Optimal Financial Contracts and The Dynamics of Insider Ownership*

Charles Himmelberg
Federal Reserve Bank of New York

Vincenzo Quadrini
New York University, CEPR and NBER

December, 2002

Abstract

This paper characterizes the optimal dynamic contract between risk-averse entrepreneurs and risk-neutral investors in an infinite horizon setting with information asymmetry. Under certain conditions, the optimal contract can be implemented with the entrepreneur's ownership of some of the firm's shares (insider ownership). This ownership evolves over time and tends to decrease as the entrepreneur's wealth increases. Another result of the paper is that the concentration of ownership declines with the degree of investor protection. These results are supported by the findings of several empirical studies.

JEL: D82, E22, G32

Key words: Optimal contracts, ownership structure, firm dynamics

*Corresponding address: Vincenzo Quadrini, Department of Economics, Stern School of Business, New York University, 44 West Fourth Street, New York, NY 10012. E-mail: vquadrin@stern.nyu.edu

1 Introduction

The ownership structure of a firm is central to its investment decision and dynamics. At the same time, the ownership structure is likely to change as the firm evolves over time. The close connection between the dynamics of the firm and its ownership is well characterized in newly public firms. As shown by Mikkelson, Partch, & Shah (1997), the share ownership of “insiders” tends to decline over time after the firm has become public. Furthermore, the insiders’ ownership declines faster for firms that experience higher growth. Another finding about the ownership structure of the firm is that countries with higher degree of investors’ protection are characterized by lower concentration of ownership as shown by La-Porta, de Silanes, Shleifer, & Vishny (1998). The question we ask in this paper is whether these patterns are captured by the properties of optimal contracts between the subjects that control the resources in the firm—the insiders—and those who provide the funds—the external investors.

We study the investment and financing problem of a firm in which a risk-averse entrepreneur (the insider) has an information advantage about the operation of the firm over external investors. More specifically, the shock to the firm’s revenues is publicly observed only with some probability. This information advantage allows the entrepreneur to divert part of the firm resources and generate a private benefit which is increasing in the invested capital. Within this framework we characterize the optimal contract, that is, the contract that maximizes the investors’ wealth subject to promise-keeping and incentive-compatibility constraints.

Under particular assumptions about the shock to the firm’s revenue, the optimal contract can be implemented with the entrepreneur’s ownership of a fraction of the firm’s shares (entrepreneur’s stake in the firm or insider ownership). The entrepreneur’s ownership evolves over time according to the performance of the firm and tends to decline as the firm becomes larger and older. These dynamics features of the ownership structure replicates the above mentioned empirical findings of Mikkelson et al. (1997).

These results derive from the existence of information asymmetries that generate moral hazard problems. When the shock is publicly revealed with probability one and information is symmetric, the optimal contract would guarantee the efficient production scale at any point in time. Moreover, the entrepreneur would enjoy a perfect consumption smoothing given that investors are risk neutral. With information asymmetry, however, this contract is not incentive-compatible, and the entrepreneur must bear at least some of the firm's risk through the ownership of the firm. As the entrepreneur's wealth increases, the contract can be made incentive-compatible with a smaller insider ownership and the entrepreneur will fully diversify once his or her wealth is sufficiently large.

The full diversification of portfolio in the limit crucially depends on the assumption that there is a positive probability with which the shock becomes public information. In this case, if an entrepreneur is caught diverting, he or she can be punished with the confiscation of a large portion of wealth. This assumption implies that the incentive to divert decreases as the entrepreneur's wealth increases. Therefore, for wealthy entrepreneurs the contract can be made incentive-compatible with a smaller insider ownership. By contrast, when the shock is never publicly revealed, the entrepreneur would never be able to implement a full diversification of portfolio.

The probability with which the shock gets publicly revealed effectively determines the degree of investor's protection. Another result of the paper is that higher is this probability and lower is the concentration of ownership. This is consistent with the higher concentration observed in countries with lower institutional protection of investors. La-Porta et al. (1998) show that in common law countries there is greater investors' protection and lower concentration of ownership.

The result that the optimal contract can be implemented through the entrepreneur's ownership of some of the firm's shares only applies when the shock takes two values. When the shock takes more than two values, the optimal contract can be implemented with the additional use of stock options. The use of stock options as an implementation mechanism is also studied in Aseff & Santos (2001) but in a different environment.

From a methodological point of view, the paper relates to two branches of existing literature. The first branch studies the optimal consumption insurance among risk-averse agents when individual endowments or efforts are unobservable.[1] The second branch assumes risk-neutral agents and studies the optimal investment schedule that maximizes the resources generated by the firm.[2] The current paper combines the main features of these two

[1]Examples of these studies include Atkeson & Lucas (1992, 1995), Clementi & Cooley (2000), Cole & Kocherlakota (1997), Green (1987), Phelan (1995), Phelan & Townsend (1991), Spear & Srivastava (1987), Thomas & Worrall (1990), Wang (1995, 1997).

[2]Examples of these studies include Clementi & Hopenhayn (1998), DeMarzo & Fisherman (2000) and Quadrini (1999).

branches and the contract solves the trade-off between the optimal consumption insurance—as entrepreneurs are risk averse—and the optimal investment schedule—as resources depend on investment. In this respect the paper is also related to Atkeson (1991), Marcet & Marimon (1992) and Castro, Clementi, & MacDonald (2001).

The plan of the paper is as follows. Section 2 describes the basic model and Section 3 characterizes the optimal contract. Section 4 shows how this contract can be implemented through the entrepreneur's ownership of some of the firm's shares and Section 5 studies a numerical example. Section 6 extends the basic model along several dimensions and Section 7 concludes.

2 The Model

Consider an entrepreneur-manager with lifetime utility:

$$E\sum_{t=0}^{\infty}\beta^t U(c_t) \tag{1}$$

where $c_t \geq 0$ is consumption and β is the intertemporal discount factor. The function U is (i) strictly increasing and concave; (ii) $\lim_{c\to\infty} U'(c) > \alpha > 0$; (iii) $\lim_{c\to 0} U(c) = -\infty$. The assumption that the marginal utility from consumption is bounded away from zero will be motivated below. The assumption that the utility is unbounded below will be relaxed in Section 6.

The entrepreneur has an investment opportunity that requires an initial set up investment κ, which is sunk, and generates revenues according to:

$$y = F(k, z) \tag{2}$$

Here y is the cash revenue, k is the input of capital, z is an i.i.d. shock with probability distribution $p(z)$. The function F is strictly increasing, concave, differentiable and $\lim_{k\to 0} F_k(k, z) = \infty$. The input of capital is decided one period in advance before the realization of the shock. The main analysis of the paper is conducted by assuming that the shock takes only two values, that is, $z \in \{z_1, z_2\}$. Section 6 will extend the model by allowing for more than two shocks.

The shock is observed with probability one by the entrepreneur. Investors, instead, observe the shock only with probability η, which is constant in the model. This information asymmetry allows the entrepreneur to divert part of the revenues. Suppose that the entrepreneur observes $z = z_2$. By claiming that $z = z_1$, he or she can divert the revenue $D(k) = F(k, z_2) - F(k, z_1)$. If the shock is not revealed, the publicly observed revenue will be the difference between the true revenue and the part that is diverted, that is, $F(k, z_2) - D(k) = F(k, z_1)$. Notice that the decision to divert revenues is taken before the shock is publicly revealed with probability η.

The diverted revenue provides utility to the entrepreneur additive to the utility of consumption. More specifically, if the entrepreneur diverts $D(k)$, the additional utility from diversion is $\alpha D(k)$ and the current total utility is $U(c) + \alpha D(k)$. Given this particular specification of the return from diversion, the assumption $\lim_{c\to\infty} U'(c) > \alpha$ guarantees that diversion is always inefficient.[3] Section 6, however, will discuss the alternative assumption in which the diverted revenue is additive to consumption and the current utility is $U(c + \alpha D(k))$. In this case the assumption that the marginal utility from consumption is bounded away from zero can be relaxed.

The investment project is financed by entering into a contractual relationship with a risk-neutral investor. The investor can be thought of as a financial intermediary that discounts future payments at the market interest rate r. We denote by $\delta = 1/(1 + r)$ the discount factor for the investor and assume that $\delta \geq \beta$. This can be considered a general equilibrium property which is usually satisfied in models with uninsurable idiosyncratic risks. Finally, we also assume that the value of the contract for the entrepreneur cannot be smaller than q_R (reservation value). Any value of the contract below q_R is not enforceable. The investor, instead, commits to fulfill any future obligation (one-side commitment). However, the assumption of commitment from the investor is not important if we allow for bonding.[4]

3 The optimal contract

With symmetric information, that is, $\eta = 1$, the input of capital will be always at the optimal level $\bar{k}$. This input of capital is determined by the first order condition $\delta \sum_z F_k(\bar{k}, z)p(z) = 1$. With information asymmetry,

[3]As we will see later, this particular specification of the return from diversion guarantees that the feasible set for the optimal contract is convex which is convenient to establish some of the basic properties of the contract.

[4]For bonding we mean the ability of the entrepreneur to access a "riskless" and "publicly observable" investment at the market interest rate.

however, the input of capital may be lower than $\bar{k}$. The goal of this section is to characterize the optimal contract when the shock is publicly observed only with probability $\eta < 1$ (information asymmetry).

3.1 Recursive formulation of the long-term contract

We start the analysis of the optimal contract by ignoring, for the moment, the issue of renegotiation. In this contract the parties commit to fulfill the terms of the contract in all possible contingencies even if ex-post it could be beneficial for both parties to change these terms. We will refer to this contract as the "long-term contract". We will show then in the next section that this contract is free from renegotiation.

Define $\hat{z}$ the shock inferred from the observation of the revenue. Of course, if the entrepreneur does not divert revenues, the inferred shock is the true realization, that is, $\hat{z} = z$. Furthermore, define e the dummy variable that takes the value of one when the entrepreneur is caught diverting, that is, when the shock is publicly revealed and $\hat{z} \neq z$. The contract can be made conditional on the whole history of $\hat{z}$ and e. We denote the history by $\mathbf{h}_t = \{\hat{z}_1, e_1, ..., \hat{z}_t, e_t\}$. The structure of the contract is as follows. At the end of each period and for each history $\mathbf{h}_t$, the investor anticipates the input of capital k. In the next period, after the observation of the revenue and the shock if publicly revealed, the revenue will be distributed in part to the entrepreneur and in part to the investor.

The contractual problem under commitment is formalized as the maximization of the investor's value, subject to the constraint that the entrepreneur does not divert revenues and subject to the participation constraint. The problem is made stationary by introducing promised utility as a state variable. Denote the promised utility *after* current consumption by q. Then for each q, the contract will choose the input of capital k and the next period entrepreneur's consumption and continuation utility. Consumption and continuation utility are conditional on the shock inferred from the observation of revenues and on diversion. We will denote by $c(\hat{z})$ and $q(\hat{z})$ the current consumption and continuation utility when the entrepreneur is not caught diverting. This can happen either because the shock is not publicly revealed or, if revealed, the entrepreneur had not diverted revenues. Moreover, we will denote by c^D and q^D the consumption and continuation utility when the entrepreneur is caught diverting. This can happen only if the shock is publicly revealed.

