January 2010

Leveraged Buyouts, Long Term Relationship and Financial Contracting

Lei Gao
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

John M. Olin School of Business

Dissertation Examination Committee:
Ohad Kadan, Co-Chair
David K. Levine, Co-Chair
Tat Chan
Radhakrishnan Gopalan
Lubomir Litov
Juan Pantano

LEVERAGED BUYOUTS, LONG TERM RELATIONSHIP AND FINANCIAL CONTRACTING

by

Lei Gao

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2010

Saint Louis, Missouri

ABSTRACT OF THE DISSERTATION

Leveraged Buyouts, Long Term Relationship and

Financial Contracting

by

Lei Gao

Doctor of Philosophy in Business Administration

Washington University in St. Louis, 2010

Professor Ohad Kadan, and

Professor David K. Levine, Chairperson(s)

Abstract: In the first chapter of my dissertation, I study private equity sponsored leveraged buyouts. In this study, I identify and estimate two components of the created value: benefits of debt and influence of private equity. In the second chapter of my dissertation, I study repeated games with private monitoring in continuous time. I apply stochastic filtering theory to characterize the evolution of players' beliefs, which is difficult to do in discrete time. In the third chapter of my dissertation, I study the connection between financial contracting and staged financing in venture capital.

Acknowledgements

I would like to express the deepest appreciation to my committee chairs, Professor Ohad Kadan and Professor David K. Levine. I have been amazingly fortunate to have Professor Kadan as my advisor who gave me the freedom to explore on my own, and at the same time the guidance to recover when my steps faltered. My committee co-chair, Professor Levine, has been always there to listen and give advice. I have benefited greatly from his deep insights into economics.

I am grateful to my committee members, Professor Radhakrishnan Gopalan, Professor Lubomir Litov, Professor Tat Chan, and Professor Juan Pantano. They introduced me to fascinating research topics in corporate finance and structural econometrics. This dissertation would not be possible without their persistent help and support.

I am also thankful to the system staff who maintained all the machines so efficiently that my computer programs can run smoothly. Scott Ladewig, Wendell Fry, and the rest of the team have done an excellent job.

Many friends have helped me stay sane through these difficult years. Their support and care helped me overcome setbacks and stay focused on my graduate study. I greatly value their friendship and I deeply appreciate their belief in me.

Most importantly, none of this would have been possible without the love and patience of my parents. I would like to express my heart-felt gratitude to my family. My dissertation is dedicated to my dear grandfather, who lost his long battle with cancer.

Contents

List of Tables

List of Figures

Chapter 1

Value Created by Private Equity: Evidence from Two-Sided Matching in Leveraged Buyouts

1.1 Introduction

Private equity firms, more specifically buyout funds, acquire companies or divisions of companies using large amounts of debt – the so called leveraged buyouts (LBOs). Many factors contribute to value creation in LBOs. According to previous research, debt (Kaplan [17], Ivashina and Kovner [14]) and monitoring (Smith [31], Kaplan [18], DeGeorge and Zeckhauser [6], Holthausen and Larcker [13]) are two possible sources of value. Debt creates value directly through tax benefits, and indirectly through mitigated agency problems (DeAngelo, DeAngelo, and Rice [5]). Monitoring includes all direct influence of private equity firms on target companies, such as suggesting new strategies and recruiting seasoned executives. Despite the importance of private equity as an asset class in the economy, little is known about the economic magnitude of the value created through debt and the value created through monitoring. This is because these two types of value are entangled together empirically.[1]

This paper develops an econometric model of the LBO market to identify and estimate the value created through debt and the value created through monitoring. The paper further analyzes how the values vary with the characteristics of private equity firms and target companies. It also provides main motivations for the practice of club deals[2] studied by Officer, Ozbas, and Sensoy [27].

There are two challenges in the estimation. First, the value that can be created by the specific match between private equity firms and target companies is known only to the bidding firms. This information cannot be systematically observed in the data. Moreover, different bidders may have different valuations of the same target company. In the framework of classical linear regressions, the dependent variable is *missing*. Second, the mutual selection by private equity firms and target companies causes an endogeneity problem. Private equity firms search for investment opportunities and compete against each other to acquire the most attractive target companies;

[1]Debt interest payments extract undistributed cash out of the target company and keep the management from investing in negative NPV projects for private benefits (Lehn and Poulsen [22]). In the meantime, private equity investors can achieve the same goal by closely monitoring the company's investment decisions.

[2]A club deal is the deal in which the acquirer is a private equity consortium.

meanwhile boards of directors of the target companies select the winning bids for the best interest of shareholders. The deal outcomes – which private equity firm acquires which target company – are endogenously decided by the choices of both sides.

To solve these two problems, this paper introduces a structural model based on a two-sided matching model, which describes the mutual selection by private equity firms and target companies.[3] In the matching model, a match between private equity firms and target companies is a potential LBO deal that creates value, and the value is measured by the expected total return on the target company's equity (abbreviated as *the total return*). The total return is the private equity firm's valuation of the target company's equity divided by its pre-buyout price. Since the total returns cannot be observed, they are calculated by a latent equity return function, a function of the characteristics of private equity firms and target companies.

One of the goals of this study is to estimate the parameters in the latent equity return function. The value created through debt and the value created through monitoring are calculated by these parameters. The estimation is based on the stability of match outcomes in the LBO market. In a realized LBO deal, the observed deal premium is the return on equity for the existing shareholders of the target company; the excess of the total return over the

[3]Strictly speaking, existing shareholders of the target companies, represented by boards of directors, choose selection mechanisms, e.g. ascending-bid or sealed-bid auctions (Cramton and Schwartz [4]), and take the highest bid under Delaware law (the predominant corporate law in the US). The advantage of estimating a matching game is that economists do not have to model matching procedures in matching games (Myerson [26]); therefore, the estimation does not require any information related to the auctions in terms of format. But the matching model in this paper implicitly includes this winning rule: target companies form preference on private equity firms based on the highest bids the latter can offer – the potential deal premia.

deal premium is the return for the private equity firm on each dollar value the firm pays the shareholders. Since the latent equity return function calculates the total return for each feasible match between private equity firms and target companies, the function together with the deal premium data tell us how private equity firms and target companies share the total returns. For the match outcomes to be stable, neither the private equity firms nor the target companies would have incentive to deviate from their current match, and this incentive compatibility is supported by an appropriate sharing mechanism of the total returns. The estimation of the structural model boils down to searching for the parameter values of the function that lead to this sharing mechanism.

I estimate the model using data on LBOs of public companies sponsored by private equity during the period from 1986 to 2007. The estimation first isolates the two components of debt and monitoring from macroeconomics factors and industry specific factors, which unfortunately cannot be identified by the model. Then using the estimated latent equity return function, I calculate the two types of values in the total return. Out of the expected total returns on the target companies' equity before deal announcements, the value created through debt is 3.1%, which is statistically not different from zero; and the value created through monitoring is 7.8%, which is statistically positive and significant.

I further investigate how the values vary with the characteristics of private equity firms or target companies, and over time. In LBOs sponsored by older private equity firms, debt destroys value. For private equity firms, more debt leads to higher private equity investment returns, but too much debt destroys value and reduces the total return of the potential match, which in turn reduces the probability of winning the bidding competitions. Older private equity firms can pursue the excess borrowing strategy because they create more value through monitoring, which compensates the loss of value on debt.

In general, the value created by the match is higher if the private equity firms invest in poorly performing companies or small companies. Private equity investments in the late 1990s are less profitable than those in other periods. This may be because the private equity industry in the late 1990s became better organized and more efficient than in the 1980s. The time effect is no longer significant after the year 2001.

This paper also analyzes the motivation for club deals using the structural model. In the existing literature, there are many explanations for the phenomenon of private equity consortia, such as: to obtain better debt financing, to share risks, and to alleviate bidding competitions. In order to find out the main reason for this practice, the paper constructs a measure called the Coalition Contribution Index (CCI). The index of a private equity firm measures its popularity among the investing firms that decide to submit joint bids. A firm that better fits the motivation for joint bidding has a higher index regardless of what the motivation is. So, if the dominant motivation is related to debt financing, we will find connections between the debt and the index. Otherwise, there is no such relation.

The estimation identifies such a relation, suggesting that private equity firms form consortia to bring better debt financing to club deals. In addition to this finding, the estimated latent equity return function allows us to examine the difference in sharing of the total returns between the private equity firms and the target companies in club deals and in single firm deals. I find no strong evidence supporting the view that private equity firms have an advantage in sharing total returns against target company existing shareholders in club deals.

The contribution of this paper is both in the empirical findings and in the structural approach. First, this is the first paper to directly estimate the value created by private equity at deal level. The paper estimates that the value created through debt is statistically zero, which rules out the possibility that private equity *fund* returns are the result of wealth reallocation from bondholders to private equity investors. Second, the structural approach in this paper can be generalized to broader research topics in empirical corporate finance, such as recruitment of corporate managers and corporate financial decisions in a competitive environment.

1.2 Related Literature

Guo, Hotchkiss, and Song [11], Acharya, Hahn, and Kehoe [1] study value creation in LBOs. One of their key approaches is to form peer industry groups and use their financial data as benchmarks. Intuitively speaking, if private equity investors are actively involved in management of post-buyout companies, their abilities and experience may affect the magnitude of the value that could be created. In contribution to existing literature, the matching model in my paper considers this heterogeneity of investors.

Sørensen [32] also applies structural modeling and uses a two-sided matching model. In studying venture capital industry, Sørensen [32] finds that companies backed by more experienced venture capitalists (VCs) are more likely to go public. This is due to sorting effect and direct influence by VCs. Sørensen [32] identifies and estimates these two effects. My paper studies a different research question, and focuses on total values created by specific matches in LBO markets and their cross-sectional variation.

There is a growing literature studying private equity returns at the fund level, such as, Jones and Rhodes-Kropf [16], Ljungqvist and Richardson [23], Kaplan and Schoar [19], Phalippou and Gottschalg [28], Jegadeesh, Kraeussl, and Pollet [15]. Many factors contribute to the value created in LBOs that leads to private equity fund returns. These sources of value includes macroeconomic factors, debt, and direct influence of private equity investors on target companies, i.e., monitoring.

Gompers and Lerner [9] and [10] find that macroeconomic factors affect the capital flows into private equity, and they also affect the valuation of individual deals. Kaplan [17] studies the tax benefits of debt in LBOs. In the study of company post-buyout performance, Smith [31], Kaplan [18], DeGeorge and Zeckhauser [6], Holthausen and Larcker [13], Guo, Hotchkiss, and Song [11] find evidence that the monitoring by private equity firms improves the operating performance of companies.

Ljungqvist, Richardson, and Wolfenzon [24] have analyzed the investment behaviors of buyout funds in LBO markets. They examine what factors affect a private equity fund's investment decisions. They find different investment behaviors by established funds and younger funds. Related to their research, I study how the value created in LBOs vary with the characteristics of the private equity firms.

1.3 The Structural Model

Papers in other areas of research have successfully applied the structural modeling approach (Berry, Levinsohn, and Pakes [2]). In the study of LBO markets, traditional linear regression models cannot solve the endogeneity problem mentioned in section 2.1. One solution to the endogeneity problem is instrumental variable (IV) method. A widely used instrument is the distance between investors and companies. But it may not be a valid instrument in the study of LBOs. For instance, a target company may register in a U.S. State with lower corporate tax rates, have headquarters in the Midwest, and have multiple manufacturing facilities outside North America. The distance can hardly be a good measure of the cost of monitoring the company.

The matching in the LBO market is many-to-many. A private equity firm is a collection of funds, and each fund invests in a portfolio of companies. The seller of a company may choose one private equity firm as the buyer, or a consortium of private equity firms. Due to this many-to-many feature in the selection process, the choice models, such as multinomial Probit model, are difficult to implement. First, the maximum number of companies that a private equity firm can invest in is unobservable. Second, anecdotal evidence suggests that the formation of teams in submitting joint bids follows a pattern similar to social networks, and the number of possible team formations grows exponentially with the total number of firms. The structural model in this paper can easily solve these problems.

1.4 The Model

1.4.1 Model Setup

Consider a market for leveraged buyouts, in which there are a finite set of public companies U (the "targets"), indexed by $i \in U$, and a finite set of private equity firms D (the "acquirers"), indexed by $a \in D$. The index i also identifies the leveraged buyout deal of target company i. In this market, private equity firms are searching for profitable deals and they can acquire more than one companies. In the mean time, companies will be sold to a private equity firm or a consortium of firms that offers the highest bid. This creates a market of many-to-many matching.[4]

Several exogenous objects characterize this matching market. The space of matches is $U \times D$. Let $\mu = \langle a, i \rangle$ be a match between private equity firm a and company i. For a match to be feasible the private equity firm's inception year must be before the transaction year. And for each feasible match there is a total value calculated by the latent equity return function $r_{\langle a,i \rangle}$. An assignment A is a finite collection of matches, which is a subset of all matches in $U \times D$. Given each assignment $A = \{\mu_1, \mu_2, \ldots\}$, let $T = \{t_{\mu_1}, t_{\mu_2}, \ldots\}$ be the set of transfers for all matches in A. Each $t_\mu \in \mathbb{R}^+$ represents a deal premium that a private equity firm pays to a target company, and t_μ specifies how the firm and the target company share the total value r_μ. For example, in KKR's buyout of Nabisco, Nabisco pre-announcement stock price is p_0; KKR believes its equity worth p_2 after due diligence; and the deal is closed

[4] I will follow most of the notations in Fox [7].

at price p_1. So $t_\mu = p_1/p_0$ and $r_\mu = p_2/p_0$.[5]

I give an example to illustrate this setting. Three private equity firms appear in the history of Ohio Mattress Company (under the name Sealy Corporation after 1990). They are Gibbons Green Van Amerongen, Bain Capital, and Kohlberg Kravis Robert & Company in 1989, 1997, and 2004 buyout deals, respectively. In any given year, say 1992, all three matches are feasible because all three firms were founded well before 1992. Each match creates a total value measured by the total return $r_{\langle a,i\rangle}$ which is a function of the characteristics of the firm and the company and is unobservable to econometricians. Feasible matches have the potential to be developed into real transactions. In the completed transaction between Gibbons Green and Ohio Mattress Company, $r_{\langle a,i\rangle}$ is the expected total return that Gibbons Green can create on each dollar value that it pays for the equity of Ohio Mattress Company, and $t_{\langle a,i\rangle}$ is the actual deal premium paid by Gibbons Green in the transaction,[6] which is astounding 94.29 percent. For a deal to be attractive, the individual rationality condition is $1 \leq t_{\langle a,i\rangle} \leq r_{\langle a,i\rangle}$.[7]

Let $Q : U \cup D \rightarrow \mathbb{N}^+$ be the set of quotas, where $q_a^d \in Q$ is the quota of a private equity firm a and $q_i^u \in Q$ is the quota of a target company i. The quota q_a^d represents the maximum number of companies a private equity firm can acquire, and the quota q_i^u represents the maximum number of private equity firms possible in forming a consortium. Let X be the collection of all outcome relevant exogenous characteristics for the transaction, the acquiring

[5] I use log returns in later empirical estimation.

[6] The return r is a relative measure adjusted by a relevant discount factor. The model cannot explain what this discount factor is.

[7] The markup $r_{\langle a,i\rangle}/t_{\langle a,i\rangle}$ depends on allocations of bargaining power between private equity firms and target companies, or intensity of bidding competitions.

firm, and the target company. The matching market can be represented by a combination of (D, U, Q, X, A, T).

Let M_i be the maximum debt (including senior and subordinate debt) to EBITDA multiple of the transaction in which the target is the company i. Let Y_i be the transaction period dummy, I_i be the target company industry dummy, R_i be the target company's EBITDA to total assets ratio, V_i be the target company's market value of equity before the deal. Also define E_a as the private equity firm's age at the time of the transaction, and S_a as the firm's cumulative committed capital under management. The payoff relevant exogenous elements are a collection of vectors $X = \{x_i^u = (M_i, Y_i, I_i, R_i, V_i)', i \in U; x_a^d = (E_a, S_a)', a \in D\}$.

For any given pair $\langle a, i \rangle$ in a match outcome (A, T), it is the combination of the private equity firm's productivity and the target company's assets (tangible, intangible, human capital, etc.) that contributes to the total returns. So the latent equity return function is affected by both the target companies and the private equity firms' characteristics. The model in this paper can identify the linear effects from target companies and the cross interaction effects by target companies and private equity firms. This is sufficient for the study because private equity firms as stand alone entity cannot create any value. The latent equity return function for a one-to-one matched transaction is defined as the following.

Definition 1. *The latent equity return function of a transaction* $\langle a, i \rangle$ *for*

$i \in U$ *and* $a \in D$ *is*

$$\begin{aligned} r_{\langle a,i\rangle} = \; & \exp\Big[\big(\alpha_1 M_i + \alpha_2 M_i^2 + \alpha_3 I_i + \alpha_4 Y_i + \alpha_5 R_i + \alpha_6 V_i\big) \\ & + \big(\beta_0^1 M_i + \beta_1^1 R_i + \beta_2^1 V_i\big) \cdot E_a + \big(\beta_0^2 M_i + \beta_1^2 R_i + \beta_2^2 V_i\big) \cdot S_a + \varepsilon_{\langle a,i\rangle}\Big]. \end{aligned} \tag{1.1}$$

The error term $\varepsilon_{\langle a,i\rangle}$ contains two main components: measurement errors and deviation of the latent equity return function from its "correct" form. The age E_a and the size S_a are proxies for the private equity firm's experience and ability, respectively. Measurement errors occur when partners in young private equity firms have previous buyout experience. In addition to these measurement errors, there is a residual, which is a collection of higher order terms of the exogenous variables in X. Fortunately, the residual does not affect the estimation for marginal effects and cross-interaction effects of those variables. I do not specify probability distributions for $\varepsilon_{\langle a,i\rangle}$, since the estimation is semi-parametric.

The LBOs studied in this paper are in general many-to-many matching. Recently, it is more common for firms to form consortia and submit joint bids. A target company may be acquired by more than one private equity firms; a private equity firm can acquire many companies. This is a typical many-to-many matching market, which is different from marriage matching, where one-to-one restriction is legally binding. I generalize the latent equity return function from a one-to-one matching market to a many-to-many matching market. Let C^d denote a coalition of private equity firms (a coalition may have only one firm). When target company i is acquired by private equity firms C^d, it generates a collection of total returns $r_{\langle a,i\rangle}$, where $a \in C^d$. Each $r_{\langle a,i\rangle}$ is calculated by the latent equity return function given as the following.

Definition 2. *The latent equity return function of a transaction $r_{\langle a,i\rangle}$, in which target company i is acquired by private equity firm a, where firm a is in a coalition C^d, is*

$$\begin{aligned} r_{\langle a,i\rangle} = \;& \exp\Bigg[\left(\alpha_1 M_i + \alpha_2 M_i^2 + \alpha_3 I_i + \alpha_4 Y_i + \alpha_5 R_i + \alpha_6 V_i\right) \\ & + \left(\beta_0^1 M_i + \beta_1^1 R_i + \beta_2^1 V_i\right) \cdot E_a + \left(\beta_0^2 M_i + \beta_1^2 R_i + \beta_2^2 V_i\right) \cdot S_a \\ & + \lambda_{-a} \left(\left(\beta_0^1 M_i + \beta_1^1 R_i + \beta_2^1 V_i\right) \cdot \sum_{a' \in C^d, a' \neq a} E_{a'} + \left(\beta_0^2 M_i + \beta_1^2 R_i + \beta_2^2 V_i\right) \cdot \sum_{a' \in C^d, a' \neq a} S_{a'} \right) \\ & + \varepsilon_{\langle a,i\rangle} \Bigg]. \end{aligned} \tag{1.2}$$

The coefficient λ_{-a} cannot be identified, so β^1s and β^2s remain the same for firm a's bidding partners $a' \in C^d$.

The summation of age E_a and capital under management S_a of private equity firms in one coalition measures their joint experience and ability. An alternative measure is the maximum values in each category among the members of a coalition. But this alternative can be easily rejected by the data, since otherwise, all private equity firms should team together with the most experienced and the largest firms, such as Bain Capital or KKR, which is not the case: elite firms tend to join those with comparable experience and size.

This functional form is log additively separable in the characteristics of private equity firms, E_a and S_a. It has an advantage in estimation because the cross-interaction effects will not be affected by whether it is a single firm deal or a club deal.

The estimation is semi-parametric. I will estimate the parameters in the latent equity return function, but specification of the probability distributions of the error terms is unnecessary (please see Fox [7]).

For the estimation purpose, I assume that, in the latent equity return function, the maximum debt to EBITDA multiple associated with a deal is unchanged regardless which private equity firm wins the bid. This assumption can be justified by the rationale that a company's cash flow must be strong enough to service post-buyout debt repayment and interest payment requirements, and the amount of debt a company can sustain in a leveraged buyout is restricted by the company's financial conditions. The benefit or cost of extra debt that is taken in a deal is then related to the winning private equity firm's characteristics.[8]

1.4.2 Private Equity Consortium Analysis

Recently, private equity firms often form consortia in corporate takeovers (Boone and Mulherin [3]). This type of deals are called *club deals*. A major criticism of private equity consortia is that joint bids in club deals tend to be less aggressive, which is against boards of directors' initial purpose of choosing auctions – to create highest possible bids for shareholders. This topic is important not only in academia, but also in policy, public interests,

[8]This point can also be explained by the practice of stapled finance, which is common in M&A markets after the year 2001. Stapled finance is a loan commitment arranged by the seller for whoever wins the bidding contest (Povel and Singh [29]), and the winning firm has an option but not obligation to choose this loan commitment. The size of the loan can be an indicator for the actual size of debt financing.

and courts. In this paper, I apply the same empirical technique to identify private equity firms' main motivation to form private equity consortia. I will leave systematic analysis to future research.

In the existing literature, there are some explanations for the practice of joint bidding. A group of private equity firms join together can obtain more debt financing with better terms; and they may also reduce the fierceness of bidding competitions. A third explanation is risk sharing: private equity firms form a consortium in one buyout transaction due to restrictions in their investment agreements with Limited Partners, which mandate the upper bound of the proportion of a certain fund that can be invested in one deal. As believed by the Limited Partners, joint investments can limit risk exposure of the injected equity by each consortium member. From an individual firm point of view, the main difference between these competing explanations is whether an extra consortium participant is an active investor or a passive free rider. So when a target company is matched with many private equity firms, this is equivalent to saying, whether adding one more firm to the acquiring team would significantly affect the performance of other investment partners.

I introduce a concept of Coalition Contribution Index (CCI) to measure popularity of a private equity firm in a consortium independent of its motivations, and then examine how CCI is interacting with the characteristics of the target companies in the latent equity return function. A private equity firm's Coalition Contribution Index measures the willingness of other private equity firms in accepting this firm as their investment partner. The CCI of a private equity firm is defined as the total number of its investment partners normalized by the total number of deals participated by that firm. The formation of coalitions is organization of the buyer side in LBO markets. And the Index indicates the popularity of a private equity firm in the process of this formation among its peers.

Direct study of formation of coalitions is cumbersome. It is technically infeasible to examine all possible coalitions empirically because of its large quantity of possible combinations, and we can only observe a limited number of realized coalitions. The Coalition Contribution Index is an alternative approach to study coalitions. The Index is assigned to each private equity firm, which is intrinsic to that particular firm. This index measures a firm's characteristics when it is interacting with other firms in a coalition, and it remains fixed when this firm is moved from one coalition to another. The Coalition Contribution Index enters the latent equity return function as interacting terms with the characteristics of the target companies.

Definition 3. *The latent equity return function of a transaction $\langle a, i \rangle$, in which target company i is acquired by private equity firms a, is*

$$\begin{aligned} r_{\langle a,i \rangle} = \quad & \exp\Big[\big(\alpha_1 M_i + \alpha_2 M_i^2 + \alpha_3 I_i + \alpha_4 Y_i + \alpha_5 R_i + \alpha_6 V_i\big) \\ & + \big(\beta_0^1 M_i + \beta_1^1 R_i + \beta_2^1 V_i\big) \cdot E_a + \big(\beta_0^2 M_i + \beta_1^2 R_i + \beta_2^2 V_i\big) \cdot S_a \\ & + \big(\beta_0^3 M_i + \beta_1^3 R_i + \beta_2^3 V_i\big) \cdot CCI_a + \varepsilon_\mu\Big]. \end{aligned} \tag{1.3}$$

The functional form for many-to-many matching is defined similar to (1.2). The empirical examination of the reason for consortia is to look at how the Index is interacting with the companies' characteristics. If a firm that joins a consortium can bring better debt financing and other expertise, it will have statistically significant connections with the maximum debt to EBITDA multiple, and this relation is positive. Otherwise, the Contribution Index will not have any such connection.

The Coalition Contribution Index is constructed in the same method as the concept of *degree* in graph theory. The graph theory has been applied in the study of social networks, and the study of networks of venture capital firms (Hochberg, Ljungqvist, and Lu [12]). The Contribution Index measures the weight of a private equity firm in coalitions. Similar to the social network measure, this concept can be easily illustrated in the same way by both matrix and graph representations, which are standard in the study of multilateral relations. Table 1.4 and Figure 1.2 are illustrations of an example, in which there are 6 transactions and 7 prominent private equity firms.

In the matrix representation, a symmetric matrix indicates the ties between private equity firms (Table 1.4). Two firms have a connection if they have been in at least one transaction together. In the setting of this paper, the matrix is undirected, that is, we do not distinguish lead investors from others. The diagonal elements are replaced with total number of deals in which the corresponding firms are involved. So the Coalition Contribution Index is computed as the sum of a column or a row's non-diagonal elements divided by the diagonal elements. In the graph representation of a square symmetric matrix, a node represents a private equity firm, and lines leading outward from that node represent its connections with other firms. The number of connections is defined as *degree* of that node. After normalizing by the number of that firm's overall transactions, we obtain its Coalition Contribution Index.

1.4.3 Equilibrium Concept

In matching games, agents compete against each other trying to match with attractive partners. For an outcome to be in equilibrium, the agents should not have incentive to deviate from their matched partners. The most widely used equilibrium concept is pairwise stable equilibrium. Research by Roth [30] suggests that natural experiments in UK medical intern markets have proved robustness of pairwise stable matchings. The estimation in this paper relies on the pairwise stability in LBO outcomes. In this section, I describe matching decision rules, agent preferences, and incentive constraints.

I follow the standard marriage matching decision rules with the one-to-one restriction relaxed. In the LBO matching markets, I assume, private equity firms search for potential target companies; owners of the companies that agree to sell organize formal auctions or solicit bids ("beauty contests"); the private equity firms submit bids; boards of directors of the companies select the winning bids. This is consistent with empirical observations that private equity firms submit bids and post-buyout plans through bidding contests and deals have to be approved by the target companies' board of directors and shareholders.

I do not explicitly model preferences of private equity firms or target companies, but their preferences depend on how they share total values created by matches. As described in section 1.4.1, t_μ specifies how they share the total return in the form of t_μ and r_μ/t_μ, which are log additive. In a realized transactions μ, t_μ is the actual return for existing shareholders of the target company, and they may not know the true value of r_μ.[9] In its preference,

[9] At initial stage of a selling process, shareholders of a target company do not know

a target company ranks private equity firms based on the magnitudes of t_μs the latter can offer. A private equity firm's payoff is positively related to the markup r_μ/t_μ, and negatively related to the amount of equity used to finance the deal. Private equity firms also rank target companies in their preferences. Allocations of bargaining power or intensity of competitions decide the location of t_μ on the interval of $[1, r_\mu]$.

In pairwise stable equilibrium, there is a set of incentive compatibility and individual rationality conditions: for any two pairs of matches in the match outcome with two target companies and two private equity firms, both the private equity firm and the target company are willing to participate in their deal (the "IR" constraint), and neither target company has incentive to switch to the other company's acquirer (the "IC" constraint). These conditions are summarized by a collection of inequalities.

Definition 4. *A feasible outcome (A, T) is a pairwise stable equilibrium of this matching game if:*

1. *For all $\langle a, i\rangle \in A$, $\langle b, j\rangle \in A$, $\langle b, i\rangle \notin A$, and $\langle a, j\rangle \notin A$,*

$$t_{\langle a,i\rangle} \geq r_{\langle b,i\rangle}/(r_{\langle b,j\rangle}/t_{\langle b,j\rangle}). \tag{1.4}$$

2. *For all $\langle a, i\rangle \in A$,*

$$t_{\langle a,i\rangle} \geq 1, \tag{1.5}$$

private equity bidders' r_μs. If board of directors of the target company is faithful (Cramton and Schwartz [4]), it will design a mechanism, such as first price seal bid auction, so that shareholders can infer the true r_μs after private equity bidders submitting their bids. This is explained by the revelation principle. However, r_μs are unobservable to econometricians neither before nor after the bidding.

and

$$r_{\langle a,i\rangle} \geq t_{\langle a,i\rangle}. \tag{1.6}$$

Part 1 of the definition states the incentive compatible constraint: company i would not deviate from its current buyer firm a to firm b even if firm b is indifferent of switching. In the inequality (1.4), $r_{\langle b,j\rangle}/t_{\langle b,j\rangle}$ is the actual markup that firm b can receive in the realized transaction $\langle b,j\rangle$; keep the value of firm b's current return unchanged and suppose firm b and company i are matched together, $r_{\langle b,i\rangle}/(r_{\langle b,j\rangle}/t_{\langle b,j\rangle})$ is the maximum return that firm b can pay to company i if b acquires i without hurting its current markup; if this value is less than company i's current return $t_{\langle a,i\rangle}$, i would not walk away from its current transaction and choose firm b instead of firm a. Note that the same inequality must hold for company j's incentive simultaneously.

To explain part 1 in the language of marriage matching games, in an outcome of a marriage matching market, for any two couples $\langle 1,1\rangle$ and $\langle 2,2\rangle$, where man 1 matches with woman 1 and man 2 matches with woman 2, and suppose women select men. These two couples are pairwise stable if the following condition is true, woman 1 would not marry man 2 instead of man 1 even if man 2 weakly prefers woman 1 to his current match woman 2; the same condition holds for woman 2.

A caveat is that the LBO matching markets are not efficient. It may happen that private equity firms and target companies which can create highest total values cannot be matched together. This is because fund sizes of a private equity firm are fixed and the number of deals it can participate in is limited. This limitation corresponds to the quota q_a^d assigned to each private equity firm in the model.

Some may wonder if pairwise stability is a realistic equilibrium concept in LBO markets, since quite often in a very short period, there are a group of private equity investors searching for investment opportunities in one industry and a group of companies searching for buyers in the mean time. Group stability is a straightforward equilibrium concept for these markets, since group stability requires stability among a group of private equity firms and a group of companies. The following proposition shows that it is sufficient to look at pairwise stable assignment outcomes.

Proposition 1. *In the matching game described above, an assignment outcome* (A, T) *is group stable if and only if it is pairwise stable.*

Next section introduces a feasible estimation procedure based only on deal match outcomes.

1.4.4 Related Inequalities

This paper uses a modified maximum score estimator originally studied by Fox [7]. This method makes empirical estimation of a matching model computationally feasible. In contrast, a direct approach requires checking the inequalities given in section 1.4.3 to predict possible matches and calculating likelihood of match assignments, then maximizing this likelihood with respect to parameters. The computation burden of the second approach is prohibitively heavy.

The maximum score estimator relies only on deal outcomes to estimate model parameters. The estimator searches for parameter values to maximize stability of observed LBO outcomes. In this paper, the model can estimate the actual scale of each parameter and there is no parameter normalization . This is because the deal premium t_μs are observable, which specify how target companies and private equity firms share the total return r_μs. By including deal premium data, the maximum score estimator scales parameter values and thus scales the latent equity return function in searching for solutions. In this way, the paper can estimate the actual magnitude of the value created through debt and the value created through monitoring.

The proposition states that the pairwise stable equilibrium exists and is unique under the assumptions: private equity firms submit bids; target companies select the winning bids; and each private equity firm has a quota on the total number of companies it can acquire.

Proposition 2. *The pairwise stable equilibrium of the matching game given above exists and is unique.*

The existence of the matching equilibrium is vital for the estimation, while the uniqueness is less important, since the estimator only checks the incentive compatibility condition, which is a necessary condition and must be satisfied by any pairwise stable equilibrium.

The essence of the maximum score estimator in this paper is the following. Section 1.4.1 introduces the latent equity return function, which calculates the expected total return that can be created by each feasible match between a target company and a private equity firm. The firm's inception year must be before the transaction year for a match to be feasible. Then I select two match outcomes, $\langle a, i\rangle$ and $\langle b, j\rangle$, that are realized transactions, and both

$\langle b, i \rangle$ and $\langle a, j \rangle$ are feasible but are not match outcomes. The set of match outcomes is a subset of all feasible matches. Whether these two transactions, $\langle a, i \rangle$ and $\langle b, j \rangle$, are pairwise stable is decided by the values of trial parameters in the latent equity return function. The estimation by this maximum score estimator is the set of parameter values that lead to the highest number of pairwise stable pairs for a given range of latent equity return values.

By Definition 4, the maximum score objective function is given by

$$\mathcal{Q} = \sum_{\langle a,i \rangle, \langle b,j \rangle \in A} \mathbf{1}\Big[\ \mathcal{I}\big(a, i; b, j\big)\ \Big], \tag{1.7}$$

where the group of inequalities $\mathcal{I}(a, i; b, j)$ are

$$\log(t_{\langle a,i \rangle}) \geq \log(r_{\langle b,i \rangle}) - \Big(\log(r_{\langle b,j \rangle}) - \log(t_{\langle b,j \rangle})\Big); \tag{1.8a}$$

$$\log(t_{\langle b,j \rangle}) \geq \log(r_{\langle a,j \rangle}) - \Big(\log(r_{\langle a,i \rangle}) - \log(t_{\langle a,i \rangle})\Big). \tag{1.8b}$$

This pair of inequalities is the work horse of the estimation, and I can estimate the parameters in the latent equity return function using the observed outcomes of the LBO markets. The indicator functions $\mathbf{1}[\cdot]$ are equal to 1 when the condition $\mathcal{I}$ is true and 0 otherwise. The estimator searches for parameter values that maximize the objective function. The estimation procedure will be outlined in the next section. This estimator does not impose the individual rationality constraints either for the target companies or for the private equity firms, because the transfer t_μ is a very noisy measure, and it is redundant to incorporate this part of error into the score function.

The general latent return function r_μ of a transaction $\mu = \langle a, i \rangle$ can be written as

$$r_\mu = \exp\Big[\, f_i\big(X_i\big) + f_a\big(X_a\big) + f_{a,i}\big(X_i \cdot X_a\big) + \varepsilon_\mu \,\Big], \tag{1.9}$$

where X_i and X_a are characteristic vectors for the target company i and the private equity firm a. Function f_i and f_a are functions solely of the characteristics of company i and firm a, respectively, while function $f_{a,i}$ is a function of cross products of characteristics X_i and X_a. Now, re-write equation (1.8b) as

$$\log(t_{\langle b,j\rangle}) - \log(t_{\langle a,i\rangle}) \geq \log(r_{\langle a,j\rangle}) - \log(r_{\langle a,i\rangle}). \tag{1.10}$$

The general latent return function r_μ is log additive separable in functions f_i, f_a, and $f_{a,i}$, and we can see that all terms in function f_a that are related to firm a are canceled out in equation (1.10). By the Theorems in Fox [7], the latent equity return function given by Definition 1 and 2 can be identified.

