THE RIMINI CENTRE FOR ECONOMIC ANALYSIS

WP 33-08

GIAN LUCA CLEMENTI
Stern School of Business, New York University, USA
and The Rimini Centre for Economics Analisys, Italy

THOMAS COOLEY
Stern School of Business, New York University and NBER, USA

SONIA DI GIANNATALE
Centro de Investigaci´on y Docencia Econ´omicas, M´exico D.F., Mexico

"A THEORY OF FIRM DECLINE"

The Rimini Centre for Economic Analysis
Legal address: Via Angherà, 22 – Head office: Via Patara, 3 - 47900 Rimini (RN) – Italy
www.rcfea.org - secretary@rcfea.org

A Theory of Firm Decline*

Gian Luca Clementi† Thomas Cooley‡ Sonia Di Giannatale §

This version: November 16, 2008
(Download the most recent version)

Abstract

We study the problem of an investor that buys an equity stake in an entrepreneurial venture, under the assumption that the former cannot monitor the latter's operations. The dynamics implied by the optimal incentive scheme is rich and quite different from that induced by other models of repeated moral hazard. In particular, our framework generates a rationale for firm decline. As young firms accumulate capital, the claims of both investor (outside equity) and entrepreneur (inside equity) increase. At some juncture, however, even as the latter keeps on growing, capital and firm value start declining and so does the value of outside equity. The reason is that incentive provision becomes costlier as inside equity grows. In turn, this leads to a decline in the constrained–efficient level of effort and therefore to a drop in the return to investment. In the long run, the entrepreneur gains control of all cash–flow rights and the capital stock converges to a constant value.

Key words. Principal–Agent, Moral Hazard, Hidden Action, Incentives, Firm Dynamics.

JEL Codes: D82, D86, D92, G32.

*We are very grateful to Dave Backus, Heski Bar–Isaac, Alberto Bisin, Andre De Souza, Kose John, and Tom Sargent, as well seminar attendants at the Minneapolis Fed, NYU, the 2008 Midwest Macro Conference in Philadelphia, SED Meeting in Cambridge, and EEA conference in Milan, for their comments and suggestions. All remaining errors are our own responsibility.

†Department of Economics, Stern School of Business, New York University and RCEA. Email: clem@nyu.edu. Web: http://pages.stern.nyu.edu/~gclement

‡Department of Economics, Stern School of Business, New York University and NBER. Email: tcooley@stern.nyu.edu. Web: http://pages.stern.nyu.edu/~tcooley/

§Centro de Investigación y Docencia Económicas, México D.F., México.
Email: sonia.digiannatale@cide.edu.
Web: http://www.cide.edu/investigador/profile.php?IdInvestigador=179

1 Introduction

Think of an investor who provides seed financing to an entrepreneurial firm. The success of her investment will depend crucially on the effort that the entrepreneur puts in. Yet, the incentives of the two individuals are not perfectly aligned, as the entrepreneur bears the whole cost of such effort while sharing the pecuniary returns with the investor. When it is difficult to monitor the entrepreneur's conduct, the investor will implement an incentive scheme that links rewards to observables. However, if the entrepreneur is risk–averse, incentive provision will be costly.

This problem can be conveniently cast as a model of infinitely repeated bilateral exchange with hidden action, along the lines of Spear and Srivastava (1987) and Wang (1997). The only caveat is that in their models there is no notion of important features of firm dynamics such as production and capital accumulation. The purpose of this paper is to explicitly model both of them. We do so by assuming that the entrepreneur is equipped with a production function that exhibits decreasing returns and is hit by multiplicative shocks whose probability distribution depends in a natural way on entrepreneurial effort.

The incentive scheme chosen by the investor belongs to the set of constrained–efficient allocations, which in turn consist of sequences of effort provision, payouts, and investment, that maximize the value of the investor's claim (outside equity) for given rewards to the entrepreneur.

Our analysis generates three key insights. The first is that the marginal value of investment is decreasing in the value of the entrepreneur's claim (inside equity). The intuition is as follows. When the entrepreneur's utility function belongs to the CRRA class, incentive provision is costlier, the greater the value of her claim to the venture's cash flows. In turn this means that, everything else equal, the constrained–efficient level of effort provision decreases as the entrepreneur accumulates wealth. By reducing the likelihood of a high productivity shock, this results in a lower marginal value of investment.

The second result, a corollary to the first, is what we call capital overshooting. When agents discount future utility flows at the same rate and the entrepreneur's RRA coefficient is strictly less than 1, the stochastic process for firm size converges to a singleton. This result is in common with other models in the literature on repeated moral hazard. What is different here is that convergence may be from above. It is well known that young firms tend to be relatively small. This means that their marginal product of capital is high, and so are their investment rates. Capital grows. At the same time, optimal incentive provision implies that on average inside equity also grows over time. These two forces have countervailing effects on effort, but the first dominates. Effort provision increases and the distribution of the shock improves. Firm value also increases. However, as inside equity increases, the marginal value of investment decreases. Eventually, capital and effort start declining, and so does firm value. In the limit all cash flows accrue to the entrepreneur. Capital, effort, and inside equity are constant and strictly positive.

Finally, our theory produces a novel justification for the limited span of control argument, which became popular thanks to the pioneering work by Lucas (1978). Such argument consists in positing that the amount of productive resources that it is efficient to put under a manager's control is bounded. In our scenario, this outcome results from the interaction between informational friction and capital accumulation. Limiting the amount of resources under management is the optimal response to the rise in the cost of incentive provision that follows from the increase in entrepreneurial wealth.

Consistently with Rogerson (1985a), we find that the dynamics we have just described is sensitive to our assumptions on preferences. When the entrepreneur is either relatively more impatient or her RRA coefficient is greater than 1, the model allows for a non–degenerate stationary distribution of firm size, firm value, and its split between the two agents.

We consider two extensions of our framework. In one, productivity shocks are persistent. Success in the current period improves the firm's future prospects. This version of the model has a stationary distribution where the value of outside equity is still zero, but effort, investment, and inside equity are time–varying. In the other extension, we assume that the marginal cost of effort depends non–trivially on capital. We find that the qualitative features of firm dynamics we have just described survive in this more general setup.

This paper contributes to a small but growing literature that explores the implications of moral hazard for firm dynamics. Albuquerque and Hopenhayn (2004) and Cooley, Marimon, and Quadrini (2004) consider scenarios where the entrepreneur has limited commitment, while Clementi and Hopenhayn (2006), Brusco and Ropero (2007), Quadrini (2003) and DeMarzo, Fishman, He, and Wang (2008) study the case of hidden information.

Our framework has also close ties to models that allow for capital accumulation in environments where market incompleteness is caused by moral hazard. Among these, the closest work is by Bohacek (2005), who provides conditions under which an economy à la Atkeson and Lucas (1995) admits a stationary and ergodic distribution of consumption.[1] Other papers in this class include Marcet and Marimon (1992), Khan and Ravikumar (2001), and Espino (2005).

Finally, our paper also belongs to a large literature, started by Holmstrom (1979)'s seminal contribution, that analyzes constrained–efficient allocations in principal–agent models with hidden action. Our work is part of the more recent tradition, started by Rogerson (1985b), that explicitly considers repeated relationships. In particular, our model extends Spear and Srivastava (1987), allowing for production and capital accumulation. Notice that a number of papers in this tradition, among which Wang (1997) and Clementi, Cooley, and Wang (2006), have interpreted the principal–agent relationship as one between shareholders and executives. This alternative interpretation is also valid for our model.

The reminder of the paper is organized as follows. The model is introduced in Section 2. The properties of the optimal contract are characterized in Section 3. We find that the most efficient way of building intuition for those properties is to briefly consider two special cases of our environment, namely one with no dynamics and one with dynamics but no capital accumulation. Section 4 is dedicated to comparative statics exercises. We then consider two extensions. In Section 5, we allow for the marginal utility cost of effort to depend on the level of the capital stock. In Section 6 we assume that shocks are autocorrelated. Finally, Section 7 concludes.

2 Model

Time is discrete and is indexed by $t = 1, 2, ...$ There are two agents, who we will refer to as investor and entrepreneur, respectively. The latter is endowed with a production technology, that produces a homogeneous good with capital as the only input. Output (y_t) is given by

$$y_t = \theta_t f(k_t),$$

where $k_t \in [\underline{k}, \overline{k}] \in \Re_+$ and where $\theta_t \in \Theta \subseteq \Re_+$ is a random variable distributed according to the time–invariant distribution function $G(\theta_t|a_t)$. The variable $a_t \in A \equiv$

[1] Our model is quite similar to his component planner problem. The main difference between such problem and ours is that Bohacek (2005)'s component planner is not bound by a period–by–period limited liability constraint. In other words, component planners rent capital on spot markets.

$[\underline{a}, \overline{a}] \in \Re_+$ denotes entrepreneurial effort. We assume that G has a density denoted by g, which is twice continuously differentiable with respect to a, and that Θ is compact.

While the output of the production process is public information, the effort exerted by the entrepreneur is her private information.

At time $t = 0$, the investor provides the entrepreneur with capital k_0. We assume that any further investment must be financed with resources produced internally. Therefore it must be the case that, at all t,

$$c_t + x_t \leq \theta_t f(k_t),$$

where $c_t \geq 0$ is the entrepreneur's consumption and $x_t \geq 0$ represents investment. The law of motion for capital is the usual one:

$$k_{t+1} = (1 - \delta)k_t + x_t,$$

where $\delta \in (0, 1)$ denotes the depreciation rate. The last two conditions imply the following resource constraint:

$$c_t \leq \theta_t f(k_t) + (1 - \delta)k_t - k_{t+1}.$$

We assume that the investor is risk–neutral, while the entrepreneur is risk–averse. The latter's static preferences are represented by the utility function $u(c_t) - a$.[2] The function u is assumed to be bounded, strictly increasing and strictly concave. Agents discount future utility streams at the common rate $\beta \in (0, 1)$.[3]

We allow investor and entrepreneur to employ history–dependent pure strategies. If we let h^0 denote the empty history, then the history at time $t \geq 1$ is given by the sequence $h^t = h^0 \cup \{(\theta_s, k_s)\}_{s=1}^t$. The investor's task is to offer the entrepreneur an incentive scheme (contract) $\sigma = \{a_t(h^{t-1}, k_t), c_t(h^t)\}_{t=1}^{\infty}$ and a sequence of contingent capital levels $\left\{k_{t+1}(h^{t-1}, k_t)\right\}_{t=1}^{\infty}$. This notation reflects the assumption, typical in neoclassical macroeconomics, that investment is chosen at the beginning of every period, before the realization of the shock. The entrepreneur's strategy consists of the sequence $\{a_t(h^{t-1}, k_t)\}_{t=1}^{\infty}$.

The continuation profile of a contract σ from date $t+1$ on, given h^t, k_{t+1}, is denoted as $\sigma|h^t, k_{t+1}$. Conditional on the entrepreneur following the action recommendation

[2] All the formal results that follow are proven for a more general specification of the entrepreneur's preferences, the details of which are given in Appendix A. In particular, such specification allows for the entrepreneur's disutility from effort to depend on the size of the capital stock she manages. This scenario will be considered in Section 5.

[3] In Section 4 we will relax this restriction by considering the case in which the entrepreneur is relatively impatient.

given by $\sigma|h^t, k_{t+1}$, her continuation value is denoted by $\omega(\sigma|h^t, k_{t+1})$. The investor's is denoted by $v(\sigma|h^t, k_{t+1})$.

A contract σ is said to be *feasible* if, at all times and after any history, effort recommendations belong to the set A and the resource constraint is satisfied. More formally,

Definition 1 *A contract σ is feasible if, for all $t \geq 1$,*

$$a_t(h^{t-1}, k_t) \in A, \forall\; h^{t-1}, \tag{1}$$

and

$$0 \leq c_t(h^t) \leq \theta_t f(k_t) + (1-\delta)k_t - k_{t+1}(h^{t-1}, k_t), \forall\; h^t. \tag{2}$$

The incentive compatibility constraint (3) rules out one–shot deviations from the investor's effort recommendation plan at all dates and after all histories.

Definition 2 *A contract σ is incentive compatible if, $\forall t \geq 1$ and $\forall\; h^{t-1},\; k_t$,*

$$a_t(h^{t-1}, k_t) \in \arg\max_a \int_\theta \left\{u(c_t(h^t)) - a + \beta\omega(\sigma|h^t, k_{t+1})\right\} g(\theta_t|a)d\theta_t. \tag{3}$$

The fact that the set A is a connected subset of $\Re_+$ suggests that rewriting (3) as a first–order condition may be handy. In the literature, this is known as the first–order approach, which is not universally valid. To ensure its validity, we follow Rogerson (1985a) and Spear and Srivastava (1987) in assuming that the Monotone Likelihood Ratio Property and the Convexity of the Conditional Distribution Condition hold.

Let Ω be the set of pairs (k,ω) such that there exists a feasible and incentive compatible contract that delivers ω, given k. That is, for $\Delta \in \Re^2$ non–empty and compact, let

$$\Omega \equiv \left\{(k,\omega) \in \Delta \mid \exists\, \sigma \quad \text{s.t.} \quad (1),(2),(3), k_0 = k, \text{ and } \omega(\sigma|h^0) = \omega\right\}.$$

In Appendix A we show that Ω can be recovered by adapting the algorithm developed by Abreu, Pierce, and Stacchetti (1990). Then, for every $(k,\omega) \in \Omega$, we define the set of the investor's expected discounted utilities that can be generated by a feasible and incentive compatible contract:

$$\Phi(k,\omega) = \left\{v(\sigma|h^0) \mid (1),\, (2),\, (3), k_0 = k, \text{ and } \omega(\sigma|h^0) = \omega\right\}.$$

For given (k,ω), the investor's problem is to choose a feasible and incentive compatible contract σ that attains the maximum element in $\Phi(k,\omega)$.[4] Denote such element as $v^*(k,\omega)$. Proposition 2 shows that the function $v^*(k,\omega)$ is a fixed point of an

[4]Proposition 1 in Appendix A proves that $\Phi(k,\omega)$ is compact.

operator, denoted as T, which maps the space of bounded and continuous functions $v : \Omega \to \Re$ into itself, with the sup norm, and is given by

$$T(v)(k,\omega) \equiv \max_{a^*,k',c(\theta),\omega'(\theta)} \int_\Theta \left\{\theta f(k) - c(\theta) - k' + (1-\delta)k + \beta v(k',\omega'(\theta))\right\} g(\theta|a^*)d\theta$$

$$\text{s.t.} \int_\Theta \left\{u(c(\theta)) - a + \beta\omega'(\theta)\right\} g(\theta|a^*)d\theta = \omega, \tag{4}$$

$$a^* \in \arg\max_a \int_\Theta \left\{u(c(\theta)) - a + \beta\omega'(\theta)\right\} g(\theta|a)d\theta, \tag{5}$$

$$0 \leq c(\theta) \leq \theta f(k) - k' + k(1-\delta) \ \ \forall\theta \in \Theta, \tag{6}$$

$$a^* \in A, \tag{7}$$

$$(k', \omega'(\theta)) \in \Omega \ \ \forall\theta \in \Theta. \tag{8}$$

Since the operator T satisfies Blackwell's sufficient conditions for a contraction, the contraction mapping theorem ensures that the fixed point is unique. Solving for it also yields policy functions for recommended effort $a(k,\omega)$, entrepreneur's cash flows $c(k,\omega,\theta)$, and continuation utility $\omega'(k,\omega,\theta)$, which can be used to recover the constrained–efficient contract in a straightforward manner.

Unfortunately an analytical characterization of the constrained Pareto–optimal contract is not possible. For this reason we make assumptions about functional forms and parameters (see below) and characterize the contract by means of numerical methods. The algorithms that were designed to approximate the set Ω and the value function are described in Appendix B.

2.1 Numerical Implementation

Our functional forms assumptions for preferences and technology are standard in the macroeconomics literature. We posit that $u(\cdot)$ belongs to the CRRA class, i.e. $u(c) = \frac{c^{1-\chi}}{1-\chi}$, $\chi > 0$, $\chi \neq 1$. The production function is $f(k) = k^\alpha$, $\alpha \in (0,1)$. Furthermore, we assume that $A = [0,\bar{a}]$ and $\Theta = \{\theta_l, \theta_h\}$, with $\theta_h > \theta_l$ and $G(\theta_l|a) = e^{-a}$.[5] While this choice of conditional distribution is dictated mostly by tractability, it has appealing features. The probability of a good outcome is zero if no effort is exerted, and goes to 1 as effort goes to infinity. Furthermore, the marginal effect of effort on the probability of success is decreasing in the effort itself. From now on, all variables that are contingent on the shock realization will be denoted with the subscripts l or h.

[5] The upper bound $\bar{a}$ will be chosen so as to ensure it never binds.

Under the assumptions listed above, the Bellman equation is as follows:

$$v(k,\omega) = \max_{a^*,k',\{c_i,\omega_i\}_{i=h,l}} (1-e^{-a^*})[\theta_h k^\alpha - c_h + \beta v(k',\omega_h)] + e^{-a^*}[\theta_l k^\alpha - c_l + \beta v(k',\omega_l)] + k(1-\delta) - k' \tag{P}$$

$$\text{s.t.} \quad (1-e^{-a^*})\left[u(c_h)+\beta\omega_h\right] + e^{-a^*}\left[u(c_l)+\beta\omega_l\right] - a^* = \omega, \tag{9}$$

$$a^* \in \arg\max\,(1-e^{-a})\left[u(c_h)+\beta\omega_h\right] + e^{-a}\left[u(c_l)+\beta\omega_l\right] - a, \tag{10}$$

$$0 \le c_i \le \theta_i k^\alpha - k' + (1-\delta)k \;\; \forall\, i = h,l, \tag{11}$$

$$a^* \in A, \tag{12}$$

$$\left(k',\omega_i\right) \in \Omega \;\; \forall\, i = h,l. \tag{13}$$

The parameter values are reported in Table 1. Even though we set β, χ, α, and δ to values that are standard in the macroeconomics literature, we wish to emphasize that by no means should this be considered a calibration exercise.

The comparative statics exercises presented in Section 4 indicate that the choice of χ is particularly relevant. In our benchmark, we set $\chi = 0.5$ because of the rapidly amassing experimental evidence in favor this value.[6]

$\underline{k}$	$\bar{k}$	β	χ	α	δ	θ_h	θ_l	$\underline{\omega}$
0	3.5	0.95	0.5	0.3	0.1	1.5	0.4	10.5

Table 1: Parameter Values.

Figure 1 depicts the equilibrium value set Ω and the value function. As expected, the range of values that is feasible and incentive compatible to promise the entrepreneur is wider, the higher the level of capital.

The value function is nothing but the expected discounted value of the cash flows that will accrue to the investor. For this reason, it can be thought of as outside equity. The illustration in the right panel shows that it is strictly increasing in the level of capital and strictly decreasing in the entrepreneur's promised utility ω. When the value function is differentiable, these properties follow from the application of the envelope theorem. In all of our computations, the value function is also globally concave. However, since standard sufficient conditions for concavity are not satisfied (the constraint set of the dynamic problem is not convex), we cannot assert this as a general property.

[6] See for example Choi, Fisman, Gale, and Kariv (2007), Goeree, Holt, and Palfrey (2002), and Holt and Laury (2002).

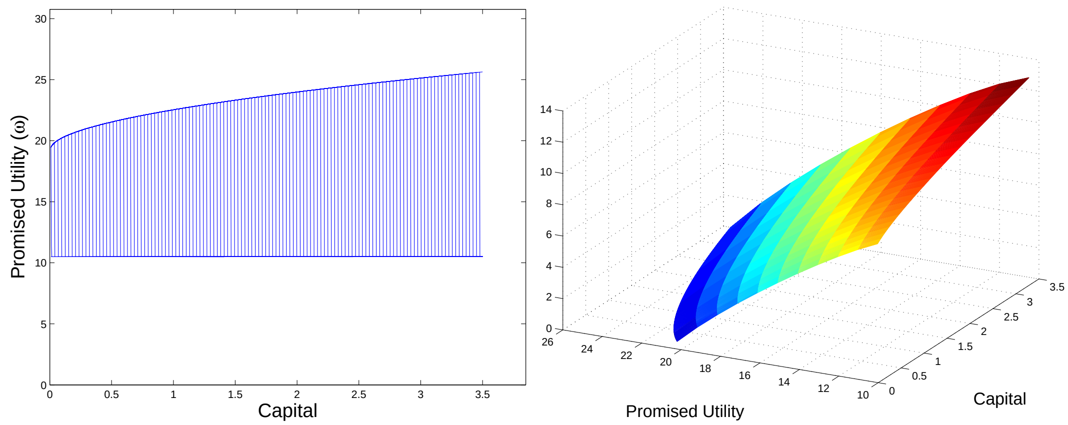

Figure 1: Set Ω and Value Function.

Given a pair $(k, \omega) \in \Omega$, we denote as $C(k, \omega)$ the expected discounted cost to the investor of delivering ω to the entrepreneur, when the current capital stock is k. The function $C(k, \omega)$ is the solution to the following functional equation:

$$C(k, \omega) = (1 - e^{-a^*})[c_h^* + \beta C(k'^*, \omega_h^*)] + e^{-a^*}[c_l^* + \beta C(k'^*, \omega_l^*)],$$

where the asterisks designate the optimal choices generated by problem (P). The function $C(k, \omega)$ is rendered in Figure 2.

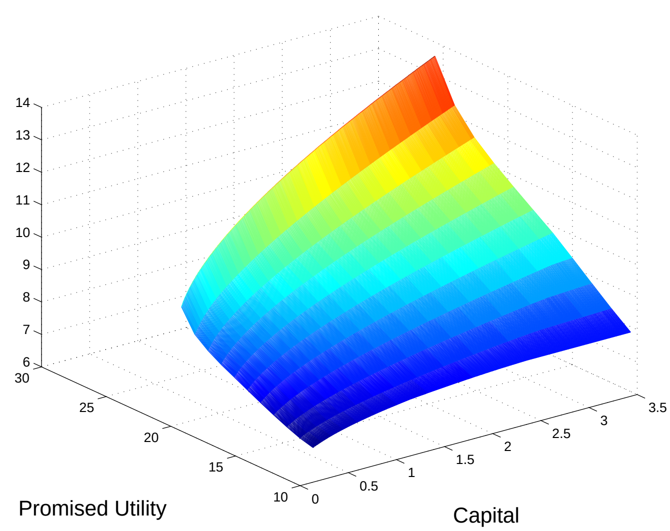

Figure 2: Inside Equity.

Notice that $C(k, \omega)$ is also the expected discounted value of the cash flows that will accrue to the entrepreneur. For this reason, we will often refer to it as the value of inside equity. It is intuitive that it is increasing in both capital and promised utility.

3 Properties of the Optimal Contract

The purpose of this section is to illustrate the properties of the optimal incentive scheme and its implications for firm dynamics. We find it convenient to start by illustrating the optimal contracts that obtain in two simpler environments. In Section 3.1 we will consider the static case, i.e. the scenario where the relationship between investor and entrepreneur lasts only one period. In Section 3.2 we will analyze the case of infinitely repeated interaction without capital accumulation. Finally, in Section 3.3, we will tackle the general case.

3.1 The Static Case

Let $u_i \equiv c_i^{1-\chi}/(1-\chi)$. With some abuse of notation, also define the inverse of the utility function as $c(u) = [(1-\chi)u]^{1/(1-\chi)}$. It yields the cost to the investor of delivering a certain level of utility and is strictly increasing and strictly convex.

We find it convenient to let the investor choose utilities rather than cash–flows. Then, whenever the recommended effort is strictly positive, her optimization program writes as

$$\max_{u_h, u_l} (1 - e^{-a^*}) \left[\theta_h k^\alpha - c(u_h)\right] + e^{-a^*} \left[\theta_l k^\alpha - c(u_l)\right] - \delta k,$$

$$\text{s.t.} \quad (1 - e^{-a^*}) u_h + e^{-a^*} u_l - a^* = \omega, \tag{14}$$

$$a^* = \log(u_h - u_l), \tag{15}$$

$$0 \leq u_i \leq u(\theta_i k^\alpha - \delta k), \quad \forall\, i = h, l.$$

The effort chosen by the entrepreneur is a simple, monotone increasing function of the utility spread $s \equiv u_h - u_l$. A higher effort is implementable only by increasing the gap between contingent rewards. Using (14) and (15) to express u_h and u_l as functions of s, the problem simplifies further to

$$\max_{s} \left(1 - \frac{1}{s}\right) \left[\theta_h k^\alpha - c(u_h)\right] + \frac{1}{s} \left[\theta_l k^\alpha - c(u_l)\right] - \delta k,$$

$$\text{s.t.} \quad u_h = 1 + \omega + \log(s),$$

$$u_l = 1 + \omega + \log(s) - s,$$

$$0 \leq u_i \leq u(\theta_i k^\alpha - \delta k) \quad \forall\, i = h, l.$$

Notice that the choice of s pins down the payoffs of the lottery offered to the entrepreneur $\{1+\omega+\log(s), 1+\omega+\log(s)-s\}$, along with the the respective probabilities $\{1-1/s, 1/s\}$.

Necessary condition for an interior solution to the above optimization problem is

$$\frac{1}{s^2}[\theta_h - \theta_l]k^\alpha = \frac{1}{s^2}\left[c(u_h) - c(u_l)\right] + \frac{s-1}{s^2}\left[c'(u_h) - c'(u_l)\right]. \tag{16}$$

The left–hand side is the marginal gain in expected revenues resulting from the increase in the probability of a good outcome. The right–hand side is the marginal increase in the cost of compensating the entrepreneur. The first term reflects the increased probability of awarding u_h rather than u_l. The second term reflects the marginal impact on the expected cost arising from an increase in the risk imposed on the entrepreneur. By strict concavity of the utility function, this terms is also positive.

Condition (16) can be used to characterize the comparative statics of s with respect to ω and k, respectively. Notice first that a larger k implies a strictly higher marginal revenue, but has no effect on the marginal cost. Therefore, the choice of s will be strictly increasing in k. Now consider the effect of an increase in ω. The marginal revenue does not change. The first term of the marginal cost always increases. This follows from strict concavity of the utility function and from the fact that higher ω implies higher utility in both states of nature. The effect on the second term depends on the third derivative of the inverse of the utility function. It will be non–negative as long as $\chi \geq 1/2$. This means that the latter inequality is also a sufficient condition, although not necessary, for the optimal choice of s to be decreasing in ω. In fact, in the case of our parameterization s is decreasing in ω for much lower values of χ.

We conclude that in the static model (i) higher capital elicits higher effort and (ii) awarding larger spreads s is costlier the higher the level of promised utility. It follows that the recommended level of effort decreases with ω.

3.2 Dynamics without Capital Accumulation

We now consider the case in which the time horizon is infinite, but there is no capital accumulation. This scenario is very close to those analyzed by Spear and Srivastava (1987) and Wang (1997). It differs from the former in that we impose limited liability, i.e. the entrepreneur's cash–flow must be non–negative. It differs from the latter, since Wang assumes the effort choice to be binary. With slight abuse of notation, let now $s \equiv u_h + \beta\omega_h - (u_l + \beta\omega_l)$. When recommended effort is strictly positive, it is still the case that $a^* = \log(s)$. The value function $v(\omega)$ solves the following functional equation:

$$v(\omega) = \max_{s,\omega_h,\omega_l} \left(1 - \frac{1}{s}\right) [\theta_h k^\alpha - c(u_h) + \beta v(\omega_h)] + \left(\frac{1}{s}\right) [\theta_l k^\alpha - c(u_l) + \beta v(\omega_l)] - \delta k,$$

$$\text{s.t. } u_h = \omega + 1 + \log(s) - \beta\omega_h,$$

$$u_l = \omega + 1 + \log(s) - s - \beta\omega_l,$$

$$0 \leq u_i \leq u(\theta_i k^\alpha - \delta k) \;\; \forall\, i = h, l,$$

$$\omega_i \in \Omega \;\; \forall\, i = h, l.$$

Differentiating with respect to s yields the analogue of condition (16):

$$\frac{1}{s^2}[\theta_h - \theta_l]k^\alpha + \frac{\beta}{s^2}[v(\omega_h) - v(\omega_l)] = \frac{1}{s^2}\left[c(u_h) - c(u_l)\right] + \frac{s-1}{s^2}\left[c'(u_h) - c'(u_l)\right]. \quad (17)$$

The forces that shape the optimal spread of utilities across states are essentially the same as in the static case. The only difference is that the marginal benefit of increasing effort also depends on the difference between the investor's contingent continuation values.

The main innovation is that now the investor has the opportunity to distribute utility awards both across states and over time. If the value function is differentiable, necessary conditions for an interior solution are

$$c'(u_i) = \left(1 - \frac{1}{s_i}\right) c'(u_{ih}) + \left(\frac{1}{s_i}\right) c'(u_{il}), \;\; i = h, l, \quad (18)$$

where s_i, u_{ih}, and u_{il} denote next period's choices contingent on the current state of nature being i. Condition (18) is analogous to condition 2.2 in Rogerson (1985b). As in Rogerson (1985b), the value of χ determines the dynamics of entrepreneurial consumption. For $\chi < 1$, payments to the entrepreneur follow a sub–martingale, while for $\chi > 1$ they follow a super–martingale. This result, however, does not necessarily generalize to scenarios where the limited liability constraint binds. This is relevant to us, as in our parameterization that constraint does bind for large enough ω.

Figure 3 plots the optimal policies for the entrepreneur's current and promised utility, respectively. The investor strives to equate the marginal costs of rewarding the entrepreneur across periods. For this reason all utility awards, both current and future, are increasing in ω. If the intuition gained in the static case generalized to the one under analysis here, one would also expect the spread between contingent utility awards to decrease with ω, both in the present and in the future. According to Figure 3, this happens for relatively low ω, but not for larger values. For relatively high levels of the state variable, the gap between u_h and u_l actually widens. The reason is that, as just argued, the limited liability constraint $u_l \leq u(\theta_l k^\alpha - \delta k)$ binds. This implies a tension between the objectives of minimizing the distance between marginal costs of utility awards across periods and across states. When the limited liability constraint binds, raising the entrepreneur's expected current payoff necessarily requires increasing the spread in current utilities. Our conjecture is that if the investor was able to borrow in the bad state of nature, so that the resources available to her were the same as in the good state, optimal recommended effort and the utility spread s would converge to zero as ω goes to the upper bound of Ω. That is, eventually the entrepreneur would achieve perfect insurance (and would exert not effort).[7]

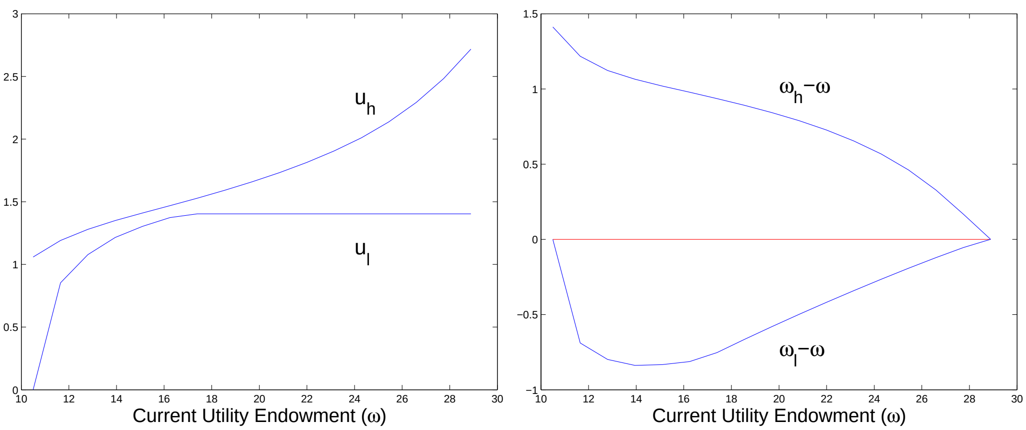

Figure 3: Compensation Policies.

In spite of the fact that, as argued above, the sufficient condition for entrepreneurial consumption to be a sub–martingale does not hold always, our simulations show that ω converges to the maximum element in the set Ω. By the limited liability constraint, this also implies that the investor's payoff converges to zero, its lower bound.

3.3 Dynamics with Capital Accumulation

We now turn to the general case with capital accumulation. Figure 4 depicts the policy functions for current and promised utility. In the left panel, we have plotted $u_i(k, \omega)$, $i = h, l$. In the right panel, we have pictured the contingent variation in promised utility $\omega_i(k, \omega) - \omega$, $i = h, l$. For given capital stock, the entrepreneur's

[7]This modification would require writing the limited liability constraints as $0 \leq c_i \leq [1 - e^{-a^*}]\theta k^\alpha + e^{-a^*} k^\alpha - \delta k \ \ \forall \, i = h, l$.

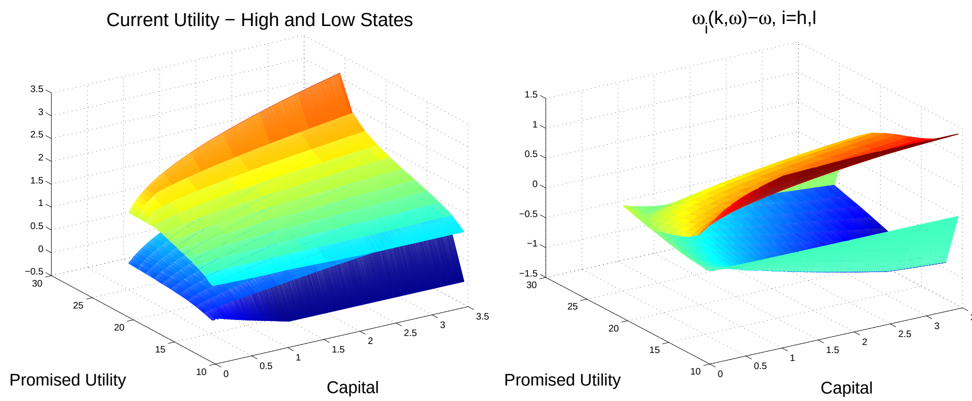

Figure 4: Compensation Policies.

contingent compensation schedules display the same qualitative features as those in the case with no accumulation. The spread between continuation utilities appears to be decreasing in ω, while the spread between current utilities is decreasing in ω for low values and increasing for high values. Once again, this is due to the fact that the limited liability constraint binds in the bad state of nature.

The optimality conditions for s and u_i, $i = h, l$, are the analogues of equations (17) and (18). Refer to the left panel of Figure 5. Consistent with the intuition developed in previous sections, recommended effort is increasing in the capital stock and decreasing in the level of promised utility ω.