Because there is a monotone relation between consumption and the period utility of the entrepreneur, we will use the period utility as a choice variable of the contract. We will denote by $u(z) = U(c(z))$ the period utility conditional on the shock z. Moreover, we will denote by $C(u(z))$ the inverse of the utility function, that is, $c(z) = C(u(z)) = U^{-1}(u(z))$. The function C returns the cost of utility $u(z)$. The contractual problem can be written as:

$$V(q) \quad = \quad \max_{k,u(z),q(z)} \Big\{ -k + \delta \sum_z \Big[F(k,z) - C(u(z)) + V(q(z))\Big]p(z)\Big\} \quad (3)$$

subject to

$$q \;=\; \beta \sum_z \Big[u(z) + q(z)\Big]p(z) \quad (4)$$

$$u(z_2) + q(z_2) \geq (1-\eta)\cdot\Big[u(z_1) + q(z_1) + \alpha D(k)\Big] + \eta\cdot\Big[u^D + q^D\Big] \quad (5)$$

$$u(z) + q(z) \geq q_R; \quad u^D + q^D \geq q_R \quad (6)$$

The function $V(q)$ is the end-of-period value of the contract for the investor. This value results from the current flow, $-k + \delta E[F(k,z) - C(u(z))]$, plus the discounted next period value, $\delta EV(q(z))$.

Equation (4) is the promised-keeping constraint and equation (5) is the incentive-compatibility constraint. The incentive-compatibility constraint imposes that, when the shock is high, the entrepreneur will get an expected lifetime utility at least as large as the utility he or she will get if resources are diverted. When the entrepreneur diverts resources, his or her value depends on whether the shock gets publicly revealed. If the shock is not revealed, the utility from consumption is augmented by the utility from the diverted revenue, $\alpha D(k)$. However, if the shock is publicly revealed and the entrepreneur is caught diverting, the current and continuation utilities are set to u^D and q^D. It is trivial to prove that in the long-term contract the optimal strategy conditional on diversion is $u^D + q^D = q_R$. This is because this strategy minimizes the entrepreneur's incentive to divert revenues. Therefore, in what follows we will take this optimal strategy as given. Notice that when the shock is publicly revealed, the diverted revenue is recovered and does not enter the entrepreneur's utility. The case in which the revenue can not be recovered will be discussed in Section 6.

Constraints (6) impose limited liability. Here the assumption is that the entrepreneur can leave the firm and get the reservation value q_R at any point in time. Therefore, $u(z) + q(z) \geq q_R$ and $u^D + q^D \geq q_R$.[5] In the contractual problem we should also impose the non-negativity of capital and consumption, that is, $k \geq 0$ and $C(u(z)) \geq 0$. However, given the properties of the revenue and utility functions, these constraints are never binding.

Before establishing the existence of a solution to problem (3), we impose a further restriction to the revenue function. Let $x = D(k) = F(k, z_2) - F(k, z_1)$. We can then define k as a function of x, that is, $k = D^{-1}(x)$ and the present value of revenues net of the cost of capital can be expressed as:

$$\pi(x) = -D^{-1}(x) + \delta \sum_z F(D^{-1}(x), z)p(z) \tag{7}$$

This is simply the function $-k + \delta \sum_z F(k, z)p(z)$ after substituting k with $D^{-1}(x)$.

Assumption 1 *The revenue function $F(k, z)$ is such that $\pi(x)$ is concave.*

An example in which $\pi(x)$ is concave is when the revenue function takes the form $F(k, z) = (1 - \omega)k + zf(k)$, with f strictly increasing and concave. In this case $x = (z_2 - z_1)f(k)$ and it can be verified that $\pi(x)$ is strictly concave in x. We then have the following proposition.

Proposition 1 (Optimal long-term contract) *The contractual problem (3) is a contraction and there is a unique fixed point $V(q)$. The solution to problem (3) is unique and $V(q)$ is strictly concave and differentiable. Moreover, there exists $\underline{\underline{q}}$ for which $V(q)$ is strictly decreasing for $q > \underline{\underline{q}}$.*

[5] This constraint then implies that $q(z) \geq \beta q_R$. To see this, let's observe that in the next period $u(z') + q(z') \geq q_R$ for each value of z', which implies that $\beta E(u(z') + q(z')) \geq \beta q_R$. Here the prime denotes the next period variable. Given that promise-keeping must also hold in the next period, that is, $q(z) = \beta E(u(z') + q(z'))$, the last constraint imposes $q(z) \geq \beta q_R$. This constraint, however, is never binding. In fact, if $q(z) = \beta q_R$, then the next period promise-keeping constraint implies that $u(z_1) + q(z_1) = u(z_2) + q(z_2) = q_R$. But then the incentive-compatibility constraint (5) requires $\alpha D(k') = 0$, which is possible only if the next period capital is zero. Given the structure of the revenue function this can not be optimal.

PROOF: *Appendix A.*

Figure 1 plots the value of the long-term contract for the investor. According to proposition 1 this function is concave. Moreover, the function is strictly decreasing for q greater than the threshold $\underline{\underline{q}}$.

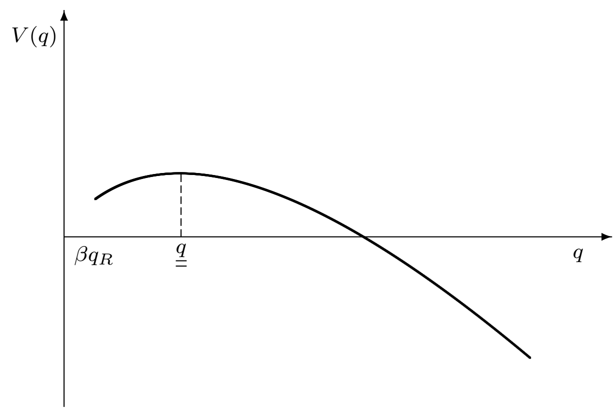


Figure 1: Investor's value as a function of the entrepreneur's value.

3.2 Renegotiation-proof of the long-term contract

The function $V(q)$ represents the utility frontier of the set of all possible allocations that can be reached with long-term contracts. The fact that the function $V(q)$ is not decreasing for $q < \underline{\underline{q}}$ implies that the utility frontier is not downward sloping. Therefore, the optimal contract is not necessarily free from renegotiation (see, Fudenberg, Holmstrom, & Milgrom (1990)). To check whether the long-term contract is free from renegotiation, we have to answer two questions. First, are values of q smaller than $\underline{\underline{q}}$ ever reached in the long-term contract? Second, if the entrepreneur is caught diverting, would the parties mutually benefit from increasing $u^D + q^D$ above q_R? This is formalized in the following definition of renegotiation-proofness for a recursive contract.

Definition 1 (Renegotiation-profness) *A recursive contract is free from renegotiation if (i) for all z there is not $\hat{q} > q(z)$ such that $V(\hat{q}) \geq V(q(z))$; and (ii) there is not $\hat{u}$ and $\hat{q}$, with $\hat{u} + \hat{q} > q_R$ such that $V(\hat{q}) \geq V(q^D)$.*

Let's start addressing the first question or point (i) in the above definition. It is obvious that, if in the execution of the contract future values of q are smaller than $\underline{\underline{q}}$—that is, q is in the region for which $V(q)$ is increasing—the parties would both benefit from raising the entrepreneur's value to $\underline{\underline{q}}$. The next proposition, however, establishes that for any initial value of q, future values are never smaller than $\underline{\underline{q}}$.

Proposition 2 *For any initial value of q, the solution of the optimal contract is such that $q(z) \geq \underline{\underline{q}}$ for all z.*

Proof: *Appendix B.*

Regarding the second question, we have to answer whether the strategy $u^D + q^D = q_R$ is credible. With commitment, it is clearly optimal to set $u^D + q^D = q_R$. This strategy is optimal because it acts as a threat to discourage the entrepreneur from diverting revenues. However, if the entrepreneur is actually caught diverting, it is not obvious whether ex-post this strategy is still optimal. The next proposition establishes the renegotiation-proofness of this strategy.

Proposition 3 *The strategy $u^D + q^D = q_R$ is free from renegotiation.*

Proof: *Appendix C.*

The intuition for the result is simple. Given the promised utility q_R in case of diversion, this utility will be delivered in part by paying current consumption which provides current utility u^D and in part with continuation utility q^D. What constrains the new input of capital is the continuation utility. Because the utility function is unbounded below, we can always set the continuation utility to $q^D = \underline{\underline{q}}$ and the current utility to $u^D = q_R - \underline{\underline{q}}$, so that the continuation contract is restarted at a point in which the utility frontier is not upward sloping. A similar intuition applies to proposition 2.

It is important to point out that propositions 2 and 3 depend crucially on the assumption that the utility from consumption is unbounded below. If the utility function is bounded below, then it is possible that future values of q are smaller than $\underline{\underline{q}}$. This case will be studied in Section 6.

3.3 First order conditions

Given the concavity and differentiability properties established in proposition 1, the optimal contract can be characterized using first order conditions. Denote by γ, λ and $\mu(z)$ the Lagrange multipliers for the incentive compatibility constraint, the promised-keeping constraint and the limited liability constraint. The Lagrangian can be written as:

$$\begin{aligned} \mathcal{L} &= \pi(x) + \delta \sum_z \Big[- C(u(z)) + V(q(z))\Big] p(z) \qquad (8) \\ &+ \gamma\beta\Big[u(z_2) + q(z_2) - (1-\eta)\cdot\Big(u(z_1) + q(z_1) + \alpha\, x\Big) - \eta\cdot q_R\Big] \\ &+ \lambda\Big[\beta \sum_z \Big(u(z) + q(z)\Big)p(z) - q\Big] \\ &+ \delta \sum_z p(z)\mu(z)\Big[u(z) + q(z) - q_R\Big] \end{aligned}$$

where the choice variable k has been replaced with the variable $x = D(k)$. Accordingly, the discounted expected net revenues, $-k + \delta\sum_z F(k,z)p(z)$, has been replaced by the function $\pi(x) = -D^{-1}(x) + \delta\sum_z F(D^{-1}(x), z)p(z)$. By assumption 1, $\pi(x)$ is concave. Therefore, the derivative is decreasing and it is equal to zero when the firm employs the optimal input of capital $\bar{k}$.

The first order conditions (after some rearrangement) can be written as:

$$\begin{aligned} \pi_x &= \gamma\cdot\alpha\beta(1-\eta) \qquad (9) \\ \lambda - (\delta/\beta)[C'(u(z_1)) - \mu(z_1)] &= \gamma\cdot(1-\eta)/p(z_1) \qquad (10) \\ \lambda - (\delta/\beta)[C'(u(z_2)) - \mu(z_2))] &= -\gamma/p(z_2) \qquad (11) \\ \lambda - (\delta/\beta)[\lambda(z_1) - \mu(z_1)] &= \gamma\cdot(1-\eta)/p(z_1) \qquad (12) \\ \lambda - (\delta/\beta)[\lambda(z_2) - \mu(z_2)] &= -\gamma/p(z_2) \qquad (13) \end{aligned}$$

The Lagrange multiplier $\mu(z)$ is greater than zero if $u(z)+q(z) = q_R$ while γ is positive if the incentive-compatibility constraint is satisfied with equality. Because a positive γ implies a suboptimal input of capital (see condition (9)), the incentive-compatibility constraint is binding if $k < \bar{k}$ (which is equivalent to $x < \bar{x} = D(\bar{k})$).