In comparison, Fox [7] introduces a concept of local production function. In the leveraged buyout context of this paper, the local return function for deal outcomes, $\langle a, i \rangle$ and $\langle b, j \rangle$, is defined as

$$f_{a,i;b,j} = \log(r_{\langle a,i\rangle}) + \log(r_{\langle b,j\rangle}). \tag{1.11}$$

And the corresponding maximum score objective function is given by

$$\mathcal{Q} = \sum_{\langle a,i\rangle,\langle b,j\rangle \in A} \mathbf{1}\Big[\; f_{a,i;b,j} > f_{b,i;a,j} \;\Big]. \tag{1.12}$$

The indicator functions $\mathbf{1}[\cdot]$ are equal to 1 when the inequality in brackets is true and 0 otherwise. The estimates of the parameters are those which maximize the objective function. This estimator can only identify the parameters in the function $f_{a,i}$, the cross interaction terms.

Note that if we sum inequality (1.8a) and inequality (1.8b), we get the inequality in equation (1.12). The return functions and the transfers, which satisfy conditions (1.8a) and (1.8b), must satisfy the inequality in equation (1.12). So the condition specified by local return function is only a necessary condition, while $\mathcal{I}(a, i; b, j)$ is both sufficient and necessary in terms of incentive compatibility.

1.4.5 Estimation

The estimation is a two-step procedure combining linear regressions and maximization of the score objective function.

The main step of the estimation is to search for parameter values that maximize the score objective function. The paper uses Differential Evolution (DE) method, an algorithm developed by Storn and Price [34]. DE is a stochastic direct search method that is designed to search for global maxima of objective functions; and it can handle objective functions with jumps. The objective function in the maximum score estimator is a step function, which is piecewise constant and piecewise continuous. If a global optimum exists, it will occur in a closed and bounded subspace of the parameter space. And at least one optimum will be at a jump of the objective function. Since the existence of global optimum is known (there are 1320, 3538, and 9082 feasible pairs of matches in the three merger waves respectively, the summation of which is the upper bound for the optimum), DE is a proper algorithm for the estimation in this paper.

Because the score objective function is a step function, there is a continuum of global optima in a very large space (there are 23 or 26 parameters), and it is difficult to decide which set of parameter values is relevant. Regarding this, the paper adopts a simple two-step estimation strategy. In the first step, I suppose target companies and acquiring firms split the total return by a fixed fraction λ, $0 < \lambda < 1$. Since deal premium is observable, this gives a set of total returns. λ's are chosen sporadically between 0 and 1 to avoid bias. Using OLS regression, we have the initial values of the parameters for the estimation. These initial values are further perturbed by noises with normal distributions to form an augmented set of initial values for the estimation. In the second step, the estimation runs a controlled DE algorithm. Controlled DE algorithm is in the sense that the average value of the latent equity return function is kept within a pre-determined range, for example, ten times the average deal premium, which corresponds to the situation that the existing shareholders of the target companies obtain 10% of the total return.

The computation is lengthy due to the estimation nature of structural models. I use subsampling to compute confidence intervals. 75% deals are randomly drawn from the full samples for one computation procedure, so that these random draws can spread through three LBO waves. Over $1,000$

such calculations are computed for confidence intervals. A small number of estimations that cause the latent equity return function well exceed the pre-determined range are discarded.

1.5 Data Description

1.5.1 Description of Sample

The data of LBO transactions come from the SDC Platinum database owned by Thomson Financial. This database is used in several previous studies of LBOs, e.g., Guo, Hotchkiss, and Song [11], and Officer, Ozbas, and Sensoy [27]. I focus on the LBOs sponsored by private equity firms.

First, I extract all completed transactions with announcement date between May 1986 and February 2007 in which the deal value is greater than $100 million, and the buyer owns more than 50% of the target's shares outstanding after the transaction. Among these transactions, the targets are identified by SDC as publicly traded companies, and the deals are characterized as "financial sponsors" with "leveraged buyout" acquisition technique. I also restrict the sample to those target companies with available financial data and reported acquisition premium. The initial screening leaves a sample of 297 possible LBOs. Each deal is examined manually to ensure it satisfies the above search criterions.

Then, by reading the comprehensive leveraged buyout reports of these deals, I identify the transactions sponsored by private equity firms or with participation of these firms. A financial buyer is treated as a private equity firm when it is: i) described as private equity firm, leveraged buyout firm, or venture capital in the business description of the acquirer or synopsis of the transaction; ii) reported in the May 2007 issue of Private Equity International (PEI) magazine; iii) claimed to be specialized in leveraged buyouts in its online profile. The sample includes divisions of the investment banks JP Morgan, Merrill Lynch, Morgan Stanley, and First Boston Credit Suisse, that provide equity financing in buyouts. In some cases, a founding partner leaves an established buyout firm and starts a new firm in the same business. For example, although Mr. Kohlberg is involved in the transactions sponsored by Kohlberg Kravis Roberts & Company (KKR) before his resignation in 1987, Kohlberg & Company and KKR are treated as two different private equity firms. I do not use the LBO and "going private" flags provided by SDC, since Officer, Ozbas, and Sensoy [27] reports missing data of the deals sponsored by private equity firms.

Data on loan terms come from Loan Pricing Corporation's DealScan database. Borrowers' names, loan spread and maximum debt to EBITDA multiple are observable in this database. Both loan spreads and debt to EBITDA multiples are widely used deal measures in academic research and in practice. In general, private equity firms sponsor loans with the target companies as the borrowers. To match the deals from SDC and the loans from DealScan, I first search target company names in DealScan, and then identify debt financing in LBO transactions by restricting deal purpose as LBO and limiting the time frames.

Finally, the deals in the sample are divided into three waves according to the year in which they are announced: wave from 1986 to 1991, from 1992 to

2001, and from 2002 to 2007. This partition of transactions is corresponding to the merger and buyout waves in the 1980s and the one from late 1990s to recent credit crunch (spanning across the bust of tech bubble in 2001). The three waves of deals will be estimated together but as three separate matching markets so that we can examine if there exists difference in value creation between these periods.

1.5.2 Description of Variables

From the sample, I form a deal matching matrix with the private equity firms and the target companies as rows and columns. Each entry in this matrix is a dummy variable which is 1 if the private equity firm and the target company corresponding to that entry are involved in a buyout transaction, and 0 otherwise. This is equivalent to saying that the dummy is 1 if they are in a match and 0 if they are not matched. Then, the sum of each row is the number of transactions completed by the private equity firm corresponding to that row; and the sum of each column is the number of financial buyers in one transaction, which is a club deal if the sum is greater than 1.

The observed firm characteristics are the inception year, and the cumulative committed capital raised. Age of a firm at the time of its transaction is a reasonable measure for its experience (please see Gompers [8], and Hochberg, Ljungqvist, and Lu [12]). The cumulative committed capital raised is the measure of a firm's ability. Private equity firms are organized as collections of funds. When the partners of a firm (the General Partners, the GPs) are seeking capital commitment from investors (the Limited Partners, the

LPs), the target fund size is an indicator of their intrinsic abilities. This is because under-commitment is detrimental to a firm's reputation and over-commitment will cause investors' concerns of suboptimal investing behaviors of the fund managers. The combination of the two measures, cumulative committed capital and age, can distinguish long time prominent private equity firms from those who are less active but have been in the LBO industry for a significantly long time, such as KKR versus Kelso & Company.

The observed exogenous characteristics for the target companies are the deal announcement date, industry, market value, and financial conditions before the transactions. The companies are divided into industries according to the Fama-French 12-industry portfolio categories. The target's performance is measured by the ratio of EBITDA to total assets for the fiscal year prior to the deal announcement. This is the measure of a company's fundamental characteristics. When the EBITDA data is missing for a target company, it is filled by the data from its major competitors in the same industry adjusted by total assets.

Characteristics of the buyout debt financing are loan spreads and debt to EBITDA multiples. Both measures are key terms in practice and in academic research (e.g., Ivashina and Kovner [14]). This paper will focus on the maximum debt to EBITDA multiple as the main description for the debt financing. Due to computation burden, the structural model in this paper is not able to analyze debt financing in more details. As pointed out by Kaplan and Stein [20], it is equally crucial to consider buyout loans contractual features – seniority, maturity, and the division between public and private lenders. I hope that advanced computation technology and econometric methods in the near future will allow us to address this research question.

1.6 Results

1.6.1 The Latent Equity Return Function

Section 1.4.5 describes the procedure of the estimation in details. The latent equity return function in the form of equation (1.2) is estimated with the industry and merger wave dummy variables included. In the latent equity return generated by the matching between a private equity firm and a target company, the financing benefit in LBO deals is estimated at 3.1%, and the management benefit by the investing firms is estimated at 7.8%. The null hypothesis that these two benefits are statistically equal is rejected at 1% level. Table 1.5 reports the estimation results. The empirical model in this paper is different from the classical regression model, and the interpretation is also different.

Not surprisingly, the coefficient of the maximum debt to EBITDA multiple is positive (0.0210) and statistically significant. The debt does affect the latent returns positively and significantly. The coefficient of the cross interaction term between the multiple and the age of the private equity firm is positive at 0.0002 and statistically significant at 10% level. This coefficient is the second order cross derivative of the latent equity return function with respect to the multiple and the firm age. It indicates that, at the average age level of private equity firms, the amount of debt borrowed by the private equity firms is increasing in firm's age. Older firms are established – their ability to screen for better deals and their ability to manage post-buyout companies are known to lenders. Banks provide bridge loans in hope to generate profits in post-buyout refinancing (e.g., issuance of high yield bonds, also known as junk bonds). Institutional investors, such as pension funds, are major high yield bond buyers. Established private equity firms benefit from their relationship with these lenders and this relationship is built upon the firms' solid investment track records.

The model can identify linear effects of target companies' characteristics on total returns of LBO transactions. Both coefficients of the term EBITDA on total assets (ROA) and the term log market value in the latent equity return function are negative and significant. The private equity investments in poorly performing companies or small companies are more profitable than those investments in good companies or large companies. Intuitively speaking, there is larger margin for operational improvement in poorly performing companies. This is also an indirect evidence of the important role that private equity firms are playing in LBO deals. As to the size effect of target companies, there is more asymmetric information problem associated with small public traded companies. Private equity investors as private investors do not have to reveal strategic information to the public, so they can invest in information opaque small companies and explore this advantages.

The investment period matters in the total return of a LBO deal. The estimation includes the time dummy variables for the LBO waves in the late 1990s and after the year 2001. If we compare the two latent equity return functions, one that does not contain the Coalition Contribution Index (CCI) with the one that does, *ceteris paribus*, a transaction in the late 1990s is less profitable than the one in the 1980s. This is true in both estimations. A transaction in the first few years after 2001 may also generate less total return, but this is no longer true when the CCI is included in the estimation. This is consistent with what we have observed in the history of the LBO sector of the private equity industry. In the late 1990s, an increasing number of private equity investors are competing for a limited number of investment opportunities, which drives down the average expected total returns. After the year 2001, less competent private equity investors have left the market and the industry is better organized. In the estimated latent equity return function with the CCI, the deals in the LBO wave after 2001 no longer have significant negative time effect.

In the estimation reported by Table 1.6, I examine which characteristics of target companies are more important in the private equity firms' deal choices for given levels of their age and size. I run "horse race" on the cross interaction terms between companies' return on assets, market value and firms' age, size. The data samples of return on assets and market value are converted into percentiles so that they are in the same unit for comparison. The latent equity return function is first estimated with the converted companies' characteristics. The coefficients for the cross interaction terms between companies' return on assets, market value and firms' age are available. Then I take difference of the absolute values of these two coefficients. The term with larger value wins the "horse race". For the significance of the difference, I randomly draw 75% deals from the full sample, and calculate the difference for each random draw. This gives the empirical distribution of the difference, and I can examine the significance from this distribution.

It is the same procedure for the cross interaction terms between companies' return on assets, market value and firms' size.

If I hold firm size fixed, the market value (size) of the target companies is a factor more important than the companies' financial conditions (return on assets) in the firms' decision to pursue the deal. This is because when the competition among private equity firms became fierce in the late 1980s and late 1990s, some private equity firms were focusing more on larger companies to avoid direct bidding competitions against younger and smaller firms. And large amount of commitment from institutional investors as Limited Partners also lead to bigger buyout funds and made large deals possible. In summary, size of the target companies is the first criterion that a private equity firm uses to screen for investment opportunities.

Given a firm's age, the estimation shows that target companies' financial conditions (return on assets) is more important than its size, but this relation is not significant. This is caused by the heterogeneity in the background of private equity firms. Some relatively younger firms are in fact founded by veterans from prestigious investment banks, who can raise large buyout funds and target bigger companies. But for some firms, although they have been in the private equity industry for a long period, they only restrict their focus in middle market on mid-size companies.

1.6.2 Contributions by "Congenial" Firms

In this section, I analyze whether the Coalition Contribution Index (CCI) is an important factor in the latent equity return function, and how it relates to the characteristics of the target companies. The Index enters the latent equity return function as cross interaction terms with debt financing, the maximum debt to EBITDA multiple, and the target company's characteristics, such as, return on total assets and market value. The estimation indicates a significant connection between high CCI and high debt to EBITDA multiple. This suggests, the main motivation for joint bidding practice is that private equity partners can bring better debt financing to club deals. Table 1.7 reports the estimation results.

In the estimated latent equity return function that includes CCI, the coefficient of the cross interaction term between the debt to EBITDA multiple and the Index is positive significant. This indicates that more "congenial" or popular firms tend to be associated with higher leverage in their buyout deals. Note that the estimation is over all buyout transactions, including both club deals and single firm deals. The estimated coefficient says that the firms with high CCI raise more debt financing in their transactions. During the due diligence process of each deal, the bidders project the target company's future cash flow and estimate the maximum debt that the post-buyout company can sustain. If a private equity firm bidder can raise cheaper debt financing, this will give that firm an advantage in the bidding competition for the target company. And the finding that these firms are well accepted by other private equity firms in club deals suggest that firms join together in order to obtain better debt financing.

It is true that sometimes private equity firms submit joint bids because of other concerns, but they may not be the dominant reasons. If the only purpose of club deals is to alleviate the severeness of bidding competitions, the CCI is a very conservative measure. The CCI is a special descriptive measure for private equity firms' unobservable characteristics, which to some extent indicates how well a firm is accepted by other private equity firms in club deals. The Index is constructed as the total number of other private equity firms which have been investment partners with this firm divided by the total number of deals this firm has been involved. If private equity firms intend to reduce competitions by joint bidding, the ability and the experience of partner firms are less important factors in formation of teams. In this situation, firms are more concerned with whether bidding partners are comfortable co-investors. Then social networks form among these firms, and bidders tend to join a smaller set of firms with closer social ties. However, we still find significant connection between CCI and debt financing.

The coefficient for the cross interaction term between target companies' return on assets and the CCI is also positive and significant. Intuitively, for the post-buyout company to enjoy tax benefit of debt, there has to be reasonable cash flow under tax shield. And strong cash flow can help the company to service more debt. This is an indirect evidence for the conjecture that the Index indicates a firm's ability to raise debt financing.

Several private equity firms team together may be able to acquire mammoth companies. And because the investment agreements between Limited Partners and General Partners usually mandate the maximum portion of a fund that can be invested in one company, fund managers of private equity firms can invest in bigger companies in club deals than by themselves. An interesting finding by the estimation of the latent equity return function is that the investment consortium makes the deal possible, but it is not a value creator. The coefficient for the cross interaction term between the market value of the target companies and the CCI is insignificant although positive.

1.6.3 Club Deals versus Single Firm Deals

The empirically estimated latent equity return function allow us to examine the difference between the performance of private equity firms in club deals and their performance in single firm deals. The paper restricts its focus on private equity firms that have participated in at least one leveraged buyout consortium. The goal of the analysis in this section is to answer the question, conditioned on the knowledge that a private equity firm has been involved in club deals, on average, what is the performance difference of that firm between club deals and single firm deals. There are firms in the sample which are never involved in club deals. These firms self-select to acquire target companies by themselves for many reasons, such as, their emphasis on management autonomy, i.e., some firms that need dominant control to implement a particular investment strategy do not want to share controls with other firms. Including these firms in the comparison will cause bias.

One difficulty of the comparison is that the model in this paper cannot identify linear effects from private equity firms in the latent equity return function, and the empirical estimated function is only a partial form of the "true" function. But, if we assume that the true equity return function is in the form of equation (1.9), then de-mean within the collection of all transactions by a specific firm can eliminate that firm's linear effect. De-mean normalizes all firms' expected total returns to one, which is equivalent to keeping all firms at the same starting point before a race. All deals finished by these firms are divided into two groups, club deals and single firm deals.

If private equity firms form consortium in order to alleviate bidding competitions, we should observe no significant difference in the latent returns calculated by the latent equity return function between club deals and single firm deals; while in the meantime, private equity firms can retain larger portions of the total value created in club deals. Opposite to this view, if private equity firms can bring better financing to a bidding consortium and share investment risks among themselves, we should observe no significant advantage held by private equity firms in sharing total returns with the target companies' existing shareholders, and more importantly, these firms may generate relatively higher expected total returns in club deals.

The paper uses a two-sample Kolmogorov-Smirnov test for the performance comparison analysis. Because the estimation of the latent equity return function is semiparametric – the probability distributions of the error terms are not specified, standard Chi-square test for comparative purposes is less favorable, which requires grouping observations into intervals but small number of intervals can change test statistics dramatically. K-S test is an alternative distribution free test of goodness of fit. As described in the model of this paper, the latent equity return function calculates the total expected equity return on any feasible match between a target company and a private equity firm, and this return plus some random shock, which is unknown to outsiders, gives the firm's *ex ante* belief of the prospect of the deal. Since each firm's average expected returns on all its deals is normalized to one, an intuitive method of comparison is to examine whether the returns for club deals and those for single firm deals are being drawn from the same probability distribution. This is related to comparison of value distributions from two separate sample sets. K-S test does not require the number of the observations in the two sample sets to be equal.

Denote the empirical cumulative probability distribution functions for the latent returns in club deals and in single firm deals as $F_c(\cdot)$ and $F_s(\cdot)$ respectively. The two-sample K-S test statistic is

$$D_{n,n'} = \sup_x \left| F_{c,n}(x) - F_{s,n'}(x) \right|,$$

where n and n' are their sample sizes. The Kolmogorov distribution is defined as the distribution of the random variable

$$K = \sup_{t \in [0,1]} \left| B(t) \right|,$$

where $B(t)$ is the Brownian bridge. The null hypothesis is that the two data samples come from the same probability distribution. The alternative hypothesis can be either the two data samples come from two different probability distributions, or one data sample comes from a distribution which stochastically dominates (or is stochastically dominated by) the second distribution. Depending on the alternative hypothesis, the test can be either two-tail or one-tail. Denote K_α as the critical values of the Kolmogorov distribution. The null hypothesis is rejected at level α if

$$\sqrt{\frac{nn'}{n+n'}} D_{n,n'} > K_\alpha.$$

(please see Massey [25] and Stephens [33].)

The paper finds that private equity firms generate higher latent returns in club deals statistically and have no significant advantage in sharing total returns. I have performed K-S test on both overall total returns and the private equity firms' share of returns, with the two data sample sets as club deal returns and single firm deal returns. Graphically, the empirical cumulative distribution function of the returns from club deals lies below that of the returns from single firm deals. Figure 1.4 and Figure 1.5 report the test results.

In the analysis of total values of latent equity return function, the null hypothesis of equal distribution is rejected at 1% significance level with p-value being 0.00. Although there are only small variations of the returns in both club deals and single firm deals, returns in club deals strongly dominate the returns in single firm deals, in the sense of first order stochastic dominance, which is illustrated in Figure 1.4 that the empirical cumulative distribution function of the returns in club deals lies below that of the returns in single firm deals. From the consortium analysis, we know that higher Coalition Contribution Index is associated with higher maximum debt to EBITDA multiple, and debt financing contributes insignificant economic value to leveraged buyout transactions. Combining this result with the finding in this section, we have seen strong evidence supporting that private equity firms do generate higher returns in club deals from a social welfare maximizer point of view.

To examine whether private equity firms have advantage in sharing total returns with target companies' existing shareholders in club deals, we subtract deal premia from the latent returns and perform K-S test on the firms' portion of the investment returns. The null hypothesis cannot be rejected at 5% significance level and the p-value is 0.0576. This result contributes greatly to our understanding of private equity firms' investment strategies. Many believe that private equity firms form consortia in order to facilitate collusion in the auctions of target companies. This is not surprising since joint bidding reduces the number of bidders. But this view fails to consider the situation that a handful of strong and serious bidders can generate higher total returns and higher deal premia than a pool of weak and less serious bidders.

1.7 Final Remarks

This paper develops a structural model based on a two-sided matching model to study LBOs sponsored by private equity, and it identifies and estimates the value created through debt and the value created through monitoring. I find that monitoring by private equity firms creates greater value than debt in private equity sponsored LBOs. My study also finds that private equity firms form consortia in LBOs in order to bring better debt financing to club deals. There is, however, no strong evidence supporting the view that these deals impede bidding competitions.

Separating the value created through debt and the value created through monitoring is important for understanding private equity sponsored LBOs. If the value created through debt is positive, any investor, who can acquire a company using large amounts of debt, is able to benefit from high leverage. Different from debt, the value created through monitoring is the value created by the activities of private equity investors. This value is specific to the joint relation between private equity firms and target companies, the matching relation.

The model in this paper solves an endogeneity problem caused by mutual selection by private equity firms and target companies in LBO markets. This structural approach in empirical corporate finance is of independent interest. In situations where researchers cannot observed economic agents' choice sets and decision making, or cannot identify directions of causality, it is difficult to apply traditional linear regression models. Structural models combine theoretical models with empirical estimation to address these problems. It may help further our understanding of the behaviors of investors and corporate managers.

1.8 Tables and Figures

Table 1.1: Summary Statistics: Private Equity Firms

The sample consists of 101 private equity firms founded between 1968 and 2004 ("inception year"). A financial buyer is classified as a private equity firm when it is: i, described as private equity firm, leveraged buyout firm, or venture capital in its business description or in deal synopsis; ii, reported in the May 2007 issue of Private Equity International (PEI) magazine; iii, claimed to be specialized in leveraged buyouts in its profile. The cumulative capital under management is the aggregate size of funds raised by a private equity firm for the purpose of conducting leveraged buyout deals.

		Cumulative Capital under Management (bil)			
Inception Year	Obs.	Min.	Max.	Mean	Median
1968	2	12.0	12.0	12.0	12.0
1969	2	1.8	3.0	2.4	2.4
1971	1	30.0	30.0	30.0	30.0
1972	1	1.4	1.4	1.4	1.4
1974	1	20.0	20.0	20.0	20.0
1975	1	2.5	2.5	2.5	2.5
1976	1	46.7	46.7	46.7	46.7
1978	4	1.0	11.0	5.9	5.8
1980	2	4.9	8.0	6.5	6.5
1981	1	2.0	2.0	2.0	2.0
1982	2	1.0	3.0	2.0	2.0
1983	4	2.5	12.5	5.5	3.5
1984	6	0.1	75.0	18.3	8.8
1985	4	6.0	51.0	23.9	19.3
1986	4	0.4	38.0	20.6	22.0
1987	3	2.8	85.5	30.7	3.7
1988	5	0.5	3.7	2.3	3.0
1989	6	0.6	22.0	8.3	6.5
1990	3	2.0	37.0	13.8	2.5
1991	3	2.0	3.3	2.7	2.8
1992	7	0.8	46.7	13.0	3.0
1993	2	2.0	7.5	4.8	4.8
1994	3	0.6	14.3	5.3	1.0
1995	9	0.5	24.0	5.2	2.0
1996	3	1.2	5.0	2.6	1.6
1997	3	0.5	2.0	1.3	1.5
1998	6	1.5	29.5	7.1	2.9
1999	4	0.2	14.0	4.0	0.9
2000	4	1.5	9.0	5.4	5.5
2001	1	8.0	8.0	8.0	8.0
2004	2	1.9	4.0	2.9	2.9
Total	101	0.1	85.5	9.6	3.2

Table 1.2: Summary Statistics: Target Companies

The sample consists of 208 completed leveraged buyout transactions sponsored by private equity firms from May 1986 to February 2007. Deals are identified by the target companies. Deal years are broken down into three waves according to the macro merger waves, from 1986 to 1991, from 1992 to 2001, and from 2002 to 2007. The data of maximum debt come from Loan Pricing Corporation's DealScan database, which includes credit lines, revolvers, and senior bank loans. The maximum debt to EBITDA multiples are trimmed to non-negative values. Return on assets is the ratio of EBITDA over total fixed assets. Size of the target companies is the market value of all equity outstanding 4 weeks before the deal announcement date.

		Maximum Debt to EBITDA Ratio			
Deal Year	Obs.	Min.	Max.	Mean	Median
1986-1991	52	-18.26	47.31	8.10	6.82
1992-2001	72	0.55	23.95	5.47	4.84
2002-2007	84	-311.27	34.89	7.39	6.82
1986-2007	208	-311.27	47.31	6.90	5.86

		Return on Total Assets			
Deal Year	Obs.	Min.	Max.	Mean	Median
1986-1991	52	-0.047	0.357	0.162	0.160
1992-2001	72	0.059	0.330	0.171	0.158
2002-2007	84	-0.022	0.295	0.135	0.126
1986-2007	208	-0.047	0.357	0.154	0.152

		Market Value 4 Weeks before Ann. (mil)			
Deal Year	Obs.	Min.	Max.	Mean	Median
1986-1991	52	31.8	13,593.3	903.6	300.3
1992-2001	72	67.8	4,605.3	402.0	216.6
2002-2007	84	46.9	18,206.2	2,582.0	830.8
1986-2007	208	31.8	18,206.2	1,407.8	383.7

Table 1.3: Deal Matching Matrix

Table 1.3 is a sub-matrix of the deal matching matrix, which illustrates the organization of the matching markets and their outcomes. The first row lists all target companies in the order of deal announcement dates, from May 1986 to February 2007. The first column lists all private equity firms which are first arranged by the years of inception, from the oldest to the newest, then ranked by size of the cumulative capital under management in descending order. Other entries of the matrix are either 0 or 1. If a private equity fund is the actual acquirer of a target company, the entry corresponding to that fund's parent firm and the target company is 1, and 0 otherwise. The sum of each row assigned to a private equity firm is the total number of deals in which the firm is involved. The sum of each column assigned to a target company is the number of private equity firm buyers. The sum is greater than one if it is a club deal. The matching matrix format is designed for sub-sampling, i.e., the sub-matrix for all deals which are completed by the private equity firms that are ever involved in club deals, is the collection of rows that have non-zero entries in the columns with greater than one summation.

	...	Safeway	Multicare	Panamsat	AMC	SunGard	Freescale	...
...								
Carlyle Group		0	0	1	0	0	0	
Bain Capital		0	0	0	0	1	0	
Blackstone		0	0	0	0	0	1	
KKR		1	0	1	0	0	0	
TPG		0	1	0	0	0	0	
Apollo Mgt.		0	0	0	1	0	0	
Permira		0	0	0	0	0	1	
...								

Table 1.4: Consortium Formation Matrix

This symmetric matrix represents the formation of the private equity consortium in the example of Table 1.3. The diagonal elements are the number of deals each firm has completed. Given a private equity firm, the sum of non-diagonal elements corresponding to that row or column represents the number of firms who have connections with the given firm. The connections are undirected, so the matrix is symmetric, which means the model does not distinguish which firm is the lead investor in a club deal. The Coalition Contribution Index (CCI) is computed as the number of connections normalized by the total number of deals conducted by that firm, that is, the sum of all non-diagonal entries of a row or a column divided by the value of the diagonal element in that row or column. For example, in this matching matrix, KKR is involved in 2 transactions with one of them a club deal, so its index is 0.5; while Bain Capital has conducted only one single firm deal, so its index is 0.

	...	Carlyle	Bain Capital	Blackstone	KKR	TPG	Apollo	Permira	...
...									
Carlyle Group		1	0	0	1	0	0	0	
Bain Capital		0	1	0	0	0	0	0	
Blackstone		0	0	1	0	0	0	1	
KKR		1	0	0	2	0	0	0	
TPG		0	0	0	0	1	0	0	
Apollo Mgt.		0	0	0	0	0	1	0	
Permira		0	0	1	0	0	0	1	
...									

Table 1.5: Latent Equity Return Function

The sample consists of 101 private equity firms, 208 leveraged buyout transactions of target companies with deal value $100 million or above during the period between May 1986 and February 2007. The data also include the matching outcomes between the private equity firms and the target companies. Given a feasible match between a private equity firm and a target company, the latent equity return function is a function of the characteristics of both the private equity firm and the target company, which calculates the maximum expected equity return the deal can possibly generate if the firm actually acquires the company. Private equity firms' characteristics are firm's age at the time of transaction and cumulative capital under management, which is the aggregate dollar value of leveraged buyout funds raised by the firm. Target companies' characteristics are return on total assets and market value. The maximum debt to EBITDA multiple is assumed to remain the same regardless which firm is the final winner of the bids. The model can identify linear terms of the company's characteristics and cross interactions between the firm and the company's characteristics in the latent equity return function. Target companies are categorized by Fama-French 12 industry portfolios according to their four digit SIC code at the time of transactions. All leveraged buyout deals are grouped into three markets, from 1986 to 1991, from 1992 to 2001, and from 2002 to 2007 following the three merger waves. The function is estimated by a modified maximum score estimator. The initial parameter values in the latent equity return function are obtained by OLS regressions with re-scaled deal premia as the dependent variable, under the assumption that the private equity firms and the target company existing shareholders share the deal surplus at a fixed ratio. The initial values are then perturbed by error terms with normal probability distributions. For a pair of match outcomes between private equity firms and target companies, if the latent equity return function with the trial parameter values allow this pair of matches satisfying pairwise stability, the estimator scores 1 on this pair; otherwise, it scores 0. A controlled Differential Evolution algorithm is run to search for parameter values that increase the value of the score objective function. The estimation of the parameters are the values that maximize the score objective function within a pre-specified range. Finding a global optimum is unnecessary. We use subsampling and computing two-tail confidence intervals to calculate significance levels of the parameters. In the report, ***, **, and * denote significance at 1%, 5%, and 10% level, respectively.

Table 1.5 of Latent Equity Return Function

		95% Confidence Interval	
	Coeff.	Lower Bound	Upper Bound
MULTIPLE	0.0210**	0.0127	0.0528
MULTIPLE ^{2}	7.38E-05**	2.15E-05	3.15E-04
INDUSTRY_1	0.8372**	0.2864	1.3053
INDUSTRY_2	0.7367*	-0.0431	1.1858
INDUSTRY_3	0.8962***	0.4003	1.2234
INDUSTRY_5	0.7006*	-0.0902	0.9130
INDUSTRY_6	0.8209**	0.3327	1.0886
INDUSTRY_7	0.9726***	0.5294	2.1573
INDUSTRY_8	0.8991*	-0.3425	1.2984
INDUSTRY_9	0.8169**	0.5972	1.1120
INDUSTRY_10	0.8063**	0.4425	1.0878
INDUSTRY_11	0.7680***	0.4612	0.8725
INDUSTRY_12	0.7872**	0.3563	0.9351
WAVE_2	-0.2127*	-0.4671	0.0265
WAVE_3	-0.2862*	-0.5169	0.3549
ROA	-0.1653*	-0.6812	0.0008
ln MARKET_VALUE	-0.0196***	-0.1654	-0.0089
MULTIPLE $*$ ln AGE	0.0002*	-0.0002	0.0006
ROA $*$ ln AGE	0.0344*	-0.0131	0.0573
ln MARKET_VALUE $*$ ln AGE	-0.0007**	-0.0015	-0.0001
MULTIPLE $*$ ln SIZE	-0.0019**	-0.0119	-0.0008
ROA $*$ ln SIZE	-0.0388	-0.0671	0.0809
ln MARKET_VALUE $*$ ln SIZE	0.0019***	0.0016	0.0251

Table 1.6: Relative Importance

The sample consists of 101 private equity firms, 208 leveraged buyout transactions of target companies with deal value \$100 million or above during the period between May 1986 and February 2007. The data also include the matching outcomes between the private equity firms and the target companies. Given a feasible match between a private equity firm and a target company, the latent equity return function is a function of the characteristics of both the private equity firm and the target company, which calculates the maximum expected equity return the deal can possibly generate if the firm actually acquires the company. Private equity firms' characteristics are firm's age at the time of transaction and cumulative capital under management, which is the aggregate dollar value of leveraged buyout funds raised by the firm. Target companies' characteristics are return on total assets and market value. The maximum debt to EBITDA multiple is assumed to remain the same regardless which firm is the final winner of the bids. Target companies are categorized by Fama-French 12 industry portfolios according to their four digit SIC code at the time of transactions. All leveraged buyout deals are grouped into three markets, from 1986 to 1991, from 1992 to 2001, and from 2002 to 2007 following the three merger waves. The function is estimated by a modified maximum score estimator. The initial parameter values in the latent equity return function are obtained by OLS regressions with re-scaled deal premia as the dependent variable. The initial values are then perturbed by error terms with normal probability distributions. For a pair of match outcomes between private equity firms and target companies, if the latent equity return function with the trial parameter values allow this pair of matches satisfying pairwise stability, the estimator scores 1 on this pair; otherwise, it scores 0. A controlled Differential Evolution algorithm is run to search for parameter values that increase the value of the score objective function. The estimation of the parameters are the values that maximize the score objective function within a pre-specified range. To examine the relative importance of target companies' characteristics in the decisions of forming matches, we run "horse race" on the cross interactions between firm and company characteristics. Companies' return on total assets and market value are re-scaled into percentiles, such that they are in the same unit. To determine which factor is more important in match formation given firm age at the time of transaction, absolute values of the coefficients of the term ROA·ln(AGE) are subtracted by absolute values of the coefficients of the term ln(MV)·ln(AGE). If the result is positive, then ROA wins the "horse race"; otherwise, the factor of market values wins. Similarly, we can examine which factor is more important in match formation given firm size. We use subsampling and computing two-tail confidence intervals to calculate significance levels of the differences between coefficients. In the report, ***, **, and * denote significance at 1%, 5%, and 10% level, respectively.