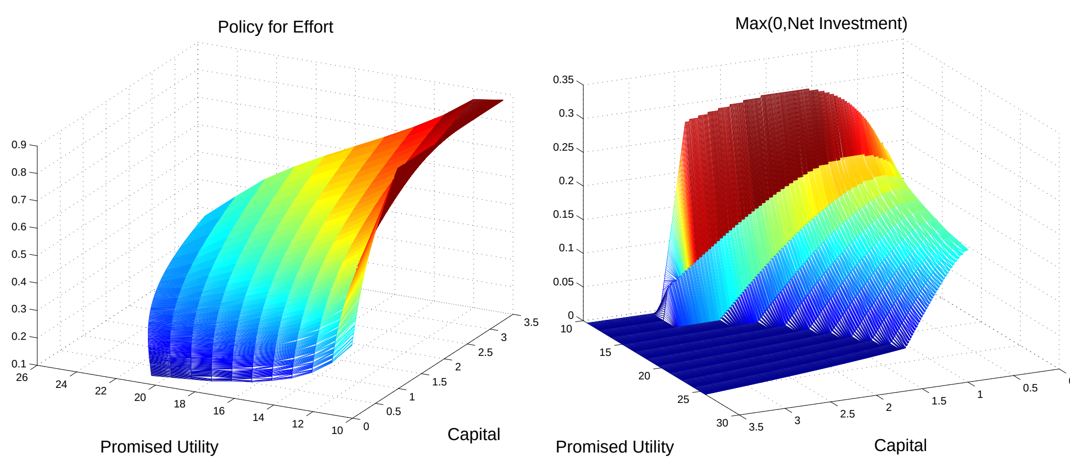

Figure 5: Policy Functions for Effort and Net Investment.

The novelty is capital accumulation. The policy function for net investment is rendered in the right panel of Figure 5, were we plotted $\max[0, k'(k,\omega) - k]$. The most interesting feature is that, for given capital, net investment is declining in ω. We can use the optimality conditions to understand why this is the case.

When the value function is differentiable, the first–order condition for the capital choice is

$$\left(1-\frac{1}{s}\right)\frac{\partial v(k',\omega_h)}{\partial k'}+\frac{1}{s}\frac{\partial v(k',\omega_l)}{\partial k'}\geq\frac{1}{\beta}. \tag{19}$$

By the envelope theorem, it follows that

$$\frac{\partial v(k,\omega)}{\partial k}=\left[\left(1-\frac{1}{s}\right)\theta_h+\frac{1}{s}\theta_l\right]\alpha k^{\alpha-1}+(1-\delta)>0. \tag{20}$$

Differentiating with respect to ω yields

$$\frac{\partial^2 v(k,\omega)}{\partial k\partial\omega}=\frac{1}{s^2}(\theta_h-\theta_l)\alpha k^{\alpha-1}\frac{\partial s}{\partial\omega}.$$

Since $\partial s/\partial\omega<0$, the latter says that the marginal value of increasing capital is decreasing in ω. The higher ω, the lower the utility spread s and the probability of success. In turn, this leads to a lower marginal value of investment.

Now fix k and consider the effect of increasing ω on the optimal choice of investment. This is tantamount to differentiating (19) with respect to ω. The first term, $\frac{1}{s^2}\left[\frac{\partial v(k',\omega_h)}{\partial k'}-\frac{\partial v(k',\omega_l)}{\partial k'}\right]\frac{\partial s}{\partial\omega}$, is positive. An higher ω today leads to a lower utility spread s, which in turn increases the probability of a bad outcome. Since $\omega_h>\omega_l$, this means that the effect on the marginal gain is positive. On the other hand, since ω_h and ω_l are strictly increasing in ω, the second term $\left(1-\frac{1}{s}\right)\frac{\partial^2 v(k',\omega_h)}{\partial k'\partial\omega_h}\frac{\partial\omega_h}{\partial\omega}+\frac{1}{s}\frac{\partial^2 v(k',\omega_l)}{\partial k'\partial\omega_l}\frac{\partial\omega_l}{\partial\omega}$ is negative. In our simulations, the latter effect dominates. Investment decreases with ω.

The simulation of the system starting from an arbitrary point in the state space sheds more light on the forces shaping its dynamics. Refer to Figure 6. The state space is partitioned in four subsets. The policy for capital is such that $k'(k,\omega)>k$ in regions A and B, and $k'(k,\omega)<k$ otherwise. The dynamics of promised utility is such that in regions B and C, $\omega_h(k,\omega)>\omega>\omega_l(k,\omega)$. In region A, ω always increases. In region D, it always decreases. In light of our previous discussion, it is not surprising that the locus separating regions B and C, along which net investment is identically zero, is downward sloping for most values of k.

For relatively low initial conditions, the scatter plot illustrates the dynamics of the state variables. The paths followed by the other relevant variables are shown in Figure 7. When k and ω are relatively low, the marginal product of capital is high and providing incentives is relatively inexpensive. Therefore the returns to investment

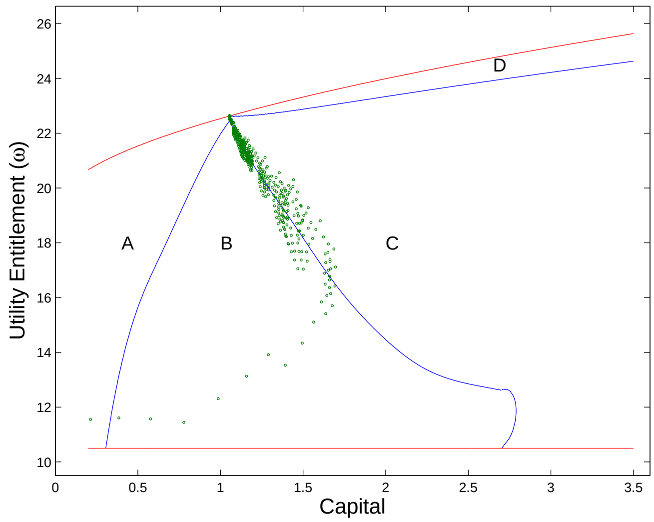

Figure 6: Sample Path.

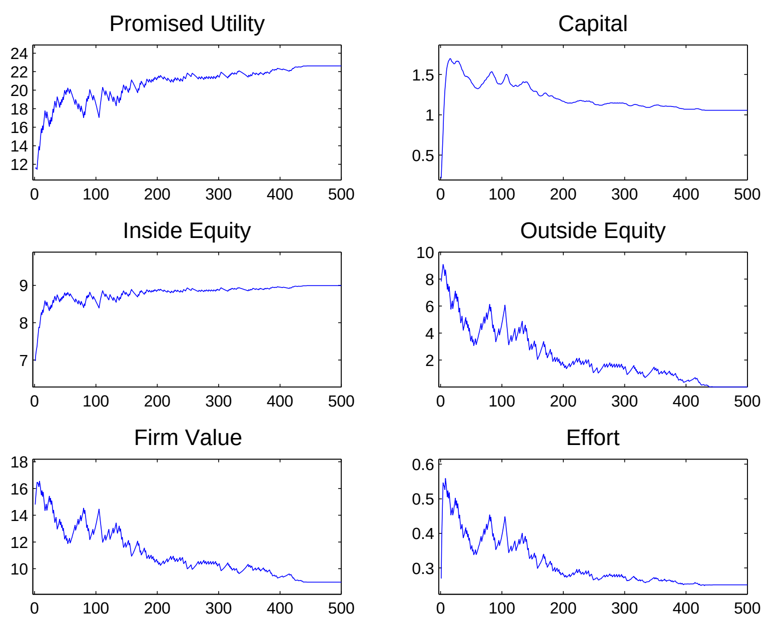

Figure 7: Sample Path.

are high. Capital grows in both states of nature. During the transition towards the locus separating the partitions B an C, all other variables also increase on average. However, increases in both ω and k lead to a progressive reduction in the marginal gain from capital accumulation. Once reached the locus, incentive provision becomes so expensive to discourage investment. From that moment onwards, positive shocks lead to contemporaneous increases in current and future payouts to the entrepreneur, and to lower investment and lower effort in the next period. The opposite is true, conditional on negative shocks. Therefore, inside equity, effort, and capital are positively correlated with each other. The continuation value of the investor's claim is negatively correlated with all of them. Because of the sub–martingale property described above, on average ω increases and capital decreases. As a consequence, effort and outside equity also decrease on average.

Eventually the system converges to a stationary point where outside equity is zero and the entrepreneur's promised utility lies on the upper contour of the set Ω. The constrained–efficient arrangement prescribes that in the limit the entrepreneur controls the totality of the cash–flow rights. Notice further that, even in such predicament, she will not achieve full insurance and therefore her effort will be strictly positive. The latter result is a consequence of our assumptions on limited liability. As already noted in Section 3.2, if the investor had access to a competitive insurance market, the entrepreneur would end up receiving full insurance and exerting no effort.

Staring at Figure 8 confirms what we have learned so far. We have initialized the system by assigning to the state variables the same initial conditions used to construct Figure 7. Then we have conducted a large number of 80–period long simulations. Figure 8 reports the simple averages. Early on all variables tend to increase. However, once ω reaches a certain value, effort, capital, profits and the value of the investor's claim start decreasing.

We close this Section by emphasizing two implications of our theory that we find of particular interest. The first is that it provides a novel rationale for the assumption of limited span of managerial control, i.e. for the idea that the amount of productive resources that it is efficient to put under a manager's control is bounded (see Lucas (1978)). Even in the case in which the production function displayed constant returns to scale, our theory would still predict that firm size converges. This is the case because the process for ω would still be a sub–martingale, and promised utility would eventually become so high (and incentive provision so costly) to discourage further investment.

The second property is what we call capital overshooting. Our theory predicts that, on the transition to its rest point, capital overshoots its long–run value. As far as we know, this feature is unique to our framework. In other papers that study firm dynamics in environments with repeated hidden information, capital is always weakly

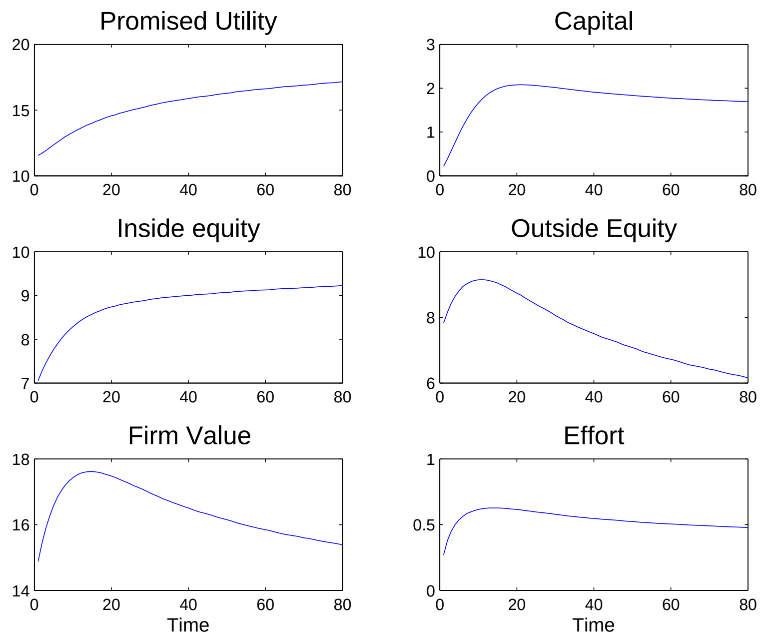

Figure 8: Average Dynamics.

lower then its long–run value.[8]

4 Comparative Statics

In this section, we document how the optimal contract and the implied dynamics change when we select alternative values for either the coefficient of relative risk aversion χ, or the entrepreneur's discount factor, or the support of the conditional distribution Θ.

4.1 The case of $\chi > 1$

The analysis conducted so far shows that the optimal contract is back–loaded. That is, the payoff to the entrepreneur is stochastically increasing over time. How general is this feature?

In a simpler environment, without capital accumulation, Rogerson (1985b) showed that it is not robust to increases in the relative risk aversion beyond 1 ($\chi > 1$, in our notation).[9]

[8]See for example Brusco and Ropero (2007), Clementi and Hopenhayn (2006), Quadrini (2003) and DeMarzo, Fishman, He, and Wang (2008).

[9]See his Proposition 3.

