = \frac{f^h_i(x_{it}) f^l_i(x_{it})}{f^h_i(x_{it}) f^l_i(x_{it}) + f^l_i(x_{it}) f^h_i(x_{it})}.
\]

If firm \( i \) is the high type, then its assets size follows distribution \( f^h_i \) and its opponent \(-i\), by exclusivity, must be the low type and therefore follows distribution \( f^l_i \). It is of particular importance to explicitly derive investors’ “mistaken belief”, which is firm \( l \)'s log likelihood ratio of being the high type, \( \rho_{lt} = \log(\frac{p^I_{lt}}{1-p^I_{lt}}) \):

\[
\rho_{lt} = \log\left(\frac{f^h_l(x_{lt}) f^l_i(x_{ht})}{f^l_l(x_{ht}) f^h_i(x_{lt})}\right) = \log\left(\frac{x_{lt}}{x_{lt}}\right) \frac{\mu_h - \mu_l}{\sigma^2}. \tag{3}
\]

\( \rho_{lt} \) measures how optimistic that the investors are about the low type through observing asset sizes in a pooling IPO. A large value of \( \rho_{lt} \) implies an extremely wrong belief deviating from the truth. By equation (1), \( d\rho_{lt} = -\frac{(\mu_h - \mu_l)^2}{\sigma^2} \) \( dt + \frac{\mu_h - \mu_l}{\sigma} dB_t \), where \( dB_t = \frac{1}{\sqrt{2}} dB_t - \frac{1}{\sqrt{2}} \) \( dB_{ht} \). The mistaken belief is expected to decrease strictly over time. If players are sufficiently patient, asset sizes will truthfully reveal types as \( \lim_{t \to \infty} \text{E}(\rho_{lt}) = -\infty \). The first part is \( \log(\frac{x_{lt}}{x_{ht}}) \). Intuitively, a larger realized size of the low type will induce a more optimistic belief by investors. The second part measure how precise asset sizes are for discerning companies. Firm \( h \)'s log-likelihood would just be the opposite of \( \rho_{lt} \) and henceforth neglected for saving the notation.\(^{17}\) Initial sizes are \( x_{ht0} = x_{lt0} = 1 \) so that \( \rho_{lo} = 0 \), which matches the fact that both firms are equally likely to be the high type in prior.

\( \rho_{lt} \) is a key state variable. Because it essentially quantifies the mispricing of shares through Equation (2). Therefore, firm strategies will be naturally impacted by \( \rho_{lt} \). As in common real

\(^{16}\)This is because:

\[
\log\left[\frac{f^h_l(x_{lt}) f^l_i(x_{ht})}{f^l_l(x_{ht}) f^h_i(x_{lt})}\right] = \log\left[\exp\left(-\frac{(\log x_{lt} - (\mu_h - \mu_l)^2)}{\sigma^2 t}\right)\right] = \log\left[\exp\left(-\frac{(\log x_{ht} - (\mu_h - \mu_l)^2)}{\sigma^2 t}\right)\right] = \log\left[\frac{x_{lt}}{x_{ht}}\right] \frac{\mu_h - \mu_l}{\sigma^2}.
\]

\(^{17}\)By addition rule for mutually exclusive events, \( P^I_{ht} + P^I_{lt} = 1 \). Therefore \( \rho_{ht} = \log(\frac{p^I_{ht}}{1-p^I_{ht}}) = \log(\frac{1-p^I_{lt}}{p^I_{lt}}) = -\rho_{lt}. \)
option model, this is reflected as threshold strategies, which require that $\sigma^i$ can be determined by verifying whether $\rho_{lt}$ is above or below certain thresholds. More importantly, since firm returns are linear in its own size and $\rho_{lt}$ only changes the marginal rate of return, these thresholds must be constant. The equilibrium strategy is stationary in belief and unrelated with 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 $x_{lt}$.

Lastly, I do not need to consider the equilibrium strategy of a single stand-alone IPO. Given that the low type has a dominant strategy of mimicking, there exists no equilibria in which $\sigma^i_h(x_{ht}, x_{lt}) > \sigma^i_l(x_{ht}, x_{lt})$ with positive measures. The flip side also implies the possible stand-alone IPO, which is not followed, is initiated 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. The procedure of solving the equilibrium is as follows. First, I guess that there are two constant thresholds $\beta$ and $\eta$, $\eta > 0 > \beta$ 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 $(\beta, \eta)$. Lastly, I verify that the guessed strategy is indeed an equilibrium by showing there exists no profitable deviations.

### 3.2. Two-threshold Equilibrium

Figure 1 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 $\rho_{lt}$ is a function of $log(\frac{x_{lt}}{x_{ht}})$. Thus a constant $\rho$-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 in realization, investors almost perfectly identify the true types correctly. The low type has no incentives to wait as the overpricing in pooling IPOs is 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 luck 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, knowing that it would be mimicked by the low type and suffer underpricing in pooling IPO. In the light area between them, belief is in the intermediate region so both types postpone exercising the option.
Given the equilibrium strategy, the next step is to quantify firm returns in different ending outcomes. The payoffs when firm $h$ makes acquisitions are

$$R_{hm}^{m}(x_{ht}, x_{lt}) = (H - \alpha) x_{ht} + (H - L - \gamma_l) x_{lt},$$

$$R_{lm}^{m}(x_{ht}, x_{lt}) = (L - \alpha + \gamma_l) x_{lt}.$$

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 expanding 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 a pooling IPO are

$$R_{hm}^{l}(x_{ht}, x_{lt}) = \left( H - \alpha - \frac{\alpha (H - L) \exp(\rho_l)}{H + L \exp(\rho_l)} \right) x_{ht},$$

$$R_{lm}^{l}(x_{ht}, x_{lt}) = \left( L - \alpha + \frac{\alpha (H - L) \exp(\rho_l)}{H \exp(\rho_l) + L} \right) x_{lt}.$$

The pooling payoff consists of two parts. The first part is the NPV of expansion and the second part is the discount or premium due to mispricing. As common, the high type is
priced at a discount and thus its total payoff is lower than NPV. The low type benefits and earns additional premium. Therefore pooling IPO is a “winning” scenario for the low type.

3.3. Value Function

This section shows the existence and uniqueness of \((\eta, \beta)\). The beginning step is to lay out 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 \(\rho_{lt}\) but \(V_i\) is a function of asset sizes. Thus the smooth pasting conditions are not clear.

However, the linearity of model setup generates great tractability, which implies \(V_i\) is homogeneous of degree one in its own assets size \(x_{it}\). Therefore the value function satisfies \(V_i(x_{it}, x_{-it}) = x_{it}V_i(1, x_{-it}/x_{it})\). Since \(\rho_{lt}\) is a function of \(\log(x_{lt}/x_{ht})\), \(V_i(1, x_{-it}/x_{it})\) is essentially a function of belief and 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 for public companies. Alternatively, when \(x_{it}\) is referred as total sales, then \(J_i\) mirrors the price-to-sales ratio for public companies. Both ratios are commonly used for valuing stocks and IPOs.

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 (B.1) and (B.2) at Appendix B. While these equations are tedious and intensive in algebra, the sources of waiting value can be divided into three parts. The first part is a size effect. Since the ending payoffs are linear in firm sizes, increasing \(x_{it}\) will provide larger benefit in both exit scenarios. The second part is belief effect. Fluctuations of \(\rho_{lt}\) affect underpricing and overpricing in pooling IPO. Besides if belief \(\rho_{lt}\) is close to thresholds \((\eta, \beta)\), the probability that the game ends will increase. The last source is the cross effect between size and belief. A larger realized size \(x_{it}\) will induce a more optimistic belief by investors on firm \(i\). Then it becomes more likely to realize its winning scenario.

The ending scenarios generate four value matching conditions by equations (B.3) to (B.6). There are two smooth pasting conditions representing the optimality of thresholds. It is important to distinguish the “decision maker” in different cases. First consider the pooling IPO case. If firm \(h\) moves first, it must be indifferent between waiting and initiating for IPO, knowing that it would be followed by the low type. Firm \(l\) has a dominant strategy by mimicking and therefore has no indifferent conditions binding at this case. Thus the “decision maker” in pooling IPO 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. Acquisition is a winning scenario for the high type since its true quality is perfectly revealed.
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 (B.7) and (B.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 $J_l(\rho_{lt})$:

\[
(\mu_h - r)J_h(\rho_{lt}) - \left(\frac{\mu_h - \mu_l}{2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2}\right)J_h'(\rho_{lt}) + \frac{1}{2} \frac{(\mu_h - \mu_l)^2}{\sigma^2}J_h''(\rho_{lt}) = 0 \tag{4}
\]

\[
(\mu_l - r)J_l(\rho_{lt}) - \left(-\frac{\mu_h - \mu_l}{2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2}\right)J_l'(\rho_{lt}) + \frac{1}{2} \frac{(\mu_h - \mu_l)^2}{\sigma^2}J_l''(\rho_{lt}) = 0 \tag{5}
\]

It is straightforward to show $J_h(\rho_{lt}) = C_1h.exp(\theta_{1h}\rho_{lt}) + C_2h.exp(\theta_{2h}\rho_{lt})$ and $J_l(\rho_{lt}) = C_1l.exp(\theta_{1l}\rho_{lt}) + C_2l.exp(\theta_{2l}\rho_{lt})$. $\theta_{ij}$s are known constants as the roots of the characteristic functions in equations (4) and (5). The four constants $C_{ij}$s are the coefficients that will be pinned down together with the two free-boundaries $\eta$ and $\beta$ 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 $\frac{L-\alpha+\frac{\alpha(H-L)}{H+L}}{L-\alpha+\gamma}$ and $\frac{H-\alpha+H-L-\gamma}{H-\alpha-\frac{\alpha(H-L)}{H-L}}$ represents the ratio of player’s winning payoff to its losing one if h’s stopping strategy is $\eta \to 0$\(^{18}\). The assumption guarantees that $l$ is sufficiently more resistant in waiting than $h$ 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 $\frac{L-\alpha+\frac{\alpha(H-L)}{H+L}}{L-\alpha+\gamma} \gg \frac{H-\alpha+H-L-\gamma}{H-\alpha-\frac{\alpha(H-L)}{H+L}} > 1$.

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

3.4. Existence and Uniqueness

Figure 2 illustrates the main idea of the two-threshold strategy. In both figures, the horizontal axis is the mistaken belief $\rho_{lt}$ and the vertical axis is 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,

\(^{18}\)The boundary value 0 is mechanically selected as initial belief at $t = 0$ is $\rho_t = 0$. Technically a negative threshold $\eta$ can exist with other assumptions. But then initial condition must be adjusted. Otherwise the game ends into pooling IPO immediately at $t = 0$. 

This figure plots the simulated value functions $J_i$s. 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.

Consider acquisitions first. Suppose the high type wants to acquire the low type before $\rho_{lt}$ falls below $\beta$, 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 $\rho_{lt}$ decreases, the case when investors become more pessimistic about the low type. This is because the belief is farther away from the threshold that triggers pooling, meaning the chances that the low type could deceive investors reduces. At the point when $\rho_{lt}$ is equal to $\beta$, the discounted payoff of waiting for pooling equals to the current acquisition offer. As now firm $l$ becomes indifferent, it accepts acquisition by firm $h$.

The logic of pooling IPO follows similarly. Before belief $\rho_{lt}$ crosses $\eta$, high type’s waiting value is strictly higher than its payoff in pooling IPO. Thus it has no incentive to go stand-alone IPO first and let the low type imitate. So why does $h$ becomes indifferent at $\eta$? At any moment firm $h$ is balancing between two types of costs. At one hand, it could always go public without acquisition and suffer a short-term underpricing cost in pooling. At the other hand, it could choose to wait the mistaken belief to adjust. This will take additional time and waiting generates delaying cost. If the current belief is extremely mistaken and far away from $\beta$, adjustment takes so long that cost in postponing expansion outweighs the cost
of underpricing. This explains why the high type will give up if $\rho_l$ is greater or equal to $\eta$.

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

(i) Merger: If $(x_{ht}, x_{lt}) \in M$, $\sigma^A_{ht}(x_{ht}, x_{lt}) = \sigma^T_{lt}(x_{ht}, x_{lt}) = 1$. Otherwise $\sigma^A_{ht}(x_{ht}, x_{lt}) = \sigma^T_{lt}(x_{ht}, x_{lt}) = 0$.

(ii) Pooling IPO: If $(x_{ht}, x_{lt}) \in P$, $\sigma^I_{ht}(x_{ht}, x_{lt}) = \sigma^I_{lt}(x_{ht}, x_{lt}) = 1$. Otherwise $\sigma^I_{ht}(x_{ht}, x_{lt}) = \sigma^I_{lt}(x_{ht}, x_{lt}) = 0$.

(iii) For all $(x_{ht}, x_{lt})$, $\sigma^A_{lt}(x_{ht}, x_{lt}) = \sigma^T_{ht}(x_{ht}, x_{lt}) = 0$.

The strategies in Theorem 3 are essentially unique. First, at boundaries of $P$ and $M$, players are indifferent between waiting and stopping. Therefore, there also exists a continuum of mixed strategies that generate the same level of payoffs for the decision-making type but strictly lower payoffs for the other type at the boundary. However, firms payoffs differ from the strategies in Theorem 3 with zero measure. Second, when acquisition happens with zero conditional probability, it is sufficient to have one type’s strategy stops the deal from happening. For example, the high type could also adopt a strategy with always bidding, $\sigma^A_{ht}(x_{ht}, x_{lt}) \equiv 1$. But the low type will reject optimally so these strategies will also not impact the equilibrium payoffs.

To show this is an equilibrium, it remains to verify that no firms will deviate. At boundaries, the decision-making type is indifferent by smooth pasting conditions and the other type’s optimality follows by dominance. Thus it only requires to verify that there is no deviation inside boundaries.

First, in the IPO stage, no firm will deviate for stand-alone IPO alone when $\rho_l \in (\beta, \eta)$. The reason is that conditional on observing a stand-alone single IPO, the refined off-equilibrium belief under D1 is such deviating firm is the low type for certainty. 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, an action of investor is an equity contract with different share price. This can be pinned down by a pseudo belief $\tilde{\rho}$ as there is a one-to-one mapping function from belief to share price. Suppose the current belief is $\rho' \in (\beta, \eta)$. As shown in Figure 2, firm $h$’s equilibrium waiting value is higher than the IPO offer defined with $\rho'$ (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 current IPO offer and therefore the lowest
deviating offer has a belief $\tilde{\rho}_l < \rho'$. This implies the offers, i.e. actions, 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 by such a deviation, investors will believe the deviating firm is the low type for sure.

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 and 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 transfer 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:

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

This implies it is impossible to find a transfer payment that satisfies both (i) and (ii) at the same time. The economics behind is that in expectation $\rho_{lt}$ drifts down strictly. Thus the exogenous learning process works towards the acquisition outcome 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.

**Theorem 4** The strategies characterized in Theorem 3 is a Markov Perfect Bayesian Equilibrium.

Theorem 4 implies that acquisition is a good signal of company quality in IPO. It is a selection process that only keeps the high type afterwards. On the contrary, when pooling IPO happens, both firms are not acquiring and appear indistinguishable. In this case both the high and low types are mixing in the composition of IPOs. 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.

My result emphasizes on connecting the underlying quality of firms with their acquisition roles. M&A happens when high type’s assets size is sufficiently larger than the low type’s. This sounds like a trivial story as we share the common prior that large firms acquire smaller ones. But the emphasis goes beyond this observation. In fact, M&A deals will not happen if
on the contrary low type has sufficiently large size. Instead both types would file for IPO. “Big-acquirer” and “small-target” situation superficially coincides with the timing that the low type finds resistance unprofitable.

A sideline implication is that valuation of a single startup alone does not fully determine the timing of IPO. $\rho_l$ is a function of assets size ratios. 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, 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 $\rho_l \in (\beta, \eta)$.

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, $(L - \alpha) x_l$. This is strictly lower than the acquisition offer $(L - \alpha + \gamma_l) x_l$. As a result, the low type will only exit in takeover by the high type. Thus, this game becomes a standard real option model in which firm $h$ decides when to acquire.

While $h$ is balancing between waiting or making an acquisition offer right away, $\rho_l$ is surprisingly still a relevant state variable. However, it now purely quantifies the ratio of asset sizes, which in turn affects firm $h$’s acquisition payoff because larger $x_l$ generates larger synergy. The question is whether $h$ will delay its acquisition as in the two-threshold equilibrium. The following proposition shows that, information asymmetry is the only reason that deters efficiency reallocation.

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

This is because adverse selection in IPO markets creates an outside option for firm $l$ as pooling becomes possible. The resistance of the low type generates a waiting zone. On the contrary, when there is no information asymmetry, firm $l$ will accept the acquisition offer regardlessly. In this case, firm $h$ would delay only if waiting expects to generate larger synergy. In other words, firm $h$ would prefer a relatively larger size of firm $l$. However at any

\footnote{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}
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 difference highlights the fundamental influence from adverse selection in the second-best case.

4. Extensions

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 \( \mu_i \) in a simultaneous subgame. In this section, I show that making those investments will hurt 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 observed 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 current solid assets size is driven by capital injections rather than underlying quality. Then investors will update belief more conservatively given the same realization of sizes.

To see it in the model, recall that the belief process follows

\[
\mathrm{d}\rho_{lt} = -\frac{(\mu_h - \mu_l)^2}{\sigma^2} \mathrm{d}t + \frac{\mu_h - \mu_l}{\sigma} \mathrm{d}B_t.
\]

Imagine that firm \( h \) increases its growth rate \( \mu_h \). 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 \( \mu_l \) is just the opposite.

Decreasing the mistaken belief \( \rho_{lt} \) 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 \( \mu_h \) 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 \( l \) is similar. A larger \( \mu_l \) decreases the drift (beneficial for firm \( l \)) and volatility (costly for firm \( l \)) simultaneously.

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

The first statement indicates that as $\mu_h$ increases, the threshold for pooling IPO, $\eta$, is closer to the initial belief but the threshold for acquisition, $\beta$, is further. This indicates the game is more likely to end in the pooling IPO case. Therefore the expected payoff for firm $h$ is lower. The second statement implies that as $\mu_l$ 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 $\beta$. By equation (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 and becomes greater than its payoff as being acquired. 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 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 and is strictly smaller than the payoff of pooling IPO at original $\eta$. Firm $h$ would have been better off if he went public earlier than $\eta$, which explains why $d\beta/d\mu_h < 0$.

One related question is about increasing growth rates of the industry as a whole. There are 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.²⁰ The previous result highlights that individually scaling up too quickly is 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 $\delta$ and increasing $\mu_h$ and $\mu_l$ by the same degree will lower both $\eta$ and $\beta$. The game ends more likely in the pooling IPO case.

The consequence of growth fights is that the threshold for pooling IPO $\eta$ is closer initially but the threshold for acquisition $\beta$ is further. Therefore, pooling IPO 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 (4) and (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 rate of $\mu_i$ (size effect), 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.

4.2. Market Volatility

Startups in an emerging industry often face huge uncertainty associated with growth. In the model, this is reflected by more volatile assets size and larger $\sigma^2$. How will startups respond if the uncertainty increases? Increased market volatility has mathematically the same effect as reducing the wedge between $\mu_h - \mu_l$. Size difference becomes a less precise signal and investors tend to be more conservative in updating their belief.

Just as Section 4.1 shows, lowering the drift value benefits firm $l$ but hurts firm $h$. Yet reduced volatility of belief makes firm $l$ more likely to accept the offer. As shown before, the volatility effect dominates the drift effect. Thus as firm size becomes more volatile, firm $l$ is worse off as the acquisition threshold is closer to the initial belief but pooling IPO threshold is further.

**Proposition 8** The two thresholds shift upward simultaneously as $\sigma$ increases, i.e., $d\eta/d\sigma^2$, $d\beta/d\sigma^2 > 0$. Acquisition becomes more likely. As a result, ex ante firm $h$ is better off while firm $l$ is worse off, $\partial V_h(x_{l0}, \rho_{lo})/\partial \sigma^2 > 0$ and $\partial V_l(x_{l0}, \rho_{lo})/\partial \sigma^2 < 0$.

It may seem surprising that the high type is better off when the exogenous learning channel becomes less precise. However, in the model there are two groups of asset takers, the passive uninformed buyers and the (possible) acquirer firm $h$. As firm $h$ perfectly knows the type of firm $l$, it is not affected by the reduced precision of signals. In other words, increased volatility actually increases the information advantage by firm $h$ to IPO investors. What indeed happens is a shift of trading opportunity from the uninformed buyers to informed
Figure 3: The Effect of Increased Volatility

This figure plots the simulated value functions $J_i$s before and after $\sigma^2$ increases. 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 line are equilibrium waiting functions before increase. Dash-dot lines are equilibrium waiting functions after increase.

buyers. With increased information asymmetry, IPO investors would rationally trade less aggressively. This makes firm $l$’s outside option decreases in value. So now the high type can take over the low type more easily.

The proposition states that when investors are less informative about startup quality, it is less likely that the game ends in a pooling IPO wave. Empirically this corresponds to a decrease in number of IPOs at time when information asymmetry is huge, consistent with findings in Lowry (2003).

4.3. Nash Bargaining

In this section I endogenize 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 firm $h$ suffers maximal underpricing cost below its NPV. Denote $\Delta$ as the markup that firm $l$ gets after bargaining, which is determined through

$$
\max_{\Delta} \Delta^{1-\xi} \left( (H - \alpha) (x_{ht} + x_{lt}) - \Delta - (L - \alpha) x_{lt} - \left( H - \frac{H}{L} \alpha \right) x_{ht} \right)^{\xi} \\
= \max_{\Delta} \Delta^{1-\xi} \left( (H - L) x_{lt} + \alpha \frac{H - L}{L} x_{ht} - \Delta \right)^{\xi}. \tag{6}
$$
\( \xi \geq 0 \) is the bargaining power of \( h \). The Nash Bargaining solution \( \Delta^* \) that solves equation (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 \( h \) expand. Second, \( l \) is in 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

\[
R^\text{NB}_h (x_{ht}, x_{lt}) = \left( H - \alpha - (1 - \xi) \alpha \frac{H - L}{L} \right) x_{ht} + \xi (H - L) x_{lt},
\]

\[
R^\text{NB}_l (x_{ht}, x_{lt}) = (L - \alpha + (1 - \xi) (H - L)) x_{lt} + (1 - \xi) \alpha \frac{H - L}{L} x_{ht}.
\]

In equation (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. At the extreme case when \( \xi = 0 \), firm \( h \) earns no synergy and suffers an extreme underpricing cost as if it is regarded as the low type for sure. Therefore it will never initiate a takeover on firm \( l \). In effect, when firm \( h \) has a small bargaining power, the only possible equilibrium outcome is pooling IPO. The high type optimally chooses stopping time \( \tau \) to go public first and let the low type mimic.

**Theorem 9** There exists \( \bar{\xi} \) such that if \( \xi \leq \bar{\xi} \), firm \( h \) never acquires and goes for stand-alone IPO when it moves first with belief \( \rho_{lt} \leq \eta^\text{NB} \leq 0 \). 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 IPO. 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.\(^21\)

**Theorem 10** 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.

\(^{21}\)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.
5. Empirical Implications

In this section I test the empirical implications of baseline model. Section 5.1 explains the construction of sample and discusses model implications in line with existing literature. Section 5.2 tests that IPOs with private acquisitions have lower first day return and Section 5.3 shows they also have better long run performance. Section 5.4 simulates a marketing timing framework where investors are fully rational.

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

Figure 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.

Private M&A data are from SDC Merger and Acquisition database, 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 substantial amount of acquisitions made by private companies. For example, Figure 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 acquisition is trending up in recent years. The average portion of private acquirers is 44.2%, which is comparable to both Maksimovic et al. (2013) and Netter et al. (2011). Among these private deals, around one-third of them happens where the involved companies were competitors in the same industry. Given the M&A sample, I select 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 SDC Platinum and SDC Merger and Acquisition provide CUSIP for issuers and acquirers, I cannot directly use it as the identifiers. This is because both CUSIP and company name documented in Merger and Acquisition database are historical upon the deal time. It is very likely that by the IPO time, the company was assigned a new CUSIP and possibly changed its firm name. For example, Snap Inc has CUSIP id “83304A” in issuance database whereas it has the following name and CUSIP combinations in M&A deals: “Snapchat Inc, 7A2488”, “Snap Inc, 7A2488”, “Snap Inc, 83304A”, “Snap Inc, 9E4450”.

I take the following procedures to potentially match cases such as “Snapchat Inc, 7A2488” to “Snap Inc, 83304A”. First I standardize both issuer’s and acquirer’s company name using NBER Patent Data Project’s file. Then I put each standardized company name in a separate name set $NS_j$ indexed by 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.

Issuers and acquirers are matched based on the final converged key. I drop all acquisitions if either acquirer has a public company status or the deal announcement date is weakly later than the issuance date. 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.

---

22 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 has CRSP price data with sample of completed mergers and acquisitions announced between 1992 and 2009.

23 I drop acquisition deals that happen on the IPO issue date to distinguish from rollup IPOs.
Figure 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.

Stock return data are taken from CRSP, merged first through CUSIP and then manual supplements. IPOs founding dates are downloaded from Jay Ritter’s website or searched through Google.

Figure 5 plots number of IPOs in different categories. In general there are more IPOs without private M&As. The ratio between acquiring IPOs to non-acquiring is approximately from 1:1 to 1:2. The number of acquiring IPOs is quite stable, around 50 per year, except the recent depression in 2008 and 2009. In each year, more than half of the acquiring IPOs have made acquisitions on their competitors in the same industry.

Before I move to the empirical tests, I discuss two implications of the model related with 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 consistent with the fact that private acquisitions respond less business cycles and less wavelike. In fact, 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 growth is strong, 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 since they believe IPOs opportunities are
near future. This forces the more efficient firms to withdraw their acquisition offer. In the end, we observe less private acquisitions and more non-acquiring IPOs.

The second implication is that due to information imperfections, efficient reallocations are blocked. Therefore, a firm with higher productivity may not be able to acquire the 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 perfect 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.

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, we should observe that IPOs with acquisitions have 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. Besides univariate comparison, I also estimate the following linear multivariate regression model:

\[
Firstret_{i,t} = \beta MA + \gamma Controls + \eta_t + \mu_j + \epsilon_{i,t}
\]  

(9)

In Equation (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 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. I also
consider including IPO year fixed effects $\eta_t$ and industry fixed effects $\mu_j$. If $\mu_j$ is specified, then \textit{HighTech} is dropped due to multilineararity.

[Table 1 Here]

Table 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. Since those variables also potentially help investors discern company qualities, it is possible that the lower underpricing is driven by those signals rather than previous acquisitions. The two groups are similar in other dimensions.

[Table 2 Here]

Column (2) in Table 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 3 Here]

Table 3 provides the regression results. Having acquisitions significantly reduces underpricing in all specifications except Column (1) 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.

To cut the clutter I omitted estimations on control variables but they largely conform with previous literature. As in Loughran and Ritter (2004), firm age significantly reduces, and operation in high tech industry increases first-day return. Total assets is insignificant due to its common component with firm age and becomes negatively significant if age variable is dropped. In line with recent literature (Loughran and McDonald, 2013, for eg.), I document significant increase in underpricing if IPO is VC backed. Lastly, though insignificant, having more bookrunners slightly reduces but listing in Nasdaq increases underpricing.

To summarize, I find having acquisitions before IPO indeed significantly reduces first-day return and increases proceeds for the issuing company. Consistent with model implication,
private acquisition is a positive signal even after controlling for correlation with firm size, age and VC fundings.

5.3. Longterm Stock Return

The model highlights that private acquisition is related with the better fundamental quality of startups. The fact that having acquisitions is viewed as a good signal and boosts offer price is due to the underlying efficiency of future growth. In other words, investors expect these companies to have better performance in the long run. 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. As mentioned in the introduction, 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). Their results indicate that consolidation of markets alone is not sufficient to sustain long term growth. Rational investor should also not reward private acquisitions if they purely help acquirers of random qualities with more monopoly power.

On the contrary, Columns (4)—(6) in Table 2 document consistently better performance in terms of three-year or four-year buy-and-hold return for IPOs with private acquisitions. In the whole samples, acquiring IPOs outperform by around 9% and those with competitor acquisitions have even larger premium. The result survives after adjusting by market returns or restricting to high tech industries. In the whole industry sample, all types of IPOs seem to underperformed than market. I decompose the full sample into different time periods and find that weaker returns concentrates in IPOs that go public between 2004—2007. Post 2007 market adjusted returns are positive for IPOs with competitor acquisitions.

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 IPOs concentrate before periods of low market performance. These evidence seem 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 expected loss in the low types is compensated by the underpayment of the high types. Investors are break-even given their belief as in Schultz (2003) with zero expected profit.
But realized performance of the new public firms varies with the quantity of new IPO firms. In an IPO wave, there are more pooling IPOs. Thus, the post-wave return rates would be mixed by two types. Off the wave, there are more private acquisitions and the more efficient firms are self-selected. The variation of performance is due to the different characteristics of firms that enter into the primary on or off the wave, i.e. a composition effect, which coincides with the quantity 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 prior to that.

To illustrate this idea, I simulate the model in the following way.\(^{24}\) Recall that 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 those total returns in a perpetuity fashion. The per period return for good firm is \(r_h = r(H - \alpha) / \alpha\) and for bad 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 \(t = k \cdot \Delta t\), \((dB_{nh}^n, dB_{nl}^n)\) is drawn independently for the \(n^{th}\) pair of startups. If \(\rho_{lt}^n\) hits the thresholds \(\beta\) or \(\eta\), then the pairs take strategies in Theorem 3. Meanwhile, a new pair of startups will fill in the vacancy left by the newly public firms. At \(k^{th}\) step of simulation, I document the number of new public firms and the average per period 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 \(\eta\) and \(\beta\) 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\(^{th}\) since the distribution of \(\rho_{lt}^{1500}\) among the potential pairs is stationary at then. Figure 6 illustrates the result of simulation. 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 a period where 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 \(n\). The return difference between the top and bottom decile is -2.49\% with a \(p\)-value less than 1\%. Without referring to the model, this pattern indeed seems like poor firms take advantage of investors optimism in hot IPO markets.

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

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

\(^{24}\)This simulation is a qualitative exercise intended to illustrate the pattern of pseudo market timing.
Investors are no longer purely Bayesian learners. At any moment, they first update \( \rho_{lt} \) 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 \( \delta \) could be due to inherent misconception of investors or firm \( l \)'s marketing strategies. Given the new belief, the Bellman Equation of waiting functions \( J_i \) follows:

\[
(\mu_h - r) J_h (\rho_{lt}) - \left( \frac{\mu_h - \mu_l}{2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2} - \delta \right) J_h' (\rho_{lt}) + \frac{1}{2} \frac{(\mu_h - \mu_l)^2}{\sigma^2} J_h'' (\rho_{lt}) = 0 \tag{10}
\]

\[
(\mu_l - r) J_l (\rho_{lt}) - \left( -\frac{\mu_h - \mu_l}{2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2} - \delta \right) J_l' (\rho_{lt}) + \frac{1}{2} \frac{(\mu_h - \mu_l)^2}{\sigma^2} J_l'' (\rho_{lt}) = 0 \tag{11}
\]

Unlike changing \( \sigma^2 \) and \( \mu_i \), a biased belief by \( \delta \) only creates drift effect. The mistaken \( \rho_{lt} \) now becomes more persistent and 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 \( \eta' < \eta \). Compare this result with Section 4.1 and 4.2. Irrationality generates more opportunities for the less efficient types to join in IPO waves. On the contrary, rational investors will account for manipulation of beliefs and adjust their way of learning.
Proposition 11  The two thresholds shift downward simultaneously as overoptimism $\delta$ is imposed, i.e., $d\eta/d\delta, d\beta/d\delta < 0$. Pooling IPO becomes more likely. As a result, ex ante firm $l$ is better off and firm $h$ is worse off $\partial V_l(x_{l0}, \rho_{l0})/\partial \delta > 0$ and $\partial V_l(x_{l0}, \rho_{l0})/\partial \delta < 0$.

The main takeaway of this section is that realized poor performance of new public companies may not be necessarily due to irrationality of investors. Instead, it could be due to a composition effect. However, investor overoptimism indeed generates market timing opportunities for the less efficient startups. The empirical documents of market timing may be a mix of the above two channels.