To understand the importance of the incentive-compatibility constraint, we have rearranged the first order conditions so that the Lagrange multiplier γ is only on the right-hand side of equations (9)-(13). If the incentive-compatibility constraint is not binding, $\gamma = 0$ and all the right-hand-side terms of equations (9)-(13) are zero. This will be the case when $\eta = 1$ and the shock is public information. In this case equation (9) tells us that the input of capital will be at the optimal level and the subsequent equations (10)-(13) will take a standard format. On the other hand, if the incentive-compatibility constraint is binding, the Lagrange multiplier γ is positive and the input of capital will be at a suboptimal level.

Together with the incentive-compatibility, promise-keeping and limited liability constraints, conditions (9)-(13) characterize the properties of the optimal contract. The next proposition summarizes some of these properties.

Proposition 4 *There exits $\bar{q} > \beta q_R$, such that:*

(a) If $q < \bar{q}$, then $k < \bar{k}$, $u(z_1) < u(z_2)$ and $q(z_1) < q(z_2)$.

(b) If $q \geq \bar{q}$, then $k = \bar{k}$, $u(z_1) = u(z_2)$ and $q(z_1) = q(z_2)$.

PROOF: *Appendix D.*

According to the proposition, when the promised utility q is low, consumption and continuation utility depend on the realization of the shock. As the promised utility becomes sufficiently large, however, the incentive compatibility constraint is no longer binding and the entrepreneur enjoys a perfect consumption smoothing, as long as q remains above $\bar{q}$.

The intuition for this result is straightforward. If the entrepreneur is caught diverting when q is large, the punishment for diverting resources is very high because he or she will get the reservation value q_R and loose a large portion of the lifetime utility accumulated until that point. On the other hand, when $q \approx q_R$, the expected punishment cannot be large, which makes difficult to provide full insurance and make the contract incentive-compatible. It is important to emphasize that full insurance is possible—as long as q remains above the threshold $\bar{q}$—only if the probability η is positive. If the shock is never publicly revealed, the only way to punish the entrepreneur from diversion is by making consumption and next period promised utility conditional on the performance of the firm also for large values of q. In Section 4 these properties of the optimal contract will be interpreted as an optimal portfolio allocation for the entrepreneur.

3.4 Initial stage of the contract

The initial condition of the contract is given by a promised utility q_0 which depends on the initial wealth of the entrepreneur. To simplify the analysis we assume that the entrepreneur's wealth is initially zero, which approximates the fact that new firms have limited internal funds relative to their financial needs. Moreover, assuming that the entrepreneur has all the bargaining power, the value of q_0 is determined as the solution to the following problem:

$$\begin{aligned} q_0 \quad &= \quad \arg\max \{q\} \\ &\quad\quad \text{s.t. } \; V(q) \geq \kappa \end{aligned} \tag{14}$$

where κ is the fixed set up investment which is sunk. Because the function $V(q)$ is decreasing above the threshold $\underline{\underline{q}}$ (see proposition 1), the problem has a unique solution. The determination of q_0 is shown in Figure 2.

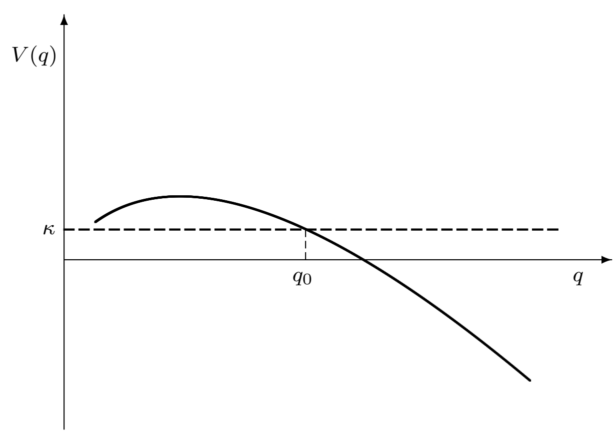


Figure 2: Determination of the initial promised utility q_0.

The initial set up cost κ plays an important role in determining the initial size of the firm: larger is κ and smaller is the initial entrepreneur's value. The smaller value of q, then, implies a smaller initial size of the firm.

4 Optimal dynamic shareholder contract

The goal of this section is to show how the optimal contract can be implemented as a dynamic shareholder contract in which the entrepreneur owns part of the firm's shares and this ownership evolves over time with the entrepreneur's wealth. Let's first define some new variables.

The optimal contract defines a sequence of state contingent payments to the entrepreneur, $c_t = C(u(z_t))$, and to the investor, $\tau_t = F(k_{t-1}, z_t) - C(u(z_t)) - k_t$. The sum of payments to the entrepreneur and the investor is equal to the dividends distributed by the firm, that is, $d_t = c_t + \tau_t = F(k_{t-1}, z_t) - k_t$. Given the sequence of dividends and payments, we can determine the firm's value, P, the entrepreneur wealth, W^E, and the investor' wealth in the firm, W^I. The firm's value (stock market value) is given by:

$$P(q_{t-1}, z_t) \;=\; E_t \sum_{j=t}^{\infty} \delta^{j-t} d_j \;=\; E_t \sum_{j=t}^{\infty} \delta^{j-t} \Big(F(k_{j-1}, z_j) - k_j \Big) \tag{15}$$

The notation makes explicit the dependence of the firm's value from the previous continuation utility q_{t-1} and the current realization of the shock z_t. As we have seen in the previous section, the entrepreneur's promised utility is the sufficient state of the contract before the realization of the shock. After the realization of the shock the sufficient states also include the shock z_t. The entrepreneur's and investor's wealth are defined as:

$$W^E(q_{t-1}, z_t) \;=\; E_t \sum_{j=t}^{\infty} \delta^{j-t} c_j \tag{16}$$

$$W^I(q_{t-1}, z_t) \;=\; E_t \sum_{j=t}^{\infty} \delta^{j-t} \tau_j \;=\; P(q_{t-1}, z_t) - W^E(q_{t-1}, z_t) \tag{17}$$

We can think of the entrepreneur's wealth as being invested in two types of assets: firm's shares and investments outside the firm. Because we are abstracting from the market risk, the investment outside the firm is equivalent to investing in a "riskless bond", although we should interpret it more broadly as an investment which is not correlated with the firm's idiosyncratic risk.

In this environment the entrepreneur starts period t with s_{t-1} shares (the shares owned by the investors are $1 - s_{t-1}$). In addition he or she owns assets b_{t-1} outside the firm with return $1/\delta - 1$. The subscript $t-1$ derives from the fact that the number of shares and the wealth invested outside the firm were decided in the previous period. Therefore, conditional on the two values of the shock, the entrepreneur's wealth can be written as:

$$W^E(q_{t-1}, z_1) \;=\; b_{t-1} + s_{t-1} \cdot P(q_{t-1}, z_1) \tag{18}$$

$$W^E(q_{t-1}, z_2) \;=\; b_{t-1} + s_{t-1} \cdot P(q_{t-1}, z_2) \tag{19}$$

Given the value of $W^E(q_{t-1}, z_t)$ defined in (16), these two conditions allow us to determine the values of b_{t-1} and s_{t-1}. In particular, subtracting equation (18) from (19) and rearranging we get:

$$s_{t-1} = \frac{W^E(q_{t-1}, z_2) - W^E(q_{t-1}, z_1)}{P(q_{t-1}, z_2) - P(q_{t-1}, z_1)} \tag{20}$$

Therefore, using the values of $W^E(q_{t-1}, z_1)$ and $W^E(q_{t-1}, z_2)$ found in the solution of the optimal contract, we are able to derive the portfolio composition of the entrepreneur, the ownership structure of the firm and how this evolves over time. We then have the following proposition.

Proposition 5 (Portfolio diversification) *Let $EW^E = \delta \sum_z W^E(q_t, z)p(z)$ be the wealth of the entrepreneur after current consumption. If $0 < \eta < 1$, there exists a wealth threshold $\overline{W}$ for which $s > 0$ only if $EW^E < \overline{W}$. If $\eta = 0$, $s > 0$ for any value of EW^E. If $\eta = 1$, $s = 0$ for any value of EW^E.*

PROOF: *Let's observe first that the entrepreneur's wealth must be an increasing function of q. Therefore, the threshold $\overline{W}$ corresponds, in utility terms, to the threshold $\bar{q}$ defined in proposition 4. According to this proposition, for $q < \bar{q}$ (which is equivalent to $EW^E < \overline{W}$), we have that $c(z_1) < c(z_2)$ and $q(z_1) < q(z_2)$. This implies that the entrepreneur's wealth depends on the shock z defined in (16). Equation (20) then implies that $s > 0$. The same proposition also establishes that for $q \geq \bar{q}$ (which is equivalent to $EW^E \geq \overline{W}$), $c(z_1) = c(z_2)$ and $q(z_1) = q(z_2)$. If next period consumption and continuation utility do not depend on z, the entrepreneur's wealth must also be independent of z. Equation (20) then implies that $s = 0$. If $\eta = 0$, however, it is obvious that the entrepreneur's wealth must be conditional on z for every level of EW^E. Finally, if $\eta = 1$, the incentive-compatibility constraint is never binding. Consequently, consumption and continuation utility do not depend on z for any q, and $s = 0$. Q.E.D.*

The above proposition establishes that the entrepreneur will fully diversify his or her portfolio only if the entrepreneur is sufficiently wealthy. Before reaching $\overline{W}$, the entrepreneur invests part of the wealth in the shares of the firm. The intuition behind this result is related to the intuition of Proposition 4 we have seen in the previous section. If the entrepreneur is caught diverting when EW^E (or q) is large, the punishment for diverting resources is very high because he or she will loose a large portion of the wealth accumulated until that point. In fact, the entrepreneur utility in case of diversion will be the minimum value q_R which corresponds to a low level of wealth. On the other hand, when $q \approx q_R$, the expected punishment cannot be large, which makes difficult for the contract to insure incentive-compatibility and provide full insurance.

It is important to emphasize that a positive value of the probability η is crucial to get this result. If $\eta = 0$, the entrepreneur will never fully diversify as stated in the proposition. This can be easily seen by using the first order conditions derived before. When the entrepreneur's portfolio is fully diversified, $c(z_1) = c(z_2)$ and $q(z_1) = q(z_2)$. Moreover, full diversification can only be reached if capital is at the optimal level $\bar{k}$. Otherwise, $\gamma > 0$ and consumption must depend on the shock. Assuming that $u(z_1) = u(z_2)$ and $q(z_1) = q(z_2)$, and using the promised-keeping constraint $q = \beta E[u(z)+q(z)]$, the incentive-compatibility constraint can be written as:

$$\frac{q}{\beta} \geq \left(\frac{1-\eta}{\eta}\right) \alpha\, D(\bar{k}) + q_R \tag{21}$$

This expression is satisfied only if q is sufficiently large which implies that full diversification can be obtained only if the entrepreneur is sufficiently wealthy. However, if $\eta = 0$, this condition will never be satisfied, and the entrepreneur will never fully diversify his or her portfolio. On the other hand, when $\eta = 1$, condition (21) is always satisfied. In this case there is no asymmetry in information and the contract would guarantee the optimal input of capital and provide a perfect consumption smoothing to the entrepreneur. More in general, the threshold $\overline{W}$ above which $s = 0$ decreases with η. It is important to observe that the threshold $\overline{W}$ will be reached only if $\delta = \beta$. When $\delta > \beta$, even if we start from $EW^E > \overline{W}$, the entrepreneur's wealth will fall below $\overline{W}$ with probability 1 in a finite number of periods.