Our numerical results tell us that the same holds true here. Figure 9 depicts the frequency distribution that obtains by running a very long simulation of our model for $\chi = 2$. Differently from the benchmark case discussed above, the optimal contract implies a non–degenerate stationary distribution, with most of the mass clustered around a downward–sloping curve which is the analogue of the locus partitioning the set Ω in Figure 6. Once again, this feature reflects the fact that the marginal gain from investing is decreasing in the promised utility ω.

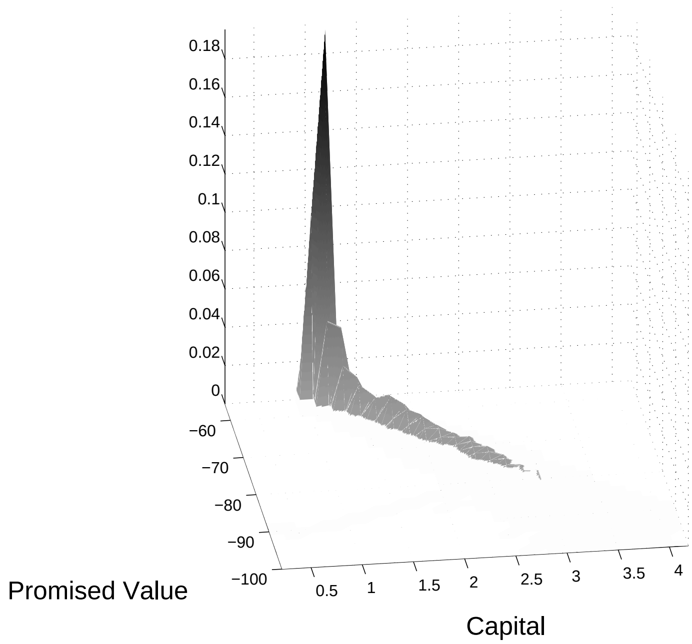

Figure 9: Stationary Distribution for $\chi = 2$.

The intuition behind this result is rather simple. As the risk aversion coefficient increases, the intertemporal elasticity of substitution decreases. In turn, this implies that distorting intertemporally the entrepreneur's consumption profile by postponing her consumption is costlier. Constrained–optimality requires the lender to leave more resources to the entrepreneur early on.

4.2 An Impatient Entrepreneur

Throughout Section 3 it was assumed that entrepreneur and investor discount future utility flows at the same rate. Here we consider the case in which the entrepreneur discounts future utility at the rate $\rho < \beta$. To start with, Ω, the set of implementable utility levels is now smaller. The reason is that incentive provision is more expensive.

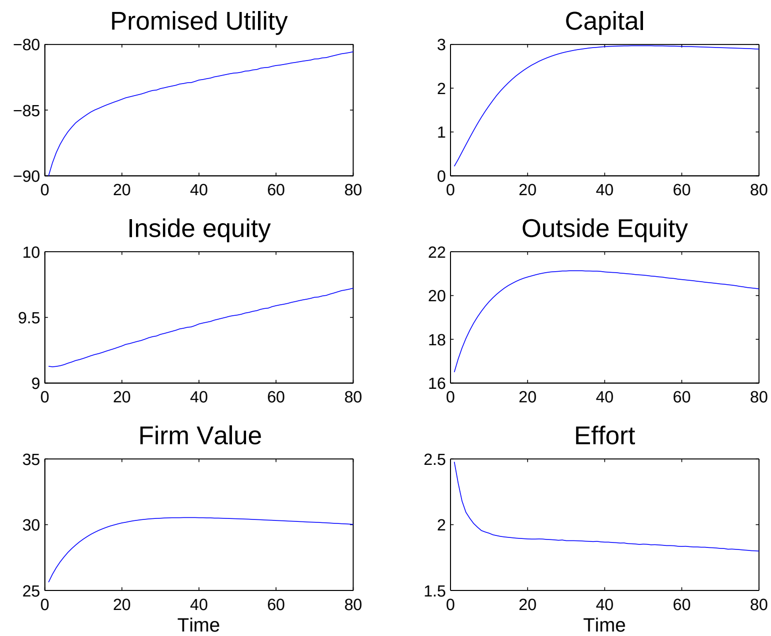

Figure 10: Average Dynamics for $\chi = 2$.

In a sense, the investor's ability to spread compensation over time is compromised.

Figure 11 compares the policy functions with those obtained in the benchmark case. Given the entrepreneur's preference for early consumption, the optimal contract calls for a change in the time profile of her cash flows in favor of the current period. In part, this is accomplished by increasing the distribution to the entrepreneur in the low state. Since the limited liability constraint binds, this translates into fewer resources available for investment. The distribution contingent on high state also increases, to an extent that increases the income risk imposed on the entrepreneur. In turn this increases the marginal cost of eliciting effort. This is why the latter drops.

The implications for dynamics is that, similarly to the case of $\chi > 1$, the ergodic set is no longer degenerate. There will be a stationary distribution, whose shape resembles that depicted in Figure 9.

4.3 The Role of Risk

Figure 12 gives a rendering of the impact of simultaneously lowering θ_h and raising θ_l, thereby reducing the gap between the two realizations of the productivity shock. Once again, the policy functions are compared to those that obtain in the benchmark case of which in Section 3.

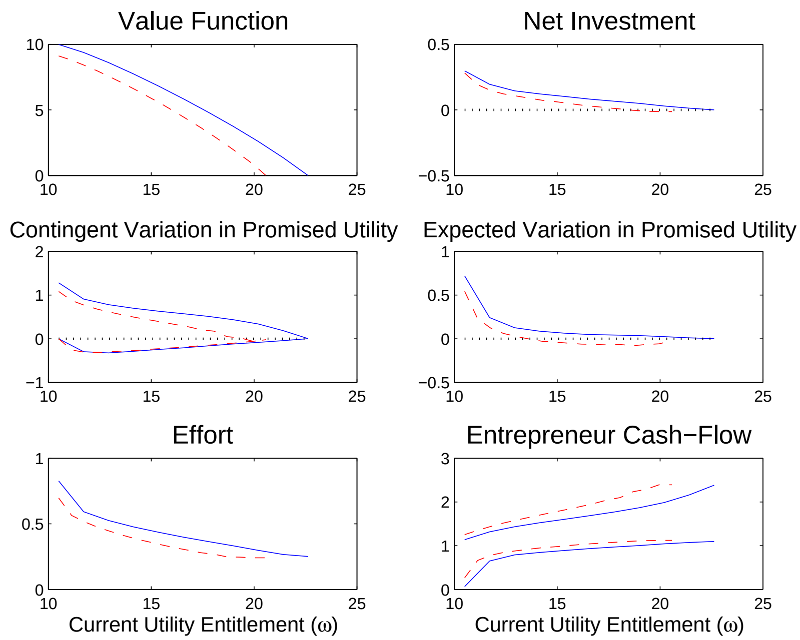

Figure 11: Policy Functions. Solid: benchmark. Dashed: $\rho = 0.945$

Net investment increases. The reason is that, for higher θ_l, more resources are available for distribution to the entrepreneur in the current period, conditional on low shock realization. For given intertemporal distribution of payoffs to the entrepreneur, this allows to narrow the gap between current conditional cash flows accruing to the entrepreneur. In turn this means that, for all continuation values ω', the marginal gain from investment will be higher.

The decrease in effort is the net result of two opposing forces. On the one end, as just argued, a higher θ_l results in a lower marginal cost of eliciting effort. On the other hand, the marginal benefit of effort also drops, as the gap $\theta_h - \theta_l$ closes. The latter turns out to dominate.

5 On Capital and the Utility Cost of Effort

So far we have assumed that the marginal impact of effort on expected revenues is proportional to $e^a f(k)$, an expression which is clearly increasing in the size of the firm. However, we have posited that the marginal utility cost of exerting effort is invariant with respect to firm size. If we think of one unit of effort as the accomplishment of one task, it is easy to think of situations in which the latter assumption looks sensible. Think, for example, at the hiring of a new factory floor technician. The effort required

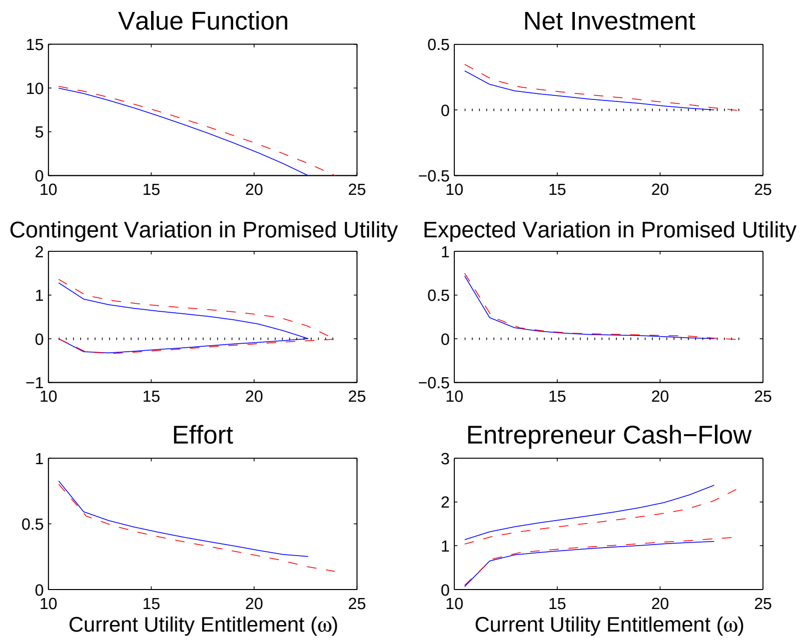

Figure 12: Policy Functions. Solid: benchmark. Dashed: $\theta_h = 1.45$, $\theta_l = 0.45$.

should be independent of the number of technicians already employed. However, there are also instances in which the same assumption is not as appealing. This is the case, for example, when the decision involves the purchase of a new software for budget control. The idea is that the larger the firm, the wider the set of software systems with which a new one has to dialogue. Therefore the associated cost should be higher.[10]

In this section, we allow for a more general specification of the utility cost of effort. The entrepreneur's preferences are now given by $u(c) - ak^{\varphi}$. The marginal cost of exerting effort will be increasing or decreasing with size, depending on the sign of the parameter φ.

Recall that in the benchmark scenario, everything else equal, recommended effort is increasing in the level of capital. This is the case, because capital only impacts (positively) the marginal revenue effect of increasing the utility spread s. Introducing the interaction of capital and effort in the entrepreneur's utility function has the potential of enriching the range of the model's empirical implications. When $\varphi > 0$, the marginal utility cost of effort is increasing in capital. In turn, that will make effort less attractive for the entrepreneur. The opposite will happen for $\varphi < 0$. This intuition is confirmed by the expression for optimal effort. Given the utility spread s

[10]See Baker and Hall (2004) for an insightful digression on this issue.

and the level of capital k, the entrepreneur will choose $a^*(k, \omega) = \log(s) - \varphi \log(k)$. The first–order condition for s now reads

$$[\theta_h - \theta_l]k^\alpha + \beta[v(\omega_h) - v(\omega_l)] = [c(u_h) - c(u_l)] + (s - k^\varphi)\left[c'(u_h) - c'(u_l)\right], \quad (21)$$

with

$$\omega_h = \frac{1}{\beta}\left[\omega + k^\varphi(1 - \log(k^\varphi)) + k^\varphi \log(s) - u_h\right],$$
$$\omega_l = \frac{1}{\beta}\left[\omega + k^\varphi(1 - \log(k^\varphi)) + k^\varphi \log(s) - s - u_l\right].$$

The comparative statics of s with respect to k is now more involved and nothing ensures that utility spread and effort are increasing in capital. In fact, we will show an example in which both are lower, the smaller the capital stock.

