When to Fire a CEO: Transience and Ergodicity in Long-term Contracts

Stephen E. Spear *
GSIA, Carnegie Mellon University
Pittsburgh, PA 15213
ss1f@andrew.cmu.edu

Cheng Wang*
Department of Economics
Iowa State University
Ames, IA 50011
chewang@iastate.edu
and
GSIA, Carnegie Mellon University
Pittsburgh, PA 15213
chewang@andrew.cmu.edu

First draft: June, 2000; Current version: March 2002
GSIA Working Paper 2002-E5

Abstract

The repeated agency model has been widely applied to a number of interesting and important problems in economics, though in many instances, the fact that the standard model generates transient dynamics limits the usefulness of the results obtained from the model for the simple reason that purely transient dynamc phenomena are empirically irrelevant since they cannot be systematically observed and studied. In this paper, we show to embed the standard long-term contracting model in an economic environment in which the inherently transient dynamics of the model are transformed into dyanmics which are stationary and ergodic. We then use the model to study CEO termination and issues of corporate takeover.

*This paper was initially circulated under the title "When to Terminate a Long-Term Contract?" We thank seminar participants at the Arizona State University, the University of Western Ontario, and the 2000 SED meeting in Costa Rica for helpful discussions

1 Introduction

The repeated agency model first introduced by Green [3] and Spear and Srivastava [9] has proven to be a workhorse in areas of economics as diverse as the study of labor markets, markets for insurance, financial markets, macroeconomics, and monetary economics. Given the wide application of the model, it is somewhat surprising that economists haven't addressed a well-known flaw in the model, identified independently by Green [3], Phelan and Townsend [6], and Thomas and Worrall [12], who showed that all of the trajectories generated by the dynamics of the basic model are transient. In the absence of external constraints on the possible payoffs in the long-term contract, one party or the other ends up owning everything. Essentially, the compensation scheme generates incentives (in the form of current and deferred payoffs) as a random walk. Models which generate transient dynamics pose obvious problems in terms of empirical verification of the model's validity, which is why we characterize this anomoly in the model as a flaw. Our purpose in this paper, then, is to reexamine the basic model in an attempt to show how it can be embedded in an ergodic framework which is amenable to empirical verification.

While there have been some efforts in the literature to address the problem of transient dynamics in the repeated agency model, the fixes imposed on the model have tended to be somewhat *ad hoc*.

Atkeson and Lucas [2], for example, provides a remedy for the transient dynamics problem by imposing a lower bound to the expected utility that the agent must receive from the beginning of each period on, and they find the solution to the optimal dynamic contracting problem implies a non-degenerate steady-state distribution. As the paper points out, though, the utility bound they impose is an additional constraint imposed upon the efficiency problem and is not derived from any additional efficiency considerations.

Phelan [5] argues that the transient dynamics are due to the assumption made in all existing models that agents make binding commitments to the terms of the long-term contract. He argues that if agents are free to leave a contract if it specifies too small a payoff (or, on the firm side, too great a promised compensation) at any point in time, then agents will recognize that any viable long-term contract must respect the bounds imposed by the possibility of termination, and hence, the contract will incorporate liability bounds as constraints on the long-term contract. In Phelan's model, this results in agents' signing contracts they never leave, and, as in Atkeson and Lucas, the dynamics of the contract are ergodic.

While Phelan's approach is less *ad hoc* in the sense of positing a economic rationale for imposing liability constraints on the contract, the approach is subject to the same criticism as the Atkeson-Lucas approach that the liability constraints are simply imposed on the efficiency problem rather than being derived from some additional efficiency considerations.

In this paper, we argue that the problem of transient equilibrium dynamics is due to a fundamental misinterpretation of the commitment requirements imposed on agents in most of the existing literature on dynamic contracting.

From an economic perspective (i.e. in terms of efficiency), the only important commitment requirements in the model are those the agents make to the incentive mechanism, rather than to the contractual relationship itself. It is the incentive compatible structure of the incentive scheme, together with an implied social mechanism for enforcing the promise-keeping constraints associated with deferred compensation that differentiate the long-term contract from the static one. The term of the contract is essentially irrelevant, although in the simplest case where agents bind themselves to each other forever, the recursive formulation of the model is particularly simple. But full long-term commitment to a contract isn't needed to make the recursive approach work; rather, full long-term commitment is an additional constraint imposed upon the agents' efficiency problem which can actually reduce agents' welfare. In practice, workers are laid off, CEOs are fired, and long-term contracts are often terminated according to prespecified mechanisms, either explicitly or implicitly. In the existing models of dynamic contracting, including Phelan [5], even if the agents can commit to the *ex post* terms of the contract (i.e., *ex post*, no party will renege because the continuation of the contract specifies a payoff that is smaller than that he could obtain from outside the contract), it may not be optimal for them to bind themselves to each other forever *ex ante*.

Our view is that a convex, stationary economic environment together with an infinite horizon will generally make it possible to implement models of dynamic contracting in recursive form. In addition to this, though, we believe a key missing ingredient in many previous applications of the repeated agency approach has been the specification of the relevant stationary *economic* environment of the model which will yield observable stationary dynamics for the model's equilibrium. For the standard version of the repeated agency model, we believe that the missing specification is that of an external labor market, and an optimal contract termination mechanism associated with it, and we will develop the model of this paper around that hypothesis. The assumption of an external market allows either party to the long-term contract to leave the contract, provided the direct compensation, and promise-keeping constraints of the contract are met in the terminal period. The dissolution of the contract then allows the two parties to return to the labor market to seek a new match. In this framework, the distinction between short-run (or static) contracts and long-run contracts is simply a matter of whether the parties to the contract agree to interact over more than a single period.

The model we develop here is a variant of the standard repeated agency model as it has been applied to models of CEO compensation (see, e.g. Wang [13]). We modify the model by assuming that there is an external labor market consisting of a pool of potential CEOs willing to work for the firm at a given starting value for future discounted expected utility. Under this simple closure of the model, two types of terminations emerge. Under one scenario, the CEO is fired after a sequence of bad outputs and she becomes too poor to be punished effectively. Under the second scenario, the CEO is forced out after a sequence of good outputs and she becomes too expensive to motivate. Here the well-known economic phenomenon of severance pay emerges (and, in the specific case of CEOs, we interpret this as the so-called "golden parachute" component of the CEO's contract). With further enrichment of the underlying economic environment of the model, we believe the framework will prove fruitful in studying such things as corporate takeover and the effects of agency costs on firm evolution.

We note that there can be other ways of closing the model so as to eliminate the transience of its outputs. Smith and Wang [8] examine long-term contracts between entrepreneurs and intermediaries in an overlapping generations model with long- (but finitely-) lived agents. In this model, even though the long-term contract generates transient dynamics, the finite lives of the agents in the model truncate the transient behavior, while the renewal of young agents in the model "restarts" the contracting dynamics, yielding an equilibrium process which is stationary. Similarly, Taub and Chade [11] show that in a model of multilateral risk-sharing in continuous time, the long-term contract generates features which resemble asset markets, although consumption under the simplest specification of the contract is transient. Taub and Chade show that the transience can be eliminated by modifying the contract to limit the insurance function of the contract sufficiently that the threat of punishment through losses will prevent agents who receive repeated good draws from defecting from the stationary risk-sharing contract. The procedure by which they limit the insurance fund has features that can be interpreted as equity holding by agents. Aiyagari and Alvarez [1] show that dynamic monitoring of long-term contracts (in the context of a model of unemployment insurance) can also yield non-transient equilibrium dynamics, although this conclusion is clouded somewhat by the fact that they impose bounds on consumption possibilities in the model, so that all dynamics unfold on a compact state space.

The paper is organized as follows. In Section 2 we lay out the basic model in a 2-period setting, and in Section 3, we study the incentives for termination in the context of this simple 2-period model. The 2-period setting is useful because it makes the various interactions of the information asymmetry and the workings of the incentive scheme particularly transparent, and makes clear when the parties to the contract will find it optimal to terminate their relationship. This in turn points up the kinds of stationary economic environments that are capable of supporting on-going contracting relationships with termination and replacement. In section 4, we analyze the fully dynamic version of the model.

2 The model

We start with a two period version of the model. A fully dynamic model will be studied in Section 4. Let time be denoted $t = 1, 2$. There is one consumption good in each period, and consumption takes place at the end of each period. There is one firm (principal) and there is a competitive supply of workers. (In the context of CEO compensation, we call the principal the shareholders of the firm and the agent its CEO.) The firm and the workers are all expected utility maximizers, and do not discount future utility. The firm's period utility function is $v(c) = c$, and the worker's period utility function is $H(c, a) = u(c) - \phi(a)$, where $c \in \mathbb{C}$ is consumption, $a \in \mathbb{A}$ is effort, and $\mathbb{C}$ and $\mathbb{A}$ are the sets of all possible consumption and effort levels.

Each period, the firm can employ one worker. The effort that the employed worker makes is observed only by the worker himself. That is, there is moral hazard. By choosing effort a_t in period t, the worker effectively chooses a probability distribution over the random output $\theta^t \in \Theta$, where θ^t is output in period-t and Θ is the set of all such outputs. We assume that $\Theta = \{\theta_1, \theta_2\}$ where $\theta_1 < \theta_2$. Let $\Pi_i(a) = Prob\{\theta^t = \theta_i | a_t = a\}$, $\theta \in \Theta_i$.