Given the above results, we would infer that larger is η, and smaller is the entrepreneur's stake in the firm. This is clearly shown in Section 5 with a numerical example. If we interpret η as an index of the investors' protection which in turn depends on the type of existing institutions, then countries with institutions more protective of investors' rights should have a lower concentration of ownership. This finding is supported by the empirical study of La-Porta et al. (1998).

Remark: The above contractual problem can be reinterpreted as the optimal portfolio problem solved by an entrepreneur who internalizes the impact of the ownership structure on the firm's value. At the end of each period, after the payments of dividends, the entrepreneur decides how to allocate his or her wealth between the firm's shares and the market portfolio. Because the entrepreneur is risk averse, he or she would prefer to invest only in the market portfolio. However, the entrepreneur also realizes that the ownership structure has an impact on the firm's value (the price at which he or she can sell the shares). Consequently, the optimal portfolio choice optimizes this trade-off and in each period the optimal number of shares that the entrepreneur decides to own is given by (20).

5 Numerical example

This section studies a parameterized version of the model. The model is solved numerically and the computational procedure is described in Appendix G. The goal of this example is not to provide a quantitative evaluation of the model but to illustrates some of its qualitative properties.

The market discount factor is set to $\delta = 0.94$ and the subjective discount factor to $\beta = 0.939$. The utility function is specified as $u(c) = c + \log(c)$ and the diversion parameter α is set to 1. The revenue function is specified as $F(k, z) = (1 - \omega)k + A + zk^{\theta}$. The parameter θ determines the return to scale and we set it to 0.99 so that the production function is close to constant return to scale. This model is very similar to a model in which $\theta = 1$ and the scale of the firm is subject to an upper bound. The parameter ω is the cost of production per unit of capital and it is equal to 0.25. Assuming that the input of labor is proportional to the input of capital, ωk is interpreted as the cost of labor plus capital depreciation. The shock takes the values $z_1 = 0.5\,\bar{z}$ and $z_2 = 1.5\,\bar{z}$ with equal probability. The value of $\bar{z}$ is such that the unconstrained input of capital is 100. The shock gets publicly revealed with probability $\eta = 0.1$. The initial set up cost κ is assumed to be zero and the initial size of firms is controlled by choosing A. We choose A such that the initial size of new firms is about half the optimal size. The reservation value is set to zero. The full set of parameter values are reported in Table 1.

Ownership dynamics: Panel (a) of Figure 3 plots the input of capital as a function of the entrepreneur's total wealth. This plot shows that the size

Table 1: Parameter values.

Market discount rate	$\delta = 0.94$
Intertemporal discount factor	$\beta = 0.939$
Revenue function $F(k, z) = (1 - \omega)k + A + zk^{\theta}$	$\omega = 0.25$, $A = 1$, $\theta = 0.99$
Revenue shock	$z \in \{0.166, 0.498\}$
Probability that the shock is publicly revealed	$\eta = 0.1$
Diversion value $\alpha D(k)$	$\alpha = 1$
Reservation value q_R	$q_R = 0$

of the firm is increasing in the entrepreneur's wealth. As the entrepreneur becomes wealthier, the incentive-compatibility constraints are relaxed and the firm expands in size.

The second panel of Figure 3 plots the next period wealth as a function of current wealth, for each realization of the shock. The 45 degree line is also plotted. Before the entrepreneur's wealth reaches a certain level, the next period wealth increases when the shock is high and decreases when the shock is low. Because the size of the firm is correlated with the entrepreneur's wealth, this property implies that the firm expands when the current cash flow is high and contracts when the current cash flow is low.

As discussed in Section 4, the entrepreneur's wealth is in part invested in the firm's shares and in part invested outside the firm. Panel (c) of Figure 3 plots the fraction of the entrepreneur's wealth invested in the firm and panel (d) the shares of the firm directly owned by the entrepreneur. These two graphs show that the entrepreneur diversifies his or her portfolio as the wealth increases above a certain threshold. In particular, the entrepreneur invests a smaller fraction of wealth in the firm and owns a smaller number of firm's shares once his or her wealth is sufficiently large. When the entrepreneur's wealth is very small, however, the diversification of the entrepreneur's portfolio decreases as the wealth increases. The reason is because for low levels of wealth the input of capital is also low. At this production scale the marginal productivity of capital is very sensitive to changes in k. Consequently, high fluctuations of capital are inefficient. Because the input of capital depends on the entrepreneur's wealth, in order to reduce the volatility of capital we have to reduce the volatility of the entrepreneur's wealth. This is obtained by reducing the insider ownership.

The next step is to relate the ownership structure to the size and the age of the firm. This is shown in Figure 4. As can be seen from the first panel, the shares owned by the entrepreneur decline once the firm has reached a certain size. This derives directly from the properties of the model shown in Figure 3: the fact that the input of capital is increasing in the entrepreneur's wealth (panel (a)) and the fact that the entrepreneur's shares are decreasing in wealth above a certain threshold (panel (d)).

The bottom panel plots the entrepreneur's shares as a function of the firm's age. To generate this graph we simulate an initial mass of entrant firms over time. More specifically, we assume that at time zero there is a large mass (infinite number) of entrants with the same initial size. This initial size is about half the optimal size. Although all firms have the same initial size, over time they become heterogeneous. The data reported in the bottom panel of Figure 4 is the average over all firms of the same age.

As can be seen from the graph, the entrepreneur's shares (stake in the firm) decreases with the age of the firm. This follows from the assumption that new entrant firms are initially small and because small they tend to grow on average.

The pattern of the shares owned by the entrepreneur is sensitive to the probability η with which the shock is publicly revealed. Figure 5 shows the shares owned by the entrepreneur as a function of wealth and age for different values of η. Panel (a) shows that the entrepreneur owns less share (for each level of wealth) when η is larger. Panel (b) also shows that the insider ownership is lower for each age class of firms when η is larger.

To understand these properties let's notice first that the cost of being caught diverting is higher when the entrepreneur's wealth is higher. This cost is given by the loss of wealth accumulated up to that point. Consequently, when the entrepreneur is wealthier, the contract can be made incentive compatible with the entrepreneur facing a lower risk—that is, better consumption smoothing. This is accomplished by reducing the ownership of the firm's shares. Now if we increase η, we also increase the expected cost of being caught diverting for each level of wealth. Therefore, the increase in η further reduces the incentive of the entrepreneur to divert resources and the contract can be made constrained efficient with a lower insider ownership.

The parameter η can be interpreted as an index of the investors' protection determined by the institutional environment. For example, we can think of higher values of η as characterizing the institutional environment of countries in which firms have more demanding obligations in the disclosure of information. If we take this interpretation of the parameter η, we should observe lower ownership concentration in countries with better institutional protection of investors. This finds empirical support in the work of La-Porta et al. (1998).

Other dynamics properties: The properties of the optimal contract emphasized above imply other properties of the investment behavior and dynamics of firms which are summarized below.

- The expected productivity of capital (return on assets) and the Tobin's q are decreasing in the firm's size and age.
- The investment rate (ratio between expected next period capital and current capital) is decreasing in size and age.
- The growth rate of the firm as well as the variability of growth are both decreasing in size and age.

These properties have simply intuitions. The marginal productivity of capital and the Tobin's q are decreasing in size because the revenue function displays decreasing returns to scale. As a result of the high marginal productivity of capital, small firms have high investment rates and growth faster. At the same time, however, they also experience greater variability of growth. The age dependence derives from the fact that younger firms are smaller.

6 Extensions

This section considers several extensions of the model. Subsection 6.1 extends the analysis to the case in which the shock takes more than two values. Subsection 6.2 studies the case in which the utility function is bounded below. Subsection 6.3 discusses the case in which the diverted revenue cannot be recovered if the entrepreneur is caught diverting. Subsection 6.4 considers the alternative assumption in which the diverted revenue is additive to consumption.

6.1 More than two shocks

Assume that the shock takes N values, that is, $z \in \{z_1, ..., z_N\}$. Denote by $D(k, z_i, z_j) = F(k, z_j) - F(k, z_i)$ the revenue diverted if the shock is z_j

and the entrepreneur announces z_i, with $j \geq i$. The incentive-compatibility constraints are:

$$u(z_j) + q(z_j) \;\geq\; (1-\eta) \cdot \Big[u(z_i) + q(z_i) + \alpha D(k, z_i, z_j)\Big] + \eta \cdot q_R \qquad (22)$$

with $j = 2, ..., N$ and $i = 1, ..., j-1$.

The contractual problem is still (3) but the incentive-compatibility constraint (5) is replaced by this new set of constraints. To simplify the analysis, we now assume that the revenue function takes a particular form:

Assumption 2 *The revenue function is $F(k,z) = (1-\omega)k + z \cdot f(k)$, where $\omega \geq 0$ is constant and f is strictly increasing, concave and differentiable.*

This assumption is convenient to establish the uniqueness of the solution to the optimal contract as stated in the next proposition.

Proposition 6 (Optimal long-term contract) *Under assumption 2, the contractual problem is a contraction and there is a unique fixed point $V(q)$. The solution is unique and $V(q)$ is strictly concave and differentiable. Moreover, there exists $\underline{\underline{q}}$ for which $V(q)$ is strictly decreasing for $q > \underline{\underline{q}}$ and the long-term contract is free from renegotiation.*

Proof: *Appendix E.*

This proposition is the equivalent of propositions 1, 2 and 3 for the case in which the shock takes only two values. Furthermore, a similar version of proposition 4 holds. Therefore, the optimal contract maintains the properties we have seen before. However, now the simple share ownership of the entrepreneur is no longer sufficient to implement the optimal contract. A more complex managerial compensation is required. The remaining part of this section will show how the use of stock options can implement the optimal contract.

Consider the simplest case in which the shock takes only three values. We then have the following proposition.