Letting the utility cost of effort depend on size impacts all features of the constrained–efficient contract. By the envelope condition, we now have that

$$\frac{\partial v(k,\omega)}{\partial k} = \left[\left(1 - \frac{k^\varphi}{s}\right)\theta_h + \frac{k^\varphi}{s}\theta_l\right]\alpha k^{\alpha-1} + (1-\delta) - \left[\left(1 - \frac{k^\varphi}{s}\right)c'(u_h) + \frac{k^\varphi}{s}c'(u_l)\right]\varphi a^* k^{\varphi-1},$$

Compare this expression with (20), its counterpart in the benchmark scenario. Recall that $a \geq 0$ implies $s \geq k^\varphi$. Increasing capital has still a positive effect on revenues, given by the first two addenda. The last addendum reflects the effect on the costs of delivering compensation. Its sign will be that of the parameter φ. When $\varphi < 0$, i.e. when the marginal utility cost of effort is decreasing in size, the marginal impact on costs will be negative. It follows that we will have $\frac{\partial v(k,\omega)}{\partial k} \geq 0$ as in Section 3. However, when $\varphi \geq 0$, increasing capital will adversely affect that cost. In principle, there may exist levels of φ large enough that the marginal gain of adding capital is indeed negative.

Figure 13 illustrates how the system's dynamics changes when we consider strictly positive values for φ. It shows averages for a large number of 80–period long simulations, in the cases of $\varphi = 0,\ 0.1$, and 0.3, respectively. As expected, a larger φ means slower capital accumulation and lower payoffs for the investor. With the exception of the first few periods, the entrepreneur's payoff and its effort exertion are also lower.[11]

6 Auto–Correlated Shocks

The model we have analyzed so far assumes that a successful performance of the entrepreneur only affects the probability distribution of the shock θ in the same period.

[11] For $\varphi = 0.3$, effort is actually increasing in φ. This is the only novel qualitative feature resulting from $\varphi > 0$. The reason is that for $k < 1$, the marginal utility cost of increasing effort is decreasing in φ.

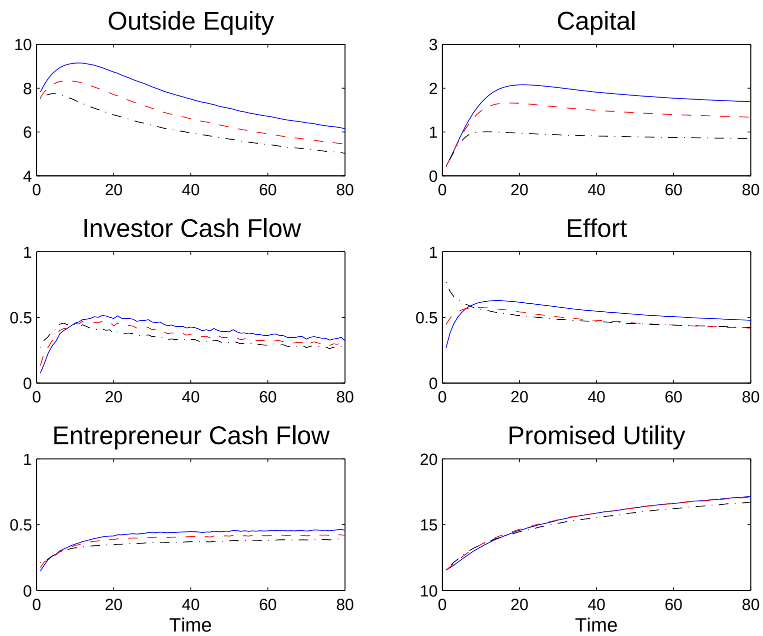

Figure 13: Average Dynamics. Dashed: $\varphi = 0.1$. Dash–Dot: $\varphi = 0.3$

Although it is shared by most of the literature on dynamic hidden action models,[12] one may find this assumption to be particularly removed from reality. In this section we address this concern by assuming that a successful outcome also alters the probability distribution in the future. We posit that next period's distribution conditional on success in the current period, $G(\theta'|a', \theta_h)$, stochastically dominates the same distribution conditional on failure, i.e. $G(\theta'|a', \theta_l)$, for all a'. In the numerical implementation, we assume that $prob(\theta' = \theta_h|\theta = \theta_i) = 1 - e^{-\psi_i a}$, with $\psi_h > \psi_l$.[13]

Figure 14 illustrates value and policy functions along the ω dimension, for given capital stock. Solid lines refer to $\psi_h = 1.4$, while dashed line refer to $\psi_l = 0.8$. All the other parameters are as described in Table 1. As expected, effort, investment, and outside equity are lower conditional on $\psi = \psi_l$. Perhaps more interestingly, cash–flows accruing to he entrepreneur both in the current and future periods are higher, conditional on $\psi = \psi_l$. That is, $c_i(k, \omega, \psi_l) > c_i(k, \omega, \psi_h)$ and $\omega_i(k, \omega, \psi_l) > \omega_i(k, \omega, \psi_h)$ for all ω and for $i = h, l$. The result that $c_l(k, \omega, \psi_l) > c_l(k, \omega, \psi_h)$ is a direct consequence of two facts: (i) the limited liability constraint binds in the low

[12]For example, see Spear and Srivastava (1987) and Wang (1997).

[13]Notice that our modeling choice is different from that of Fernandes and Phelan (2000). In their case, next period's distribution depends on current effort (which is private information), rather than on current realization (public information).

state of nature and (ii) investment is lower conditional on $\psi = \psi_l$.

Contingent variations in promised utility are lower when $\psi = \psi_h$ because of the negative complementarity between the two state variables, which we have described in Section 3. Since the marginal gain of raising ω is decreasing in k and the choice of k' is higher when $\psi = \psi_h$, the optimal choices of ω_h and ω_l will have to be lower.

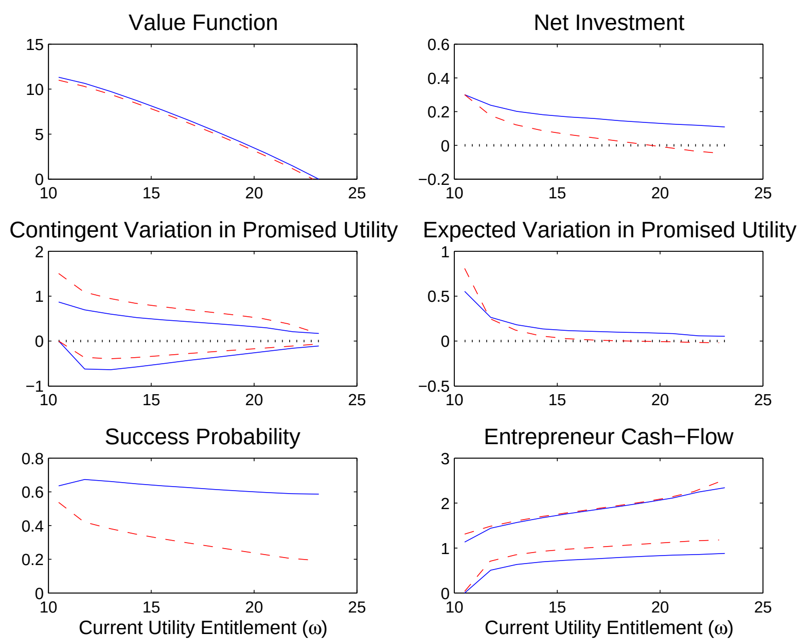

Figure 14: Policy Functions. Case with State–Dependent Shock Distribution

Figure 15 illustrates the effects of persistence on the dynamics of all relevant variables. Similarly to the cases considered above, the value of outside equity converges to zero and the entrepreneur ends up controlling all cash–flow rights. However, the ergodic set for the state variables is now a non–degenerate subset of Ω's upper contour. As the investor's payoff settles down to its long–run value, the other variables are time–varying. Since a good shock today implies a more favorable probability distribution tomorrow, promised utility, investment, and recommended effort are all higher following a good shock.

7 Conclusion

In this paper we have characterized the firm dynamics implied by constrained–efficient contracts between a risk–neutral investor and a risk–averse entrepreneur under the assumption that the latter's effort is not publicly observable.

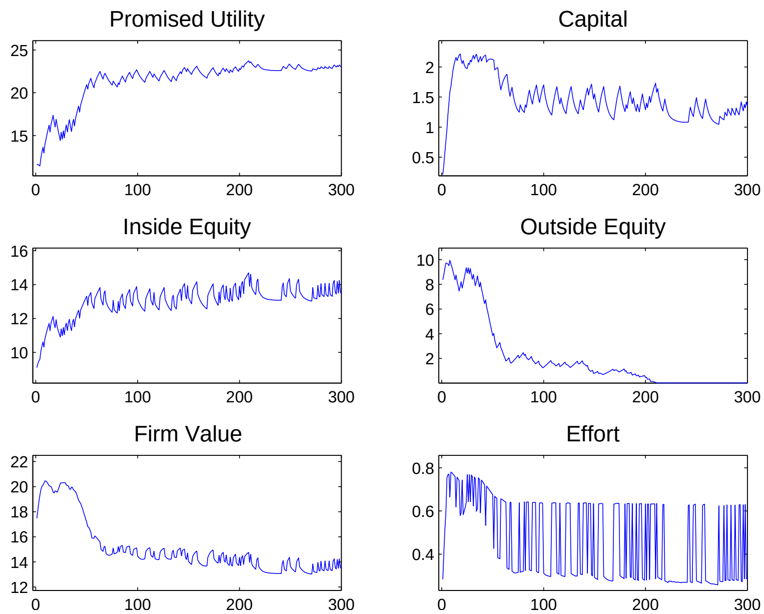

Figure 15: Sample Path. Case with State–Dependent Shock Distribution

A robust feature of the model is the sub–modularity of the value function. That is, the marginal gain from investing declines with the level of promised utility. This happens because the cost of incentive provision is increasing in ω. In turn, this means that the higher ω, the lower constrained–efficient effort and the probability of success, and therefore the lower the return to capital accumulation.

Consistent with a widely cited result by Rogerson (1985b), when the entrepreneur's relative risk aversion coefficient is less than 1 and the two agents are equally impatient, on average ω and the value of inside equity grow over time. Because its marginal product of capital is relatively high, a small firm will see its capital grow over time. However, because of the sub–modularity property, the rise in the value of inside equity will imply a drop in the return to investment, which eventually will lead to a decline in the capital stock. In this sense, our theory provides a rationale for firm's decline. This feature distinguishes our model from others in which repeated moral hazard shapes firm dynamics. In those, firm size never overshoots its long–run level.

Interestingly, the constrained–efficient contract prescribes that in the long run the entrepreneur becomes the only claimant to the firm's cash flows.

A key mechanism in our theory is that providing incentives to exert effort becomes costlier as the manager increases her stake in the firm. We believe that it would be interesting to extend our framework by allowing for the possibility of termination.

That is, by empowering the investor to liquidate the entrepreneur and hire someone else to run operations. The insights provided by Spear and Wang (2005) may prove useful in carrying out this task.

A Proofs and Lemmas

The results that follow are shown under the assumption that the entrepreneur's preferences are represented by the utility function $u(c_t, m(a_t)l(k_t))$, which is assumed to be bounded, strictly increasing and strictly concave in c_t, and strictly decreasing in $m(a_t)l(k_t)$. We assume that $m(a_t)$ is increasing and convex in a_t and that $l : \Re_+ \to \Re_+$. We also assume that $u(c_t, m(a_t)l(k_t))$ is additively separable in the arguments c_t and $m(a_t)l(k_t)$.