The model takes as given that, at the beginning of period 1, there is an ongoing relationship between the firm and an initial worker. This relationship specifies that the firm owes the worker a certain level of expected utility. At the end of period 1, the firm can fire the initial worker and replace him with a new worker. The process of replacement involves no costs to either the firm or the worker. If the worker is fired but not replaced, then production does not take place in that period and the firm has net payoff equal to $\eta = 0$.

For convenience we assume that once a worker leaves the firm, he can never be reemployed. We also assume that each worker in each period has a reservation utility equal to $w_0 = u(0) - \phi(0)$.

3 Optimal contracting

3.1 Contracting in the second period

We first address the problem of optimal contracting in the second period. Suppose that, at the beginning of period 2, the firm must deliver a promised expected utility equal to w (no more and no less) to a worker who may or may not have worked for the firm in period 1. Now given w, the firm can follow one of the two strategies. The firm can employ the worker in period 2, or fire the worker, compensate him with cash and replace him with a new worker. If the firm employs the worker in period 2, then the contract takes the form of (s_1, s_2), where s_i is the worker's compensation if his output is θ_i.

The firm's optimization problem can be formulated as follows. The firm's value function $V(w)$ is given by

$$V(w) = \max\{V_{\mathrm{r}}(w),\ V_{\mathrm{f}}(w)\}$$

where $V_{\mathrm{r}}(w)$ is the firm's value function conditional on employing the worker:

$$V_{\mathrm{r}}(w) = \max_{s_1, s_2} \sum \Pi_i(a)(\theta_i - s_i)$$

subject to $s_i \in \mathbb{C}$, the promise-keeping constraint $\sum \Pi_i(a) u(s_i) - \phi(a) = w$, and the incentive constraint $a \in \arg\max_{a' \in \mathbb{A}} \{\sum \Pi_i(a') u(s_i) - \phi(a'\}$, and $V_{\mathrm{f}}(w)$ is the firm's value function conditional on firing the worker and replacing him with a new worker:

$$V_{\mathrm{f}}(w) = \max\{V_{\mathrm{r}}(w) | w \geq w_0\} - u^{-1}[w + \phi(0)].$$

Note that when the worker is fired, he will not exert any effort in the second period, he is paid cash compensation $u^{-1}(w+\phi(0))$, and he is replaced by a new worker who is promised an expected utility w_* at which the firm's expected utility is maximized. That is, $V_\mathrm{r}(w_*) = \max\{V_\mathrm{r}(w) | w \geq w_0\}$. We assume that $V_\mathrm{r}(w_*) > 0$, so if the initial worker is fired, he is always replaced. Let (s_1^*, s_2^*) be the optimal contract that solves the maximization problem that defines $V_\mathrm{f}(w)$.

Let $\Phi \subseteq [0,\infty)$ be the set of expected utilities of the worker which can be supported by a feasible and incentive compatible one period contract, and decompose this set as $\Phi \equiv \Phi_\mathrm{r} \cup \Phi_\mathrm{f}$ where $\Phi_\mathrm{r} \equiv \{w \in \Phi: \ V_\mathrm{r}(w) \geq V_\mathrm{f}(w)\}$ and $\Phi_\mathrm{f} \equiv \{w \in \Phi: \ V_\mathrm{r}(w) < V_\mathrm{f}(w)\}$. That is, if the initial worker is promised utility w in period 2 and $w \in \Phi_\mathrm{r}$, then the initial worker is retained; otherwise, he is fired. Note here we are implicitly making the assumption that if the firm is indifferent between firing or retaining the worker, it retains the worker.

3.2 The optimal long-term contract

We now move on to define the optimal long-term contract. As stated earlier, we take as given that the firm is in an on-going relationship with the initial worker which dictates that the firm must deliver to the worker expected utility of at least x. We now describe the form of the contract that governs this relationship. The contract can potentially last for two periods, but may be terminated in one period. To keep the mathematical formulation of the long-term contract simple, we will assume Φ_r and Φ_f are both convex subsets of R, and that V_r and V_f are concave functions. With these qualifications, a long-term contract σ offered at the beginning of period 1 takes the following form:

$$\sigma = \{P_{ik}, \ (c_{ik}, w_{ik}), \ i=1,2, \ k=f,r\}.$$

Here, P_{ik} is the probability that the worker's employment status in period 2 (k) is r (retained) or f (fired), conditional on the realization of period 1 output being θ_i. In addition, c_{ik} is the worker's compensation in period 1 if output is θ_i and the worker's period 2 employment status is k. Finally, w_{ik} is the worker's expected utility at the beginning of period 2 if output in period 1 is θ_i and the worker's period 2 employment status is k.

The following problem characterizes the optimal contract that promises expected utility exactly equal to w to the worker:

$$\max_{a, P_{ik}, c_{ik}, w_{ik}} \sum_{i=1,2} \Pi_i(a) \sum_{k=f,r} P_{ik}[\theta_i - c_{ik} + V(w_{ik})] \tag{1}$$

subject to

$$a \in \mathbb{A}, \ P_{ik} \geq 0, \ \sum_k P_{ik} = 1, \ c_{ik} \in \mathbb{C}, \ w_{ik} \in \Phi_k, \ \forall i,k \tag{2}$$

$$\sum_{i=1,2} \Pi_i(a) \sum_{k=f,r} P_{ik}(u(c_{ik}) + w_{ik}) - \phi(a) = w, \tag{3}$$

$$a \in \arg\max_{a' \geq \mathbb{A}} \left\{ \sum_{i=1,2} \Pi_i(a') \sum_{k=f,r} P_{ik}(u(c_{ik}) + w_{ik}) - \phi(a') \right\}. \qquad (4)$$

Let $\sigma^* = \{a^*(x); P^*_{ik}(x), c^*_{ik}(x), w^*_{ik}(x)\}$ denote the solution to the problem above. Clearly σ^* is the optimal long-term contract we are looking for.

The end this section, note that we have so far conditioned the problem of optimal contracting on the assumption that the initial worker is employed in period 1. That is, we have chosen not to discuss the initial worker's employment status in period 1. This allows us to focus on termination at the end of period 2. It is clear however that if we take one step further, we can carry on the analysis to decide whether or not the initial worker should be employed in period 1. That is, we can allow termination to occur at the beginning of period 1 as well.[1]

3.3 Two types of terminations

3.3.1 Terminate when the worker is too poor to punish

In this section, we study a version of the model which allows us to describe the type of termination which occurs when the agent's promised utility is too low to support the desired effort. Assume workers are risk neutral, i.e., $u(c) = c$. Consumption must be non-negative: $c \in \mathbb{C} = [0, \infty)$. There are only two effort levels: $\mathbb{A} = \{a_1, a_2\}$, where $0 = a_1 < a_2$, with $\phi(a_1) = 0$ and $\phi(a_2) = \psi > 0$. So the workers' period reservation utility is $w_0 = 0$.

Assume $0 \leq \Pi_2(a_1) < \Pi_2(a_2) < 1$. Let $\overline{\theta}(a_i) = \Pi_1(a_i)\theta_1 + \Pi_2(a_i)\theta_2$. Assume $\overline{\theta}(a_1) \leq 0$ and $\overline{\theta}(a_2) > \psi$, so the high effort is associated with positive net returns, and the low effort is associated with negative net returns. Moreover, we assume that $\overline{\theta}(a_1)$ is sufficiently low that it is never optimal to implement the low effort.

We will focus on contracts that promise workers non-negative expected utilities. This is justified by the fact that workers are able to walk away from contracts that promise less than their reservation utilities. Now fix w, where

[1] To be specific, let $U_{\rm r}(x)$ denote the firm's value function conditional on retaining the initial worker. Suppose the initial worker is fired at the beginning of period 1, then the firm's value is

$$U_{\rm f}(x) = \max\{U_{\rm r}(x) | w \geq w_0\} - 2u^{-1}[x/2 + \phi(0)].$$

Here, when the initial worker is fired, he will not exert any effort in periods 1 and 2, and it is optimal to pay him cash compensation equal to $u^{-1}(x/2 + \phi(0))$ in both periods and to replace him with a new worker who is promised an expected utility w_{**}, where w_{**} is the worker's promised utility at which the firm's expected utility is maximized subject to $x \geq 0$. That is, $U_{\rm r}(w_{**}) = \max\{U_{\rm r}(x), x \geq 2w_0\}$. We assume $U_{\rm r}(w_{**}) > 0$. Let $U(x) = \max\{U_{\rm r}(x),\ U_{\rm f}(x)\}$. Let $\Psi \equiv \Psi_{\rm r} \cup \Psi_{\rm f}$ where $\Psi_{\rm r} \equiv \{x \in \Psi :\ U_{\rm r}(x) \geq U_{\rm f}(x)\}$ and $\Psi_{\rm f} \equiv \{x \in \Psi :\ U_{\rm r}(x) < U_{\rm f}(x)\}$. That is, if the initial worker is promised utility is x and $x \in \Psi_{\rm r}$, then the initial worker is retained; otherwise, he is fired. Notice that the set Ψ may not be convex and the function $U(x)$ may not be concave. Thus in general, the firm's optimal termination policy regarding the initial worker at the beginning of period 1 is determined by a lottery with Ψ as its support. For brevity we do not spell out the optimization problem fully here.

$w \geq 0$. Consider a one period contract (s_1, s_2), where $s_1, s_2 \geq 0$, that implements effort $a = a_2$ and promises expected utility equal exactly to w to the worker. That is, (s_1, s_2) satisfies the following incentive and promise-keeping constraints:

$$(1 - \Pi_2(a_2))s_1 + \Pi_2(a_2)s_2 - \psi \geq (1 - \Pi_2(a_1))s_1 + \Pi_2(a_1)s_2,$$

$$w = (1 - \Pi_2(a_2))s_1 + \Pi_2(a_2)s_2 - \psi.$$

Here the incentive constraint requires that $s_2 - s_1 \geq \frac{\psi}{\Pi_2(a_2) - \Pi_2(a_1)}$ which, given the non-negativity of s_1 and the promise-keeping constraint, implies that $w \geq \underline{w}$ where

$$\underline{w} = \frac{\Pi_2(a_1)}{\Pi_2(a_2) - \Pi_2(a_1)}\psi. \tag{5}$$

Thus, if $w \geq \underline{w}$ then $a^* = a_2$ is implementable, and the optimal contract is $s_1^*(w) = 0$, $s_2^*(w) = \frac{w+\psi}{\Pi_2(a_2)}$ and $V_{\mathrm{r}}(w) = \bar{\theta}(a_2) - \psi - w$, $\forall w \geq \underline{w}$.[2] On the other hand, if $w < \underline{w}$, then the high effort cannot be implemented and $V_{\mathrm{r}}(w) = \bar{\theta}(a_1) - w$.[3] Suppose the worker is paid cash in the amount of w to leave the firm, and the firm hires a new worker, then its value is $V_{\mathrm{f}}(w) = \bar{\theta} - \psi - \underline{w} - w$. we therefore have:

Proposition 1. $V_{\mathrm{r}}(w) < V_{\mathrm{f}}(w)$ if and only if $w < \underline{w}$, and $V_{\mathrm{r}}(w) \geq V_{\mathrm{f}}(w)$ if and only if $w \geq \underline{w}$. The firm's value function is:

$$\begin{aligned} V(w) &= \bar{\theta}(a_2) - \psi - \underline{w} - w, \;\; w < \underline{w}, \\ &= \bar{\theta}(a_2) - \psi - w, \;\; w \geq \underline{w}. \end{aligned}$$

This proposition reflects the main idea we are trying to convey in this section: termination is a necessary punishing device if the contract must make the worker sufficiently poor in period 2 (his expected utility sufficiently low, lower than $\underline{w}$.)[4]

Following from the definitions in Section 3.1, we have $\Phi_{\mathrm{f}} = [0, \underline{w})$, $\Phi_{\mathrm{r}} = [\underline{w}, \infty)$. We also have $V(0) = V(\underline{w})$.

We now move on to determine the optimal long-term contract. Here, we solve for the optimal long-term contract that implements the high effort, promises expected utility x to the worker, where $x \geq 0$. Now fix x. Notice that given

[2] We assume $V(\underline{w}) = \bar{\theta}(a_2) - \psi - \underline{w} > 0$, so that in the one period setting, if the worker is paid the efficiency wage, then the firm makes a positive profit. Without this assumption, it is optimal to shut down the firm.

[3] Clearly, $\underline{w} \geq 0$, and $\underline{w} = 0$ if and only if $\Pi_2(a_1) = 0$. Notice that if $\underline{w} > 0$, then $\underline{w}$ is the "efficiency wage (utility)" in the sense of Shapiro and Stiglitz [7] in that, first, $\underline{w}$ is strictly greater than the worker's reservation utility; second, any expected utility below $\underline{w}$ cannot support the high effort.

[4] It should perhaps be noted that Proposition 1 does not mean that the one period model is enough for thinking about termination. Notice Proposition 1 requires that the worker is promised expected utility equal exactly to w.

$V(\underline{w}) > 0$, whenever the initial worker is fired, he is always replaced. The optimization problem is:

$$\max_{(P_{ik}, c_{ik}, w_{ik})} \left\{ \sum_i \Pi_i(a_2) \sum_{k=f,k} P_{ik}[\theta_i - c_{ik} + V(w_{ik})] \right\} \tag{6}$$

subject to

$$c_{ik} \geq 0, \ \ \forall i, k, \tag{7}$$

$$0 \leq P_{ik} \leq 1, \ \sum_k P_{ik} = 1, \ \ \forall i, k, \tag{8}$$

$$w_{ik} \in \Phi_k, \ \ \forall i, k, \tag{9}$$

$$\sum_{i=1,2} \Pi_i(a_2) \sum_{k=f,r} P_{ik}(c_{ik} + w_{ik}) - \psi = x \tag{10}$$

$$\sum_{i=1,2} \Pi_i(a_2) \sum_{k=f,r} P_{ik}(c_{ik} + w_{ik}) - \psi \geq \sum_{i=1,2} \Pi_i(a_1) \sum_{k=f,r} P_{ik}(c_{ik} + w_{ik}). \tag{11}$$

The above incentive constraint can be rewritten as

$$\sum_k P_{2k}(c_{2k} + w_{2k}) - \sum_k P_{1k}(c_{1k} + w_{1k}) \geq \frac{\psi}{\Pi_2(a_2) - \Pi_2(a_1)}. \tag{12}$$

Proposition 2. Consider the optimal termination policy associated with an optimal contract that promises to deliver expected utility of at least x to the initial worker. Then the firm's optimal termination policy is $P^*_{2f}(x) = 0$ for all x, and

$$\begin{aligned} P^*_{1f}(x) &= 1, \ \ 0 \leq w \leq \underline{w}, \\ &= 2 - \frac{x}{\underline{w}}, \ \ \underline{w} < x < 2\underline{w}, \\ &= 0, \ \ x \geq 2\underline{w}. \end{aligned}$$

The proof of Proposition 2 is quite lengthy and appears in the Appendix. The first thing that Proposition 2 shows is that, given $\underline{w} > 0$, termination is a decreasing function of the agent's initial promised utility x. This is the theme of this section. If $x \geq 2\underline{w}$, then the optimal contract involves no termination. If $w \in [\underline{w}, 2\underline{w})$, then it is optimal to terminate the long-term contract with a positive probability in the low output state, and this probability is a decreasing function of x.

Proposition 2 also states that a higher $\underline{w}$ implies a higher rate of termination: For fixed x, P_{1f}^* is an increasing function of $\underline{w}$. Why is termination an increasing function of $\underline{w}$? Remember $\underline{w}$ is the minimum expected utility to which the initial worker must be entitled, if he is retained in period 2 and if he is given incentives to make the high effort. Suppose $\underline{w}$ increases. Other things being equal, the firm's ability to punish the initial worker in the state of low output is weakened. This then implies that the firm must rely more on using the threat of termination to motivate the worker.

We have the following corollary of Proposition 2 which makes some intuitive predictions of the model. The corollary states that the turnover rate of the initial worker at the end of period 1 is higher if the problem of moral hazard is more severe.

Corollary 3. The probability of termination increases as ψ increases, increases as $\Pi_2(a_1)$ increases, but decreases as $\Pi_2(a_2)$ decreases.

The corollary holds because it is clear from equation (5) that

$$\frac{\partial \underline{w}}{\partial \Pi_2(a_1)} > 0,\ \frac{\partial \underline{w}}{\partial \Pi_2(a_2)} < 0,\ \frac{\partial \underline{w}}{\partial \psi} > 0. \tag{13}$$

The Cost of Termination Note that in the physical environment of the model, there is not an exogenously imposed cost associated with the event of termination. However, the firm does incur a cost whenever it fires the initial worker. This cost is equal to $\underline{w}$, which is the "extra" net compensation that the firm gives to the new worker. This cost of termination is the model's endogenous variable. The following duality theorem states that a contract is optimal if and only if it minimizes the expected cost of firing.

Proposition 4. $\{P_{ik}^*, c_{ik}^*, w_{ik}^*\}$ is an optimal contract if and only if it solves the following programming problem:

$$\min_{P_{ik}, c_{ik}, w_{ik}} \{\Pi_1(a_2)P_{1f} + \Pi_2(a_2)P_{2f}\} \tag{14}$$

subject to (7)-(11).

Proof. Let $M(x)$ denote the maximand of Problem (6).

$$M(x) = 2\overline{\theta}(a_2) - \sum \Pi_i(a_2) \sum P_{ik}c_{ik} + \sum \Pi_i(a_2)[p_{if}V(w_{if}) + p_{ir}V(w_{ir})].$$

But $w_{if} = 0$ and $V(0) = V(\underline{w})$, so

$$M(w) = 2\overline{\theta}(a_2) - \sum \Pi_i(a_2) \sum P_{ik}c_{ik} + \sum \Pi_i(a_2)[p_{if}V(\underline{w}) + p_{ir}V(w_{ir})].$$

Now given $V(w) = \overline{\theta}(a_2) - \psi - w$ for $w \geq \underline{w}$, and $w = \sum \Pi_1(a_2) \sum_k P_{ik}[c_{ik} + w_{ik}] - \psi$ with $w_{if} = 0$, we then have:

$$M(w) = 2(\overline{\theta}(a_2) - \psi) - w - [\Pi_1(a_2)P_{1f} + \Pi_2(a_2)P_{2f}]\underline{w}.$$

Thus maximizing $M(x)$ subject to (7)-(11) is equivalent to minimizing $[\Pi_1(a_2)P_{1f} + \Pi_2(a_2)P_{2f}]\underline{w}$ subject to the same constraints.■

That termination is costly is useful for thinking about why termination is a decreasing function of x, as Proposition 2 shows. When x is higher, incentives can be more readily obtained by way of manipulating the compensation scheme and thus more costly termination can be avoided.

Remark

The ideas developed in this section are partly based on the early work of Stiglitz and Weiss [10]. The innovation here is that the worker can be replaced after he is fired, and this innovation, as we discussed earlier in the paper, is essential for what we are after in this paper. This innovation also allows us to discuss the cost of termination as an endogenous variable of the model and show its effects on termination. And of course the model here also allows for analytical solution that permits a much more precise expression of the idea that termination is optimal when the worker becomes too poor to punish.

3.3.2 Terminate when the worker is too rich to motivate

We now move on to describe the second type of termination which occurs when the agency costs of providing incentives to the agent exceed the cost to the firm of terminating the agent and replacing him with a new one. We will use a different version of the model for this case. Specifically, we will depart from the assumption of risk neutrality, and we will not need the assumption that compensation must be non-negative. The main assumption we use here is that workers are risk averse. That is, we assume $u(c)$ is well defined on $\mathbb{R}$ and $u(c)$ is strictly concave.

For mathematical convenience, we restrict the firm to deterministic termination policies. That is, we assume P_{ik} to be either 0 or 1. Clearly, this will not affect the firm's problem in the second period (the one-period problem). In particular, it will not affect the sets Φ, Φ_{r}, Φ_{f} and the value function $V(w)$. But a long-term contract now takes the simpler form of $\{c_i, w_i\}$, where c_i and w_i are the worker's current period consumption and promised utility in state θ_i respectively. Here, if $w_i \in \Phi_{\mathrm{r}}$, then the worker is retained. If $w_i \in \Phi_{\mathrm{f}}$, then the worker is fired. The firm's problem in the first period is characterized by

$$U_{\mathrm{r}}(w) = \max_{a, c_i, w_i} \sum_{i=1,2} \Pi_i(a)[\theta_i - c_i + V(w_i)] \tag{15}$$

subject to

$$a \in \mathbb{A}, \;\; c_i \in \mathbb{C}, \; w_i \in \Phi = \Phi_{\mathrm{r}} \cup \Phi_{\mathrm{f}}, \; \forall i, \tag{16}$$

$$\sum_{i=1,2} \Pi_i(a)[u(c_i) + w_i] - \phi(a) = w, \tag{17}$$

$$a \in \arg\max_{a' \geq \mathbb{A}} \left\{ \sum_{i=1,2} \Pi_i(a')[u(c_{ik}) + w_{ik}] - \phi(a') \right\}. \tag{18}$$

We now go back to characterize the one-period value function functions $V_f(w)$ and $V_r(w)$. Let $X(a) = \Pi_2(a)$.

Assumption 5. (i) $\mathbb{A} = [0, \infty)$. (ii) $\phi : \mathbb{A} \to \mathbb{R}_+$ is twice continuously differentiable and satisfies: $\phi(0) = 0$, $\phi'(0) = 0$; $\phi'(a) > 0$, $\forall a > 0$; $\phi'(\infty) = \infty$; $\phi''(a) > 0$, $\forall a$. (iii) $X(a)$ is twice continuously differentiable, strictly increasing, and satisfies $X(0) = 0$, $X(\infty) = 1$; for all a, $X'(a) > 0$, $X''(a) < 0$.

Under the above assumptions, we have

$$V_r(w) = \max_{a, c_1, c_2} \{\bar{\theta}(a) - (1 - X(a))c_1 - X(a)c_2\} \tag{19}$$

subject to

$$u(c_2) - u(c_1) = \frac{\phi'(a)}{X'(a)}, \tag{20}$$

$$(1 - X(a))u(c_1) + X(a)u(c_2) - \phi(a) = w, \tag{21}$$

where the first constraint is the incentive constraint in it first-order form[5], and the second constraint is promise-keeping. Now by substituting the incentive and promise-keeping constraints into the objective function, we have

$$V_r(w) = \max_a \left\{ \bar{\theta}(a) - (1 - X(a))u^{-1}[w + G(a)] - X(a)u^{-1}\left[w + G(a) + \frac{\phi'(a)}{X'(a)}\right] \right\}, \tag{26}$$

where $G(a) \equiv \phi(a) - X(a)\frac{\phi'(a)}{X'(a)}$. By the envelope theorem then,

$$V_r'(w) = -(1 - X(a^*))(u^{-1})'[w + G(a^*)] - X(a^*)(u^{-1})'\left[w + G(a^*) + \frac{\phi'(a^*)}{X'(a^*)}\right]. \tag{27}$$

[5] The first order approach to incentive compatibility is valid here. Let

$$K(a) \equiv (1 - X(a))u(c_1) + X(a)u(c_2) - \phi(a). \tag{22}$$

Then

$$K'(a) = [u(c_2) - u(c_1)]X'(a) - \phi'(a), \tag{23}$$

$$K''(a) = [u(c_2) - u(c_1)]X'(a) - \phi'(a). \tag{24}$$

Given the assumptions on $X(a)$ and $\phi(a)$, the optimal effort satisfies $a > 0$, which in turn implies $c_2 > c_1$, and hence the following holds,

$$K''(a) < 0, \;\; \forall a, \;\; K'(0) > 0, \;\; K'(\infty) = -\infty. \tag{25}$$

So the problem $\max_{a \in A} K(a)$ has a unique interior solution that is determined by $K'(a) = 0$.

Clearly, $V_r'(w) < 0$ for all w.

Recall that new workers are willing to work for $w_0 = u(0) - \phi(0)$ in period 2. Then, given $V_r'(w) < 0$ for all w,

$$V_f(w) = V_r(w_0) - u^{-1}(w).$$

Clearly,

$$V_r(w_0) = V_f(w_0).$$

Proposition 6. Suppose the function $(u^{-1})'$ is increasing and convex. Then for all $w > w_0$, $V_f(w) > V_r(w)$, and for all $w < w_0$, $V_f(w) < V_r(w)$.

Proof. For all $w \in \mathbb{R}$, let

$$F(w) = V_f(w) - V_r(w) = [V_r(w_0) - V_r(w)] - u^{-1}(w).$$

Now $F(w_0) = 0$. So to prove the proposition we need only show $F'(w) > 0$ for all w, or

$$V_r'(w) + (u^{-1})'(w) < 0.$$

Given equation (23), in order to show that the above holds, we must show

$$(1 - X(a^*))(u^{-1})'[w + G(a^*)] + X(a^*)(u^{-1})'\left(w + G(a^*) + \frac{\phi'(a^*)}{X'(a^*)}\right) > (u^{-1})'(w).$$

But the function $(u^{-1})'$ is increasing and convex, so that

$$\begin{aligned}
& (1 - X(a^*))(u^{-1})'[w + G(a^*)] + X(a^*)(u^{-1})'\left(w + G(a^*) + \frac{\phi'(a^*)}{X'(a^*)}\right) \\
\geq \; & (u^{-1})'\left[((1 - X(a^*))[w + G(a^*)] + X(a^*)\left(w + G(a^*) + \frac{\phi'(a^*)}{X'(a^*)}\right)\right] \\
= \; & (u^{-1})'[w + \phi(a^*)] \\
> \; & (u^{-1})'(w).
\end{aligned}$$

This proves the proposition.■

An example of the function $u(c)$ that satisfies the condition of the proposition is $u(c) = -\exp(-c)$, where $u^{-1}(w) = -\log(-w)$ where $w < 0$ and $(u^{-1})'(w) = -w^{-1}$.

Proposition 6 shows if the agent is too rich (his promised utility is higher than his reservation utility w_0, then the firm is better off replacing him with a new worker who earns the reservation utility w_0. Rich workers (whose promised utility is high) must be replaced because they are hard to motivate. They are hard to motivate because around the average level of consumption that the firm must provide for them, the marginal utility of consumption is low, and so any given amount of consumption increase (decrease) implies less utility gain (loss) to him than to a "poorer" woker.

4 Wage dynamics and contract termination in a fully dynamic model

We now consider the infinite horizon version of model where time is denoted $t = 1, 2, ...$ The firm's preferences are

$$\mathsf{E}_0 \sum_{t=1}^{\infty} (1-\delta)\delta^{t-1} c_t,$$

and the worker's preferences are

$$\mathsf{E}_0 \sum_{t=1}^{\infty} (1-\delta)\delta^{t-1} [u(c_t) - \phi(a_t)],$$

where $\delta \in [0, 1)$ is the discount factor. The workers are risk averse in consumption and so the utility function u is concave.

To keep the problem tractable, we again focus on pure strategy termination rules. Under this assumption, a fully dynamic contract can be expressed recursively as

$$\{c_i(w), w_i(w),\ w \in \Phi = \Phi_r \cup \Phi_f\},$$

with the following interpretation. Suppose at the beginning of date t a worker is promised utility w and he is employed this period. Let his output in this period be θ_i. Then his compensation is $c_i(w)$, and his expected utility from next period on is $w_i(w)$, where if $w_i(w) \in \Phi_r$, he continues to be employed, otherwise he is fired. Here the set Φ is the set of promised utilities of the worker that can be supported.

Let $V : \Phi \to R$ denote the firm's value function. The optimal dynamic contract solves the following Bellman equation.

$$V(w) = \max\{V_{\mathrm{r}}(w),\ V_{\mathrm{f}}(w)\}$$

where

$$V_{\mathrm{r}}(w) = \max_{a, c_i, w_i} \sum \Pi_i(a)[(1-\delta)(\theta_i - c_i) + \delta V(w_i)] \tag{28}$$

subject to $c_i \geq 0$, $w_i \in \Phi$,

$$\sum \Pi_i(a)[(1-\delta)u(c_i) + \delta w_i] - (1-\delta)\phi(a) = w, \tag{29}$$

$$a \in \arg\max_{a \in \mathbb{A}} \left\{ \sum \Pi_i(a)[(1-\delta)u(c_i) + \delta w_i] - (1-\delta)\phi(a) \right\}, \tag{30}$$

and

$$V_{\mathrm{f}}(w) = \max\{V_{\mathrm{r}}(w), w \geq w_0\} - u^{-1}[w + \phi(0)], \tag{31}$$

and

$$\Phi = \Phi_{\rm r} \cup \Phi_{\rm f}$$

where

$$\Phi_{\rm r} = \{w \in \Phi : \; V_{\rm r}(w) \geq V_{\rm f}(w)\} ,$$

$$\Phi_{\rm f} = \{w \in \Phi : \; V_{\rm r}(w) < V_{\rm f}(w)\} .$$

The optimal contract that solves the Bellman equation above fully characterizes the equilibrium of the model. Specifically, in equilibrium the starting expected utility of a new worker is equal to

$$\underline{w} = \arg\max_{w \in \Phi_r} V_r(w) \tag{32}$$

Moreover, when an old worker is fired, he gets a lump-sum compensation equivalent to a constant stream of $u^{-1}[w + \phi(0)]$ in each period of his life.

The compensation paid on termination has an obvious correlative in real contracting situations. In the context of CEO compensation, this payment is the CEO's "golden parachute", which is paid out when the CEO is either forced to retire or the firm is taken over. We note that in most real world contracts, the parachute payment is generally negotiated ahead of time, so that the endogenously determined termination payment in our model would serve, in a model of take-over, to determine whether the CEO and/or other high level corporate officers were actually let go (with the contractual parachute payment being made), or kept on in some alternative capacity if the contractual parachute payment exceeded the endogenously determined optimal termination payment. Note that the model predicts that the size of the parachute should depend positively on w, which in turn depends on the CEO's history of performances. This property of the model is in accord with the observed fact that the CEOs generally get larger parachute payments than more junior officers, and that the CEOs for larger, more successful firms get larger parachute payments than those of smaller, less successful firms.

Analytic solutions to the dynamic programming problem above are hard to obtain, so we rely mainly on numerical simulations to shed light on the relationship between compensation dynamics and termination. Specifically, consider the following parameterization of the model: $H(c, a) = \sqrt{0.1 + c} - a^2$, $c \geq 0$, $\mathbb{A} = [0, \infty)$, $\Theta = \{0, 1\}$, $\Pi_2(a) = 1 - \exp(-a)$, $a \geq 0$, $\delta = 0.95$. Here $w_0 = \sqrt{0.1}$.

Figure 1 shows the value functions $V_{\rm r}(w)$ and $V_{\rm f}(w)$. Notice the equilibrium expected utility of the new worker, $\underline{w}$, is strictly greater than the worker's reservation utility w_0. Figure 2 shows the law of motion for the worker's promised utility w. Figure 3 shows the worker's current compensation as a function of his expected utility w and output θ. Figure 4 shows the worker's effort as a function of w.

As Figure 1 shows, the infinite horizon model here embeds both types of terminations we discussed in the previous section. The first type of termination occurs when the worker's promised utility is sufficiently low and the worker becomes too poor to punish. The second type of termination occurs when the worker's promised utility is sufficiently high and worker is too rich to motivate.

The first type of termination is involuntary because the utility of the worker who is fired is strictly lower than the promised utility of the new worker who the firm hires to replace him. (The new worker starts out with promised utility $\underline{w}$.) The second type of termination is not involuntary because the worker who is terminated receives from the firm a constant stream of consumption which makes him strictly better off than the new worker the firm hires to replace him. Another difference between the two types of terminations is that the first type always occurs after the agent produces a low output whereas the second type occurs after a high output. In a sense, termination of the first type is a punishing device, while termination of the second type is a rewarding device.

Now suppose a termination has just occured and a new worker starts out with the equilibrium starting expected utility $\underline{w}$. If he produces a low output in the first period, then he is fired immediately. If he produces a high output in the first period, then his current compensation is positive, he is retained and promised a utility strictly higher than $\underline{w}$. Suppose in the following periods the worker continues to produce high outputs. Then his expected utility continues to rise, current compensation increases, and he is terminated in period 8, that is, after seven consecutive high outputs. The CEO is terminated with a "golden parachute" that is equivalent to a stream of constant compensation payments.

Suppose, instead, the worker produces a sequence of six high outputs and then a low ouput. Then he will be retained in period 8. Let $\tilde{w}$ denote the new worker's expected utility after he has produced six high outputs. We have the following interesting result: Conditional on $\tilde{w}$, the worker is terminated if the next output is high and he is retained if the next output is low. This seemingly counter-intuitive result has a natural explanation in our model. What happens here is, if output is high, then the worker must be promised more utility in the next period which in turn makes him too expensive to motivate, and so he should be replaced. On the other hand, when output is low in the current period, the worker will be promised a lower utility in the next period, but that makes him less expensive to motivate in the future and so he should be retained.

Suppose we follow a worker who starts with a relatively high expected utility and then produces a sequence of low outputs. Then in each of the following periods, his current compensation is lower than before, and his promised expected utility declines. The worker will experience a sequence of current compensation and promised utility reductions, until his promised utility is sufficiently low and he is fired after another low output. Note this time when the CEO is fired, the size of his "golden parachute" is zero.

A new worker is much more likely to be laid off involuntarily than an old worker. Specifically, a new worker is laid off immediately after one low output, whereas it takes many periods of low output before an old worker is finally fired. On the other hand, an older worker (whose promised expected utility is necessarily higher than $\underline{w}$) is more likely to experience a termination of the second type in which he is given a parachute payment to essentially retire from the firm.

Finally, we simulate the model for a large number of periods, and we keep track of the expected utilities of the employed workers. Workers each start with the market promised utility $\underline{w}$ and work for the firm until they are fired and replaced by a new worker. For each worker, in the long-run, termination occurs with probability one. We then get an ergodic distribution of promised utilities over a bounded range. This follows immediately from the fact that the sequence of compensations and promised utility payments lies in a compact set, so that the cluster points of the sequence constitute the support of the ergodic distribution, while the relative frequencies with which each cluster point is hit constitute the ergodic probabilities. One can also establish this result directly by noting that the incentive mechanism constitutes a random walk with reflecting barriers. The ergodic distribution is shown in Figure 5.

The fact that we obtain an ergodic limiting result for the model illustrates our main point: while the incentive mechanism of the long-term contract generates transient dynamics, the economic environment in which we embed the mechanism (in this case a renewable labor market) allows the option of optimal termination, which in turn generates an observable, ergodic outcome amenable to empirical verficiation. We note that in the same numerical example we have just examined, if we impose, as most existing models do, the constraint that the firm and the worker bind themselves to each other forever, then the limiting distribution of the worker's promised utility is degenerate: it converges to w_0, which is an absorbing state, with probability one.

5 Concluding Remarks

In this paper, we have explored the question of how to embed the incentive structures of the standard repeated agency model in an economic environment that transforms the normally transient behavior of the optimal long-term contract into one which is ergodic, and therefore subject to empirical verification. For the basic model we considered here, this environment involes the specification of an external labor market, and the relaxation of the standard constraints which bind the parties to the long-term contract to each other forever. This combination of options external to the contract and the option to terminate the contract then yield a model whose equilibria are stationary.

While we have focused on one particular interpretation of the model – that of the relationship of a CEO to the shareholders of a firm – the approach we have taken can clearly be enriched. One could, for example, enrich the specification of the labor market to permit agents to have different *ex ante* abilities or reservation utilities, or in which agents engage in search for desirable contracts. On the firm side of the model, one could extend the decision-making chores of the agent to allow for capital accumulation or project development in order to model the effects of technological change and the implied comparative advantages that such change can confer on different managers. A richer model of the firm could also allow analysis of such things as the free cash flow theory of takeover (see, e.g. [4]) and related issues in corporate finance.

References

[1] Aiyagari, S. Rao and F. Alvarez, Efficient dynamic monitoring of unemployment insurance claims, University of Chicago working paper, July 1995.

[2] Atkeson, A. and R. Lucas (1995), Efficiency and equality in a simple model of efficient unemployment insurance, *Journal of Economic Theory*, 66, 64-88.

[3] Green, E. (1987), Lending and the smoothing of uninsurable income, in *Contractual Arrangements for Intertemporal Trade*, University of Minnesota Press, Minneapolis, MN (E.C. Prescott and N. Wallace, eds.)

[4] Jensen, M.C. (1988), Takerovers: their causes and consequences, *Journal of Economic Perspectives*, Vol. 2, Num. 1, 21-48

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

[6] Phelan, C., and R. Townsend (1991), Computing multiperiod information-constrained optima, *Review of Economic Studies*, 58, 853-881.

[7] Shapiro, C. and J.E. Stiglitz (1984), Equilibrium unemployment as a worker-discipline device, *American Economic Review*, 74, 433-444.

[8] Smith, A.A. and C. Wang (2000), Dynamic credit relationships in general equilibrium, GSIA Working Paper #2000-27.

[9] Spear, S.E. and S. Srivastava (1987), On repeated moral hazard with discounting, *Review of Economics Studies*, 54, 599-617.

[10] Stiglitz, J.E. and A. Weiss(1983), Incentive effects of termination: applications to the credit and labor markets *American Economic Review* 73, 912-927.

[11] Taub, B. and H. Chade (2001), Segmented risk-sharing in a continuous-time setting, *Economic Theory*, forthcoming.

[12] Thomas, J. and R. Worrall (1990), Income fluctuations and asymmetric information: an example of the principle-agent problem, *Journal of Economic Theory*, 51, 367-390.

[13] Wang, C. (1997), Incentives, CEO compensation, and shareholder wealth in a dynamic agency model, *Journal of Economic Theory*, 76, 72-105.

Appendix

In this appendix we prove Proposition 2. We use a sequence of lemmas whose proofs are also provided later in this appendix.

Lemma 1 It is always optimal to set $w_{if} = 0$.

Lemma 1 shows if a worker is fired, then his consumption and utility in period 2 is zero.

Lemma 2 (i) It is always optimal to set $c_{ir} = 0$. (ii) It is always optimal to set $c_{if} = 0$ if $P_{ir} > 0$.

Lemma 2 implies that wecan restrict attention to contracts which satisfy $c_{if} P_{if} = 0$. The argument employed in the proof of Lemma 2 reflects the notion that firing is costly, and it is always optimal to minimize termination whenever possible. We will come back to this point later.

Lemma 3 It is never optimal to fire the worker with probability one. That is $P_{1f} = P_{2f} = 1$ is not optimal.

Lemma 4 It is never optimal to have $P_{1r} > P_{2r} = 0$.

Thus by Lemmas 3 and 4, $P_{2r} \neq 0$ at the optimum. Moreover, given the lemmas, the problem of optimal contracting can be simplified as:

$$\begin{aligned} U_r(w) \quad = \quad & \max_{c_{1f}, P_{ik}, w_{ir}} \{(\Pi_1(a_2)[\theta_1 + P_{1f}(-c_{1f} + V(\underline{w})) + P_{1r}V(w_{1r})] \\ & +\Pi_2(a_2)[\theta_2 + P_{2f}V(\underline{w}) + P_{2r}V(w_{2r})]\} \end{aligned} \tag{33}$$

subject to $c_{1f} \geq 0$, $0 \leq P_{ik} \leq 1$, $w_{ir} \geq \underline{w}$,

$$c_{1f} P_{1r} = 0 \tag{34}$$

$$P_{2r} w_{2r} - (P_{1f} c_{1f} + P_{1r} w_{1r}) \geq \frac{\psi}{\Pi_2(a_2) - \Pi_2(a_1)} \tag{35}$$

$$w = \Pi_1(a_2)(P_{1f} c_{1f} + P_{1r} w_{1r}) + \Pi_2(a_2) P_{2r} w_{2r} - \psi \tag{36}$$

Lemma 5 (i) If $P^*_{ir} w^*_{ir} < \underline{w}$, then it must hold that $P^*_{ir} < 1$, $w^*_{ir} = \underline{w}$. (ii) If $P^*_{ir} < 1$, then it must hold that $w^*_{ir} = \underline{w}$. (iii) If $w^*_{ir} > \underline{w}$, then it must hold that $P^*_{ir} = 1$.

Lemma 6 If $P^*_{2r} < 1$, then it is necessary that $P^*_{1r} = 0$.

Lemma 7 $P^*_{2r} = 1$.

We now solve for P^*_{1r}. Go back to the program (20). With $P^*_{2r} = 1$, the incentive and promise-keeping constraints imply

$$w \geq \Pi_1(a_2)(P_{1f} c_{1f} + P_{1r} w_{1r}) + \frac{\Pi_2(a_2)\psi}{\Pi_2(a_2) - \Pi_2(a_1)} - \psi \geq \underline{w}.$$

Thus if $w < \underline{w}$, then $a = a_2$ cannot be implemented. If $w = \underline{w}$, then by the above equation it must hold that $c^*_{1f} = 0$ and $P^*_{1r} = 0$. In addition w^*_{2r} can be solved from the promise-keeping constraint:

$$w_{2r}^{*}=\frac{w+\psi}{\Pi_2(a_2)}.$$

Note that here the incentive constraint holds as equality.

Now all that is left is to solve for P_{1r}^{*} for $w>\underline{w}$.

Lemma 8 If $w>\underline{w}$, then $P_{1r}^{*}>0$.

Given Lemma 8, for $w>\underline{w}$ we can write the problem of optimal contract in the following form:

$$U(w)=\max_{P_{1r},w_{2r}}\{\overline{\theta}(a_2)+\Pi_1(a_2)V(\underline{w})+\Pi_2(a_2)V(w_{2r})\} \tag{37}$$

subject to $0\leq P_{1r}\leq 1$, $w_{2r}\geq\underline{w}$,

$$w_{2r}-P_{1r}\underline{w}\geq\frac{\psi}{\Pi_2(a_2)-\Pi_2(a_1)}, \tag{38}$$

$$w=\Pi_1(a_2)P_{1r}\underline{w}+\Pi_2(a_2)w_{2r}-\psi. \tag{39}$$

Notice that the constraint $w_{2r}\geq\underline{w}$ is never binding, because it is always implied by $0\leq P_{1r}\leq 1$ and the incentive constraint. It is also clear that the program has a solution if and only if $w\geq w_A$, where

$$w_A=\Pi_1(a_2)0\underline{w}+\Pi_2(a_2)\frac{\psi}{\Pi_2(a_2)-\Pi_2(a_1)}-\psi=\underline{w}$$

It is also clear from the figure that $P_{1r}^{*}(w)=1$ if and only if $w\geq w_B$ where

$$w_B=\Pi_1(a_2)1\underline{w}+\Pi_2(a_2)(\underline{w}+\tilde{w})-\psi=2\underline{w}.$$

Moreover, for $w=\underline{w}$, $P_{1r}^{*}=0$, $w_{2r}^{*}=\tilde{w}$. For $w\in(\underline{w},2\underline{w})$, $P_{1r}^{*}\in(0,1)$, and (P_{1r}^{*},w_{2r}^{*}) solve the following equations

$$w_{2r}^{*}=P_{1r}^{*}\underline{w}+\tilde{w},$$

$$w=\Pi_1(a_2)P_{1r}^{*}\underline{w}+\Pi_2(a_2)w_{2r}^{*}-\psi.$$

In other words, for $w\in(\underline{w},2\underline{w})$,

$$w_{2r}^{*}=w-\underline{w}+\frac{\psi}{\Pi_2(a_2)-\Pi_2(a_1)},$$

$$P_{1r}^{*}=\frac{\underline{w}-\Pi_2(a_2)\tilde{w}+\psi}{\underline{w}}=\frac{w}{\underline{w}}-1,$$

and thus of course $P^*_{1f}(w) = 2 - w/\underline{w}$. The proposition is proven.

Proof Lemma 1: We need only show that for any long-term contract $\sigma = \{P_{ik}, c_{ik}, w_{ik}, i = 1, 2, k = f, r\}$, there exists a contract $\hat{\sigma} = \{\hat{P}_{ik}, \hat{c}_{ik}, \hat{w}_{ik}, i = 1, 2, k = f, r\}$, which implements the same effort, the same expected utilities for the initial worker and the firm as σ does, and satisfies $\hat{w}_{if} = 0$.

Suppose $w_{if} > 0$, for some $i \in \{1, 2\}$. Construct the contract $\hat{\sigma}$ in the following way. Let

$$\hat{c}_{if} = c_{if} + w_{if},\ \ \hat{c}_{jl} = c_{jl},\ jl \neq if,$$

$$\hat{w}_{if} = 0,\ \ \hat{w}_{jl} = w_{jl},\ jl \neq if,$$

Now notice that given risk neutrality, $V(w_{ik}) = V_{\text{f}}(w_{ik}) = A - w_{ik}$ where A is some constant. It is clear that $\hat{\sigma}$ implements the same effort, it promises the same level of expected utility to the agent, and it gives the firm the same expected utility as σ does.■

Proof Lemma 2: (i) is obvious. To show (ii), suppose $P_{ir} \neq 0$ but $c_{if} > 0$, then the firm is weakly better off by reducing c_{if} to zero and increasing w_{ir} while keeping the worker's expected utility fixed.■

Proof of Lemma 3: Suppose the optimal contract has $P_{1f} = P_{2f} = 0$, then the rest of the optimal contract is characterized by

$$U(w) = \max_{c_{1f}, c_{2f}} \left\{ \overline{\theta}(a_2) - \sum \Pi_i(a_2) c_{if} + V(\underline{w}) \right\} \tag{40}$$

subject to $c_{if} \geq 0$,

$$c_{2f} - c_{1f} \geq \frac{\psi}{\Pi_2(a_2) - \Pi_2(a_1)}$$

$$w = \sum \Pi_i(a_2) c_{if} - \psi.$$

Now consider a contract σ' defined by

$$P'_{1r} = 0,\ c'_{1f} = c_{1f},\ P'_{2r} > 0,\ c'_{2f} = 0,\ w'_{2r} = \underline{w},$$

where P'_{2r} is set to satisfy $P'_{2r} w'_{2r} = c_{2f}$. Notice that $c_{2f} > 0$ for otherwise the incentive constraint cannot hold. Now, compared to σ, the contract σ' implements the same level of effort and promises the same utility to the agent. However, under contract σ', the firm's expected utility is strictly higher than under σ, a contradiction.■

Proof of Lemma 4: Suppose $P_{1r} > P_{2r} = 0$ holds at the optimum. Then the optimal contract is characterized by the following optimization problem:

$$U(w) = \max_{c_{2f}, P_{1k}, w_{1r}} \{(\Pi_1(a_2)[\theta_1 + P_{1f} V(\underline{w}) + P_{1r} V(w_{1r})] + \Pi_2(a_2)[\theta_2 - c_{2f} + V(\underline{w})]\} \tag{41}$$

subject to $c_{2f} \geq 0$, $0 \leq P_{1k} \leq 1$, $w_{1r} \geq \underline{w}$,

$$c_{2f} + \underline{w}) - (P_{1f}\underline{w} + P_{1r}w_{1r}) \geq \frac{\psi}{\Pi_2(a_2) - \Pi_2(a_1)} \tag{42}$$

$$w = \Pi_1(a_2)(P_{1f}\underline{w} + P_{1r}w_{1r}) + \Pi_2(a_2)(c_{2f} + \underline{w}) \tag{43}$$

notice that the incentive constraint implies $c_{2f} > (P_{1f}\underline{w} + P_{1r}w_{1r}) - \underline{w} \geq 0$.

Now, consider a contract σ' which is otherwise identical to σ except $w'_{2r} = \underline{w}$, $c'_{2r} = 0$, and P'_{2r} is determined by $P'_{2r}w_{2r} = c_{2f} > 0$. The contract σ' implements the same effort, promises the same expected utility to the worker, makes the firm strictly better off. This is a contradiction.■

Proof of Lemma 5: Notice that this program contains a sub-program that governs the optimal relationship between P_{ir} and w_{ir}:

$$\max_{P_{ik}, w_{ir}} \{(1 - P_{ir})V(\underline{w}) + P_{ir}V(w_{ir})\} \tag{44}$$

subject to $P_{ir} \in [0, 1]$, $w_{ir} \geq \underline{w}$, and

$$P_{ir}w_{ir} = \Delta, \tag{45}$$

where $\Delta(> 0)$ is some given constant. But this problem in turn is equivalent to the following problem:

$$\max_{P_{ik}, w_{ir}} \{P_{ir}\underline{w} - \Delta\}$$

subject to $P_{ir} \in [0, 1]$, $w_{ir} \geq \underline{w}$, and $P_{ir}w_{ir} = \Delta$. Clearly, if $\Delta < \underline{w}$, then the optimal combination of P_{ir} and w_{ir} has

$$P_{ir} < 1, \quad w_{ir} = \underline{w};$$

and if $\Delta > \underline{w}$, then

$$P_{ir} = 1, \quad w_{ir} > \underline{w};$$

and finally, if $\Delta = \underline{w}$, then $P_{ir} = 1$, $w_{ir} = \underline{w}$. These results are summarized in Lemma 5.■

Proof of Lemma 6: Suppose $P_{2r} < 1$. Remember that $P_{2r} > 0$. Now suppose $P_{1r} > 0$, then $c_{1f} = 0$. But $P_{2r} < 1$ implies $w_{2r} = \underline{w}$, and thus, by Lemma 4.4, the incentive constraint reads

$$\underline{w} - (P_{1f}\underline{w} + P_{1r}w_{1r}) \geq \frac{\psi}{\Pi_2(a_2) - \Pi_2(a_1)},$$

which in turn implies

$$P_{1f}\underline{w} + P_{1r}w_{1r} < \underline{w},$$

and hence $P_{1r}w_{1r} < \underline{w}$, and so again by Lemma 5 it must hold that $w_{1r} = \underline{w}$. This is a contradiction for it leads to a violation of the incentive constraint.■

Proof of Lemma 7: Suppose $P_{2r} < 1$, then $P^*_{1f} = 0$, and $w_{2r} = \underline{w}$. The incentive constraint now reads

$$[\Pi_2(a_2) - \Pi_2(a_1)](P_{2r}\underline{w} - c_{1f}) \geq \psi,$$

which, given $c_{1f} \geq 0$, implies

$$P_{2r}\underline{w} \geq \frac{\psi}{\Pi_2(a_2) - \Pi_2(a_1)}.$$

But this can never hold, because

$$P_{2r}\underline{w} < \underline{w} = \frac{\Pi_2(a_1)\psi}{\Pi_2(a_2) - \Pi_2(a_1)} < \frac{\psi}{\Pi_2(a_2) - \Pi_2(a_1)},$$

a contradiction.■

Proof of Lemma 8: Let $w > \underline{w}$. Suppose $P^*_{1r} = 0$. Then the optimal contract is characterized by the following problem:

$$U(w) = \max_{c_{1f}, w_{2r}} \{\overline{\theta}(a_2) + \Pi_1(a_2)[-c_{1f} + V(\underline{w}] + \Pi_2(a_2)V(w_{2r})\} \tag{46}$$

subject to $c_{1f} \geq 0$, $w_{2r} \geq \underline{w}$,

$$[\Pi_2(a_2) - \Pi_2(a_1)](w_{2r} - c_{1f}) \geq \psi, \tag{47}$$

$$w = \Pi_1(a_2)c_{1f} + \Pi_2(a_2)w_{2r} - \psi. \tag{48}$$

Now suppose $c_{1f} = 0$. Then from the promise-keeping constraint:

$$w_{2r} = \frac{w + \psi}{\Pi_2(a_2)},$$

and the incentive constraint is

$$w_{2r} \geq \frac{\psi}{\Pi_2(a_2) - \Pi_2(a_1)}.$$

Clearly the incentive constraint is not binding, for $w > \underline{w}$ implies

$$\frac{w + \psi}{\Pi_2(a_2)} > \frac{\psi}{\Pi_2(a_2) - \Pi_2(a_1)}.$$

Now reduce w_{2r} by $\Delta > 0$ (this is possible for $w_{2r} > 0$ by the incentive constraint), increase P_{1r} from 0 to $\delta > 0$, where Δ and δ sufficiently small to make the incentive constraint hold, and are chosen to make the promise-keeping constraint hold, i.e.,

$$(1 - \Pi_2(a_2))\delta\underline{w} - \Pi_2(a_2)\Delta = 0.$$

But this strictly increases the firm's expected utility, a contradiction.

Suppose $c_{if} > 0$, then reducing c_{1f} and increasing P_{1r} can make the firm strictly better off.■

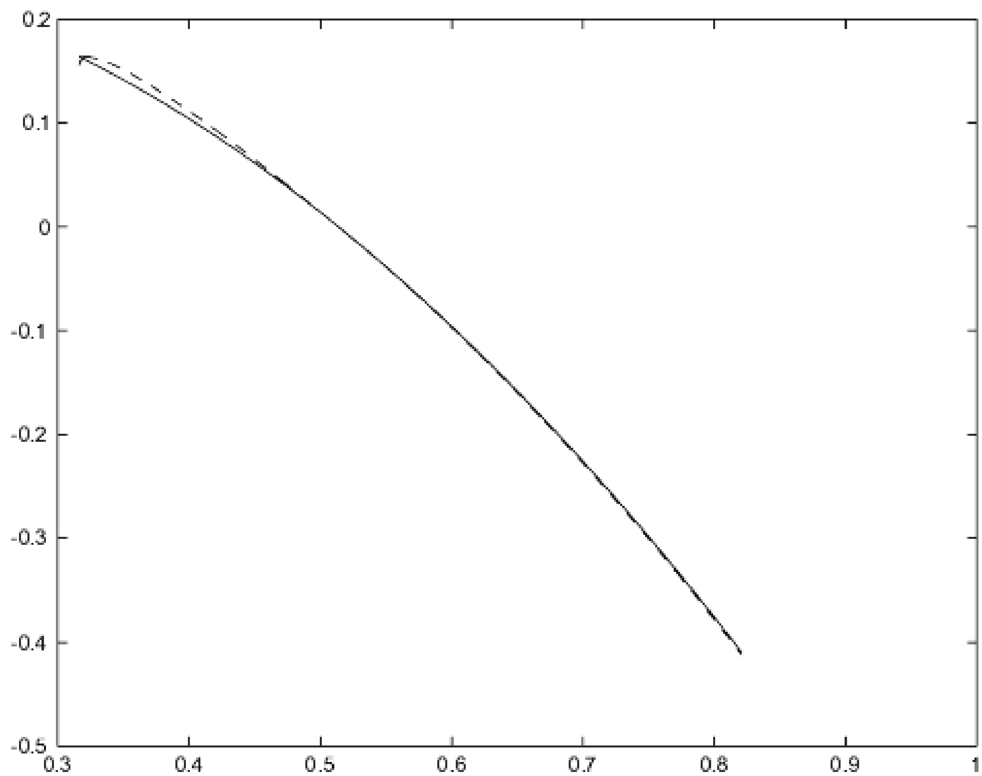

Figure 1: The value functions $V_r(w)$ (solid line) and $V_f(w)$ (dashed line)

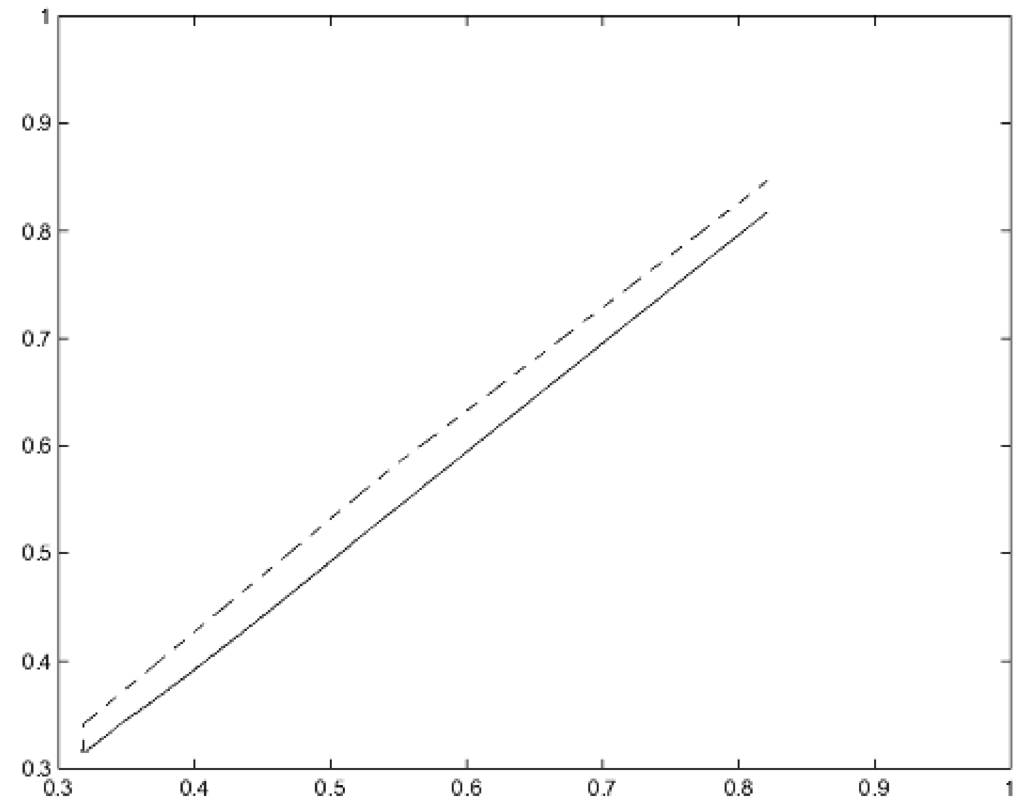

Figure 2: The law of motion for the state variable w. Here the solid line is for $w_1(w)$ and the dashed line is for $w_2(w)$.

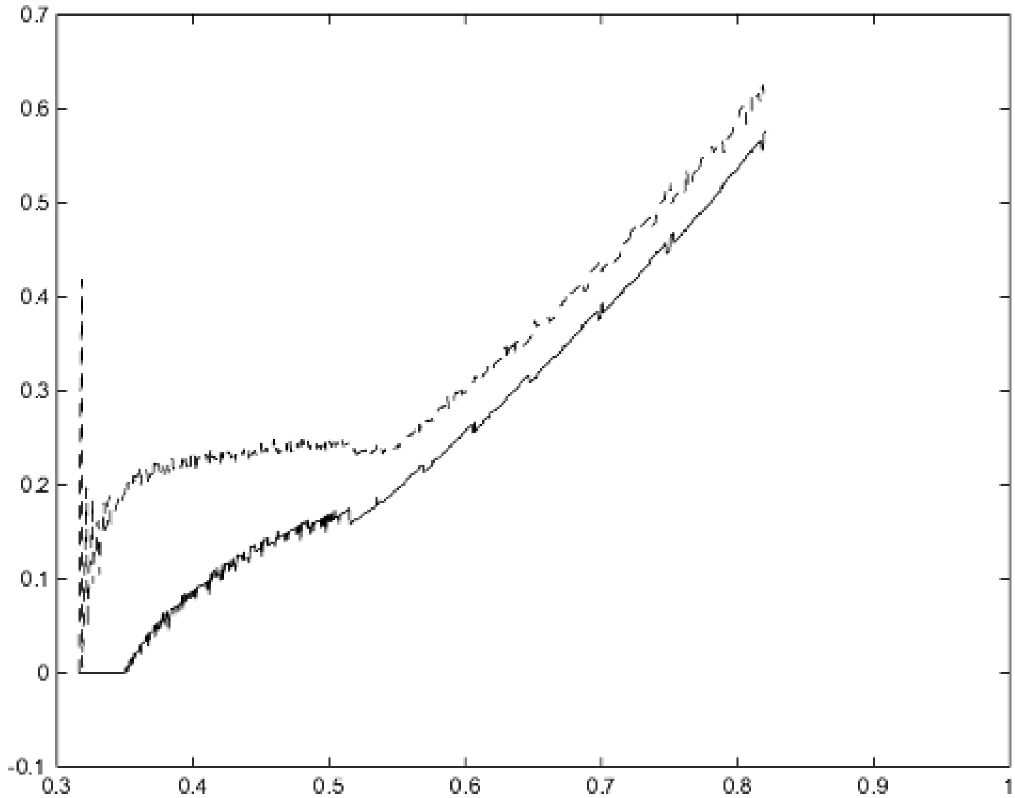

Figure 3: The CEO's consumption as a function of w. Here the solid line is for $c_1(w)$ and the dashed line is for $c_2(w)$.

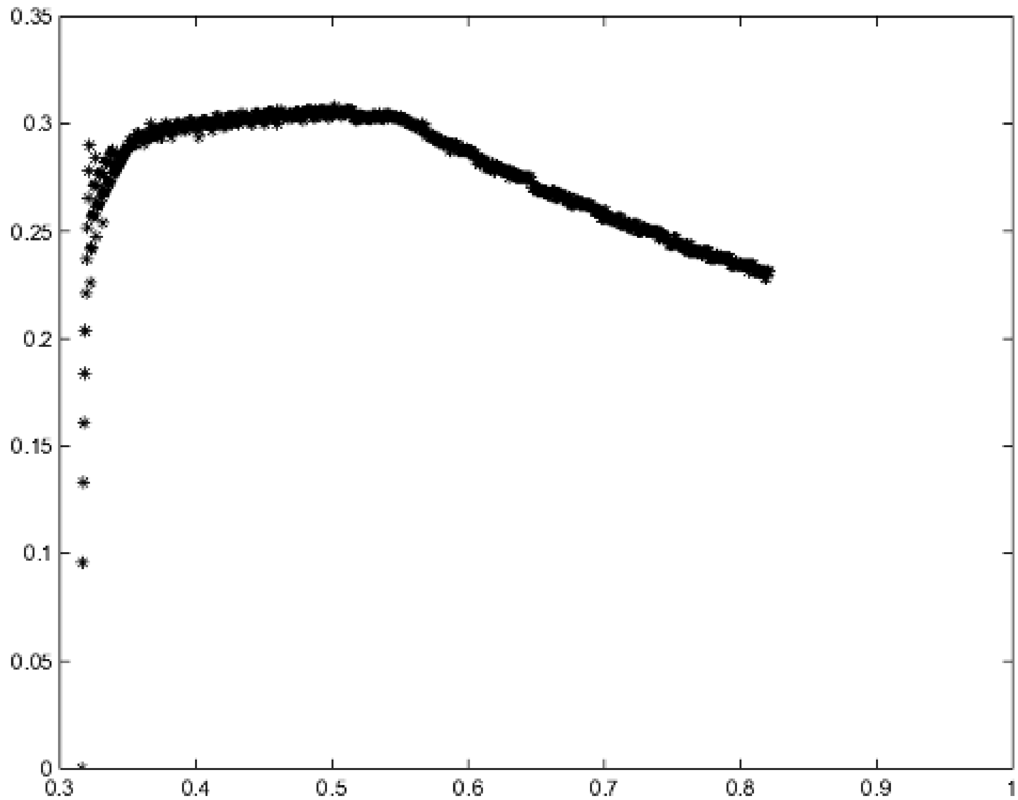

Figure 4: The optimal effort as a function of w.

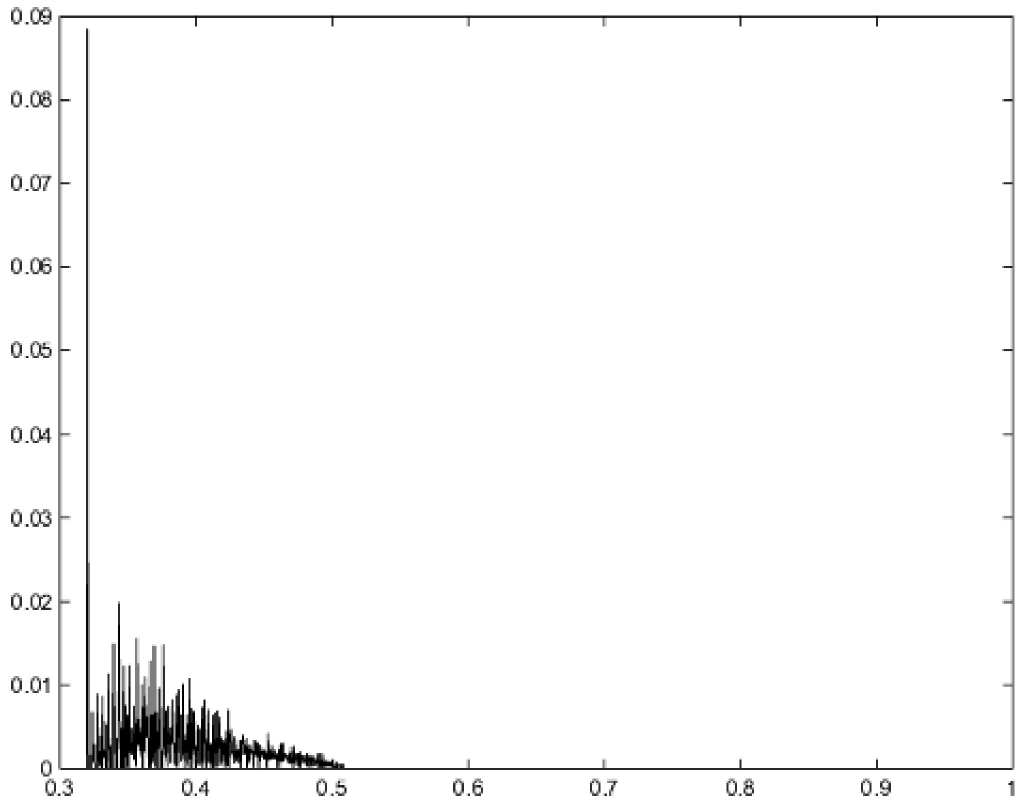

Figure 5: The limiting distribution of expected utilities.