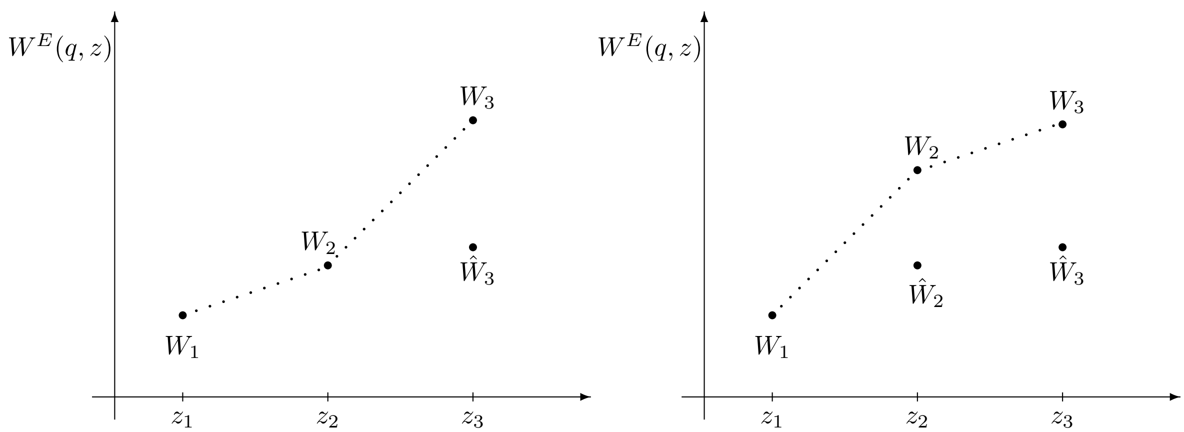
Proposition 7 (Stock options) *Assume that $z \in \{z_1, z_2, z_3\}$ and denote by $P(q,z)$ the market price of the firm's shares in the optimal contract when the pre-shock promised utility is q and the realization of the shock is z. Then the optimal contract can be implemented with the entrepreneur ownership of $s(q)$ shares and the option to buy $s^o(q)$ shares at some price $P^o(q) \geq P(q, z_1)$.*

The proof of the proposition is simple and it is illustrated in Figure 7. In the optimal contract, the entrepreneur's wealth after the realization of z is $W^E(q,z)$. Because the shock takes only three values, the entrepreneur's wealth also takes three values (for given q). These values are plotted in Figure 7 and they are denoted by W_1, W_2 and W_3. Two cases are considered: in the first case $W^E(q,z)$ is a convex function of z while in the second case it is concave.

Let's consider first the case plotted in panel (a). With the simple ownership of shares, we can replicate only two of the three values that $W^E(q,z)$ should take. Suppose that the entrepreneur's allocation of wealth is such that $b+s(q)\cdot P(q,z_1)=W^E(q,z_1)=W_1$ and $b+s(q)\cdot P(q,z_2)=W^E(q,z_2)=W_2$. The values of b and $s(q)$ are determined as in Section 4. Although this composition of the entrepreneur's portfolio replicates W_1 and W_2 as required by the optimal contract, it does not replicate $W^E(q,z_3)=W_3$. Let $\hat{W}_3=b+s(q)\cdot P(q,z_3)$ be the entrepreneur's wealth generated by this portfolio composition when the shock is z_3 and let's assume that $\hat{W}_3<W_3$ as plotted in the graph. To make sure that the entrepreneur's wealth is W_3 when the shock is z_3 the entrepreneur could be offered the option to buy $s^o(q)$ at the price $P^o(q)$, where the price satisfies $P(q,z_2)<P^o(q)<P(q,z_3)$. The number of option shares offered to the entrepreneur is determined by the condition $s^o(q)[P(q,z_3)-P^o(q)]=W_3-\hat{W}_3$. Notice that the market price $P(q,z)$ is the price after the entrepreneur has exercised the option. Because the share price is increasing in z, the option will be exercised only if $z=z_3$ and the entrepreneur's wealth will evolve according to the recommendation of the optimal contract.

Let's consider now the second case plotted in panel (b) of Figure 7. In this case we have to change the structure of the option. Here b and $s(q)$ are chosen to replicate only W_1, that is, $W_1=b+s(q)\cdot P(q,z_1)$. The values of $\hat{W}_2=b+s(q)\cdot P(q,z_2)$ and $\hat{W}_3=b+s(q)\cdot P(q,z_3)$ are smaller than W_2 and W_3, that is, $\hat{W}_2<W_2$ and $\hat{W}_3<W_3$. By adding an option to purchase $s^o(q)$ shares at the price $P^o(q)$ which satisfies $P(q,z_1)<P^o(q)<P(q,z_2)$, we can replicate the points W_2 and W_3. Here we need the right combination of share ownership $s(q)$, option shares $s^o(q)$ and option price $P^o(q)$. Notice that, after exercising the option, the entrepreneur is not constrained to keep the previous holding of shares. Therefore, after trading, the new shares holding is not necessarily $s(q)+s^o(q)$. Also notice that there are many share ownership and option schemes that implement the optimal contract.

When the shock takes more than three values, the implementation of



(a) $W^E(z)$ is convex

(a) $W^E(z)$ is concave

Figure 7: Contract implementation with stock options.

the optimal contract requires more complex option schemes. For example, it could be implemented by giving the option to purchase different sets of shares at different prices. However, independently of how the optimal contract is implemented, the general properties do not change and the entrepreneur will eventually reduce the wealth dependence from the performance of the firm as his or her wealth increases. Broadly interpreted this means that the insider ownership tends to decline and the entrepreneur's portfolio becomes more diversified on average over time.

6.2 Utility is bounded below

A key assumption for the renegotiation-proofness of the long-term contract established in Section 3.2 was that the utility function is unbounded below. We now relax this assumption and we assume that $U(0) = \underline{u} > -\infty$. The most important consequence of making this alternative assumption is that the long-term contract—that is, the contract that the parties commit not to renegotiate in future periods—may not be free from renegotiation.

To see why the contract may not be free from renegotiation, consider the event in which the entrepreneur is caught diverting. According to the optimal long-term contract, the entrepreneur value will be $u^D + q^D = q_R$. If the utility function were unbounded below, q^D could be set to $\underline{\underline{q}}$—which is the value of q above which the investor's value is decreasing—and there is no gain for the investor from renegotiating the contract. However, if the utility function is bounded below, this may not be feasible. In fact, in order to set $q^D = \underline{\underline{q}}$, the current utility should be set to $u^D = q_R - \underline{\underline{q}}$. But $q_R - \underline{\underline{q}}$ may be smaller than $\underline{u}$, and therefore, it would not be feasible to set $q^D = \underline{\underline{q}}$. In this case the investor would gain from renegotiating the contract. In particular, the value of the contract for the entrepreneur will be increased to $\underline{u} + \underline{\underline{q}} > q_R$ and the threat of $u^D + q^D = q_R$ in case of diversion is not credible.

A similar argument shows that the contract could be renegotiated also in contingencies in which the entrepreneur is not caught diverting. More specifically, let's assume that along the execution of the contract, the promised value conditional on z_1 is binding, that is, $\tilde{q}(z_1) = u(z_1) + q(z_1) = q_R$. After the announcement of the shock in this contingency, the optimal policy is to set $q(z_1) \geq \underline{\underline{q}}$ and $u(z_1) = q_R - \underline{\underline{q}}$. If however $q_R - \underline{\underline{q}} < \underline{u}$, this is not feasible and the contract will be renegotiated.

The derivation of the renegotiation-proof contract is based on the following idea. In those contingencies in which the contract is renegotiated, renegotiation implies an increase in the continuation value. Anticipating this, there is not reason to set in advance a low promised value when this value will be increased ex-post. Therefore, we can impose an additional constraint for which $u(z) + q(z)$ must always be greater than some lower bound $q_{min} \geq q_R$. This constraint is in addition to the lower bound for the current utility, that is, $u(z) \geq \underline{u}$. The contractual problem can been written as:

$$V(q) \;=\; \max_{k,u(z),q(z)} \Big\{ -k + \delta \sum_z \Big[F(k,z) - C(u(z)) + V(q(z))\Big]p(z) \Big\} \quad (23)$$

subject to

$$q \;=\; \beta \sum_z \Big[u(z) + q(z)\Big]p(z) \quad (24)$$

$$u(z_2) + q(z_2) \geq (1-\eta)\cdot\Big[u(z_1) + q(z_1) + \alpha D(k)\Big] + \eta \cdot \Big[u^D + q^D\Big] \quad (25)$$

$$u(z) + q(z) \geq q_{min}; \quad u^D + q^D \geq q_{min} \tag{26}$$

$$u(z) \geq \underline{u}; \quad u^D \geq \underline{u} \tag{27}$$

The following proposition defines the renegotiation-proof contract.

Proposition 8 *The optimal and renegotiation-proof contract is characterized by some lower bound q_{min}.*

PROOF: *Appendix F.*

A similar result is derived by Wang (2000) in a finite horizon model and by Quadrini (1999) in a model with risk neutral agents.

The imposition of the lower bound q_{min} does not change the main properties of the optimal contract. Therefore, all the previous results apply. In particular, if the shock takes only two values, the optimal contract can be implemented with the entrepreneur's ownership of some of the firm's shares. With more than two shocks, the basic shares ownership could be complemented with stock options. Moreover, if the probability η is positive, the entrepreneur's stake in the firm will converge to zero (full diversification) once the entrepreneur's wealth has reached a certain level.

6.3 Diverted revenues cannot be recovered

In the analysis conducted so far, we have assumed that in the case in which the entrepreneur is caught diverting, the diverted revenues is recovered. Given this assumption, the incentive-compatibility constraint was:

$$u(z_2) + q(z_2) \;\geq\; (1-\eta)\cdot\Big[u(z_1) + q(z_1) + \alpha D(k)\Big] + \eta \cdot q_R \tag{28}$$

From this equation we can see that if the shock is public information, that is, $\eta = 1$, this constraint simply imposes that $u(z_2) + q(z_2) \geq q_R$. Assuming that the surplus that a firm can generate is sufficiently large, this constraint is never binding and the firm will be always operated at the optimal scale $\bar{k}$ (frictionless economy).

Let's consider now the case in which the revenues can not be recovered. The incentive-compatibility constraint becomes:

$$u(z_2) + q(z_2) \ \geq \ (1-\eta) \cdot \Big[u(z_1) + q(z_1)\Big] + \alpha D(k) + \eta \cdot q_R \qquad (29)$$

If $\eta < 1$, it is easy to show that the properties of the model do not change qualitatively. However, if $\eta = 1$ and there is no asymmetry in information, the firm is not necessarily operated at the optimal scale $\bar{k}$ as in the case in which the diverted revenue can be recovered. In fact, if we set $\eta = 1$, the incentive-compatibility constraint becomes:

$$u(z_2) + q(z_2) \ \geq \ \alpha D(k) + \eta \cdot q_R \qquad (30)$$

Therefore, even if the shock is public information, the capital invested in the firm constraints the feasibility of the values promised to the entrepreneur. In this respect the model resembles a limited enforceability model in which diversion is perfectly observable but can not be prevented. Therefore, when $\eta < 1$ the model combines some features of the limited enforceability model with some features of the pure moral hazard model. The general properties, however, remain unaffected.

6.4 Diverted revenues additive to consumption

An alternative formulation of the benefit from diversion is to assume that the diverted revenues are additive to consumption. Denote by $c(z_1)$ the consumption recommended by the contract when the shock is $z = z_1$. If the shock is $z = z_2$ and the entrepreneur diverts revenues, the current utility is $U(c(z_1) + D(k))$. With this alternative assumption, the incentive-compatibility constraint becomes:

$$U(c(z_2)) + q(z_2) \ \geq \ (1-\eta) \cdot \Big[U\Big(c(z_1) + D(k)\Big) + q(z_1)\Big] + \eta \cdot q_R \qquad (31)$$

The contractual problem is still given by (3) but with this new incentive-compatibility constraint. Notice that now we do not need to impose that the marginal utility of consumption is bounded away from zero.