Proposition 1 *$\Phi(k, w)$ is compact, $\forall(k, \omega) \in \Omega$.*

Proof. Fix the pair (k, ω). We already know that $\Phi(k, w)$ is bounded. It is left to prove that it is also closed. Let $\{V_n\} \subseteq \Phi(k, w)$, where $V_n \to V_\infty$ when $n \to \infty$. We need to show that $V_\infty \in \Phi(k, w)$. In words, we need to demonstrate that there exists a contract σ_∞ that satisfies (1), (2), (3), $\omega(\sigma_\infty|h^0) = w$, and $v(\sigma_\infty|h^0) = V_\infty$. Now we will construct such an optimal contract σ_∞. By the definition of $\Phi(k, w)$, there exists a sequence of contracts $\{\sigma_n\} = \{a_t^n(h^{t-1}), c_t^n(h^t)\}$ and capital $\{k_{t+1}^n(h^{t-1})\}$, where the constraints (1), (2), and (3), $\omega(\sigma_n|h^0) = w$ are satisfied for every n. Therefore

$$V_\infty = \lim_{n\to\infty} \sum_{t=1}^{\infty} \beta^{t-1} \int [\theta_t f(k_t) - c_t^n(h^t) - k_{t+1}^n(h^{t-1}) + (1-\delta)k_t] g(\theta_t | a_t^n(h^{t-1})) dh^t$$

For $t = 1$, notice that $\{a_1^n(h^0), c_1^n(h^1)\}$ and $\{k_2^n(h^0)\}$ are finite collections of bounded sequences. Therefore, there exist collections of subsequences $\{a_1^{n_q}(h^0), c_1^{n_q}(h^1)\}$ and $\{k_2^{n_q}(h^0)\}$ such that

$$\lim_{n_q\to\infty} a_1^{n_q}(h^0) = a_1^\infty(h^0), \lim_{n_q\to\infty} c_1^{n_q}(h^1) = c_1^\infty(h^1), \text{ and } \lim_{n_q\to\infty} k_2^{n_q}(h^0) = k_2^\infty(h^0).$$

We now consider $t = 2$. Notice that $\{a_2^n(h^1), c_2^n(h^2)\}$ and $\{k_3^n(h^1)\}$ are finite collections of bounded sequences, and we can define $\{a_2^\infty(h^1), c_2^\infty(h^2)\}$ and $\{k_3^\infty(h^1\}$ similarly as we did for $t = 1$. If we iterate this procedure for $t = 3, 4, ...$, and let $\sigma_\infty = \{a_t^\infty(h^{t-1}), c_t^\infty(h^t)\}$ along with $k = \{k_{t+1}^\infty(h^{t-1})\}$, then it is easy to verify that the constructed contract σ_∞ is what we desired for. ■

Proposition 2 *$v^*(k, \omega) = T(v^*)(k, \omega)$.*

Proof. Fix ω, the lifetime discounted utility ensured by the optimal contract to the agent, and k, the optimal capital level of the firm. First, we show that

$T(v^*)(k,\omega) \le v^*(k,\omega)$. This inequality is true if there exists a feasible and incentive compatible contract σ such that $\omega(\sigma|h^0) = \omega$ and $v(\sigma|h^0) = T(v^*)(k,\omega)$. The desired contract σ can be constructed in the following way. Let $a(k,\omega)$, $c(\theta,k,\omega)$, $k'(k,\omega)$, and $\omega'(\theta,k,\omega)$ denote the solution of the maximization problem associated with the definition of $T(v^*)(k,\omega)$. Now, let $a_1(h^0) = a(k,\omega)$, $c_1(h^1) = c(\theta_1,k,\omega)$, and $k_2(h^0) = k'(k,\omega)$, $\forall h^1$ along with $k_1 = k_0$. For the realization of θ in $t = 1$, denoted θ_1 for the purpose of this proof, there exists a feasible and incentive compatible contract σ_{θ_1} that ensures a level of expected discounted utility $\omega'(\theta_1,k,\omega)$ to the agent, and $v^*(k'(k,\omega),\omega'(\theta_1,k,\omega))$ to the principal. Thus, we can say that $\sigma|h^1 = \sigma_{\theta_1}$, $\forall h^1$. It is obvious that the constructed contract σ is what is desired.

We now need to show that $v^*(k,\omega) \le T(v^*)(k,\omega)$. Let σ^* be an optimal contract that ensures a level of expected discounted utility of ω to the agent, given k. In consequence, we can say that

$$v^*(k,\omega) = v(\sigma^*|h^0),$$

or

$$v^*(k,\omega) = \int_\Theta \left\{\theta_1 f(k_1) - c_1^*(\theta_1) - k_2^*(h^0) + (1-\delta)k_1 + \beta v^*(k_2^*(h^0),\sigma^*|h^1)\right\} g(\theta|a_1^*(h^0))d\theta,$$

or, finally,

$$v^*(k,\omega) \le T(v^*)(k,\omega),$$

where the last inequality is obtained by letting $a(k,\omega) = a^*(h^0)$, $c(\theta,k,\omega) = c_1^*(\theta_1)$, $\omega'(\theta,k,\omega) = \omega'^*|h^0)$ along with $k'(k,\omega) = k_2^*(h^0)$ and $k_1 = k_0$. This solution satisfies the constraints (4), (5), (6), (7), and (8). ■

A.1 The APS Algorithm

In this Appendix, we show how results from Abreu, Pierce, and Stacchetti (1990) can be adapted to our environment in order to define an iterative algorithm that allows for the numerical approximation of the set Ω.

For any arbitrary $\Sigma \in \Re^2$, define the operator B as

$$B(\Sigma) = \{(k,w) \mid \exists \left\{a, k', c(\theta), \omega'(\theta)\right\} \text{ s.t. } (4), (5), (6), (7), \text{ and } (k', w\omega'(\theta)) \in \Sigma, \forall\theta\}$$

Notice that the operator B is monotone, i.e. $\Sigma_1 \subseteq \Sigma_2$ implies that $B(\Sigma_1) \subseteq B(\Sigma_2)$. Following Abreu, Pierce, and Stacchetti (1990), we say that Σ is self–generating if $\Sigma \subseteq B(\Sigma)$.

Proposition 3 *(a) Ω is self–generating. (b) If Σ is self–generating, then $B(\Sigma) \subseteq \Omega$.*

Proof. To prove (a), let $(k, \omega) \in \Omega$. There exists a contract $\sigma = \{a_t(h^{t-1}), c_t(h^t)\}$ and a sequence $\{k_{t+1}(h^{t-1})\}$ which satisfy the constraints (1), (2), (3), and $\omega(\sigma|h^0) = \omega$. We now say that

$$a(k,\omega) = a_1(h^0); \quad k'(k,\omega) = k_2(h^0); \;\; c(\theta,k,\omega) = c_1(\{\theta\}),\ \forall\theta; \;\; \omega'(\theta,k,\omega) = \omega_2(\sigma|\{\theta\}),\ \forall\theta.$$

It is obvious that $\{a(k,\omega), c(\theta,k,\omega), k'(k,\omega), \omega'(\theta,k,\omega)\}$, defined above, satisfies the constraints (4), (5), (6), (7), and (8). Therefore, $(k,\omega) \in B(\Omega)$, which demonstrates that (a).

To prove (b), let Σ be self–generating, and let $(k,\omega)_{h^0} \in B(\Sigma)$. We have to construct a contract $\sigma = \{a_t(h^{t-1}), c_t(h^t)\}$ and a sequence $k_{t+1}(h^{t-1}) = k_{h^0}$ that satisfy the constraints (1), (2), (3), and $\omega(\sigma|h^0) = \omega_{h^0}$. We construct such a contract recursively. First, there exist $\{a(k_{h^0},\omega_{h^0}), c(\theta,k_{h^0},\omega_{h^0}), k'(k_{h^0},\omega\_h^0), \omega'(\theta,k_{h^0},\omega_{h^0})\}$ that satisfies (5), (7), (8), and

$$\int_\Theta \left\{u(c(\theta,k_{h^0},\omega_{h^0}), m(a(k_{h^0},\omega_{h^0}))l(k)) + \beta\omega'(\theta,k_{h^0},\omega_{h^0})\right\} g(\theta|a(k_{h^0}, w\omega_{h^0}))d\theta = \omega_{h^0},$$
$$0 \le c(\theta,k_{h^0},\omega_{h^0}) \le \theta f(k) - k_{h^0} + (1-\delta)k.$$

For $t = 1$, let $a_1(h^0) = a(k_{h^0},\omega_{h^0})$ and $c1(h^1) = c(\theta_1,k_{h^0},\omega_{h^0})$, $\forall h^1$. Also, let $k'_{h^0} = k_{h^1} = k'(k_{h^0},\omega_{h^0})$ and $\omega_{h^1} = \omega'(\theta,k_{h^0},\omega_{h^0})$, $\forall h^1$. Notice that $(k_{h^1},\omega_{h^1}) \in \Sigma \in B(\Sigma)$ implies the existence of $\{a(k_{h^1},\omega_{h^1}), c(\theta,k_{h^1},\omega_{h^1}), k'(k_{h^1},\omega_{h^1}), \omega'(\theta,k_{h^1},\omega_{h^1})\}$ that satisfies (5), (7), (8), and

$$\int_\Theta \left\{u(c(\theta,k_{h^1},\omega_{h^1}), m(a(k_{h^1},\omega_{h^1}))l(k)) + \beta\omega'(\theta,k_{h^1},\omega_{h^1})\right\} g(\theta|a(k_{h^1},\omega_{h^1}))d\theta = \omega_{h^1},$$
$$0 \le c(\theta,k_{h^1},\omega_{h^1}) \le \theta f(k) - k_{h^1} + (1-\delta)k.$$

We can iterate for $t = 2,3,4,...$ to construct the complete profile σ. We can then observe that, by construction, for any arbitrary $t \ge 0$ and h^t,

$$\omega(\sigma|h^t) - \omega_{h^t} = \int_\Theta \beta[\omega(\sigma|h^{t*1}) - \omega_{h^{t*1}}]g(\theta|a(k_{h^t},\omega_{h^t}))d\theta_{t*1}$$

Since $0 < \beta < 1$ and the utilities are bounded, the above equation implies that

$$\omega(\sigma|h^t) = \omega_{h^t} \quad \forall\, t \ge 0 \;\text{ and }\; \forall\, h^t.$$

Hence, the contract that we have constructed is what is desired. ■

Now define a closed and bounded set Ω_0 such that $\Omega \subseteq \Omega_0$. Recall that $k \in [\underline{k}, \bar{k}] \in \Re_+$. Let

$$\Omega_0 = \{(k, \omega) \mid k \in [\underline{k}, \bar{k}] \text{ and } \omega \in [\underline{\omega}, \bar{\omega}(k)]\},$$

where $\underline{\omega}$ is set to an arbitrarily small positive number. We do not need to assume that $u(\cdot)$ is positive–valued, since a positive monotone transformation can achieve this if needed. Also, let $\overline{\omega}(k) \equiv \frac{u(c,m(\underline{a})l(k))}{1-\beta}$, where $\underline{a} = \min\{A\}$ and $c = \overline{\theta} f(k)$, $\overline{\theta} = \max\{\Theta\}$.

Proposition 4 shows that (i) the set Ω is a fixed point of the operator B and that (ii) the sequence constructed by iterating on B starting with Ω_0 converges to the set Ω.

Proposition 4 *(a)* $\Omega = B(\Omega)$. *(b) Let* $X_0 = \Omega_0$*, and let* $X_{n+1} = B(X_n)$*, for* $n = 0, 1, 2, ...$ *Then,* $\lim\limits_{n \to \infty} X_n = \Omega$.

Proof. Part (a) is obvious. To show part (b), we will first show that the sequence $\{X_n\}$ is convergent. Obviously, $B(X_0) \subseteq X_0$. Next, we operate B on both sides of this expression and obtain $X_{n+1} = B(X_n) \subseteq X_n$, $\forall n$, because B is monotone increasing. Hence, $\{X_n\}$ is a bounded and monotone decreasing set sequence with $X_\infty = \lim\limits_{n \to \infty} X_n = \bigcap\limits_{n=0}^{\infty} X_n$. Now, we show that $\Omega \subseteq X_\infty$. Given that $\Omega \subseteq X_0$, the monotonicity property of B ensures that $B(\Omega) \subseteq B(X_0)$. However, it must be true that $\Omega = B(\Omega)$, by part (a), and $B(X_0) = X_1$, by construction. Then, $\Omega \subseteq X_1$. By iteration we obtain $\Omega \subseteq X_n$, $\forall n \geq 0$, and consequently, $\Omega \subseteq X_\infty$. Now, we demonstrate that $X_\infty \subseteq \Omega$. Given the properties of the sequence $\{X_n\}$, we have that $B(X_\infty) = X_\infty$. Hence, X_∞ is self–generating, and $X_\infty = B(X_\infty) \subseteq \Omega$. ■

B Algorithm

In this section we provide a brief description of the algorithm that was used to compute a numerical approximation to the value function $v(k, \omega)$. Given that the equilibrium value set Ω is not square, it is not efficient to approximate the value function by means of bi–dimensional splines. For this reason, we will restrict the choice of capital to a finite number of levels and approximate the value function on the ω dimension by means of cubic splines.

We start by defining a fine grid for the capital stock. Denote it as $\mathcal{K} \equiv \{k_j\}_{j=1}^{n_k}$ and let the related set of indexes be $\mathcal{J} \equiv \{j\}_{j=1}^{n_k}$. The upper bound of $\mathcal{K}$ must be chosen in such a way that the corresponding net investment will be negative for all ω. For this to be the case, it is sufficient to set it equal to the efficient capital stock when

$\theta = \theta_h$ with probability 1 in all periods. That is, we let $k_{n_k} = \left(\frac{\alpha\theta_h}{\delta}\right)^{\frac{1}{1-\alpha}}$. The next task consists in approximating the equilibrium value set of the transformed problem.

B.1 Approximation of the Equilibrium Value Set

From the analysis conducted in Section 2, it follows that for every $j \in \mathcal{J}$, the set of feasible and incentive compatible values will be given by an interval $[\underline{\omega}, \bar{\omega}_j] \in \Re_+$. This means that our task reduces to approximate the mapping $\Omega : \mathcal{K} \to \Re_+$ which is given by the sequence $\{\bar{\omega}_j\}_{j\in\mathcal{J}}$. The mapping Ω can be shown to be increasing and strictly concave.

Following Abreu, Pierce, and Stacchetti (1990), we start by defining an initial guess $\Omega_0 = \{\bar{\omega}_{0j}\}_{j\in\mathcal{J}}$. We impose that Ω_0 is weakly increasing, weakly concave, and such that $\bar{\omega}_{0j} \geq \bar{\omega}_j$ for all j. These requirements are satisfied by letting $\bar{\omega}_{0j} = \frac{u(\theta_h k_{n_k}^\alpha - \delta k_{n_k})}{1-\beta}$. Then, for every $j, q \in \mathcal{J}$ such that $\theta_l k_j^\alpha + k_j(1-\delta) - k_q \geq 0$, we compute

$$b_{jq} \equiv \max_{a,\{u_i,\omega_i\}_{i=h,l}} \left(1 - e^{-a}\right)[u_h + \beta\omega_h] + e^{-a}[u_l + \beta\omega_l] - a \tag{22}$$
$$\text{s.t. } 0 \leq u_i \leq u(\theta_i k_j^\alpha + k_j(1-\delta) - k_q),$$
$$\underline{\omega} \leq \omega_i \leq \bar{\omega}_{nq}$$

and

$$\bar{\omega}_{n+1,j} \equiv \max_j \{b_{jq}\}. \tag{23}$$

The operator defined by (22)–(23) generates a sequence $\{\Omega_n\}$ that converges to Ω. Our approximation will be Ω_m such that $||\Omega_m - \Omega_{m-1}||_\infty < 10.0^{-8}$.

Notice that, conditional on effort being zero, the above optimization problem simplifies to

$$\max_{u_l,\omega_l} \; u_l + \beta\omega_l$$
$$\text{s.t. } 0 \leq u_l \leq u(\theta_l k^\alpha + k(1-\delta) - k_j),$$
$$\underline{\omega} \leq \omega_l \leq \bar{\omega}_{nj}.$$

Obviously the solution calls for $u_l = u(\theta_l k_j^\alpha + k_j(1-\delta) - k_q)$ and $\omega_l = \bar{\omega}_{nq}$. Alternatively, when effort is strictly positive, $a = \log(s)$, where $s \equiv (u_h + \beta\omega_h - u_l - \beta\omega_l)$. The optimization problem then becomes

$$\max_{s,u_h,\omega_h} u_h + \beta\omega_h - 1 - \log(s)$$
$$\text{s.t. } 0 \leq u_i \leq u(\theta_i k_j^\alpha + k_j(1-\delta) - k_q),$$
$$\underline{\omega} \leq \omega_i \leq \bar{\omega}_{nq}.$$

In this case the solution calls for $u_i = u(\theta_i k_j^\alpha + k_j(1-\delta) - k_q)$ and $\omega_i = \bar{\omega}_{nq}$.

B.2 Approximation of the Value Function

For every $j \in \mathcal{J}$, we define a coarse grid $\mathcal{Z}_j = \{\omega_{jz}\}_{z=1}^{n_\omega}$ over the interval $[\underline{\omega}, \bar{\omega}_j]$. We also define an initial guess for the value function: $v_{0j} : \mathcal{Z}_j \to \Re_+$. For all other $\omega \in [\underline{\omega}, \bar{\omega}_j]$, the guess is approximated by a cubic spline which we denote as $v_{0j}(\omega)$. We impose that $v_{0j}(\omega)$ is decreasing and concave in ω for all $j \in \mathcal{J}$ and that the function is increasing and concave in capital. Then, for all z and every $j, q \in \mathcal{J}$ such that $\theta_l k_j^\alpha + k_j(1-\delta) - k_q \geq 0$, we compute

$$\begin{aligned} d_{jzq} \equiv \max_{a^*, \{u_i, \omega_i,\}_{i=h,l}} & (1-e^{-a^*})[\theta_h k^\alpha - c(u_h) + \beta v_{nq}(\omega_h)] + e^{-a^*}[\theta_l k^\alpha - c(u_l) + \beta v_{nq}(\omega_l)] \\ & + k_j(1-\delta) - k_q, \end{aligned} \tag{24}$$

$$\begin{aligned} \text{s.t. } & (1-e^{-a^*})[u_h + \beta\omega_h] + e^{-a^*}[u_l + \beta\omega_l] - a^* = \omega_{jz}, \\ & a^* = \arg\max\, (1-e^{-a})[u_h + \beta\omega_h] + e^{-a}[u_l + \beta\omega_l] - a, \\ & 0 \leq u_i \leq u(\theta_i k_j^\alpha + k_j(1-\delta) - k_q) \;\; \forall\, i = h, l, \\ & a^* \geq 0, \\ & \underline{\omega} \leq \omega_i \leq \bar{\omega}_q \;\; \forall\, i = h, l, \end{aligned}$$

and

$$v_{n+1,jz} \equiv \max_q \{d_{jzq}\}. \tag{25}$$

The operator defined by (24)–(25) generates a sequence $\mathcal{V}_n \equiv \{v_{nj}\}_{j\in\mathcal{J}}$. Our approximation of the value function on the grid will be $\mathcal{V}_m$ such that $||\mathcal{V}_m - \mathcal{V}_{m-1}||_\infty < 10.0^{-8}$. Notice that when recommended effort is positive, the above optimization problem simplifies to

$$\begin{aligned} \max_{s, \{u_i\}_{i=h,l}} & \left(1 - \frac{1}{s}\right)[\theta_h k^\alpha - c(u_h) + \beta v_{nq}(\omega_h)] + \frac{1}{s}[\theta_l k^\alpha - c(u_l) + \beta v_{nq}(\omega_l)] \\ & + k_j(1-\delta) - k_q, \end{aligned}$$

$$\begin{aligned} \text{s.t. } & \omega_h = [\omega_{jz} + 1 + \log(s) - u_h]/\beta, \\ & \omega_l = [\omega_{jz} + 1 + \log(s) - s - u_l]/\beta, \\ & 0 \leq u_i \leq u(\theta_i k_j^\alpha + k_j(1-\delta) - k_q) \;\; \forall\, i = h, l, \\ & \underline{\omega} \leq \omega_i \leq \bar{\omega}_q \;\; \forall\, i = h, l. \end{aligned}$$

References

Abreu, D., D. Pierce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1040–1063.

Albuquerque, R., and H. Hopenhayn (2004): "Optimal Lending Contracts and Firm Dynamics," *Review of Economic Studies*, 71, 285–315.

Atkeson, A., and R. Lucas (1995): "Efficiency and Equality in a Simple Model of Efficient Unemployment Insurance," *Journal of Economic Theory*, 66, 64–98.

Baker, G., and B. Hall (2004): "CEO Incentives and Firm Size," *Journal of Labor Economics*, 22(4), 767–798.

Bohacek, R. (2005): "Capital Accumulation in Private Information Economies," *Topics in Macroeconomics*, 5(1), Article 24.

Brusco, S., and E. Ropero (2007): "Financing Constraints and Firm Dynamics with Durable Capital," Working paper, SUNY at Stony Brook.

Choi, S., R. Fisman, D. Gale, and S. Kariv (2007): "Consistency and Heterogeneity of Individual Behavior under Uncertainty," *American Economic Review*, 97(5), 1921–1938.

Clementi, G. L., T. Cooley, and C. Wang (2006): "Stock Grants as a Commitment Device," *Journal of Economic Dynamics and Control*, 30(11), 2191–2216.

Clementi, G. L., and H. Hopenhayn (2006): "A Theory Financing Constraints and Firm Dynamics," *Quarterly Journal of Economics*, 121(1), 229–265.

Cooley, T., R. Marimon, and V. Quadrini (2004): "Aggregate Consequences of Limited Contract Enforceability," *Journal of Political Economy*, 112(4), 817–847.

DeMarzo, P., M. Fishman, Z. He, and N. Wang (2008): "Dynamic Agency and the q Theory of Investment," Working paper, Graduate School of Business, Stanford University.

Espino, E. (2005): "On Ramsey's Conjecture: Efficient Allocations in the Neoclassical Growth Model With Private Information," *Journal of Economic Theory*, 121, 192–213.

FERNANDES, A., AND C. PHELAN (2000): "A Recursive Formulation for Repeated Agency with History Dependence," *Journal of Economic Theory*, 91, 223–47.

GOEREE, J., C. HOLT, AND T. PALFREY (2002): "Quantal Response Equilibrium and Overbidding in Private–Value Auctions," *Journal of Economic Theory*, 104(1), 247–272.

HOLMSTROM, B. (1979): "Moral Hazard and Observability," *Bell Journal of Economics*, 10, 74–91.

HOLT, C., AND S. LAURY (2002): "Risk Aversion and Incentive Effects," *American Economic Review*, 92(5), 1644–1655.

KHAN, A., AND B. RAVIKUMAR (2001): "Growth and Risk–Sharing with Private Information," *Journal of Monetary Economics*, 47, 499–521.

LUCAS, R. E. J. (1978): "On The Size Distribution of Business Firms," *Bell Journal of Economics*, 9(2), 508–523.

MARCET, A., AND R. MARIMON (1992): "Communication, Commitment, and Growth," *Journal of Economic Theory*, 58, 219–249.

QUADRINI, V. (2003): "Investment and Liquidation in Renegotiation–Proof Contracts with Moral Hazard," *Journal of Monetary Economics*, 51, 713–751.

ROGERSON, W. (1985a): "The First–Order Approach to Principal–Agent Problems," *Econometrica*, 53, 1357–1968.

——— (1985b): "Repeated Moral Hazard," *Econometrica*, 53, 69–76.

SPEAR, S., AND S. SRIVASTAVA (1987): "On Repeated Moral Hazard with Discounting," *Review of Economic Studies*, 54, 599–617.

SPEAR, S., AND C. WANG (2005): "When to Fire a CEO: Optimal Termination in Dynamic Contracts," *Journal of Economic Theory*, 120, 239–256.

WANG, C. (1997): "Incentives, CEO Compensation, and Shareholder Wealth in a Dynamic Agency Model," *Journal of Economic Theory*, 76, 72–105.