Managerial Incentives and the Role of Advisors in the Continuous-Time Agency Model*

Keiichi Hori† **and Hiroshi Osano**‡

March 8, 2012

Revised: September 25, 2012

Second Revised: February 20, 2013

Final Revised: April 24, 2013

*The authors would like to thank Eric Chou, Yuichi Fukuta, Nobuhiko Hibara, Shinsuke Ikeda, Shingo Ishiguro, Takayuki Ogawa, Yoshiaki Ogura, Jie Qin, Reijiro Samura, Luca Taschini, the editor (Pietro Veronesi), an anonymous referee, and seminar participants at the University of Tokyo, Osaka University, Ritsumeikan University, the 3rd Florence–Ritsumeikan Workshop on "Modeling and pricing finance and insurance assets in a risk management perspective", the 2011 Asian Meeting of the Econometric Society, and the 5th Japan–Taiwan Contract Theory Conference for helpful comments and suggestions on an earlier version of this paper. This work was supported from the Japan Securities Scholarship Foundation.
Correspondence to: Hiroshi Osano, Institute of Economic Research, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan.
E-mail: osano@kier.kyoto-u.ac.jp; **Tel:** +81-75-753-7131; **Fax:** +81-75-753-7138.

†Faculty of Economics, Ritsumeikan University (khori@ec.ritsumei.ac.jp).

‡Institute of Economic Research, Kyoto University (osano@kier.kyoto-u.ac.jp).

Managerial Incentives and the Role of Advisors in the Continuous-Time Agency Model

Abstract

We explore a continuous-time agency model with double moral hazard. Using a venture capitalist (VC)–entrepreneur relationship where the VC both supplies costly effort and chooses the optimal timing of the initial public offering (IPO), we show that optimal IPO timing is earlier under double moral hazard than under single moral hazard. Our results also indicate that the manager's compensation tends to be paid earlier under double moral hazard. We derive several comparative static results, notably that IPO timing is earlier when the need for monitoring by the VC is smaller and when the volatility of cash flows is larger.

JEL Classification Codes: D82, D86, G24, G34, M12, M51.

Keywords: two-sided moral hazard, IPO timing, managerial compensation, dynamic incentives, spin-offs.

1. Introduction

This paper examines the situation in which the joint provision of costly effort by a manager and an advisor initially improves the productivity of a project in a firm. However, after a substantial upward shift in productivity is achieved through irreversible investment, the role of the advisor ends, and only the manager remains to provide ongoing costly effort to improve the productivity of the project. Because the timing of investment corresponds with the timing of the change in the role of the advisor, optimal timing of investment needs to consider the effect of this change. A difficulty may therefore arise if the manager's costly effort is unobservable. In this case, not only the advisor's costly effort but also the manager's compensation may be controlled so as to provide the manager with appropriate incentives. In addition, the advisor's costly effort may be unobservable. Consequently, we must consider the provision of incentives for effort for both the manager and the advisor when solving the optimal timing problem of investment.

This particular situation typically arises when a venture capitalist (VC) exerts substantial effort in monitoring managerial activity, providing managerial advice to entrepreneurs, and choosing the optimal timing of both the initial public offering (IPO) and any investment financed by the IPO. Indeed, the monitoring and advisory role of the VC through ongoing long-term involvement can increase the value of its portfolio companies.[1] The VC often participates directly in management by holding one or more positions on the board of directors. The VC also specializes in a particular industry and uses those industry contacts to help entrepreneurs to perform various business activities.[2] Furthermore, some VCs employ consulting staff that are involved in the management of companies in their portfolios. However, it is difficult to verify the intensity of such monitoring and advising efforts provided by the

[1] The monitoring and advisory role of the VC is empirically documented by Barry et al. (1990), Gompers and Lerner (1999), Hellmann and Puri (2002), and Kaplan and Strömberg (2004).

[2] For example, the VC helps to shape and recruit the management team, to shape the strategy and the business model before and after investing, to assist in production, to line up suppliers, and to develop customer relations.

VC.[3] The VC can harvest investments in private companies in one of two ways: namely, by taking the firm public via an IPO or by selling the firm to another company (mergers and acquisitions (M&A)). In fact, we can interpret the IPO process in this paper as an M&A process if the acquiring company's activity or investment has a synergetic effect on the acquired firm's productivity. Hence, we use the terms of the "IPO" and "M&A" processes for the VC exit route interchangeably.[4]

To capture this dynamic problem, we develop a continuous-time agency model with double moral hazard, in which a risk-neutral agent (manager) provides unobservable value-adding effort and a risk-neutral principal (advisor) also contributes unobservable value-adding effort as well as choosing the optimal timing of changing her role with irreversible investment. The basic question that we address is how the optimal timing of changing the advisor's role and the dynamic properties of optimal incentive provision are characterized under double moral hazard, compared with those under single moral hazard. We further explore the comparative statics regarding the timings of changing the advisor's role and compensating the manager, and provide several empirical implications.

In our basic model, the project generates cumulative cash flows that are affected by both the principal's and the manager's efforts. There are complementarities between the value-adding efforts of the principal and the manager. The principal also chooses the timing of exiting; that is, the timing of the IPO and investment for achieving an upward shift in the

[3] The literature on optimal venture capital contracts makes a similar assumption. See Casamatta (2003), Schmidt (2003), Repullo and Suarez (2004), Inderst and Müller (2004), and Hellmann (2006). However, they use a discrete-time agency model with two or three periods, which is not suitable for the analysis of either IPO timing or dynamic optimal incentive provision.

[4] A similar situation occurs when buyout funds attempt to make private firms in their portfolio go public or when banks attempt to return bankrupt firms under their administration to public ownership. Another example is where a large established firm sells all of its claims in a joint venture with a small innovative firm or where a parent firm spins off one of its subsidiaries to its own shareholders and makes the subsidiary a new entity that is managed independently of the parent firm. With joint venture dissolution, as the large established firm provides resources in areas such as manufacturing, distribution, and marketing, the large-established-firm–small-innovative-firm relationship acts like a VC–entrepreneur relationship. With corporate spin-offs, the parent firm's CEO–subsidiary manager relationship corresponds to a VC–entrepreneur relationship if the CEO's objective is to maximize the initial shareholders' payoff where the CEO is not involved in the management of the subsidiary following a spin-off. In both cases, restructuring costs need to be expended at the time of the dissolution or spin-off. Evidence of operating performance improvements following spin-offs is in Daley, Mehrotra, and Sivakumar (1997) and Desai and Jain (1999).

productivity of the project. Following the IPO, new outside investors own the firm and induce the manager to exert an unobservable effort to operate the project, although they cannot themselves expend any effort to operate the project.

Now, suppose as a benchmark the single moral hazard case in which the principal's effort is observable and verifiable. Although the manager's effort is unobservable, he has an incentive to increase effort if his value depends more strongly on output. Hence, increasing volatility of the manager's value is required to attain a higher level of effort from the manager. However, this implies that poor results are met with penalties. As the manager's limited liability precludes negative wages, the project has to be terminated inefficiently once his future discounted payoff—that is, his continuation value—hits his outside option. Thus, there exists a trade-off between more incentive provision and inefficient termination. This also means that paying cash to the manager earlier makes future inefficient liquidation more likely, as it reduces the manager's continuation payoff. Hence, the possibility of inefficient liquidation causes deferred compensation. Furthermore, if costly investment is irreversible, the possibility of inefficient termination makes IPO execution more costly because the investment cost is less likely to be recovered. Hence, there exists another trade-off between earlier IPO timing and inefficient termination. This trade-off creates an option value of waiting for an IPO. In addition, the change in the governance mechanism with the IPO needs to be considered in choosing the timing of the IPO, because this substantially affects the merit of the IPO via the significant effect on the manager's incentives.

We next suppose the double moral hazard case in which the principal's effort is unobservable. Then, the principal may reduce her effort ex post because she may not be able to commit to secure an appropriate ex post incentive for herself. The possibility of a decrease in the principal's effort may reduce the option value of waiting for the IPO (or the principal's expected profit obtained by waiting for the IPO), thereby advancing the IPO timing. This possibility may also have an adverse effect on the cost of compensating the manager for his effort and may make earlier compensation more costly under double moral hazard. However, we will show that this final intuition is incorrect.

The main results of the model are as follows. Our first main result is that optimal IPO timing is earlier under double moral hazard than under single moral hazard. Intuitively, the principal's effort supplied prior to the IPO is smaller under double moral hazard than under single moral hazard. The reason is that under double moral hazard, the principal has an ex post incentive to reduce her effort level, relative to that optimally obtained under single moral hazard, because she cannot internalize the externality effect of her effort on the manager's incentives when choosing the level of effort after the contract has been offered. This reduces the option value of waiting for the IPO, thus advancing the IPO timing.

The second main result is that the manager's compensation tends to be paid earlier under double moral hazard than under single moral hazard. Intuitively, the manager's compensation is never paid before the IPO under either double or single moral hazard. Given that the cash payment threshold is located after the IPO and that only single moral hazard situation arises after the IPO, the compensation payment timing is earlier under double moral hazard because the IPO timing is earlier under double moral hazard.

We carry out comparative statics for the IPO timing and the manager's compensation timing under double moral hazard. In particular, we show that the IPO is more likely to be brought forward when the need for monitoring by the VC is smaller and the volatility of cash flows is larger. Our results provide testable predictions about IPO timing and vesting provisions for the manager in early-stage start-up firms, management buyout firms, or firms reorganized following bankruptcy. These predictions can also be tested for the dissolution of joint ventures and for corporate spin-offs.

Practically, some VCs retain a fraction of the firm's shares and continue to hold their board positions even after the IPO. If the principal retains a portion of the equity claims in the firm and continues to provide effort after the IPO, a multiagent problem may arise after the IPO. We extend our model to this case and show that most of our main results still hold.

Our work in this paper is related to the growing literature on continuous-time agency models using the martingale techniques developed by Sannikov (2008).[5] Philippon and Sannikov (2007) extend the cash diversion model in DeMarzo and Sannikov (2006) by considering an endogenous shift in the timing in the mean of cash flows by irreversible investment or IPO. However, in Philippon and Sannikov (2007), the principal supplies no effort. Indeed, the main difference between our model and the aforementioned continuous-time agency models is that our model deals with a double moral hazard situation in which the principal exerts unobservable effort before making an irreversible investment and exiting the firm. This enables us to investigate how the principal's moral hazard affects the optimal timing of organizational change and the optimal properties of dynamic contracts, and to explore the interplay between organizational change and the optimal properties of dynamic contracts. In particular, unlike Philippon and Sannikov (2007), we can capture the effect of a change in the ownership structure or governance on the optimal IPO timing and compensation profile for the manager.[6] Furthermore, if the principal does not completely exit with the IPO, we are able to explore a continuous-time multiagent model following the IPO.

Several existing studies have implications for firm characteristics during IPOs. Pagano and Röell (1998) examine the effect of ownership structure on the decision to go public. However, their model is static. Benninga, Helmantel, and Sarig (2005) and Pástor, Taylor, and Veronesi (2009) employ a dynamic IPO model by trading off the diversification benefits of going public against the benefits of private control. Both of these models suggest that going public is optimal when cash flow or the firm's expected future profitability is sufficiently high. However, they do not use the continuous-time agency model, nor do they explore the effect of the change in managerial control with the IPO (VC advice and monitoring vs. market discipline) on the IPO timing or the manager's compensation profile.

[5] See DeMarzo and Sannikov (2006), Biais, Mariotti, Plantin, and Rochet (2007), He (2009), Biais, Mariotti, Rochet, and Villeneuve (2010), Hoffmann and Pfeil (2010), Jovanovic and Prat (2010), Piskorski and Tchistyi (2010), He (2011), and DeMarzo, Fishman, He, and Wang (2012). The continuous-time agency model began with Holmstrom and Milgrom (1987). However, unlike recent models, they assume that the agent can receive a lump-sum payment only at the end of an exogenous finite time interval.