Table 1.6 of Relative Importance

	Coeff.	95% Confidence Interval	
		Lower Bound	Upper Bound
MULTIPLE	0.0203*	-0.0034	0.0315
MULTIPLE ^{2}	9.64E-05**	3.14E-05	4.84E-04
INDUSTRY_1	0.8094**	0.4061	1.3586
INDUSTRY_2	0.4681**	0.3680	2.1081
INDUSTRY_3	0.8650***	0.6004	1.3786
INDUSTRY_5	0.6012*	-0.0719	0.9976
INDUSTRY_6	0.8042***	0.6398	1.0065
INDUSTRY_7	0.9183**	0.3647	1.2204
INDUSTRY_8	0.9129*	-1.8227	1.7818
INDUSTRY_9	0.7842***	0.6601	1.1420
INDUSTRY_10	0.7970**	0.3901	1.1611
INDUSTRY_11	0.6765***	0.4178	0.9574
INDUSTRY_12	0.8440***	0.5605	0.9305
WAVE_2	-0.1539*	-0.3075	0.0069
WAVE_3	-0.3294*	-0.5688	0.1658
ROA	-0.0405*	-0.1307	0.0051
ln MARKET_VALUE	-0.2161***	-1.5822	-0.1572
MULTIPLE $*$ ln AGE	0.0004*	-0.0006	0.0008
ROA $*$ ln AGE	0.0083*	-0.0013	0.0449
ln MARKET_VALUE $*$ ln AGE	-0.0073*	-0.0148	0.0019
MULTIPLE $*$ ln SIZE	-0.0021***	-0.0037	-0.0014
ROA $*$ ln SIZE	-0.0067	-0.0100	0.0033
ln MARKET_VALUE $*$ ln SIZE	0.0215***	0.0186	0.1532
	Diff(Coeff.)		
HORSE RACE on ln AGE	0.0006	-0.0115	0.0341
HORSE RACE on ln SIZE	-0.0375***	-0.1510	-0.0109

Table 1.7: Private Equity Consortium Analysis

The sample consists of 101 private equity firms, 208 leveraged buyout transactions of target companies with deal value \$100 million or above during the period between May 1986 and February 2007. The data also include the matching outcomes between the private equity firms and the target companies. Given a feasible match between a private equity firm and a target company, the latent equity return function is a function of the characteristics of both the private equity firm and the target company, which calculates the maximum expected equity return the deal can possibly generate if the firm actually acquires the company. Private equity firms' characteristics are firm's age at the time of transaction and cumulative capital under management, which is the aggregate dollar value of leveraged buyout funds raised by the firm. Firms' characteristics also include the Coalition Contribution Index (CCI), which measures a firm's weight or importance in formation of bidding coalitions. A firm's CCI is computed as the total number of firms that have connections with the given firm normalized by the total number of deals that firm is involved in. Target companies' characteristics are return on total assets and market value. The maximum debt to EBITDA multiple is assumed to remain the same regardless which firm is the final winner of the bids. Target companies are categorized by Fama-French 12 industry portfolios according to their four digit SIC code at the time of transactions. All leveraged buyout deals are grouped into three markets, from 1986 to 1991, from 1992 to 2001, and from 2002 to 2007 following the three merger waves. The function is estimated by a modified maximum score estimator. The initial parameter values in the latent equity return function are obtained by OLS regressions with re-scaled deal premia as the dependent variable. The initial values are then perturbed by error terms with normal probability distributions. For a pair of match outcomes between private equity firms and target companies, if the latent equity return function with the trial parameter values allow this pair of matches satisfying pairwise stability, the estimator scores 1 on this pair; otherwise, it scores 0. A controlled Differential Evolution algorithm is run to search for parameter values that increase the value of the score objective function. The estimation of the parameters are the values that maximize the score objective function within a pre-specified range. Finding a global optimum is unnecessary. We use subsampling and computing two-tail confidence intervals to calculate significance levels of the parameters. In the report, ***, **, and * denote significance at 1%, 5%, and 10% level, respectively.

Table 1.7 of Private Equity Consortium Analysis

	Coeff.	95% Confidence Interval	
		Lower Bound	Upper Bound
MULTIPLE	0.0277	-0.0302	0.0338
MULTIPLE ^{2}	6.61E-05	-2.19E-04	3.20E-04
INDUSTRY_1	0.8438**	0.3219	1.2248
INDUSTRY_2	0.8956**	0.1090	2.6855
INDUSTRY_3	0.7715**	0.1150	1.0424
INDUSTRY_5	0.8060*	-0.0085	3.5647
INDUSTRY_6	0.7603***	0.6728	1.2889
INDUSTRY_7	0.9574**	0.2759	1.4759
INDUSTRY_8	0.8860**	0.3564	2.3435
INDUSTRY_9	0.8086***	0.5958	1.0162
INDUSTRY_10	0.7800***	0.2191	1.2170
INDUSTRY_11	0.6440**	0.2760	1.5692
INDUSTRY_12	0.9011***	0.5028	1.0185
WAVE_2	-0.1968**	-0.3892	-0.0202
WAVE_3	0.3609	-1.1504	0.2218
ROA	0.0573	-0.2626	0.2824
ln MARKET_VALUE	-0.0615***	-0.2432	-0.0250
MULTIPLE ∗ ln AGE	0.0002*	-0.0000	0.0002
ROA ∗ ln AGE	0.0340	-0.1151	0.1203
ln MARKET_VALUE ∗ ln AGE	-0.0008*	-0.0013	0.0042
MULTIPLE ∗ ln SIZE	-0.0036*	-0.0067	0.0007
ROA ∗ ln SIZE	-0.0448**	-0.1407	-0.0021
ln MARKET_VALUE ∗ ln SIZE	0.0089	-0.0180	0.0262
MULTIPLE ∗ CCI	0.0060**	0.0054	0.0491
ROA ∗ CCI	0.0871*	-0.1146	0.1588
ln MARKET_VALUE ∗ CCI	0.0060	-0.0241	0.8018

Table 1.8: Firm Size Effect on Performance

The sample consists of 101 private equity firms, 208 leveraged buyout transactions of target companies with deal value \$100 million or above during the period between May 1986 and February 2007. The data also include the matching outcomes between the private equity firms and the target companies. To examine the performance difference between big firms and small firms, the full sample of private equity firms are divided into two groups with equal number according to their cumulative capital under management. Two latent equity return functions are estimated separately for the two groups. For any feasible match between a private equity firm and a target company, the latent equity return function is a function of the characteristics of both the private equity firm and the target company, which calculates the maximum expected equity return the deal can possibly generate if the firm actually acquires the company. Private equity firms' characteristics are firm's age at the time of transaction and cumulative capital under management, which is the aggregate dollar value of leveraged buyout funds raised by the firm. Target companies' characteristics are return on total assets and market value. The maximum debt to EBITDA multiple is assumed to remain the same regardless which firm is the final winner of the bids. The model can identify linear terms of the company's characteristics and cross interactions between the firm and the company's characteristics in the latent equity return function. Target companies are categorized by Fama-French 12 industry portfolios according to their four digit SIC code at the time of transactions. All leveraged buyout deals are grouped into three markets, from 1986 to 1991, from 1992 to 2001, and from 2002 to 2007 following the three merger waves. Industry and market dummy variables are included in the estimation but not reported. The functions are estimated by the modified maximum score estimator. The initial parameter values in both estimations are the estimated values for the full sample latent function. For a pair of match outcomes between private equity firms and target companies, if the latent equity return function with the trial parameter values allow this pair of matches satisfying pairwise stability, the estimator scores 1 on this pair; otherwise, it scores 0. In order to estimate the function for a particular firm size group, in the pair of matches under examination of pairwise stability, there must be at least one private equity firm from that size group, big or small. Then the controlled Differential Evolution algorithm is run to search for parameter values that increase the value of the score objective functions. The estimation of the parameters are the values that maximize the score objective function within a pre-specified range. Finding a global optimum is unnecessary. We use subsampling and computing two-tail confidence intervals to calculate significance levels of the results. In the report, ***, **, and * denote significance at 1%, 5%, and 10% level, respectively.

Table 1.8 of Firm Size Effect on Performance

		95% Confidence Interval	
Big Firms	Coeff.	Lower Bound	Upper Bound
MULTIPLE	0.0210**	0.0127	0.0528
MULTIPLE ^{2}	7.38E-05**	2.15E-05	3.15E-04
ROA	-0.1653*	-0.6812	0.0008
ln MARKET_VALUE	-0.0196***	-0.1654	-0.0089
MULTIPLE $*$ ln AGE	0.0002*	-0.0002	0.0006
ROA $*$ ln AGE	0.0344*	-0.0131	0.0573
ln MARKET_VALUE $*$ ln AGE	-0.0007**	-0.0015	-0.0001
MULTIPLE $*$ ln SIZE	-0.0019**	-0.0119	-0.0008
ROA $*$ ln SIZE	-0.0388	-0.0671	0.0809
ln MARKET_VALUE $*$ ln SIZE	0.0019***	0.0016	0.0251

		95% Confidence Interval	
Small Firms	Coeff.	Lower Bound	Upper Bound
MULTIPLE	0.0166**	0.0127	0.0528
MULTIPLE ^{2}	1.31E-04**	2.15E-05	3.15E-04
ROA	-0.1708*	-0.6812	0.0008
ln MARKET_VALUE	-0.0212***	-0.1654	-0.0089
MULTIPLE $*$ ln AGE	-0.0002*	-0.0002	0.0006
ROA $*$ ln AGE	0.0463*	-0.0131	0.0573
ln MARKET_VALUE $*$ ln AGE	-0.0011**	-0.0015	-0.0001
MULTIPLE $*$ ln SIZE	-0.0161**	-0.0119	-0.0008
ROA $*$ ln SIZE	-0.0505	-0.0671	0.0809
ln MARKET_VALUE $*$ ln SIZE	0.0007***	0.0016	0.0251

Table 1.9: Firm Age Effect on Performance

The sample consists of 101 private equity firms, 208 leveraged buyout transactions of target companies with deal value $100 million or above during the period between May 1986 and February 2007. The data also include the matching outcomes between the private equity firms and the target companies. Since private equity firms may be involved in multiple deals, to examine the performance difference when firms are young and when firms are more experienced, the observed firm company pairs rather than firms alone are separated into two age groups with equal number, the old and the young. Two latent equity return functions are estimated separately for the two groups. For any feasible match between a private equity firm and a target company, the latent equity return function is a function of the characteristics of both the private equity firm and the target company, which calculates the maximum expected equity return the deal can possibly generate if the firm actually acquires the company. Private equity firms' characteristics are firm's age at the time of transaction and cumulative capital under management, which is the aggregate dollar value of leveraged buyout funds raised by the firm. Target companies' characteristics are return on total assets and market value. The maximum debt to EBITDA multiple is assumed to remain the same regardless which firm is the final winner of the bids. The model can identify linear terms of the company's characteristics and cross interactions between the firm and the company's characteristics in the latent equity return function. Target companies are categorized by Fama-French 12 industry portfolios according to their four digit SIC code at the time of transactions. All leveraged buyout deals are grouped into three markets, from 1986 to 1991, from 1992 to 2001, and from 2002 to 2007 following the three merger waves. Industry and market dummy variables are included in the estimation but not reported. The function is estimated by the modified maximum score estimator. The initial parameter values in both estimations are the estimated values for the full sample latent function. For a pair of match outcomes between private equity firms and target companies, if the latent equity return function with the trial parameter values allow this pair of matches satisfying pairwise stability, the estimator scores 1 on this pair; otherwise, it scores 0. In order to estimate the function for a particular firm age group, in the pair of matches under examination of pairwise stability, there must be at least one private equity firm from that age group, old or young. Then the controlled Differential Evolution algorithm is run to search for parameter values that increase the value of the score objective functions. The estimation of the parameters are the values that maximize the score objective function within a pre-specified range. Finding a global optimum is unnecessary. We use subsampling and computing two-tail confidence intervals to calculate significance levels of the results. In the report, ***, **, and * denote significance at 1%, 5%, and 10% level, respectively.

Table 1.9 of Firm Age Effect on Performance

Old Firms	Coeff.	95% Confidence Interval	
		Lower Bound	Upper Bound
MULTIPLE	0.0533**	0.0127	0.0528
MULTIPLE ^{2}	1.19E-04**	2.15E-05	3.15E-04
ROA	-0.4634*	-0.6812	0.0008
ln MARKET_VALUE	-0.0098***	-0.1654	-0.0089
MULTIPLE $*$ ln AGE	0.0011*	-0.0002	0.0006
ROA $*$ ln AGE	0.0838*	-0.0131	0.0573
ln MARKET_VALUE $*$ ln AGE	-0.0004**	-0.0015	-0.0001
MULTIPLE $*$ ln SIZE	-0.0087**	-0.0119	-0.0008
ROA $*$ ln SIZE	0.0041	-0.0671	0.0809
ln MARKET_VALUE $*$ ln SIZE	0.0062***	0.0016	0.0251

Young Firms	Coeff.	95% Confidence Interval	
		Lower Bound	Upper Bound
MULTIPLE	0.0210**	0.0127	0.0528
MULTIPLE ^{2}	7.38E-05**	2.15E-05	3.15E-04
ROA	-0.1653*	-0.6812	0.0008
ln MARKET_VALUE	-0.0196***	-0.1654	-0.0089
MULTIPLE $*$ ln AGE	0.0002*	-0.0002	0.0006
ROA $*$ ln AGE	0.0344*	-0.0131	0.0573
ln MARKET_VALUE $*$ ln AGE	-0.0007**	-0.0015	-0.0001
MULTIPLE $*$ ln SIZE	-0.0019**	-0.0119	-0.0008
ROA $*$ ln SIZE	-0.0388	-0.0671	0.0809
ln MARKET_VALUE $*$ ln SIZE	0.0019***	0.0016	0.0251

Table 1.10: Value Creation Analysis

In the latent equity return function, I group all terms that contain the maximum debt to EBITDA multiple M_i into the category of "debt"; group all terms related to the characteristics of private equity firms E_a and S_a (but not M_i) into the category of "monitoring". Using the estimated parameters for the full sample and the subsamples under different firm characteristics, i.e., age and size, and using average values of the variables in those terms, I calculate the estimated value created through debt and the estimated value created through monitoring.

$$\log\ r_{\langle a,i\rangle} = \left(\alpha_1 M_i + \alpha_2 M_i^2 + \alpha_3 I_i + \alpha_4 Y_i + \alpha_5 R_i + \alpha_6 V_i\right) + \left(\beta_0^1 M_i + \beta_1^1 R_i + \beta_2^1 V_i\right) \cdot E_a + \left(\beta_0^2 M_i + \beta_1^2 R_i + \beta_2^2 V_i\right) \cdot S_a + \varepsilon_{\langle a,i\rangle} \tag{1.13}$$

This calculation is based on two key assumptions: (i) private equity firms that have no experience and no ability do not create value; (ii) private equity firms as stand alone entities do not create value.

	Within Total Value Created by an Average Match	
	Debt	Monitoring
Full Sample	3.1%	7.8%
Large PE Firms	2.2%	8.3%
Small PE Firms	-50.7%	-2.4%
Older PE Firms	-14.0%	56.2%
Younger PE Firms	3.0%	7.7%

Figure 1.1: Firm Age Distribution

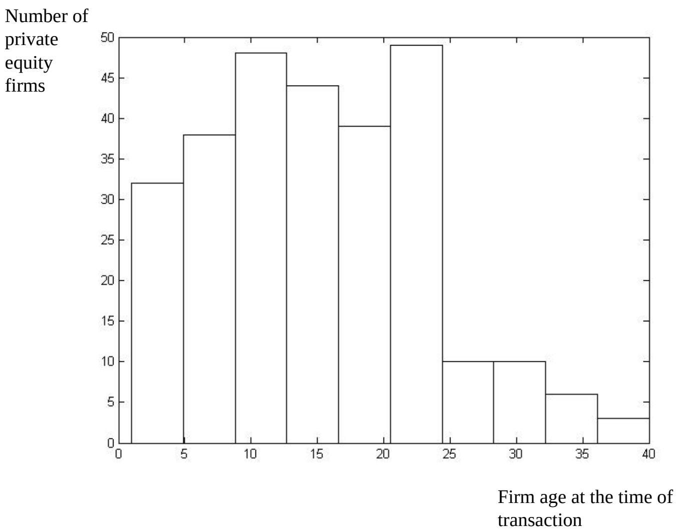

Figure 1.1 illustrates the age distribution of private equity firms at the time of transactions. Since a firm may be active in more than one deal, some private equity firms are counted multiple times; while several firms may also join together to acquire a target company in a club deal, so some deals are also counted multiple times. There are 279 deal-age observations overall. These observations are grouped into 10 4-year intervals, and the figure shows the number of firms in each interval.

Figure 1.2: Consortium Formation Graph

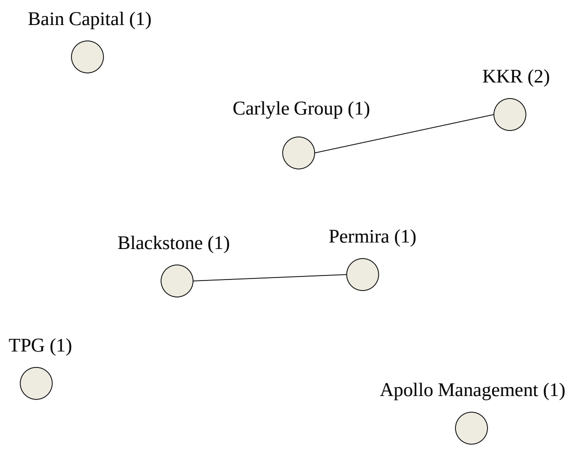

Figure 1.2 is the graph equivalent of the Table 1.4 on formation of private equity consortia in leveraged buyout deals. It is an illustrative selection of 6 deals conducted by 7 prominent private equity firms. Nodes on the graph represent private equity firms, and lines connecting the nodes indicate those firms that are partners in deals. The graph is undirected since lines do not show which firm is the lead investor. The number in the parenthesis is the number of deals that private equity firm has participated in this subsample. The number of lines leading out of a node is defined as the *degree* of the corresponding private equity firm in this graph. The degree of a private equity firm measures the extent how much other private equity firms are willing to accept this particular firm as a bidding partner. The Coalition Contribution Index (CCI) of a private equity firm is the degree normalized by the total number of deals that firm is involved in.

Figure 1.3: Differential Evolution Algorithm

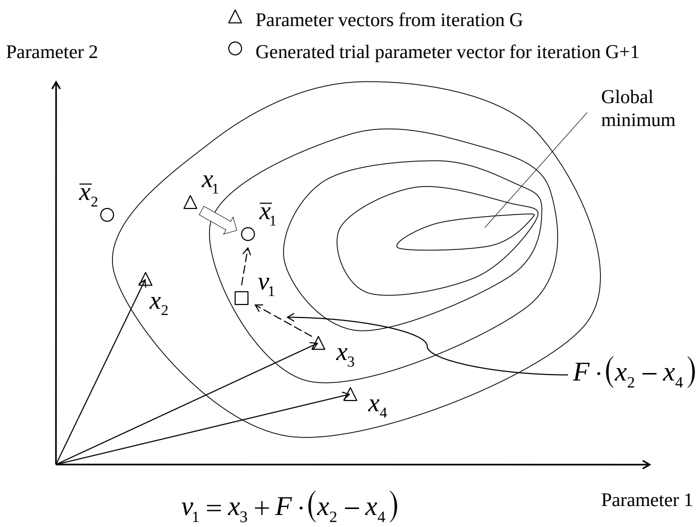

Figure 1.3 is a graph illustration of the Differential Evolution algorithm introduced by Storn and Price [34]. DE algorithm is a stochastic direct search method for global optimum. Suppose we are searching a global minimum of a cost function, and we start from N initial trial parameter vectors. Given N, $N \geq 4$, parameter vectors of dimension D in iteration G, the N new vectors in iteration $G+1$ are generated as the following. For vector x_1, randomly select 3 vectors other than x_1 in the same iteration G, e.g., x_2, x_3, and x_4. Generate vector v_1 as $x_3 + F \cdot (x_2 - x_4)$, where F is a random number strictly between 0 and 2. A trial vector $\bar{x}_1$ is generated by random selecting entries from x_1 and v_1. At least one entry from v_1 must be selected. If $\bar{x}_1$ yields a smaller cost function value than x_1, $\bar{x}_1$ replaces x_1 as a new parameter vector in iteration $G+1$; otherwise, x_1 is retained. x_2, ..., x_N are treated with the same procedure from iteration G to iteration $G+1$. In this example, $\bar{x}_1$ replaces x_1 while x_2 is retained in the iteration $G+1$.

Figure 1.4: Single Firm Deal versus Club Deal: Raw Returns

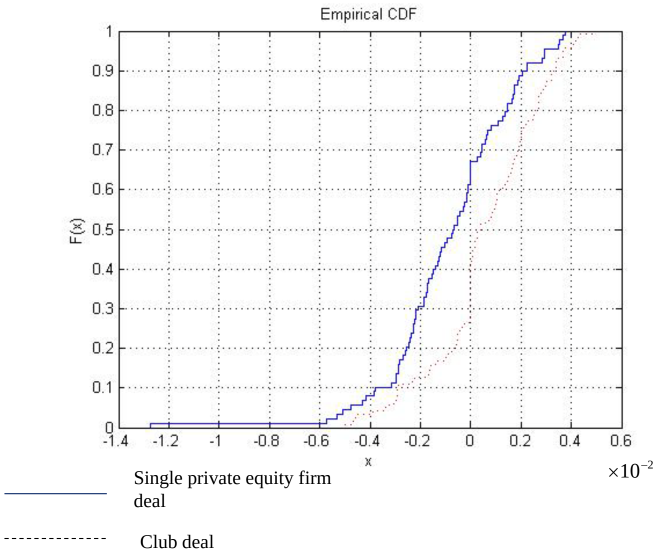

Figure 1.4 draws two empirical cumulative distribution functions of the values of the latent equity return functions grouped by whether the transactions are club deals or single firm deals. The private equity firms which have joined at least one consortium are first selected from the full sample. Then the deals conducted by those firms are selected, which include single firm deals but the acquiring firms are involved in club deals of other target companies. The values of the corresponding latent equity return functions are de-meaned at firm level to eliminate firm linear effect, since the model cannot identify this effect. A Kolmogorov-Smirnov test is performed to examine whether these two groups of values are coming from a same probability distribution, which is the null hypothesis. The alternative hypothesis is that the returns by club deals are first order stochastic dominating the returns by single firm deals, or equivalently speaking, the distribution function of club deal latent equity return values is smaller in general. The test is an one-tail K-S test. The null hypothesis is rejected at 1% level with p-value equal to 0.00 and the test statistics is 0.3531.

Figure 1.5: Single Firm Deal versus Club Deal: Firm Surplus

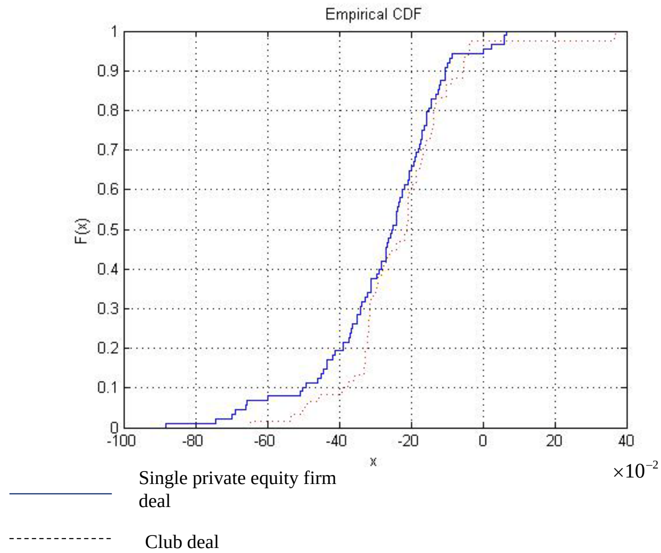

Figure 1.5 draws two empirical cumulative distribution functions of the private equity firms' portion of the deal surplus grouped by whether the transactions are club deals or single firm deals. The private equity firm's portion of the deal surplus is the actual expected return a firm can create from a deal after paying the target company's existing shareholders. The private equity firms which have joined at least one consortium are first selected from the full sample. Then the deals conducted by those firms are selected, which include single firm deals but the acquiring firms are involved in club deals of other target companies. The values of the corresponding latent equity return functions are first de-meaned at firm level to eliminate firm linear effect, then subtracted by log one plus deal premia. A Kolmogorov-Smirnov test is performed to examine whether these two groups of values are coming from a same probability distribution, which is the null hypothesis. The alternative hypothesis is that the returns by club deals are first order stochastic dominating the returns by single firm deals, or equivalently speaking, the distribution function of firm actual returns from club deals is smaller in general. The test is an one-tail K-S test. The null hypothesis cannot be rejected at 5% level. The p-value is 0.0576 and the test statistics is 0.1837. The null hypothesis cannot be rejected either in a two-tail K-S test when the alternative hypothesis is that the two empirical distribution functions are coming from two different probability distributions.

Chapter 2

Repeated Games with Learning and Private Monitoring in Continuous Time

2.1 Introduction

This paper analyzes a class of two-player games in continuous time with learning of a state parameter and imperfect private monitoring. Each individual player's payoff depends on a common state parameter of the nature. The true value of this parameter is unknown to both players throughout the horizon of these games. But players can infer its value by observing public signals distorted by other players' actions and Brownian motions. In addition to the public signals, each player observes imperfect private information about the opponent's actions. Without this private monitoring, players have no incentive to cooperate. Private monitoring filters out distortions in public signals caused by the other player's actions, and leads to possible cooperation. The objective of this paper is to characterize the equilibrium set by means of the stochastic filtering theory.

Understanding of repeated games with imperfect private monitoring is clumsy although not impossible, since these games lack the *recursive structure*, and checking incentives at each stage requires complex *statistical inference* (Kandori [41]). This paper does not claim that it completely solves this open problem, instead, it does propose an approach in a continuous time setting which relies heavily on the stochastic filtering theory. At the expense of simplicity, the continuous time setting and the existence of the unknown state parameter greatly improve the tractability of the model, which cannot be achieved in discrete time repeated games.

In this class of games I consider, both players are facing a common uncertain parameter, and its corresponding the public signals are distorted by players' actions and Brownian motions. Meanwhile, players' private information also contains noise in the form of Brownian motions. The public and the private information of all the players generate a filtration of σ-algebras, such that the information structure of each player is a filtration of coarser σ-algebras. Once the game is started, each player has her own private information and shares the public information with the other player. A player's private strategies depend on all the history of the game available to that particular player including her perceived value of the state parameter. This leads to private strategies as functionals of all known sample pathes. From the view of the stochastic filtering theory, one player's statistical inference of the opponent's actions and the state parameter is a *projection* of the general sample pathes to her own information sets, as represented by her own coarser σ-algebras. This avoids analyzing players' beliefs in the probability distribution form which is known to be dauntingly complicated in the imperfect private monitoring case.

The paper shows the existence and the uniqueness of this information structure. And it has also proved a certainty equivalence principle, which is the counterpart of the separation and certainty equivalence principle in the optimal control theory. This principle validates the concept of Nash equilibrium with learning, in the sense that, players conjecture each other's actions and infer the value of the state parameter, then play the best responses; in the Nash equilibrium outcomes, both the conjecture and the inference are correct (sequential).

In a game without private monitoring, because of the existence of the unknown state parameter and players' actions can distort the opponent's statistical inferences, players have no incentive to cooperate, and playing instant stage game Nash equilibrium throughout the game is the only equilibrium outcome. To be more precise, cooperation requires at some point of time some player does not play best responses to the other player's actions. Without private monitoring, public signals are common knowledge, and by the certainty equivalence principle, these signals can be separated into two components, a stochastic process for the state parameter and a deterministic process indicating players' actions. If this player deviates, the deviation contaminates the stochastic process perceived by her opponent, and it can decoy the opponent into choosing actions desirable by the former player. Foreseeing this, no player would cooperate.

The case of continuous time repeated games with imperfect private monitoring is more interesting. A special linear form of the public signals and availability of private monitoring allow the players to filter out distortions by the opponent's actions from the public signals. This makes learning of the state parameter possible. And by the certainty equivalence principle, the games can be transformed into a class of continuous time repeated games with imperfect (public) monitoring, which is similar to the one studied by Sannikov [32], but the key difference here is the persistent heterogeneous information sets when players have private information.

Given any action set of each player, the set of vectors of feasible players' actions is a subset of a two (the number of players) dimensional Euclidean space. This set is fixed once given. The dynamics of the game are partially driven by the perceived values of the state parameter by each player. The perceived values generate a sample path as the game continues, which is shown to be connected to a collection of Brownian motions. With probability one, this sample path will force the equilibrium actions hit the boundary of the set given above. This is equivalent to a passage time problem, and its solution in turn is equivalent to the solution of a harmonic equation with boundary conditions specific to any given game.

This class of games resembles many real economic practices. In reality there are many situations in which a small number of agents are competing in a long-term relationship with uncertain state of the nature, and the agents receive imperfect private information about the opponents' actions. Consider the following example in financial markets. Financial economists have long been searching for theoretical foundations of the interactions among privately informed investors in a risky environment. Suppose there are a group of large hedge funds investing in an emerging market. The investment opportunities are limited and the true aggregate market investment returns are unknown to all investors. Each investor's activities not only affect the other investors' returns, but also influence their learning of the true characteristics of this emerging market. The financial reports by the local news media and governments are public information observable by all investors. Additionally, investors can obtain some private information about each other's activities through social networks. This is a typical setting of continuous learning and private monitoring.

This paper is a natural but nontrivial extension of Sannikov [32]. Sannikov [32] studies a class of continuous time games, in which players' choices of actions are not directly observable. In his model, all possible signals are public information, so it is sufficient to analyze the set of all payoff pairs achievable by public perfect equilibria. My paper introduces unknown state parameter and imperfect private monitoring to repeated games in a continuous time setting. All three factors combined lead to tractability of statistical inferences by the players, and allow feasible characterization of equilibrium sets technically.

Major proofs and mathematical derivations can be found in the appendix.

2.2 Literature Review

Kandori [41] presents a brief introduction to repeated games with imperfect private monitoring.

2.3 The Model

Consider a two player repeated game in continuous time with learning and imperfect private monitoring. The nature is characterized as a complete probability space $(\Omega, \mathcal{F}, P)$, with $\{\mathcal{F}_t\}_{t=0}^{\infty}$ an increasing family of σ-algebras and $\mathcal{F}_0$ contains all sets of probability measure 0.

The state of the nature $\{X_t\}$ follows a stochastic process

$$dX_t = 0, \quad \text{i.e. } X_t = X_0; \quad \mathbb{E}X_0 = \mu > 0, \quad \mathbb{E}(X_0 - \mu)^2 = \theta^2 > 0. \tag{2.1}$$

Players' instantaneous payoffs will be depending on this state parameter.

At any moment $t \in [0, \infty)$, players i chooses action k_t^i from a compact convex subset $\mathcal{A}^i$ of $\mathbb{R}^+$. Players' actions cannot be observed directly, but they can be revealed in the following two ways.

A public signal is given in the stochastic differential equation form:

$$dY_t = (X_t - \mu(k_t))dt + dZ_t, \quad Y_0 = 0; \tag{2.2}$$

where $k_t = (k_t^1, k_t^2)$, $k_t^i \in \mathcal{A}^i$, $i = 1, 2$, and (Z_t) is a standard Wiener process. Then player i's payoff on an interval of length dt at time t is

$$c_i(k_t^i)dt + b_i(k_t^i)dY_t.$$

Besides the public signal, player i can receive private information related to the other player's actions, in the form of a stochastic process

$$dY_t^{(i)} = \mu(k_t)dt + dZ_t^{(i)}, \quad Y_0^{(i)} = 0. \tag{2.3}$$

Another form of private monitoring could be

$$dY_t^{(i)} = k_t^{-i}dt + dZ_t^{(i)}$$

instead of equation (2.3), but they are technically equivalent up to a functional transformation, while using equation (2.3) will have the advantage of more straight forward mathematical derivation.

Player i's information structure can be written as

$$\begin{bmatrix} dY_t^{(i)} \\ dY_t \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} X_t dt + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \mu(k_t)dt + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} dZ_t^{(i)} \\ dZ_t \end{bmatrix}. \tag{2.4}$$

$(Z_t^{(1)})$, $(Z_t^{(2)})$ are standard Wiener processes, and $(Z_t^{(1)})$, $(Z_t^{(2)})$, (Z_t) are assumed to be pairwise independent.

Let r be discount factor, then player i's total discounted payoff for a profile of strategies $\kappa_t = (\kappa_t^1, \kappa_t^2)$, $t \in [0, \infty)$, is given in the Itô integral form[1]:

$$
\begin{aligned}
J_i(\kappa^i(\cdot), \kappa^{-i}(\cdot)) &= r \int_0^\infty e^{-rt}(c_i(k_t^i)dt + b_i(k_t^i)dY_t) \\
&= r \int_0^\infty e^{-rt}\Big(c_i(k_t^i)dt + b_i(k_t^i)\big(X_t - \mu(k_t)\big)\Big)dt + r \int_0^\infty e^{-rt} b_i(k_t^i)dZ_t.
\end{aligned}
\tag{2.5}
$$

It can be shown later in this paper that there exist τ^i, a stopping time for player i, when she stops learning. And (τ^i, τ^j) is a structural breaking point for the game.

Let

$$
g_i(k_t) = c_i(k_t^i)dt + b_i(k_t^i)\big(X_t - \mu(k_t)\big)
$$

be player i's expected payoff flow at time t.

Define the continuation value of player i based on all player i's information at time t as

$$
\begin{aligned}
W_t^i(\kappa) &= \mathbb{E}_t\left[r \int_t^\infty e^{-r(s-t)}\big(c_i(k_s^i)ds + b_i(k_s^i)dY_s\big) \;\Big|\; \mathcal{F}_t^{Y,Y^{(i)}}, \kappa_s, s \in [t, \infty) \right] \\
&= \mathbb{E}_t\left[r \int_t^\infty e^{-r(s-t)} g_i(k_s)ds \;\Big|\; \mathcal{F}_t^{Y,Y^{(i)}}, \kappa_s, s \in [t, \infty) \right].
\end{aligned}
\tag{2.6}
$$

The σ-algebra $\mathcal{F}_t^{Y,Y^{(i)}}$ reflects all player i's information both public and private. Player i infers the value of the state parameter and her opponent's possible actions, and this statistical inference affects the expected discounted payoff at time t.

[1] In this paper, the notation κ_t emphasizes the functional form of the players' strategies, while k_t is the realization of these strategies, which is a history of continuous actions after κ_s, $0 \le s \le t$, acting on the sample pathes in the players' information sets.

2.4 Statistical Inferences: A Functional Analysis Approach

When stage game outcomes are distorted by noises and monitoring is imperfect, statistical inference is essential for the players to choose future actions. The paper relies on continuous stochastic filtering theory in formulating players' beliefs. One advantage which continuous time filtering has over traditional Bayesian approach is that classical Bayesian statistical inference requires reducing the stochastic processes into discrete grids, which is inconvenient for dynamic programming. But, the most important contribution of the filtering theory is that it provides a tractable mathematical representation of the players' beliefs.

Based on Fujisaki-Kallianpur-Kunita equation of the filtering theory, the method used in this paper connects the time t conditional expectation of the opponent's realized private history to a subspace of the Hilbert space spanned by all finite, real linear combinations of all feasible history realizations in a player's information set at time t. In the simplest setting, the linear estimation of a single decision maker problem, a generic element in this Hilbert space is a stochastic integral over sample pathes in the decision maker's information set. Thus, strategies are in fact a series of functionals indexed by time t mapping from the information sets to the action spaces.

2.5 The Existence Theorem

In continuous time setting, the evolution of the game and the private monitoring can be represented by stochastic processes. The players' strategies in general are history dependent, which creates *feedback* effect in those processes. In this section, I will show the existence and uniqueness of the signal processes associated with the game and the monitoring, this in turn builds the foundation for studying the information structures of repeated games in continuous time.

Suppose player 1 and 2 are playing a continuous time game with time $t \in [0, \infty)$. The nature is characterized as a complete probability space $(\Omega, \mathcal{F}, P)$ on which is given a q-dimensional standard Wiener process (Z_t). Here $q = m + n_1 + n_2$.

Define the notation square bracket of time t in subscript as $U_{[t]} = \{U_s, 0 \leq s \leq t\}$ for any given stochastic process U_t, $t \geq 0$. Then the state of the nature is following a m-dimensional state process (X_t) determined by the stochastic differential equation

$$dX_t = A(t, X_{[t]}, k_t^1, k_t^2)dt + B(t, X_{[t]}, k_t^1, k_t^2)dZ_t, \tag{2.7}$$

where k_t^i is player i's action at t.

Player i can receive a n_i-dimensional observation process $(Y_t^{(i)})$ determined by the stochastic differential equation

$$dY_t^{(i)} = a_i(t, X_{[t]}, k_t^1, k_t^2)dt + b_i(t, k_t^1, k_t^2)dZ_t. \tag{2.8}$$

Then the true state of the nature is described by an increasing family $\{\mathcal{F}_t^{X,Z}\}_{t=0}^{\infty}$ of σ-algebras. And the information structure for player i is given by $\{\mathcal{F}_t^{Y^{(i)}}\}_{t=0}^{\infty}$, also an increasing family of σ-algebras.

Let the action sets $\mathcal{A}^i$, $i = 1, 2$, be a family of complete, separable metric spaces, from which player 1 and player 2 choose actions at any moment $t \geq 0$. Players' strategies are defined as following:

Definition 5. *A player i's strategy is a profile of functionals $\{\kappa_t^i\}_{t\geq 0}$ which map sample pathes of $Y_t^{(i)}$ to $\mathcal{A}^i$:*

$$\kappa_t^i : \ Y_{[t]}^{(i)} \ \longmapsto \ k_t^i \in \mathcal{A}^i. \tag{2.9}$$

Then, for a given pair of player 1 and player 2's strategies, $\{\kappa_t^1\}_{t\geq 0}$ and $\{\kappa_t^2\}_{t\geq 0}$, the equation (2.7) and equation (2.8) are reduced to

$$dX_t = A(t, X_{[t]}, Y_{[t]})dt + B(t, X_{[t]}, Y_{[t]})dZ_t, \tag{2.10a}$$

$$dY_t^{(i)} = a_i(t, X_{[t]}, Y_{[t]})dt + b_i(t, Y_{[t]})dZ_t, \quad i = 1, 2, \tag{2.10b}$$

where (Y_t) is the (n_1+n_2)-dimensional process $(Y_t^{(1)}, Y_t^{(2)})$. Then by Theorem 8.2.1. in Kallianpur [40], the following claim is true.

Theorem 1. ***(Existence of Observation)*** *For any given pair of players' strategies, the stochastic equations (2.10b) for i = 1, 2, have a unique solution $(Y_t) = (Y_t^{(1)}, Y_t^{(2)})$.*

In general, player i's total discounted payoff for a profile of strategies κ

will be given in the form

$$J_i(\kappa^i(\cdot),\kappa^{-i}(\cdot)) = r\int_0^\infty e^{-rt} f_i(t,X_t,k_t^i,k_t^{-i})dt + r\int_0^\infty e^{-rt} h_i(t,X_t,k_t^i,k_t^{-i})dY_t. \tag{2.11}$$

For simplicity, this paper only considers the case in which the state parameter and players' actions are separable and additive in the signals and the payoffs.

Now back to the original game given by equations (2.1), (2.2), and (2.3). I return to the original notations where Y_t is the public signal and $Y_t^{(i)}$ is player i's private information, respectively. We have,

Corollary 1. *For any given pair of players' strategies, the stochastic equations (2.2) and (2.3) have a unique solution* $(Y_t, Y_t^{(1)}, Y_t^{(2)})$.

Let the true state of nature be given by an increasing family $\{\mathcal{F}_t^{X,Z}\}$ of σ-algebras. Both player 1 and player 2 can commonly observe a signal process $\{Y_t\}_{t\geq 0}$. Denote $\mathcal{G}_t^\flat := \mathcal{F}_t^Y$, the σ-algebras generated by $Y_{[t]} = \{Y_s, 0 \leq s \leq t\}$ when $\mu(k_t)$ is forced to be 0. Let $\mathcal{G}_t^\kappa$ be the σ-algebras when the process $\{Y_t\}_{t\geq 0}$ involves the $\mu(k_t)$ term. Players' actions and strategies satisfy the following property.

Let $\mathcal{A}^i$ be a compact convex subset of $\mathbb{R}^+$, the nonnegative real numbers, for all $t \geq 0$ and $i = 1, 2$. Let $\mathcal{K}^i$ be a set of piecewise continuous functions k_t^i of time t with values in $\mathcal{A}^i$, each function k_t^i being the player's actions on the time interval $[0,\infty)$[2], which is right continuous and has left limits at jumps (RCLL). Assume $\mathcal{K}^i$ has the following property. If k_t^i is in $\mathcal{K}^i$ and

[2]The paper distinguishes actions from strategies in the following way. For given time t, κ_t^i as a strategy is a functional mapping sample pathes (e.g. $Y_{[t]} = \{Y_s, 0 \leq s \leq t\}$) to the set $\mathcal{A}^i$, while for $t \in [0,\infty)$, k_t^i as an player's actions is the realization of those strategies, which is a function mapping time t to the set $\mathcal{A}^i$.

for $l = 1, \dots, L$, $k_l^i \in \mathcal{A}^i$ and $t_l \le t < t_l + h_l$ are non-overlapping intervals intersecting $[0, \infty)$ then

$$\tilde{k}_t^i = \begin{cases} k_l^i & \text{if } t_l \le t < t_l + h_l \\ k_t^i & \text{if } t \in [0, \infty) \text{ and } \notin \text{ one of the intervals } t_l \le t < t_l + h_l \end{cases} \tag{2.12}$$

is in $\mathcal{K}^i$. This allows the approximation of players' actions by piecewise constant actions (see Fleming and Rishel [38], p. 24).

Denote the game as $\Gamma(s, X_s)$ when it starts at time s and the initial state is X_s.

Definition 6. *A* ***Nash Equilibrium*** *of the game* $\Gamma(0, X_0)$ *is a pair of strategy profiles* $\{\kappa_t^{1*}, \kappa_t^{2*}\}_{t=0}^{\infty}$, *such that for any other strategy profile* κ^i,

$$J_i(\kappa^i(\cdot), \kappa^{-i*}(\cdot)) \le J_i(\kappa^{i*}(\cdot), \kappa^{-i*}(\cdot)), \quad i = 1, 2.$$

Definition 7. *A pair of strategy profiles is* ***Subgame Perfect Equilibrium*** *of the game* $\Gamma(0, X_0)$ *if it is Nash equilibrium for all subgames* $\Gamma(s, X_s)_{s\ge 0}$, *when the strategies are restricted to the subgames.*

In the original game given by equations (2.1), (2.2), and (2.3), the linearity of the stochastic processes in the information structure leads to the following theorem on certainty equivalence principle, related to the one in the stochastic control theory.

Let $\widehat{X}_t^{(i)}$ denote player i's expected value of X_t at time t. Denote $\widehat{X}_t := (\widehat{X}_t^{(1)}, \widehat{X}_t^{(2)})$. Let $\widehat{\Gamma}(s, \widehat{X}_s)_{s\ge 0}$ be the games in which player i chooses actions with respect to the process $\widehat{X}^{(i)}$.

Theorem 2. ***(The Certainty Equivalence Principle)*** *The sets of subgame perfect Nash equilibria of the game* $\Gamma(s, X_s)$ *and the game* $\widehat{\Gamma}(s, \widehat{X}_s)$*,* $s \geq 0$*, are identical.*

Please see appendix for the proof of Theorem 2.

Proposition 3. *The set of subgame perfect Nash equilibria of the game* $\widehat{\Gamma}(s, \widehat{X}_s)_{s\geq 0}$ *is nonempty.*

Please see appendix for the proof of Proposition 3.

Corollary 2. *A Nash equilibrium of the game* $\Gamma(s, X_s)_{s\geq 0}$ *is the fixed point in the space* $\mathcal{C}[0, \infty)$*, all continuous functions on the interval* $[0, \infty)$*. Player* i *infers* $\mathbb{E}[\kappa_t^j | \mathcal{F}_t^{Y,Y^{(i)}}]$ *and* $\widehat{X}_t^{(i)}$ *from the public signal* Y_t *and the private signal* $Y_t^{(i)}$*.* κ_t^i *is the best response chosen as an element in* $\mathcal{C}[0, \infty)$ *with respect to the inferences.*

Proof of Corollary 2. Using equation (2.1), (2.2), and (2.3), player i cannot observe the following system equations,

$$dX_t = 0; \tag{2.13a}$$

$$dY_t^{(j)} = \mu(k_t^i, k_t^j)dt + dZ_t^{(j)}. \tag{2.13b}$$

By Proposition 3, κ_t^j is a functional of $Y_{[t]}$ and $Y_{[t]}^{(j)}$ with realization k_t^j in $\mathbb{R}^+$, so the system equations (2.13) can be reduced to

$$dX_t = 0; \tag{2.14a}$$

$$dY_t = \Big(X_t - \mu\big(k_t^i, k_t^j(Y_{[t]}, Y_{[t]}^{(j)})\big)\Big)dt + dZ_t; \tag{2.14b}$$

$$dY_t^{(j)} = \mu\big(k_t^i, k_t^j(Y_{[t]}, Y_{[t]}^{(j)})\big)dt + dZ_t^{(j)}. \tag{2.14c}$$

Equation (2.14b), the public signal, enters the system equations because player j relies on this signal to infer $\widehat{X}_t^{(j)}$.

Player i's observable signals are

$$dY_t = \Big(X_t - \mu\big(k_t^i, k_t^j(Y_{[t]}, Y_{[t]}^{(j)})\big)\Big)dt + dZ_t; \tag{2.15a}$$

$$dY_t^{(i)} = \mu\big(k_t^i, k_t^j(Y_{[t]}, Y_{[t]}^{(j)})\big)dt + dZ_t^{(i)}. \tag{2.15b}$$

By Theorem 1, the observations $Y_t^{(i)}$ and $Y_t^{(j)}$ exist and are unique (applying the theorem twice, for player i and then for player j because of the symmetric form of the information structure). Then as in the proof of Theorem 2, $\widehat{X}_t^{(i)}$ and $\widehat{X}_t^{(j)}$ exist and are unique. By the proof of Proposition 3, the claim follows.

Q.E.D.

2.6 Learning without Private Monitoring

Proposition 4. *In a game $\Gamma(s, X_s)_{s\geq 0}$ without private monitoring, the only Nash equilibrium is the one in which all players play instant stage game Nash equilibrium. This equilibrium is a Perfect Public Equilibrium.*

Please see appendix for the proof of Proposition 4.

2.7 Analysis of the Monitoring Game

Proposition 5. ***(Representation and Promise Keeping)*** *A stochastic process W_t^i is the continuation value $W_t^i(\kappa)$ of player i under a strategy profile κ if and only if there exist progressive measurable processes $\beta^i = \{\beta_t^{i0}, \beta_t^{i1}, \mathcal{F}_t^{Y,Y^{(i)}}; 0 \leq t < \infty\}$ satisfying*

$$\mathbb{E}\int_0^T \left(\beta_t^{ih}\right)^2 dt < \infty; \quad h = 1, 2 \tag{2.16}$$

for every $0 < T < \infty$, and a square integrable process ξ_t^i orthogonal to the linear functional space spanned by $\{Y, Y^{(i)}\}$, such that for all $t \geq 0$, W_t^i satisfies the following stochastic equation

$$\begin{aligned} W_t^i = W_0^i + r\int_0^t \left(W_s^i - g_i(k_s)\right)ds + r\int_0^t \beta_s^{i0}\Big(dY_s - \widehat{X}_s^{(i)}ds + \mu(k_s^i, \hat{k}_s^j)ds\Big) \\ + r\int_0^t \beta_s^{i1}\Big(dY_s^{(i)} - \mu(k_s^i, \hat{k}_s^j)ds\Big) + r\int_0^t \xi_s^i ds \end{aligned} \tag{2.17}$$

Proposition 6. ***(Incentive Compatibility)***

Proposition 7. *The set of subgame perfect Nash equilibria of the game $\Gamma(s, X_s)_{s\geq 0}$ can be characterized by a harmonic equation with game specific boundary conditions.*

Please see appendix for the proof of Proposition 7.

2.8 An Example

First consider the simple case, game $\Gamma(0, X_0)$ given by equation (2.1), (2.2), and (2.5), with $T \in (0, +\infty)$, or $T = +\infty$.

First consider the stochastic inference problem when there is no distortion by the investors' actions k_t^i in the state and observation processes:

$$dX_t = 0, \quad \text{i.e. } X_t = X_0; \quad \mathbb{E}X_0 = \mu > 0, \quad \mathbb{E}(X_0 - \mu)^2 = \theta^2 > 0; \tag{2.18}$$

and the corresponding observations

$$dY_t^\flat = X_t dt + \sigma dZ_t, \quad Y_0^\flat = 0. \tag{2.19}$$

By the Kalman-Bucy filter model (see Øksendal [46] Theorem 6.10. and Example 6.11.), and let $S(t)$ be the error (co)variance, $S(t) = \mathbb{E}[(X_t - \widehat{X}_t)^2]$, then $S(t)$ satisfies the (deterministic) Riccati equation

$$\frac{dS}{dt} = -\frac{1}{\sigma^2} S^2, \quad S(0) = \theta^2$$

i.e.

$$S(t) = \frac{\theta^2 \sigma^2}{\sigma^2 + \theta^2 t}, \quad t \geq 0.$$

This gives the stochastic differential equation for the observation $\widehat{X}_t$,

$$d\widehat{X}_t = -\frac{\theta^2}{\sigma^2 + \theta^2 t} \widehat{X}_t dt + \frac{\theta^2}{\sigma^2 + \theta^2 t} dY_t^\flat, \quad \widehat{X}_0 = \mathbb{E}X_0 = \mu \tag{2.20}$$

then

$$d\Big(\widehat{X}_t \exp\Big(\int_0^t \frac{\theta^2}{\sigma^2+\theta^2 s}ds\Big)\Big) = \exp\Big(\int_0^t \frac{\theta^2}{\sigma^2+\theta^2 s}ds\Big)\frac{\theta^2}{\sigma^2+\theta^2 t}dY_t^\flat$$

which gives

$$\widehat{X}_t = \frac{\sigma^2\mu + \theta^2 Y_t^\flat}{\sigma^2+\theta^2 t}, \quad t \geq 0. \tag{2.21}$$

By equation (2.19), the stochastic differential equation (2.20) for the observation $\widehat{X}_t$ can also be written as

$$d\widehat{X}_t = \frac{\theta^2}{\sigma^2+\theta^2 t}(dY_t^\flat - \widehat{X}_t dt) = \frac{\theta^2}{\sigma^2+\theta^2 t}dN_t, \tag{2.22}$$

where N_t is the *innovation process* defined by

$$dN_t = (X_t - \widehat{X}_t)dt + \sigma dZ_t. \tag{2.23}$$

And by Øksendal [46] Lemma 6.7., N_t is a Gaussian process with orthogonal increments, and $\mathbb{E}[N_t^2] = t$ (thus N_t is a generalized Brownian motion).

Define the process $Y_t^\dagger$ as the solution of the following ordinary differential equation

$$dY_t^\dagger = (-k_t^1 - k_t^2)dt, \quad Y_0^\dagger = 0. \tag{2.24}$$

Then

$$Y_t = Y_t^\flat + Y_t^\dagger. \tag{2.25}$$

The method in the proof of the following results is standard in optimal control theory, which can be found in Davis [37], Fleming and Rishel [38].

Theorem 3. ***(The Separation Principle)*** *Inference of the model parameter is independent of the other investor's strategies in the statistical sense.*

Proof of Theorem 3.

Q.E.D.

Let $\kappa^i_{t,n}$ be the instant stage game Nash equilibrium strategy, $\kappa^i_{t,n} = (1/3)\widehat{X}_t$. Let $\kappa^i_{t,c}$ denote the cooperative strategy which is investing at the half of the monopoly investment level $\kappa^i_{t,c} = (1/4)\widehat{X}_t$. Here $i = 1, 2$. More generally, the cooperative strategy could be the investment levels at $\kappa^i_{t,c} = \lambda\widehat{X}_t$ with $\lambda \in [1/4, 1/3)$. These strategies are well defined because $\widehat{X}_t$ depends on $Y_{[t]}$ by (2.19), (2.20), (2.24), and (2.25).

Proposition 8. *The game $\Gamma(0, X_0)$ has a unique Nash equilibrium. In this equilibrium, investor 1 and investor 2 choose symmetric investment strategies $\kappa^i_{t,n}$ and play instant stage game Nash equilibrium at any time $t \in [0, T]$. This equilibrium is subgame perfect.*

Please see appendix for the proof of Proposition 8.

Proposition 9. *Investor i's investment strategy κ^i_t is given in the following form:*

$$\kappa^i_t = \alpha(t) + \int_0^t \beta(t,s)dY_s, \tag{2.26}$$

where

$$\alpha(t) = \frac{1}{3}\mu\sigma^{\frac{2}{3}}(\sigma^2 + \theta^2 t)^{-\frac{1}{3}}, \tag{2.27}$$

and

$$\beta(t,s) = \frac{1}{3}\theta^2(\sigma^2 + \theta^2 t)^{-\frac{1}{3}}(\sigma^2 + \theta^2 s)^{-\frac{2}{3}}. \tag{2.28}$$

Please see appendix for the proof of Proposition 9.

2.9 Proofs

2.9.1 Proof of Theorem 2

Multiply both sides of the equation (2.4), player i's information vector, by $[1, 1]$. We have

$$d(Y_t^{(i)} + Y_t) = X_t dt + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} dZ_t^{(i)} \\ dZ_t \end{bmatrix}. \tag{2.29}$$

Combine equation (2.1) and (2.29), then use the multi-dimensional Kalman-Bucy filter, Øksendal [46] Theorem 6.16., we have the solution $\widehat{X}_t^{(i)} = \mathbb{E}[X_t | \mathcal{F}_t^{Y_t, Y_t^{(i)}}]$ of this filtering problem.

Let $S(t)$ be the error (co)variance, $S(t) := \mathbb{E}[(X_t - \widehat{X}_t^{(i)})^2]$, then $S(t)$ satisfies the (deterministic) Riccati equation

$$\frac{dS}{dt} = -\frac{1}{2} S^2, \quad S(0) = \theta^2$$

i.e.

$$S(t) = \frac{2\theta^2}{2 + \theta^2 t}, \quad t \geq 0.$$

This gives the stochastic differential equation for the observation $\widehat{X}_t^{(i)}$,

$$d\widehat{X}_t^{(i)} = -\frac{\theta^2}{2 + \theta^2 t} \widehat{X}_t^{(i)} dt + \frac{\theta^2}{2 + \theta^2 t} (dY_t^{(i)} + dY_t), \quad \widehat{X}_0^{(i)} = \mathbb{E} X_0 = \mu \tag{2.30}$$

then

$$d\Big(\widehat{X}_t^{(i)} \exp\Big(\int_0^t \frac{\theta^2}{2+\theta^2 s}ds\Big)\Big) = \exp\Big(\int_0^t \frac{\theta^2}{2+\theta^2 s}ds\Big)\frac{\theta^2}{2+\theta^2 t}(dY_t^{(i)} + dY_t)$$

which gives

$$\widehat{X}_t^{(i)} = \frac{2\mu + \theta^2(Y_t^{(i)} + Y_t)}{2+\theta^2 t}, \quad t \geq 0. \tag{2.31}$$

By the stochastic differential equation for $dY_t^{(i)} + dY_t$, the equation (2.30) for the observation $\widehat{X}_t^{(i)}$ can also be written as

$$d\widehat{X}_t^{(i)} = \frac{\theta^2}{2+\theta^2 t}(dY_t^{(i)} + dY_t - \widehat{X}_t^{(i)}dt) = \frac{\theta^2}{2+\theta^2 t}dN_t^{(i)}, \tag{2.32}$$

where $N_t^{(i)}$ is the *innovation process* defined by

$$dN_t^{(i)} = (X_t - \widehat{X}_t^{(i)})dt + dZ_t^{(i)} + dZ_t. \tag{2.33}$$

And by Øksendal [46] Lemma 6.7., $N_t^{(i)}$ is a Gaussian process with orthogonal increments, and $\mathbb{E}[(N_t^{(i)})^2] = 2t$ (thus $N_t^{(i)}$ scaled by $1/\sqrt{2}$ is a one dimensional Brownian family, as in Karatzas and Shreve [42] Definition 5.8., page 73).

Given player j's strategy κ^j, and fix this strategy, then by Definition 5, the definition of players' strategies, and the footnote on players' actions, k_t^j is progressively measurable with respect to the σ-algebras $\{\mathcal{F}_t^{Y,Y^{(j)}}\}_{t\geq 0}$. Since the processes $Y^{(j)}$ and $Y^{(i)}$ are players' private information and assumed to be independent, player i chooses best responses to $\mathbb{E}[\kappa_t^j|\mathcal{F}_t^{Y,Y^{(i)}}]_{t\geq 0}$.

Define $\widehat{J}_i(\kappa^i, \kappa^j)$ as the following

$$\widehat{J}_i(\kappa^i(\cdot), \kappa^j(\cdot)) = r\int_0^{\tau^i} e^{-rt}(c_i(k_t^i)dt + b_i(k_t^i)(\widehat{X}_t^{(i)} - \mu(k_t)))dt + r\int_0^{\tau^i} e^{-rt} b_i(k_t^i)dZ_t, \tag{2.34}$$

for i, $j \in \{1, 2\}$. And denote the new games with these payoff forms as $\widehat{\Gamma}(s, \widehat{X}_s)_{s\geq 0}$. To show that the two sets of subgame perfect Nash equilibria coincide, it is sufficient to show that player i's best response to $\mathbb{E}[\kappa_t^j|\mathcal{F}_t^{Y,Y^{(i)}}]_{t\geq 0}$ based on $\widehat{X}^{(i)}$ can do no worse than the best response to κ^j based on X. Both k_t^i and $\widehat{X}_t^{(i)} = \mathbb{E}[X_t|\mathcal{F}_t^{Y,Y^{(i)}}]_{t\geq 0}$ are progressively measurable with respect to the σ-algebras $\{\mathcal{F}_t^{Y,Y^{(i)}}\}_{t\geq 0}$. And the drift term of the public signal $X_t - \mu(k_t)$ is in separable linear form. Note also the stationarity of J_i and $\widehat{J}_i$, the claim is true following Fleming and Rishel [38] Theorem 11.1., page 194.

Q.E.D.

2.9.2 Proof of Proposition 3

By Theorem 2, this also implies the set of Nash equilibria of the game $\Gamma(s, X_s)$ is nonempty.

Player i's private monitoring of player j's actions enables player i inferring the value of X_t without the distortion by player j's behavior. Player j's actions are progressively measurable with respect to σ-algebras $\{\mathcal{F}_t^{Y,Y^{(j)}}\}_{t\geq 0}$, but as to player i's best responses per se, player i behaves as if the game is played over the information structure spanned by σ-algebras $\{\mathcal{F}_t^{Y,Y^{(i)}}\}_{t\geq 0}$. So at any moment t, player i's incremental payoff is

$$c_i(k_t^i) + b_i(k_t^i)\big(\widehat{X}_t^{(i)} - \mu(k_t)\big)dt. \tag{2.35}$$

Consider this as the payoff for the stage game, then the strategy for the proof is to show that there exists a profile of modified stochastic processes in $\mathcal{K} = (\mathcal{K}^1, \mathcal{K}^2)$ such that when this profile of strategies is restricted to time t, it is Nash equilibrium for the stage games at time t almost surely.

The action space $\mathcal{A}^i$ is compact, connected in $\mathbb{R}$, then $\mathcal{A} = \mathcal{A}^1 \times \mathcal{A}^2$ is compact, connected in the space $\mathbb{R}^2$. Using the argument in the proof of the existence of Nash equilibrium in normal form games, player i's best responses to player j's action choices is a 1-dimensional manifold in $\mathcal{A}$. Then the Nash equilibrium of the stage game, when payoff is given by equation (2.35), is the intersection of these two manifolds by player 1 and 2. The intersection is either 0-dimensional submanifold (discrete points in $\mathcal{A}$), or 1-dimensional submanifold of these two manifolds parameterized by $\widehat{X}_t^{(1)}$ and $\widehat{X}_t^{(2)}$. By Theorem 1 and the proof of Theorem 2, the two manifolds are parameterized by $Y_t^{(1)} + Y_t$ and $Y_t^{(2)} + Y_t$, which are continuous processes. Then the two manifolds are continuously parameterized by time t.

It is straight forward to show that any sequence of best responses of the corresponding games converges to the best responses. Then the Nash equilibria of stage games parameterized by time t is a 1-dimensional manifold in the space $\mathcal{A} \times [0, \infty)$. Choose the Nash equilibrium obtained by a sequence of games approaching from the right. Since jump points in time t are countable, the strategy constructed is RCLL, which means it is a valid strategy in $\mathcal{K}^i$. (To be made mathematically precise later.)

Q.E.D.

2.9.3 Proof of Proposition 4

Without private monitoring, players can only learn the value of the state process $\{X_t\}$ through the public signal (2.2). However, this public signal is distorted by the other player's actions, and players do not have private signals to correct this distortion. Following the argument in the proof of Corollary 2, players' strategies are defined as in game $\Gamma(s, X_s)_{s\geq 0}$ but the noise in the private signals have infinite variance. So the strategies are perfect public, and in any equilibrium both players' strategies are known to each other although they are not directly observable.

To show that players are playing instant stage game Nash equilibrium. Suppose this is not true and there exists a time interval $[t, t+dt)$, such that players are *not* playing stage game Nash equilibrium. This time interval exists because strategies are piecewise continuous, and the measure of this interval is strictly positive. On this time interval, player i is not playing best response to player j's actions, for i, $j \in \{1, 2\}$.

Let $\kappa_t^* = (\kappa_t^{i*}, \kappa_t^{j*})$ denote the original strategies. On the interval $[t, t+dt)$, let $\kappa_t^\diamond = (\kappa_t^{i\diamond}, \kappa_t^{j\diamond})$ denote consistent playing of stage game Nash equilibrium. $\kappa_t^* \neq \kappa_t^\diamond$ on $[t, t+dt)$.

Define the process $Y_t^\flat$

$$dY_t^\flat = X_t dt + dZ_t, \quad Y_0^\flat = 0. \tag{2.36}$$

By the Kalman-Bucy filter model (see Øksendal [46] Theorem 6.10. and Example 6.11.), and let $S(t)$ be the error (co)variance, $S(t) = \mathbb{E}[(X_t - \widehat{X}_t)^2]$, then $S(t)$ satisfies the (deterministic) Riccati equation

$$\frac{dS}{dt} = -S^2, \quad S(0) = \theta^2$$

i.e.

$$S(t) = \frac{\theta^2}{1+\theta^2 t}, \quad t \geq 0.$$

This gives the stochastic differential equation for the observation $\widehat{X}_t$,

$$d\widehat{X}_t = -\frac{\theta^2}{1+\theta^2 t}\widehat{X}_t dt + \frac{\theta^2}{1+\theta^2 t} dY_t^\flat, \quad \widehat{X}_0 = \mathbb{E}X_0 = \mu \tag{2.37}$$

then

$$d\left(\widehat{X}_t \exp\left(\int_0^t \frac{\theta^2}{1+\theta^2 s} ds\right)\right) = \exp\left(\int_0^t \frac{\theta^2}{1+\theta^2 s} ds\right)\frac{\theta^2}{1+\theta^2 t} dY_t^\flat$$

which gives

$$\widehat{X}_t = \frac{\mu + \theta^2 Y_t^\flat}{1+\theta^2 t}, \quad t \geq 0. \tag{2.38}$$

By equation (2.36), the stochastic differential equation (2.37) for the observation $\widehat{X}_t$ can also be written as

$$d\widehat{X}_t = \frac{\theta^2}{1+\theta^2 t}(dY_t^\flat - \widehat{X}_t dt) = \frac{\theta^2}{1+\theta^2 t} dN_t, \tag{2.39}$$

where N_t is the *innovation process* defined by

$$dN_t = (X_t - \widehat{X}_t)dt + dZ_t. \tag{2.40}$$

And by Øksendal [46] Lemma 6.7., N_t is a Gaussian process with orthogonal increments, and $\mathbb{E}[N_t^2] = t$ (thus N_t is a Brownian family).

Define the process $Y_t^\dagger$ as the solution of the following ordinary differential equation

$$dY_t^\dagger = -\mu(k_t^i, k_t^j)dt, \quad Y_0^\dagger = 0. \tag{2.41}$$

Then

$$Y_t = Y_t^\flat + Y_t^\dagger. \tag{2.42}$$

Player i's continuation value in expectation depends on her perceived state parameter and both players' strategies, $W_t^i(\widehat{X}_t^i, \kappa_t^i, \kappa_t^j)$. If in the original strategies $\kappa_t^* = (\kappa_t^{i*}, \kappa_t^{j*})$, the realized players' actions are not mutually best responses as in stage games on the time interval $[t, t+dt)$, the possible benefits from deviation are two folds: the direct benefit is the gain from short period cheating behavior, and the indirect benefit is the distortion of the opponent's belief on X_t.

Equally dividing $[t, t+dt)$ into L subintervals, and suppose the players can only play constant actions k_t on each subinterval. Further assume these actions are right continuous and have left limits (RCLL), so the strategies with the realized actions in this form are in $\mathcal{K}^i$. The next step is to approximate the possible nontrivial deviations with these strategies.

Q.E.D.

2.9.4 Proof of Proposition 9

Note that

$$Y_t^\flat = Y_t - Y_t^\dagger = Y_t + \int_0^t (\kappa_s^1 + \kappa_s^2) ds = Y_t + 2\int_0^t \kappa_s^i ds,$$

since the investors' strategies are symmetric in the common learning case.

Then from equation (2.21),

$$\widehat{X}_t = \frac{\sigma^2\mu + \theta^2 Y_t^\flat}{\sigma^2 + \theta^2 t} = \frac{\sigma^2\mu + \theta^2 Y_t + 2\theta^2 \int_0^t \kappa_s^i ds}{\sigma^2 + \theta^2 t}.$$

By Lemma **??** and Proposition 8, $\kappa_t^i = (1/3)\widehat{X}_t$, so

$$3(\sigma^2 + \theta^2 t)\kappa_t^i = \sigma^2\mu + \theta^2 Y_t + 2\theta^2 \int_0^t \kappa_s^i ds.$$

Differentiating both sides,

$$3\theta^2\kappa_t^i dt + 3(\sigma^2 + \theta^2 t) d\kappa_t^i = 2\theta^2\kappa_t^i dt + \theta^2 dY_t;$$

or in the stochastic differential equation form

$$d\kappa_t^i = -\frac{\theta^2}{3(\sigma^2 + \theta^2 t)}\kappa_t^i dt + \frac{\theta^2}{3(\sigma^2 + \theta^2 t)} dY_t, \quad \kappa_0^i = \frac{1}{3}\mu. \tag{2.43}$$

Using the same method for the closed form solution of $\widehat{X}_t$,

$$d\left(\kappa_t^i \exp\left(\int_0^t \frac{\theta^2}{3(\sigma^2 + \theta^2 s)} ds\right)\right) = \exp\left(\int_0^t \frac{\theta^2}{3(\sigma^2 + \theta^2 s)} ds\right) \frac{\theta^2}{3(\sigma^2 + \theta^2 t)} dY_t$$

which gives

$$d\Big(\kappa_t^i(\sigma^2+\theta^2 t)^{\frac{1}{3}}\Big) = \frac{\theta^2}{3(\sigma^2+\theta^2 t)^{\frac{2}{3}}}dY_t,$$

and

$$\kappa_t^i = \frac{1}{3}\mu\sigma^{\frac{2}{3}}(\sigma^2+\theta^2 t)^{-\frac{1}{3}} + \int_0^t \Big(\frac{1}{3}\theta^2(\sigma^2+\theta^2 t)^{-\frac{1}{3}}(\sigma^2+\theta^2 s)^{-\frac{2}{3}}\Big)dY_s, \quad \kappa_0^i = \frac{1}{3}\mu. \tag{2.44}$$

This finishes the proof.

Q.E.D.

Chapter 3

Financial Contract Design and Staging in Venture Capital

3.1 Introduction

Financial contracts play a key role in coordinating investment behaviors by venture capitalists and entrepreneurs in venture capital backed companies (Kaplan and Strömberg [66], and [67]). In practice, staged financing is widely used by venture capitalists in their investments (Sahlman [31], Gompers [59], and Lerner [69]). What's the connection between this financing procedure and the associated financial contracts? The paper shows that, venture capitalists create option value in corporate decision making by financial contract design, and the implementation of these contracts leads to the staging of venture capital.

Venture capital investment processes involve private contracting and intense negotiations between the investors and the entrepreneurs. Both the quality of the project and the ability of the entrepreneur are vital for the success of the venture. In flexible contracts, such as, a short term open-ended financial contract, in which certain clauses are excluded from the contract and left for future negotiations, the venture capitalist may not be able to protect the previous investments due to hold-up by the entrepreneur in later round negotiations. In rigid contracts, such as, a long term contingent contract, the venture capitalist may not have the flexibilities in important corporate decision making. The paper will show that the optimal contract – the short term open-ended financial contract – creates option value for the venture capitalist which dominates the cost of later stage negotiations. Moreover, strategic allocation of investments at different stages of financing will reduce the negotiation cost.

When the venture seems promising in going public and the entrepreneur appears competent, the entrepreneur becomes a scarce human resource and will have bargaining power over sharing surplus with the venture capitalist. This is a typical situation where the entrepreneur can "hold-up" the venture capitalist after the initial round of the investment. It is possible that a long term contingent contract with lump sum capital infusion will mitigate this agency problem, but under some investment conditions with complex information structure, this contractual form may be suboptimal.

The venture capitalist can gather information about the prospect of the venture during an investment process. The information about the feasibility of the innovation (or the technology), and about the managerial ability of the entrepreneur, etc., is naturally multi-dimensional. Anticipating the information update, the venture capitalist might have a list of alternatives:

abandoning the venture, replacing the entrepreneur by a professional manager, for instance. But, for some of the contingencies, it may be either unlawful to write them in contracts, or difficult to describe *ex ante* and difficult to prove their occurrence in court *ex post*. When these noncontractible contingencies exist, it may be optimal to exclude other contingencies and use short term contracts with full expectation of future negotiations. Moreover, if the ability to choose these alternatives in the future has option value, a short term noncontingent (open-ended) contract will materialize this value by letting the venture capitalist decide when to exercise this option.

This contractual solution is indebted to the "conventional wisdom" that given the contracts are incomplete (partially), it is optimal to choose entirely noncontingent contracts. This question is addressed in the study of employment contracts. In the multitask principal-agent problems studied by Holmstrom and Milgrom [65], when the principal has either several independent tasks or a single task with multi-dimensional aspects for the agent to perform, the principal often will pay fixed wages although objective output measures are available and the agent is responsive to incentive pay. I will extend this theory to financial contracting in this paper.

Short term contract also gives the venture capitalist the option to efficiently adjust venture ownership structures. On one hand, the venture capitalist values ownership because the rights of corporate decision making are embedded in the owners' rights, and the ownership gives the venture capitalist bargaining power in possible future negotiations. On the other hand, ownership functions as an incentive for the entrepreneur to exert effort. Also, at the early stage of financing, acquiring information about the prospect of the venture is more critical than providing the entrepreneur incentives. Considering these factors, the venture capitalist would choose to retain ownership in the beginning, and later decide whether to transfer it to the entrepreneur according to future situations of the investment.

The model in this paper is close to Che and Sákovics [56]'s dynamic theory of hold-up. Che and Sákovics [56] develops a dynamic model of investment and bargaining, in which both parties can continue to invest if agreement is not reached in the previous negotiation. As an extension, my model incorporates investments, negotiations, and contracting. The venture capitalist chooses the contractual form before investment, and both the venture capitalist and the entrepreneur have to decide how to invest with intertwined negotiations and information arrivals.

Specifically, the extensive form of the model has the following structure. A venture capitalist and an entrepreneur together start a new venture, and intend to launch IPO eventually. The venture capitalist supplies capital investment, and the entrepreneur exerts effort. The quality of the technology owned by the entrepreneur, and the entrepreneur's managerial ability are uncertain but decisive for a successful IPO. Information partially resolving these uncertainties will be available during the investment process. The venture capitalist chooses the contractual form, determines whether to negotiate, decides how to invest, and the entrepreneur solves how to exert effort both before and after the information arrivals. At the last stage, an exogenously given investment bank examines both the technology and the manager, and announces whether the venture is qualified for an IPO.

This paper takes the financial contract design approach to explain investment behaviors in venture financing, which has been shown to be a powerful tool by Hellmann [61]. In explaining how convertible securities can be used to settle disagreement between the venture capitalist and the entrepreneur on the timing of exit, Hellmann [61] first finds the optimal contracts, then shows how venture capitalists can use convertible securities to implement these contracts. My paper is an application of this methodology.

The rest of the paper is organized as follows. Section 3.2 reviews the related literature and discusses the differences between the existing literature and this paper. Section 3.3 introduces the model. Section 3.4 outlines the contracting possibilities. Section 3.5 studies investment behaviors and related inefficiencies. Although this is the benchmark model, it provides two thresholds for investment decisions in the cases of short term and long term contracts. Section 3.6 studies the case where both the feasibility of the technology and the entrepreneur's managerial ability are uncertain, and explains why short term open-ended contracts are optimal despite these inefficiencies. This section establishes the main results. Section 3.7, 3.8, and 3.9 provide an example, empirical predictions, and conclusion, respectively.

3.2 Related Literature

This paper studies the situation in which both agency problem and option value coexist. Neher [70] provides an explanation for staged financing from the agency perspective. The venture capitalist divides the total investments into consecutive rounds, so that the investment of inalienable human capital by the entrepreneur in the previous financing round can be used as collateral for the following round, and this mitigates the hold-up problem by the entrepreneur. The key difference between Neher [70] and this paper is that, in this paper, the venture capitalist trades off protections from hold-up for the options in corporate decision making. This is motivated by the empirical observation that, in some software and pharmaceutical companies backed by venture capital, the most valuable assets are human capital, and the companies have little liquidation value even at the time of IPO.

The theory presented in this paper is related to the "real option" model on investment under uncertainty. The real option view is effective in evaluating the situations of sequential information revelation, especially in corporate R&D projects. Different from these situations, several factors contribute to a successful venture capital investment (Lerner [68]), and new information about these factors may arrive simultaneously, for example, information about the managerial ability of the entrepreneur and results of clinical trials. The model predicts upward distortion in the initial rounds of investments, which is consistent with recent empirical findings by Puri and Zarutskie [71]. This phenomenon cannot be explained by the real option model.

The paper assumes symmetric information based on the following reasons: first, if the entrepreneur has information advantage over the venture capitalist, the Revelation Principle suggests that truth telling mechanisms can be designed to reduce this asymmetry. In Cornelli and Yosha [57], convertible securities can be designed to prevent window dressing behaviors by the entrepreneur. Second, although staged financing facilitates information acquisition, the ultimate goal of the venture capitalist is to make appropriate corporate decisions after acquiring new information. The paper is a complement to the research by Gompers [59], Admati and Pfleiderer [51], Bolton and Scharfstein [54]. Gompers [59] examines how asymmetric information affects the structure of staged venture capital investments. Admati and Pfleiderer [51], Bolton and Scharfstein [54] also investigate information asymmetries and financial contracting in the financing of an entrepreneurial venture, but staging is exogenous in their models.

Although there may not be technical short term and long term contracts in real world venture financing, the purpose of using these concepts is to emphasize the scope of the investment horizons. [1]. The analysis in this paper lays out the theoretical foundations for the design of term sheets and the valuation of each class of private equities. This is a connection between economic theory and real world practice as addressed by Kaplan and Strömberg [66], [67].

This approach inevitably raises the question on the optimality between short term and long term contracts. Fudenberg, Holmstrom, and Milgrom [58] studies the sufficiency of short term contracts and provides several prerequisite conditions, including availability of public information for contracting and equal borrowing capabilities. My paper departs from their research since most of their assumptions are violated in venture financing and bargaining power might shift from the venture capitalist to the entrepreneur under some circumstances.

The short term open-ended contract in this paper is different from those in the relational contracting literature. In relational contracts (Baker, Gibbons, and Murphy [52]), economic behaviors based on informal agreements

[1] I thank professor Josh Lerner for pointing this out.

are enforced because of the expectations of future relationships. In venture financing, although the presence of short term contracts does not preclude the possibility of future interactions between the venture capitalist and the entrepreneur, both parties have the freedom to exit the investment unilaterally. The situation becomes severe when the limited liability constraint of the entrepreneur is binding.

The paper is closely related to Chan, Siegel, and Thakor [55] in modeling the learning process of the entrepreneur's ability. When the signal indicates that the entrepreneur is less competent and the venture has little probability of going public with the entrepreneur as the manager, the venture capitalist considers operating the venture with a professional manager. Hermalin and Weisbach [63] studies CEO replacement as part of the negotiation process between the CEO and the board. But, in venture financing, the venture capitalist decides whether to keep the entrepreneur before they negotiate over sharing of surplus from a successful IPO. If the venture capitalist decides to keep the entrepreneur as the manager, the entrepreneur becomes a scarce human resource and will have bargaining power. This negotiation process is modeled after Binmore, Rubinstein, and Wolinsky [53]'s bargaining games.

Hellmann [60] studies the allocations of control rights in venture financing. His paper argues that the entrepreneur voluntarily relinquishes control rights to the venture capitalist so that the venture capitalist will have incentive to search for a better management team. My model focuses on a different aspect of management replacing – negotiations – in a dynamic setting, because, in venture financing, allocations of control rights are not always clear-cut: conflicts of interests are often settled through negotiations.

3.3 The Model

3.3.1 Model Description

In this section, I describe the setting of the model. Then in section 3.3.2, I will outline the extensive form of the game.

Consider the relationship between a venture capitalist (as "he" solely for model description convenience), denoted by V, and an entrepreneur (as "she"), denoted by E. Both V and E are assumed to be risk neutral and there is no discounting. E is penniless and has limited liability, but she has a technology or a business idea and wants to start a new venture. V is wealthy and looking for an investment opportunity. V will receive his final payoffs through dividend, ownership of the venture, or proceeds from liquidation of the sunk capital investment; while E will receive her final payoffs through wage and the venture ownership. Either the dividend or the wage will be paid regardless the outcome of the investment, and they are pecuniary transfer between V and E. The payoffs in the form of ownership can only be realized in the case of a successful IPO.

It is assumed that V has all the bargaining power in the beginning and makes take-it-or-leave-it offer to E. Throughout the model, I treat "venture" and "company" as two equivalent terms and use them interchangeably.

Once the company is started, its quality depends on both the quality of E's technology, and the managerial ability of E. These two factors are uncertain to both V and E. They need to start the company and invest to resolve these uncertainties. So the investment is a learning process. The uncertainties are modeled as the following.

The prior distribution of the quality of E's technology, type δ_T, is normal with mean zero and variance $1/h_T$ (h_T is the precision of the distribution). The prior distribution of the managerial ability of E, type δ_E, is normal with mean zero and variance $1/h_E$. Both of these two distributions are common knowledge to V and E. I follow Holmstrom [64], Hermalin and Weisbach [63] by assuming that E knows only the distribution of her ability. The reason for this assumption is that entrepreneurs in venture capital backed young, start-up companies generally have limited experience as managers.

The arguments throughout this paper are valid without the normal distribution assumption as long as Bayesian information updating is applicable for their underlying probability distributions. For simplicity, I assume the distributions for δ_T and δ_E are independently distributed. The assumption that the quality of the technology and E's ability are independent from each other is without loss of generality. For instance, the market demand for an invention that greatly improves fuel efficiency depends on world oil prices. It is very unlikely that the inventor's ability of managing a small company is correlated with fluctuations of world oil prices.

In combination, the quality of the whole venture, type δ_C, is defined as $\delta_C = \min\{\delta_T, \delta_E\}$. The true value of δ_C will not be revealed until the end of the game, and it will be revealed by an exogenously given underwriter, an investment bank. In general, one can assume that δ_C is a given function of both δ_T and δ_E. It is also possible that the quality of the technology and the ability of the manager may affect the venture success in many different ways. For example, both factors are vital for a success in high-tech and bio-tech industries. But in fast food and service industries, the two factors might be substitute to each other, since in these industries advanced innovations are less important and high quality management can certainly lead to high venture valuation despite mediocre underlying business ideas (such as changing the size of hamburgers from regular to bite-size). It will become obvious later that the assumption $\delta_C = \min\{\delta_T, \delta_E\}$ greatly simplifies the calculation and makes the model tractable.

There is an exogenously given threshold δ^*, such that this venture capital backed company is qualified for IPO if and only if $\delta_C > \delta^*$. The economic interpretation of δ^* is very rich. This threshold might be lower when macro economy or a particular industrial sector is in boom, and higher in downturns, because demand for new technologies varies with many economic factors, which are beyond the control of venture capitalists and entrepreneurs. δ^* might be different for different industries, since my model normalizes the means of δ_T and δ_E to zero.

The intuition of separating quality of a venture into a combination of quality of technologies and ability of managers is based on empirical observations in venture financing. Venture capital backed Federal Express Corporation pre-IPO history is a good illustration (Gompers [59]). The company was built around an innovative concept of package distribution system, but the company performed well below expectations initially, until the venture capitalists intervened extensively in its management. Eventually, Federal Express Corporation went public in 1978.

δ_C, δ_T, and δ_E together characterize the information structure faced by V, E, and later professional managers and investment banks in this investment process.

The value of the venture depends on V's investment, E or her replacement's effort (to be specified below), and the company's type δ_C. V can invest k on critical physical assets at any time and in multiple times before any exit decision on IPO or liquidation. The capital k contributes to the value of the company a factor $Q(k)$. The investments are cumulative, in the sense that if V invests k and then k', the factor will be $Q(k+k')$. Assume that the investment is sunk and V cannot disinvest the existing capital, except for a liquidation at loss.

ASSUMPTION 1: The factor $Q(k)$, which V's capital investment k contributes to the value of the company, satisfies $Q(0) = 0$, $\lim_{k\to 0+} Q'(k) = +\infty$, $Q'(\cdot) > 0$, $\lim_{k\to +\infty} Q'(k) = 0$ and $Q''(\cdot) < 0$.

The assumption that $\lim_{k\to 0+} Q'(k) = +\infty$ will simplify proofs. All claims will remain the same as long as $\lim_{k\to 0+} Q'(k)$ is sufficiently large.

It is common knowledge that V and E can together receive a public signal, x, about the quality of the technology. The signal x is verifiable and contractible. x is normally distributed with a mean equal to the technology's true quality, type δ_T, and a variance equal to $1/h_x$. In the meantime, V and E also acquire a public signal, y, about the ability of E. y is nonverifiable in court and cannot be written in contracts. Assume y is normally distributed with a mean equal to E's true ability, type δ_E, and a variance equal to $1/h_y$. Assume that the random variables $x - \delta_T$ and $y - \delta_E$ are independently distributed[2].

[2] I follow Hermalin and Weisbach [63]'s approach in modeling how players update their beliefs about the technology and E's ability after new information is observed. Note that both V and E can observe the realization of the signal y, but this state variable is not contractible *ex ante*, and its realization cannot be verified in court *ex post*.

The signal x for the technology is contractible, because personal beliefs of the quality of the technology lie in the category of objective assessment. In practice, granted patent, FDA approval, report of marketing research are all verifiable in court and noisily indicate future prospect of the technology. However, the judgement wether E is a competent manager lies in the category of subjective assessment, which is often distorted by personal biases. So court might not accept y as valid argument for E's ability.

Assume V can replace E with a professional manager, denoted by M (as "she"). The prior distribution of the managerial ability of M, type δ_M, is also normal with mean zero and variance $1/h_M$. The distributions of the abilities of the professional manager and the entrepreneur are identical and independent, $h_M = h_E$. This is a strong assumption, since professional managers in general are experienced corporate veterans, comparing to entrepreneurs who might have little track record in managing medium or large companies. By this assumption, V is indifferent to who manages the venture in the beginning until new information about E's ability arrives. Following the classic career concern model, I assume both E and M's abilities are fixed throughout their career.

Both E and M can contribute effort in addition to the company's quality δ_C by a factor $\mu(e)$, at cost e. The factor $\mu(e)$ is deterministic. E and M's effort e are homogeneous and cumulative, so that if the effort inputs are e and then e', together their contribution to the company's valuation is $\mu(e + e')$. Assume that E cannot input negative effort, or in other words, sabotage.

ASSUMPTION 2: The productivity of E and M, $\mu(\cdot)$, satisfies $\mu(0) = 0$, $\mu'(\cdot) > 0$, $\lim_{e\to+\infty} \mu'(e) = 0$ and $\mu''(\cdot) < 0$.

Now I can define the value of the company $v(k, e)$.

ASSUMPTION 3: In a successful IPO after V has invested capital k, and E, M have invested effort e, the company's market value $v(k, e)$ is $Q(k)(1 + \mu(e))$. The company's market value is 0 in case of a failed IPO.

Assumption 3 says the public market represented by an investment bank will reveal the venture's true quality δ_C. The assumption that the company has market value 0 when it withdraws from IPO seems extreme. But in the venture financing context, it has the following reasons. First, a venture capital backed company that fails to go public after five years of operation is generally mediocre, and it barely generates enough cash flow to compensate venture capitalists outside opportunity costs. Second, withdrawal from IPO by a young company causes severe reputation damage. Third, since venture capital funds are closed-end, and venture capitalists as fund managers share the proceeds with fund contributors, a portfolio company which remains private negatively affects the calculation of fund returns.

Note that the company's value v is super-modular: $v_{ke}(k, e) \geq 0$ for all $(k, e) \geq (0, 0)$, which means that V's investments and E's effort are weak complements, globally. v has two components: $Q(k)$ and $1 + \mu(e)$. The assumption that $Q(0) = 0$ causes $v(0, \cdot) = 0$ reflects the fact that V's initial investment is necessary to start a new company. In the mean time, the quality of the technology, E's ability, both E and M's effort play a value enhancing role. The form $1 + \mu(e)$ is the counterpart of log growth rate in accounting literature.

A caveat is that the signals for information update on the uncertainties over the technology and the manager's ability are separated from the managers' efforts in this model. This is different from Hermalin and Katz [62], in which signals indicate the agent's effort input level in a moral hazard problem.

ASSUMPTION 4: The company's liquidation value is determined by a factor q. And this value is qk if the sunk capital investment level is $k \geq 0$. The liquidation process is irreversible, which means once the company is ceased from operation, it cannot be re-opened. Assume that $q < \min\{\inf_k Q'(k), 1\}$, for all k in feasible investment region.

The IPO process is modeled as follows. At the end of the game, an investment bank conducts evaluation of the company and compares the true value of δ_C with δ^*[3]. The investment bank informs the company whether it is qualified for IPO. The company's value is $Q(k)(1+\mu(e))$ if IPO is successful. Otherwise, the company remains private.

Only when V with full ownership of the company can decide to liquidate the venture, and V retains the company's entire liquidation value qk. E has no ability to liquidate the company. This is based on empirical observations that the venture capitalists usually have strong social networks which help them to recover the past investments to some extent. While this is a disadvantage of the entrepreneurs who might only have technologies, inventions, or simply business ideas. Full ownership of the company in this model is in the general sense. It not only has the meaning that sole owner of a property can lawfully liquidate this asset, but also means that the company's board of directors can vote for liquidation according to corporate bylaws.

[3]Alternatively, I could assume that E and M's effort improves the company's probability of going public, by comparing $\delta_C + \mu(e)$ and δ^*, or E and M's effort shortens the pre-IPO period. These modeling alternatives are mathematically equivalent.

3.3.2 Model Outline

The game has multiple stages with the following timing (see Figure 3.1).

1. At the start of the game, V chooses a format of financial contract and signs an initial contract with E. The contract specifies the amount of E's compensation ω_E payable upon termination of E as a manager. The contract also describes the ownership structure α_V, α_E, and its possible future variation.

2. V invests capital k_0 and E exerts effort e_0.

3. The realization of signal x for the technology δ_T occurs. V and E observe a nonverifiable signal y of E's ability δ_E.

4. The ownership structure is determined according to the initial contract. If liquidation is chosen by V, no further contract is necessary. If further investment is chosen, V decides whether to keep E or to replace E by M.

If V decides to continue investing with E as manager, then

5E. V and E renegotiate the existing contract or sign a new contract, that specifies possibly new compensation for E.

6E. V (weakly) increases the capital investment to a new level k_1, and E (weakly) increases the effort input to a new level e_1.

7E. The true value of δ_C is revealed by an investment bank. If the company is qualified for IPO, the payoff is distributed according to the effective contract. Otherwise, the company is terminated with value 0.

If V decides to replace E with a professional manager M, then

5M. E leaves with severance payment. V and M sign a contract, that specifies M's compensation.

6M. V (weakly) increases the capital investment to a new level k_1, and M (weakly) increases the effort input to a new level e_1.

7M. The true value of δ_C is revealed by an investment bank. If the company is qualified for IPO, the payoff is distributed between V and M according to the effective contract. Otherwise, the company is terminated with value 0.

3.3.3 Information Updating

Given the structure of the quality of the venture, it is easy to see that

$$\Pr(\delta_C > \delta^*) \;=\; \Pr(\delta_T > \delta^*) \cdot \Pr(\delta_E > \delta^*).$$

Let p denote the value above, which is the prior probability of the company going public. Define p_T, p_E, and p_M as the prior probabilities of each type being above the threshold δ^*:

$$\begin{aligned} p_T &= \Pr(\delta_T > \delta^*), \\ p_E &= \Pr(\delta_E > \delta^*), \\ p_M &= \Pr(\delta_M > \delta^*). \end{aligned}$$

Since δ_T and δ_E are normally distributed, and the two random variables associated with the signals x, y: $x-\delta_T$ and $y-\delta_E$, are by assumption independently distributed, the posterior estimation δ'_T of δ_T is a normal distribution with mean

$$\hat{\delta}_T = \frac{0h_T + xh_x}{h_T + h_x} = \frac{xh_x}{h_T + h_x},$$

and precision

$$\hat{h}_T = h_T + h_x.$$

And the posterior estimation δ'_E of δ_E is a normal distribution with mean

$$\hat{\delta}_E = \frac{0h_E + yh_y}{h_E + h_y} = \frac{yh_y}{h_E + h_y},$$

and precision

$$\hat{h}_E = h_E + h_y.$$

Thus $\delta'_C = \min\{\delta'_T, \delta'_E\}$ if E is the manager, or $\delta'_C = \min\{\delta'_T, \delta_M\}$ if M is the manager.

Define $\hat{p}_T$ and $\hat{p}_E$ as the posterior probabilities of each type being above the threshold δ^*:

$$\begin{aligned} \hat{p}_T &= \Pr(\delta'_T > \delta^*), \\ \hat{p}_E &= \Pr(\delta'_E > \delta^*). \end{aligned}$$

After observation of signals x and y, if E continues to be the manager, then the company will eventually go public with probability

$$p \;:=\; \Pr(\delta'_C > \delta^*) \;=\; \Pr(\delta'_T > \delta^*) \cdot \Pr(\delta'_E > \delta^*) \;=\; \hat{p}_T \hat{p}_E.$$

If V successfully replaces E by M, then the company will eventually go public with probability

$$p \;:=\; \Pr(\delta'_C > \delta^*) \;=\; \Pr(\delta'_T > \delta^*) \cdot \Pr(\delta_M > \delta^*) \;=\; \hat{p}_T p_M.$$

This information structure tells us that the signal x is *not* informative on when V should fire E. However, realization of signal x affects the posterior probability of the quality of the technology being above the threshold δ^*, and this probability in turn affects both V and E's investment behaviors.

After observation of signal y, the ability of E becomes less uncertain, so the threshold for signal y of firing or keeping E will be above the expected ability of M, which is 0 by assumption. This observation is given in the following lemma.

Lemma 1. *When the technology and the entrepreneur's ability are both uncertain, for any given investment level k, effort level e, any realization of signal x and existing initial contract in any form, there exists a threshold $y^* > 0$ such that if $y \leq y^*$, the venture has less probability going public when the entrepreneur is the manager instead of a professional manager.*

Because of the information structure, there are one to one correspondences between signals and the probabilities of going public. Let ϕ, Φ denote the probability density function and probability distribution function of the standard normal random variable.

Lemma 2. *The probability $\hat{p}_T$ is a smooth, strictly increasing function of the signal x. The prior probability distribution of x before stage 3 induces a prior* *probability distribution of $\hat{p}_T$. The same is true for signal y and probability $\hat{p}_E$.*

Let ρ_T, ρ_E denote the probability density functions for $\hat{p}_T$, $\hat{p}_E$, respectively.

By lemma 2, studying how signals affect investments is reduced to studying how probabilities of a successful IPO affects investments. Conversely, strategies based on probabilities can be easily transformed into strategies based on signals.

3.3.4 Manager Replacing

Different from publicly traded companies in which monitoring and influence to corporate decision making by shareholders are collective (and often ineffective), also different from family owned companies in which owners, managers are bound by blood and marriages, majority of venture capital backed companies are financed by issuing private equities to a small group of venture capitalists. In these companies, ownership structures are written in legal documents one way or another to avoid future power struggles when conflicts of interests occur. In the language of contract theory, co-ownership of critical physical assets by two economic agents is an extreme form of long term contract between these two agents. The duration of the clause on ownership is indefinite until one or both of them decide to terminate this economic relationship, or to replace the existing contract by a (weakly) Pareto improved new ownership structure.

If E's ability of being a corporate manager is uncertain, the venture capitalist V wish to be able to replace E when it is optimal for him to do so instead of status quo based on new information. In the situation that V possesses full ownership of the company, the contract between V and E is essentially an employer-employee contract, in which V provides financing and E contributes human capital. Further more, V has contractual and legal rights[4] to exclude E from operating and managing the company if V is the sole owner of the company.

However, the mechanism of venture capital investment procedures complicates the story. Since venture capital funds are closed-end, V as a fund manager will not stay with the portfolio company forever, and has to unload this ownership sometime in the future after the initial investment. Additionally, entrepreneurs play an important role in young, start-up companies, so V needs to provide E incentives for effort. The widely used solutions are vesting schedules which grant entrepreneurs restricted stocks[5]. This can be considered as a contractual solution of defining and transforming ownership structures in private companies financed by venture capital.

Then what would happen if V and E share the ownership of the company? The situations in which V and E co-own a private company is different from those between shareholders of a publicly traded company. The decision

[4] An example of contractual rights is a clause explicitly written in the contract which says the venture capitalist is able to remove the entrepreneur from the manager position unilaterally, given the occurrence or non-occurrence of some pre-specified events. An example of legal rights is that the venture capitalist has full control of the corporate board and is able to vote against the entrepreneur according to corporate bylaws.

[5] A caveat is that the venture capital fund itself may be publicly traded: the venture capital funds organized by master limited partnership. Divisions of publicly traded companies may also be dedicated to venture capital investments: IBM's Venture Capital Group, and Intel Capital. But these are quite different from the private ownership of portfolio companies.

making process is closer to a negotiation process than a simple majority voting process, since neither can V force E to contribute effort, nor can E force V to invest. When the technology seems promising, or E appears to be a manager with high ability, E will have bargaining power in negotiations with V on sharing surplus, because either E owns the technology (for example, patents will only be granted to inventors not investors in U.S.) or E becomes scarce human resource. But, when it turns out that E has low ability, V can no longer freely fire E since E herself is also an owner of the company. In this case, it is costly for V to replace E by a professional manager M.

In designing venture financing contracts, V balances the tradeoffs between providing E incentives and having effective rights of replacing E when necessary. V's main challenge is to decide what will be written explicitly in the contract, and what will be excluded from the contract on purpose and kept for future negotiations. The next step of the paper is to show the contractual solutions under different situations in venture financing when moral hazard and multidimensional uncertainties coexist.

3.4 Contracting Possibilities

In order for the financial contracts to coordinate the investment behaviors by V, E, and M, there are two basic questions which need to be answered: what will be written in the contract and the duration of the contract. The latter is equivalent to the choice between short term contracts and long term contracts, since the initial decision to select short term contracts with negotiations will either lead to an investment process governed by a sequence of short term contracts, or results in an early termination of the investment.

Since E and M's effort is non-contractible, neither can the venture's valuation $v(k, e) = Q(k)(1 + \mu(e))$ be written in the contract, V considers the signal x and the event of IPO in this contract design problem. V compensates E and M's effort by granting them ownership of the venture. This is due to the fact that signal x based pecuniary compensation is futile in providing E incentives, because E's effort cannot affect the realization of the signal x. This form of compensation corresponds to the widely adopted practice that the venture capitalists grant the entrepreneurs restricted stocks through a variety of vesting schedules. The monetary transfer is either in the form of wage paid to E and M by V, or in the form of dividend paid to V by E and M, so that E and M's individual rationality and limited liability constraints are satisfied.

Ownership structures affect management replacing decisions. When V is the sole owner of the company, the contract between V and E is essentially an employment contract. E does not have unfair dismissal rights in this case, so it is less costly for V to replace E by M, but E won't have incentive to exert effort. When V and E share the ownership of the company, replacing E by M is costly for V. It will be in the form of severance payment.

Let α_V, α_E, $\alpha_M \in [0, 1]$ denote V, E, M's proportion of the ownership of the company. Let ω_E and ω_M be the monetary transfer from V to E and M. ω_E, ω_M are wages if they are greater or equal to zero, and they are dividends paid to V if less than zero. And let s be the severance payment from V to E.

Different from existing contract theory literature, the challenges to V

are not only to decide the compensations and ownership structures based upon verifiable signals and events, but also to decide what will be written in contract explicitly and what will be excluded from contract for future negotiations. When there is no binding contract clause on some particular subspaces of the strategy spaces of V, E, or M, this paper uses subgame perfect equilibrium (SPE) as the solution concept. To be more specific, I will consider SPE in Markov strategies – Markov perfect equilibria (MPE).

3.5 Investment Behaviors and Related Inefficiencies

First consider the case of one dimension uncertainty about the technology, and the signal x for the technology is contractible. Suppose the ability of E is certain and is common knowledge, so either $p_E = 0$ or $p_E = 1$. In the former, V has no incentive to invest. So let us look at the interesting case where $p_E = 1$. V and E now concern about whether the type of the technology is above the threshold, $\delta_T > \delta^*$. Denote its probability as $\hat{p}_T$, and its prior distribution is given by ρ_T in lemma 2. The time line of the game follows stage 1, 2, 3, 5E, 6E, and 7E (see Figure 3.2).

At stage 1, V chooses between short term contract and long term contract. The usage of short term contract is to govern the initial investment process by a contract whose duration is up until the realization of signal x, then V, E negotiate a new contract if both sides decide to continue the venture. A long term contract is a contract signed at stage 1, which governs the entire investment process until the IPO stage.

Note that the state space of the signal x is perfectly foreseeable and the signal itself is contractible *ex ante* at stage 1. Intuitively, the optimal long term contract should be a contingent contract based on signal x, and V extracts all the surplus. E accepts the contract in the beginning as long as the individual rationality and limited liability constraints are satisfied.

The search for optimal short term contracts can be solved by backward induction. When information arrived at stage 3 reveals that the technology is promising, then at stage 5E, E will have bargaining power in the negotiation of the new contract on sharing surplus with V, since E owns the technology. Foreseeing this, V requires higher signal x to compensate the loss of surplus. The following propositions will rigorously prove these observations. An interesting result is that when V is forced to use a sequence of short term contracts, the optimal contracts will no longer be contingent on the signal x.

Introducing notations: let $k_0 \geq 0$, $e_0 \geq 0$ be V, E's initial capital and effort investment levels at stage 2, $k' \geq 0$, $e' \geq 0$ be their incremental investment at stage 6E, $0 \leq p \leq 1$ be the probability of going public. Define the interim expected payoff as

$$U(k_0, e_0; k', e'; p) = Q(k_0 + k')(1 + \mu(e_0 + e'))p - k' - e',$$

and the optimal interim expected payoff as

$$\mathcal{U}(k_0, e_0; p) = \max_{k' \geq 0, e' \geq 0} Q(k_0 + k')(1 + \mu(e_0 + e'))p - k' - e'.$$

Define the interim expected surplus as

$$S(k_0, e_0; k', e'; p) = Q(k_0 + k')(1 + \mu(e_0 + e'))p - k' - e' - Q(k_0)(1 + \mu(e_0))p,$$

and the optimal interim expected surplus as

$$\mathcal{S}(k_0, e_0; p) = \max_{k' \geq 0, e' \geq 0} Q(k_0 + k')(1 + \mu(e_0 + e'))p - k' - e' - Q(k_0)(1 + \mu(e_0))p.$$

The first step is to find out how V and E's incremental investment behaviors vary with IPO probability p given k_0 and e_0. Define the solution pair (k, e) of the optimization problem

$$\max_{k \geq 0, e \geq 0} Q(k)(1 + \mu(e))p - k - e \tag{3.1}$$

as the first best investment frontier when the parameter p ranging from 0 to 1.

Lemma 3. *For any given initial investment levels k_0 and e_0 at stage 2, there exist $\underline{x}$, $\overline{x}$ with $\underline{x} \leq \overline{x}$ and the corresponding $\underline{p}$, $\overline{p}$, $\underline{p} \leq \overline{p}$ given by $p = \hat{p}_T p_E = \hat{p}_T$ and $\hat{p}_T = f(x)$ in the proof of lemma 2. $\underline{p} = \overline{p}$ only when (k_0, e_0) is on the first best investment frontier. Then after the realization of signal x at stage 3,*

1. *if $p > \overline{p}$, the venture capitalist and the entrepreneur will increase the investment levels to the first best investment frontier;*

2. *if $\underline{p} < p \leq \overline{p}$, the venture capitalist chooses $k' = 0$ and the entrepreneur overinvests effort, the interim expected final payoff $\mathcal{U}(k_0, e_0; p) - k_0 - e_0$*

is strictly less than the payoff from the first best investment frontier of the corresponding p;

3. *if $p \leq \underline{p}$, $k' = 0$ and $e' = 0$, the interim expected final payoff $\mathcal{U}(k_0, e_0; p) - k_0 - e_0$ is strictly less than the payoff from the first best investment frontier of the corresponding p.*

If the initial contract at stage 1 is short term and V decides to continue investing after realization of signal x, V and E will negotiate a new contract. I model the negotiations between V and E as a Nash bargaining game with double moral hazard (bilateral investments by both V and E). This paper follows the approach in Hermalin and Weisbach [63], but there is a delicate difference in the determination of disagreement points between this model and theirs. Since the negotiation is after both V and E choose to continue investing in the venture, the disagreement points are decided by their minmax actions as Nash rational threats instead of the threats which they can each carry out independently.

Before the negotiation, V's investment k_0, E's effort e_0 are sunk. And, the signal x, the probability p, are common knowledge to both V and E. Suppose the short term contract signed at stage 1 mandates the ownership structure at the end of stage 3 to be α_V, α_E for V and E respectively. I now calculate V and E's Nash rational threats, and the disagreement points decided by these threats. Let u_V, u_E, d_V, d_E be their interim expected payoffs and disagreement points respectively.

The value of the company $v(k, e)$ is super-modular: $v_{ke}(k, e) \geq 0$ for all $(k, e) \geq (0, 0)$, and $v_{ke}(k, e)$ is strictly greater than zero for $(k, e) > (0, 0)$, which means that V's investments and E's effort are weak complements, globally. So V's optimal investment strategy k' in response to E's further effort input e' is decreasing if E reduces e', and then V's interim expected payoff will be reduced given any existing ownership structure α_V, α_E. E's Nash rational threat is to shirk, $e' = 0$. Similarly, V's Nash rational threat is to withhold any further investment, $k' = 0$.

The disagreement points for V and E are

$$d_V = \alpha_V Q(k_0)(1 + \mu(e_0))p,$$

and

$$d_E = \alpha_E Q(k_0)(1 + \mu(e_0))p.$$

At stage 5E, V signs a new contract with E which specifies E's compensation and a new ownership structure. The contract is composed of ω_E, α'_V, and α'_E, where $\alpha'_V + \alpha'_E = 1$. ω_E is E's wage if $\omega_E \geq 0$, and it is dividend paid to V if $\omega_E < 0$. So the Nash bargaining solution with moral hazard is the choice of k', e', ω_E and α'_E which solves

$$\max_{k', e', \omega_E, \alpha'_E} \quad (u_V - d_V)(u_E - d_E), \tag{3.2}$$

with

$$u_V = (1 - \alpha'_E)Q(k_0 + k')(1 + \mu(e_0 + e'))p - k' - \omega_E,$$

and

$$u_E = \omega_E + \alpha'_E Q(k_0 + k')(1 + \mu(e_0 + e'))p - e'.$$

There are two steps to solve (3.2): first step, find the optimal k', e' to maximize V and E's joint surplus $S(k_0, e_0; k', e'; p)$ given ω_E, α'_E; second step, find the optimal ω_E, α'_E to maximize (3.2). Lemma 3 solves the first step. Then $\bar{\omega}_E$ in the proof of lemma 3 is chosen to split the surplus so that (3.2) is maximized. And E's IR constraint decides ω_E.

Note that $\mathcal{S} := \mathcal{S}(k_0, e_0; p)$ is $(u_V - d_V) + (u_E - d_E)$ when both V and E choose optimal incremental investment levels, so we have the Nash bargaining solution

$$u_V = d_V + \frac{1}{2}\mathcal{S},$$

and

$$u_E = d_E + \frac{1}{2}\mathcal{S}.$$

Lemma 3 only considers V and E's continuing investment behaviors conditioned on V choosing to continue. There has to be nonnegative surplus for V and E to share anyhow. If $\alpha_V = 1$ in the existing ownership structure, V as the sole owner of the company, has a choice to liquidate the venture and recoup qk_0. V searches for the optimal short term contracting strategy and the optimal long term contracting strategy, then chooses the one with higher *ex ante* expected payoff at stage 1.

Proposition 10. *The optimal short term contracting strategy is composed of two short term contracts, phase I and phase II. Phase I contract covers stage 1, 2, and 3; phase II contract covers stage 5E, 6E, and 7E.*

1. *In phase I contract, $\alpha_E = 0$, $\omega_E = 0$, the venture capitalist invests k_0 at stage 2, but the entrepreneur does not exert effort, $e_0 = 0$;*

2. *there exists a threshold x_s^* such that after stage 3, the venture capitalist will choose liquidation if $x \leq x_s^*$ and continuation if $x > x_s^*$;*

3. *in phase II contract, $\alpha_E = 1$, ω_E is chosen so that the venture capitalist and the entrepreneur share the joint surplus.*

Proposition 10 highlights an interesting effect on the venture capitalist's behavior caused by possible bargaining power the entrepreneur may obtain during the investment process. The venture capitalist invests in the very beginning and retains the full ownership of the company, so that he can have advantages in the negotiations with the entrepreneur, in anticipation that if the entrepreneur has higher ability than average professional managers, the venture capitalist cannot force the entrepreneur to stay, and extracting all the surplus becomes difficult.

Proposition 11. *The venture capitalist's optimal long term financial contract is option like. There exists a threshold x_l^*, such that the venture capitalist chooses to continue investing after stage 3 and transfer the ownership to the entrepreneur if the signal $x > x_l^*$; the venture capitalist keeps the ownership and waits to liquidate the company if $x \leq x_l^*$. the venture capitalist extracts all the surplus, and the entrepreneur does not exert effort until the venture capitalist decides to continue investing.*

When the investment process is governed by a sequence of short term contracts, E's bargaining power increases once V chooses to continue investing. The following proposition describes the inefficiency in two folds: (i), V inputs more initial capital in the case of short term contracts; (ii), after the arrival of new information, the technology with the quality in some range cannot receive financing from V in the case of short term contracts.

Proposition 12. *When only the technology is uncertain and the signal x is contractible, the thresholds x_s^*, x_l^* in proposition 10 and 11 satisfies $x_s^* > x_l^*$, such that*

1. *if the investment process is governed by a sequence of short term contracts, the venture capitalist will continue investing and transfer the ownership of the venture to the entrepreneur when $x > x_s^*$;*

2. *if the investment process is governed by a long term contract, the venture capitalist will continue investing and transfer the ownership of the venture to the entrepreneur when $x > x_l^*$.*

In terms of initial capital investments, $k_{0s} > k_{0l} \geq 0$.

This gives the optimal contractual choice for this investment problem.

Proposition 13. *When only the technology is uncertain and the signal x is contractible, as the venture capitalist's strategies, the optimal long term contract weakly dominates a sequence of optimal short term contracts for each realization of signal x.*

3.6 Main Results

What are the contracting behaviors, when the uncertainties are multidimensional, and not every signal is contractible? The degree of the contractibilities of the signals is mixed, as discussed in the section of model description, the signal x for the technology and the event of IPO are contractible, while the signal y for E's ability are not contractible. Would it still be optimal to write a contingent contract on x when signal y is available? This section will show that the contract incompleteness in one dimension of the uncertainties causes the incompleteness in the other dimension, even though the latter is contractible.

Consider the investment process following the complete time line: stage 1, 2, 3, and 4, then if V decides to continue investing with E as the manager, the process evolves along stage 5E, 6E, and 7E; otherwise, the game follows 5M, 6M, 7M. In the beginning at stage 1, V chooses the contractual form, either a sequence of short term contracts, or a long term contract. If the venture capital investment activities are coordinated by a sequence of short term contracts, V is expecting to negotiate a new contract with E or M when previous contract expires. If the full investment period is covered by a long term contract, after new information arrives, V can either renegotiate the existing contract with E, or he can sign a new contract with M, but replacing E might be costly. We now look at these two cases separately.

3.6.1 Short Term Contracts

First consider the case of short term contracts. Applying backward induction, suppose there is a contract initiated at stage 1 and effective until stage 3. At stage 2, V invests capital k_0 and E exerts effort e_0. Both capital investment and effort investment are sunk. At stage 3, signals x and y are realized and their values are common knowledge after realization. Suppose the contract from stage 1 mandates the ownership structure after stage 3 to be α_V, α_E for V and E respectively. Since the contract signed at stage 1 is no longer effective at stage 4 by assumption, V chooses the manager for continuation and negotiates a new contract with the chosen manager at stage 4.

When the signal y is sufficient low, V has intention to replace E with M. If V and E share the ownership of the company, $\alpha_E > 0$, it is difficult for V to remove E from the manager's position. The solution is V providing E a package of severance compensation in exchange for E to leave office. More specifically, V repurchases E's portion of ownership stake of the company, plus necessary pecuniary compensation. At the negotiation table, V will address E as follows:

" Look! We all know that you are less competent than a professional manager. Our chance of going public is slim if you stay. If you remain as a manager, I will not invest a penny beyond k_0. Then the best payoff you can receive in expectation is

$$\max_{e'_E \geq 0} \alpha_E Q(k_0)(1 + \mu(e_0 + e'_E))p_E - e'_E. \tag{3.3}$$

I can either pay you s, or reduce your stake α_E, and let M run the company. Your expected payoff remains the same, so why don't you leave. "

After E is replaced by M, V signs a contract with M. Since M is selected from a group of candidates who have identical perceived management abilities, M has no bargaining power and V extract all surplus from M (see Hermalin and Weisbach [63] p. 104 after Lemma 2).

The form of the company's value $v(k, e)$ indicates that managers' effort input is affected by their ownership stakes, V's capital investment, and the probability of a successful IPO. E and M's incentives of exerting effort are provided by sharing ownership with V, and V would extract as much surplus as possible. Intuitively, ownership should be awarded to the more productive manager, that is, to the manager with higher perceived ability. The following proposition verifies this intuition.

Proposition 14. *If the technology and the entrepreneur's ability are both uncertain, when the signal reveals that the entrepreneur is not a competent manager, and the venture capitalist intends to replace the entrepreneur by a professional manager, the venture capitalist will repurchase all of the entrepreneur's ownership stake at the price given by (3.3).*

Proposition 14 is consistent with widely adopted practice in venture financing: the venture capitalists usually retain the right to repurchase the entrepreneurs' shares (restricted stocks) upon termination of the financial contract. This also provides an explanation why the venture capitalists normally spread granting "sweet" equities to the entrepreneurs in vesting schedules throughout the whole investment processes.

The majority of venture capital backed companies are financed by issuing private equities to venture capitalists. Financing is conducted in separated, consecutive rounds. In case the company fails to reach certain thresholds (the counterpart of signal x in this context) in a given period, if the venture capitalists agree to continue financing, a new class of private equities will be issued at significantly lower prices. Additionally, restricted stocks granted to the entrepreneur are normally deposited in eschew accounts, so the vesting is intentionally back-loaded. These practices will dramatically reduce the founder's share of ownership when the company performs poorly[6].

The next step is to search for the optimal contracting strategies when the investment process is coordinated by a sequence of short term contracts. In this process, investing, contracting, and negotiations are intertwined. The paper continues to use the Nash bargaining model in negotiations as the one in section 3.5. If V decides to continue investing with E as the manager, E will have bargaining power over sharing the surplus with V. But the professional manager M is assumed to have no bargaining power.

As mentioned in the model description, at stage 4, V's investment k_0, E's effort e_0 are sunk, and signals x, y are common knowledge to both V and E. Suppose that the short term contract signed at stage 1 mandates the ownership structure at the end of stage 3 to be α_V, α_E for V and E respectively. I now calculate V and E's Nash rational threats, and the disagreement points decided by these threats. The same as the solution concept in section 3.5, the disagreement points are decided by their minmax actions as Nash rational threats, since the negotiation is after both V and E agree to continue investing in the venture.

Let $k' \geq 0$, $e'_E \geq 0$ denote V and E's capital and effort investment strategies at stage 6E. Let u_V, u_E, d_V, d_E be their interim expected payoffs and disagreement points respectively.

[6]The model does not consider tax benefits. The reverse vesting schedules become more and more popular recently mainly because of tax benefits.

Using the following notations as in section 3.5,

$$U(k_0, e_0; k', e'; p),\ \mathcal{U}(k_0, e_0; p),\ S(k_0, e_0; k', e'; p),\ \mathcal{S}(k_0, e_0; p),$$

with $p = \hat{p}_T \hat{p}_E$ when E is the manager. Similarly, $p = \hat{p}_T p_M$ when M is the manager.

By the same argument in section 3.5, the disagreement points for V and E are

$$d_V = \alpha_V Q(k_0)(1 + \mu(e_0))\hat{p}_T \hat{p}_E,$$

and

$$d_E = \alpha_E Q(k_0)(1 + \mu(e_0))\hat{p}_T \hat{p}_E,$$

respectively.

At stage 5E, V signs a new contract with E which specifies E's compensation and a new ownership structure. The contract is composed of ω_E, payable at stage 7E, specified ownership structure α'_V, α'_E, where $\alpha'_V + \alpha'_E = 1$. ω_E is E's wage if $\omega_E \geq 0$, and it is dividend paid to V if $\omega_E < 0$. So the Nash bargaining solution with moral hazard is the choice of k', e'_E, ω_E and α'_E which solves

$$\max_{k', e'_E, \omega_E, \alpha'_E} (u_V - d_V)(u_E - d_E), \tag{3.4}$$

with

$$u_V = (1 - \alpha'_E)Q(k_0 + k')(1 + \mu(e_0 + e'_E))\hat{p}_T \hat{p}_E - k' - \omega_E,$$

and

$$u_E = \omega_E + \alpha'_E Q(k_0 + k')(1 + \mu(e_0 + e'_E))\hat{p}_T \hat{p}_E - e'_E.$$

The two steps to solve (3.4) remain the same: (i), find the optimal k', e'_E to maximize V and E's joint surplus $S(k_0, e_0; k', e'; p)$ given ω_E, α'_E; (ii), find the optimal ω_E, α'_E to maximize (3.4). Since all agents are risk neutral, and E's ownership stake plays a major role in providing E incentives, it is easy to see that when $\alpha'_E = 1$, V and E choose k', e'_E so that the overall capital and effort investment levels will maximize the joint surplus. Then ω_E is chosen to split the surplus so that (3.4) is maximized.

The calculation of ω_E is straightforward. We already know that $\alpha'_E = 1$ in the optimal solution. Let k^*, e^*_E be the solutions for $\mathcal{S}(k_0, e_0; p)$. They exist by assumption 1 and 2. And, $\mathcal{S} := \mathcal{S}(k_0, e_0; p)$ is exactly the value of $(u_V - d_V) + (u_E - d_E)$, so we have the Nash bargaining solution

$$u_V = d_V + \frac{1}{2}\mathcal{S},$$

and

$$u_E = d_E + \frac{1}{2}\mathcal{S}.$$

At stage 4, V calculates the maximal expected payoff when either E or M is manager, then decides whether it is optimal to continue the investment or choose the liquidation. In the choice of manager, since E will have bargaining power over sharing surplus, but M will not have this power, V will demand higher ability level from E. This is true for general stage 2 investment levels k_0, e_0, and interim ownership structure α_V, α_E at stage 4, but for the reason of proving the main result, I only need the following special case.

Lemma 4. *Suppose the venture capitalist invests capital k_0, but the en-*

trepreneur does not exert effort $e_0 = 0$ *at stage 2, and suppose the venture capitalist is the sole owner of the company until stage 4,* $\alpha_V = 1$*, then given the signal realization* x, y*, there exists a threshold* $\tilde{y}_{E,s}(k_0, x)$ *such that it is optimal to replace the entrepreneur by a professional manager if* $y \leq \tilde{y}_{E,s}(k_0, x)$*.* $\tilde{y}_{E,s}(k_0, x) \geq y^*$ *for any* k_0*, and* x*, where* y^* *is given in lemma 1.*

Let $\mathcal{X} \times \mathcal{Y}$ denote the signal space for x and y. In the space $\mathcal{X} \times \mathcal{Y}$, let $\Pi_{L,E}$, $\Pi_{L,M}$, $\Pi_{I,E}$, and $\Pi_{I,M}$ denote the regions of the signals in which it is optimal to liquidate the venture with E, M as the manager, and to invest with E, M as the manager, respectively. Since there are one-to-one, monotonic correspondences between the signals x, y and the IPO probabilities $\hat{p}_T$, $\hat{p}_E$, I can use these notations to denote the regions of updated beliefs without causing confusion.

Lemma 5. *Under the assumptions of lemma 4, the regions for each optimal decision,* $\Pi_{L,E}$*,* $\Pi_{L,M}$*,* $\Pi_{I,E}$*, and* $\Pi_{I,M}$*, are given in Figure 3.5.*

The following proposition characterizes the optimal contracts and the equilibrium when the venture capital investment process is governed by a sequence of short term contracts.

Proposition 15. *Suppose both the technology and the entrepreneur's ability are uncertain, with signal* x *being contractible but not signal* y*. The optimal short term contracting strategy is composed of two short term contracts, phase* I *and phase* II*. Phase* I *contract covers stage 1, 2, and 3; phase* II *contract covers stage 5E, 6E, and 7E, or stage 5M, 6M, and 7M. The phase I contract is not contingent on signal* x*.*

1. *In phase I contract,* $\alpha_E = 0$*,* $\omega_E = 0$*, the venture capitalist invests* k_0 *at stage 2, but the entrepreneur does not exert effort,* $e_0 = 0$*;*

2. *given stage 2 sunk investments* k_0 *and* $e_0 = 0$*, and stage 3 signal realizations* x*,* y*, then at stage 4, the venture capitalist chooses a professional manager as the manager if* $y \leq y^*_{E,s}(x)$*, and the entrepreneur as the manager otherwise;*

3. *also at stage 4, after the venture capitalist has chosen the manager, he decides to continue investing, or to wait for liquidation;*

4. *if the venture capitalist decides to continue the investment at stage 4, then in the phase II contract, full ownership will be granted to the manager.*

Proposition 15 says that in the existence of noncontractible signal y, it is optimal to exclude the clauses which are contingent on signal x. Theoretically, the ability to replace the entrepreneur by a professional manager gives the venture capitalist option value, and exclusion of contingent clauses on signal x gives the venture capitalist further option value on when to exercise this option. Moreover, the venture capitalist chooses the initial investment and the initial ownership structure of the company so that he will have advantages in the possible negotiations with the entrepreneur in the future.

The proposition also shows that for some signal realizations, the venture capitalist should replace the entrepreneur by a professional manager, even if he is seeking liquidation eventually. This seems counterintuitive at first. For the venture capitalists, this is a balance between providing the manager incentives and protecting their owner rights. The prospect of the venture is the key factor for decision making. Generally speaking, IPO and liquidation are two modes used by venture capitalists to exit the financing. The exit decisions depend on the outlook of the venture, which in turn decides the transferal of the venture ownership. This situation is typical in venture capital backed pharmaceutical companies, where the payoff distributions are highly skewed. These companies are often operated by seasoned professional managers recruited by venture capitalists, and the venture capitalists hold most shares outstanding at the time of IPO.

3.6.2 Long Term Contracts

In this section, I will describe the optimal long term contingent contract. Since the venture capitalist has all the bargaining power in the beginning of the investment, to show it is indeed the best strategy for the venture capitalists to choose short term contracts and leave the contracts open for future negotiation, it is sufficient to show the optimality of short term contracts, by comparing the performances of the short term contracts and the long term contracts.

As discussed in section 3.5, the advantage of a long term contract is that it eliminates the possibility of (re)negotiation so that E have no chance to demand increasing share of surplus from V in the middle of investment. However, this is not true when there is additional information during the investment process and this information, which is orthogonal to the other contractible signal, cannot be described in the initial contract.

Suppose V employs long term contingent contract in the beginning, then after both signals on the technology and E's ability are revealed, if V decides to continue investing with E as the manager, and if V has to renegotiate the existing contract with E, E will obtain bargaining power similar to the case of short term contracts, in which V and E negotiate a new second phase contract. This is because the existing contract will be used as a reference point, and the claim is no longer a mere assumption.

Proposition 16. *Whenever there is a renegotiation of the existing long term contract, the entrepreneur possesses bargaining power and shares positive fraction of the surplus with the venture capitalist.*

In search for the optimal contract in the category of long term contingent ones, signal x and the event of IPO are contractible. And, the signal x in the contract decides the allocation of ownership and the wage (dividend) payable to (by) E. In the mean time, the signal y, which is uncorrelated to x, cannot be contracted upon, so the contingent clauses on x is written based on the probability distribution of y.

After the realizations of the signals, there is a possibility of renegotiation. If E's interim expected payoff is positive at stage 5E, she indeed has all the surplus and will not accept any alternative contractual offer from V, then V loses all surplus to E in this case. On the other hand, by proposition 16, if E's interim expected payoff is negative at stage 5E, E threats to quit, V then will have surplus at $p = \hat{p}_T p_M$ at most, so V will propose a new contract to share the surplus with E. Optimally, V chooses a contract in the beginning of the investment with full intention to renegotiate this contract during the process.

For any realization of signal x, there is a possibility that choosing M as the manager is optimal. If E has ownership by the initial contract, it is optimal for V to repurchase E's shares and grant them to M, but this is costly. There are two ways to reduce this cost, decreasing the initial investment k_0, and delaying transferring ownership to E. This leads to degeneracy of the long term contingent contract.

Proposition 17. *Suppose both the technology and the entrepreneur's ability are uncertain, with signal x being contractible but not signal y. The optimal long term contract is degenerated, in the sense that it is not contingent on the signal x. In the beginning of the investment, the venture capitalist retains all the ownership, $\alpha_V = 1$, and there is no wage or dividend payment, $\omega_E = 0$. The venture capitalist signs the initial contract with full intention for future renegotiation.*

Since the equilibrium outcomes will be the same when the investment is governed by a sequence of short term contracts and a degenerated long term contract, we have:

Proposition 18. *When both the technology and the entrepreneur's ability are uncertain, with signal x being contractible but not signal y. The optimal contracting strategy for the venture capitalist is a sequence of short term contracts with interim negotiations.*

By the result of proposition 15, E does not exert effort, $e_0 = 0$, initially at stage 2. It is interesting to find out V's initial investment behavior under the anticipation of possible future negotiation between V and E. Since the optimal long term contingent contract is degenerated, I only need to solve for the case when the investment is governed by a sequence of short term contracts given $e_0 = 0$.

From the proof of lemma 4, let $\mathcal{M}(k_0, x)$ denote the set of p_M's such that $\tilde{p}_{E,s}(k_0, x) \leq 1$ for given k_0 and x. Define

$$\tilde{p}_M := \min\{1/2, \inf_{k_0,x} \sup_{p_M} \mathcal{M}(k_0, x)\}.$$

If $\tilde{p}_M = 0$ is perceived *ex ante* at stage 1, V will replace E by M with certainty. Then $k_0 = 0$, and no contract is necessary. Now consider the more interesting case $\tilde{p}_M > 0$.

Proposition 19. *When $p_M < \tilde{p}_M$, the venture capitalist contributes positive initial investment, $k_0 > 0$, for a better bargaining position in possible negotiations later with the entrepreneur.*

3.7 An Example

General forms of production functions could be

$$Q(k) := k^m;$$

and

$$\mu(e) := (e + i_e)^n - i_e^n,$$

where $0 < m, n < 1$, $i_e \geq 0$. Here, I will use $m = n = 1/2$, and $i_e = 0$ to illustrate the model, that is, $Q(k) = \sqrt{k}$, and $\mu(e) = \sqrt{e}$.

The first best investment frontier is the solution pair (k, e) for (3.1), and they are

$$k = \Big(\frac{2p}{4-p^2}\Big)^2, \quad e = \Big(\frac{p^2}{4-p^2}\Big)^2.$$

Since $\lim_{e\to 0} \mu'(e) = +\infty$, $p_0 = 0$ which is defined in the proof of lemma 3. Then, given V and E's sunk capital and effort investments (k_0, e_0) at stage 2, solving equations (3.9) and (3.10), we have

$$\overline{p} = \frac{4\sqrt{k_0}}{1+\sqrt{1+4k_0}}, \quad \underline{p} = \sqrt{\frac{4\sqrt{e_0}}{1+\sqrt{e_0}}}.$$

And finally,

$$\mathcal{U}(k_0, e_0; p) = \begin{cases} \frac{p^2}{4-p^2} + k_0 + e_0, & \text{if } p > \overline{p}; \\ p\sqrt{k_0} + \frac{p^2}{4}k_0 + e_0, & \text{if } \underline{p} < p \le \overline{p}; \\ \sqrt{k_0}(1+\sqrt{e_0})p, & \text{if } p \le \underline{p}. \end{cases}$$

An interesting property of $\mathcal{U}$ is that $\mathcal{U}$ as a function of p is differentiable at $p = \overline{p}$. To see this, at the point $p = \overline{p}$, we have

$$p = \frac{4\sqrt{k_0}}{1+\sqrt{1+4k_0}}, \quad \text{or,} \quad \sqrt{k_0} = \frac{2p}{4-p^2}.$$

Consider the right and left derivatives of $\mathcal{U}(k_0, e_0; p)$ with respect to p.

$$(\frac{p^2}{4-p^2} + k_0 + e_0)' = \frac{2p}{4-p^2} + \frac{2p^3}{(4-p^2)^2},$$

and

$$(p\sqrt{k_0} + \frac{p^2}{4}k_0 + e_0)' = \sqrt{k_0} + \frac{p}{2}k_0 = \frac{2p}{4-p^2} + \frac{2p^3}{(4-p^2)^2}.$$

So $\mathcal{U}(k_0, e_0; p)$ as a function of p is differentiable at $p = \overline{p}$.

3.8 Empirical Implications

In the beginning of each round of venture financing, the venture capitalist and the entrepreneur negotiate over tentative term sheets, preliminary agreements on investor rights, voting rights, and issuance of a new class of private equities. The essence of comparing long term contingent contracts with short term open-ended contracts is to analyze the question of how investors balance between rigidity and flexibility in financial contracts and agreements. This view builds a bridge connecting contract theory with actual practice of financing. Detailed contract clauses provide rigidity, while staged financing with negotiable agreements or open-ended contracts provide flexibility. This theory can be tested by examining what is included and what is excludes in contracts, and the variation of clauses from stage to stage.

The model introduced by this paper is different from the real option theory in the information structure. The real option theory explains the situation in which information arrives sequentially, such as drug research: the results of laboratory studies are followed by the results of the clinical trials. However, there are also situations where information arrives in parallel. For example, there may be no distinguishable sequentiality between the information about the market reaction to an innovation and the information about the entrepreneur's managerial abilities. Under this circumstance, the paper predicts certain contracting and negotiation behaviors. I find upward investment distortion in initial rounds of staged financing, which remains to be further tested.

The paper has provided some guidelines in examining the venture capitalists' investment strategies. Generally speaking, each venture capitalist's alternative action is a form of protection from downside risks. When the entrepreneur seems to be less competent as a manager, the venture capitalist would consider recruiting a seasoned professional manager. Also, liquidation is a choice when the venture has little probability of going public. The venture capitalists would design contracts to secure decision making options so that these alternative choices will be kept open in future investments. When the existing investments are reasonably protected, the venture capitalist then considers strategies to better capture upside payoffs.

The model predicts that it is optimal to repurchase all shares held by the entrepreneur upon termination of the employment, but one limitation of this model is that it does not consider behavioral factors. In empirical studies, it is necessary to separate the observations of actual contracts from implementation of these contracts. In practice, although the venture capitalist could hold contractual rights to repurchase all of the entrepreneur's shares upon termination of the contract, anecdotal evidence suggests that there are possibly psychological factors involved – the venture capitalist would let the entrepreneur remain to be a shareholder out of sympathy.

3.9 Conclusion

This paper presents a dynamic model, which incorporates contracting, negotiations, and investments in venture financing. The model explains that implementation of optimal short term open-ended financial contracts leads to staged financing in venture capital investments. For each category of down side risks about the investment, the venture capitalist could have corresponding alternatives to mitigate these risks. And signals related to these risks will be revealed to both the venture capitalist and the entrepreneur during the investment process. But the information structure may be complicated and some of the information may be difficult to described in the contract. The venture capitalist would choose short term open-ended contracts so that the options of choosing these alternatives could be kept open in the future. This theory is fundamentally different from the "real option" theory, where waiting creates option value in a model of investment under uncertainty.

Staged financing gives the venture capitalist the option to tailor the ownership structure of a privately-held venture-capital-backed company according to information update. The venture capitalist values ownership because there are control rights naturally imbedded in ownership, but ownership also functions as an incentive for the entrepreneur to exert effort. The paper predicts that the general rule would be, ownership of the company will gradually shift from the venture capitalist to the entrepreneur if additional information indicates a higher probability of success. Otherwise, the venture capitalist retains ownership to protect existing investments. This rule can be easily extended to the situation where information updates occur in consecutive stages.

The paper offers a novel view in which investment, ownership structure, and existing long term contracts function as reference points in negotiations between the venture capitalist and the entrepreneur. Staged financing is costly for the venture capitalist because as the venture is developing, if the prospect of the venture appears promising, the bargaining power of the entrepreneur becomes stronger in sharing venture surplus. When this happens, protection of previous investments is less of concern to the venture capitalist. Instead, the venture capitalist allocates capital investments, chooses ownership structures, and design initial contracts in order to have considerable leverage in later negotiations with the entrepreneur over sharing surplus.

There are many natural extensions to my model. The model can be used to explain when innovations should be financed internally through company's R&D projects, and when innovations should be financed externally by specialized investors. Another possible extension of the model would be one that incorporates geographical factors and social networks among the investors.

3.10 Proofs

3.10.1 Proof of Lemma 1

Because the expected gain of V, E, or M is $\alpha Q(k)(1+\mu(e))\Pr(IPO)$ in a separable form, where α is the fraction of ownership, the question boils down to comparing different $\hat{p}$ under E and M's management. Since $\hat{p}_T$ is the same in each situation, I only need to compare $\hat{p}_E = \Pr(\delta'_E > \delta^*)$ and

$\hat{p}_M = \Pr(\delta_M > \delta^*)$.

By the calculation in section 3.3.3,

$$\Pr(\delta'_E > \delta^*) \;=\; \sqrt{\frac{\hat{h}_E}{2\pi}} \int_{\delta^*}^{\infty} e^{-\frac{\hat{h}_E}{2}(t-\hat{\delta}_E)^2} dt \;= \\ \sqrt{\frac{h_E + h_y}{2\pi}} \int_{\delta^*}^{\infty} \exp\Big[-\frac{h_E + h_y}{2}\big(t - \frac{y h_y}{h_E + h_y}\big)^2 \Big] dt,$$

and

$$\Pr(\delta_M > \delta^*) \;=\; \sqrt{\frac{h_M}{2\pi}} \int_{\delta^*}^{\infty} e^{-\frac{h_M}{2}t^2} dt.$$

$\Pr(\delta'_E > \delta^*)$ is a continuous, strictly increasing function of y, while $\Pr(\delta_M > \delta^*)$ is a constant function with respect to y.

We also have

$$\lim_{y \to +\infty} \Pr(\delta'_E > \delta^*) = 1 > \Pr(\delta_M > \delta^*) > 0,$$

and

$$\lim_{y \to -\infty} \Pr(\delta'_E > \delta^*) = 0.$$

The existence of y^* follows from Intermediate Value Theorem. The value of y^* is unique for each given k, e, x, and ownership structure. It can be numerically calculated from Implicit Function Theorem by equating $\Pr(\delta'_E > \delta^*)$ and $\Pr(\delta_M > \delta^*)$. $y^* > 0$ holds, because $h_E = h_M$, $h_y > 0$ implies $\hat{h}_E > h_M$.

Q.E.D.

3.10.2 Proof of Lemma 2

Note that

$$
\begin{aligned}
\hat{p}_T &= \Pr(\delta'_T > \delta^*) \\
&= \sqrt{\frac{\hat{h}_T}{2\pi}} \int_{\delta^*}^{\infty} e^{-\frac{\hat{h}_T}{2}(t-\hat{\delta}_T)^2} dt \\
&= \sqrt{\frac{h_T + h_x}{2\pi}} \int_{\delta^*}^{\infty} \exp\Big[-\frac{h_T + h_x}{2}(t - \frac{x h_x}{h_T + h_x})^2\Big] dt \\
&= 1 - \Phi\Big(\sqrt{h_T + h_x}(\delta^* - \frac{x h_x}{h_T + h_x})\Big).
\end{aligned}
$$

So $\hat{p}_T$ is a smooth, strictly increasing function of signal x. Denote this function as $\hat{p}_T = f(x)$. The prior probability density function for x is $\phi(\sqrt{h_T h_x/(h_T + h_x)}x)$ by assumption. So the prior probability density function for $\hat{p}_T$ is given by

$$
\rho_T(\hat{p}_T) = \frac{\phi(\sqrt{h_T h_x/(h_T + h_x)} f^{-1}(\hat{p}_T))}{f'(f^{-1}(\hat{p}_T))},
$$

by change of variables formula for density functions of random variables. The proof is the same for signal y and probability $\hat{p}_E$.

Q.E.D.

3.10.3 Proof of Lemma 3

Consider a general moral hazard problem faced by V,

$$
\max_{k \geq 0, e \geq 0, 0 \leq \alpha_E \leq 1, \omega_E} (1 - \alpha_E) Q(k)(1 + \mu(e)) p - k - \omega_E, \tag{3.5}
$$

subject to E's IC constraint

$$e \in \underset{e' \geq 0}{\operatorname{argmax}} \ \omega_E + \alpha_E Q(k)(1 + \mu(e'))p - e', \tag{3.6}$$

and IR constraint

$$\omega_E + \alpha_E Q(k)(1 + \mu(e))p - e \geq \bar{\omega}_E. \tag{3.7}$$

By assumption 2, $\mu(\cdot)$ is strictly concave, $\mu'(\cdot)$ is strictly decreasing and goes to 0 as e goes to infinity, the question whether we can replace E's IC constraint by first order condition depends on the marginal productivity of both V and E, along with the probability of IPO. In general, E's IC constraint is equivalent to either $e = 0$ or the first order condition:

$$e\Big[\alpha_E Q(k)\mu'(e)p - 1\Big] = 0. \tag{3.8}$$

Since the solution varies with parameter p, and E's optimal effort level is nondecreasing in p and bounded below by 0, there exists a $p_0 \in [0, 1]$ such that the IC constraint is $e = 0$ when $p \leq p_0$ and it is the first order condition when $p > p_0$. p_0 depends on the marginal productivity of both V and E. If $p_0 = 1$, then $\underline{p} = \overline{p} = p_0 = 1$, since V has no incentive to transfer ownership to E. If $p_0 < 1$, then it is obvious that $\underline{p} \geq p_0$ by the same reason. So it is valid to replace E's IC constraint by first order condition of (3.6).

It is easy to see that when $\alpha_E = 1$, V and E's investment levels are the solutions of (3.1). This is the first best investment frontier, the upper left investment curve in Figure 3.3. Then ω_E is chosen such that E's IR constraint is binding.

Given V and E's sunk capital and effort investments (k_0, e_0) at stage 2, and suppose both V and E are rational in the sense that they will not invest beyond the maximum capital investment level for $p = 1$ and the first best investment frontier. Define $\overline{p}$ as the solution of the pair of equations

$$\begin{cases} Q'(k_0)(1+\mu(e))p - 1 = 0, \\ Q(k_0)\mu'(e)p - 1 = 0, \end{cases} \tag{3.9}$$

with p, e being unknown variables. Define $\underline{p}$ as the solution of the pair of equations

$$\begin{cases} Q'(k)(1+\mu(e_0))p - 1 = 0, \\ Q(k)\mu'(e_0)p - 1 = 0, \end{cases} \tag{3.10}$$

with p, k being unknown variables. The solutions for p of equations (3.9) and (3.10) exist and are nonnegative. Take equations (3.10) for example,

$$p = \frac{1}{Q'(k)(1+\mu(e_0))}$$

is increasing from 0 to positive infinity as k goes from 0 to infinity, and

$$p = \frac{1}{Q(k)\mu'(e_0)}$$

is decreasing from positive infinity to a constant as k goes from 0 to infinity. $\underline{p}$, $\overline{p} \leq 1$ because I assume both V and E are rational and they will not invest (k_0, e_0) beyond the first best, and I also assume $p_0 < 1$. $\underline{p} \leq \overline{p}$ since E would not exert effort beyond the first best investment frontier given V's capital investment k_0.

In Figure 3.4, as $p > \overline{p}$, both V and E will increase their investment levels to the first best. When $\underline{p} < p \leq \overline{p}$, V has invested k_0 which is overinvesting, but he cannot disinvest, so the optimal incremental investment is $k' = 0$. E will make incremental investment exceeding the first best, since

$$\mu'(e) = \frac{1}{Q(k_0)p},$$

V is overinvesting, and V, E's investments are strictly complements outside boundary $k = 0$ and $e = 0$, globally. When $p \leq \underline{p}$, both V and E have overinvested, so $k' = 0$, $e' = 0$, and the allocation of ownership is no longer important as to the investment per se.

Q.E.D.

3.10.4 Proof of Proposition 10

The time line of the game is stage 1, 2, 3, 5E, 6E, and 7E. Negotiations and contracting will be conducted at stage 1 and 5E. Investments will be made simultaneously by both V and E at stage 2 and 6E. Signal x will be realized at stage 3. All uncertainties of the technology, the managers' abilities, and IPO will be resolved at stage 7E.

The posterior probability of IPO perceived by V and E from stage 3 on is $p = \hat{p}_T\hat{p}_E$, and the stage 1 prior probability distribution of p is given by $\rho = \rho_T\rho_E$ since the technology, E's ability, and the signals are assumed to be independent to each other. In this section, E's ability is certain and is above the threshold δ^*, so $\hat{p}_E = 1$, $p = \hat{p}_T$, and $\rho = \rho_T$.

Since the contracts are short term, let α_V, α_E be the ownership structure after investments k_0, e_0, and the signal x. If V, E's contracting and investing behaviors will be optimal at stage 5E, 6E, then V's interim expected payoff from IPO is

$$\begin{aligned} u_V &= d_V + \frac{1}{2}\mathcal{S} \\ &= \alpha_V Q(k_0)(1+\mu(e_0))p + \frac{1}{2}\mathcal{S}(k_0, e_0; p) \\ &= \frac{1}{2}\mathcal{U}(k_0, e_0; p) + (\alpha_V - \frac{1}{2})Q(k_0)(1+\mu(e_0))p. \end{aligned}$$

Then V would choose α_V as high as possible at stage 1. Additionally, V will have liquidation choice if $\alpha_V = 1$. So $\alpha_V = 1$ is optimal in stage 1 short term contract.