With this alternative formulation of the benefits from diversion, it is more difficult to prove the concavity of the function $V(q)$. This is because the feasible set is not convex. However, if this function is concave, then all the results proved in the previous sections hold and the optimal contract displays the same qualitative properties.

In the remaining part of this section we will show a numerical example in which this property is satisfied. In addition to assuming that the diverted revenues are additive to consumption, we also assume that the utility function is $U(c) = \log(c)$. With the exception of the utility function, we use the same parametrization of Section 5.

Figure 8 plots the entrepreneur's shares as a function of wealth and age, for different values of the probability η. As can be seen from this figure, the properties of the model are very similar to the properties of the model analyzed in the previous sections.

7 Conclusion

This paper studies the optimal contract between risk-averse entrepreneurs and risk-neutral investors in a model with repeated moral hazard. When the shock to the firm's revenues takes only two values, the optimal contract can be implemented with the entrepreneur's ownership of a fraction of the firm's shares. This ownership, then, evolves over time according to the performance of the firm. Initially, the entrepreneur's wealth is small and the optimal contract requires that the entrepreneur faces part of the firm's risk by holding a large stake in the firm. When the entrepreneur's wealth is sufficiently large, however, the insider ownership declines. This follows from the fact that it is easier to make the contract incentive compatible as the entrepreneur's wealth increases. Because the entrepreneur's wealth increases on average over time, it is shown that the insider ownership tends to decline with the age of the firm. The paper also shows that higher is the degree of investors' protection—as measured by the ability of investors to extract information about the performance of the firm—and lower is the concentration of ownership. These results are consistent with the ownership dynamics observed in cross-section and cross-country empirical studies.

When the shock takes more than two values, the optimal contract can not be implemented with the simple ownership of the firm's shares. In this case, the implementation of the optimal contract can be obtained with the addition of stock options. This is consistent with the growing practise of using stock options for managerial compensation.

A Proof of proposition 1

Let's rewrite first the contractual problem by using $x = D(k)$ as a choice variable.

$$T(V)(q) = \max_{x,u(z),q(z)} \left\{ \pi(x) + \delta \sum_z \Big[-C(u(z)) + V(q(z)) \Big] p(z) \right\} \tag{32}$$

subject to

$$q = \beta \sum_z \Big[\Big(u(z) + q(z) \Big) \Big] p(z) \tag{33}$$

$$u(z_2) + q(z_2) \geq (1-\eta) \cdot \Big[u(z_1) + q(z_1) + \alpha x \Big] + \eta \cdot q_R \tag{34}$$

$$u(z) + q(z) \geq q_R \tag{35}$$

All the constraints are linear functions of the states and choice variables. This insures that the feasible set is convex. The verification of the Blackwell conditions then shows that the mapping T is a contraction and there is a unique function V.

The concavity of V derives from the concavity of $\pi(x)$ and from the fact that the mapping T preserves concavity. The concavity of $\pi(x)$ and $V(q)$ then implies that the solution is unique. The differentiability can be proved by verifying the conditions of Theorem 9.10 in Stokey, Lucas, & Prescott (1989).

Let's prove now that there is some $\underline{\underline{q}} \geq \beta q_R$ for which V is decreasing in $q \geq \underline{\underline{q}}$. It is sufficient to prove that there are two points q_a and q_b, with $q_a < q_b$, for which $V(q_a) > V(q_b)$. Associated with q_a and q_b there are solutions $(x_a, u_a(z), q_a(z))$ and $(x_b, u_b(z), q_b(z))$ respectively. We want to show that $V(q_a)$ cannot be smaller than $V(q_b)$. In fact with $q = q_a$ there is a feasible solution, different from $(x_a, u_a(z), q_a(z))$, for which $V(q_a) > V(q_b)$. Consider the solution $(\tilde{x}_a, \tilde{u}_a(z), \tilde{q}_a(z))$ constructed as follows: $\tilde{x}_a = x_b$, $\tilde{q}_a(z) = q_b(z)$, $\tilde{u}_a(z_2) = u_b(z_2)$ and $\tilde{u}_a(z_1)$ is determined by the equation:

$$q_a = \beta \sum_z \Big[\Big(\tilde{u}_a(z) + \tilde{q}_a(z) \Big) \Big] p(z) \tag{36}$$

This is the promised-keeping constraint. It is easy to verify that the solution $(\tilde{x}_a, \tilde{u}_a(z), \tilde{q}_a(z))$ also satisfies the incentive compatibility constraint:

$$\tilde{u}_a(z_2) + \tilde{q}_a(z_2) \geq (1-\eta) \cdot \Big[\tilde{u}_a(z_1) + \tilde{q}_a(z_1) + \alpha \tilde{x}_a \Big] + \eta \cdot q_R \tag{37}$$

Moreover, if q_a is sufficiently large and q_b is sufficiently close to q_a, the limited liability constraint $\tilde{u}_a(z_1) + \tilde{q}_a(z_1) > q_R$ is also satisfied. Therefore, this solution is feasible. Because $EC(\tilde{u}_a(z)) < EC(u_b(z))$, the value of V associated with this solution must be greater than $V(q_b)$. Finally notice that this result assumes that q_a and q_b are sufficiently large. Therefore, we can not conclude that the function $V(q)$ is decreasing for all feasible values of q. *Q.E.D.*

B Proof of proposition 2

Let $(x, u(z), q(z))$ be a solution of problem (32) for some q and assume that $q(z_1) < \underline{\underline{q}}$. We want to show that this cannot be the optimal solution.

Consider the alternative solution $(\tilde{x}, \tilde{u}(z), \tilde{q}(z))$, where $\tilde{x} = x$, $\tilde{u}(z_2) = u(z_2)$ and $\tilde{q}(z_2) = q(z_2)$, $\tilde{q}(z_1) = \underline{\underline{q}}$ and $\tilde{u}(z_1)$ satisfies $u(z_1) + q(z_1) = \tilde{u}(z_1) + \tilde{q}(z_1)$. In this alternative solution we only change the current and continuation utilities when the shock is low. This alternative solution is obviously feasible. At the same time it satisfies $C(\tilde{u}(z_1) < C(u(z_1))$ and $V(\tilde{q}(z_1)) > V(q(z_1))$. Therefore, the solution $q(z_1) < \underline{\underline{q}}$ cannot be optimal. Notice that finding the value of $\tilde{u}(z_1)$ that satisfies $u(z_1)+q(z_1) = \tilde{u}(z_1)+\underline{\underline{q}}$ is always possible because the function $U(c)$ is unbounded below. *Q.E.D.*

C Proof of proposition 3

Suppose that the entrepreneur is caught diverting. In this case the promise utility is $u^D + q^D = q_R$. Because $U(c)$ is not bounded below, we can set $q^D = \underline{\underline{q}}$ and $u^D = q_R - q^D$. Therefore, we are able to deliver the promised value q_R by choosing a continuation utility at least as big as $\underline{\underline{q}}$ and the investor would not gain by increasing the promised value $u^D + q^D$ above q_R.

D Proof of proposition 4

Let's observe first that the incentive-compatibility constraint $u(z_2) + q(z_2) \geq (1 - \eta) \cdot [u(z_1) + q(z_1) + \alpha\, x] + \eta \cdot q_R$ implies that $u(z_2) + q(z_2) > u(z_1) + q(z_1)$. Because $u(z_1) + q(z_1) \geq q_R$, it is obvious that $u(z_2) + q(z_2)$ is always greater than q_R as long as $x > 0$. Let's also observe that λ is increasing in q. In fact, from the envelope condition we have $\lambda = -V_q$. Because V is concave, $-V_q$ is increasing in q. Therefore, λ must also be increasing in q. Finally, because q constraints the set of feasible choices, the Lagrange multiplier $\mu(z_1)$ can not be increasing in q (and λ). At some point, when q is above some threshold $\bar{q}$, the incentive-compatibility constraint is no longer binding. At this point $\gamma = 0$ and it can be verified from (9) that $k = \bar{k}$, and from (10)-(13) that $u(z_1) = u(z_2)$ and $q(z_1) = q(z_2)$.

When $q < \bar{q}$, it must be that $\lambda(z_1) < \lambda(z_2)$, which in turn implies that $u(z_1) < u(z_2)$ and $q(z_1) < q(z_2)$. This can be shown by contradiction. Suppose that $\lambda(z_1) > \lambda(z_2)$. Then $q(z_1) > q(z_2)$. Furthermore, conditions (10)-(13) imply $C'(u(z_1)) = \lambda(z_1)$ and $C'(u(z_2)) = \lambda(z_2)$ and $u(z_1) > u(z_2)$. Therefore, $u(z_1) + q(z_1) > u(z_2) + q(z_2)$. But then we could improve the contract by decreasing $u(z_1)$ and increasing $u(z_2)$. In fact, the convexity of $C(u)$ makes the expected cost of the current promised utility smaller and, as can be seen from the incentive-compatibility constraint, allows us to increase x. Therefore, the solution must be $\lambda(z_1) < \lambda(z_2)$, which in turn implies $u(z_1) < u(z_2)$ and $q(z_1) < q(z_2)$. At the same time, $\lambda(z_1) < \lambda(z_2)$ is possible only if $\gamma > 0$. From condition (9) we then see that $k < \bar{k}$. *Q.E.D.*

E Proof of proposition 6

Define $x = f(k)$. Therefore, $D(k, z_j, z_i) = \alpha(z_j - z_i)x$ and $\pi(x) = -f^{-1}(x) + \delta \sum_z F(f^{-1}(x), z)p(z)$. The contractual problem can be written as:

$$T(V)(q) = \max_{x,u(z),q(z)} \left\{ \pi(x) + \delta \sum_z \Big[- C(u(z)) + V(q(z)) \Big] p(z) \right\} \tag{38}$$

subject to

$$q = \beta \sum_z \Big[\Big(u(z) + q(z) \Big) \Big] p(z) \tag{39}$$

$$u(z_j) + q(z_j) \geq (1 - \eta) \cdot \Big[u(z_i) + q(z_i) + \alpha(z_j - z_i)x \Big] + \eta \cdot q_R \tag{40}$$

$$u(z) + q(z) \geq q_R \tag{41}$$

with $j = 2, ..., N$ and $i = 1, ..., j - 1$. All the constraints are linear functions of the states and choice variables. This insures that the feasible set is convex. The verification of the Blackwell conditions then shows that the mapping T is a contraction and there is a unique function V.