[6] Our result—that the greater the need for the VC's monitoring role, the higher the IPO timing threshold—is novel. Furthermore, Philippon and Sannikov (2007) suggest that the IPO threshold is increasing in volatility. By contrast, our result shows that the IPO threshold may be decreasing in the risk of the project.

The paper is organized as follows. Section 2 describes the basic model. Sections 3 and 4 derive an optimal contract under single and double moral hazard. Section 5 considers the comparative statics. Section 6 allows the principal not to exit completely with the IPO. Section 7 provides empirical implications for our results, and Section 8 concludes. All proofs appear in the Appendix.

2. Basic Model

We now consider a continuous-time model in which a principal hires an agent to operate a firm and has an opportunity to improve the firm's output process substantially by undertaking irreversible investment K when the principal exits. In the subsequent analysis, we take as an example an IPO for a private company presently under the control of a VC or buyout fund. The principal (the VC or buyout fund) hires a manager and then has an opportunity to exit with an IPO. The investment cost K is paid by the funds financed from new outside investors in the IPO.

Let X_t denote the cumulative cash flows produced by the firm up to time t. The total output X_t evolves according to

$$dX_t = A_t dt + \sigma dZ_t, \tag{1}$$

where A_t is the drift of the cash flows, σ is the instantaneous volatility, and $Z = \{Z_t, \mathcal{F}_t; 0 \leq t < \infty\}$, which is a standard Brownian motion on the complete probability space $(\Omega, \mathcal{F}, \mathbb{Q})$. Specifically, before the investment with the IPO, A_t is represented by $A_t = a_{At} + \zeta a_{Pt} + \alpha a_{At} a_{Pt}$, where a_{At} is the manager's effort, a_{Pt} the principal's effort, ζ a constant nonnegative scale adjusted parameter, and α a constant nonnegative complementarity parameter. As the principal makes an effort to monitor and advise the manager in order to improve the firm's business,[7] the efforts of the principal and the manager are interrelated before the IPO.

[7] In the monitoring–auditing model of Townsend (1979), the informed agent asks for costly state verification so that the realization of the project returns is made known to the uninformed agent. As the costly verification in equilibrium occurs only in the low-revenue states, the monitoring and auditing in his model naturally corresponds to bankruptcy proceedings. Hence, the monitoring and auditing cannot increase the

However, after the investment with the IPO, for $\mu > 1$, $A_t = \mu a_{At}$. If the IPO is undertaken, the principal exits, and the firm is now owned by new outside investors. Thus, $a_{Pt} = 0$. However, the scale of the firm expands because of the investment; thus, $\mu > 1$.

The principal and new outside investors are risk neutral and discount the flow of profit at rate r. The principal's effort is a stochastic process $\{a_{Pt} \in A_P,\ 0 \leq t \leq \tau_I\}$ progressively measurable with respect to $\mathcal{F}_t$, where the set of her feasible effort levels A_P is compact with the smallest element 0, and τ_I is the time of the IPO. The principal's effort cost function $g(a_{Pt})$—measured in the same units as the flow of profit—is an increasing, convex, and C^2 function. We normalize $g(0) = 0$.

The manager is risk neutral, but a negative wage is ruled out by limited liability. The manager also discounts his consumption at γ $(> r)$.[8] If the manager's savings interest rate is lower than the principal's discount rate and if the manager is risk neutral, DeMarzo and Sannikov (2006) show that there is an optimal contract in which the manager maintains zero savings. Hence, in this model, the manager can be restricted to consuming what the principal pays him at any time. For simplicity, we assume that at each point of time, the manager can choose either to shirk ($a_{At} = 0$) or to work ($a_{At} = 1$), where $A_A = \{0, 1\}$. Because working is costly for the manager or shirking results in a private benefit, we also suppose that the manager receives a flow of private benefit equal to hdt if he shirks.

The total output process $\{X_t, 0 \leq t < \infty\}$ is publicly observable by all the agents. However, neither the principal nor new outside investors can observe $\{a_{At} \in A_A,\ 0 \leq t < \infty\}$ or the flow of the manager's private benefit, whereas the manager may observe and verify $\{a_{Pt} \in A_P,\ 0 \leq t \leq \tau_I\}$ or may not observe $\{a_{Pt} \in A_P,\ 0 \leq t \leq \tau_I\}$.

For $0 \leq t \leq \tau_I$, the principal signs a contract with the manager at time $t = 0$ that specifies

productivity of the firm. By contrast, in our model, the principal, who is uninformed about the manager's effort, monitors and advises the manager, who is informed or uninformed about the principal's effort, as discussed in the introduction. As a result, the monitoring and advising in our model can increase the productivity of the firm although it cannot make the manager's effort known to the principal.

[8] If the principal and the manager are equally patient when the manager is risk neutral, the principal can postpone payments to the manager indefinitely. This possibility precludes the existence of an optimal contract. See Biais, Mariotti, Rochet, and Villeneuve (2010).

how the manager's nondecreasing cumulative consumption, C_t, depends on the observation of X_t.[9] If the manager can observe and verify $\{a_{Pt} \in A_P,\ 0 \leq t \leq \tau_I\}$, the contract can also stipulate how a_{Pt} depends on the observation of X_t. Let $\Pi = \{C_t, a_{Pt} : 0 \leq t \leq \tau_I\}$ denote the contract prior to the IPO if the manager can observe and verify $\{a_{Pt} \in A_P,\ 0 \leq t \leq \tau_I\}$; and $\Pi_D = \{C_t : 0 \leq t \leq \tau_I\}$ denote the contract prior to the IPO if the manager cannot observe $\{a_{Pt} \in A_P,\ 0 \leq t \leq \tau_I\}$. On the other hand, for $\tau_I \leq t < \infty$, new outside investors sign a contract with the manager at time $t = \tau_I$ that specifies how C_t depends on the observation of X_t. Let $\overline{\Pi} = \{C_t : \tau_I \leq t < \infty\}$ denote the contract after the IPO. The principal and new outside investors are able to commit to any such contract. In addition, the principal determines how the time of the IPO, τ_I, depends on the observation of X_t and how the time when the contract is terminated before the IPO, τ_{T_0}, depends on the observation of X_t. New outside investors also determine how the time when the contract is terminated after the IPO, τ_{T_1}, depends on the observation of X_t.

Now, for any contract Π (or Π_D) and $\overline{\Pi}$ and for any time $(\tau_I, \tau_{T_0}, \tau_{T_1})$, suppose that the manager chooses an effort-level process $\{a_{At} \in A_A,\ 0 \leq t < \infty\}$. The manager's total expected payoff at $t = 0$ is then

$$E\left\{\int_0^{\tau_{T_0 I}} e^{-\gamma t}\left[dC_t + h \cdot (1 - a_{At})dt\right] + \mathbf{1}_{\tau_I \geq \tau_{T_0}} e^{-\gamma \tau_{T_0}} R\right.$$

$$\left. + \mathbf{1}_{\tau_I < \tau_{T_0}} e^{-\gamma \tau_I}\left[\int_{\tau_I}^{\tau_{T_1}} e^{-\gamma(t-\tau_I)}\left[dC_t + h \cdot (1 - a_{At})dt\right]\right] + \mathbf{1}_{\tau_I < \tau_{T_0}} e^{-\gamma \tau_{T_1}} R\right\}, \tag{2}$$

where $\tau_{T_0 I} = \min(\tau_{T_0}, \tau_I)$; $\mathbf{1}_{\tau_I \geq \tau_{T_0}} = 1$ or 0 ($\mathbf{1}_{\tau_I < \tau_{T_0}} = 0$ or 1) according to $\tau_I \geq \tau_{T_0}$ or $\tau_I < \tau_{T_0}$; and the manager receives expected payoff R from an outside option when the contract is terminated, irrespective of whether the IPO occurs. In addition, for any contract Π (or Π_D) and $\overline{\Pi}$ and for any time $(\tau_I, \tau_{T_0}, \tau_{T_1})$, suppose that the principal and the manager choose effort-level processes $\{a_{Pt} \in A_P,\ 0 \leq t \leq \tau_I\}$ and $\{a_{At} \in A_A,\ 0 \leq t < \infty\}$. Then, the

[9] Because a negative wage is ruled out, C_t must be nondecreasing.

principal's total expected profit at $t = 0$ is

$$E\left\{\int_0^{\tau_{T_0 I}} e^{-rt}\left[(a_{At} + \zeta a_{Pt} + \alpha a_{At}a_{Pt} - g(a_{Pt}))\,dt - dC_t\right] + \mathbf{1}_{\tau_I \geq \tau_{T_0}} e^{-r\tau_{T_0}} L_0 \right.$$
$$\left. + \mathbf{1}_{\tau_I < \tau_{T_0}} e^{-r\tau_I}\left[\int_{\tau_I}^{\tau_{T_1}} e^{-r(t-\tau_I)}(\mu a_{At}dt - dC_t) - K\right] + \mathbf{1}_{\tau_I < \tau_{T_0}} e^{-r\tau_{T_1}} L_1\right\}, \tag{3}$$

where L_0 (L_1) is the expected liquidation payoff of the principal (new outside investors) when the contract is terminated before (after) the IPO. Note that at $t = \tau_I$, new outside investors buy the firm via the IPO. The principal anticipates an incentive for the manager to choose his effort level after the IPO and sets the IPO price reasonably. Because of arbitration, the principal's profit from the IPO equals the total expected payoff of new outside investors at $t = \tau_I$ exclusive of the IPO cost, represented by the third and fourth terms in (3).[10]

3. Optimal Contract under Single Moral Hazard

In this section, we assume that the manager can observe and verify $\{a_{Pt} \in A_P,\ 0 \leq t \leq \tau_I\}$. The game is solved through backward induction.

3.1. Optimal contract after the IPO.—

Consider the case where the IPO has already taken place. The contracting problem is then to find a combination of an incentive-compatible contract and termination timing, $(\overline{\Pi}, \tau_{T_1})$, and an incentive-compatible choice of the manager's effort process, $\{a_{At} \in A_A,\ \tau_I \leq t \leq \tau_{T_1}\}$, that maximize the expected profit of new outside investors subject to delivering to the manager a required payoff W_I at the IPO, where W_I is defined by the manager's continuation value W_t at $t = \tau_I$ given below. The manager's effort process is incentive compatible if it maximizes his total expected payoff from τ_I onward given $(\overline{\Pi}, \tau_{T_1})$; that is, if it maximizes $E_{\tau_I}\left\{\int_{\tau_I}^{\tau_{T_1}} e^{-\gamma(t-\tau_I)}\left[dC_t + h\cdot(1 - a_{At})dt\right] + e^{-\gamma(\tau_{T_1}-\tau_I)}R\right\}$.

[10] In corporate spin-offs, the shares of the subsidiary are distributed on a pro rata basis to the parent firm's shareholders. If the principal is the parent firm's CEO whose preferences therefore coincide with those of the parent firm's shareholders, her objective function is still given by (3) because she is not involved in the management of the spun-off subsidiary.

For any $(\overline{\Pi}, \tau_{T_1})$, the manager's continuation value W_t is his future expected discounted payoff at time t, given that he will follow $\{a_{At} \in A_A,\, \tau_I \leq t \leq \tau_{T_1}\}$:

$$W_t = E_t \left\{ \int_t^{\tau_{T_1}} e^{-\gamma s} \left[dC_s + h \cdot (1 - a_{As}) ds \right] + e^{-\gamma(\tau_{T_1} - t)} R \right\}. \tag{4}$$

The manager's effort process is then specified by $\{a_A(W_t) : \tau_I \leq t \leq \tau_{T_1}\}$.[11] Denote by $G(W)$ the value function of new outside investors before deducting the investment cost K (the highest expected present value of the profit in an optimal contract). For simplicity, we assume that $G(W)$ is concave. The formal proof for the concavity of $G(W)$ is provided in the Appendix. We assume that implementing the manager's high effort $a_A(W_t) = 1$ at any $t \in [\tau_I,\, \tau_{T_1}]$ is optimal for new outside investors, and provide a sufficient condition for the optimality of implementing this action in the Appendix.

The optimal contract is now derived using the technique in DeMarzo and Sannikov (2006) and Sannikov (2008). Specifically, for any $(\overline{\Pi}, \tau_{T_1})$, there exists a progressively measurable process $\{Y_t, \mathcal{F}_t : \tau_I \leq t \leq \tau_{T_1}\}$ in $\mathcal{L}^*$ such that[12]

$$dW_t = \{\gamma W_t - h \cdot [1 - a_A(W_t)]\} dt - dC_t + Y(W_t)[dX_t - \mu a_A(W_t) dt]. \tag{5}$$

The evolution of W_t in (5) includes the sensitivity of W_t to output, $Y(W_t)$. This suggests that we can characterize the manager's incentive compatibility by considering $Y(W_t)$. Indeed, implementing $a_A(W_t) = 1$ is incentive compatible if and only if $Y(W_1) \geq \beta_1$ for $t \in [\tau_I,\, \tau_{T_1}]$, where $\beta_1 \equiv \min\{y : y\mu \geq h\} = \frac{h}{\mu}$.

As proved in DeMarzo and Sannikov (2006), the marginal cost of delivering W can never exceed the cost of an immediate transfer in terms of the utility of new outside investors; $G'(W) \geq -1$. Define W^{++} as the lowest value of W such that $G'(W) = -1$. Then, it is optimal to set the manager's compensation as $dC = \max(W - W^{++}, 0)$.

[11] As the manager's cumulative consumption C_t is defined as a mapping from the history of $\mathcal{F}_t$ to a nondecreasing process, it may be more general than a function of W_t. Hence, we simply write it by C_t.

[12] A process Y is in $\boldsymbol{\mathcal{L}}^*$ if $E\left[\int_0^t Y_s^2 ds\right] < \infty$ for all $t \in [0, \infty)$.

Now, the following proposition summarizes the optimal contract after the IPO.

Proposition 1: *For the manager's starting value $W_I \in [R, W^{++}]$, the optimal contract is characterized by the unique concave function $G(W)$ that satisfies the Hamilton–Jacobi–Bellman (HJB) equation*

$$rG(W) = \max_{Y \geq \beta_1} \mu + G'(W)\gamma W + \frac{G''(W)}{2} Y^2 \sigma^2, \tag{6}$$

with $\beta_1 = \frac{h}{\mu}$ and boundary conditions $G(R) = L_1$, $G'(W^{++}) = -1$, and $G''(W^{++}) = 0$.[13] *When $W_t \in [R, W^{++})$, $dC_t = 0$. When $W_t = W^{++}$, payments dC_t cause W_t to reflect at W^{++}. If $W_t > W^{++}$, an immediate payment $W_t - W^{++}$ is made. The contract is terminated at time τ_{T_1} when W_t hits R for the first time. The optimal contract then attains profit $G(W_I)$ for new outside investors.*[14]

As $G''(W) < 0$, new outside investors dislike volatility in W and optimally choose the sensitivity of W to output; that is, $Y = \beta_1 = \frac{h}{\mu}$ in (6). Using (6), $G'(W^{++}) = -1$, and $G''(W^{++}) = 0$, we obtain $rG(W^{++}) + \gamma W^{++} = \mu$; that is, payment to the manager is postponed until the new outside investors' and the manager's required expected returns exhaust the available expected cash flows generated after the IPO.

3.2. Optimal contract before the IPO.—

In this case, the contracting problem is to find a combination of an incentive-compatible contract and IPO and termination timing, $(\Pi, \tau_I, \tau_{T_0})$, and an incentive-compatible manager's effort process, $\{a_{At} \in A_A,\ 0 \leq t < \infty\}$, that maximize the expected profit of the principal subject to delivering to the manager an initial required payoff W_0. The manager's

[13] The first boundary condition is the value-matching condition, which implies that the principal must terminate the contract to hold the agent's reservation value, R. The second boundary condition is the smooth-pasting condition, which guarantees the optimal choice of W^{++}. The third boundary condition is the super contract condition for the optimal choice of W^{++}, which requires that the second derivatives match at the boundary.

[14] For any starting value of $W_I > W^{++}$, $G(W_I)$ is an upper bound on the total expected profit of new outside investors. However, this case can be excluded because $F'(W_I) = G'(W_I) > -1$, as is shown in Proposition 2. If $W_I < R$, the manager never participates in the contract.

effort process is incentive compatible if it maximizes his total expected payoff defined by (2), given $(\Pi, \tau_I, \tau_{T_0})$.

Given W_t as in (4), the processes of the manager's and principal's efforts can be specified by $\{a_A(W_t),\ a_P(W_t) : 0 \leq t \leq \tau_{T_0 I}\}$. Denote by $F(W)$ the value function of the principal. To facilitate our discussion, we assume that $F(W)$ is concave. The formal proof for the concavity of $F(W)$ is provided in the Appendix.

We again assume that implementing $a_A(W_t) = 1$ at any $t \in [0,\ \tau_{T_0 I}]$ is optimal for the principal, and provide a sufficient condition for such optimality in the Appendix. We also make the following assumption, which ensures that the IPO does not take place at time 0 if W_0 is not sufficiently large.

Assumption 1: $K > \max(L_0, L_1)$.

As in Section 3.1, we can characterize the evolution of W_t, the manager's incentive compatibility, and the manager's compensation as follows. First, for any $(\Pi, \tau_I, \tau_{T_0})$, there exists a progressively measurable process $\{Y_t, \mathcal{F}_t : 0 \leq t \leq \tau_{T_0 I}\}$ in $\mathcal{L}^*$ such that

$$dW_t = \{\gamma W_t - h \cdot [1 - a_A(W_t)]\}\, dt - dC_t + Y(W_t)\{dX_t - [a_A(W_t) + \zeta a_P(W_t) + \alpha a_A(W_t) a_P(W_t)]dt\}. \tag{7}$$

Second, implementing $a_A(W_t) = 1$ is incentive compatible if and only if $Y(W_t) \geq \beta_0(a_P(W_t))$ for $t \in [0, \tau_{T_0 I}]$, where $\beta_0(a_P) = \min\{y : y \cdot (1 + \alpha a_P) \geq h\} = \frac{h}{1+\alpha a_P}$. Finally, the optimality of the contract implies $F'(W) \geq -1$. Hence, it is optimal to pay the manager according to $dC = \max(W - W^+, 0)$, where W^+ is defined as the lowest value of W such that $F'(W) = -1$. In fact, as in Proposition 2 below, we will show that $W^+ \geq W_I$. Thus, the manager's compensation is zero before the IPO.

The following proposition summarizes the optimal contract before the IPO.

Proposition 2: *Suppose that the manager can observe and verify $\{a_{Pt} \in A_P,\ 0 \leq t \leq \tau_{T_0 I}\}$. For any starting value $W_0 \in [R, W_I]$, the optimal contract is characterized by the* *unique concave function $F(W)$ $(\geq G(W) - K)$ that satisfies the HJB equation*

$$rF(W) = \max_{a_P \in A_P, Y \geq \beta_0(a_P)} 1 + (\zeta + \alpha)a_P - g(a_P) + F'(W)\gamma W + \frac{F''(W)}{2} Y^2 \sigma^2, \tag{8}$$

with $\beta_0(a_P) = \frac{h}{1+\alpha a_P}$ and boundary conditions $F(R) = L_0$, $F(W_I) = G(W_I) - K$, and $F'(W_I) = G'(W_I)$.[15] *When $W_t \in [R, W_I]$, $dC_t = 0$. This means $W_I \leq W^+$. The IPO occurs when W_t reaches W_I $(> R)$, or the contract is terminated when W_t hits R, whichever happens sooner. After the IPO, the continuation contract is given by Proposition 1 at the starting value W_I. The optimal contract then provides profit $F(W_0)$ to the principal.*[16]

As $F''(W) < 0$, the principal optimally chooses $Y = \beta_0(a_P) = \frac{h}{1+\alpha a_P}$ in (8).

We now discuss the optimal solution in more detail. Inspecting the results of Propositions 1 and 2 shows that the manager's compensation is never paid before the IPO, but is paid with a cash payment immediately after the IPO when $W_t > W^{++}$. The intuition is that if the manager receives compensation before the IPO (that is, $W^+ < W_I$), the principal must pay $W_t - W^+$ to the manager immediately. As this immediate payment causes W_t to be brought back to W^+ for any $W_t \in (W^+, \infty)$, it follows from $W^+ < W_I$ that the principal's profit from the IPO always becomes smaller than her expected profit obtained by waiting for the IPO. Hence, the IPO would never be done.

For the optimal choice of a_P, it follows from $\beta_0(a_P) = \frac{h}{1+\alpha a_P}$ that if $a_P > 0$, the first-order condition leads to

$$\begin{aligned} \zeta + \alpha &= g'(a_P(W)) - F''(W)\sigma^2 \beta_0(a_P(W))\beta_0'(a_P(W)) \\ &\leq g'(a_P(W)) \qquad \text{for all } W \leq W_I \text{, with strict inequality only if } \alpha > 0. \end{aligned}$$

This implies that the marginal productivity of a_P is smaller than its marginal cost when α

[15] The first and second boundary conditions are the value-matching conditions, while the third boundary condition is the smooth-pasting condition that guarantees the optimal choice of W_I. Note that the IPO cost is K, which is deducted from $G(W_I)$.

[16] For the starting value of $W_0 > W_I$, the IPO takes place immediately or contracts with positive profit do not exist. If $W_0 < R$, the manager never participates in the contract.

> 0. Intuitively, an increase in a_P relaxes the manager's incentive-compatibility constraint because of the complementarity effect of the principal's effort ($\beta_0'(a_P(W)) < 0$ for $\alpha > 0$). Hence, an increase in a_P reduces the cost of the principal exposing the manager to income uncertainty to provide incentive. Thus, to provide the manager with appropriate less costly incentives using the complementarity effect, the principal increases a_P by a level at which the marginal productivity of a_P is smaller than its marginal cost.

To conclude this section, we comment on the IPO timing. If the IPO is implemented, the investment cost arises at the time of the IPO. However, there is a risk of losing value if the contract is terminated. Because $L_1 < K$ implies that the liquidation value cannot compensate for the investment cost, it is inefficient to plan the IPO when there is a higher probability of liquidation. As a result, a sufficiently small W_t close to R raises the risk of losing value upon termination and reduces the IPO price, thereby making it impossible for the principal to recover the total cost of the IPO.[17] Hence, it is optimal to execute the IPO only when the manager accumulates a sufficiently large W_t. In addition, the optimal IPO timing is affected by both the agency problem and the change in the governance mechanism. For example, if the change in the governance mechanism with the IPO affects the post-IPO agency relation and the post-IPO firm value, it will affect the timing of the IPO.

4. Optimal Contract under Double Moral Hazard

In this section, we assume that the manager cannot observe $\{a_{Pt} \in A_P, 0 \leq t \leq \tau_{T_0 I}\}$. This assumption can be justified because it is difficult to verify the intensity of the monitoring and advising efforts provided by the VC. Even in this case, a different formulation of the contract problem is required only before the IPO because the principal does not provide any effort after the IPO. Hence, Proposition 1 still holds.

Before the IPO, the principal's design for the incentive scheme must address the manager's

[17] On the other hand, if W_0 is sufficiently large, and if contracts with positive profit exist, it is optimal for the principal to execute the IPO at $t = 0$ immediately because she can recover the cost of the IPO.

incentive problem and her own incentive problem.[18] Thus, the optimal contracting problem prior to the IPO is to find a combination of an incentive-compatible contract and IPO and termination timing, $(\Pi_D, \tau_I, \tau_{T_0})$, an incentive-compatible manager's effort process, $\{a_{At} \in A_A,\ 0 \leq t \leq \tau_{T_0 I}\}$, and an incentive-compatible principal's effort process, $\{a_{Pt} \in A_P,\ 0 \leq t \leq \tau_{T_0 I}\}$, that maximize the expected profit of the principal subject to delivering to the manager an initial required payoff W_0. The manager's effort process is incentive compatible if it maximizes his total expected utility defined by (2), given $(\Pi_D, \tau_I, \tau_{T_0}, \{a_{Pt} \in A_P,\ 0 \leq t \leq \tau_{T_0 I}\})$, while the principal's effort process is incentive compatible if it maximizes her total expected payoff defined by (3), given $(\Pi_D, \tau_I, \tau_{T_0}, \{a_{At} \in A_A,\ 0 \leq t \leq \tau_{T_0 I}\})$. Denote by $\widetilde{F}(W)$ the value function of the principal in this case. For simplicity, we assume that $\widetilde{F}(W)$ is concave; the formal proof is shown in the Appendix. We still assume that implementing $a_A(W_t) = 1$ at any $t \in [0, \tau_{T_0 I}]$ is optimal for the principal, and provide a sufficient condition for such optimality in the Appendix.

Indeed, the discussion in Section 3.2 still applies for the evolution of W_t and the incentive-compatibility constraint for the manager. However, the incentive-compatible principal's effort process is determined by the maximizer to the following maximization problem, given the optimal level of Π_D, τ_I, τ_{T_0}, $\{a_{At} \in A_A,\ 0 \leq t \leq \tau_{T_0 I}\}$, and $\widetilde{F}(W)$:

$$a_P = \underset{a_P \in A_P}{\arg\max} \quad 1 + (\zeta + \alpha) a_P - g(a_P) + \widetilde{F}'(W)\gamma W + \frac{\widetilde{F}''(W)}{2} \left[Y(W)\right]^2 \sigma^2. \tag{9}$$

Here, $Y(W)$ is given by $a_P(W)$: $Y(W) = \beta_0(a_P(W))$, where $a_P(W)$ is the recommended principal's effort at each point of W.[19] Note that after offering the contract, the principal can optimally choose her effort at each point of W without considering the manager's incentive-

[18] In our model, if the principal makes dC_t sufficiently large at some t' to penalize herself severely when the cash flows are very low, then W_t for $t > t'$ will be smaller than R. Thus, the possibility of contract termination implies that the principal cannot arbitrarily reduce the low outcome range in which she should take the penalty by increasing the penalty amount. Hence, in our dynamic setting, in contrast to the static double moral hazard model, such as that in Kim and Wang (1998), the double moral hazard case cannot arbitrarily closely approach the single moral hazard case, even though there is no upper bound for the wage contract.

[19] In other words, the recommended $Y(W)$ is set equal to $\beta_0(a_P(W))$ under the optimal contract.

compatibility constraint, if the manager cannot observe $\{a_{Pt} \in A_P,\, 0 \leq t \leq \tau_{T_0 I}\}$. To ensure that $a_P > 0$ under the optimal contract, we assume that $g'(0) < \zeta + \alpha$. Then, as the right-hand side of (9) is concave with respect to a_P, (9) is rewritten as $\zeta + \alpha = g'(a_P)$, or $a_P = g'^{-1}(\zeta + \alpha) = \psi(\zeta + \alpha)$.

Now, repeating a procedure similar to that in the proof of Proposition 2, we can derive the following proposition, which summarizes the optimal contract under double moral hazard.

Proposition 3: *Suppose that the manager cannot observe $\{a_{Pt} \in A_P,\, 0 \leq t \leq \tau_{T_0 I}\}$. For any starting value $W_0 \in [R,\, W_I]$, the optimal contract is characterized by the unique concave function $\widetilde{F}(W)$ $(\geq G(W) - K)$ that satisfies the HJB equation*

$$r\widetilde{F}(W) = \max_{a_P = \psi(\zeta+\alpha),\, Y \geq \beta_0(a_P)} 1 + (\zeta + \alpha)a_P - g(a_P) + \widetilde{F}'(W)\gamma W + \frac{\widetilde{F}''(W)}{2} Y^2 \sigma^2, \qquad (10)$$

with $\beta_0(a_P) = \frac{h}{1+\alpha a_P}$ and boundary conditions $\widetilde{F}(R) = L_0$, $\widetilde{F}(W_I) = G(W_I) - K$, and $\widetilde{F}'(W_I) = G'(W_I)$. When $W_t \in [R,\, W_I]$, $dC_t = 0$. This means that $W_I \leq W^+$. The IPO takes place when W_t reaches W_I, or the contract is terminated when W_t hits R, whichever happens sooner. After the IPO, the continuation contract is given by Proposition 1 at the starting value W_I. Then, the optimal contract attains profit $\widetilde{F}(W_0)$ for the principal.[20]

Again, as $\widetilde{F}''(W) < 0$, the principal optimally chooses $Y = \beta_0(a_P) = \frac{h}{1+\alpha a_P}$ in (10).

We next compare the optimal choices of W^{++}, W_I, and a_P under double moral hazard with those under single moral hazard. Under double (or single) moral hazard, let $\widetilde{W}^{++*}$ and $\widetilde{W}_I^*$ (or W^{++*} and W_I^*) denote the corresponding cash payment and value-maximizing IPO thresholds, respectively.

We begin by discussing the optimal choices of W^{++} and a_P. First, as in single moral hazard, the manager is never paid before the IPO under double moral hazard. Furthermore, even under double moral hazard, the choice rule of W^{++} is given by Proposition 1. As $G(W)$ is the same irrespective of single or double moral hazard, we see $\widetilde{W}^{++*} = W^{++*}$. Second,

[20]For the starting value of $W_0 > W_I$, the IPO is immediately executed or contracts with positive profit do not exist.

because the optimality implies $\zeta + \alpha = g'(a_P)$ under double moral hazard but $\zeta + \alpha < g'(a_P)$ under single moral hazard if $\alpha > 0$, it follows from $g'' > 0$ that a_P is smaller under double moral hazard than under single moral hazard.

The above discussion leads to the following proposition.

Proposition 4: *(i) The manager is never paid before the IPO, irrespective of single or double moral hazard. In addition, the optimal cash payment threshold is not affected by the contractibility of the principal's effort:* $\widetilde{W}^{++*} = W^{++*}$.
(ii) Double moral hazard induces the principal to undersupply her own effort relative to the case of single moral hazard.

Several remarks on Proposition 4 are in order. First, Philippon and Sannikov (2007) suggest that under single moral hazard, the manager is not paid until a certain period after the IPO. Proposition 4(i) confirms their finding even under double moral hazard. This finding also suggests that the optimal contract is not linear under the continuous-time agency model with double moral hazard. By contrast, in the static agency model with double moral hazard, Romano (1994) and Bhattacharyya and Lafontaine (1995) show that a simple linear contract with a fixed fee implements the second-best outcome when the agent is risk neutral. Our result depends on the possibility of contract termination under the manager's limited liability, which is not considered in their model.

Second, for the choice of a_P, if the manager cannot observe a_P, the principal cannot commit to considering the manager's incentive-compatibility constraint in choosing her own effort after the compensation contract has been offered. Hence, the principal cannot internalize the external effect of her own effort on the manager's incentives. As a result, the noncontractibility of the principal's effort leads to an undersupply of a_P.

We next explore how the contractibility of the principal's effort affects the IPO threshold. The result is novel because Philippon and Sannikov (2007) consider only the single moral hazard case.

Proposition 5: *For all* $W \geq R$, $\widetilde{F}(W) < F(W)$. *In addition, the optimal IPO timing is* *earlier under double moral hazard than under single moral hazard:* $\widetilde{W}_I^* < W_I^*$.

Intuitively, $\widetilde{F}(W)$ is smaller than $F(W)$ because the principal cannot provide an appropriate ex post incentive for herself under double moral hazard. However, $G(W)$ is the same, irrespective of single or double moral hazard. This implies that the net expected payoff of the principal obtained when postponing the IPO is always smaller under double moral hazard than under single moral hazard ($\widetilde{F}(W) - [G(W) - K] < F(W) - [G(W) - K]$ for all $W \geq R$). Because the IPO does not take place until the expected present value of the principal's profit ($\widetilde{F}(W)$ or $F(W)$) meets her profit from the IPO ($G(W) - K$), it follows from the boundary conditions of (8) and (10) and the concavity of $\widetilde{F}(W)$, $F(W)$, and $G(W)$ that the IPO threshold is lower under double moral hazard than under single moral hazard.

Combining Propositions 4(i) and 5, we obtain the following proposition.

Proposition 6: *The manager's compensation tends to be paid earlier under double moral hazard than under single moral hazard.*

Hence, the manager's compensation as a function of performance history depends on whether the principal's effort is observable, even though the manager's compensation as a function of W is not (These points are suggested by an anonymous referee).

5. Comparative Statics under Double Moral Hazard

Using Propositions 1 and 3, we compute the comparative statics on the IPO and managerial compensation policies in the optimal contract under double moral hazard. Table 1 summarizes our results. The derivation procedures are provided in the Appendix. As in Proposition 6, by combining the results of $\widetilde{W}^{++*}$ and $\widetilde{W}_I^*$ in Table 1, we also derive the following result: the larger is α (or the larger is μ), the later the manager's compensation tends to be paid (or earlier if $\frac{\partial G(W)}{\partial \mu} > \frac{1}{r}$ *for all* $W \geq R$ and if $d\mu$ is sufficiently small).

The intuition for the results of $\widetilde{W}^{++*}$ is as follows. First, as discussed after Proposition 1, the payment to the manager is postponed until the new outside investors' and the manager's required expected returns, $rG(W) + \gamma W$, exhaust the available expected cash flows, μ. Indeed, an increase in μ not only raises the stream of the expected cash flows, μ, but also reduces the cost of the principal exposing the manager to income uncertainty to provide incentives because it reduces the variation in W by relaxing the manager's incentive-compatibility constraint. Thus, when μ increases, $rG(W)$ increases. On the other hand, an increase in μ raises $\mu - \gamma W$. However, if $\frac{\partial G(W)}{\partial \mu}$ is greater than $\frac{1}{r}$, an increase in μ raises $rG(W)$ more than $\mu - \gamma W$. Thus, $\widetilde{W}^{++*}$ needs to be reduced because $rG'(\widetilde{W}^{++*}) + \gamma = \gamma - r > 0$. Second, an increase in σ^2 reduces $G(W)$ because it directly increases the variation in W. This induces new outside investors to choose the higher $\widetilde{W}^{++*}$.

We next explain the intuition for the results of $\widetilde{W}_I^*$. First, an increase in α directly increases $\widetilde{F}(W)$ because it raises the stream of expected cash flows $(1 + (\zeta + \alpha)a_P)$ and also reduces the cost of exposing the manager to income uncertainty by reducing the variation in W through the relaxation of the manager's incentive-compatibility constraint. However, an increase in α does not affect $G(W)$. As the principal's profit obtained when postponing the IPO becomes larger relative to her profit generated from the IPO, $\widetilde{W}_I^*$ increases with α.

Second, an increase in μ raises $G(W)$. This effect also increases $\widetilde{F}(W)$ because the principal's profit from the IPO is equal to $G(W) - K$. In fact, if $d\mu$ is sufficiently small, $G(W)$ increases with μ more rapidly than $\widetilde{F}(W)$ near the IPO threshold. The reason is that the direct effect of an increase in μ only raises the expected cash flows of the firm after the IPO, thus increasing $G(W)$. This implies that near the IPO threshold, the principal's profit generated from the IPO $(G(W) - K)$ is more sensitive to μ than her profit obtained when postponing the IPO $(\widetilde{F}(W))$. Hence, if $d\mu$ is sufficiently small, the principal is more likely to speed up the IPO timing as μ increases: $\widetilde{W}_I^*$ is decreasing in μ.[21]

Third, an increase in σ^2 reduces $G(W)$. This effect also decreases $\widetilde{F}(W)$ because the

[21] Clementi (2002) builds a discrete-time model of the optimal IPO decision and suggests that private firms with a larger positive productivity shock are more likely to go public. Benninga, Helmantel, and Sarig (2005) consider a simple value function model of optimal IPO timing. They predict that firms go public when cash flow is high. Pástor, Taylor, and Veronesi (2009) develop a continuous-time model of the optimal IPO decision with learning about average profitability under no moral hazard. They show that if both investors and the entrepreneur are risk averse, an IPO is more likely at some prespecified time when the expected profitability and the volatility of profitability are higher. However, their qualitative results on the likelihood of an IPO depend on the assumption that IPO timing is exogenously given.

principal's profit from the IPO decreases. On the other hand, an increase in σ^2 directly reduces $\widetilde{F}(W)$ because it increases the variation in W. Note that the first effect on $G(W)$ increases $\widetilde{W}_I^*$, whereas the second and third effects on $\widetilde{F}(W)$ decrease $\widetilde{W}_I^*$. In fact, if r and $d\sigma^2$ are sufficiently small, the first and second effects cancel each other out. Thus, only the third effect remains near the IPO threshold. Hence, $\widetilde{W}_I^*$ is decreasing in σ^2.[22]

One of the interesting results obtained here is that the IPO threshold $\widetilde{W}_I^*$ is decreasing in σ^2 if r and $d\sigma^2$ are sufficiently small. Philippon and Sannikov (2007) extend the continuous-time cash diversion model of DeMarzo and Sannikov (2006). They assume that a standard Brownian motion represents a signal regarding the action of the manager rather than the cumulative cash flow. They indicate that the IPO threshold is increasing in the volatility of the signal. By contrast, we show that the IPO threshold is decreasing in the risk of the project (equal to the volatility of the cumulative cash flow) under certain conditions. Intuitively, in their cash diversion model, there is no difference between the incentive-compatibility constraints for the manager before and after the IPO. However, in our double moral hazard model, differences exist between these constraints before and after the IPO. The differences mainly result from the complementarity between the principal's and the manager's efforts before the IPO. As a result, and in contrast to the model in Philippon and Sannikov (2007), the net expected profit of the principal obtained when postponing the IPO ($\widetilde{F}(W)$) is decreasing in σ^2 for all $W \geq R$ in our model, thus generating the different result.

6. Model of the VC Not Exiting the Firm Following the IPO

In the previous sections, we considered the case where the principal (the VC) exits completely with the IPO. In this section, we assume that for reasons not modeled in this analysis, the VC retains a fraction of ω ($< \overline{\omega}$) of the firm's shares and provides effort to increase the productivity of the firm even after the IPO. However, the firm is under the control of new outside investors because the VC must set ω sufficiently small in order to recover the funds

[22] See footnote 21.

invested for future investment opportunities or to finance the investment cost K for expanding the scale of the firm. Furthermore, as the new outside investors cannot observe either the manager's effort or the principal's effort, they need to design an optimal incentive scheme in a multiagent environment following the IPO, regardless of whether we assume single or double moral hazard before the IPO.

Here, we assume that the drift of the cash flows after the IPO is given by $A_t^N = \mu(a_{At} + \zeta a_{Pt} + \alpha a_{At} a_{Pt})$. Then, the manager's total expected payoff at $t = 0$ is still written by (2), whereas the principal's total expected payoff at $t = 0$ is

$$
\begin{aligned}
& E\Bigg\{ \int_0^{\tau_{T_0 I}} e^{-rt} \left[(A_t - g(a_{Pt}))\, dt - dC_t \right] + \mathbf{1}_{\tau_I \geq \tau_{T_0}} e^{-r\tau_{T_0}} L_0 \\
& + \mathbf{1}_{\tau_I < \tau_{T_0}} \left[e^{-r\tau_I} \left[\int_{\tau_I}^{\tau_{T_1}} e^{-r(t-\tau_I)} \left[(\omega A_t^N - g(a_{Pt}))dt - \omega dC_t \right] - \omega K \right] + e^{-r\tau_{T_1}} \omega L_1 \right] \\
& + \mathbf{1}_{\tau_I < \tau_{T_0}} \left[e^{-r\tau_I} (1-\omega) \left[\int_{\tau_I}^{\tau_{T_1}} e^{-r(t-\tau_I)} \left(A_t^N dt - dC_t \right) - K \right] + e^{-r\tau_{T_1}} (1-\omega) L_1 \right] \Bigg\}.
\end{aligned} \tag{11}
$$

Note that the profit of the principal from the IPO equals the total expected payoff for new outside investors at $t = \tau_I$, which is shown by the terms in the third line in (11).

In designing the optimal incentive scheme after the IPO, for simplicity we assume that there is no colluding agreement between the principal and the manager, and that if there are multiple effort processes of $\{(a_{At}, a_{Pt}) \in A_A \times A_P, \tau_I \leq t < \infty\}$ as Nash equilibrium for the principal and the manager, new outside investors can implement their most preferred effort processes from among those effort processes. In addition, new outside investors cannot offer the principal or the manager a contract contingent on the manager's report of a_P even when the manager can observe a_P. This assumption can be justified because such contracts are not found in practice. Assumption 1 is still assumed to hold. Again, we assume that implementing $a_A(W_t) = 1$ is always optimal.[23] In the Appendix, we provide a formal characterization of the optimal contract. Here, we briefly summarize the results, and

[23] The sufficient condition for such optimality is shown in Hori and Osano (2013).

discuss whether our main results in Sections 4 and 5 are unaffected by this extension.

First, Propositions 4–6 regarding the difference between the features of the optimal contracts under single and double moral hazard continue to hold. Second, Table 2 gives the comparative statics under double moral hazard in this situation, where G_O (G_P) and $\widetilde{W}_O^{++*}$ denote the value function of new outside investors (the principal) and the optimal cash payment threshold. The main difference between the results in Tables 1 and 2 is that Table 2 includes the effect of α even after the IPO and the effects of ω before and after the IPO. Combining the results of $\widetilde{W}_O^{++*}$ and $\widetilde{W}_I^*$ in Table 2, we also derive the following result: The larger is μ, the earlier the manager's compensation tends to be paid if $\frac{\partial G_O(W)}{\partial \mu} > \frac{\partial[(1-\omega)A^{N**}(W)]}{\partial \mu}$ for all $W \geq R$ and if $d\mu$ is sufficiently small, where $A^{N**}(W) = \mu[1 + (\zeta + \alpha)a_P^{**}(W)]$.

The intuition for the results of α after the IPO can be explained in a way similar to that in the case of μ in Section 5. The effect of α on $\widetilde{W}_I^*$ becomes more complicated because an increase in α raises $G_O(W) + G_P(W)$. However, if r and $d\alpha$ are sufficiently small, the increasing effect of α on $\widetilde{F}(W)$ dominates the increasing effect of α on $G_O(W) + G_P(W)$. Thus, $\widetilde{W}_I^*$ increases with α. For the result of ω on $\widetilde{W}_I^*$, an increase in ω increases $G_O(W) + G_P(W)$ because it not only raises the stream of expected cash flows by inducing the higher a_P after the IPO but also reduces the variation in W by providing the manager with more incentive to work through the complementarity effect of the higher a_P. In fact, an increase in ω increases $\widetilde{F}(W)$ because the sum of the principal's profits at and after the IPO equals $G_O(W) + G_P(W) - K$. However, if $d\omega$ is sufficiently small, the principal's profit generated from the IPO ($G_O(W) + G_P(W) - K$) is more sensitive to μ than that obtained when postponing the IPO ($\widetilde{F}(W)$) near the IPO threshold. Hence, $\widetilde{W}_I^*$ is decreasing in ω.

To summarize, the possibility of the VC not exiting the firm completely with the IPO invokes a multiagent setting when new outside investors cannot observe either the VC's or the manager's effort. However, most of our main conclusions continue to hold, even in this case. Therefore, our main results do not depend on the assumption that the principal (VC) completely exits with the IPO.

7. Empirical Implications

We begin by proposing empirical implications for the IPO timing, using the comparative static results regarding the IPO threshold.

(A) An IPO is likely to be earlier the less the need for the VC's monitoring role, the greater the increment in future expected cash flows, the higher the volatility of cash flows, and the larger the equity claim of the VC after the IPO.

Using US manufacturing firm data, Chemmanur, He, and Nandy (2010) report that firms with higher sales growth and higher total factor productivity are more likely to go public. This finding is consistent with our implication that the IPO is likely to be earlier when the increment in future expected cash flows increases. Chemmanur, He, and Nandy (2010) also suggest that firms in industries characterized by riskier cash flows are more likely to go public. Hsu (2013) shows that VCs shorten the incubation period and take their portfolio companies public when technological change at the industry level is greater. These findings are consistent with our implication that the IPO is likely to be earlier the higher the volatility of cash flows. Our implications concerning the need for the VC's monitoring role and the post-IPO equity position of the VC provide new testable predictions. The need for the VC's monitoring role can be measured by several indexes. For example, Lee and Wahal (2004) suggest that venture financing is disproportionately provided to firms in technology-intensive industries, in particular software and commercial biological research. They also indicate that VCs generally take smaller and younger firms public. These findings imply that the VC's monitoring role is needed more in firms in technology-intensive industries, smaller firms, and younger firms. Hence, our theory predicts that an IPO is likely to be later in these kinds of firms, where the VC's effort is more important.

A large number of IPOs do not necessarily have any VC contract relationships. Although our present model cannot suitably deal with the decision as to whether an entrepreneur and a VC make a contract relationship, our theory implies that firms with a high initial continuation value of the manager W_0 ($> W_I$) undertake the IPO immediately. If larger and mature firms have a higher W_0, this suggests that larger and mature firms need not receive the VC's support for undertaking the IPO, which is consistent with the empirical findings of Lee and Wahal (2004).[24]

Our theory also derives new implications for the managerial compensation profile, using the comparative static results regarding the cash payment and IPO thresholds.

(B) Managerial compensation will tend to be paid later, the smaller the increment in future expected cash flows, if the discounted expected profit of new outside investors to equity is more sensitive to this increment than the firm's expected productivity. In addition, managerial compensation will tend to be paid later after the IPO, the higher the volatility of cash flows, and the less acute the need for the VC's monitoring role if the discounted expected profit of new outside investors to equity is more sensitive to this need than the firm's expected productivity.

It is common for VC–entrepreneur relationships to include vesting provisions (see Kaplan and Strömberg (2003)). Our implications are novel in providing testable predictions that time vesting is used more the smaller the increment in future expected cash flows, and that it is also used more after the IPO the larger the volatility of cash flows and the less the need for the VC's monitoring role.

Under our optimal contract, the manager cannot receive any payment before the IPO nor any lump-sum payment at the IPO. Instead, the manager can receive a lump-sum payment only after the IPO. Practically, this delay of payment can be interpreted as a lockup period during which the manager may not sell his shares for a period of time after the IPO. In

[24]There are several empirical findings of long-run post-IPO underperformance (for example, see Ritter (1991) and Loughran and Ritter (1995) for the negative abnormal returns after the IPO, and Pástor, Taylor, and Veronesi (2009) for the lower profitability after the IPO). Even in our model, we show that $\widetilde{F}(W) \geq G(W) - K$ (or $\widetilde{F}(W) \geq G_O(W) + G_P(W) - K$) for all $W \geq R$ under the optimal contract. Hence, if the firm must promise to expend K at the time of the IPO but real expenditure evolves only gradually after the IPO and if the manager's continuation value W can be fully controlled under the estimating equation, our model suggests that the post-IPO performance of the firm is not better than the pre-IPO one from the point of view of the negative abnormal returns after the IPO.

any case, our model indicates that the contract can specify a large payment for the manager around the IPO. Furthermore, the VC has recently been more likely to exit through the sale of its portfolio firm to another company (M&A) than through the IPO.[25] As mentioned in the introduction, our model can be applied to the case where the VC exits through M&A. In this case, the manager is less likely to receive a large lump-sum payment at the exit of the VC.[26] Thus, the result that the manager does not receive any lump-sum payment at the exit of the VC does not restrict the explanatory power of our model.

Because the role of the VC can also be played by buyout funds or banks, our statements (A) and (B) with the associated arguments apply not only to IPOs or M&A for early-stage start-up firms but also to IPOs or M&A involving the relisting of management buyout firms or firms reorganized following bankruptcy.[27] Our model also allows us to derive new implications for the dissolution of joint ventures or corporate spin-offs. In the case of corporate spin-off, a parent company separates one of its subsidiaries from itself, sells the equity of the subsidiary to its own shareholders, and makes the subsidiary a new entity that is managed independently of the parent company. Thus, if the parent company's CEO (principal) attempts to maximize the expected payoff of the existing shareholders, the process is well characterized by the model in which the principal completely exits upon the spin-off timing. Daley, Mehrotra, and Sivakumar (1997) and Desai and Jain (1999) find that the number of focus-increasing spin-offs is more than twice that of non-focus-increasing spin-offs. Ahn and Walker (2007) also report that diversified firms with more effective corporate governance are more likely to conduct spin-offs. These findings are consistent with our hypothesis that a spin-off is likely to be earlier the smaller the need for the parent company's involvement.

[25] See Ball, Chiu, and Smith (2011) and Yearbook 2012, National Venture Capital Association.

[26] A VC states that the case in which the manager receives a large lump-sum payment at M&A is usually the case in which he quits the firm at M&A. Hence, except for this golden-parachute case, the manager is less likely to receive a large lump-sum payment at the exit of the VC via M&A. As our model focuses on the case in which the manager remains in the firm even after the VC exits, we can neglect this kind of golden-parachute case.

[27] M&A is also the most common route for the exit of buyout funds. See Kaplan and Strömberg (2009).

8. Conclusion

We explore a continuous-time agency model with double moral hazard in which an agent provides unobservable effort whereas a principal (advisor) supplies unobservable effort and chooses an optimal exit timing with irreversible investment. We take an example in which a VC or buyout fund and a manager jointly exert an unobservable level of effort to develop a project before the firm goes public to finance irreversible investment and the VC or buyout fund exits. We show that the optimal IPO timing is earlier under double moral hazard than under single moral hazard. Our results also indicate that the manager's compensation tends to be paid earlier under double moral hazard. We derive several comparative static results; in particular, the IPO timing is earlier when the need for monitoring by the VC is smaller and the volatility of cash flows is larger. Furthermore, we argue that even where the VC or buyout fund does not completely exit with the IPO, most of our main results remain unaffected. In addition, our model can be applied to not only the dissolution of joint ventures but also corporate spin-offs.

Appendix

Proof of Proposition 1: Our contract model after the IPO is essentially similar to the hidden effort model of DeMarzo and Sannikov (2006, Section III). The difference is that the manager's action increases cash flows in our model, whereas it decreases cash flows in DeMarzo and Sannikov (2006). Hence, we prove the statement of this proposition by applying a procedure similar to that of DeMarzo and Sannikov (2006).

The remaining problem is to derive a sufficient condition for the optimality of implementing $a_A(W_t) = 1$ at any $t \in [\tau_I, \tau_{T_1}]$. When the contract induces the manager to shirk, his promised payoff would evolve according to $dW_t = (\gamma W_t - h)dt - dC_t + \beta_1 \sigma dZ_t$. For $a_A(W_t) = 1$ to be optimal at any $t \in [\tau_I, \tau_{T1}]$, the expected payoff rate of new outside investors from letting the manager shirk would be less than that under our existing contract for all $W \in [R, W^{++}]$:

$$rG(W) \geq G'(W)(\gamma W - h). \tag{A1}$$

The manager's and the principal's expected payoffs if the manager shirks forever are given by $W^S = \frac{h}{\gamma}$ and $G^S = 0$. Using this, we rewrite (A1) as

$$G(W) + \frac{\gamma}{r} G'(W)(W^S - W) \geq 0. \tag{A2}$$

Applying a procedure similar to that of Proposition 8 in DeMarzo and Sannikov (2006), we can prove that a sufficient condition for (A2) is $\frac{\gamma}{r} G(W^S) + (1 - \frac{\gamma}{r}) G(W^{\max G}) \geq 0$, where $W^{\max G}$ denotes the value of W that achieves the greatest value of $G(W)$ in the range of $W \in [R, W^{++}]$. ■

Proof of Proposition 2: First, we can prove that the evolution of W_t and the incentive-compatibility condition are represented by (7) and $Y(W_t) \geq \beta_0(a_P(W_t))$ for $t \in [0, \tau_{T_0 I}]$ using a procedure similar to that in the proofs for Propositions 1 and 2 in the Appendix in Sannikov (2008).

We next consider the following HJB equation:

$$rF(W) = \max_{a_P \in A_P, Y \geq \beta_0(a_P), dC} 1 + (\zeta + \alpha) a_P - g(a_P) - dC + F'(W)\gamma W + \frac{F''(W)}{2} Y^2 \sigma^2. \tag{A3}$$

Then, we prove the regulatory properties of $F(W)$ in (A3).

Lemma A1: *The solutions to (A3) exist and are unique and continuous in initial conditions $F(W)$ and $F'(W)$. In addition, initial conditions with $F''(W) < 0$ result in a concave function.*

Proof: Using a procedure similar to that in the proof of Lemma 1 in the Appendix in Sannikov (2008), we can show the first statement. To prove the second statement, let us define $H(F, F', W) = \min_{a_P \in A_P, Y \geq \beta_0(a_P), dC} \frac{rF - 1 - (\zeta+\alpha)a_P + g(a_P) + dC - F'\gamma W}{\frac{1}{2}Y^2\sigma^2}$. Then, $F''(W) = H(F(W), F'(W), W)$. We can show that if ever $H(F, F', W) = 0$ on the path of a solution, the corresponding solution must be $F(W') = F(W) + \frac{\gamma}{r}(W' - W)F'(W)$. To see this, we must verify that $H(F + \frac{\gamma}{r}(W' - W)F', F', W') = 0$ for all W, $W' \geq R$. Indeed, $H(F + \frac{\gamma}{r}(W' - W)F', F', W') = \min_{a_P \in A_P, Y \geq \beta_0(a_P), dC} \frac{rF + \gamma(W'-W)F' - 1 - (\zeta+\alpha)a_P + g(a_P) + dC - F'\gamma W'}{\frac{1}{2}Y^2\sigma^2} = \min_{a_P \in A_P, Y \geq \beta_0(a_P), dC} \frac{rF - 1 - (\zeta+\alpha)a_P + g(a_P) + dC - F'\gamma W}{\frac{1}{2}Y^2\sigma^2} = H(F, F', W) = 0$. However, if $F(W') = F(W) + \frac{\gamma}{r}(W' - W)F'(W)$ for all W, $W' \geq R$, rearranging this relation yields $\frac{F'(W) - F(W)}{W' - W} = \frac{\gamma}{r}F'(W)$, for all W, $W' \geq R$. As $W' \to W$, this implies that $F'(W) = \frac{\gamma}{r}F'(W)$, which contradicts $\gamma > r$. Hence, the second derivative of the solution can never reach zero. We thus verify that a solution with a negative second derivative at one point must be concave everywhere. ||

We proceed to examine the existence and uniqueness of the solution that solves (A3) with the boundary conditions of (8).

Lemma A2: *There exists a unique concave function $F(W) \geq G(W) - K$ that solves (A3) with the boundary conditions of (8) for some $W_I > R$.*

Proof: Let us consider the solutions of (A3) with $F(R) = L_0$ and various slopes $G'(R) > F'(R)$. By Lemma A1, all of these solutions are unique and continuous in $F'(R)$. It also follows from Lemma A1 and $G''(W) < 0$ for $W \in [R, W^{++})$ that all of these solutions are concave if they satisfy $F(W_I) = G(W_I) - K$ and $F'(W_I) = G'(W_I)$.[28] Because $F(R) = L_0 > L_1 - K = G(R) - K$ by Assumption 1, it is evident that $W_I > R$. Now, as $G'(R) > F'(R)$, the resulting solution $F(W)$ must reach $G(W) - K$ at some point $W_I \in (R, \infty)$, as typically shown in Figure 1. ||

We next prove that $dC_t = 0$ for any $t \in [0, \tau_{T_0 I}]$.

[28] If $F''(W) \geq 0$ for all $W \geq R$, then $F(W) > G(W) - K$ for all $W \in (W_I, W^{++}]$, which is a contradiction.

Lemma A3: $W_I \leq W^+$. Thus, $dC_t = 0$ for all $W_t \in [R,\, W_I]$

Proof: If $W_I > W^+$, then $W_I - W^+$ must be paid as soon as the IPO is done. Thus, the principal's profit from the IPO is reduced to $G(W^+) - K$. However, it follows from $W_I > W^+$ and $F(W) > G(W) - K$ for all $W \in [R,\, W_I)$ that $F(W^+) > G(W^+) - K$, which means that the IPO should not be done. Thus, we must have $W_I \leq W^+$. Then, it follows from $F'(W) \geq -1$ and the definition of W^+ that $dC_t = 0$ for all $W_t \in [R,\, W_I]$. $\|$

Now, we conjecture an optimal contract from the solution of (8).

Lemma A4: *Consider the unique solution $F(W) \geq G(W) - K$ that solves (8) and satisfies the boundary conditions of (8) for some $W_I > R$. Let $a_P : [R,\, W_I] \to A_P$ and $Y : [R,\, W_I] \to [\beta_0,\, \infty]$ be the maximizers in (8); in particular, $Y(W_t) = \beta_0(a_P(W_t)) = \frac{h}{1+\alpha a_P(W_t)}$. For any starting condition $W_0 \in [R,\, W_I]$, there is a unique in the sense of probability law weak solution to the following equation:*

$$dW_t = \gamma W_t dt + \beta_0(a_P(W_t))\{dX_t - [1 + (\zeta + \alpha)a_P(W_t)]dt\}, \tag{A4}$$

until time $\tau_{T_0 I}$. The process of the manager's and the principal's efforts $\{a_A(\cdot), a_P(\cdot)\} \equiv \{a_A(W), a_P(W) : W \in [R, W_I]\}$ is defined by $a_{At} = 1$, $a_{Pt} = a_P(W_t)$, and $Y_t = \beta_0(a_P(W_t))$ for $t \in [0,\, \tau_{T_0 I})$ and for $t = \tau_{T_0 I}$ and $W_{\tau_{T_0 I}} = W_I$; and $a_{At} = 0$, $a_{Pt} = 0$, and $Y_t = 0$ for $t \geq \tau_{T_0 I}$ and $W_{\tau_{T_0 I}} = R$. This process is incentive compatible, and has value W_0 to the manager and the expected present value of the profit $F(W_0)$ to the principal.

Proof: There is a weak solution of (A4) that is unique in the sense of probability law because the drift of W_t is bounded on $[R,\, W_I]$ and the volatility is bounded above 0 by $\frac{h}{1+\alpha a_P(W_t)}\sigma$. Note that $a_P(W_t) \in A_P$ and A_P is a compact set. Now, using a procedure similar to that in the proof of Proposition 3 in the Appendix in Sannikov (2008), we can verify that the principal obtains the profit $F(W_0)$. $\|$

We finally show the following lemma.

Lemma A5: *Consider a concave solution $F(W)$ of the HJB equation that satisfies (8) with the boundary conditions of (8). Any incentive-compatible contract Π achieves the expected present value of the profit of at most $F(W_0)$.*

Proof: Applying the standard procedure of justifying the optimal contract (for example, see DeMarzo and Sannikov (2006), Sannikov (2008), and DeMarzo, Fishman, He, and Wang (2012)), we can prove this lemma (see Hori and Osano (2013)). $\|$

Combining Lemmas A1–A5, we establish the statement of Proposition 2.

The remaining problem is to derive a sufficient condition for the optimality of implementing $a_A(W_t) = 1$ at any $t \in [0, \tau_{T_0 I}]$. Repeating a procedure similar to that of the proof of Proposition 1 in this Appendix, we can prove that a sufficient condition for the optimality is $\frac{\gamma}{r}F(W^S) + (1 - \frac{\gamma}{r})F(W^{\max F}) \geq 0$, where $W^{\max F}$ denotes the value of W that achieves the greatest value of $F(W)$ in the range of $[R, W_I]$.[29] ■

Proof of Proposition 3: The proof is similar to that of Proposition 2 in this Appendix. Hence, we omit most of the proof and only provide the following sufficient condition for the optimality of implementing $a_A(W_t) = 1$ at any $t \in [0, \tau_{T_0 I}]$: $\frac{\gamma}{r}\widetilde{F}(W^S) + (1 - \frac{\gamma}{r})\widetilde{F}(W^{\max \widetilde{F}}) \geq 0$, where $W^{\max \widetilde{F}}$ denotes the value of W that achieves the greatest value of $\widetilde{F}(W)$ in the range of $[R, W_I]$. ■

Proof of Proposition 5: Comparing (10) with (8), we see that the only difference between the HJB equations and boundary conditions under double and single moral hazard is that the right-hand side of (10) is maximized subject to the principal's moral hazard constraint at each level of W, whereas the right-hand side of (8) is not. Hence, $\widetilde{F}(W) < F(W)$ for all $W \geq R$. In addition, using the boundary conditions of (8) and (10), $G''(W) < 0$, $F''(W) < 0$, $\widetilde{F}''(W) < 0$, $F(W) > G(W) - K$ for $W \in [R, W_I^*)$, and $\widetilde{F}(W) > G(W) - K$ for $W \in [R, \widetilde{W}_I^*)$, it follows that $\widetilde{W}_I^* < W_I^*$ if $\widetilde{F}(W) - G(W) < F(W) - G(W)$ for all $W \geq R$. Thus, it is evident from $\widetilde{F}(W) < F(W)$ that $\widetilde{W}_I^* < W_I^*$. ■

Proofs for the Comparative Static Results in Section 5: We begin with the case after the IPO. Applying procedures similar to those in the proofs for Lemmas 4 and 6 in the Appendix in DeMarzo and Sannikov (2006) or the proof for Lemma 3 in the Appendix in

[29] For simplicity, we assume that $\zeta < g'(0)$. Hence, if the manager shirks forever, it is not optimal for the principal to supply any of her own effort.

He (2009), we obtain

$$\frac{\partial G(W)}{\partial \theta} = E\left\{\int_{\tau_I}^{\tau_{T_1}} e^{-r(t-\tau_I)}\left[\frac{\partial \mu}{\partial \theta} + \frac{\partial \gamma W_t}{\partial \theta} G'(W_t) + \frac{1}{2}\frac{\partial\left(\frac{h\sigma}{\mu}\right)^2}{\partial \theta} G''(W_t)\right] dt\right.$$

$$\left. + e^{-r\left(\tau_{T_1}-\tau_I\right)} \left.\frac{\partial L_1}{\partial \theta}\right|_{W_t = W_{\tau_{T_1}}} \right| W_{\tau_I} = W \Bigg\}, \qquad \theta = \mu, \sigma^2. \tag{A5}$$

Note that $\beta_1 = \frac{h}{\mu}$ and $G(R) = L_1$ at $W = R$. It follows from (A5) that $\frac{\partial G(W)}{\partial \mu} > 0$ and $\frac{\partial G(W)}{\partial \sigma^2} < 0$. It also follows from (6), $G'(W^{++}) = -1$, and $G''(W^{++}) = 0$ that $rG(W^{++}) + \gamma W^{++} = \mu$. Differentiating this condition with $\gamma > r$, $G'(W^{++}) = -1$, $\frac{\partial G(W)}{\partial \mu} > 0$, and $\frac{\partial G(W)}{\partial \sigma^2} < 0$, we show that $\frac{\partial W^{++*}}{\partial \mu} < 0$ if $\frac{\partial G(W)}{\partial \mu} > \frac{1}{r}$ for all $W \geq R$, and that $\frac{\partial W^{++*}}{\partial \sigma^2} > 0$.

We next consider the case before the IPO. Repeating the procedure used above, we have

$$\begin{aligned}\frac{\partial \widetilde{F}(W)}{\partial \theta} &= E\left\{\int_0^{\tau_{T_0 I}} e^{-rt}\left[\frac{\partial\left[1+(\zeta+\alpha)a_P(W_t) - g(a_P(W_t))\right]}{\partial \theta} + \frac{\partial(\gamma W_t)}{\partial \theta}\widetilde{F}'(W_t)\right.\right.\\ &\quad + \frac{1}{2}\frac{\partial\left[\frac{h\sigma}{1+\alpha a_P(W_t)}\right]^2}{\partial \theta}\widetilde{F}''(W_t)\, dt + \mathbf{1}_{\tau_I \geq \tau_{T_0}} e^{-r\tau_{T_0}}\frac{\partial L_0}{\partial \theta}\end{aligned}$$

$$\left. + \mathbf{1}_{\tau_I < \tau_{T_0}} e^{-r\tau_I} \left.\frac{\partial\left[G(W_t) - K\right]}{\partial \theta}\right|_{W_t = W_{\tau_I}} \right| W_0 = W \Bigg\}, \qquad \theta = \mu, \sigma^2, \alpha. \tag{A6}$$

Note that $\beta_0(a_P) = \frac{h}{1+\alpha a_P}$, $\zeta + \alpha = g'(a_P)$, $\widetilde{F}(R) = L_0$ at $W = R$, and $\widetilde{F}(W_{\tau_I}) = G(W_{\tau_I}) - K$ at $W = W_{\tau_I}$. It follows from (A5) and (A6) that $\frac{\partial \widetilde{F}(W)}{\partial \mu} > 0$ and $\frac{\partial \widetilde{F}(W)}{\partial \sigma^2} < 0$. Given $\zeta + \alpha = g'(a_P)$, we obtain $\frac{\partial \widetilde{F}(W)}{\partial \alpha} > 0$. It also follows from the boundary conditions of (10), $\widetilde{F}''(W) < 0$, $G''(W) < 0$, and $\widetilde{F}(W) > G(W) - K$ for all $W \in [R, W_I)$ that for $\theta = \alpha$, μ, and σ^2, $\frac{\partial \widetilde{W}_I^*}{\partial \theta} \gtreqless 0$ if $\frac{\partial\left[\widetilde{F}(W) - G(W)\right]}{\partial \theta} \gtreqless 0$ for all $W \in [R, W_I)$. In particular, when $\theta = \alpha$, we obtain $\frac{\partial G(W)}{\partial \alpha} = 0$. Thus, using $\frac{\partial \widetilde{F}(W)}{\partial \alpha} > 0$, this implies that $\frac{\partial \widetilde{W}_I^*}{\partial \alpha} > 0$. Regarding $\theta = \mu$ and σ^2, we focus on the perturbation in the small neighborhood of $W = W_{\tau_I}$ ($\equiv W_I$) when $\tau_I < \tau_{T_0}$. For $\theta = \mu$, it follows from (A5) and (A6) that $\left.\frac{\partial\left[\widetilde{F}(W) - G(W)\right]}{\partial \mu}\right|_{W = W_{\tau_I}} \simeq \left(e^{-r\tau_I} - 1\right)\left.\frac{\partial G(W)}{\partial \mu}\right|_{W = W_{\tau_I}} < 0$ if $d\mu$ is sufficiently small. Thus, $\frac{\partial \widetilde{W}_I^*}{\partial \mu} < 0$ if $d\mu$ is sufficiently small. For $\theta = \sigma^2$, it follows from (A5) and (A6) that if r and $d\sigma^2$ are sufficiently small, then $\left.\frac{\partial\left[\widetilde{F}(W) - G(W)\right]}{\partial \sigma^2}\right|_{W = W_{\tau_I}} \simeq (e^{-r\tau_I}$

$- 1) \left.\frac{\partial G(W)}{\partial \sigma^2}\right|_{W=W_{\tau_I}} + \frac{1-e^{-r\tau_I}}{r}\frac{1}{2}\left[\frac{h}{1+\alpha a_P(W_{\tau_I})}\right]^2 \widetilde{F}''(W_{\tau_I})$. In fact, on the right-hand side of the above relation, the first term has a higher order of r than the second term. As this implies $\left.\frac{\partial\left[\widetilde{F}(W)-G(W)\right]}{\partial \sigma^2}\right|_{W=W_{\tau_I}} < 0$, we have $\frac{\partial \widetilde{W}_I^*}{\partial \sigma^2} < 0$ if r and $d\sigma^2$ are sufficiently small. ■

Appendix for Section 6: Regardless of single or double moral hazard before the IPO, we can prove that there exists a progressively measurable process $\{Y_t, \mathcal{F}_t : \tau_I \leq t \leq \tau_{T_1}\}$ in $\mathcal{L}^*$ such that

$$dW_t = \{\gamma W_t - h \cdot [1 - a_A(W_t)]\}\, dt - dC_t + Y(W_t)\{dX_t - \mu[a_A(W_t) + \zeta a_P(W_t) + \alpha a_A(W_t) a_P(W_t)]dt\},$$

for every $t \in [\tau_I, \tau_{T_1}]$. Then, the incentive-compatibility constraint for the manager implementing the high effort after the IPO is summarized by $Y = \eta_1(a_P)$, where $\eta_1(a_P) \equiv \min\{y : y\mu(1 + \alpha a_P) \geq h\} = \frac{h}{\mu(1+\alpha a_P)}$.

Let $G_O(W)$ and $G_P(W)$ denote the value functions of new outside investors and the principal after the IPO, respectively. The optimal compensation policy is still determined by $dC = \max(W - W_O^{++}, 0)$, where W_O^{++} is the lowest value of W such that $G_O'(W) = -1$. As $G_O(W)$ does not depend on whether there is single or double moral hazard before the IPO, and neither does W_O^{++}.

Then, we can derive the following HJB equations after the IPO. First, for any $W_I \in [R, W_O^{++}]$, there is the unique concave function $G_O(W)$ that satisfies

$$rG_O(W) = \max_{a_P = \psi(\omega\mu(\zeta+\alpha)), Y \geq \eta_1(a_P)} (1-\omega)\mu\left[1 + (\zeta + \alpha)a_P\right] + G_O'(W)\gamma W + \frac{G_O''(W)}{2} Y^2\sigma^2, \quad \text{(A7)}$$

where $\psi(\omega\mu(\zeta + \alpha))$ is given below (A9), and the boundary conditions are represented by $G_O(R) = (1 - \omega)L_1$, $G_O'(W_O^{++}) = -1$, and $G_O''(W_O^{++}) = 0$.[30] Second, for any $W_I \in [R, W_O^{++}]$, there is the unique concave function $G_P(W)$ that satisfies

$$rG_P(W) = \max_{a_P \in A_P} \omega\mu\left[1 + (\zeta + \alpha)a_P\right] - g(a_P) + G_P'(W)\gamma W + \frac{G_P''(W)}{2} Y^2\sigma^2, \qquad \text{(A8)}$$

with the boundary conditions of $G_P(R) = \omega L_1$ and $G_P'(W_O^{++}) = 0$ (for the derivation procedure of (A8) and the boundary conditions, see Hori and Osano (2013)).

[30] As the smooth-pasting and super contract conditions are the same as those in the model of DeMarzo and Sannikov (2006), we can prove the concavity of $G_O(W)$ as in Proposition 1.

Now, the principal optimally chooses her effort after the IPO as

$$a_P = \arg\max_{a_P \in A_P} \omega\mu \left[1 + (\zeta + \alpha)a_P\right] - g(a_P) + G'_P(W)\gamma W + \frac{G''_P(W)}{2} \left[Y(W)\right]^2 \sigma^2. \tag{A9}$$

Note that $Y(W) = \eta_1(a_P(W)) = \frac{h}{\mu[1+\alpha a_P(W)]}$, where $a_P(W)$ is given by the recommended principal's effort at each point of W after the IPO. If we assume that $\omega\mu(\zeta + \alpha) > g'(0)$, the incentive-compatibility constraint for the principal after the IPO is represented by $\omega\mu(\zeta+\alpha) = g'(a_P)$ or $a_P = g'^{-1}(\omega\mu(\zeta + \alpha)) = \psi(\omega\mu(\zeta + \alpha))$.

Similarly, we derive the HJB equations before the IPO. Note that the evolution of W_t is still given by (7), irrespective of single or double moral hazard. Hence, if the manager can observe and verify $\{a_{Pt} \in A_P,\ 0 \leq t \leq \tau_I\}$, the unique concave function $F(W)$ ($\geq G_O(W) + G_P(W) - K$) satisfies

$$rF(W) = \max_{a_P \in A_P, Y \geq \beta_0(a_P)} 1 + (\zeta + \alpha)a_P - g(a_P) + F'(W)\gamma W + \frac{F''(W)}{2} Y^2\sigma^2, \tag{A10}$$

with $\beta_0(a_P) = \frac{h}{1+\alpha a_P}$ and the boundary conditions of $F(R) = L_0$, $F(W_I) = G_O(W_I) + G_P(W_I) - K$, and $F'(W_I) = G'_O(W_I) + G'_P(W_I)$. If the manager cannot observe $\{a_{Pt} \in A_P,\ 0 \leq t \leq \tau_I\}$, the unique concave function $\widetilde{F}(W)$ ($\geq G_O(W) + G_P(W) - K$) satisfies

$$r\widetilde{F}(W) = \max_{a_P = \psi(\zeta+\alpha), Y \geq \beta_0(a_P)} 1 + (\zeta + \alpha)a_P - g(a_P) + \widetilde{F}'(W)\gamma W + \frac{\widetilde{F}''(W)}{2} Y^2\sigma^2, \tag{A11}$$

where $\psi(\zeta + \alpha) = g'^{-1}(\zeta + \alpha)$ and the boundary conditions are given by $\widetilde{F}(R) = L_0$, $\widetilde{F}(W_I) = G_O(W_I) + G_P(W_I) - K$, and $\widetilde{F}'(W_I) = G'_O(W_I) + G'_P(W_I)$. ■

References

Ahn, S., and M. D. Walker. 2007. Corporate Governance and the Spinoff Decision. *Journal of Corporate Finance* 13: 76–93.

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

Barry, C. B., C. J. Muscarella, J. W. Peavy III, and M. R. Vetsuypens. 1990. The Role of Venture Capital in the Creation of Public Companies. *Journal of Financial Economics* 27: 447–471.

Benninga, S., M. Helmantel, and O. Sarig. 2005. The Timing of Initial Public Offerings. *Journal of Financial Economics* 75: 115–132.

Bhattacharyya, S., and F. Lafontaine. 1995. Double-Sided Moral Hazard and the Nature of Share Contracts. *RAND Journal of Economics* 26: 761–781.

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

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

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

Chemmanur, T. J., S. He, and D. K. Nandy. 2010. The Going-Public Decision and the Product Market. *Review of Financial Studies* 23: 1855–1908.

Clementi, G. L. 2002. IPOs and the Growth of Firms. mimeo, Carnegie Mellon University.

Daley, L., V. Mehrotra, and R. Sivakumar. 1997. Corporate Focus and Value Creation: Evidence from Spinoffs. *Journal of Financial Economics* 45: 257–281.

DeMarzo, P. M., M. J. Fishman, Z. He, and N. Wang. 2012. Dynamic Agency and the q Theory of Investment. *Journal of Finance* 67: 2295–2340.

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

Desai, H., and P. C. Jain. 1999. Firm Performance and Focus: Long-run Stock Market Performance Following Spinoffs. *Journal of Financial Economics* 54: 75–101.

Gompers, P. A., and L., Joshua. 1999. *The Venture Capital Cycle.* Cambridge: MIT Press.

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

He, Z. 2011. A Model of Dynamic Compensation and Capital Structure. *Journal of Financial Economics* 100: 351–360.

Hellmann, T. 2006. IPOs, Acquisitions, and the Use of Convertible Securities in Venture Capital. *Journal of Financial Economics* 81: 649–679.

Hellmann, T., and M. Puri. 2002. Venture Capital and the Professionalization of Start-up Firms: Empirical Evidence. *Journal of Finance* 57: 169–198.

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

Holmstrom, B., and P. Milgrom. 1987. Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica* 55: 303–328.

Hori, K., and H. Osano. 2013. Managerial Incentives and the Role of Advisors in the Continuous-Time Agency Model. Discussion Paper, Kyoto University and Ritsumeikan University.

Hsu, H.-C. S. 2013. Technology Timing of IPOs and Venture Capital Incubation. *Journal of Corporate Finance* 19: 36–55.

Inderst, R., and H. M. Müller. 2004. The Effect of Capital Market Characteristics on the Value of Start-up Firms. *Journal of Financial Economics* 72: 319–356.

Jovanovic, B., and J. Prat. 2010. Dynamic Incentive Contracts under Parameter Uncertainty. CEPR Discussion Paper No. 8136, Centre for Economic Policy Research.

Kaplan, S. N., and P. Strömberg. 2003. Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts. *Review of Economic Studies* 70: 281–315.

Kaplan, S. N., and P. Strömberg. 2004. Characteristics, Contracts, and Actions: Evidence from Venture Capitalist Analyses. *Journal of Finance* 59: 2177–2210.

Kaplan, S. N., and P. Strömberg. 2009. Leveraged Buyouts and Private Equity. *Journal of Economic Perspectives* 25: 121–146.

Kim, S. K., and S. Wang. 1998. Linear Contracts and the Double Moral Hazard. *Journal of Economic Theory* 82: 342–378.

Lee, P. M., and S. Wahal. 2004. Grandstanding, Certification and the Underpricing of Venture Capital Backed IPOs. *Journal of Financial Economics* 73: 375–407.

Loughran, T., and J. R. Ritter. 1995. The New Issue Puzzle. *Journal of Finance* 50: 23–51.

Pástor, L., L. A. Taylor, and P. Veronesi. 2009. Entrepreneurial Learning, the IPO Decision, and the Post-IPO Drop in Firm Profitability. *Review of Financial Studies* 22: 3005–3046.

Philippon, T., and Y. Sannikov. 2007. Real Options in a Dynamic Agency Model, with Applications to Finance Development, IPOs, and Business Risk. NBER Working Paper No. 13584, National Bureau of Economic Research.

Piskorski, T., and A. Tchistyi. 2010. Optimal Mortgage Design. *Review of Financial Studies* 23: 3098–3140.

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

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

Romano, R. E. 1994. Double Moral Hazard and Resale Price Maintenance. *RAND Journal of Economics* 25: 455–466.

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

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

Spear, S. E., and C. Wang. 2005. When to Fire a CEO: Optimal Termination in Dynamic Contracts. *Journal of Economic Theory* 120: 239–256.

Townsend, R. M. 1979. Optimal Contracts and Competitive Markets with Costly State Verification. *Journal of Economic Theory* 21: 265–293.

Table 1: Comparative Statics under Double Moral Hazard When the VC Exits the Firm Following the IPO

θ	α	μ	σ^2
	After	the	IPO
$\partial G(W)/\partial\theta$	0	>0	<0
$\partial\widetilde{W}^{++*}/\partial\theta$	0	$<0^a$	>0
	Before	the	IPO
$\partial\widetilde{F}(W)/\partial\theta$	>0	>0	<0
$\partial\widetilde{W}_I^*/\partial\theta$	>0	$<0^b$	$<0^c$

Notes:
[a] We assume that $\frac{\partial G(W)}{\partial\mu} > \frac{1}{r}$ for all $W \geq R$.
[b] We assume that $d\mu$ is sufficiently small.
[c] We assume that r and $d\sigma^2$ are sufficiently small.

Table 2: Comparative Statics under Double Moral Hazard When the VC Does Not Exit the Firm Following the IPO

θ	α	μ	σ^2	ω
After the IPO				
$\partial G_O(W)/\partial\theta$	>0	>0	<0	?
$\partial\widetilde{W}_O^{++*}/\partial\theta$	$<0^a$	$<0^b$	>0	?
Before the IPO				
$\partial\widetilde{F}(W)/\partial\theta$	>0	>0	<0	>0
$\partial\widetilde{W}_I^*/\partial\theta$	$>0^c$	$<0^d$	$<0^c$	$<0^d$

Notes:

[a] We assume that $\frac{\partial G_O(W)}{\partial\alpha} > \frac{\partial\left[(1-\omega)A^{N**}(W)\right]}{\partial\alpha}$, where $A^{N**}(W) = \mu[1 + (\zeta + \alpha)a_P^{**}(W)]$.

[b] We assume that $\frac{\partial G_O(W)}{\partial\mu} > \frac{\partial\left[(1-\omega)A^{N**}(W)\right]}{\partial\mu}$.

[c] We assume that r and $d\theta$ are sufficiently small for $\theta = \alpha$ and σ^2.

[d] We assume that $d\theta$ is sufficiently small for $\theta = \mu$ and ω.

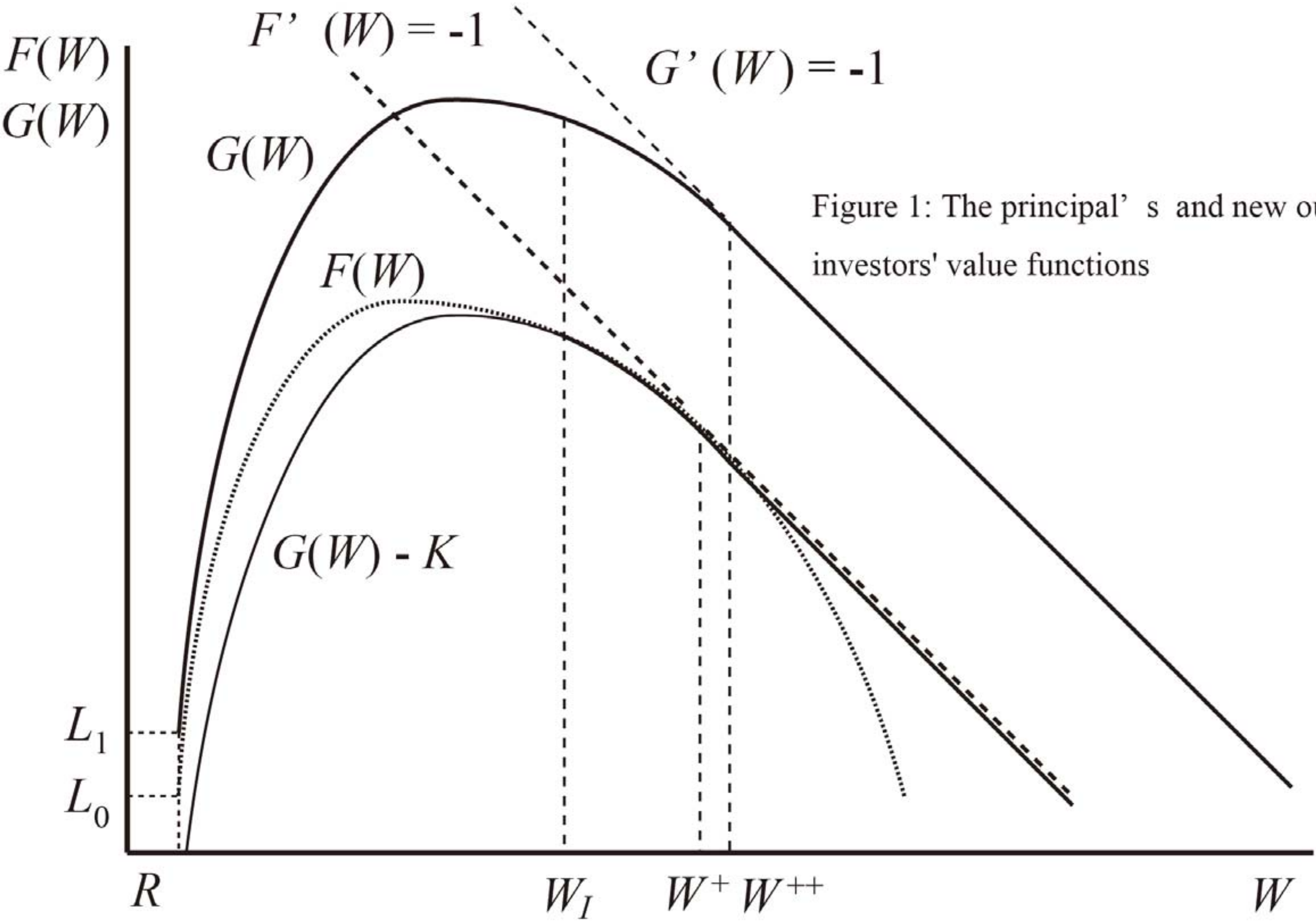

Figure 1: The principal' s and new outside investors' value functions