6. Conclusion

Acquisitions by private companies are understudied in 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 option model with information imperfections, the more efficient firm can initiate takeovers and therefore signal its quality. However, asset reallocations are blocked with certain probability. Low types may resist to sell because they could possibly pool 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.

In terms of efficiency, private acquisitions help eliminate financing uncertainty and make fundings allocated to better technology. Thus, it is of importance to understand how to increase the frequency of private acquisitions. I show that for the more efficient startups, they should not try to overscale and boost financial performance in order to impress investors. The reason is that rational investors would take the expenditure into consideration and consider that assets growth are driven by high investment rates rather than underlying technology. This in turn gives poor quality firms opportunities to imitate. Especially, private acquisitions become less possible in an cash burning arm-race.

In contrast to IPOs with acquisitions simultaneously or soon afterwards, I document that the issuing firms which have acquired their competitors previously perform better over the long term. This confirms the theoretical result that private acquisitions indicate better growth.
opportunities. Investors rationally take the deals as good signals so that these companies have less underpricing and more proceeds in IPO.
References


Maksimovic, V. and Phillips, G. (2002). Do Conglomerate Firms Allocate Resources ineffi-


Appendix A. Tables

Table 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.

<table>
<thead>
<tr>
<th></th>
<th>(1) Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Full Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Offer Price</em></td>
<td>14.11</td>
<td>6.02</td>
<td>5.00</td>
<td>85.00</td>
<td>1537.00</td>
</tr>
<tr>
<td><em>Proceeds (Mil.)</em></td>
<td>207.05</td>
<td>634.08</td>
<td>3.50</td>
<td>16006.90</td>
<td>1537.00</td>
</tr>
<tr>
<td><em>Total Assets (Mil.)</em></td>
<td>854.04</td>
<td>4744.24</td>
<td>0.20</td>
<td>137238.00</td>
<td>1475.00</td>
</tr>
<tr>
<td><em>Age</em></td>
<td>19.44</td>
<td>23.41</td>
<td>1.00</td>
<td>166.00</td>
<td>1530.00</td>
</tr>
<tr>
<td><em>VC</em></td>
<td>0.48</td>
<td>0.50</td>
<td>0.00</td>
<td>1.00</td>
<td>1537.00</td>
</tr>
<tr>
<td><em>High Tech (dummy)</em></td>
<td>0.31</td>
<td>0.46</td>
<td>0.00</td>
<td>1.00</td>
<td>1537.00</td>
</tr>
<tr>
<td><em>Bookrunners</em></td>
<td>2.68</td>
<td>1.73</td>
<td>1.00</td>
<td>13.00</td>
<td>1537.00</td>
</tr>
<tr>
<td><em>Nasdaq (dummy)</em></td>
<td>0.66</td>
<td>0.47</td>
<td>0.00</td>
<td>1.00</td>
<td>1537.00</td>
</tr>
<tr>
<td><strong>Panel B: MA = 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Offer Price</em></td>
<td>14.93</td>
<td>6.93</td>
<td>5.00</td>
<td>85.00</td>
<td>524.00</td>
</tr>
<tr>
<td><em>Proceeds (Mil.)</em></td>
<td>297.28</td>
<td>1033.66</td>
<td>7.00</td>
<td>16006.90</td>
<td>524.00</td>
</tr>
<tr>
<td><em>Total Assets (Mil.)</em></td>
<td>1508.97</td>
<td>7803.03</td>
<td>0.20</td>
<td>137238.00</td>
<td>503.00</td>
</tr>
<tr>
<td><em>Age</em></td>
<td>22.27</td>
<td>25.47</td>
<td>1.00</td>
<td>166.00</td>
<td>522.00</td>
</tr>
<tr>
<td><em>VC</em></td>
<td>0.45</td>
<td>0.50</td>
<td>0.00</td>
<td>1.00</td>
<td>524.00</td>
</tr>
<tr>
<td><em>High Tech (dummy)</em></td>
<td>0.38</td>
<td>0.49</td>
<td>0.00</td>
<td>1.00</td>
<td>524.00</td>
</tr>
<tr>
<td><em>Bookrunners</em></td>
<td>2.97</td>
<td>1.87</td>
<td>1.00</td>
<td>13.00</td>
<td>524.00</td>
</tr>
<tr>
<td><em>Nasdaq (dummy)</em></td>
<td>0.54</td>
<td>0.50</td>
<td>0.00</td>
<td>1.00</td>
<td>524.00</td>
</tr>
</tbody>
</table>
Table 2: First-day and Long-run Returns on IPOs Categorized by M&A Before

This table provides results on average first-day return, 3-year and 4-year buy-and-hold returns categorized by whether the IPO company has acquisitions before. First-day return is defined as the return of closing price in the first trading day over offer price. Buy-and-hold return is cumulative return of stocks from its first trading day till the specified ending date, excluding first-day return. Column (4) and (6) are adjusted by value-weighted market returns.

<table>
<thead>
<tr>
<th>Panel A: All Industry</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>w/o Acq</td>
<td>1,012</td>
<td>14.724</td>
<td>13.384</td>
<td>-10.268</td>
<td>19.134</td>
<td>-11.192</td>
</tr>
<tr>
<td>w/ Competitor Acq</td>
<td>339</td>
<td>12.186</td>
<td>24.735</td>
<td>-1.303</td>
<td>29.574</td>
<td>-2.822</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: High Tech</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>w/o Acq</td>
<td>282</td>
<td>18.508</td>
<td>17.246</td>
<td>-4.667</td>
<td>31.216</td>
<td>2.409</td>
</tr>
<tr>
<td>w/ Acq</td>
<td>198</td>
<td>18.773</td>
<td>29.778</td>
<td>2.871</td>
<td>41.473</td>
<td>8.110</td>
</tr>
<tr>
<td>w/ Competitor Acq</td>
<td>118</td>
<td>17.825</td>
<td>28.966</td>
<td>0.627</td>
<td>41.014</td>
<td>5.864</td>
</tr>
</tbody>
</table>

43
Table 3: Regressions of First-day Returns on Private Acquisition Indicators

This table provides estimations on the following equation:

\[ \text{Firstret}_{i,t} = \beta M A + \gamma \text{Controls} + \eta_i + \mu_j + \epsilon_{i,t} \]

\( \text{Firstret}_{i,t} \) 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.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Private Acquisitions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( MA )</td>
<td>-1.646</td>
<td>-2.945*</td>
<td>-2.661**</td>
<td>-3.136*</td>
</tr>
<tr>
<td></td>
<td>(1.299)</td>
<td>(1.614)</td>
<td>(1.292)</td>
<td>(1.612)</td>
</tr>
<tr>
<td>( Controls )</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>( Year Fixed Effects )</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>( Industry Fixed Effects )</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>#Obs</td>
<td>1536</td>
<td>1391</td>
<td>1467</td>
<td>1323</td>
</tr>
<tr>
<td>( Adj. R^2 )</td>
<td>0.00</td>
<td>0.06</td>
<td>0.04</td>
<td>0.09</td>
</tr>
</tbody>
</table>

| **Panel B: Private Acquisitions of Competitors** |           |          |          |          |
| \( M A^{comp} \) | -2.537*  | -3.554** | -3.017** | -3.725** |
|                  | (1.362)  | (1.655)  | (1.348)  | (1.660)  |
| \( Controls \)   | N        | N        | Y        | Y        |
| \( Year Fixed Effects \) | N        | Y        | Y        | Y        |
| \( Industry Fixed Effects \) | N        | Y        | N        | Y        |
| \#Obs            | 1536     | 1391     | 1467     | 1323     |
| \( Adj. R^2 \)   | 0.00     | 0.06     | 0.04     | 0.09     |
Appendix B. Omitted Equations in Section 3.3

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

\[
\begin{align*}
    rV_h(x_{ht}, \rho_{lt}) &= \mu_h x_{ht} \frac{\partial V_h(x_{ht}, \rho_{lt})}{\partial x} + \frac{1}{2} \frac{\sigma^2}{\sigma^2} x_{ht}^2 \frac{\partial^2 V_h^2(x_{ht}, \rho_{lt})}{\partial x^2} \\
    &\quad + \frac{1}{2} (\mu_h - \mu_l) x_{ht} \frac{\partial V_h^2(x_{ht}, \rho_{lt})}{\partial x} + \frac{1}{2} (\mu_h - \mu_l)^2 \frac{\partial V_h^2(x_{ht}, \rho_{lt})}{\partial \rho^2} \\
    &= \text{Size Effect} \\
    + \frac{1}{2} (\mu_h - \mu_l) x_{ht} \frac{\partial V_h^2(x_{ht}, \rho_{lt})}{\partial x} \\
    &= \text{Belief Effect} \\
    \end{align*}
\]

(B.1)

\[
\begin{align*}
    rV_l(x_{lt}, \rho_{lt}) &= \mu_l x_{lt} \frac{\partial V_l(x_{lt}, \rho_{lt})}{\partial x} + \frac{1}{2} \frac{\sigma^2}{\sigma^2} x_{lt}^2 \frac{\partial^2 V_l^2(x_{lt}, \rho_{lt})}{\partial x^2} \\
    &\quad + \frac{1}{2} (\mu_h - \mu_l) x_{lt} \frac{\partial V_l^2(x_{lt}, \rho_{lt})}{\partial x} + \frac{1}{2} (\mu_h - \mu_l)^2 \frac{\partial V_l^2(x_{lt}, \rho_{lt})}{\partial \rho^2} \\
    &= \text{Size Effect} \\
    + \frac{1}{2} (\mu_h - \mu_l) x_{lt} \frac{\partial V_l^2(x_{lt}, \rho_{lt})}{\partial x} \\
    &= \text{Belief Effect} \\
    \end{align*}
\]

(B.2)

When the game ends, firms receive their payoffs correspondingly:

\[
\begin{align*}
    V_h(x_{ht}, \eta) &= (H - \alpha - \alpha (H - L) \exp(\eta)) x_{ht} \\
    V_l(x_{lt}, \eta) &= (L - \alpha + \alpha (H - L) \exp(\eta)) x_{lt} \\
    V_h(x_{ht}, \beta) &= (H - a) x_{ht} + (H - L - \gamma_l) x_{lt} \\
    V_l(x_{lt}, \beta) &= (L - a + \gamma_l) x_{lt}.
\end{align*}
\]

(B.3 - B.6)

The two smooth pasting conditions are

\[
\begin{align*}
    \frac{\partial V_l(x_{ht}, \beta)}{\partial \beta} &= 0 \\
    \frac{\partial V_h(x_{ht}, \eta)}{\partial \eta} &= \frac{-\alpha (H - L) \exp(\eta)}{(H + L \exp(\eta))^2} x_{ht}.
\end{align*}
\]

(B.7 - B.8)
The last step is to transform all the boundary conditions in forms of $J_i$:

\[
J_h(\eta) = H - \alpha - \frac{\alpha(H - L)\exp(\eta)}{(H + L\exp(\eta))}
\]

\[
J_l(\eta) = L - \alpha + \frac{\alpha(H - L)\exp(\eta)}{(H\exp(\eta) + L)}
\]

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

\[
J_l(\beta) = L - a + \gamma_l
\]

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

\[
J_l'(\beta) = 0.
\]

Combine the above boundary conditions with equation (4) and (4) and solve $(\beta, \eta)$. 
Appendix C. Proofs

Proof of Lemma 1

Proof. Fixing a pair \((x_1, x_2)\), there exists no equilibria in which \(\sigma^A_{ht}(x_1, x_2) \sigma_T^{T}(x_1, x_2) > 0\) and \(\sigma^A_{lt}(x_1, x_2) \sigma_T^{T}(x_1, x_2) > 0\). 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,

\[
\frac{H(x_1 + x_2)(1 - s_t)}{h\text{-acquirer total net value}} = \frac{L(x_1 + x_2)(1 - s_t)}{l\text{-acquirer total net value}}.
\]

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

Next, fixing \(x_{mt}\), the following type of equilibria cannot exists: there exists two pairs \((x_1, x_2)\) and \((x'_1, x'_2)\), \(x_1 + x_2 = x'_1 + x'_2 = x_{mt}\) and \(x_1 \neq x'_1\), such that:

\[
\sigma^A_{ht}(x_1, x_2) \sigma_T^{T}(x_1, x_2) > 0, \quad \sigma^A_{lt}(x_1, x_2) \sigma_T^{T}(x_1, x_2) = 0,
\]

\[
\sigma^A_{ht}(x'_1, x'_2) \sigma_T^{T}(x'_1, x'_2) = 0, \quad \sigma^A_{lt}(x'_1, x'_2) \sigma_T^{T}(x'_1, x'_2) > 0.
\]

In other words, the high type acquires only if the sizes are \((x_1, x_2)\) and the low type acquires only if the sizes are \((x'_1, x'_2)\). Firm \(l\)'s payoff at \((x'_1, x'_2)\) is \((1 - s_t) Lx_{mt} - (H - \alpha + \gamma_{h}) x'_1\).

The following deviation is strictly profitable:

\[
\tilde{\sigma}^A_{ht}(x'_1, x'_2) \tilde{\sigma}^{T}_{lt}(x'_1, x'_2) = 1, \quad \tilde{\sigma}^A_{lt}(x'_1, x'_2) \tilde{\sigma}^{T}_{ht}(x'_1, x'_2) = 0,
\]

while the high type still gets \((H - \alpha + \gamma_{h}) x'_1\). Notice this deviation will not affect investors belief and therefore \(s_t\). This deviation generates \((1 - s_t) Hx_{mt} - (H - \alpha + \gamma_{h}) x'_1\) for firm \(l\).

Lastly, it is intuitive that there exits no equilibria in which fixing \(x_{mt}\), for all \((x_1, x_2)\), \(x_1 + x_2 = x_{mt}\),

\[
\sigma^A_{ht}(x_1, x_2) \sigma_T^{T}(x_1, x_2) = 0, \quad \sigma^A_{lt}(x_1, x_2) \sigma_T^{T}(x_1, x_2) > 0.
\]

As firm \(l\) is strictly better off by accepting takeovers by firm \(h\).

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

\[25\] Off equilibrium acquisition offer does not have to equal exogenous reservation value.
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 (3), firm $l$ must be adjusted to $x'_{lt} = \alpha x_{lt}$.

Consider the new firm $h$’s problem:

\[
\begin{align*}
\sup_{\sigma^i} & \quad \mathbb{E}^h \left( \int_t^\tau e^{-\tau} \left( R^l_i \left( x'_{ht}, x'_{lt} \right) 1^l_\tau + R^m_i \left( x'_{ht}, x'_{lt} \right) 1^m_\tau \right) \right) | x'_{ht}, x'_{lt} \\
= \sup_{\sigma^i} & \quad \mathbb{E}^h \left( \int_t^\tau e^{-\tau} \left( R^l_i \left( x'_{ht} e^{\left( \mu_h - \frac{\sigma^2}{2} \right) (\tau-t)} + \frac{1}{\sqrt{2}} \sigma B_{ht} \right), x'_{lt} e^{\left( \mu_l - \frac{\sigma^2}{2} \right) (\tau-t)} + \frac{1}{\sqrt{2}} \sigma B_{lt} \right) 1^l_\tau \\
& \quad + R^m_i \left( x'_{ht} e^{\left( \mu_h - \frac{\sigma^2}{2} \right) (\tau-t)} + \frac{1}{\sqrt{2}} \sigma B_{ht} \right), x'_{lt} e^{\left( \mu_l - \frac{\sigma^2}{2} \right) (\tau-t)} + \frac{1}{\sqrt{2}} \sigma B_{lt} \right) 1^m_\tau \right) | x'_{ht}, x'_{lt} \\
= \sup_{\sigma^i} & \quad \mathbb{E}^h \left( \int_t^\tau e^{-\tau} \left( \alpha R^l_i \left( x'_{ht} e^{\left( \mu_h - \frac{\sigma^2}{2} \right) (\tau-t)} + \frac{1}{\sqrt{2}} \sigma B_{ht} \right), x'_{lt} e^{\left( \mu_l - \frac{\sigma^2}{2} \right) (\tau-t)} + \frac{1}{\sqrt{2}} \sigma B_{lt} \right) 1^l_\tau \\
& \quad + \alpha R^m_i \left( x'_{ht} e^{\left( \mu_h - \frac{\sigma^2}{2} \right) (\tau-t)} + \frac{1}{\sqrt{2}} \sigma B_{ht} \right), x'_{lt} e^{\left( \mu_l - \frac{\sigma^2}{2} \right) (\tau-t)} + \frac{1}{\sqrt{2}} \sigma B_{lt} \right) 1^m_\tau \right) | x'_{ht}, x'_{lt} \\
= \sup_{\sigma^i} & \quad \mathbb{E}^h \alpha \left( \int_t^\tau e^{-\tau} \left( R^l_i \left( x_{ht}, x_{lt} \right) 1^l_\tau + R^m_i \left( x_{ht}, x_{lt} \right) 1^m_\tau \right) \right) | x'_{ht}, x'_{lt} \\
= \sup_{\sigma^i} & \quad \mathbb{E}^h \alpha \left( \int_t^\tau e^{-\tau} \left( R^l_i \left( x_{ht}, x_{lt} \right) 1^l_\tau + R^m_i \left( x_{ht}, x_{lt} \right) 1^m_\tau \right) \alpha x_{ht}, \alpha x_{lt} \right) \\
= \sup_{\sigma^i} & \quad \mathbb{E}^h \alpha \left( \int_t^\tau e^{-\tau} \left( R^l_i \left( x_{ht}, x_{lt} \right) 1^l_\tau + R^m_i \left( x_{ht}, x_{lt} \right) 1^m_\tau \right) \right) x_{ht}, x_{lt}
\end{align*}
\]

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 IPO case) or independent of sizes (acquisition case). Therefore the stopping time $\tau$ will not be affected.

**Proof of Theorem 3**

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

**Lemma 12** (i) $\theta_{1h} < 0 < 1 < \theta_{2h}$, $\theta_{1l} < 0 < \theta_{2l}$; (ii) $\theta_{1h} + \theta_{2h} = 2 + (\mu_h - \mu_l)^{-1}$, $\theta_{1l} + \theta_{2l} = 2 - (\mu_h - \mu_l)^{-1}$; (iii) $\theta_{2h} - \theta_{2l} = \theta_{1h} - \theta_{1l} = (\mu_h - \mu_l)^{-1}$.
Proof. By solving the characteristic function, we have

\[
\theta_{1h} = \frac{\mu_h - \mu_l}{\sigma^2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2} - \sqrt{\left(\frac{\mu_h - \mu_l}{\sigma^2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2}\right)^2 - 2\frac{(\mu_h - \mu_l)^2}{\sigma^2}(\mu_h - r)}
\]

\[
\theta_{2h} = \frac{\mu_h - \mu_l}{\sigma^2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2} + \sqrt{\left(\frac{\mu_h - \mu_l}{\sigma^2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2}\right)^2 - 2\frac{(\mu_h - \mu_l)^2}{\sigma^2}(\mu_h - r)}
\]

\[
\theta_{1l} = -\frac{\mu_h - \mu_l}{\sigma^2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2} - \sqrt{\left(-\frac{\mu_h - \mu_l}{\sigma^2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2}\right)^2 - 2\frac{(\mu_h - \mu_l)^2}{\sigma^2}(\mu_l - r)}
\]

\[
\theta_{2l} = -\frac{\mu_h - \mu_l}{\sigma^2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2} + \sqrt{\left(-\frac{\mu_h - \mu_l}{\sigma^2} + \frac{(\mu_h - \mu_l)^2}{\sigma^2}\right)^2 - 2\frac{(\mu_h - \mu_l)^2}{\sigma^2}(\mu_l - r)}
\]

Notice

\[
\left(\frac{\mu_h - \mu_l}{\sigma^2}\right)^2 + \frac{(\mu_h - \mu_l)^2}{\sigma^2} - 2\frac{(\mu_h - \mu_l)^2}{\sigma^2}(\mu_h - r)
\]

\[
= (\frac{\mu_h - \mu_l}{\sigma^2})^2 + \left(\frac{\mu_h - \mu_l}{\sigma^2}\right)^2 + 2\left(\frac{\mu_h - \mu_l}{\sigma^2}\right)(\mu_h - \mu_l)^2 - 2\frac{(\mu_h - \mu_l)^2}{\sigma^2}(\mu_h - r)
\]

\[
= (\frac{\mu_h - \mu_l}{\sigma^2})^2 + \left(\frac{\mu_h - \mu_l}{\sigma^2}\right)^2 + 2\left(\frac{\mu_h - \mu_l}{\sigma^2}\right)(\mu_h - \mu_l)^2 - 2\frac{(\mu_h - \mu_l)^2}{\sigma^2}(\mu_l - r)
\]

\[
= (\frac{\mu_h - \mu_l}{\sigma^2})^2 + \left(\frac{\mu_h - \mu_l}{\sigma^2}\right)^2 - 2\frac{(\mu_h - \mu_l)^2}{\sigma^2}(\mu_l - r)
\]

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

\[
\frac{(\mu_h - \mu_l)}{\sigma^2} = k
\]

\[
\sqrt{\frac{(\mu_h - \mu_l)^2}{\sigma^2}} = \Delta
\]

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 \( \theta_{1l} \theta_{2l} = 2(\mu_l - r) < 0, \theta_{1h} \theta_{2h} = 2(\mu_h - r) < 0. \)

\[\blacksquare\]

Proof. To show the main results, we first replace the boundary conditions with explicit expression of \( J_h \) and \( J_l: \)
\[ C_{ih} \exp(\theta_{1h}\eta) + C_{2h} \exp(\theta_{2h}\eta) = H - \alpha - \frac{\alpha(H - L)\exp(\eta)}{H + L\exp(\eta)} \] (C.1)

\[ C_{1i} \exp(\theta_{1i}\eta) + C_{2i} \exp(\theta_{2i}\eta) = L - \alpha + \frac{\alpha(H - L)\exp(\eta)}{H\exp(\eta) + L} \] (C.2)

\[ C_{ih} \exp(\theta_{1h}\beta) + C_{2h} \exp(\theta_{2h}\beta) = H - \alpha + (H - L - \gamma_i)\exp(k\beta) \] (C.3)

\[ C_{1i} \exp(\theta_{1i}\beta) + C_{2i} \exp(\theta_{2i}\beta) = L - \alpha + \gamma_i \] (C.4)

\[ C_{1h} \theta_{1h} \exp(\theta_{1h}\eta) + C_{2h} \theta_{2h} \exp(\theta_{2h}\eta) = -\frac{\alpha H(H - L)\exp(\eta)}{(H + L\exp(\eta))^2} \] (C.5)

\[ C_{1i} \theta_{1i} \exp(\theta_{1i}\beta) + C_{2i} \theta_{2i} \exp(\theta_{2i}\beta) = 0 \] (C.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 12 (iii) to change \( \theta_{1h} \) and \( \theta_{2h} \) into \( \theta_{1i} \) and \( \theta_{2i} \) into (C.1) to (C.6):

\[ C_{1h} x^{\theta_{1i}} + C_{2h} x^{\theta_{2i}} = (H - \alpha - f(x))x^{-k} \] (C.7)

\[ C_{1i} x^{\theta_{1i}} + C_{2i} x^{\theta_{2i}} = L - \alpha + g(x) \] (C.8)

\[ C_{1h} y^{\theta_{1i}} + C_{2h} y^{\theta_{2i}} = (H - \alpha + (H - L - \gamma_i)y^k)y^{-k} \] (C.9)

\[ C_{1i} y^{\theta_{1i}} + C_{2i} y^{\theta_{2i}} = L - \alpha + \gamma_i \] (C.10)

\[ C_{1h} \theta_{1h} x^{\theta_{1i}} + C_{2h} \theta_{2h} x^{\theta_{2i}} = -f'(x)x^{-k} \] (C.11)

\[ C_{1i} \theta_{1i} y^{\theta_{1i}} + C_{2i} \theta_{2i} y^{\theta_{2i}} = 0 \] (C.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 (C.10) and (C.12) to solve \( C_{1i} \) and \( C_{2i} \) and (C.7) and (C.11) to solve \( C_{1h} \) and \( C_{2h} \). Then replace the solved constants in (C.8) and (C.9) respectively.

\[
\begin{align*}
[(H - \alpha - f(x))\theta_{2h} + f'(x)](\frac{x}{y})^{-\theta_{1i}} - [(H - \alpha - f(x))\theta_{1h} + f'(x)](\frac{x}{y})^{-\theta_{2i}} \\
= (\theta_{2i} - \theta_{1i})[(H - \alpha) + (H - L - \gamma_i)y^k](\frac{x}{y})^k
\end{align*}
\] (C.13)

\[
(L - \alpha + \gamma_i)(\frac{x}{y})^{\theta_{1i}} - \theta_{1i}(\frac{x}{y})^{\theta_{2i}} = (\theta_{2i} - \theta_{1i})(L - \alpha + g(x))
\] (C.14)

Instead of solving \( y \) directly, we solve \( m = \frac{x}{y} \geq 1 \) instead as it’s easier to deal with.
\[(H - \alpha - f(x))\theta_{2h} + f'(x)m^{-\theta_{1h}} - [(H - \alpha - f(x))\theta_{1h} + f'(x)m^{-\theta_{2h}} - (\theta_{2h} - \theta_{1h})(H - \alpha)x^k] = (\theta_{2h} - \theta_{1h})(H - \alpha + \gamma_l)x^k \quad (C.15)\]

\[(L - \alpha + \gamma_l)(\theta_{2i}m^{\theta_{1i}} - \theta_{1i}m^{\theta_{2i}}) = (\theta_{2i} - \theta_{1i})(L - \alpha + g(x)) \quad (C.16)\]

**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 (C.15) and (C.16) correspondingly.

To show the part of \(m_h(x)\), the LHS of (C.15) is

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

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

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

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

As \(\theta_{1i}\theta_{2l} < 0 \text{ and } \theta_{1l} < \theta_{2l}\), 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 (C.15)

\[
\frac{\partial A}{\partial m}m^k + kAm^{k-1} = (-\frac{\partial A}{\partial x}m^k + (\theta_{2h} - \theta_{1h})(H - \alpha)x^k)dx
\]

It remains to check \(\frac{\partial A}{\partial x}\).
\[
\frac{\partial A}{\partial x} = (-f'(x)\theta_{2h} + f''(x))m^{-\theta_{1h}} - (-f'(x)\theta_{1h} + f''(x))m^{-\theta_{2h}}
\]
\[
= (-f'(x)\left(\frac{k}{2} + \Delta\right) + f''(x) - f'(x))m^{-\theta_{1h}} - (-f'(x)\left(\frac{k}{2} - \Delta\right) + f''(x) - f'(x))m^{-\theta_{2h}}
\]
\[
< (-f'(x)\left(\frac{k}{2} + \Delta\right) + f''(x) - f'(x))m^{-\theta_{1h}} - (-f'(x)\left(\frac{k}{2} + \Delta\right) + f''(x) - f'(x))m^{-\theta_{2h}}
\]
\[
= (-f'(x)\left(\frac{k}{2} + \Delta\right) + f''(x) - f'(x))(m^{-\theta_{1h}} - m^{-\theta_{2h}})
\]
\[
< 0
\]

The second line comes from Lemma 12. The last line comes from \(f''(x) < f'(x)\) for any given \(x \geq 1\) and \(-\theta_{1h} > -\theta_{2h}\). Thus

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

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 (C.16) is bounded. Thus, \(m_l\) is bounded as \(x \to \infty\). However, RHS of (C.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

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

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

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

which is Assumption 2.

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

**Step 4.** \(x^*\) 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 \( m_z \) from above at the second. It’s easy to verify \( g'(x) < 0 \) and \( \frac{\partial \theta_l \theta_l (m_h^{x_1} - m_l^{x_1})}{\partial m} > 0 \). Thus \( \frac{d^2 m_h}{dx^2} < 0 \). This implies the \( \frac{d^2}{dx^2} \) 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, \( \frac{\partial^2 A}{\partial x^2 m} = (f'(x)(\frac{k}{2} + \Delta) + f''(x) - f'(x))(-\theta_{1h}m^{-\theta_{1h} - 1} + \theta_{2h}m^{-\theta_{2h} - 1}) < 0 \). Secondly, applying Lemma 12

\[
\frac{\partial A}{\partial m} m^k + k A m^{k-1} = \theta_{2l}((H - \alpha - f(x)\theta_{1h} + f'(x))m^{\theta_{2h} - 1}
- \theta_{1l}((H - \alpha - f(x)\theta_{2l} + f'(x))m^{-\theta_{1l} - 1} - km^{\theta_{1l} - 1} - (H - \alpha)(\theta_{2l} - \theta_{1h})
\]

(C.18)

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

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

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

\[
\frac{\partial (-\frac{\partial A}{\partial m} m^k + (\theta_{2l} - \theta_{1h})(H - L - \gamma_l)k x^{k-1})}{\partial x} = \left(-\frac{\partial^2 A}{\partial x \partial m} m^k - k \frac{\partial A}{\partial x} m^{k-1}\right) \frac{\partial m}{\partial x} - \frac{\partial^2 A}{\partial x^2} m^k
\]

\[
+ (\theta_{2l} - \theta_{1h})(H - L - \gamma_l)k(k-1) x^{k-2} > 0
\]

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 \( m_l(x) \).


\[\text{Proof of Theorem 4}\]

\[\text{Proof.}\] It remains to show \( l \) could not benefit by deviating to act as an acquirer. First notice given the two thresholds \( (\beta, \eta) \) and the proposed strategy, the sum of waiting value \( V_l(x_l, \rho_l) \) at \( \rho_l \) can be expressed as

\[
SW(x_{ht}, x_{lt}, \rho_l) = [(H - \alpha)(x_{ht} + x_{lt})]E(e^{-rT(\beta)})|_{\rho_l} P(\rho_T = \beta)
+ [(H - \alpha - \frac{\alpha(H - L)exp(\eta)}{(H + L)exp(\eta)})x_{ht} + (L - \alpha + \frac{\alpha(H - L)exp(\eta)}{(H + L)exp(\eta)})x_{lt}]E(e^{-rT(\eta)})|_{\rho_l} P(\rho_T = \eta)
\]

In the above equation, \( T(\eta) \) and \( T(\beta) \) is the first hitting time of \( \rho_l \) to \( \eta \) and \( \beta \). Given that \( d\rho_l = -\frac{\mu_h - \mu_l}{\sigma} dt + \frac{\mu_h - \mu_l}{\sigma} dB_t \) 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

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

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 \((H - \alpha)(x_{ht} + x_{lt}) > (H - \alpha - \frac{\alpha(H-L)\exp(\eta)}{(H+L\exp(\eta))})x_{ht} + (L - \alpha + \frac{\alpha(H-L)\exp(\eta)}{(H\exp(\eta)+L)})x_{lt}\), it’s easy to verify that \( SW(\rho_{lt}) \) is decreasing in \( \rho_{lt} \).

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

\[
D(\rho'_{lt}) = (L - \alpha \frac{L}{H})(x_{ht} + x_{lt}) > SW(\rho'_{lt})
\]
and $D'(\rho_{it}) > 0$. Therefore $D(\eta) > SW(\eta)$, in other words

$$(L - \alpha \frac{L}{H})(x_{it} + x_{lt}) > SW(\rho_{it}) > (H - \alpha - \frac{\alpha(H - L) \exp(\eta)}{(H + \lambda \exp(\eta))})x_{it} + (L - \alpha + \frac{\alpha(H - L) \exp(\eta)}{(H + \lambda \exp(\eta))})x_{lt}$$

$\iff \frac{\alpha(H - L)}{H} > (H - L - \frac{H - L}{H} - \alpha - \frac{\alpha(H - L) \exp(\eta)}{(H + \lambda \exp(\eta))})x_{it} + \frac{\alpha(H - L) \exp(\eta)}{(H + \lambda \exp(\eta))}x_{lt}$

$\iff \frac{\alpha(H - L)}{H} > (H - L - \frac{H - L}{H} - \alpha - \frac{\alpha(H - L) \exp(\eta)}{(H + \lambda \exp(\eta))})x_{it} + \frac{\alpha(H - L) \exp(\eta)}{(H + \lambda \exp(\eta))}x_{lt} + \frac{\alpha(H - L)}{H}.$

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

\section*{Proof of Proposition 5}

\textit{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 $V^*_h(x_{ht}, \rho_{lt}) = x_{ht}J^*_h(\rho_{lt})$ and $V^*_l(x_{lt}, \rho_{lt}) = x_{lt}J^*_l(\rho_{lt})$ as the value functions of $h$ and $l$ in the first-best solution. The Bellman Equations of $J^*_l(\rho_{lt})$ are the same as $J^*_l(\rho_{lt})$ so we still have $J^*_l(\rho_{lt}) = C^*_l \exp(\theta_1 \rho_{lt}) + C^*_2 \exp(\theta_2 \rho_{lt})$.

If in equilibrium $h$ delays its financing decision, there will be two boundaries (potentially infinite) $\beta^*_1 < 0 < \beta^*_2$ where $h$ is waiting in between. The boundary conditions are:

$$J^*_h(\beta^*_1) = (H - \alpha) + (H - L - \gamma_l) \exp(k \beta^*_1)$$

$$J^*_h(\beta^*_1)' = k(H - L - \gamma_l) \exp(k \beta^*_1)$$

$$J^*_h(\beta^*_2) = (H - \alpha) + (H - L - \gamma_l) \exp(k \beta^*_2)$$

$$J^*_h(\beta^*_2)' = k(H - L - \gamma_l) \exp(k \beta^*_2)$$

Replace $\exp(\beta^*_1)$ with $y_1$ and $\exp(\beta^*_2)$ with $y_2$. We solve $C^*_1$ and $C^*_2$ from the first two equations

$$(\theta_{2h} - \theta_{1h})C^*_1 = \theta_{2h} (H - \alpha) y_1^{-\theta_{1h}} + (\theta_{2h} - k)(H - L - \gamma_l)y_1^{-\theta_{1h} + k}$$

$$(\theta_{1h} - \theta_{2h})C^*_2 = \theta_{1h} (H - \alpha) y_1^{-\theta_{2h}} + (\theta_{1h} - k)(H - L - \gamma_l)y_1^{-\theta_{2h} + k}$$

and from the last two equations
Thus, it's only possible that need to check \( \theta \).

\[ \text{Proof of Proposition 6} \]

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 \( \beta_1^* = \beta_2^* = \beta^* \), which is contradictory to the assumption of two-threshold equilibrium.

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

\[ \Box \]

**Proof of Proposition 6**

We start with the following lemma.

**Lemma 13**  
i) \( \frac{d\theta_{1l}}{d\mu_h} > 0, \frac{d\theta_{1l}}{d\mu_l} > 0, \frac{d\theta_{2h}}{d\mu_h} < 0, \frac{d\theta_{2l}}{d\mu_h} < 0 \).

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

iii) Suppose \( \mu_h = \mu_l + \delta \). Fixing \( \delta \) and increasing \( \mu_h \) and \( \mu_l \) simultaneously will increase \( \theta_{1h} \) and \( \theta_{1l} \) but decrease \( \theta_{2h} \) and \( \theta_{2l} \).

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

\[
\text{sgn}\left(\frac{d(\frac{1}{2} - \sqrt{A})k}{d\mu_h}\right)
= \text{sgn}\left(\frac{1}{2\sqrt{A}} - \frac{k}{2} + \sqrt{A}k - \frac{1}{\sqrt{Ak}}\right)
= \text{sgn}\left(\left(\sqrt{Ak} - 1\right)(1 + \frac{1}{\sqrt{Ak}} - \frac{1}{2\sqrt{A}})\right) > 0
\]

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

56
\[
\text{sgn}(\frac{d(\frac{1}{2} + \sqrt{A})k}{d\mu_h})
\]
\[
= \text{sgn}(\frac{1}{\sqrt{A}k} - \frac{1}{2\sqrt{A}} - \frac{k}{2} - \sqrt{A}k) < 0
\]

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

ii)
\[
\text{sgn}(\frac{d(\frac{1}{2} - \sqrt{A})k}{d\mu_i})
\]
\[
= \text{sgn}(\frac{1}{2\sqrt{A}} + \frac{k}{2} - \sqrt{A}k + \frac{1}{\sqrt{A}k})
\]
\[
= \text{sgn}((\sqrt{A}k + 1)(1 + \frac{1}{2}k - \sqrt{A}k)) < 0
\]

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

iii) Since \(\delta\) is fixed, thus \(k\) is unchanged but \(A\) is smaller. Thus \(\theta_{1h}\) and \(\theta_{1l}\) are larger but decrease \(\theta_{2h}\) and \(\theta_{2l}\) are smaller.

\textbf{Proof.} Rearrange equations (C.15) and (C.16) to be
\[
\left\{ [(H - \alpha - f(x))\frac{\theta_{2h}}{\theta_{2h} - \theta_{1h}} + \frac{f'(x)}{\theta_{2h} - \theta_{1h}}]m^{-\theta_{1h}} - [(H - \alpha - f(x))\frac{\theta_{1h}}{\theta_{2h} - \theta_{1h}} + \frac{f'(x)}{\theta_{2h} - \theta_{1h}}]m^{-\theta_{2h}} - (H - \alpha) \right\} m^k
\]
\[
= (H - L - \gamma_l)x^k \quad (C.19)
\]
\[
(L - \alpha + \gamma_l)(\frac{\theta_{2l}}{\theta_{2l} - \theta_{1l}}m^{\theta_{1l}} - \frac{\theta_{1l}}{\theta_{2l} - \theta_{1l}}m^{\theta_{2l}}) = L - \alpha + g(x) \quad (C.20)
\]
Take the derivative of $M_1$ w.r.t $\mu_h$

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

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

$$\frac{-\theta_{1h} m^{-\theta_{1h}} - \theta_{2h} m^{-\theta_{2h}}}{\theta_{2h} - \theta_{1h}} - m^{-\theta_{1h}} + \theta_{1h} m^{-\theta_{1h}} \log(m) < 0$$

The second line is due to $\theta_{1h} < 0$ and $\log(m) > 0$. By Lemma 13, $\frac{d\theta_{1h}}{d\mu_h} > 0$ 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 $\frac{d m^k}{d \mu_h} < 0$. For any given value of $m$, the LHS of equation (C.19) is strictly decreasing in $\mu_h$. Therefore $\frac{d m^k(x; \mu_h)}{d \mu_h} > 0$. In other words, $m_h(x)$ shifts towards northwest as $\mu_h$ increases. Applying the same kind of trick, we can show

$$\frac{\partial M_2}{\partial \mu_h} = \frac{\theta_{2l}}{\theta_{2l} - \theta_{1l}} \frac{d\theta_{2l}}{d\mu_h} (\frac{m^{\theta_{1l}} - m^{\theta_{2l}}}{\theta_{2l} - \theta_{1l}} + \log(m) m^{\theta_{1l}}) < 0$$

$$+ \frac{\theta_{1l}}{\theta_{2l} - \theta_{1l}} \frac{d\theta_{2l}}{d\mu_h} (\frac{m^{\theta_{2l}} - m^{\theta_{1l}}}{\theta_{2l} - \theta_{1l}} - \log(m) m^{\theta_{2l}}) < 0$$

Thus, both $m_h(x)$ and $m_l(x)$ shift towards northwest as $\mu_h$ increases. Denote $m^*(\mu_h)$ and $x^*(\mu_h)$ as solution of the equations. By comparing $\frac{\partial M_1}{\partial \mu_h}$ over $\frac{\partial M_2}{\partial \mu_h}$ and $\frac{\partial M_2}{\partial m}$ 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 $\frac{d x^*(\mu_h)}{d \mu_h} < 0$. Notice $y^* = \frac{x^*}{m^*} = \cot(\theta_x)$, where $\theta_x$ is the angle between the line $(x^*, m^*)$, $x^* = (0, 0)$ and $(x^*, 0) \rightarrow (0, 0)$. As $(x^*, m^*)$ shifts northwest, $\theta_x$ increases and $\cot(\theta_x)$ decreases. Hence $\frac{d y^*(\mu_h)}{d \mu_h} < 0$.

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

58
 Proof of Proposition 7

Proof. The proof is similar to Proposition 6. Denote $\theta'_{ij}$ as the changes of $\theta$ when $\mu_h$ and $\mu_l$ are simultaneously increased without changing the wedge. By Lemma 13, $\theta'_{1h} > 0$, $\theta'_{1l} > 0$, $\theta'_{2h} < 0$, 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 $\mu_h$ and $\mu_l$ increase. The rest follows argument in Proposition 6. □

Proof of Proposition 8

Proof. Using the same notation as Lemma 13, I start with showing $\frac{d\theta_{1h}}{d\sigma^2} > 0$, $\frac{d\theta_{1l}}{d\sigma^2} > 0$, $\frac{d\theta_{2h}}{d\sigma^2} < 0$, $\frac{d\theta_{2l}}{d\sigma^2} < 0$.

$$\text{sgn} \left( \frac{d(\frac{1}{2} - \sqrt{A})k}{d\sigma^2} \right)$$

$$= \text{sgn} \left( A - \frac{\sqrt{A}}{2} - \frac{12r - \mu_h - \mu_l}{2\sigma^2} - \frac{1}{k^2} \right)$$

$$= \text{sgn} \left( \frac{1}{4} + \frac{12r - \mu_h - \mu_l}{2\sigma^2} - \sqrt{A} - \frac{1}{2} \right)$$

The last line uses the fact that $A = \frac{1}{4} + \frac{2r - \mu_h - \mu_l}{\sigma^2} + \frac{1}{k^2}$. Notice function $\frac{1}{2}(x - \sqrt{\frac{1}{4} + \frac{1}{k^2} + x})$ is increasing in $x$. As $r > \mu_h > \mu_l$, $\frac{2r - \mu_h - \mu_l}{\sigma^2} > \frac{\mu_h - \mu_l}{\sigma^2} = \frac{1}{k}$. Since $\frac{1}{4} + \frac{11}{2k} - \frac{\sqrt{\frac{1}{4} + \frac{1}{k^2} + \frac{1}{2}}}{2} = 0$, $\frac{1}{4} + \frac{1}{2} \frac{2r - \mu_h - \mu_l}{\sigma^2} - \frac{\sqrt{A}}{2} > 0$. 59
\[ sgn\left(\frac{d\left(\frac{1}{2} + \sqrt{A}\right)}{d\sigma^{-2}}\right) \]

\[ = sgn\left(-A - \frac{\sqrt{A}}{2} + \frac{12r - \mu_h - \mu_l}{\sigma^2} + \frac{1}{k^2}\right) \]

\[ = sgn\left(-\frac{1}{4} - \frac{12r - \mu_h - \mu_l}{\sigma^2} - \frac{\sqrt{A}}{2}\right) < 0 \]