Then $\alpha_E = 0$ which implies $d_E = 0$. Let k^*, e^* denote V, E's optimal incremental investments at stage 6E. Then

$$Q(k_0 + k^*)\mu'(e_0 + e^*)p = 1,$$

if $e^* > 0$;

$$Q(k_0 + k^*)\mu'(e_0 + e^*)p \leq 1,$$

if $e^* = 0$, since it is optimal for E to exert positive amount of effort when the left hand side is strictly greater than 1.

E's expected payoff at stage 1 is

$$\frac{1}{2}\int_{p\in I}\Big[Q(k_0+k^*)(1+\mu(e_0+e^*))p-k^*-e^*-Q(k_0)(1+\mu(e_0))p\Big]\rho(p)dp-e_0,$$

where I is the interval for p in which V decides to continue investing.

Take first order derivative with respect to e_0 using Envelope Theorem, since both k^* and e^* are functions of e_0:

$$\frac{1}{2}\int_{p\in I}\Big[Q(k_0+k^*)\mu'(e_0+e^*)p-Q(k_0)\mu'(e_0)p\Big]\rho(p)dp-1.$$

Note that the first term of the integrand is less or equal to 1 and the second term of the integrand is positive. So the integral is strictly less than 1 regardless the interval I, which means E's optimal initial effort investment is $e_0=0$.

Given $\alpha_V=1$, $e_0=0$ after stage 3, V chooses between continuing the venture and liquidation. If V decides to continue investing, as in the beginning of the proof of lemma 3, $\alpha'_E=1$ and $\bar{\omega}_E$ is chosen such that V's interim expected payoff is

$$\begin{aligned}d_V+\frac{1}{2}\mathcal{S}(k_0,0;p) \;&=\; \frac{1}{2}\Big[\mathcal{U}(k_0,0;p)+Q(k_0)p\Big] \;=\\ &\frac{1}{2}\Big[\max_{k'\geq 0,e'\geq 0}Q(k_0+k')(1+\mu(e'))p-k'-e'+Q(k_0)p\Big],\end{aligned}$$

where $d_V=Q(k_0)p$ since $\alpha_V=1$, $e_0=0$. If V decides to liquidate the company, V's interim expected payoff is $Q(k_0)p+qk_0(1-p)$.

So V's interim expected payoff from either outcome – IPO or liquidation, is

$$\max\left\{\ \frac{1}{2}\Big[\mathcal{U}(k_0,0;p)+Q(k_0)p\Big],\ Q(k_0)p+qk_0(1-p)\ \right\}. \tag{3.11}$$

Take difference of these two payoffs with p as a parameter,

$$\frac{1}{2}\Big[\mathcal{U}(k_0,0;p)+Q(k_0)p\Big]-\Big[Q(k_0)p+qk_0(1-p)\Big]\ =\ \frac{1}{2}\mathcal{S}(k_0,0;p)-qk_0(1-p). \tag{3.12}$$

By Envelope Theorem, $\mathcal{S}(k_0,0;p)/2$ is strictly increasing in p and goes from 0 to a positive constant as p goes from 0 to 1. In the mean time, $qk_0(1-p)$ is strictly decreasing in p and goes from $qk_0>0$ to 0 as p goes from 0 to 1. Then by Intermediate Value Theorem, there exists $\tilde{p}_s(k_0)$ and thus a corresponding $\tilde{x}_s(k_0)=f^{-1}(\tilde{p}_s(k_0))$, such that V chooses liquidation if $x\le\tilde{x}_s(k_0)$ and investing if $x>\tilde{x}_s(k_0)$ after stage 3.

V's optimal initial investment level k_{0s} is the solution of

$$\max_{k_0}\int_0^1\max\left\{\ \frac{1}{2}\Big[\mathcal{U}(k_0,0;p)+Q(k_0)p\Big],\ Q(k_0)p+qk_0(1-p)\ \right\}\rho(p)dp-k_0. \tag{3.13}$$

The solution always exists since k_0 lies in a closed interval bounded by 0 and the maximal investment level of the first best investment frontier, which is compact. Then $x_s^*=\tilde{x}_s(k_{0s})$ for the optimal k_{0s}.

Q.E.D.

3.10.5 Proof of Proposition 11

Start from the same setting in the proof of proposition 10. The time line of the game is stage 1, 2, 3, 5E, 6E, and 7E. Contracting is only at stage 1.

Investments will be made simultaneously by both V and E at stage 2 and 6E. Signal x will be realized at stage 3. All uncertainties of the technology, the managers' abilities, and IPO will be resolved at stage 7E.

The posterior probability of IPO perceived by V and E from stage 3 on is $p = \hat{p}_T \hat{p}_E$, and the stage 1 prior probability distribution of p is given by $\rho = \rho_T \rho_E$ since the technology, E's ability, and the signals are assumed to be independent to each other. Suppose E's ability is certain and is above the threshold δ^*, so $\hat{p}_E = 1$, $p = \hat{p}_T$, and $\rho = \rho_T$.

Suppose the contract is long term and apply backward induction. The state space of signal x is perfectly foreseeable and x is contractible, so V would (weakly) prefer a contract contingent on the signal x, since any contract unrelated to x is the extreme form of a trivial contingent contract. Then after the realization of the signal x at stage 3, the contract is the solution of the following.

$$\max_{k' \geq 0, e' \geq 0, 0 \leq \alpha_E \leq 1, \omega_E} (1 - \alpha_E) Q(k_0 + k')(1 + \mu(e_0 + e'))p - k' - \omega_E, \tag{3.14}$$

subject to E's IC constraint

$$e' \in \operatorname*{argmax}_{e'' \geq 0} \ \omega_E + \alpha_E Q(k_0 + k')(1 + \mu(e_0 + e''))p - e'', \tag{3.15}$$

and IR constraint

$$\omega_E + \alpha_E Q(k_0 + k')(1 + \mu(e_0 + e'))p - e' \geq \bar{\omega}_E, \tag{3.16}$$

with α_E, ω_E, and $\bar{\omega}_E$ are functions of x, and it is suppressed for simplification of notations. If V chooses to continue investing, it is easy to see that the optimal solution involves $\alpha_E(x) = 1$.

Since V has all bargaining power at stage 1, V would choose $\bar{\omega}_E(x) = 0$. Given V's proposal of $\bar{\omega}_E(x) = 0$, E exerts zero effort at stage 2, $e_0 = 0$. Then it is optimal for V to be the sole owner of the company until stage 3, $\alpha_V = 1$, since V will have choice to liquidate the company when x is sufficiently low. The next step is to search for the range of x in which V will keep investing.

Given $\alpha_V = 1$, $e_0 = 0$ after stage 3, V chooses between continuing the venture and liquidation. If V decides to continue investing, $\alpha_E(x) = 1$ and $\bar{\omega}_E(x) = 0$. So V's interim expected payoff is $\mathcal{U}(k_0, 0; p)$ from IPO. If V decides to liquidate the company, V's interim expected payoff is $Q(k_0)p + qk_0(1-p)$.

So V's interim expected payoff is essentially

$$\max\Big\{\ \mathcal{U}(k_0, 0; p),\ Q(k_0)p + qk_0(1-p)\ \Big\}. \tag{3.17}$$

Take difference of these two payoffs with p as a parameter,

$$\mathcal{U}(k_0, 0; p) - \Big[Q(k_0)p + qk_0(1-p)\Big] \ = \ \mathcal{S}(k_0, 0; p) - qk_0(1-p). \tag{3.18}$$

By Envelope Theorem, $\mathcal{S}(k_0, 0; p)$ is strictly increasing in p and goes from 0 to a positive constant as p goes from 0 to 1. In the mean time, $qk_0(1-p)$ is strictly decreasing in p and goes from $qk_0 > 0$ to 0 as p goes from 0 to 1. Then by Intermediate Value Theorem, there exists $\tilde{p}_l(k_0)$ and thus a corresponding $\tilde{x}_l(k_0) = f^{-1}(\tilde{p}_l(k_0))$, such that V chooses liquidation if $x \leq \tilde{x}_l(k_0)$ and investing if $x > \tilde{x}_l(k_0)$ after stage 3.

V's optimal initial investment level k_{0l} is the solution of

$$\max_{k_0} \int_0^1 \max\left\{ \ \mathcal{U}(k_0, 0; p), \ Q(k_0)p + qk_0(1-p) \ \right\}\rho(p)dp - k_0. \tag{3.19}$$

The solution always exists since k_0 lies in a closed interval bounded by 0 and the maximal investment level of the first best investment frontier, which is compact. Then $x_l^* = \tilde{x}_l(k_{0l})$ for the optimal k_{0l}.

Q.E.D.

3.10.6 Proof of Proposition 12

In this proof, I am going to show that $k_{0s} > k_{0l} \geq 0$, and ultimately $x_s^* > x_l^*$. Then for some realization in the signal space of x, V will not choose to continue investing under short term contracts, but will do so under long term contracts. This establishes the inefficiency both in capital investments and in the choice of technology.

To show $x_s^* > x_l^*$, I will show $\tilde{p}_s(k_0) > \tilde{p}_l(k_0)$ for each $k_0 > 0$; then for the V's optimal initial investment levels, $k_{0s} > k_{0l} \geq 0$; and finally, $\tilde{p}_l(k_0)$ is (weakly) increasing in k_0 for $k_0 \geq k_{0l}$. Combine these three, it gives

$$\tilde{p}_s(k_{0s}) > \tilde{p}_l(k_{0s}) \geq \tilde{p}_l(k_{0l}).$$

Finally, by lemma 2, $f^{-1}(\cdot)$ is strictly increasing function of p, so $x_s^* > x_l^*$ from the definitions of x_s^* and x_l^*.

To show the first claim $\tilde{p}_s(k_0) > \tilde{p}_l(k_0)$ for each $k_0 > 0$, fix an arbitrary feasible $k_0 > 0$. Set the right hand side of the equation (3.12) equal to zero, then $\tilde{p}_s(k_0)$ is the solution of this new equation

$$\frac{1}{2}\mathcal{S}(k_0, 0; p) - qk_0(1-p) \;=\; 0.$$

The solution is unique because both of the terms on the left hand side are monotonically increasing. Similarly, set the right hand side of the equation (3.18) equal to zero, then $\tilde{p}_l(k_0)$ is the solution of this new equation

$$\mathcal{S}(k_0, 0; p) - qk_0(1-p) \;=\; 0.$$

The solution is unique because both of the terms on the left hand side are also monotonically increasing.

Since $k_0 > 0$, $qk_0(1-p) > 0$ when $p = 0$, while $\mathcal{S}(k_0, 0; p) = 0$ when $p = 0$. This implies $\tilde{p}_s(k_0) > 0$ and $\tilde{p}_l(k_0) > 0$. Now suppose $\tilde{p}_s(k_0) \leq \tilde{p}_l(k_0)$, then

$$\begin{aligned} qk_0(1-\tilde{p}_s(k_0)) &\geq qk_0(1-\tilde{p}_l(k_0)) = \mathcal{S}(k_0, 0; \tilde{p}_l(k_0)) \\ &> \frac{1}{2}\mathcal{S}(k_0, 0; \tilde{p}_l(k_0)) \geq \frac{1}{2}\mathcal{S}(k_0, 0; \tilde{p}_s(k_0)) = qk_0(1-\tilde{p}_s(k_0)). \end{aligned}$$

It is impossible. This finishes the proof of the first claim.

Next, prove the second claim $k_{0s} > k_{0l} \geq 0$. Let k^*, e^* denote V, E's optimal incremental investments at stage 6E as in the proof of proposition 10.

From the equation (3.13) and the equation (3.19), k_{0s} solves

$$\max_{k_0} \int_0^{\tilde{p}_s(k_0)} \Big[Q(k_0)p + qk_0(1-p)\Big]\rho(p)dp + \int_{\tilde{p}_s(k_0)}^1 \frac{1}{2}\Big[\mathcal{U}(k_0, 0; p) + Q(k_0)p\Big]\rho(p)dp - k_0;$$

k_{0l} solves

$$\max_{k_0} \int_0^{\tilde{p}_l(k_0)} \Big[Q(k_0)p + qk_0(1-p)\Big]\rho(p)dp + \int_{\tilde{p}_l(k_0)}^1 \mathcal{U}(k_0, 0; p)\rho(p)dp - k_0. \quad (3.20)$$

By definition of $\mathcal{U}$ and k^*, e^*,

$$\mathcal{U}(k_0, 0; p) \;=\; Q(k_0 + k^*)(1 + \mu(e^*))p - k^* - e^*.$$

And taking derivative with respect to k^*, we have

$$Q'(k_0 + k^*)(1 + \mu(e^*))p \;=\; 1,$$

if $k^* > 0$;

$$Q'(k_0 + k^*)(1 + \mu(e^*))p \;\leq\; 1,$$

if $k^* = 0$. Since both k^* and e^* are functions of k_0 and p, using Envelope Theorem and the two (in)equalities above,

$$Q'(k_0 + k^*)(1 + \mu(e^*))p \;\leq\; 1,$$

holds for all k_0 when the derivative is taken with respect to k_0.

Because $e_0 = 0$ and $\underline{p}$ is the minimum probability of IPO such that E has incentive to exert the incremental effort, $\tilde{p}_s(k_0) \geq \underline{p}$, where $\underline{p}$ is defined in lemma 3. Since $\tilde{p}_s(k_0)$ satisfies

$$\frac{1}{2}\mathcal{S}(k_0, 0; p) - qk_0(1-p) \;=\; 0,$$

it is a differentiable function of k_0 by Implicit Function Theorem.

The solution of (**??**) exists, since k_0 lies in a compact set. If I can show that the first order derivative of the objective function in (**??**) has a positive right limit at $k_0 = 0$, and the first order condition has a unique solution, then k_{0s} is the unique interior solution of (**??**), and the first order condition is sufficient and necessary. Taking the first order derivative of the objective function in (**??**), and setting it equal to 0, we have

$$\int_0^{\tilde{p}_s(k_0)} \Big[Q'(k_0)p + q(1-p)\Big]\rho(p)dp + \\ \int_{\tilde{p}_s(k_0)}^1 \frac{1}{2}\Big[Q'(k_0+k^*)(1+\mu(e^*))p + Q'(k_0)p\Big]\rho(p)dp \;=\; 1.$$

The term involves the derivative of $\tilde{p}_s(k_0)$ is 0 by the choice of $\tilde{p}_s(k_0)$. All terms in the integrands are nonnegative, so

$$\begin{aligned}&\int_0^{\tilde{p}_s(k_0)} \Big[Q'(k_0)p + q(1-p)\Big]\rho(p)dp + \\ &\int_{\tilde{p}_s(k_0)}^1 \frac{1}{2}\Big[Q'(k_0+k^*)(1+\mu(e^*))p + Q'(k_0)p\Big]\rho(p)dp \\ \geq &\frac{1}{2}Q'(k_0)\int_0^1 p\rho(p)dp.\end{aligned}$$

By assumption, $\lim_{k\to 0+} Q'(k) = +\infty$. So the first order derivative of the objective function in (**??**) has a positive right limit at $k_0 = 0$. Now, $Q'(\cdot)$ is strictly monotonically decreasing, and $\lim_{k\to+\infty} Q'(k) = 0$, so the first order condition has a unique solution by Intermediate Value Theorem. This finishes the proof of the claim that $k_{0s} > 0$.

If $k_{0l} = 0$, there is nothing to prove. So assuming $k_{0l} > 0$ and then for $\tilde{p}_l(k_0)$,

$$\int_0^{\tilde{p}_l(k_0)} \Big[Q'(k_0)p+q(1-p)\Big]\rho(p)dp+\int_{\tilde{p}_l(k_0)}^1 Q'(k_0+k^*)(1+\mu(e^*))p\rho(p)dp \;=\; 1. \tag{3.21}$$

This is because it is possible that $\underline{p} > 0$, then $k_{0l} > 0$ since $\tilde{p}_l(k_0) \geq \underline{p}$.

If I can show that

$$\begin{aligned}\int_0^{\tilde{p}_s(k_{0l})} \Big[Q'(k_{0l})p + q(1-p)\Big]\rho(p)dp+ \\ \int_{\tilde{p}_s(k_{0l})}^1 \frac{1}{2}\Big[Q'(k_{0l}+k^*)(1+\mu(e^*))p + Q'(k_{0l})p\Big]\rho(p)dp \;&>\; 1,\end{aligned} \tag{3.22}$$

then it would imply $k_{0s} > k_{0l}$, since all terms in the integrands are less or equal to 1 except for $Q'(\cdot)p$ terms, and $Q'(\cdot)$ is a decreasing function.

Taking the difference between the left hand side of (3.22) and the left hand side of (3.21) with $k_0 = k_{0l}$,

$$\begin{aligned}
&\int_0^{\tilde{p}_s(k_{0l})} \Big[Q'(k_{0l})p + q(1-p)\Big]\rho(p)dp + \\
&\int_{\tilde{p}_s(k_{0l})}^1 \frac{1}{2}\Big[Q'(k_{0l}+k^*)(1+\mu(e^*))p + Q'(k_{0l})p\Big]\rho(p)dp \\
&- \int_0^{\tilde{p}_l(k_{0l})} \Big[Q'(k_{0l})p + q(1-p)\Big]\rho(p)dp - \int_{\tilde{p}_l(k_{0l})}^1 Q'(k_{0l}+k^*)(1+\mu(e^*))p\rho(p)dp \\
= &\int_{\tilde{p}_l(k_{0l})}^{\tilde{p}_s(k_{0l})} \Big[Q'(k_{0l})p + q(1-p) - Q'(k_{0l}+k^*)(1+\mu(e^*))p\Big]\rho(p)dp \\
&+ \int_{\tilde{p}_s(k_{0l})}^1 \frac{1}{2}\Big[Q'(k_{0l})p - Q'(k_{0l}+k^*)(1+\mu(e^*))p\Big]\rho(p)dp.
\end{aligned}$$

Note that $Q'(k_{0l}+k^*)(1+\mu(e^*))p \leq 1$, and $Q'(k_{0l})p + q(1-p) > 1$ as long as $p \geq \tilde{p}_l(k_{0l})$, because $Q'(k_{0l})p + q(1-p)$ is an increasing function in p, and the equation (3.21) holds if $k_{0l} > 0$. By assumption, q is sufficiently small, so $Q'(k_{0l})p > 1$ as long as $p \geq \tilde{p}_s(k_{0l})$. Then this difference is strictly greater than 0. This finishes the proof of the claim that $k_{0s} > k_{0l}$.

For the third claim, want to show that $\tilde{p}_l(k_0)$ is increasing k_0 for $k_0 \geq k_{0l}$. Denote

$$\mathcal{F} \;:=\; \mathcal{S}(k_0, 0; p) - qk_0(1-p) \;=\; Q(k_0+k^*)(1+\mu(e^*))p - k^* - e^* - Q(k_0)p - qk_0(1-p).$$

Then

$$\frac{\partial \mathcal{F}}{\partial p} \;=\; Q(k_0+k^*)(1+\mu(e^*)) - Q(k_0) + qk_0 \;>\; 0,$$

and

$$\frac{\partial \mathcal{F}}{\partial k_0} \;=\; Q'(k_0+k^*)(1+\mu(e^*))p - Q'(k_0)p - q(1-p).$$

So the sign of $d\tilde{p}_l(k_0)/dk_0$ is the opposite of sign of $\partial\mathcal{F}/\partial k_0$. But we know that $\partial\mathcal{F}/\partial k_0$ is negative at $k_0 = k_{0l}$ by the condition (3.21). By assumption 1,

$$\lim_{k_0\to\infty} \frac{\partial\mathcal{F}}{\partial k_0} \;=\; -\,q(1-p) \;<\; 0.$$

And for $k_0 \geq k_{0l}$ sufficiently large, $k^* = 0$, then

$$\frac{\partial^2\mathcal{F}}{\partial k_0^2} \;\approx\; Q''(k_0)\mu(e^*)p \;<\; 0.$$

This implies that $\partial\mathcal{F}/\partial k_0 < 0$ for $k_0 \geq k_{0l}$. So $\tilde{p}_l(k_0)$ is increasing for $k_0 \geq k_{0l}$.

Finally,

$$\begin{aligned}
&\int_0^{\tilde{p}_s(k_{0s})} \Big[Q(k_{0s})p + qk_{0s}(1-p)\Big]\rho(p)dp + \\
&\int_{\tilde{p}_s(k_{0s})}^1 \frac{1}{2}\Big[\mathcal{U}(k_{0s},0;p) + Q(k_{0s})p\Big]\rho(p)dp - k_{0s} \\
\leq &\int_0^{\tilde{p}_s(k_{0s})} \Big[Q(k_{0s})p + qk_{0s}(1-p)\Big]\rho(p)dp + \int_{\tilde{p}_s(k_{0s})}^1 \mathcal{U}(k_{0s},0;p)\rho(p)dp - k_{0s} \\
< &\int_0^{\tilde{p}_l(k_{0s})} \Big[Q(k_{0s})p + qk_{0s}(1-p)\Big]\rho(p)dp + \int_{\tilde{p}_l(k_{0s})}^1 \mathcal{U}(k_{0s},0;p)\rho(p)dp - k_{0s} \\
\leq &\int_0^{\tilde{p}_l(k_{0l})} \Big[Q(k_{0l})p + qk_{0l}(1-p)\Big]\rho(p)dp + \int_{\tilde{p}_l(k_{0l})}^1 \mathcal{U}(k_{0l},0;p)\rho(p)dp - k_{0l}.
\end{aligned}$$

The first inequality is because

$$\mathcal{U}(k_{0s},0;p) \;=\; Q(k_{0s}+k^*)(1+\mu(e^*))p - k^* - e^* \;\geq\; Q(k_{0s})p.$$

The second inequality is because $\tilde{p}_l(k_{0s}) < \tilde{p}_s(k_{0s})$, and $\tilde{p}_l(k_{0s})$ is the solution for long term contract when the initial investment is k_{0s}, so $\mathcal{S}(k_{0s},0;p) >$

$qk_{0s}(1-p)$ on the interval of $(\tilde{p}_l(k_{0s}), \tilde{p}_s(k_{0s}))$. The inequality is strict because $\mathcal{S}(k_0, 0; p)$ is increasing in p. The last inequality is by the definition of k_{0l}. So the optimal contract is the long term contract.

Q.E.D.

3.10.7 Proof of Proposition 13

Suppose the quality of the technology is uncertain, but E's ability is certain and greater than the threshold δ^*. A feasible sequences of short term contracts are composed of an initial contract which coordinates V and E's investment actions up to stage 3, and a continuing contract governing the rest investment process. Using a sequence of short term contracts can destroy value in two ways: $k_{0s} > k_{0l} \geq 0$, the first best investment levels may not be feasible if $\hat{p}_T$ is very low, and $qk_0 < k_0$ in case of liquidation; E will share positive amount of surplus $\mathcal{S}(k_0, 0; p)/2$ if V decides to continue investing with E as the manager. In the mean time, given any initial contract, the state space of x, V and E's investments at stage 2, and the ownership structure based on the realization of signal x are all perfectly predictable. Then, for V, a long term contingent contract on x at stage 1 coordinating the whole investment process would perform no worse than a sequence of short term contracts for each given x.

Q.E.D.

3.10.8 Proof of Proposition 14

Suppose after stage 3, E's existing ownership stake is of proportion α_E. If V decides to remove E, it is (weakly) costly since E is now an owner-manager. In this case, V offers E a package equivalent to a severance payment, so that E is indifferent between staying and leaving, then E will leave.

Let V's offer be composed of a new proportion of ownership stake α'_E and pecuniary payment of s_E. For E to be willing to give up the manager position, the severance payment must satisfy the condition

$$s_E+\alpha'_E Q(k_0+k')(1+\mu(e_0+e'_M))\hat{p}_T p_M = \max_{e'_E\geq 0} \alpha_E Q(k_0)(1+\mu(e_0+e'_E))\hat{p}_T\hat{p}_E - e'_E,$$

where $k' \geq 0$ and $e'_M \geq 0$ are V and M's further input under M's management. The severance payment is nonnegative because E has limited liability, and her participation constraint will be violated if the expected continuation payoff is negative. In that case, E is willing to leave the company without severance payment from V.

Let e^*_E be the optimal solution of the right hand side, and this solution exists on the interval $[0,\infty)$ by assumption 2. Then the condition can be rewritten as

$$s_E+\alpha'_E Q(k_0+k')(1+\mu(e_0+e'_M))\hat{p}_T p_M = \alpha_E Q(k_0)(1+\mu(e_0+e^*_E))\hat{p}_T\hat{p}_E - e^*_E. \tag{3.23}$$

Since M has no bargaining power, V can make take-it-or-leave-it offer to

M at stage 5M, and he solves the following optimization problem.

$$\max_{s_E,\alpha'_E,\omega_M,\alpha_M,k',e'_M} (1-\alpha'_E-\alpha_M)Q(k_0+k')(1+\mu(e_0+e'_M))\hat{p}_T p_M - k' - s_E - \omega_M, \tag{3.24}$$

subject to the constraints that include M's IC constraint

$$\max_{e'_M \geq 0} \omega_M + \alpha_M Q(k_0+k')(1+\mu(e_0+e'_M))\hat{p}_T p_M - e'_M. \tag{3.25}$$

Let

$$e^*_M \in \operatorname*{argmax}_{e'_M \geq 0} \omega_M + \alpha_M Q(k_0+k')(1+\mu(e_0+e'_M))\hat{p}_T p_M - e'_M.$$

The optimization problem is also subject to M's IR constraint

$$\omega_M + \alpha_M Q(k_0+k')(1+\mu(e_0+e^*_M))\hat{p}_T p_M - e^*_M = 0, \tag{3.26}$$

and the condition (3.23).

Suppose $\hat{p} = \hat{p}_T p_M$ is high enough so that it is optimal for V to continue investing. M's IC constraint (3.25) can be replaced by first order condition as

$$\alpha_M Q(k_0+k')\mu'(e_0+e'_M)\hat{p}_T p_M - 1 = 0.$$

So V solves problem (3.24) subject to the constraints (**??**), (3.26), and (3.23).

Substituting s_E, ω_M in (3.24) from constraints (3.26) and (3.23), V solves

$$\max_{\alpha'_E,\alpha_M,k',e'_M} Q(k_0+k')(1+\mu(e_0+e'_M))\hat{p}_T p_M - k' - e'_M - \\ (\alpha_E Q(k_0)(1+\mu(e_0+e^*_E))\hat{p}_T\hat{p}_E - e^*_E)$$

subject to constraint (**??**) and $\alpha_M \in [0, 1-\alpha'_E]$.

It is easy to see that when $\alpha_M = 1$, both M's effort input e'_M and V's investment k' will increase the overall effort and investment levels to the first best, conditioned on all available information. So in E's severance payment, $\alpha'_E = 0$, and $s_E \geq 0$, since

$$s_E = \max_{e'_E \geq 0} \alpha_E Q(k_0)(1+\mu(e_0+e'_E))\hat{p}_T\hat{p}_E - e'_E \geq \alpha_E Q(k_0)(1+\mu(e_0))\hat{p}_T\hat{p}_E \geq 0.$$

Q.E.D.

3.10.9 Proof of Lemma 4

Denote $p = \hat{p}_T\hat{p}_E$ or $p = \hat{p}_T p_M$. Note that $\mathcal{U}(k_0, e_0; p) > Q(k_0)(1+\mu(e_0))p$ almost surely, with the exception that (k_0, e_0) is on the first best investment frontier corresponding to p. Now, $e_0 = 0$. V's expected payoff is

$$\mathcal{U}(k_0, 0; \hat{p}_T p_M), \text{ or } Q(k_0)\hat{p}_T p_M + qk_0(1-\hat{p}_T p_M),$$

if M is the manager;

$$\frac{1}{2}\Big[\mathcal{U}(k_0, 0; \hat{p}_T\hat{p}_E) + Q(k_0)\hat{p}_T\hat{p}_E\Big], \text{ or } Q(k_0)\hat{p}_T\hat{p}_E + qk_0(1-\hat{p}_T\hat{p}_E),$$

if E is the manager.

If liquidation leads to higher interim expected payoff for V, then it is optimal to keep E being the manager as long as $\hat{p}_E \geq p_M$, since $Q(k_0) > qk_0$ by assumption 4.

As $\hat{p}_T$ increases, V requires higher ability of E if E being a manager is desirable. In order to show this, suppose $p_M = \hat{p}_E$, then either

$$\mathcal{U}(k_0, 0; \hat{p}_T p_M) > Q(k_0)\hat{p}_T p_M + qk_0(1 - \hat{p}_T p_M) = Q(k_0)\hat{p}_T\hat{p}_E + qk_0(1 - \hat{p}_T\hat{p}_E)$$

for sufficiently high value of $\hat{p}_T$, or

$$\mathcal{U}(k_0, 0; \hat{p}_T p_M) = \mathcal{U}(k_0, 0; \hat{p}_T\hat{p}_E) > \frac{1}{2}\mathcal{U}(k_0, 0; \hat{p}_T\hat{p}_E) + \frac{1}{2}Q(k_0)\hat{p}_T\hat{p}_E$$

by the inequality $\mathcal{U}(k_0, e_0; p) > Q(k_0)(1 + \mu(e_0))p$ given in the beginning of the proof with $e_0 = 0$. So, E is preferred by V to be a manager only when $\hat{p}_E > p_M$.

V's expected payoff when E is the manager (the right hand side of the inequalities) is strictly increasing in $\hat{p}_E$ by Envelope Theorem. Consider the two equations

$$\mathcal{U}(k_0, 0; \hat{p}_T p_M) = Q(k_0)\hat{p}_T\hat{p}_E + qk_0(1 - \hat{p}_T\hat{p}_E),$$

and

$$\mathcal{U}(k_0, 0; \hat{p}_T p_M) = \frac{1}{2}\mathcal{U}(k_0, 0; \hat{p}_T\hat{p}_E) + \frac{1}{2}Q(k_0)\hat{p}_T\hat{p}_E.$$

Solve for $\hat{p}_E$ as a function of $\hat{p}_T$, then a function of x with k_0 as one of the parameters. The solution exists by Implicit Function Theorem, and denote this solution as $\tilde{p}_{E,s}(k_0, x)$. $\tilde{p}_{E,s}(k_0, x) \geq p_M$. We cannot guarantee the solution is less or equal to one. When it does, it gives the threshold $\tilde{y}_{E,s}(k_0, x)$ by a function similar to the function f^{-1} in the proof of lemma 2; otherwise, set $\tilde{y}_{E,s}(k_0, x)$ equal to $+\infty$. From the proof of lemma 1, $\tilde{y}_{E,s}(k_0, x) \geq y^*$ for any k_0, and x.

Q.E.D.

3.10.10 Proof of Lemma 5

The probability $\tilde{p}_l(k_0)$ in the proof of proposition 11 gives the threshold for $\hat{p}_T p_M$ on the choice of liquidation or investing with M as the manager. The probability $\tilde{p}_s(k_0)$ in the proof of proposition 10 gives the threshold for $\hat{p}_T \hat{p}_E$ on the choice of liquidation or investing with E as the manager. And $\tilde{p}_{E,s}(k_0, x)$ in the proof of lemma 4 gives the threshold for $\hat{p}_E$ on the choice of manager E or M. Combining all these boundary conditions, we have the Figure 3.5.

Q.E.D.

3.10.11 Proof of Proposition 15

The time line of the game is stage 1, 2, 3, 4; 5E, 6E, and 7E; or 5M, 6M, and 7M. Negotiations and contracting will be conducted at stage 1, and 5E, or 5M. Investments will be made simultaneously by both V and E at stage 2 and 6E, or V and M at stage 6M. Signal x and y will be realized at stage

3. All uncertainties of the technology, the managers' abilities, and IPO will be resolved at stage 7E, or 7M.

The posterior probability of IPO perceived by V, E, and M from stage 3 on is $p = \hat{p}_T \hat{p}_E$ ($p = \hat{p}_T p_M$), and the stage 1 prior probability distribution of p is given by $\rho = \rho_T \rho_E$ ($\rho = \rho_T$, no signal to be observed for M) since the technology, E's ability, and the signals are assumed to be independent to each other.

Since the contracts are short term, let α_V, α_E be the ownership structure after investments k_0, e_0, and the signal x and y. If V decides to continue investing at stage 4 with E as the manager, and if V, E's contracting and investing behaviors will be optimal at stage 5E, 6E, then V's interim expected payoff is

$$\begin{aligned} u_V &= d_V + \frac{1}{2}\mathcal{S} \\ &= \alpha_V Q(k_0)(1+\mu(e_0))\hat{p}_T\hat{p}_E + \frac{1}{2}\mathcal{S}(k_0, e_0; \hat{p}_T\hat{p}_E) \\ &= \frac{1}{2}\mathcal{U}(k_0, e_0; \hat{p}_T\hat{p}_E) + (\alpha_V - \frac{1}{2})Q(k_0)(1+\mu(e_0))\hat{p}_T\hat{p}_E. \end{aligned}$$

Then V would choose α_V as high as possible at stage 1. If V decides to continue investing at stage 4 with M as the manager, then V's maximal interim expectation of final payoff is

$$\mathcal{U}(k_0, e_0; \hat{p}_T p_M) - \Big[\max_{e'_E \geq 0} \alpha_E Q(k_0)(1+\mu(e_0+e'_E))\hat{p}_T\hat{p}_E - e'_E\Big],$$

which would be maximized when $\alpha_V = 1$ so that $\alpha_E = 0$. Additionally, V will have liquidation choice if $\alpha_V = 1$. So $\alpha_V = 1$ is optimal in stage 1 short term contract.

The proof for the claim that E will not exert effort, $e_0 = 0$, at stage 2 is exactly the same as the one in the proof of the proposition 10, with some necessary changes: $p = \hat{p}_T \hat{p}_E$, $\rho = \rho_T \rho_E$, and I being a region in the signal space $\mathcal{X} \times \mathcal{Y}$ instead of an interval.

Given $\alpha_V = 1$, $e_0 = 0$ after stage 3, we can apply the results in lemma 5. V makes decisions according to the allocations of the signal realizations to the regions $\Pi_{L,E}$, $\Pi_{L,M}$, $\Pi_{I,E}$, and $\Pi_{I,M}$. Since $\alpha_E = 0$, replacing E by M is costless. If V decides to continue investing, he will transfer all the ownership to the manager. If V chooses liquidation, he will retain all the ownership and then be able to liquidate the company at the last stage in failure of going public.

By backward induction, V then solves the optimal initial investment level k_{0s}, and it is the solution of

$$\begin{aligned}
\max_{k_0} \Big\{ \int_{\Pi_{L,E}} & \Big[Q(k_0)\hat{p}_T\hat{p}_E + qk_0(1-\hat{p}_T\hat{p}_E)\Big]\rho(p)dp \\
&+ \int_{\Pi_{L,M}} \Big[Q(k_0)\hat{p}_T p_M + qk_0(1-\hat{p}_T p_M)\Big]\rho(p)dp \\
&+ \int_{\Pi_{I,E}} \frac{1}{2}\Big[\mathcal{U}(k_0, 0; \hat{p}_T\hat{p}_E) + Q(k_0)\hat{p}_T\hat{p}_E\Big]\rho(p)dp \\
&\qquad\qquad + \int_{\Pi_{I,M}} \mathcal{U}(k_0, 0; \hat{p}_T p_M)\rho(p)dp - k_0\Big\}.
\end{aligned}$$

The solution always exists since k_0 lies in a closed interval bounded by 0 and the maximal investment level of the first best investment frontier, which is compact. Then $y^*_{E,s}(x) = \tilde{y}_{E,s}(k_{0s}, x)$ for the optimal k_{0s}.

Q.E.D.

3.10.12 Proof of Proposition 16

The renegotiation of existing long term contracts can only happen at stage 5E. Given any long term contract, any pair of initial investment k_0, e_0, and any signal realizations x, y, E's interim expected payoff at stage 5E can be either nonnegative or strictly negative. In the former case, E actually has all the expected surplus, and she won't accept any other new contract proposed by V. In the latter, E is protected by limited liability, so she can quit the manager's position and have outside reservation utility of at least zero by assumption. By quitting, E can receive a payoff of at least as good as the one generated by accepting V's take-it-or-leave-it offer. However, V will renegotiate the contract only when E is perceived to be better manager than M, $\hat{p}_E > p_M$. V has to recruit M as manager if E rejects V's offer, but this leads to lower expected payoff for V, since E is supported by higher perceived ability. So, if V and E have equal opportunities proposing a new contract, we conclude that V and E will share the surplus equally as predicted by the Rubinstein's bargaining model.

Q.E.D.

3.10.13 Proof of Proposition 17

Using backward induction, given the initial investments k_0, e_0 at stage 2, and signal realizations x, y at stage 3, suppose the initial contract is contingent on signal x with components $\alpha_V(x)$, $\alpha_E(x)$, and $\omega_E(x)$.

Since $\omega_E(x)$ decides the share of the surplus $\mathcal{U}(k_0, e_0; \hat{p}_T\hat{p}_E)$, which also depends on the signal realization y. If $\omega_E(x)$ grants V the exact surplus corresponding to a signal $\bar{y}$, then V will lose all the extra surplus when $y > \bar{y}$, since there is no alternative contract which Pareto improves both V and E's payoffs. By proposition 16, V and E renegotiate the existing contract when $y \leq \bar{y}$, if V plans to keep E. So, $\omega_E(x)$ is chosen to allocate $\mathcal{U}(k_0, e_0; \hat{p}_T)$ to V as if $\hat{p}_E = 1$.

Next, find the optimal $\alpha_V(x)$ and $\alpha_E(x)$. When x is low, it is optimal for V to retain all ownership, then $\alpha_V(x) = 1$ for such x. As x increases, V considers whether to transfer ownership to E. However, y is random and uncorrelated to x. When y is low enough so that choosing M as the manager is optimal, V needs to repurchase the shares from E at cost

$$\max_{e'_E} \alpha_E(x)Q(k_0)(1 + \mu(e_0 + e'_E))\hat{p}_T\hat{p}_E - e'_E$$

by proposition 14. To minimize this cost, V can either reduces k_0, or reduces $\alpha_E(x)$. Since $\lim_{k\to 0+} Q'(k) = +\infty$ and renegotiation is inevitable, $k_0 > 0$. So $\alpha_E(x) = 0$ for any x. Then $\omega_E(x) = 0$ to satisfy E's initial individual rationality constraint. The long term contract is degenerated and not contingent on signal x.

Q.E.D.

3.10.14 Proof of Proposition 19

To show that $k_{0s} > 0$, it is sufficient to show that the area of $\Pi_{I,E}$ has a strictly positive lower bound. We do not need to consider the region $\Pi_{L,E}$

because this region will no longer exist if $k_0 = 0$, where $\tilde{p}_s(0) = \tilde{p}_l(0) = 0$. The condition is sufficient because V's interim expected payoff in this region is

$$\frac{1}{2}\Big[\mathcal{U}(k_0, 0; p) + Q(k_0)p\Big],$$

and $\lim_{k \to 0+} Q'(k) = +\infty$.

Suppose $p_M < \tilde{p}_M$, then there exist a constant $\varepsilon > 0$ such that $p_M \leq \tilde{p}_M - \varepsilon$. The requirement that $p_M < 1/2$ is not a strong assumption, it is true as long as the IPO threshold $\delta^* > 0$ which is reasonable.

Now consider the segment of the function $\tilde{p}_{E,s}(k_0, x)$ which separate the regions $\Pi_{I,E}$ and $\Pi_{I,M}$. It is the solution of $\hat{p}_E$ as an implicit function of $\hat{p}_T$ given by the equation

$$\mathcal{U}(k_0, 0; \hat{p}_T p_M) = \frac{1}{2}\mathcal{U}(k_0, 0; \hat{p}_T \hat{p}_E) + \frac{1}{2}Q(k_0)\hat{p}_T \hat{p}_E.$$

For any k_0 and x (which corresponds to $\hat{p}_T > 0$), setting $p_M \to 0$ on the left hand side of the equation, then the left hand side goes to 0, while the right hand side remains strictly positive. And $\mathcal{U}(k_0, 0; \hat{p}_T p_M)$ is increasing in p_M. This justifies that the set $\mathcal{M}(k_0, x)$ is nonempty. Since $p_M < \tilde{p}_M$ by assumption, we can apply Implicit Function Theorem for any k_0 and x.

Define

$$\mathcal{F} := 2\,\mathcal{U}(k_0, 0; \hat{p}_T p_M) - \mathcal{U}(k_0, 0; \hat{p}_T \hat{p}_E) - Q(k_0)\hat{p}_T \hat{p}_E.$$

Let k_M, e_M be the solutions for $\mathcal{U}(k_0, 0; \hat{p}_T p_M)$, that is,

$$\mathcal{U}(k_0, 0; \hat{p}_T p_M) = Q(k_0 + k_M)(1 + \mu(e_M))\hat{p}_T p_M - k_M - e_M.$$

Similarly, Let k_E, e_E be the solutions for $\mathcal{U}(k_0, 0; \hat{p}_T \hat{p}_E)$. Using the Envelope Theorem,

$$\frac{\partial \mathcal{F}}{\partial \hat{p}_T} = 2Q(k_0 + k_M)(1 + \mu(e_M))p_M - Q(k_0 + k_E)(1 + \mu(e_E))\hat{p}_E - Q(k_0)\hat{p}_E;$$

$$\frac{\partial \mathcal{F}}{\partial \hat{p}_E} = -Q(k_0 + k_E)(1 + \mu(e_E))\hat{p}_T - Q(k_0)\hat{p}_T.$$

So the sign of $d\hat{p}_E/d\hat{p}_T$ is the same as the sign of $\partial\mathcal{F}/\partial\hat{p}_T$.

If $\hat{p}_E > 2p_M$, then $k_E \geq k_M$, $e_E \geq e_M$, so $\partial\mathcal{F}/\partial\hat{p}_T < 0$, then $d\hat{p}_E/d\hat{p}_T < 0$. If I can show $\hat{p}_E \leq 2p_M$ at the initial point of the segment $\tilde{p}_{E,s}(k_0, x)$ which separates the regions $\Pi_{I,E}$ and $\Pi_{I,M}$, then $\tilde{p}_{E,s}(k_0, x) \leq 1 - 2\varepsilon$, since $p_M \leq \tilde{p}_M - \varepsilon \leq 1/2 - \varepsilon$, and $d\hat{p}_E/d\hat{p}_T < 0$ once $\hat{p}_E \geq 2p_M - \tau$ for some $\tau > 0$ sufficiently small.

At the initial point, $\hat{p}_T$ and $\hat{p}_E$ satisfy

$$\begin{aligned}
&Q(k_0 + k_M)(1 + \mu(e_M))\hat{p}_T p_M - k_M - e_M \\
&= \frac{1}{2}\Big[Q(k_0 + k_E)(1 + \mu(e_E))\hat{p}_T\hat{p}_E - k_E - e_E + Q(k_0)\hat{p}_T\hat{p}_E\Big] \\
&= Q(k_0)\hat{p}_T\hat{p}_E + qk_0(1 - \hat{p}_T\hat{p}_E).
\end{aligned}$$

Then subtract $Q(k_0)\hat{p}_T\hat{p}_E$ from each equation above:

$$\begin{aligned}
& Q(k_0+k_M)(1+\mu(e_M))\hat{p}_T p_M - k_M - e_M - Q(k_0)\hat{p}_T\hat{p}_E \\
&= qk_0(1-\hat{p}_T\hat{p}_E) \\
&= \frac{1}{2}\Big[Q(k_0+k_E)(1+\mu(e_E))\hat{p}_T\hat{p}_E - k_E - e_E - Q(k_0)\hat{p}_T\hat{p}_E\Big].
\end{aligned}$$

Note that $\tilde{p}_l(k_0)$ is the solution for p from the first equality if we replace $Q(k_0)\hat{p}_T\hat{p}_E$ with $Q(k_0)\hat{p}_T p_M$, and $\tilde{p}_s(k_0)$ is the solution for p from the second equality. If I can show that $\tilde{p}_s(k_0) \leq 2\tilde{p}_l(k_0)$, then $\hat{p}_T\hat{p}_E \leq 2\hat{p}_T p_M$, and finally $\hat{p}_E \leq 2p_M$, since $\hat{p}_T\hat{p}_E = \tilde{p}_s(k_0)$ and $\hat{p}_T p_M \geq \tilde{p}_l(k_0)$. The latter inequality is because $\hat{p}_E = \tilde{p}_{E,s}(k_0, x) \geq p_M$ at that point, and $\hat{p}_T p_M$ has to be greater or equal to $\tilde{p}_l(k_0)$ for the equality to hold.

Now to show $\tilde{p}_s(k_0) \leq 2\tilde{p}_l(k_0)$ (a geometric proof). Suppose $k_0 \neq 0$. By definition, $\tilde{p}_l(k_0)$ is the solution of

$$qk_0(1-p) \;=\; \mathcal{S}(k_0, 0; p);$$

and $\tilde{p}_s(k_0)$ is the solution of

$$qk_0(1-p) \;=\; \frac{1}{2}\mathcal{S}(k_0, 0; p).$$

By Envelope Theorem, simply taking second order derivative will show that $\mathcal{S}(k_0, 0; p)$ is a convex function in p. Let $\mathcal{L}_{l,1}$ be the straight line tangent to $\mathcal{S}(k_0, 0; p)$ at the point $\tilde{p}_l(k_0)$. $\mathcal{L}_{l,1}$ also crosses $qk_0(1-p)$ at the same point. Let $\mathcal{L}_{s,1}$ be the straight line tangent to $\mathcal{S}(k_0, 0; p)$ at the point $\tilde{p}_s(k_0)$.

We know that $\tilde{p}_s(k_0) > \tilde{p}_l(k_0)$, and the function $\mathcal{S}(k_0, 0; p)$ is convex, so the intersection of $\mathcal{L}_{l,1}$ and $\mathcal{L}_{s,1}$ is strictly between $\tilde{p}_l(k_0)$ and $\tilde{p}_s(k_0)$. Let $\mathcal{L}_{l,1/2}$ and $\mathcal{L}_{s,1/2}$ be the straight lines scaled down from $\mathcal{L}_{l,1}$ and $\mathcal{L}_{s,1}$ by $1/2$. $\mathcal{L}_{s,1/2}$ is in fact tangent to $\mathcal{S}(k_0, 0; p)/2$ at the point $\tilde{p}_s(k_0)$, and crosses $qk_0(1-p)$ at the same point. The intersection of $\mathcal{L}_{l,1/2}$ and $\mathcal{L}_{s,1/2}$ has the same p value as the intersection of $\mathcal{L}_{l,1}$ and $\mathcal{L}_{s,1}$ which is strictly less than $\tilde{p}_s(k_0)$. But $\mathcal{L}_{l,1/2}$ has a smaller slope than $\mathcal{L}_{s,1/2}$ does since $\mathcal{S}(k_0, 0; p)/2$ is a convex function in p. Denote the point where $\mathcal{L}_{l,1/2}$ crosses $qk_0(1-p)$ as $\tilde{p}_{l,1/2}$. Then $\tilde{p}_{l,1/2} > \tilde{p}_s(k_0)$, because $qk_0(1-p)$ is decreasing in p, and the line $\mathcal{L}_{l,1/2}$ is below $\mathcal{S}(k_0, 0; p)/2$.

Simple algebra shows that $2\tilde{p}_l(k_0) > \tilde{p}_{l,1/2}$. To see this, denote $\mathcal{L}_{l,1} := \zeta p - \xi$, with ζ, $\xi > 0$, since $\mathcal{S}(k_0, 0; 0) = 0$. Then $\mathcal{L}_{l,1/2} := (\zeta p - \xi)/2$. Solve

$$qk_0(1-p) \;=\; \zeta p - \xi, \qquad p_1 \;=\; \frac{qk_0 + \xi}{qk_0 + \zeta};$$

Solve

$$qk_0(1-p) \;=\; \frac{1}{2}\zeta p - \frac{1}{2}\xi, \qquad p_2 \;=\; \frac{2qk_0 + \xi}{2qk_0 + \zeta}.$$

So

$$2p_1 \;=\; \frac{2qk_0 + 2\xi}{qk_0 + \zeta} \;>\; p_2 \;=\; \frac{2qk_0 + \xi}{2qk_0 + \zeta}.$$

Combine these two inequalities, $\tilde{p}_s(k_0) < 2\tilde{p}_l(k_0)$. Then $\tilde{p}_s(k_0) \le 2\tilde{p}_l(k_0)$ in general for any $k_0 \ge 0$. This finishes the proof of this claim.

Next is to show that $\tilde{p}_s(k_0)$ has an upper bound strictly less than 1. By the definition of $\tilde{p}_s(k_0)$, it is the solution of the equation

$$Q(k_0)p + qk_0(1-p) \;=\; \frac{1}{2}\Big[Q(k_0+k_E)(1+\mu(e_E))p - k_E - e_E + Q(k_0)p\Big].$$

Let $\overline{k}$ be the maximal investment level on the first best investment frontier, then $\tilde{p}_s(k_0)$ is bounded above by

$$Q(\overline{k})p + q\overline{k}(1-p) \;=\; \frac{1}{2}\Big[Q(\overline{k})(1+\mu(e_E))p - e_E + Q(\overline{k})p\Big],$$

or simplified as $2q\overline{k}(1-p) = Q(\overline{k})\mu(e_E)p - e_E$. So $\tilde{p}_s(k_0) \le \tilde{p}_s(\overline{k})$. Since $\mu(\cdot)$ is sufficiently productive such that the first best investment frontier is nonempty, $e_E > 0$ for $k_0 = \overline{k}$, which corresponds to $p = 1$. But the left hand side is 0 when $p = 1$, so $\tilde{p}_s(\overline{k}) < 1$. This finishes the proof that $\tilde{p}_s(k_0)$ has an upper bound strictly less than 1.

So the area of the region $\Pi_{I,E}$ has a lower bound $2\varepsilon(1 - \tilde{p}_s(\overline{k}))$, and by the argument in the beginning of this section, $k_{0s} > 0$.

Q.E.D.

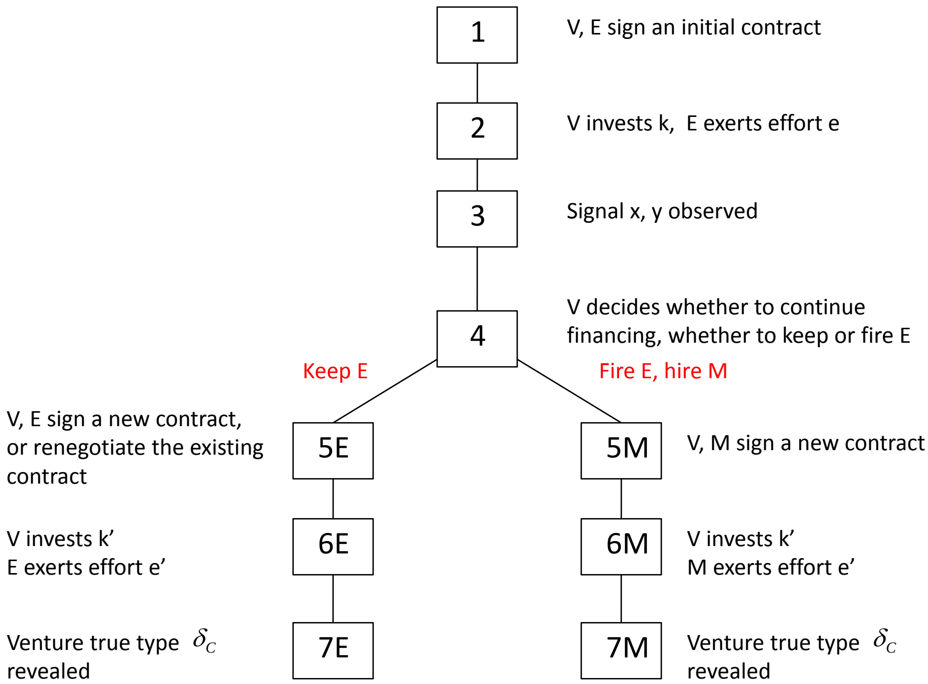

Figure 3.1: Extensive Form of the Game

3.11 Illustration

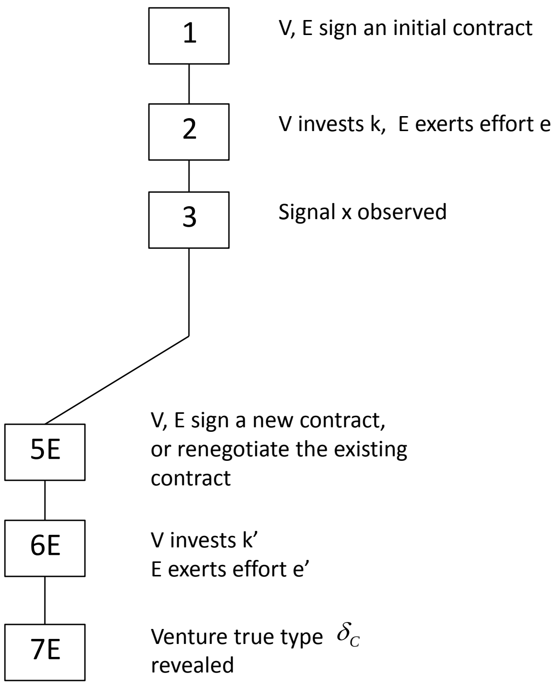

Figure 3.2: The Benchmark Model

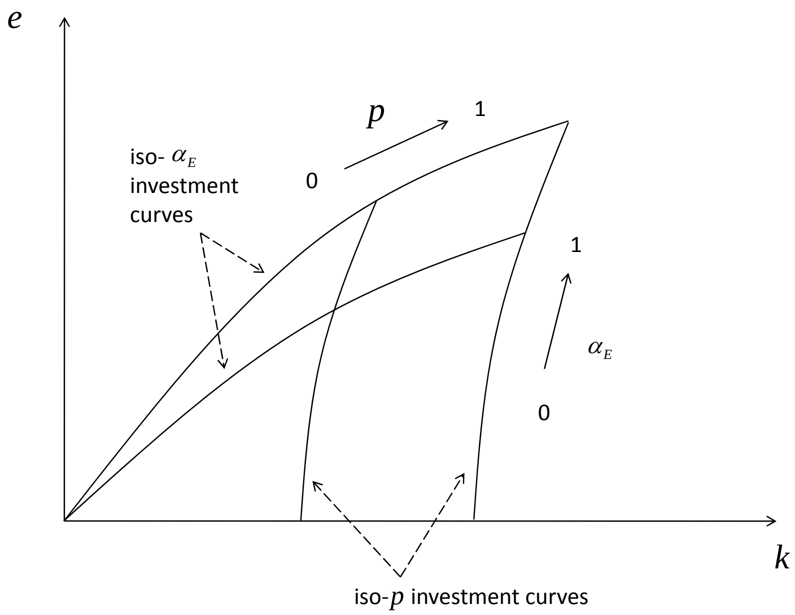

Figure 3.3: Investment Curves

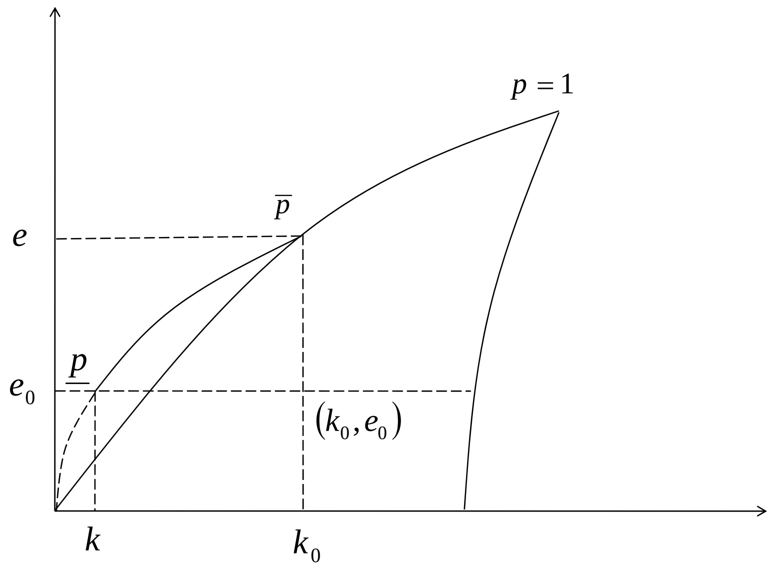

Figure 3.4: Incremental Investments

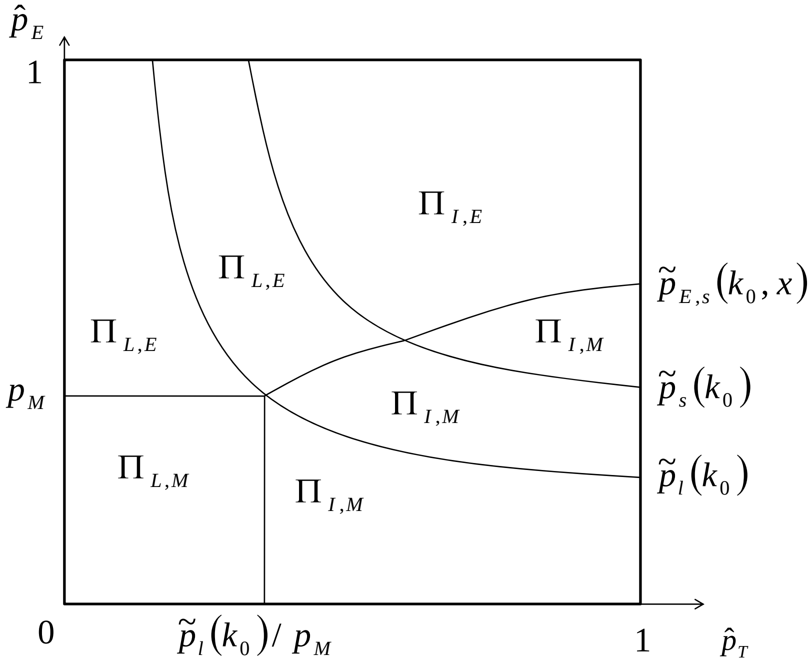

Figure 3.5: Decision Regions

Bibliography

[1] Acharya, Viral V., Moritz Hahn, and Conor Kehoe, 2009, "Corporate Governance and Value Creation: Evidence from Private Equity," Working Paper, New York University.

[2] Berry, Steve, James Levinsohn, and Ariel Pakes, 1995, "Automobile prices in market equilibrium," *Econometrica*, 63, 841890.

[3] Boone, Audra L., and J. Harold Mulherin, 2009, "Do Private Equity Consortiums Facilitate Collusion in Takeover Bidding?" SSRN Working Paper Series.

[4] Cramton, Peter, and Alan Schwartz, 1991, "Using Auction Theory to Inform Takeover Regulation," *The Journal of Law, Economics, & Organization*, V7 N1, 27-53.

[5] DeAngelo, Harry, Linda DeAngelo, and Edward M. Rice, 1984, "Going Private: Minority Freezeouts and Stockholder Wealth," *Journal of Law & Economics*, 27, 367-401.

[6] DeGeorge, Francois, and Richard Zeckhauser, 1993, "The Reverse LBO Decision and Firm Performance: Theory and Evidence," *The Journal of Finance*, 48, 1323-1348.

[7] Fox, Jeremy T., 2008, "Estimating Matching Games with Transfers," NBER Working Paper No. W14382.

[8] Gompers, Paul A., 1996, "Grandstanding in the Venture Capital Industry," *Journal of Financial Economics*, 42, 133-156.

[9] Gompers, Paul, and Josh Lerner, 1999, "An Analysis of Compensation in the U.S. Venture Capital Partnership," *Journal of Financial Economics*, 51, 3-44.

[10] Gompers, Paul, and Josh Lerner, 2000, "Money Chasing Deals? The Impact of Fund Inflows on Private Equity Valuation," *Journal of Financial Economics*, 55, 281-325.

[11] Guo, Shourun, Edith S. Hotchkiss, and Weihong Song, 2008, "Do Buyouts (Still) Create Value?" SSRN Working Paper Series.

[12] Hochberg, Yael V., Alexander Ljungqvist, and Yang Lu, 2007, "Whom You Know Matters: Venture Capital Networks and Investment Performance," *The Journal of Finance*, 62, 251-302.

[13] Holthausen, Robert W., and David F. Larcker, 1996, "The Financial Performance of Reverse Leveraged Buyouts," *Journal of Financial Economics*, 42, 293-332.

[14] Ivashina, Victoria, and Anna Kovner, 2008, "The Private Equity Advantage: Leveraged Buyout Firms and Relationship Banking," SSRN Working Paper Series.

[15] Jegadeesh, Narasimhan, Roman Kraeussl, and Joshua Matthew Pollet, 2009, "Risk and Expected Returns of Private Equity Investments: Evidence Based on Market Prices," SSRN Working Paper Series.

[16] Jones, Charles, and Matthew Rhodes-Kropf, 2003, "The Price of Diversifyable Risk in Venture Capital and Private Equity," SSRN Working Paper Series.

[17] Kaplan, Steven N., 1989, "Management Buyouts: Evidence on Taxes as a Source of Value," *The Journal of Finance*, 44, 611-632.

[18] Kaplan, Steven N., 1989, "The Effects of Management Buyouts on Operating Performance and Value," *Journal of Financial Economics*, 24, 217-254.

[19] Kaplan, Steven N., and Antoinette Schoar, 2005, "Private Equity Performance: Returns, Persistence, and Capital Flows," *The Journal of Finance*, 60, 1791-1823.

[20] Kaplan, Steven N., and Jeremy C. Stein, 1993, "The Evolution of Buyout Pricing and Financial Structure," NBER Working Paper No. W3695.

[21] Kaplan, Steven N., and Per Strömberg, 2008, "Leveraged Buyouts and Private Equity," NBER Working Paper No. W14207.

[22] Lehn, Kenneth, and Annette Poulsen, 1989, "Free Cash Flow and Stockholder Gains in Going Private Transactions," *The Journal of Finance*, 44, 771-787.

[23] Ljungqvist, Alexander, and Matthew Richardson, 2003, "The Cash Flow, Return and Risk Characteristics of Private Equity ," SSRN Working Paper Series.

[24] Ljungqvist, Alexander, Matthew Richardson, and Daniel Wolfenzon, 2007, "The Investment Behavior of Buyout Funds: Theory and Evidence," ECGI Working Paper Series in Finance No. 174/2007.

[25] Massey, Frank J., Jr., 1951, "The Kolmogorov-Smirnov Test for Goodness of Fit," *Journal of the American Statistical Association*, 46, 68-78.

[26] Myerson, Roger B., 1997, "Game Theory: Analysis of Conflict," Harvard University Press.

[27] Officer, M. S., O. Ozbas, and B. A. Sensoy, 2008, "Club Deals in Leveraged Buyouts," Marshall School of Business Working Paper No. MKT 10-08.

[28] Phalippou, Ludovic, and Oliver Gottschalg, 2009, "The Performance of Private Equity Funds," *Review of Financial Studies*, 22(4), 1747-1776.

[29] Povel, Paul and Rajdeep Singh, 2009, "Stapled Finance," SSRN Working Paper Series.

[30] Roth, Alvin E., 1991, "A Natural Experiment in the Organization of Entry-Level Labor Markets: Regional Markets for New Physicians and

Surgeons in the United Kingdom," *The American Economic Review*, 81, 415-440.

[31] Smith, Abbie J., 1990, "Corporate Ownership Structure and Performance," *Journal of Financial Economics*, 27, 143-164.

[32] Sørensen, Morten, 2007, "How Smart is Smart Money? A Two-Sided Matching Model of Venture Capital," *The Journal of Finance*, 62, 2725-2762.

[33] Stephens, Michael A., 1974, "EDF Statistics for Goodness of Fit and Some Comparisons," *Journal of the American Statistical Association*, 69, 730-737.

[34] Storn, Rainer, and Kenneth Price, 1997, "Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces," *Journal of Global Optimization*, 115, 341359.

[35] Bolton, P., and C. Harris, 1996, "Strategic Experimentation: A Revision", Discussion Paper 9627, CentER, Tilburg University.

[36] Bolton, P., and C. Harris, 1999, "Strategic Experimentation", *Econometrica*, 67, 349-374.

[37] Davis, M. H. A., 1977, *Linear Estimation and Stochastic Control (Chapman & Hall Mathematics Series)*, A Halsted Press Book, John Wiley & Sons, New York.

[38] Fleming, W. H., and R. W. Rishel, 1982, *Deterministic and Stochastic Optimal Control (Stochastic Modelling and Applied Probability)*, Springer-Verlag.

[39] Green, E. J., and R. H. Porter, 1984, "Noncooperative Collusion Under Imperfect Price Information", *Econometrica*, 52, 87-100.

[40] Kallianpur, G., 1980, *Stochastic Filtering Theory (Applications of Mathematics)*, Springer-Verlag.

[41] Kandori, M., 2002, "Introduction to Repeated Games with Private Monitoring", *Journal of Economics Theory*, 102, 1-15.

[42] Karatzas, I., and S. E. Shreve, 1991, *Brownian Motion and Stochastic Calculus*, Springer-Verlag.

[43] Keller, G., S. Rady, and M. Cripps, 2005, "Strategic Experimentation with Exponential Bandits", *Econometrica*, 73, 39-68.

[44] Morris, S., and H. S. Shin, 2002, "Social Value of Public Information", *The American Economic Review*, 92, 1521-1534.

[45] Moscarini, G., and L. Smith, 2001, "The Optimal Level of Experimentation", *Econometrica*, 69, 1629-1644.

[46] Øksendal, B., 2007, *Stochastic Differential Equations: An Introduction with Applications (Universitext)*, Springer-Verlag.

[47] Sannikov, Y., 2007, "Games with Imperfectly Observable Actions in Continuous Time", *Econometrica*, 75, 1285-1329.

[48] Stigler, G. J., 1964, "A Theory of Oligopoly", *The Journal of Political Economy*, 72, 44-61.

[49] Townsend, R. M., 1983, "Forecasting the Forecasts of Others", *The Journal of Political Economy*, 91, 546-588.

[50] Yong, J., and X.-Y. Zhou, 1999, *Stochastic Controls: Hamiltonian Systems and HJB Equations (Applications of Mathematics)*, Springer-Verlag.

[51] Admati, Anat R., Paul Pfleiderer, 1994, "Robust Financial Contracting and the Role of Venture Capitalists", *The Journal of Finance*, 49, 371-402.

[52] Baker, George, Robert Gibbons, and Kevin J. Murphy, 2002, "Relational Contracts and the Theory of the Firm", *Quarterly Journal of Economics*, 117, 39-84.

[53] Binmore, Ken, Ariel Rubinstein, and Asher Wolinsky, 1986, "The Nash Bargaining Solution in Economic Modelling", *RAND Journal of Economics*, 17, 176-188.

[54] Bolton, Patrick, David S. Scharfstein, 1990, "A Theory of Predation Based on Agency Problems in Financial Contracting", *The American Economic Review*, 80, 93-106.

[55] Chan, Yuk-Shee, Daniel Siegel, and Anjan V. Thakor, 1990, "Learning, Corporate Control and Performance Requirements in Venture Capital Contracts", *The International Economic Review*, 31, 365-381.

[56] Che, Yeon-Koo, and József Sákovics, 2004, "A Dynamic Theory of Holdup", *Econometrica*, 72, 1063-1103.

[57] Cornelli, Francesca, and Oved Yosha, 2003, "Stage Financing and the Role of Convertible Securities", *Review of Economic Studies*, 70, 1-32.

[58] Fudenberg, Drew, Bengt Holmstrom, and Paul Milgrom, 1990, "Short-Term Contracts and Long-Term Agency Relationships", *Journal of Economic Theory*, 51, 1-31.

[59] Gompers, Paul A., 1995, "Optimal Investment, Monitoring, and the Staging of Venture Capital", *The Journal of Finance*, 50, 1461-1489.

[60] Hellmann, Thomas F., 1998, "The Allocation of Control Rights in Venture Capital Contracts", *RAND Journal of Economics*, 29, 57-76.

[61] Hellmann, Thomas F., 2006, "IPOs, Acquisitions and the Use of Convertible Securities in Venture Capital", *Journal of Financial Economics*, 81, 649-679.

[62] Hermalin, Benjamin E., and Michael L. Katz, 1991, "Moral Hazard and Verifiability: the Effects of Renegotiation in Agency", *Econometrica*, 59, 1735-1753.

[63] Hermalin, Benjamin E., and Michael S. Weisbach, 1998, "Endogenously Chosen Boards of Directors and Their Monitoring of the CEO", *The American Economic Review*, 88, 96-118.

[64] Holmstrom, Bengt, 1999, "Managerial Incentive Problems: A Dynamic Perspective", *Review of Economic Studies*, 66, 169-182.

[65] Holmstrom, Bengt, and Paul Milgrom, 1991, "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", *Journal of Law, Economics, & Organization*, 7, 24-52.

[66] Kaplan, Steven N., and Per Strömberg, 2003, "Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts", *Review of Economic Studies*, 70, 281-315.

[67] Kaplan, Steven N., and Per Strömberg, 2004, "Characteristics, Contracts, and Actions: Evidence from Venture Capitalist Analyses", *The Journal of Finance*, 59, 2177-2210.

[68] Lerner, Josh, 1994, "Venture Capitalists and the Decision to Go Public", *Journal of Financial Economics*, 35, 293-316.

[69] Lerner, Josh, 1995, "Venture Capitalists and the Oversight of Private Firms", *The Journal of Finance*, 50, 301-318.

[70] Neher, Darwin V., 1999, "Stage Financing: An Agency Perspective", *Review of Economic Studies*, 66, 255-274.

[71] Puri, Manju, and Rebecca Zarutskie, 2008, "On the Lifecycle Dynamics of Venture-Capital- and Non-Venture-Capital-Financed Firms", EFA 2007 Ljubljana Meetings Paper, available at SSRN: http://ssrn.com/abstract=967841.

[72] Sahlman, William A., 1990, "The Structure and Governance of Venture-Capital Organizations", *Journal of Financial Economics*, 27, 473-521.