The concavity of V derives from the concavity of $\pi(x)$ and from the fact that the mapping T preserves concavity. The concavity of $\pi(x)$ and $V(q)$ then implies that the solution is unique. The differentiability can be proved by verifying the conditions of Theorem 9.10 in Stokey et al. (1989). The rest of the proof is as in the proof of propositions 1, 2 and 3. *Q.E.D.*

F Proof of proposition 8

For any q_{min}, the problem is well defined and it has the same properties established in proposition A. For a value of q_{min} sufficiently large, $V'(q) < 0$. To see this, consider the incentive-compatibility constraints $u_2 + q_2 \geq (1-\eta)(u_1 + q_1 + \alpha\bar{x}) + \eta q_{min}$. Substituting this constraint on the promised-keeping constraints $q = \beta \sum_z p(z)(u(z) + q(z))$ we get:

$$q \geq \beta p(z_1)(u(z_1) + q(z_1)) + \beta p(z_2)(1-\eta)(u(z_1) + q(z_1) + \alpha\bar{x}) + \beta p(z_2)\eta q_{min} \quad (42)$$

Let's set $u(z_1) + q(z_1) = q_{min}$. The above equation becomes:

$$q \geq \beta q_{min} + \beta p(z_2)\alpha\bar{x} \quad (43)$$

Of course $q \geq q_{min} - \underline{u}$. Therefore, substituting q we get:

$$q_{min}(1-\beta) \geq \underline{u} + \beta p(z_2)\alpha\bar{x} \quad (44)$$

If q_{min} is sufficiently large, this constraint is always satisfied and the production scale will always be at the optimal level $\bar{x}$. Therefore, $V'(q)$ can not be positive.

Consider now the extreme case in which $q_{min} = q_R$. We distinguish two cases. In the first case $\underline{u}+\underline{\underline{q}} < q_R$ and the long-term contract is free from renegotiation. In the second case $\underline{u}+\underline{\underline{q}} > q_R$ which allows for $q(z_1)$ to be smaller than $\underline{\underline{q}}$. Because we know that for q_{min} sufficiently large $V'(q_{min}) < 0$, there must be some $q_{min} > q_R$ such that $V'(q_{min}) = 0$. We may have more than two values of q_{min} that satisfy this. We then take the lowest value. *Q.E.D.*

G Numerical solution

Let's notice first that there is a unique correspondence between q and the Lagrange multiplier λ. If we knew the function relating these two variables, the first order conditions would be sufficient to solve the model. Denote this function by $g(q) = \lambda$. The computation procedure consists of the iteration over this function.

We discretize the state space of q and we guess the values of the function $g(q)$ at each grid point. The guessed points are then joined with step-wise linear functions. Equations (9)-(13) together with the promised-keeping, incentive compatibility and limited liability constraints provides nine conditions. With the addition of $\lambda(z_1) = g(q(z_1))$ and $\lambda(z_2) = g(q(z_2))$, we have eleven conditions. These allow us to solve—at each grid point of q—for the following eleven unknowns: x, $u(z_1)$, $u(z_2)$, $q(z_1)$, $q(z_2)$, γ, λ, $\lambda(z_1)$, $\lambda(z_2)$, $\mu(z_1)$, $\mu(z_2)$. The value of λ can then be used to update the guessed value of the function $g(q)$ at the particular grid point. The procedure is repeated until convergence in the function $g(q)$.

References

Aseff, J. G. & Santos, M. (2001). Stock optionds and managerial optimal contracts. Unpublished Manuscript, Arizona State University.

Atkeson, A. (1991). International lending with moral hazard and risk of reputation. *Econometrica*, *59*(4), 1069–89.

Atkeson, A. & Lucas, R. E. (1992). On efficient distribution with private information. *Review of Economic Studies*, *59*(4), 427–53.

Atkeson, A. & Lucas, R. E. (1995). Efficiency and equality in a simple model of efficient unemployment insurance. *Journal of Economic Theory*, *66*(1), 64–88.

Castro, R., Clementi, G., & MacDonald, G. (2001). Optimal financing contracts, investor protection, and growth. Unpublished manuscript.

Clementi, G. & Cooley, T. F. (2000). Sensitivity of CEO pay to performance wealth in a dynamic agency model. Unpublished manuscript, Carnegie Mellon University.

Clementi, G. & Hopenhayn, H. A. (1998). Optimal lending contracts and firms' survival with moral hazard. Unpublished manuscript, Carnegie Mellon University and University of Rochester.

Cole, H. L. & Kocherlakota, N. (1997). Efficient allocations with hidden income and hidden storage. Research Department Staff Report 238, Federal Reserve Bank of Minneapolis.

DeMarzo, P. M. & Fisherman, M. J. (2000). Optimal long-term financial contracting with privately observed cash flows. Unpublished manuscript, Stanford University and Northewestern University.

Fudenberg, D., Holmstrom, B., & Milgrom, P. (1990). Short-term contracts and long-term agency relationships. *Journal of Economic Theory*, *51*(1), 1–31.

Green, E. J. (1987). Lending and the smoothing of uninsurable income. In Prescott, E. C. & Wallace, N. (Eds.), *Contractual Arrangements for Intertemporal Trade*, pp. 3–25. University of Minneapolis Press, Minneapolis.

La-Porta, R., de Silanes, F. L., Shleifer, A., & Vishny, R. W. (1998). Law and finance. *Journal of Political Economy*, *106*(6), 1113–55.

Marcet, A. & Marimon, R. (1992). Communication, commitment and growth. *Journal of Economic Theory*, *58*(1), 219–249.

Mikkelson, W., Partch, M., & Shah, K. (1997). Ownership and operating performance of companies that go public. *Journal of Financial Economics*, *44*, 281–307.

Phelan, C. (1995). Repeated moral hazard and one-sided commitment. *Journal of Economic Theory*, *66*(2), 488–506.

Phelan, C. & Townsend, R. (1991). Computing multiperiod information-constrained optima. *Review of Economic Studies*, *58*(4), 853–81.

Quadrini, V. (1999). Investment and default in renegotiation-proof contracts with moral hazard. Unpublished manuscript, New York University.

Spear, S. & Srivastava, S. (1987). On repeated moral hazard with discounting. *Review of Economic Studies*, *54*(4), 599–818.

Stokey, N. L., Lucas, R. E., & Prescott, E. C. (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, Massachusetts.

Thomas, J. & Worrall, T. (1990). Income fluctuations and asymmetric information: an example of the repeated pricipal-agent problem. *Journal of Economic Theory*, *51*, 367–390.

Wang, C. (1995). Dynamic insurance with private information and balanced budgets. *Review of Economic Studies*, *62*(4), 577–595.

Wang, C. (1997). CEO compensation and shareholder wealth in a dynamic agency model. *Journal of Economic Theory*, *76*(1), 72–105.

Wang, C. (2000). Renegotiation-proof dynamic contracts with private information. *Review of Economic Dynamics*, *3*(3), 396–422.

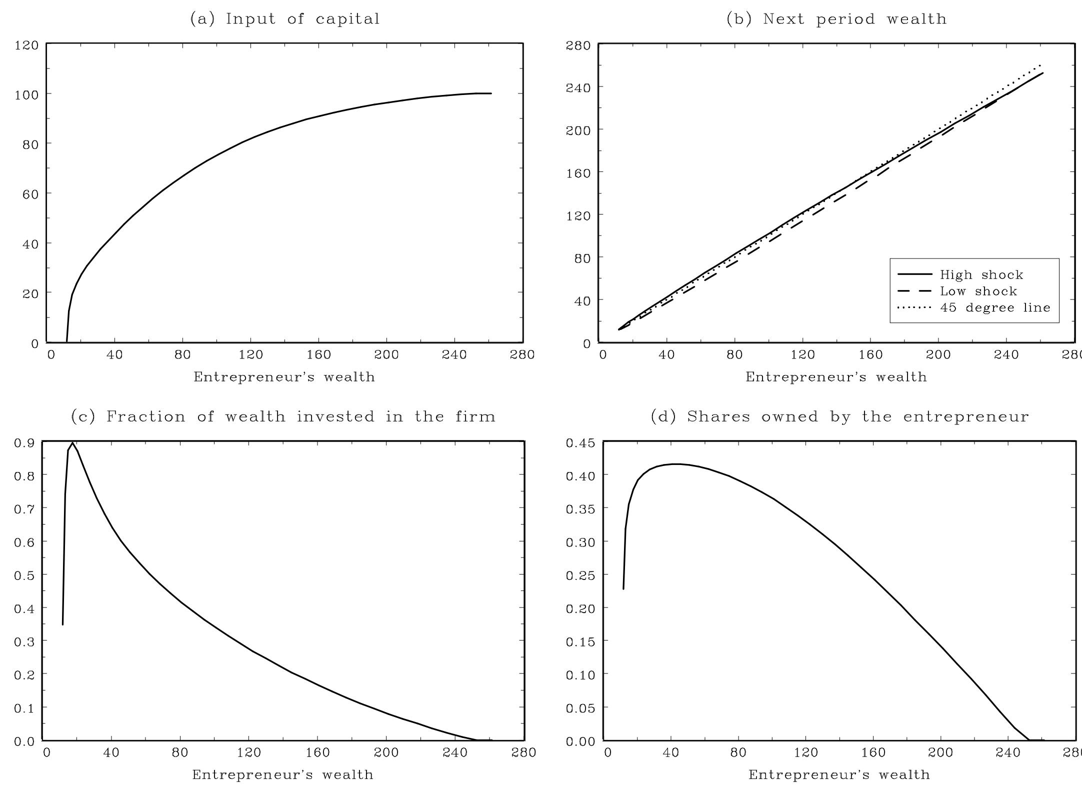

(a) Input of capital
Entrepreneur's wealth
(b) Next period wealth
High shock
Low shock
45 degree line
Entrepreneur's wealth
(c) Fraction of wealth invested in the firm
Entrepreneur's wealth
(d) Shares owned by the entrepreneur
Entrepreneur's wealth

Figure 3:

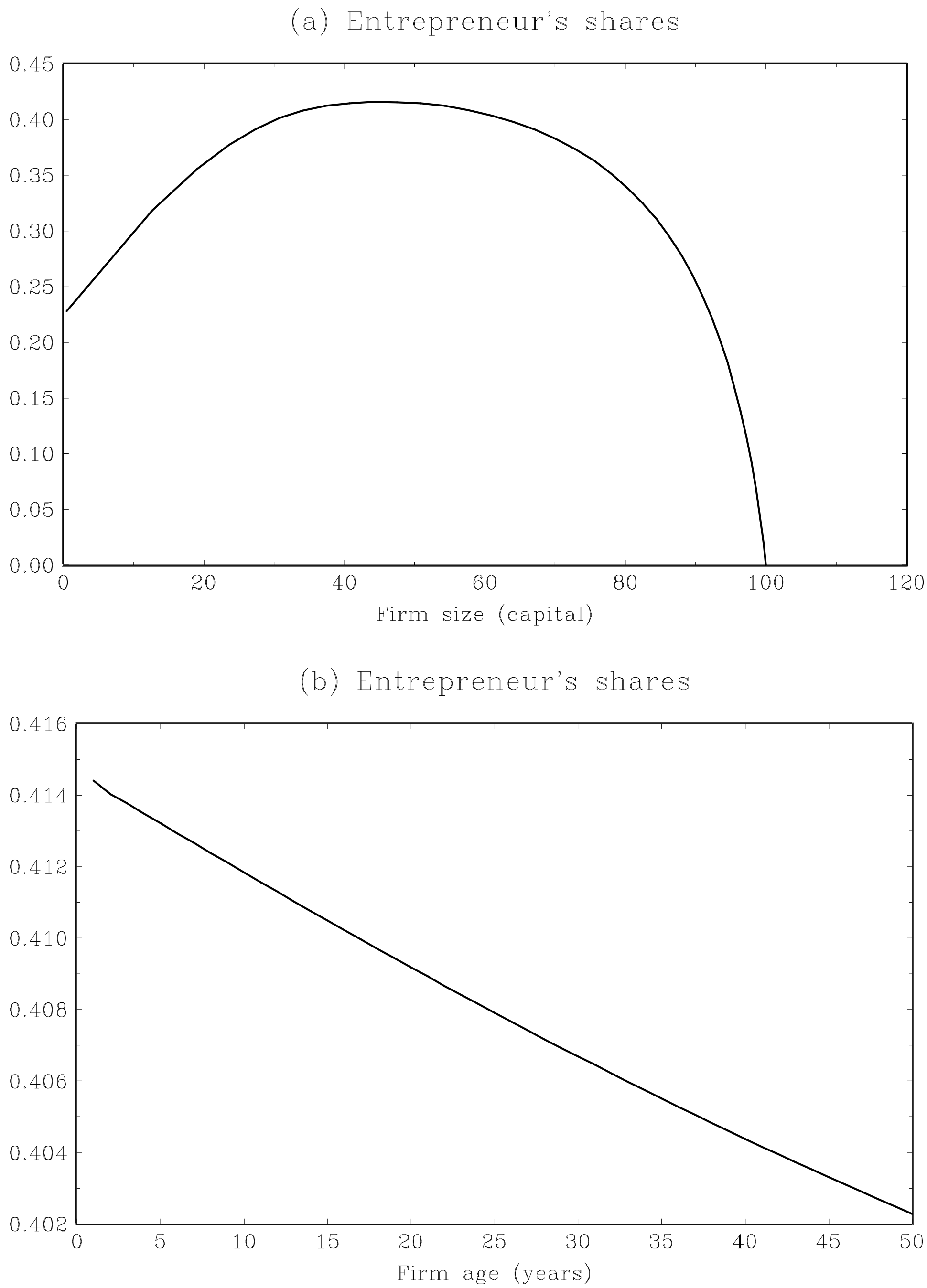
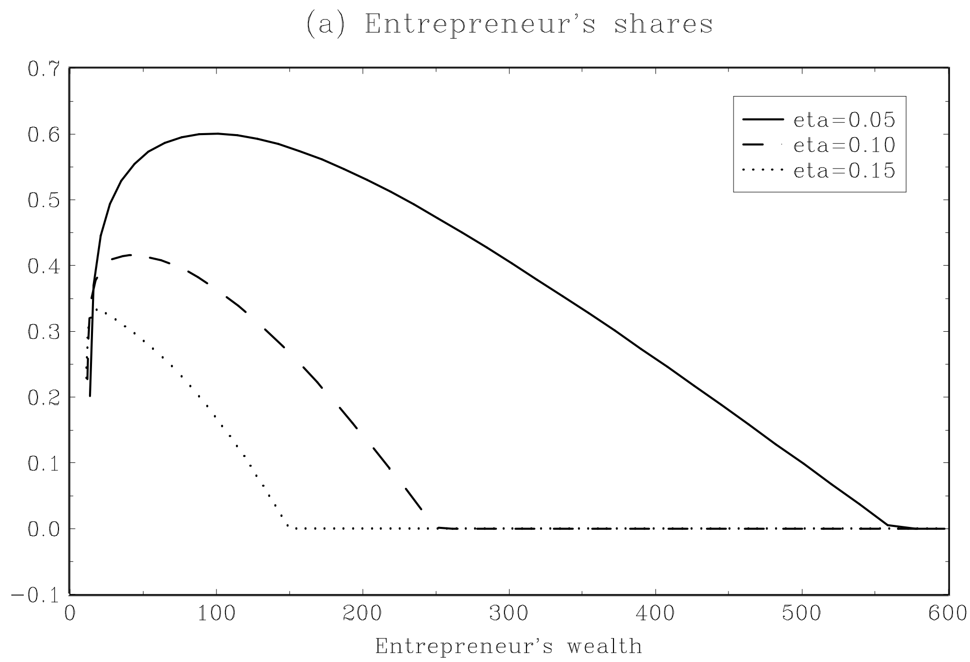

(a) Entrepreneur's shares
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
20
40
60
80
100
120
Firm size (capital)
(b) Entrepreneur's shares
0.416
0.414
0.412
0.410
0.408
0.406
0.404
0.402
0
5
10
15
20
25
30
35
40
45
50
Firm age (years)

Figure 4:


(a) Entrepreneur's shares
eta=0.05
eta=0.10
eta=0.15
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
0
100
200
300
400
500
600
Entrepreneur's wealth

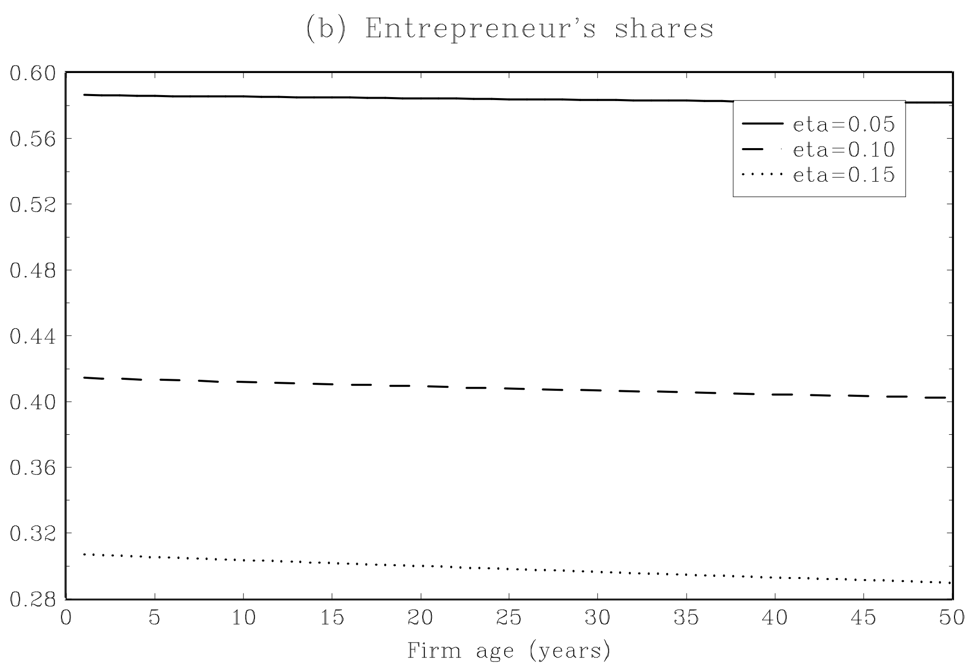

(b) Entrepreneur's shares
eta=0.05
eta=0.10
eta=0.15
0.60
0.56
0.52
0.48
0.44
0.40
0.36
0.32
0.28
0
5
10
15
20
25
30
35
40
45
50
Firm age (years)

Figure 5:

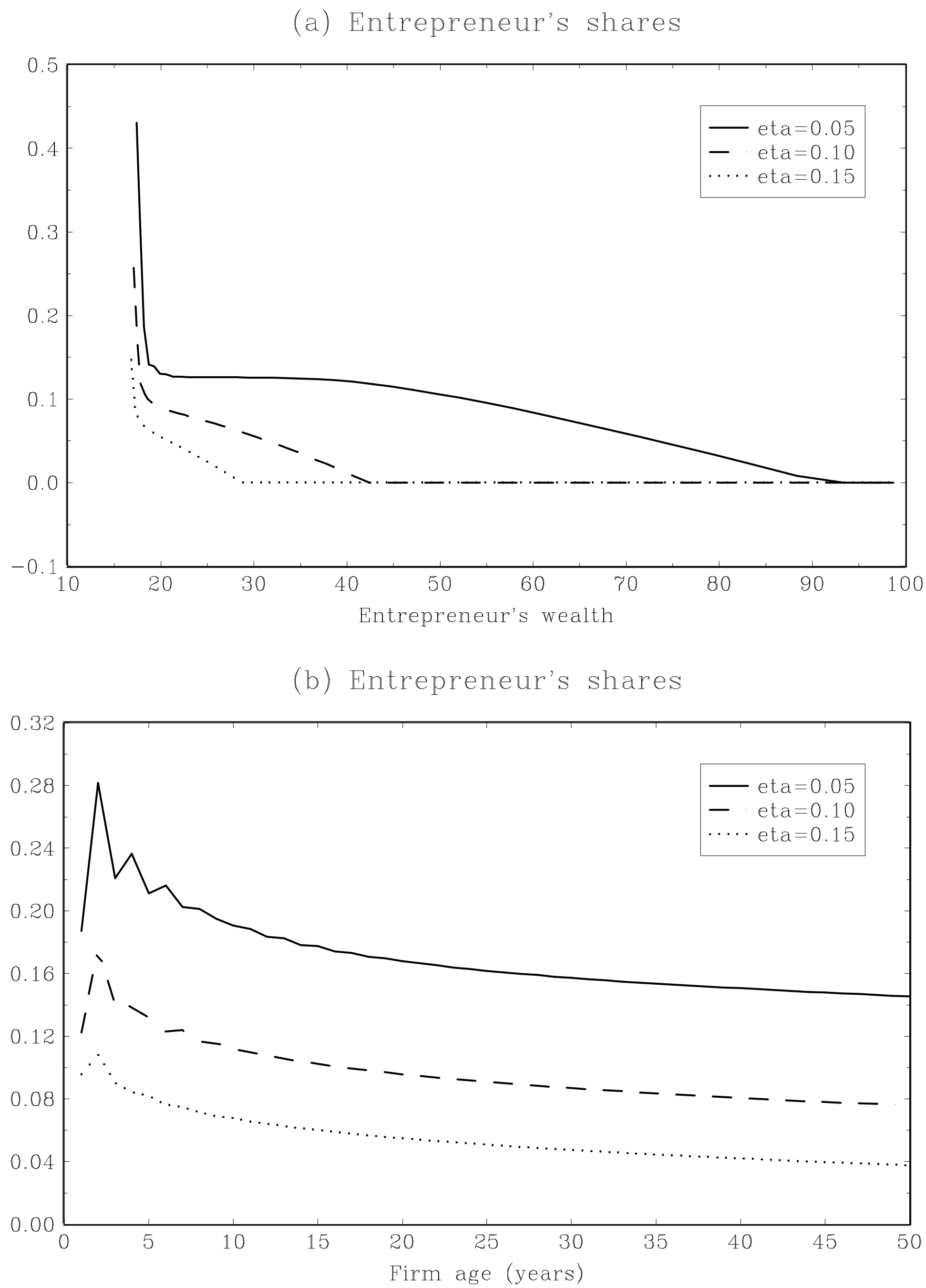

(a) Entrepreneur's shares
eta=0.05
eta=0.10
eta=0.15
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
10
20
30
40
50
60
70
80
90
100
Entrepreneur's wealth
(b) Entrepreneur's shares
eta=0.05
eta=0.10
eta=0.15
0.32
0.28
0.24
0.20
0.16
0.12
0.08
0.04
0.00
0
5
10
15
20
25
30
35
40
45
50
Firm age (years)

Figure 8: