Permanent Separations and Optimal Compensation with History-Dependent Reservation Utilities*

Stanimir Morfov†

February, 2009

Abstract

This paper considers an infinite-horizon, hidden action model characterized by limited commitment and history-dependent reservation utilities. I focus on the optimal contract allowing for permanent separations which can be triggered by any party. I prove existence and characterize the solution recursively. I compute the optimal contract for top executives and find little difference with the optimal self-enforcing contract. Namely, high effort is optimal for most but the richest managers. Compensation and future utility promise increase in both current utility and firm's stock price. The contract is less successful, however, in smoothing initial utility promises across initial price histories.

Keywords: principal-agent problem, moral hazard, dynamic contracts, executive compensation, limited commitment

Journal of Economic Literature Classification Numbers: C63, D82, G30

*I am particularly grateful to Manuel Santos for his guidance and support.

†State University - Higher School of Economics, International College of Economics and Finance, Pokrovski Bulvar, 11, 109028 Moscow, Russian Federation, smorfov@hse.ru

1 Introduction

An important strand of the dynamic agency literature has investigated the long run properties of optimal contracts. Green (1987), Thomas and Worrall (1990) and Phelan and Townsend (2001) have shown that legally enforceable contracts lead to degenerate long-term wealth distributions. Atkenson and Lucas (1995) demonstrate that if the agent's expected discounted utility is bounded from below, the model induces a non-degenerate invariant distribution. Phelan (1995) achieves a similar result by assuming limited commitment on part of the agent. In such a setting, a self-enforcing contract induces participation at every contingency, so the agent has no incentive to leave the relationship.

The techniques developed by Abreu, Pearce and Stacchetti (1990) prove useful in characterizing the space of agent's feasible continuation utilities. The optimal dynamic contract can then be represented recursively by a series of static contracts defined on this space in line with Spear and Srivastava (1987). Introducing limited commitment on part of the principal poses a different problem related to the fact that the value function of the recursive problem enters the set of constraints and standard dynamic programming fails since the operator used is no longer a contraction. Nevertheless, Rustichini (1998) based on some earlier results by Streufert (1992) shows that a simple modification of the operator allows a recursive representation of the dynamic problem.

The dynamic contracts of the form described above refer to long-term relationships which (by construction) neither party has an incentive to renege on. Spear and Wang (2005) focus on contract terminations instead and achieve stationary and ergodic dynamics by allowing for replacements of the contracted agent with a new one from a labor market pool and for golden parachutes at termination. Sleet and Yeltekin (2005) consider a dynamic model of temporary layoffs and permanent separations where their limited commitment case allows for both the firm and the employee to dissolve the relationship and pursue a fixed outside option. They introduce publicly observable shocks on firm's period profits, enlarge the state space of continuation utilities, and use a monotonic operator to recover it and the respective value function.

In this paper, I consider an environment similar to the one described in Sleet and Yeltekin (2001) where there is no profit perturbation, but the outside options are allowed to vary across (finite truncations of) the history of observables. The model is of the hidden action variety. A principal (firm) seeks a potentially infinite relationship with an agent (CEO) where the agent operates a stochastic technology transforming actions (effort) into outcomes (firm's output, revenues, or stock price). The outcomes are publicly observable, but the actions are not, so the contract should induce the proper incentives for the agent to exercise some desired sequence of actions. Both the principal and the agent cannot commit to long term relationships, so the contract should implicitly offer continuation utilities for both parties above their reservation utility values at any contingency which is actually reached. The contract may be optimally terminated at any contingency and in this case both parties will receive their respective reservation

utilities. This possibility effectively enters the set of possible strategies when signing the contract.

I characterize the contract recursively on the space of truncated histories of outcomes matched with agent's expected discounted utilities. The state space which is endogenous to the problem is further characterized in line with Abreu, Pearce and Stacchetti (1990). I parameterize and numerically compute the model in view of top executive compensation. The results show that the optimal contract estimated here is similar to the optimal incentive-compatible contract which is self-enforcing in the sense of Phelan (1995). Namely, the firm finds it optimal to induce high effort for most but the highest utility promises. Compensation and future utility promise increase in both current utility and firm's stock price. The contract partially smooths initial utility promises across initial price histories, which can be interpreted as a partial insurance against fluctuations in the manager's outside options.

The rest of the paper is structured as follows. Section 2 presents the dynamic model. Section 3 recursively characterizes the optimal contract. Section 4 computes the contract numerically and discusses the results. Section 5 concludes. The theory behind the characterization is developed in Appendix 1. The numerical results are presented in Appendix 2.

2 Dynamic model

The framework is as in Morfov (2008). A principal contracts an agent to implement a sequence of actions where the choice of an action each period is, in fact, a choice of an end-of-period probability distribution over outcomes. Since the exercise of an action brings disutility to the agent, he/she should be compensated by a monetary transfer from the principal. Everything is common knowledge but the particular action implemented which is only observed by the agent. The contract will, therefore, need to be incentive-compatible, i.e. to induce the agent to implement a particular action (or action sequence) recommended by the principal. Since both the principal and the agent are unable to commit to long term relationships, a long-term contract would require their (finite or infinite) participation. Unlike the analysis in Morfov (2008), here either party is allowed to dissolve a long-term relationship at contingencies where the contract fails to provide a continuation utility that is at least as high as this party's respective reservation utility At such a node, the relationship is terminated and both the principal and the agent consume their outside option, i.e. receive their reservation utilities. Notice the difference: before, we had the parties signing a self-enforcing contract which guaranteed their participation at every node, while here the contract ex-ante allows for termination at any node. As before, the reservation utilities are allowed to vary across some finite truncation of the history of past outcomes

Let Y, A, and W be the sets of possible outcomes, actions, and monetary transfers, which are all assumed compact subspaces of $\mathbb{R}$. In particular, $Y \in$

$\mathbb{R}_+$ is assumed finite with N distinct elements, a minimum element $\underline{y}$, and a maximum element $\overline{y}$. Let $\pi : Y \times A \to [\underline{\pi}, 1]$ describe the probability distribution of outcomes conditional on actions, where $\underline{\pi} \in (0,1)$ and $\pi(y|.)$ is continuous on A for $\forall y \in Y$. Denote by $u : W \times Y \to \mathbb{R}$ the (end-of-period) utility function of the principal which is assumed continuous, decreasing in the monetary transfer and increasing in the outcome. The utility function of the agent $\nu : W \times A \to \mathbb{R}$ is continuous, increasing in the monetary transfer and decreasing in the action. The principal and the agent discount future utility by discount factors β_P and β_A respectively, where $\beta_P, \beta_A \in (0,1)$. The reservation utilities of both parties depend on the previous θ outcomes, where θ is nonnegative integer. Formally $\underline{U} : Y^\theta \to \mathbb{R}$ is the the reservation utility of the principal and $\underline{V} : Y^\theta \to \mathbb{R}$ is the one of the agent. Let ${}^t y^\tau := (y_t, y_{t+1}, ...y^\tau)$ denote a particular sequence of outcomes realized between periods t and τ. Let ${}^t y := {}^t y^{-1}$ and $y^\tau := {}^0 y^\tau$. Let a supercontract $c := (h, a, w)$ be a plan of termination decisions, actions and compensation schemes defined on all possible contingencies. For example, at some history ${}^{-\theta} y^{t-1}$ observed in the beginning of period t, the contract terminates if $h_t\left({}^{-\theta} y^{t-1}\right) = 0$ or proceeds if $h_t\left({}^{-\theta} y^{t-1}\right) = 1$, recommends an action $a_t\left({}^{-\theta} y^{t-1}\right)$, and specifies a monetary transfer $w_t\left({}^{-\theta} y^{t-1}, y\right)$ contingent on an (end-of-period-t) outcome y for $\forall y \in Y$. Note that the supercontract is defined on every node of the tree of possible outcome histories, independent of whether the contract has been terminated before (including at) that node or not. We will refer to the supercontract as feasible when at every node ${}^{-\theta} y^{t-1}$, we have $h_t\left({}^{-\theta} y^{t-1}\right) \in \{0,1\}$, $a_t\left({}^{-\theta} y^{t-1}\right) \in A$, and $w_t\left({}^{-\theta} y^{t-1}, Y\right) \subset W$. Then, we can define the expected discounted utility of the principal at some node ${}^{-\theta} y^{\tau-1}$ given a feasible supercontract c as

$$U_\tau\left(c, {}^{-\theta} y^{\tau-1}\right) := \sum_{t=\tau}^{\infty} \beta_P^{t-\tau} \sum_{y_\tau \in Y} \cdots \sum_{y_{t-1} \in Y} [\left(1 - h_t\left({}^{-\theta} y^{t-1}\right)\right) \underline{U}_{t-\theta y^{t-1}} +$$
$$h_t\left({}^{-\theta} y^{t-1}\right) \sum_{y_t \in Y} u\left(w_t\left({}^{-\theta} y^{t-1}, y_t\right), y_t\right) \pi\left(y_t | a_t\left({}^{-\theta} y^{t-1}\right)\right)] \prod_{i=\tau}^{t-1} h_i\left({}^{-\theta} y^{i-1}\right) \pi\left(y_i | a_i\left({}^{-\theta} y^{i-1}\right)\right).$$

Given that everything is bounded, we will have that for any feasible contract c, the following holds:

$$U_\tau\left(c, {}^{-\theta} y^{\tau-1}\right) = \left(1 - h_\tau\left({}^{-\theta} y^{\tau-1}\right)\right) \underline{U}_{\tau-\theta y^{\tau-1}} + h_\tau\left({}^{-\theta} y^{\tau-1}\right) \sum_{y_\tau \in Y} [u\left(w_\tau\left({}^{-\theta} y^{\tau-1}, y_\tau\right), y_\tau\right) +$$
$$\beta_P U_{\tau+1}\left(c, {}^{-\theta} y^{\tau-1}, y_\tau\right)] \pi\left(y_\tau | a_\tau\left({}^{-\theta} y^{\tau-1}\right)\right).$$

Analogously, we can define $V_\tau\left(c, {}^{-\theta} y^{\tau-1}\right)$ as the expected discounted utility of the agent at node ${}^{-\theta} y^{\tau-1}$ and represent it recursively.

Then, the principal's problem at some initial node of contracting ${}^{-\theta} y$ (in the beginning of period 0) is:

[PPa]

$\sup_c U_0\left(c, {}^{-\theta}y\right)$ s.t.:

$$c \text{ feasible} \tag{1}$$

$$V_\tau\left(h, a, w, {}^{-\theta}y^{\tau-1}\right) \geq V_\tau\left(h, a', w, {}^{-\theta}y^{\tau-1}\right),$$
$$\forall \text{ feasible } a', \forall \textit{ non-terminal nai } \left({}^{-\theta}y\right) \tag{2}$$

$$V_\tau\left(c, \left({}^{-\theta}y^{\tau-\theta-1}, {}^{-\theta}\widetilde{y}\right)\right) \geq \underline{V}\left({}^{-\theta}\widetilde{y}\right), \forall nai\left({}^{-\theta}y\right) \tag{3}$$

$$U_\tau\left(c, \left({}^{-\theta}y^{\tau-\theta-1}, {}^{-\theta}\widetilde{y}\right)\right) \geq \underline{U}\left({}^{-\theta}\widetilde{y}\right), \forall nai\left({}^{-\theta}y\right) \tag{4}$$

where "$\forall nai\left({}^{-\theta}y\right)$" stays for "any *node after* and *including* ${}^{-\theta}y$", that is $\forall {}^{-\theta}\widetilde{y}^{\tau-1} = \left({}^{-\theta}y, \widetilde{y}^{\tau-1}\right) : \widetilde{y}^{\tau-1} \in Y^\tau, \forall \tau = 0, 1, 2, ...$ Here, we follow the convention that a function maximized over an empty set takes an arbitrarily low value, but assume that its multiplication with zero is well defined and is, in fact, 0. For example, if there does not exists a super contract that satisfies constraints (1)-(4), then the value function of [PPa] equals $\underline{U}\left({}^{-\theta}y\right)$ and the supremum is achieved at $h_0\left({}^{-\theta}y\right) = 0$, i.e. the contract is terminated at ${}^{-\theta}y$. Regarding the constraints, (1) was already discussed above, (3) and (4) are the individual rationality constraints[1] of the agent and, respectively, the principal guaranteeing them an expected discounted utility at every node at least as high as their respective reservation utility. Since, in general not all nodes will be reached, initial incentive-compatibility will not be equivalent to incentive compatibility at all nodes, just at the nodes actually reached. Therefore, from the start, we impose incentive compatibility at all nodes [constraint (2)] having in mind that what happens on nodes that are not actually reached is immaterial.

The difference from the problem analyzed in Morfov (2008), [PP], is that here the principal optimally chooses at the period of contracting when or if to terminate the contract.

3 Recursive Form

It is not difficult to prove[2] that for any feasible contract, incentive compatibility at all nodes is equivalent to Green (1987)'s temporary incentive compatibility at all nodes. A plan is temporary incentive compatible on a node if conditional

[1] These constraints are not referred to as participation constraints since they do not guarantee participation. Indeed, each party can receive his/her respective reservation utility by unilaterally terminating the relationship.

[2] Most of the theoretical results mentioned in this section are formally established in Appendix 1.

on future compliance, there is no deviation from the recommended action at this node that will make the agent strictly better off. Then, we can replace constraint (2) in [PPa] by:

$$\forall \textit{ non-terminal nai } \left({}^{-\theta}y\right), V_\tau\left(h, a, w, {}^{-\theta}y^{\tau-1}\right) \geq V_\tau\left(h, a', w, {}^{-\theta}y^{\tau-1}\right),$$

$$\forall a' : a'_\tau\left({}^{-\theta}y^{\tau-1}\right) \in A, \text{ and } \forall nai \ \left({}^{-\theta}y^{\tau-1}, Y\right), a'_t(.) = a_t(.) \tag{5}$$

Using the arguments of Morfov (2008), we can show that an optimal contract exists and it can be characterized recursively.

Since the reservation utilities depend on the previous θ outcome realizations, we will consider N^θ relevant initial histories. Let us extend the lexicographic order on the elements of Y to the elements of the set of possible initial histories Y^θ. Hereafter, all correspondences and functions with domain Y^θ will be considered as product vectors ordered by initial history.

Since the reservation utilities depend on the previous θ outcome realizations, we will consider N^θ relevant initial histories. Let us extend the lexicographic order on the elements of Y to the elements of the set of possible initial histories Y^θ. Hereafter, all correspondences and functions with domain Y^θ will be considered as product vectors ordered by initial history.

Let $\left\{V^{ICAPa}\right\}$ be the product vector of sets of possible expected discounted utilities for the agent signing a feasible incentive-compatible supercontract that is individually rational for the agent and allows for a termination at every node (an ICAPa contract). To be more precise, the supporting supercontracts should be feasible on the whole tree of possible histories, should satisfy agent's individual rationality[3] (3) at every node and (temporary) incentive compatibility (5) at all non-terminal nodes. Define $\left\{V^{IC2Pa}\right\}$ as the product vector of sets of possible expected discounted utilities for the agent signing a supercontract that triggers termination only at nodes where feasibility, temporary incentive-compatibility, agent's participation, or principal's participation (at least one of these constraints) is violated (an IC2Pa contract).[4]

For $\forall V \in \left\{V^{IC2Pa}\right\}$, let $U^{a^*}(V)$ be a product vector with a general element $U^{a^*}\left(V_{-\theta y}, {}^{-\theta}y\right)$ defined as the maximum utility the principal can get by signing an optimal supercontract of the second type at ${}^{-\theta}y$ offering $V_{-\theta y}$ to the agent. Let $\widehat{U}^{a^*}$ be the extension of U^{a^*} on $\{V^{ICAPa}\}$ s.t. for $\forall V \in \{V^{ICAPa}\}$, $\widehat{U}^{a^*}(V)$ is a product vector with a general element $\widehat{U}^{a^*}\left(V_{-\theta y}, {}^{-\theta}y\right) = U^{a^*}\left(V_{-\theta y}, {}^{-\theta}y\right)$ if

[3] Individual rationality should not be confused with participation. Terminations are practically allowed at every node.

[4] I am keeping some abbreviations from Morfov (2008), but note that here the difference is substantial. The ICAPa contract does not guarantee the participation of the agent at all. It simply allows termination at each contingency. The IC2Pa contract does not satisfy the participation constraints of both parties at all nodes. In fact, if any of the constraints (1), (3), (4), or (5) is violated, it triggers termination. That is, unlike the ICAPa contract, the IC2Pa one allows termination only when optimal.

$V_{-^{\theta}y} \in \{V^{IC2Pa}\left(^{-\theta}y\right)\}$ and $\widehat{U}^{a^*}\left(V_{-^{\theta}y},^{-\theta}y\right) := -\infty$ otherwise. Now, we define the operators used to characterize the optimal contract.

Let $\widehat{V} := \frac{\nu(\max\{W\},\min\{A\})}{1-\beta_A}$ and note that $\max\limits_{-^{\theta}y \in Y^{\theta}} \left\{\max\{V^{ICAPa}\left(^{-\theta}y\right)\}\right\} \leq \widehat{V}$. For $\forall X \in \mathbb{R}^{N^{\theta}}$ let $B^a(X)$ be a set operator such that $B^a_{-^{\theta}y}(X) := \underline{V}\left(^{-\theta}y\right) \cup \{V \in X_{-^{\theta}y} : \exists$ a (static) contract $c(V) = (a(V), w(V,Y), V_+(V,y))$ s.t.:

$$a(V) \in A \tag{6}$$

$$w(V,y) \in W, \forall y \in Y \tag{7}$$

$$\sum_{y \in Y} [\nu(w(V,y), a(V)) + \beta_A V_+(V,y)]\pi(y|a(V)) \geq$$
$$\sum_{y \in Y} [\nu(w(V,y), a'(V)) + \beta_A V_+\left(V_{-^{\theta}y}, y\right)]\pi(y|a'(V)), \forall a'(V) \in A \tag{8}$$

$$\sum_{y \in Y} [\nu(w(V,y), a(V)) + \beta_A V_+(V,y)]\pi(y|a(V)) = V \tag{9}$$

$$V_+(V,y) \in X_{-^{\theta+1}y,y} \tag{10}$$

hold}.

Note that for each initial utility level V, the static contract mentioned in the definition of the operator above specifies an action $a(V)$, a contingent transfer $w(V,.)$ from the principal to the agent and a contingent continuation utility for the agent $V_+(V,.)$, where the last two elements are defined on Y. Then, (6) and (7) are simply feasibility constraints, (8) imposes temporary incentive compatibility, (9) is a promise-keeping constraint. Note that the operator $\widetilde{B}^a$ does not impose individual rationality on part of the agent; it only requires the principal to provide him/her with a current and future utility levels from an initially specified set. Constraint (10) merely guarantees the consistency of the principal's promise. Nevertheless, iterating on $\widetilde{B}^a$ will allow us to recover the equilibrium set of utility promises at a low cost.[5] We just need to choose the proper initial guess, e.g. $\left\{\left[\underline{V}\left(^{-\theta}y\right), \widehat{V}\right]\right\}_{-^{\theta}y \in Y^{\theta}}$.

Now, for any continuos function $U : \left\{V^{ICAPa}\right\} \to \mathbb{R}$, let $T^a(U)$ be an operator such that

[5] See Appendix 1 for details.

$$T^a_{-\theta y}(U)_{\left(V_{-\theta y}\right)} :=$$

$$\max_{h\left(V_{-\theta y}\right)\in\{0,1\}} \{\left(1-h\left(V_{-\theta y}\right)\right)\underline{U}\left({}^{-\theta}y\right) + h\left(V_{-\theta y}\right)[\max_{c\left(V_{-\theta y}\right)}\{\sum_{y\in Y}[u\left(w\left(V_{-\theta y},y\right),y\right)+$$

$$\beta_P U_{-\theta+1 y,y}\left(V_+\left(V_{-\theta y},y\right)\right)]\pi\left(y|a\left(V_{-\theta y}\right)\right)\}\} \text{ s.t.:}$$

$$(6)-(9) \text{ hold, and}$$

$$V_+\left(V_{-\theta y},y\right) \in \left\{V^{ICAPa}\left({}^{-\theta+1}y,y\right)\right\},\ \forall y\in Y\} \tag{11}$$

This operator is defined on continuous functions with domain $\left\{V^{ICAPa}\right\}$. Since $\left\{V^{ICAPa}\right\}$ is compact, these functions are also bounded and form a complete metric space with the sup metric. Here, constraint (11) is the equivalent of the consistency constraint (10).

Once we have obtained $\left\{V^{ICAPa}\right\}$, the operator T^a can easily be shown to be a contraction.[6] Then, if we start with some initial guess for the value function, e.g. $U_0 = \left\{\underline{U}\left({}^{-\theta}y\right)\right\}_{-\theta y\in Y^\theta}$, and successively apply T^a, we will converge to $\widetilde{U}^{a^*}$, the product vector of maximum utilities the principal can derive from an IC2P contract that allows for inconsistent utility promises. How does the inconsistency enter the picture? Note that for any utility promise to the agent, the principal can (immediately) terminate the relationship if he/she finds proceeding it suboptimal. In general, this may happen for utility promises in $\left\{V^{ICAPa}\right\}$ such that not all their elements are agent's reservation utilities. Suppose, for example, that for some $V_{-\theta y}\in\left\{V^{ICAPa}\left({}^{-\theta}y\right)\right\}\diagdown\left\{\underline{V}\left({}^{-\theta}y\right)\right\}$, we have $\widetilde{U}^{a^*}\left(V_{-\theta y},{}^{-\theta}y\right) = \underline{U}\left({}^{-\theta}y\right)$ which is uniquely supported by for $h\left({}^{-\theta}y\right)=0$. This, however, would not be consistent with the promise to provide the agent with $V_{-\theta y}$ since the principal actually terminates the relationship and the agent only gets $\underline{V}\left({}^{-\theta}y\right) < V_{-\theta y}$. Moreover, $V_{-\theta y}$ may be a continuation utility of a unique single-round contract supporting some other initial utility promise in which case this promise would also prove inconsistent. Therefore, we may in fact be facing an "inconsistency cascade".

To deal with this problem, I use a procedure inspired by Rustichini (1998). For any upper-semicontinuos function bounded from above, $U:\left\{V^{ICAPa}\right\}\to\mathbb{R}^{N^\theta}$ and any function $H:\left\{V^{ICAPa}\right\}\to\{0,1\}^{N^\theta}$ let $\underline{T}^a(U,H)$ be an operator such that

[6] Note that we cold have alternatively defined the operator as follows:

$$T^a_{-\theta y}(U)_{\left(V_{-\theta y}\right)} := \max\{\underline{U}\left({}^{-\theta}y\right), \max_{c\left(V_{-\theta y}\right)}\{\sum_{y\in Y}[u\left(w\left(V_{-\theta y},y\right),y\right)+$$

$$\beta_P U_{-\theta+1 y,y}\left(V_+\left(V_{-\theta y},y\right)\right)]\pi\left(y|a\left(V_{-\theta y}\right)\right)\}\} \text{ s.t.:}$$

$$(6)-(9) \text{ and } (11) \text{ hold.}$$

$$\underline{T}^{a}_{-\theta y}(U,H)_{\left(V_{-\theta y}\right)} =$$

$$\begin{cases} (-\infty,0) \text{ if } V_{-\theta y} \in \{V^{ICAPa}\left({}^{-\theta}y\right)\} \diagdown \{\underline{V}\left({}^{-\theta}y\right)\} \text{ and } H\left(V_{-\theta y}, {}^{-\theta}y\right) = 0 \\ \left(T^{a}_{-\theta y}(U)_{\left(V_{-\theta y}\right)}, h\left(V_{-\theta y}\right)\right) \text{ otherwise,} \end{cases}$$

where $h\left(V_{-\theta y}\right)$ is such that the max in the definition of $T^{a}_{-\theta y}(U)_{\left(V_{-\theta y}\right)}$ is achieved.

Using this operator, we can recursively clean the inconsistent promises and modify the value function accordingly to obtain $\widehat{U}^{a^*}$. Then, we can recover U^{a^*} and $\left\{V^{IC2Pa}\right\}$.

4 Computation and Results

Since the model cannot be analyzed analytically, I resort to numerical methods. I focus on CEO compensation (i.e. the principal is a proxy for firm's shareholders, the agent is the company's top executive, the action is interpreted as effort, the monetary transfer as the manager's compensation package, and the outcomes as the company's stock price realizations). I parameterize the model as follows[7], i.e. $Y = \left\{y_{(1)}, y_{(2)}, y_{(3)}\right\} = \{0.55,\ 1.125,\ 1.7\}$, $A = \{\underline{a}, \overline{a}\} = \{0.1253,\ 0.1469\}$, $\pi\left(y_{(1)}|\underline{a}\right) = 0.1891$, $\pi\left(y_{(2)}|\underline{a}\right) = 0.7687$, $\pi\left(y_{(3)}|\underline{a}\right) = 0.0421$, $\pi\left(y_{(1)}|\overline{a}\right) = 0.1555$, $\pi\left(y_{(2)}|\overline{a}\right) = 0.7654$, $\pi\left(y_{(3)}|\overline{a}\right) = 0.0791$, $W = \left[0, \frac{y_{(3)}}{1-\beta_P}\right]$, $\beta_A = \beta_P = 0.96$, $u(w_t, y_t) = y_t - w_t$, $\nu(w_t, a_t) = \log(1 + w_t) - a_t$, $\theta = 1$, $\underline{U}(.) = 0$, $\underline{V}(.) \in$ {L,M,H}, where L := $\frac{\upsilon(\underline{w}) - \max A}{1-\beta_A}$ = -3.6725, M := 0, H := -L and assume a positive correlation between $\underline{V}(.)$ and Y. I also consider three cases: case 1a where the firm cannot offer the manager a compensation that exceeds its highest possible lifetime profit in discounted terms, i.e. $w_t \leq \frac{y_{(3)}}{1-\beta_P}$ (and $\widehat{V} = \frac{\upsilon\left(\frac{y_{(3)}}{1-\beta_P}\right) - \underline{a}}{1-\beta_A}$), case 2 where it can compensate the manager up to the highest possible stock price realization, i.e. $w_t \leq y_{(3)}$ (and $\widehat{V} = \frac{\upsilon\left(y_{(3)}\right) - \underline{a}}{1-\beta_A}$), and case 3 where it is essentially prohibited from borrowing, i.e. $w_t\left({}^{-\theta}y^{t-1}, y\right) \leq y, \forall y \in Y$ (where, as in case 2, we take $\widehat{V} = \frac{\upsilon\left(y_{(3)}\right) - \underline{a}}{1-\beta_A}$). Since the results are qualitatively similar, I focus on LMH, case 3, where LMH is the assignment of reservation values to initial histories, namely $\underline{V}\left(y_{(1)}\right)$ = L, $\underline{V}\left(y_{(2)}\right)$ = M, $\underline{V}\left(y_{(3)}\right)$ = H.

Figure 1 in the Appendix plots the value function for the optimal contract allowing for permanent separations (the IC2Pa contract). It is very similar to the value function of the optimal self-enforcing contract (the IC2P contract in the terminology of Morfov (2008) depicted together with the value function

[7] The parameterization is based on Aseff and Santos (2005) who take the value of low effort from Margiotta and Miller (2000) and use the results of Hall and Liebman (1998) to derive the stock price distribution conditional on manager's effort.

of the related augmented problem (the ICAP contract in the terminology of Morfov (2008)) on Figure 9. The main difference stems from the fact that the set of possible expected discounted utilities that can be offered to the manager is larger here than for the self-enforcing contract. Indeed, the principal can offer lower utility values to the agent supported by a future termination of the contract. These are exactly the values for which the value function fails to be monotonically decreasing, so these promises would be wiped out by renegotiation if such were possible. Note that here the partial insurance against fluctuations in the value of manager's outside options is less pronounced. This is because much lower utility promises are now supported.

Otherwise, the properties of the optimal contract allowing for separations (the IC2Pa contract) are not much different from those of the optimal self-enforcing contract (the IC2P contract). Figure 2 shows the optimal level of effort as a function of initial utility (cf. Figure 10 for the self-enforcing contract). High effort appears to be the predominant strategy for the principal except for managers endowed with higher initial utility who prove too costly to motivate. Executive pay is presented on the two dimensional Figure 3 (Figure 11 for the respective self-enforcing contract). As before, wage increases with the utility promise (compare Figure 4 with Figure 12; the graph for an outcome $y_{(3)}$ is kinked due to the binding borrowing constraint) and with the stock price realization (Figure 5 vs. Figure 13). Manager's continuation utility is plotted against initial utility and future outcome on Figure 6 (Figure 14 for the self-enforcing contract). It tends to increase in both current utility (cf. Figures 7 and 15) and firm's stock price (cf. Figures 8 and 16). The difference with the optimal self-enforcing contract is that now much lower continuation utilities are possible. The reason, as for the initial utility promises, is that they can be supported by future terminations.

5 Conclusion

In a dynamic model of hidden action with limited commitment and history-dependent reservation utilities, I recursively characterize the optimal contract allowing for permanent separations. I numerically compute the optimal contract for top executives and find little difference with the optimal self-enforcing contract. High effort appears optimal for most but the richest managers. CEO's compensation and future utility increase in both current utility and firm's stock price. The contract provides the manager with a lower level of insurance against fluctuations in the value of his/her outside options than does the respective self-enforcing contract.

References

[1] Abreu, D., Pearce, D. and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," Econometrica, 58, 1041-1063.

[2] Aseff, J. and M. Santos (2005): "Stock Options and Managerial Optimal Contracts," Economic Theory, 26 (4), 813-837.

[3] Atkeson, A. and R. Lucas (1995): "Efficiency and Equality in a Simple Model of Efficient Unemployment Insurance," Journal of Economic Theory, 66 (1), 64-88.

[4] Green, E. (1987): "Lending and the Smoothing of Uninsurable Income," in Contractual Agreements for Intertemporal Trade, ed. by E. C. Prescott and N. Wallace, Minneapolis: University of Minnesota Press, 3-25.

[5] Hall, B and J. Liebman (1998): "Are CEOs Really Paid like Bureaucrats?" Quarterly Journal of Economics, 111, 653–693.

[6] Margiotta, M and R. Miller (2000): "Managerial Compensation and the Cost of Moral Hazard," International Economic Review, 41, 669–719.

[7] Morfov, S. (2008): "Dynamic Moral Hazard with History-Dependent Participation Constraints," State University - Higher School of Economics, mimeo.

[8] Phelan, C. (1995): "Repeated Moral Hazard and One-Sided Commitment," Journal of Economic Theory, 66, 488-506.

[9] Phelan, C. and R. Townsend, (1991) "Computing Multi-Period, Information-Constrained Optima", Review of Economic Studies, 58 (5), 853-881.

[10] Rustichini, A. (1998): "Dynamic Programming Solution of Incentive Constrained Problems," Journal of Economic Theory, 78, 329-354.

[11] Sleet, C. and S. Yeltekin (2001): "Dynamic Labor Contracts with Temporary Layoffs and Permanent Separations," Economic Theory, 18, 207-235.

[12] Spear, S. and S. Srivastava (1987): "On Repeated Moral Hazard with Discounting," Review of Economic Studies, 54 (4), 599-617.

[13] Spear, S. and C. Wang (2005): "When to fire a CEO: optimal termination in dynamic contracts," Journal of Economic Theory, 120 (2), 239-256.

[14] Streufert, P. (1992): "An Abstract Topological Approach to Dynamic Programming," Journal of Mathematical Economics, 21 (1), 59-88.

[15] Thomas, J. and T. Worrall (1990): "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem," Journal of Economic Theory, 51 (2), 367-390.

APPENDIX 1

The first step towards a recursive representation is to establish that a feasible supercontract is incentive compatibility on all non-terminal nodes if and only if it is temporary incentive compatible on all non-terminal nodes.

Proposition 1 *For any c feasible, (2)⇔(5).*

Proof. It is trivial to show (2)⇒(5). In the other direction, let (5) hold, but assume that (2) is not satisfied, i.e. there is a non-terminal node ${}^{-\theta}y^{\tau-1}$ s.t. $\exists a'$ feasible on the tree with initial node ${}^{-\theta}y^{\tau-1}$ and $V_\tau\left(h, a', w, {}^{-\theta}y^{\tau-1}\right) > V_\tau\left(h, a, w, {}^{-\theta}y^{\tau-1}\right)$. We have:

$$V_\tau\left(h, a', w, {}^{-\theta}y^{\tau-1}\right) :=$$

$$\sum_{t=\tau}^{T}\beta_A^{t-\tau}\sum_{y_\tau\in Y}\cdots\sum_{y_{t-1}\in Y}[\left(1-h_t\left({}^{-\theta}y^{t-1}\right)\right)\underline{V}_{{}^{t-\theta}y^{t-1}}\prod_{i=\tau}^{t-1}h_i\left({}^{-\theta}y^{i-1}\right)\pi\left(y_i|a_i'\left({}^{-\theta}y^{i-1}\right)\right)+$$

$$\sum_{t=\tau}^{T}\beta_A^{t-\tau}\sum_{y_\tau\in Y}\cdots\sum_{y_t\in Y}v\left(w_t\left({}^{-\theta}y^{t-1}, y_t\right), a_t'\left({}^{-\theta}y^{t-1}\right)\right)\prod_{i=\tau}^{t}h_i\left({}^{-\theta}y^{i-1}\right)\pi\left(y_i|a_i'\left({}^{-\theta}y^{i-1}\right)\right)+$$

$$\beta_A^{T-\tau+1}V_{T+1}\left(h, a', w, {}^{-\theta}y^{\tau-1}\right)\prod_{i=\tau}^{T}h_i\left({}^{-\theta}y^{i-1}\right)\pi\left(y_i|a_i'\left({}^{-\theta}y^{i-1}\right)\right)$$

where the last term on the right-hand side can be made arbitrarily small by choosing T big enough. Therefore, $\exists T\in\mathbb{Z}_+$ and a feasible action plan a'' : $a_t''\left({}^{-\theta}y^{t-1}\right) = a_t'\left({}^{-\theta}y^{t-1}\right)$, $\forall {}^{-\theta}y^{t-1}\in {}^{-\theta}y^{\tau-1}\times Y^{t-\tau}$, $\forall t\leq T$, and $a_t''=a_t$ elsewhere, s.t. $V_\tau\left(h, a'', w, {}^{-\theta}y\right) > V_\tau\left(h, a, w, {}^{-\theta}y\right)$ Then, take $\tau'\in\mathbb{Z}_+ : \tau\leq\tau'\leq T$ s.t. $\exists {}^{-\theta}y^{\tau'-1} : a_{\tau'}''\left({}^{-\theta}y^{\tau'-1}\right)\neq a_{\tau'}\left({}^{-\theta}y^{\tau'-1}\right)$ and $\nexists\tau''\in\mathbb{Z}_{++} : \tau'<\tau''\leq T$: $a_{\tau''}''\left({}^{-\theta}y^{\tau''-1}\right)\neq a_{\tau''}\left({}^{-\theta}y^{\tau''-1}\right)$ for some ${}^{-\theta}y^{\tau''-1}\in {}^{-\theta}y^{\tau-1}\times Y^{\tau''-\tau}$. Define a feasible action plan a''' : $a_{\tau'}'''\left({}^{-\theta}y^{\tau'-1}\right) = a_{\tau'}\left({}^{-\theta}y^{\tau'-1}\right), \forall {}^{-\theta}y^{\tau'-1}\in {}^{-\theta}y^{\tau-1}\times Y^{\tau'-\tau}$ and $a_t'''=a_t''$ elsewhere. Then, for $\forall {}^{-\theta}y^{\tau'-1}\in {}^{-\theta}y^{\tau}\times Y^{\tau'-\tau}$ such that $h_{\tau'}\left({}^{-\theta}y^{\tau'-1}\right) = 0$, we have that $V_{\tau'}\left(h, a''', w, {}^{-\theta}y^{\tau'-1}\right) = V_{\tau'}\left(h, a'', w, {}^{-\theta}y^{\tau'-1}\right) = \underline{V}\left({}^{\tau'-\theta}y^{\tau'-1}\right)$, while if $h_{\tau'}\left({}^{-\theta}y^{\tau'-1}\right) = 1$, by (5) we obtain that $V_{\tau'}\left(h, a''', w, {}^{-\theta}y^{\tau'-1}\right)\geq V_{\tau'}\left(h, a'', w, {}^{-\theta}y^{\tau'-1}\right)$. Therefore, $V_\tau\left(h, a''', w, {}^{-\theta}y\right)\geq V_\tau\left(h, a'', w, {}^{-\theta}y\right)$. Proceeding in this way we can eliminate all the deviations (note that $\tau'\in\mathbb{Z}_+ : \tau'\leq T$) to obtain $V_\tau\left(h, a, w, {}^{-\theta}y\right)\geq V_\tau\left(h, a'', w, {}^{-\theta}y\right)$, i.e. a contradiction. ■

This operator is defined on continuous functions with domain $\left\{V^{ICAPa}\right\}$. Since $\left\{V^{ICAPa}\right\}$ is compact, these functions are also bounded and form a complete metric space with the sup metric. Here, constraint (11) is the equivalent of the consistency constraint (10).

Once we have obtained $\left\{V^{ICAPa}\right\}$ by iterating on the operator $\widetilde{B}^a$ starting with an initial guess X_0, the operator T^a can easily be shown to be a contraction.

For $\forall X \in \mathbb{R}^{N^\theta}$ let $B^a(X)$ be a set operator such that $B^a_{-^\theta y}(X) := \underline{V}\left(^{-\theta}y\right) \cup \{V \in \left[\underline{V}\left(^{-\theta}y\right), \widehat{V}\right] : \exists$ a (static) contract $c(V)$ s.t. (6)-(9) are satisfied and

$$V_+(V, y) \in X_{-^{\theta+1}y,y} \cap \left[\underline{V}\left(^{-\theta+1}y, y\right), \infty\right) \tag{12}$$

holds$\}$.

Compare this operator with $\widetilde{B}^a$. Here, $V \in \left[\underline{V}\left(^{-\theta}y\right), \widehat{V}\right]$ and constraint (12) combines individual rationality on part of the agent with a consistent continuation utility promise by the principal. The operator is more robust than $\widetilde{B}^a$ regarding the initial guess for the equilibrium set. However, given a suitably chosen initial guess, iterating on $\widetilde{B}^a$ leads to the same result at a lower computational cost.

For $\forall X = \left\{X_{-^\theta y}\right\}_{-^\theta y \in Y^\theta} : X_{-^\theta y} \in \mathbb{R}$, $\forall {}^{-\theta}y \in Y^\theta$ let $B^{a\prime}(X) := \left\{B^{a\prime}_{-^\theta y}(X)\right\}_{-^\theta y \in Y^\theta}$ with $B^{a\prime}_{-^\theta y}(X) := \{V_{-^\theta y} \in \left[\underline{V}\left(^{-\theta}y\right), \widehat{V}\right] : \exists c_R$:(6)-(9) and (10) hold at $\left(V_{-^\theta y}, {}^{-\theta}y\right)\}$. Note that the only difference between this operator and operator $\widetilde{B}^a$ defined in Section 3 is that $B^{a\prime}_{-^\theta y}(X) \subset \left[\underline{V}\left(^{-\theta}y\right), \widehat{V}\right]$, while $\widetilde{B}^a_{-^\theta y}(X) \subset X_{-^\theta y}$.

Lemma 1 $\left\{V^{ICAPa}\right\} \subset B^a\left(\left\{V^{ICAPa}\right\}\right)$.

Proof. Let $V \in \left\{V^{ICAPa}\right\}$ and fix an arbitrary ${}^{-\theta}y \in Y^\theta$. Since $V_{-^\theta y} \in \left\{V^{ICAPa}\left(^{-\theta}y\right)\right\}$, $\exists c$: (1) holds, (3) $\forall nai\left(^{-\theta}y\right)$, (5) $\forall$ non-terminal $nai\left(^{-\theta}y\right)$, and $V_0\left(c, {}^{-\theta}y\right) = V_{-^\theta y}$. By construction, $V_{-^\theta y} \in \left[\underline{V}\left(^{-\theta}y\right), \widehat{V}\right]$. If $h\left(^{-\theta}y\right) = 0$, node ${}^{-\theta}y$ is terminal, so $V_{-^\theta y} = \underline{V}\left(^{-\theta}y\right)$, and since $\underline{V}\left(^{-\theta}y\right) \in B^a_{-^\theta y}\left(\left\{V^{ICAPa}\right\}\right)$ by definition, the result is trivial. Therefore, let us assume $h\left(^{-\theta}y\right) = 1$. For $\forall y \in Y$, let $a_- := a_0\left(^{-\theta}y\right)$, $w_+(y) := w_0\left(^{-\theta}y, y\right)$, and $V_+(y) := V_1\left(c, \left(^{-\theta}y, y\right)\right)$. Given these choices, we immediately have that (9) holds. Moreover, (1) $\Rightarrow$ (6) $\cap$ (7), (5) $\Rightarrow$ (8). Note that for $\forall y \in Y$, $\left\{V^{ICAPa}\left(^{-\theta+1}y, y\right)\right\} \cap \left[\underline{V}\left(^{-\theta+1}y, y\right), +\infty\right) = V^{ICAPa}\left(^{-\theta+1}y, y\right)$. Note that for $\forall y \in Y$, the truncation of the original supercontract c to the tree with initial node $\left(^{-\theta}y, y\right)$ satisfies (1) on the tree with initial node $\left(^{-\theta+1}y, y\right)$ and is such that (3) $\forall nai\left(^{-\theta+1}y, y\right)$, (5) $\forall$ non-terminal

$nai\left({}^{-\theta+1}y,y\right)$ and $V_0\left(c_y,\left({}^{-\theta+1}y,y\right)\right)=V_1\left(c,\left({}^{-\theta}y,y\right)\right)$, which means that (12) is satisfied. Therefore, $V_{-\theta y}\in B^a_{-\theta y}\left(\left\{V^{ICAPa}\right\}\right)$. Since ${}^{-\theta}y\in Y^\theta$ was chosen randomly, this generalizes to $V\in B^a\left(\left\{V^{ICAPa}\right\}\right)$. ■

The lemma establishes that $\left\{V^{ICAPa}\right\}$ is self-generating in the terminology of Abreu, Pearce and Stacchetti (1990).

Lemma 2 *Assume* $X=\left\{X_{-\theta y}\right\}_{-\theta y\in Y^\theta}:X_{-\theta y}\subset B^a_{-\theta y}(X)$, $\forall{}^{-\theta}y\in Y^\theta$. *Then,* $B^a(X)\subset\left\{V^{ICAPa}\right\}$.

Proof. Let the condition of the lemma hold and take $V\in B^a(X)$. Fix an arbitrary ${}^{-\theta}y\in Y^\theta$. If $V_{-\theta y}=\underline{V}\left({}^{-\theta}y\right)$, it is immediate that $V_{-\theta y}\in\left\{V^{ICAPa}\left({}^{-\theta}y\right)\right\}$ since we can support it by a feasible supercontract with $h(.)=0$ on all $nai({}^{-\theta}y)$. If $V_{-\theta y}\neq\underline{V}\left({}^{-\theta}y\right)$, then $\exists c_{R,{}^{-\theta}y}\left(V_{-\theta y}\right)$: (6)-(12) hold at ${}^{-\theta}y$. By (12) and $X_{-\theta+1_{y,y}}\subset B^a_{-\theta+1_{y,y}}(X)$, we obtain that $V_{+,{}^{-\theta}y}\left(V_{-\theta y},y\right)\in B^a_{-\theta+1_{y,y}}(X)$, $\forall y\in Y$. Consider $y\in Y$. We either have that $V_{+,{}^{-\theta}y}\left(V_{-\theta y},y\right)$ can be supported by a contract $c_{R,(-\theta y,y)}\left(V_{+,{}^{-\theta}y}\left(V_{-\theta y},y\right)\right)$: (6)-(12) hold at $\left({}^{-\theta}y,y\right)$ or $V_{+,{}^{-\theta}y}\left(V_{-\theta y},y\right)=\underline{V}\left({}^{-\theta+1}y,y\right)\in\left\{V^{ICAPa}\left({}^{-\theta+1}y,y\right)\right\}$ since it can supported by a feasible supercontract with $h(.)=0$ on all $nai({}^{-\theta+1}y,y)$. Proceeding this way, we can consecutively construct a supercontract c after ${}^{-\theta}y$ s.t. (1), (3) $\forall nai\left({}^{-\theta}y\right)$, (5) $\forall$non-terminal $nai\left({}^{-\theta}y\right)$ and $V_0\left(c,{}^{-\theta}y\right)=V_{-\theta y}$. Here, it deserves noting that while (12) implies (3) on every node but the initial one, $V_{-\theta y}\in B^a_{-\theta y}(X)\subset\left[\underline{V}\left({}^{-\theta}y\right),\widehat{V}\right]$, from where (3) is also satisfied at ${}^{-\theta}y$. Therefore, $V_{-\theta y}\in\left\{V^{ICAPa}\left({}^{-\theta}y\right)\right\}$, which generalizes to $V\in\left\{V^{ICAP}\right\}$. ■

The lemma says that the image of every nonempty, self-generating set is a subset of $\left\{V^{ICAPa}\right\}$.

Proposition 2 *(a)* $B^a\left(\left\{V^{ICAPa}\right\}\right)=\left\{V^{ICAP}\right\}$*; and (b) if* $\exists X\subset\mathbb{R}^{N^\theta}$: $B^a(X)=X$*, then* $X\subset\left\{V^{ICAPa}\right\}$.

Proof of Proposition 2. (a) From Lemmas 1 and 2.
(b) It follows by Lemma 2. ■

This proposition establishes that the set of agent's expected discounted utilities supportable by an ICAPa supercontract is the largest fixed point of B^a.

Lemma 3 *Assume* $X'=\left\{X'_{-\theta y}\right\}_{-\theta y\in Y^\theta}$ *and* $X''=\left\{X''_{-\theta y}\right\}_{-\theta y\in Y^\theta}:X'_{-\theta y}\subset X''_{-\theta y}\subset\mathbb{R}$, $\forall{}^{-\theta}y\in Y^\theta$. *Then,* $B^a_{-\theta y}(X')\subset B^a_{-\theta y}(X'')$, $\forall{}^{-\theta}y\in Y^\theta$.

Proof. Trivial. ■

Lemma 4 *Assume* $X = \left\{X_{-^\theta y}\right\}_{-^\theta y \in Y^\theta}$*:* $X_{-^\theta y} \subset \mathbb{R}$ *compact,* $\forall -^\theta y \in Y^\theta$*. Then,* $B^a_{-^\theta y}(X)$ *compact,* $\forall -^\theta y \in Y^\theta$*.*

Proof. Let the condition of the lemma hold and assume $B^a_{-^\theta y}(X) \neq \emptyset$ for some $-^\theta y \in Y^\theta$. Note that $B^a_{-^\theta y}(X) \subset \left[\underline{V}\left(-^\theta y\right), \widehat{V}\right] \subset \mathbb{R}$ is bounded by definition. We should also show that it is closed. Take an arbitrary convergent sequence $\{V_i\}_{i=1}^\infty : V_i \in B^a_{-^\theta y}(X)$, $\forall i \in \mathbb{Z}_{++}$ with $V_i \underset{i\to\infty}{\to} V_\infty$. We need to prove that $V_\infty \in B^a_{-^\theta y}(X)$. If $V_\infty = \underline{V}\left(-^\theta y\right)$, then the result is trivial since $\underline{V}\left(-^\theta y\right) \in B^a_{-^\theta y}(X)$. Therefore, assume that $V_\infty \neq \underline{V}\left(-^\theta y\right)$. Then, it should be the case that there exists a subsequence $\left\{V_{i_j}\right\}_{j=1}^\infty$ of $\{V_i\}_{i=1}^\infty$, such that for every positive integer j, $V_{i_j} \in \left[\underline{V}\left(-^\theta y\right), \widehat{V}\right]$ and $\exists c_{R,i_j}$:(6)-(12) hold at $\left(V_{i_j}, -^\theta y\right)$. By $V_{i_j} \in \left[\underline{V}\left(-^\theta y\right), \widehat{V}\right]$, $\forall$ positive integer j, we obtain $V_\infty \in \left[\underline{V}\left(-^\theta y\right), \widehat{V}\right]$. By (6), (7), (12), Y finite, and $X_{-^\theta y} \subset \mathbb{R}$ compact for $\forall -^\theta y \in Y^\theta$, we have that $\left\{c_{R,i_j}\right\}_{j=1}^\infty$ is uniformly bounded, therefore $\exists$ a subsequence $\left\{c_{R,i_{j_k}}\right\}_{k=1}^\infty$ of $\left\{c_{R,i_j}\right\}_{j=1}^\infty : c_{R,i_{j_k}} \underset{k\to\infty}{\to} c_{R,\infty}$. It is immediate that $c_{R,\infty}$ satisfies (6)-(12) at $\left(V_\infty, -^\theta y\right)$. ■

Proposition 3 *Let* X_0 *compact* : $\left\{V^{ICAPa}\right\} \subset X_0 \subset \mathbb{R}^{N^\theta}$ *and* $B^a(X_0) \subset X_0$*. Define* $X_{i+1} := B^a(X_i)$ *for* $\forall i \in \mathbb{Z}_+$*. Then,* $X_{i+1} \subset X_i$*,* $\forall i \in \mathbb{Z}_+$ *and* $X_\infty := \lim_{i\to\infty} X_i = \left\{V^{ICAPa}\right\}$*.*

Proof of Proposition 3. For $\forall -^\theta y \in Y^\theta$ and $\forall i \in \mathbb{Z}_+$, denote by $X_i\left(-^\theta y\right)$ the element of X_i corresponding to initial history $-^\theta y$. By the condition of the Proposition, we have $\left\{V^{ICAPa}\left(-^\theta y\right)\right\} \subset X_0\left(-^\theta y\right) \subset \mathbb{R}$, $\forall -^\theta y \in Y^\theta$. Since by Proposition 2 (a) $B^a\left(\left\{V^{ICAPa}\right\}\right) = \left\{V^{ICAPa}\left(-^\theta y\right)\right\}$, we can apply Lemma 3 to obtain $\left\{V^{ICAPa}\left(-^\theta y\right)\right\} \subset X_1\left(-^\theta y\right) \subset \mathbb{R}$, $\forall -^\theta y \in Y^\theta$. Using $X_1 \subset X_0$ and repeating the argument, we reach $\left\{V^{ICAP}\right\} \subset X_{i+1} \subset X_i$, $\forall i \in \mathbb{Z}_+$ Then, $\{X_i\}_{i=0}^\infty$ is a sequence of non-empty, compact (by Lemma 4 since X_0 compact), monotonically decreasing (nested) sets; therefore it converges to $X_\infty = \bigcap_{i=0}^\infty X_i \supset \left\{V^{ICAPa}\right\}$ with X_∞ compact.

What remains to be shown is that $X_\infty \subset \left\{V^{ICAPa}\right\}$. By Lemma 2, it is enough to show that $X_\infty \subset B^a(X_\infty)$. Let $V \in X_\infty$. This implies that $V \in X_i$, $\forall i \in \mathbb{Z}_+$. Fix an arbitrary $-^\theta y \in Y^\theta$. If $V_{-^\theta y} = \underline{V}\left(-^\theta y\right)$, we immediately have that $V_{-^\theta y} \in B^a_{-^\theta y}(X_\infty)$,.therefore assume $V_{-^\theta y} \neq \underline{V}\left(-^\theta y\right)$.Then, $\exists c_{R,i}$:

(6)-(12) hold at $\left(V_{-^{\theta}y},{}^{-\theta}y\right)$. By (6), (7), (12), Y finite, and $X_i \subset X_0 \subset \mathbb{R}^{N^{\theta}}$ compact, $\forall i \in \mathbb{Z}_+$, we have that $\{c_{R,i}\}_{i\in\mathbb{Z}_+}$ is uniformly bounded; therefore, $\exists$ a subsequence $\left\{c_{R,i_j}\right\}_{j=0}^{\infty}$ of $\{c_{R,i}\}_{i=0}^{\infty} : c_{R,i_j} \underset{j\to\infty}{\to} c_{R,\infty}$. It is immediate that $c_{R,\infty}$ satisfies(6)-(8) at $\left(V_{-^{\theta}y},{}^{-\theta}y\right)$. Moreover, $V_{+,\infty}(y) \geq \underline{V}_{-^{\theta+1}y,y}$, $\forall y \in Y$. We also need to show that for $\forall y \in Y$, $V_{+,\infty}(y) \in X_{\infty}\left({}^{-\theta+1}y, y\right)$. Fix an arbitrary $y \in Y$ and assume, on the contrary, that $V_{+,\infty}(y) \notin X_{\infty}\left({}^{-\theta+1}y, y\right)$. Since $X_{\infty}\left({}^{-\theta+1}y, y\right) = \bigcap_{i=0}^{\infty} X_i\left({}^{-\theta+1}y, y\right) = \bigcap_{j=0}^{\infty} X_{i_j}\left({}^{-\theta+1}y, y\right)$, we have that $\exists j' \in \mathbb{Z}_+ : V_{+,\infty}(y) \notin X_{i_{j'}}\left({}^{-\theta+1}y, y\right)$. Furthermore, $\left\{X_{i_j}\right\}_{j=0}^{\infty}$ was shown to be a monotonically decreasing (nested) sequence, from where $V_{+,i_j}(y) \in X_{i_j}\left({}^{-\theta+1}y, y\right) \subset X_{i_{j'}}\left({}^{-\theta+1}y, y\right)$, $\forall j \in \mathbb{Z}_+ : j \geq j'$. Since $X_{i_{j'}}\left({}^{-\theta+1}y, y\right)$ is closed and $V_{+,i_j}(y) \underset{k\to\infty}{\to} V_{+,\infty}(y)$, we obtain that $V_{+,\infty}(y) \in X_{i_{j'}}\left({}^{-\theta+1}y, y\right)$, i.e. a contradiction is reached. This proves $V_{+,\infty}(y) \in X_{\infty}\left({}^{-\theta+1}y, y\right)$, $\forall y \in Y$. Consequently, (12) holds for $c_{R,\infty}$. Finally, note that $V_{-^{\theta}y} \in \left[\underline{V}\left({}^{-\theta}y\right), \widehat{V}\right]$ follows immediately from $V_{-^{\theta}y} \in X_1\left({}^{-\theta}y\right)$. Therefore, $V_{-^{\theta}y} \in B^{a}_{-^{\theta}y}(X_{\infty})$, which generalizes to $V \in B^{a}(X_{\infty})$. ■

The proposition says that if we start iterating on B^a taking as an initial guess some compact set X_0 that contains both $B^a(X_0)$ and $\left\{V^{ICAPa}\right\}$, we will ultimately converge to the largest fixed point of the operator, $\left\{V^{ICAPa}\right\}$. This is sufficient for obtaining $\left\{V^{ICAPa}\right\}$ since we can always take $X_0 = \left\{X_{-^{\theta}y}\right\}_{-^{\theta}y\in Y^{\theta}} : \left[\underline{V}\left({}^{-\theta}y\right), \widehat{V}\right] \subset X_{-^{\theta}y} \subset \mathbb{R}$ with $X_{-^{\theta}y}$ compact, $\forall {}^{-\theta}y \in Y^{\theta}$. However, an even more computationally efficient result exists.

For $\forall X = \left\{X_{-^{\theta}y}\right\}_{-^{\theta}y\in Y^{\theta}} : X_{-^{\theta}y} \in \mathbb{R}$, $\forall {}^{-\theta}y \in Y^{\theta}$ let $B^{a\prime}(X) := \left\{B^{a\prime}_{-^{\theta}y}(X)\right\}_{-^{\theta}y\in Y^{\theta}}$ with $B^{a\prime}_{-^{\theta}y}(X) := \{V_{-^{\theta}y} \in \left[\underline{V}\left({}^{-\theta}y\right), \widehat{V}\right] : \exists c_R$:(6)-(9) and (10) hold at $\left(V_{-^{\theta}y},{}^{-\theta}y\right)\}$. Note that the only difference between this operator and operator $\widetilde{B}^a$ defined in Section 3 is that $B^{a\prime}_{-^{\theta}y}(X) \subset \left[\underline{V}\left({}^{-\theta}y\right), \widehat{V}\right]$, while $\widetilde{B}^{a}_{-^{\theta}y}(X) \subset X_{-^{\theta}y}$.

Lemma 5 *Take* $X_0' := \left\{X_0'\left({}^{-\theta}y\right)\right\}_{-^{\theta}y\in Y^{\theta}}$ *with* $X_0'\left({}^{-\theta}y\right) := \left[\underline{V}\left({}^{-\theta}y\right), \widehat{V}\right]$, $\forall {}^{-\theta}y \in Y^{\theta}$ *and let* $X_{i+1}' := B^{a\prime}(X_i')$ *for* $\forall i \in \mathbb{Z}_+$. *Then,* $X_{i+1}' \subset X_i'$, $\forall i \in \mathbb{Z}_+$ *and* $X_{\infty}' := \lim_{i\to\infty} X_i' = \left\{V^{ICAPa}\right\}$.

Proof. We have that X_0' is compact and $\left\{V^{ICAPa}\right\} \subset X_0' \subset \mathbb{R}^{N^{\theta}}$. Note that for $\forall X \subset \mathbb{R}^{N^{\theta}}$, we have $B^{a}_{-^{\theta}y}(X) \subset B^{a\prime}_{-^{\theta}y}(X)$. Then, by Lemma 3 and Proposition 2 (a), we obtain $\left\{V^{ICAPa}\right\} \subset B^{a}(X_0') \subset B^{a\prime}(X_0')$. Using the same arguments plus the monotonicity of $B^{a\prime}$ (trivial), we have $\left\{V^{ICAPa}\right\} \subset X_i^{a\prime}$,

$\forall i \in \mathbb{Z}_+$. Moreover, by construction $B^{a\prime}(X_0') \subset X_0'$. Then, the condition $B(X_0') \subset X_0'$ is satisfied. Observe that for $\forall {}^{-\theta}y \in Y^\theta$, $X_1'({}^{-\theta}y) = \underline{V}({}^{-\theta}y) \cap \{V_{-\theta_y} \in \left[\underline{V}({}^{-\theta}y), \widehat{V}\right] : \exists c_R$ s.t. (6)-(9), (10) hold at $(V_{-\theta_y}, {}^{-\theta}y)\} = \underline{V}({}^{-\theta}y) \cap \{V_{-\theta_y} \in \left[\underline{V}({}^{-\theta}y), \widehat{V}\right] : \exists c_R$ s.t. (6)-(9), (12) hold at $(V_{-\theta_y}, {}^{-\theta}y)\} = B^a_{-\theta_y}(X_0')$ since, by construction, we have that $X_0'({}^{-\theta+1}y, y) \cap \left[\underline{V}({}^{-\theta+1}y, y), +\infty\right) = X_0'({}^{-\theta+1}y, y)$, $\forall y \in Y$. Furthermore, by $X_1' \subset X_0'$ and the monotonicity of B', we obtain $X_{i+1}' \subset X_i'$, $\forall i \in \mathbb{Z}_+$. Then, it is trivial that $X_{i+1}' = B^a(X_i')$, $\forall i \in \mathbb{Z}_+$ Therefore, Proposition 3 applies to $\{X_i'\}_{i=1}^\infty$. ■

Lemma 6 *Let $\{X_i'\}_{i=1}^\infty$ be defined as in Lemma 5. Take $\widetilde{X}_0 := X_0'$ and let $\widetilde{X}_{i+1} := \widetilde{B}^a\left(\widetilde{X}_i\right)$ for $\forall i \in \mathbb{Z}_+$. Then, $\widetilde{X}_i = X_i'$, $\forall i \in \mathbb{Z}_+$.*

Proof. Assume $\widetilde{X}_{i-1} = X_{i-1}'$ for some $i \in \mathbb{Z}_{++}$. By Lemma 5, $\emptyset \neq X_i' \subset X_{i-1}'$. Fix ${}^{-\theta}y \in Y^\theta$ and let $V \in X_i'({}^{-\theta}y)$ Then, we have $V \in \widetilde{X}_{i-1}({}^{-\theta}y)$, which together with $V \in B'_{-\theta_y}\left(\widetilde{X}_{i-1}\right)$ implies that $V \in \widetilde{B}_{-\theta_y}\left(\widetilde{X}_{i-1}\right)$. Since ${}^{-\theta}y$ and V were chosen randomly, this generalizes to $X_i' \subset \widetilde{X}_i$. Note that $\widetilde{X}_{i-1} = X_{i-1}' \subset X_0'$ by Lemma 5. Consequently, $\widetilde{B}_{-\theta_y}\left(\widetilde{X}_{i-1}\right) \subset B'_{-\theta_y}\left(\widetilde{X}_{i-1}\right)$, i.e. $\widetilde{X}_i \subset X_i'$.

We have that $\widetilde{X}_0 = X_0'$ by definition and have just shown that $\widetilde{X}_{i-1} = X_{i-1}'$ would imply $\widetilde{X}_i = X_i'$; therefore, by induction we obtain that $\widetilde{X}_i = X_i'$ for $\forall i \in \mathbb{Z}_+$. ■

Proposition 4 *(a) Take $\widetilde{X}_0 := \left\{\widetilde{X}_0({}^{-\theta}y)\right\}_{-\theta_y \in Y^\theta}$ with $\widetilde{X}_0({}^{-\theta}y) = \left[\underline{V}({}^{-\theta}y), \widehat{V}\right]$, $\forall {}^{-\theta}y \in Y^\theta$ and let $\widetilde{X}_{i+1} := \widetilde{B}^a\left(\widetilde{X}_i\right)$ for $\forall i \in \mathbb{Z}_+$. Then, $\widetilde{X}_{i+1} \subset \widetilde{X}_i$, $\forall i \in \mathbb{Z}_+$ and $\widetilde{X}_\infty := \lim_{i\to\infty} \widetilde{X}_i = \left\{V^{ICAPa}\right\}$. (b) $\widetilde{B}^a\left(\left\{V^{ICAPa}\right\}\right) = \left\{V^{ICAPa}\right\}$; and (c) if $\exists X \subset \widetilde{X}_0 : \widetilde{B}^a(X) = X$, then $X \subset \left\{V^{ICAPa}\right\}$.*

Proof of Proposition 4. (a) From Lemmas 5 and 6.

(b) Similarly to the proof of Lemma 1, we can show that $\left\{V^{ICAP}\right\} \subset \widetilde{B}_{-\theta_y}\left(\left\{V^{ICAP}\right\}\right)$. It is trivial that $\widetilde{B}_{-\theta_y}\left(\left\{V^{ICAP}\right\}\right) \subset \left\{V^{ICAP}\right\}$.

(c) Since $X \subset \widetilde{X}_0$, we can use the monotonicity of $\widetilde{B}^a$ and $\widetilde{B}^a(X) = X$ to obtain $X \subset \widetilde{X}_i$, $\forall i \in \mathbb{Z}_+$. Then, by (a), we have $X \subset \widetilde{X}_\infty = \left\{V^{ICAPa}\right\}$. ■

This proposition outlines a practical way of obtaining $\left\{V^{ICAPa}\right\}$. Namely, we start with the set $\left\{\left[\underline{V}({}^{-\theta}y), \widehat{V}\right]\right\}_{-\theta_y \in Y^\theta}$ and iterate on the set operator $\widetilde{B}^a$ until convergence in a properly defined sense is attained.

APPENDIX 2

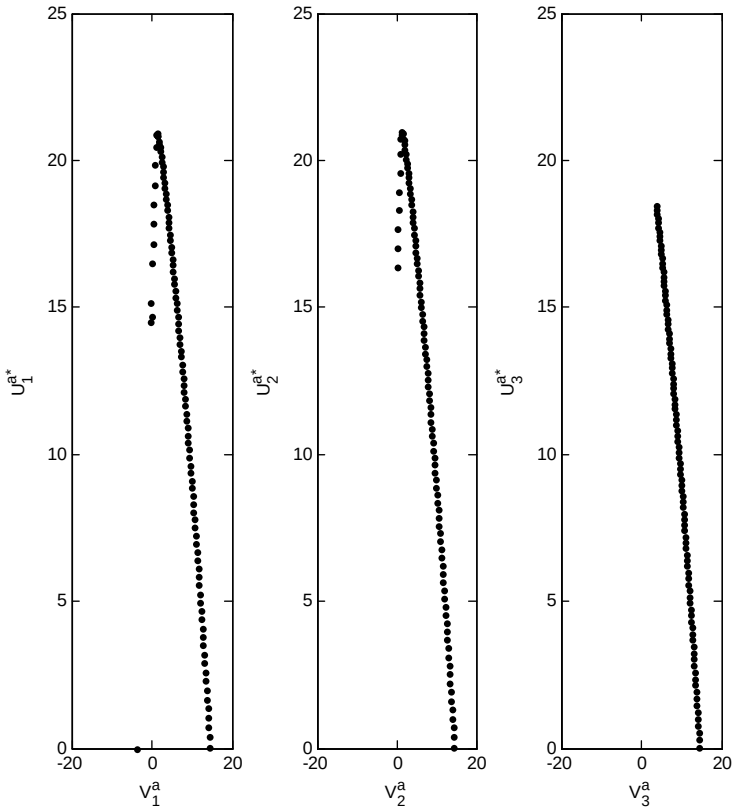

Figure 1: Value functions for the IC2Pa contract ordered by initial stock-price history: $U^{a*}(., y_i)$, $i \in \{1, 2, 3\}$ (LMH, case 3)

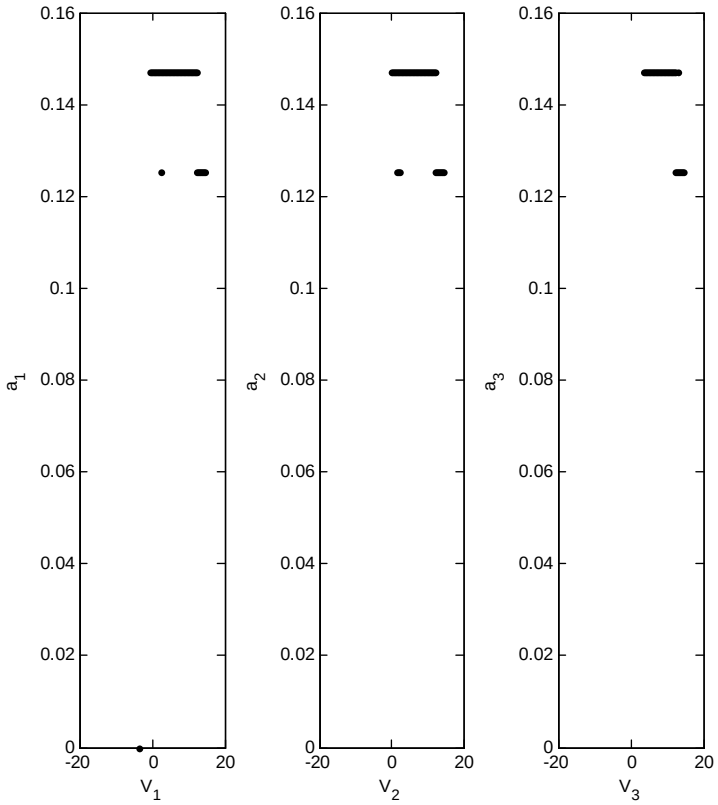

Figure 2: Optimal effort under the IC2Pa contract as a function of initial utility promise: $a_{y_i}^{a^*}$, $i \in \{1, 2, 3\}$ (LMH, case 3)

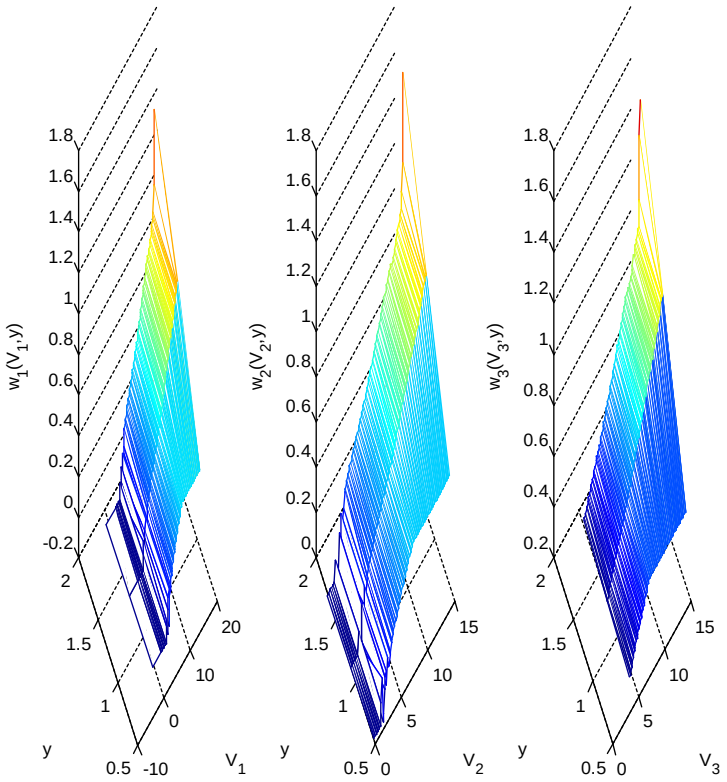

Figure 3: Optimal wage under the IC2Pa contract as a function of initial utility promise and future stock price: $w_{y_i}^{a^*}\left(V_{y_i}^a, y_k\right)$: $V_{y_i}^a \in \{V^a\left(y_i\right)\}$, $i, k \in \{1, 2, 3\}$ (LMH, case 3)

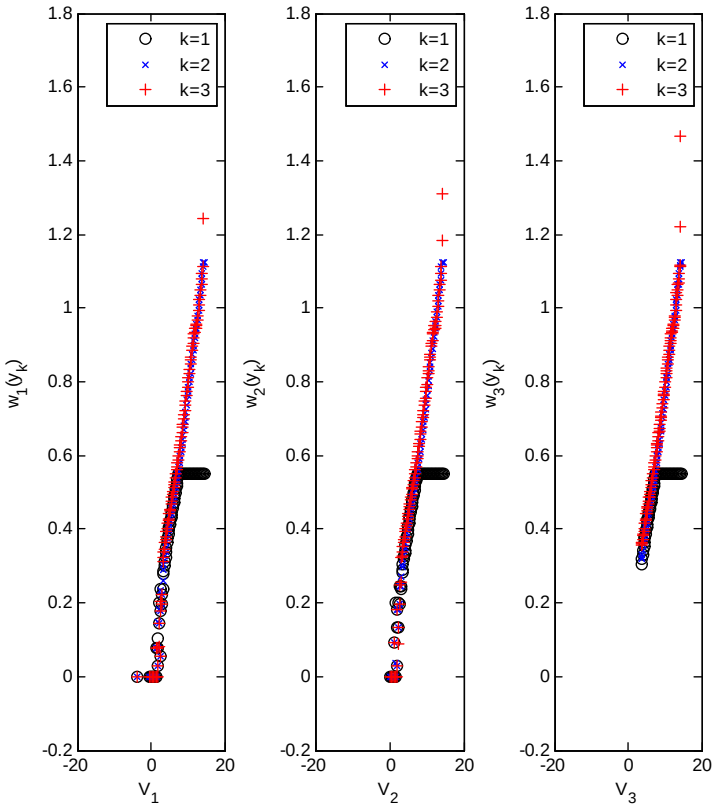

Figure 4: Optimal wage under the IC2Pa contract as a function of initial utility promise: $w_{y_i}^{a^*}(., y_k)$, $i, k \in \{1, 2, 3\}$ (LMH, case 3)

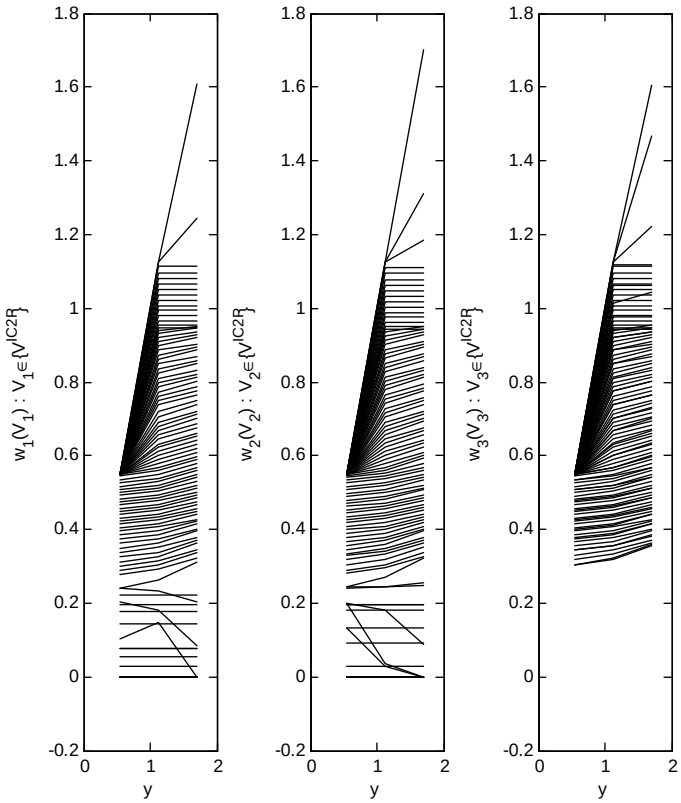

Figure 5: Optimal wage under the IC2Pa contract as a function of future stock price: $w_{y_i}^{a^*}\left(V_{y_i}^a, .\right)$: $V_{y_i}^a \in \{V^a\left(y_i\right)\}$, $i \in \{1, 2, 3\}$ (LMH, case 3)

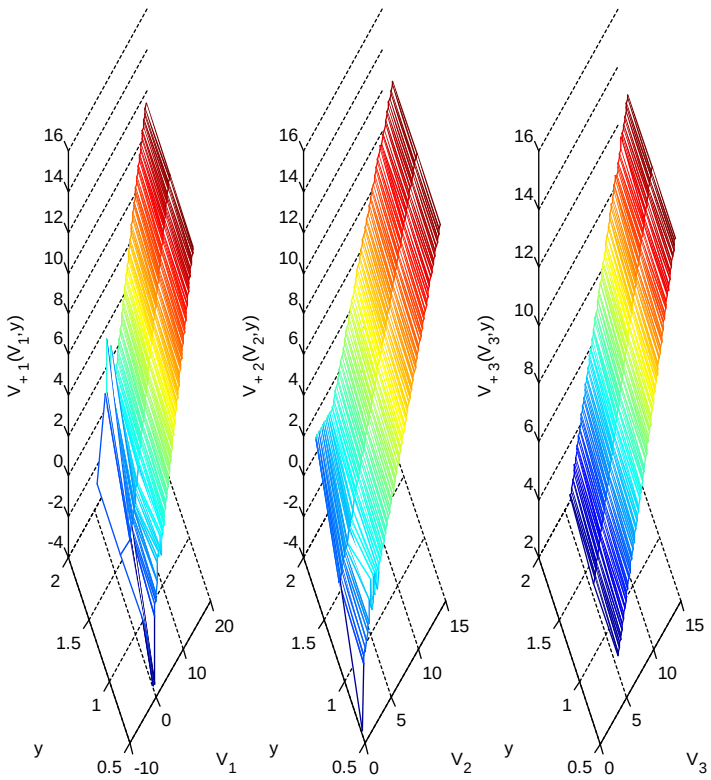

Figure 6: Optimal future utility promise under the IC2Pa contract as a function of initial utility promise and future stock price: $V^{a^*}_{+\,y_i}\left(V^a_{y_i}, y_k\right)$: $V^a_{y_i} \in \{V^a\left(y_i\right)\}$, $i, k \in \{1, 2, 3\}$ (LMH, case 3)

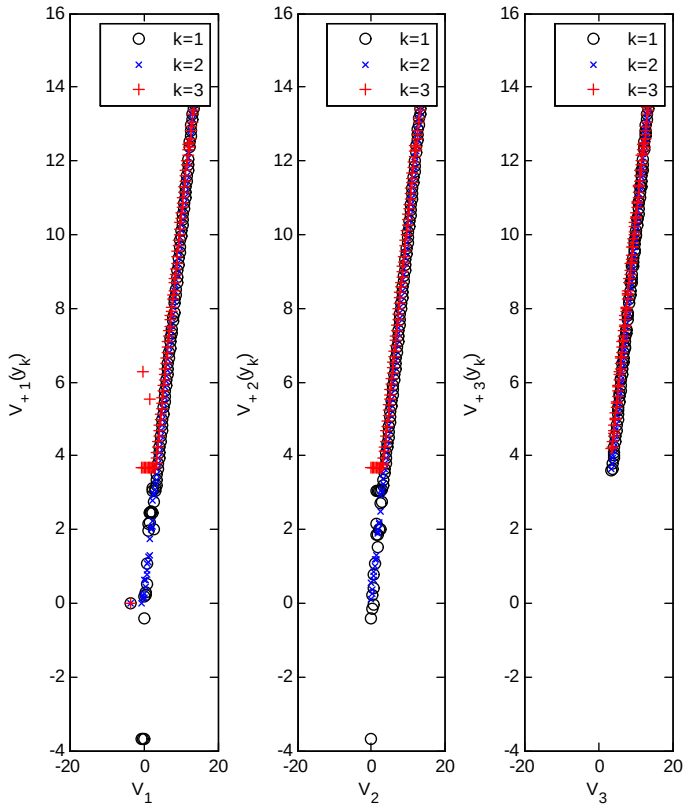

Figure 7: Optimal future utility promise under the IC2Pa contract as a function of initial utility promise: $V^{a^*}_{+\ y_i}(., y_k)$, $i, k \in \{1, 2, 3\}$ (LMH, case 3)

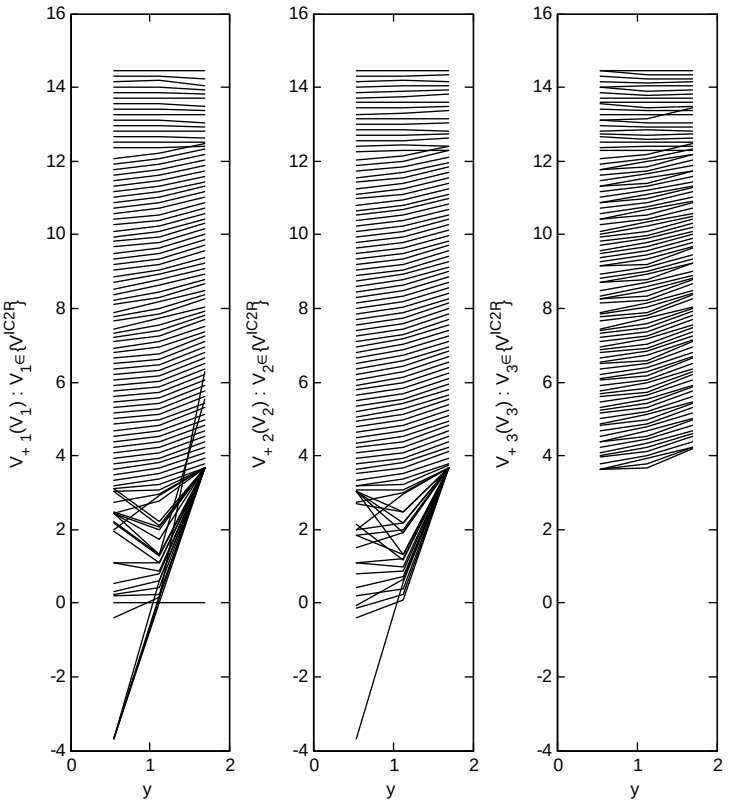

Figure 8: Optimal future utility promise under the IC2Pa contract as a function of future stock price: $V^{a^*}_{+\,y_i}\left(V^a_{y_i},.\right)$: $V^a_{y_i} \in \{V^a\left(y_i\right)\}$, $i \in \{1,2,3\}$ (LMH, case 3)

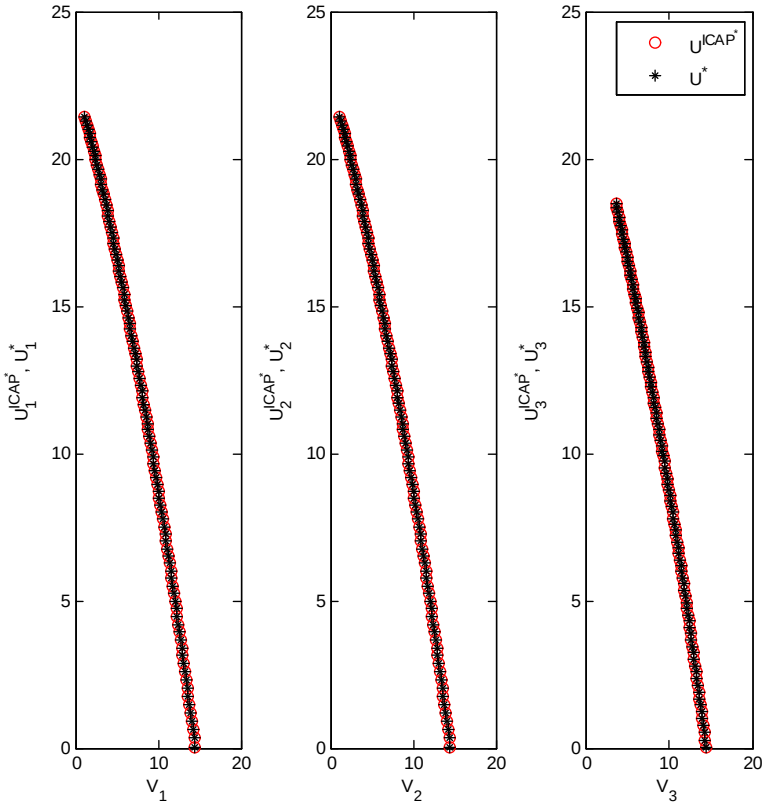

Figure 9: Value functions for the ICAP and IC2P contracts ordered by initial stock-price history: $U^{ICAP^*}(., y_i)$, $U^*(., y_i)$, $i \in \{1, 2, 3\}$ (LMH, case 3)

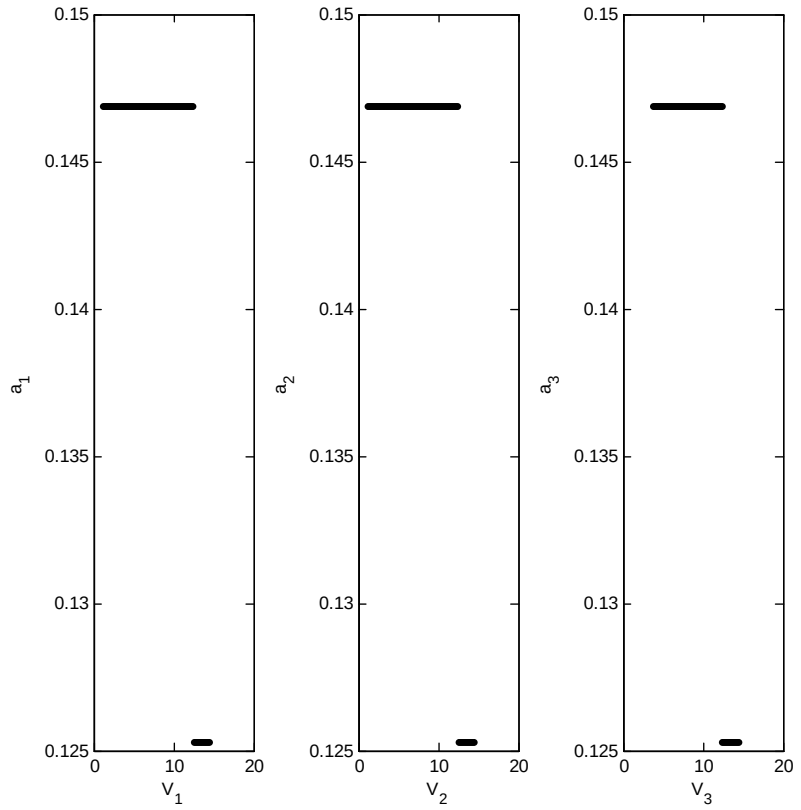

Figure 10: Optimal effort under the IC2P contract as a function of initial utility promise: $a^*_{y_i}$, $i \in \{1, 2, 3\}$ (LMH, case 3)

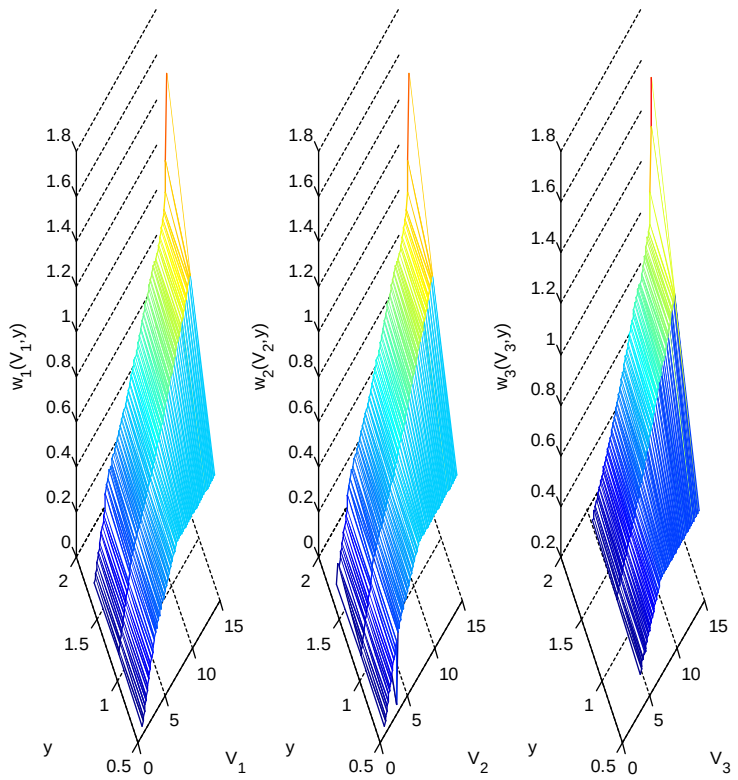

Figure 11: Optimal wage under the IC2P contract as a function of initial utility promise and future stock price: $w^*_{y_i}(V_{y_i}, y_k)$: $V_{y_i} \in \{V^{IC2P}(y_i)\}$, $i, k \in \{1, 2, 3\}$ (LMH, case 3)

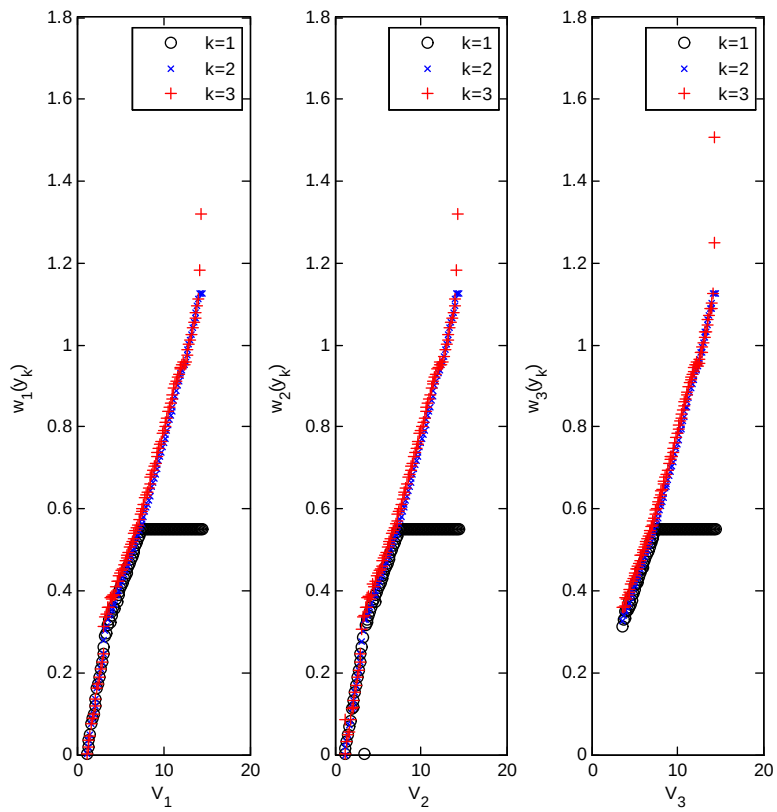

Figure 12: Optimal wage under the IC2P contract as a function of initial utility promise: $w^*_{y_i}(., y_k)$, $i, k \in \{1, 2, 3\}$ (LMH, case 3)

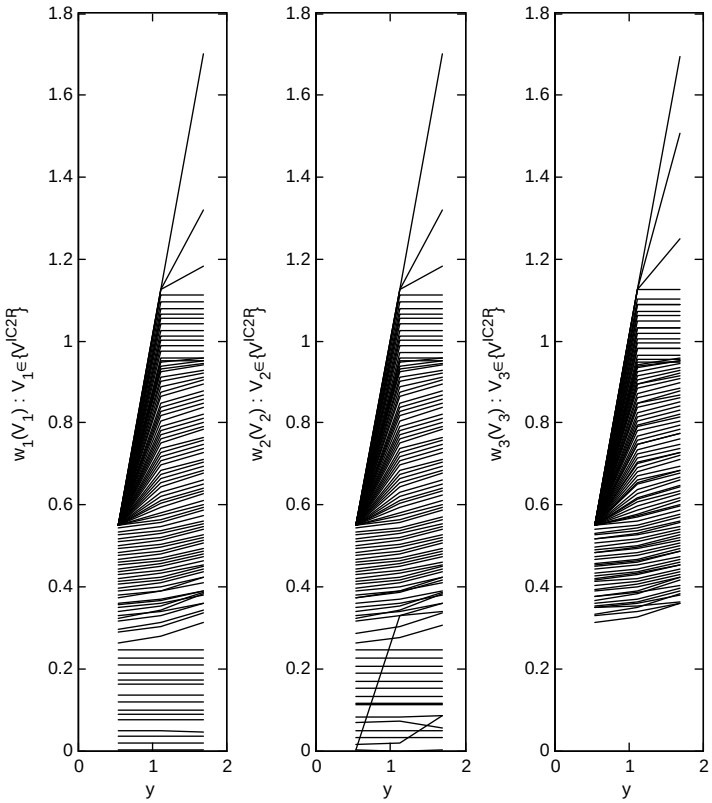

Figure 13: Optimal wage under the IC2P contract as a function of future stock price: $w^*_{y_i}\left(V_{y_i},.\right)$: $V_{y_i} \in \left\{V^{IC2P}\left(y_i\right)\right\}$, $i \in \{1,2,3\}$ (LMH, case 3)

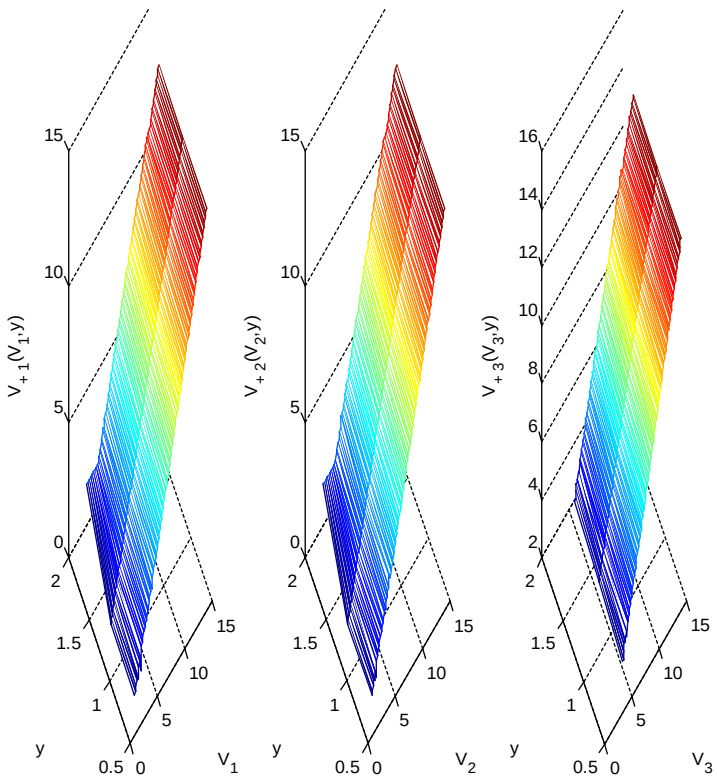

Figure 14: Optimal future utility promise under the IC2P contract as a function of initial utility promise and future stock price: $V^*_{+\ y_i}(V_{y_i}, y_k)$: $V_{y_i} \in \left\{V^{IC2P}(y_i)\right\}$, $i, k \in \{1, 2, 3\}$ (LMH, case 3)

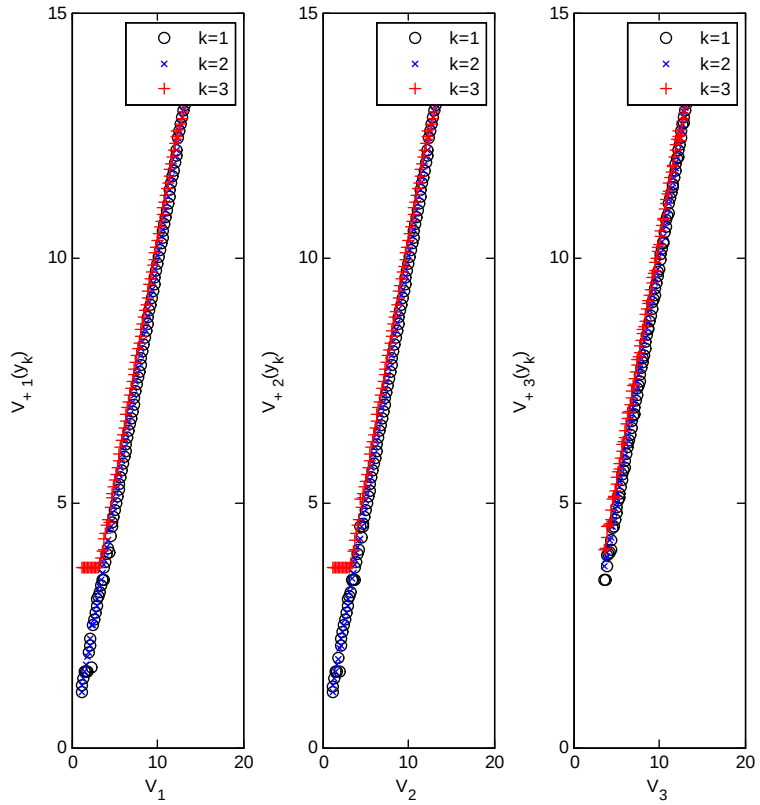

Figure 15: Optimal future utility promise under the IC2P contract as a function of initial utility promise: $V^*_{+\ y_i}(., y_k)$, $i, k \in \{1, 2, 3\}$ (LMH, case 3)

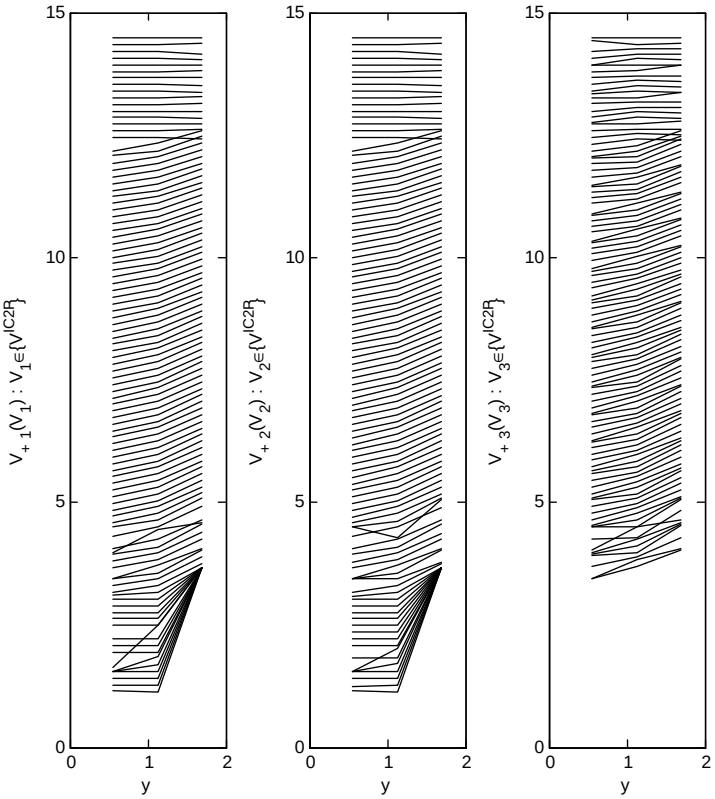

Figure 16: Optimal future utility promise under the IC2P contract as a function of future stock price: $V^*_{+\ y_i}(V_{y_i}, .)$: $V_{y_i} \in \left\{V^{IC2P}(y_i)\right\}$, $i \in \{1, 2, 3\}$ (LMH, case 3)