Then it’s easy to check all signs of \(\frac{d\theta_{ij}}{d\sigma^2}\) is just the opposite of \(\frac{d\theta_{ij}}{d\sigma^2 - 2}\). Thus \(\frac{d\theta_{1j}}{d\sigma^2} < 0, \frac{d\theta_{ij}}{d\sigma^2} < 0, \frac{d\theta_{ij}}{d\sigma^2} > 0\). Using the same trick as Proposition 6, one could show \(\frac{\partial M_1}{\partial \sigma^2} > 0\) and \(\frac{\partial M_2}{\partial \sigma^2} > 0\). The effect of increasing \(\sigma^2\) is similar to increasing \(\mu_l\) as the changes w.r.t \(\theta_{ij}\)s are the same. Both \(m_h(x)\) and \(m_l(x)\) shift towards southeast as \(\sigma^2\) increases. Then \(\frac{dm^*}{d\sigma^2} < 0\) and \(\frac{dx^*}{d\sigma^2} > 0\). As \(y^* = \frac{x^*}{m^*} = \cot(\theta_x)\) and \(\theta_x\) decreases, \(\frac{dy^*}{d\sigma^2} > 0\).

\[\square\]

**Proof of Theorem 9**

**Proof.** The statement is true if the underpricing cost in pooling IPO is larger than \(h\)'s net benefit in acquisition \(\forall \rho \leq 0\). Combine equations (??) and (7). It is equivalent to:

\[ -\frac{\alpha(H - L)\exp(\rho)}{H + L\exp(\rho)} > -(1 - \xi)\alpha \frac{H - L}{L} + \xi(H - L) \frac{x_{lt}}{x_{ht}} \]

\[\iff\]

\[ -\frac{\alpha(H - L)\exp(\rho)}{H + L\exp(\rho)} > -(1 - \xi)\alpha \frac{H - L}{L} + \xi(H - L)\exp(k\rho) \]

\[\iff\]

\[ (1 - \xi)\alpha \frac{H - L}{L} > \frac{\alpha(H - L)\exp(\rho)}{H + L\exp(\rho)} + \xi(H - L)\exp(k\rho) \]

The RHS of last line is increasing in \(\rho\). 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 \(\xi \leq \xi = \frac{\frac{\alpha - \frac{\alpha L}{H + L}}{1 + \frac{\alpha}{L}}}{1} < 1\). The remaining statement comes from solving a standard optimal problem for \(h\).

\[\square\]
Proof of Theorem 10

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

\[ C_{1h}\exp(\theta_{1h}\eta) + C_{2h}\exp(\theta_{2h}\eta) = H - \alpha - \frac{\alpha(H - L)\exp(\eta)}{H + L\exp(\eta)} \]  \hspace{1cm} (C.21)

\[ C_{1l}\exp(\theta_{1l}\eta) + C_{2l}\exp(\theta_{2l}\eta) = L - \alpha + \frac{\alpha(H - L)\exp(\eta)}{H\exp(\eta) + L} \]  \hspace{1cm} (C.22)

\[ C_{1h}\exp(\theta_{1h}\beta) + C_{2h}\exp(\theta_{2h}\beta) = H - \alpha - (1 - \xi)\alpha\frac{H - L}{L} + \xi(H - L)\exp(k\beta) \]  \hspace{1cm} (C.23)

\[ C_{1l}\exp(\theta_{1l}\beta) + C_{2l}\exp(\theta_{2l}\beta) = L - \alpha + (1 - \xi)(H - L) + (1 - \xi)\alpha\frac{H - L}{L}\exp(-k\beta) \]  \hspace{1cm} (C.24)

\[ C_{1h}\theta_{1h}\exp(\theta_{1h}\eta) + C_{2h}\theta_{2h}\exp(\theta_{2h}\eta) = -\frac{\alpha H(H - L)\exp(\eta)}{(H + L\exp(\eta))^2} \]  \hspace{1cm} (C.25)

\[ C_{1l}\theta_{1l}\exp(\theta_{1l}\beta) + C_{2l}\theta_{2l}\exp(\theta_{2l}\beta) = -k(1 - \xi)\alpha\frac{H - L}{L}\exp(-k\beta) \]  \hspace{1cm} (C.26)

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 12 (iii) to change \( \theta_{1h} \) and \( \theta_{2h} \) into \( \theta_{1l} \) and \( \theta_{2l} \) into (C.21) to (C.26):

\[ C_{1h}x^{\theta_{1l}} + C_{2h}x^{\theta_{2l}} = (H - \alpha - f(x))x^{-k} \]  \hspace{1cm} (C.27)

\[ C_{1l}x^{\theta_{1l}} + C_{2l}x^{\theta_{2l}} = L - \alpha + g(x) \]  \hspace{1cm} (C.28)

\[ C_{1h}y^{\theta_{1l}} + C_{2h}y^{\theta_{2l}} = (H - \alpha - (1 - \xi)\alpha\frac{H - L}{L} + \xi(H - L)y^k)y^{-k} \]  \hspace{1cm} (C.29)

\[ C_{1l}y^{\theta_{1l}} + C_{2l}y^{\theta_{2l}} = L - \alpha + (1 - \xi)(H - L) + (1 - \xi)\alpha\frac{H - L}{L}y^{-k} \]  \hspace{1cm} (C.30)

\[ C_{1h}\theta_{1h}x^{\theta_{1l}} + C_{2h}\theta_{2h}x^{\theta_{2l}} = -f'(x)x^{-k} \]  \hspace{1cm} (C.31)

\[ C_{1l}\theta_{1l}y^{\theta_{1l}} + C_{2l}\theta_{2l}y^{\theta_{2l}} = -k(1 - \xi)\alpha\frac{H - L}{L}y^{-k} \]  \hspace{1cm} (C.32)

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 (C.30) and (C.32) to solve \( C_{1l} \) and \( C_{2l} \) and (C.27) and (C.31) to solve \( C_{1h} \) and \( C_{2h} \). Then replace the solved constants in (C.28) and (C.29) respectively.

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

\[(H - \alpha - f(x))\theta_{2h} + f'(x)m^{-\theta_{1l}} - [(H - \alpha - f(x))\theta_{1h} + f'(x)m^{-\theta_{2l}}] \]

\[-(\theta_{2h} - \theta_{1h})(H - \alpha - (1 - \xi)\alpha\frac{H - L}{L})m^k = (\theta_{2l} - \theta_{1l})\xi(H - L)x^k \]  \hspace{1cm} (C.33)
\[(L - \alpha + (1 - \xi)(H - L))(\theta_{2l}m^{\theta_{1l}} - \theta_{1l}m^{\theta_{2l}})\]
\[+(1 - \xi)\alpha \frac{H - L}{L} x^{-k}(\theta_{2h}m^{\theta_{1h}} - \theta_{1h}m^{\theta_{2h}}) = (\theta_{2l} - \theta_{1l})(L - \alpha + g(x)) \quad (C.34)\]

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 \(\xi > \frac{1 + \frac{2}{\alpha} \frac{\sqrt{\gamma}}{H-L}}{1 + \frac{2}{\alpha} \frac{\sqrt{\gamma}}{H-L}}\), then \(\forall x \geq 1\) there exists a unique \(m_h(x) \geq 1\) and \(m_l(x) \geq 1\) that solves the equations (C.33) and (C.34) correspondingly.

The statement on \(m_h(x)\) can be shown in the same way as before. For \(m_l(x)\), the LHS of (C.34) 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 \(\xi > \frac{1 + \frac{2}{\alpha} \frac{\sqrt{\gamma}}{H-L}}{1 + \frac{2}{\alpha} \frac{\sqrt{\gamma}}{H-L}}\). \(L - \alpha + g(1)\) is the minimum of RHS. This proves the statement.

**Step 2.** \(m'_h(x) > 0\) and \(m'_l(x) > 0\).
\[
\frac{dm_h}{dx} = A(x) > 0
\]
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 \(B(m) = (L - \alpha + \gamma)\theta_{1l}\theta_{2l}(m^{\theta_{1l} - 1} - m^{\theta_{2l} - 1}) + (1 - \xi)\alpha \frac{H - L}{L} x^{-k}\theta_{1h}\theta_{2h}(m^{\theta_{1h} - 1} - m^{\theta_{2h} - 1}) > 0\).

**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 (C.34), satisfying
\[(L - \alpha + (1 - \xi)(H - L))(\theta_{2l}m_l(1)^{\theta_{1l}} - \theta_{1l}m_l(1)^{\theta_{2l}})\]
\[+(1 - \xi)\alpha \frac{H - L}{L} (\theta_{2h}m_l(1)^{\theta_{1h}} - \theta_{1h}m_l(1)^{\theta_{2h}}) = (\theta_{2l} - \theta_{1l})(L - \alpha + g(1)) \quad (C.35)\]

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
\[
\theta_{2l}m_l(1)^{\theta_{1l}} - \theta_{1l}m_l(1)^{\theta_{2l}} > \frac{(\theta_{2l} - \theta_{1l})(L - \alpha + g(1))}{[L - \alpha + (1 - \xi)(H - L)](1 - \xi)\alpha \frac{H - L}{L}} \quad (C.36)
\]
Now consider the following equation:

\[
\begin{align*}
\{(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))*m_l(1)^{\theta_{1h}} - (H - \alpha - f(1))*m_l(1)^{\theta_{2h}} & - (\theta_{2h} - \theta_{1h})(H - \alpha - (1 - \xi)\alpha \frac{H - L}{L}) \\

=(H - \alpha - f(1))((\theta_{2h} - \theta_{1h})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{align*}
\]

Thus, if we can show \((H - \alpha - f(1))((\theta_{2h} - \theta_{1h})m_l(1)^{\theta_{1h}} - \theta_{1h}m_l(1)^{\theta_{2h}}) > (\theta_{2h} - \theta_{1h})(H - \alpha - (1 - \xi)\alpha \frac{H - L}{L} + \xi(H - L))\), by monotonicity we prove \(m_h(1) < m_l(1)\). By equation (C.36), a sufficient condition is

\[
\frac{L - \alpha + g(1)}{[L - \alpha + (1 - \xi)(H - L)] + (1 - \xi)\alpha \frac{H - L}{L}} > \frac{\theta_{2h}m_l(1)^{\theta_{1h}} - \theta_{1h}m_l(1)^{\theta_{2h}}}{\theta_{2l}m_l(1)^{\theta_{1l}} - \theta_{1l}m_l(1)^{\theta_{2l}}}
\]

(C.37)

\[
\frac{H - \alpha - (1 - \xi)\alpha \frac{H - L}{L} + \xi(H - L)}{H - \alpha - f(1)}
\]

(C.38)

When \(\xi = 1\), this equation reduces to

\[
\frac{L - \alpha + g(1)}{L - \alpha} \geq \frac{H - \alpha + H - L}{H - \alpha - f(1)}.
\]

(C.39)

By continuity, if this is true, there exists a threshold \(\xi^{*}\) such that (C.38) is true whenever \(\xi \geq \xi^{*}\). Denote \(\bar{\xi} = \max\{\xi^{*}, \frac{1}{1 + \frac{\pi + \pi}{1 + \frac{\pi}{2}}}\}\). The theorem is proved.

Proof of Proposition 11

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