Is there a Golden Parachute in Sannikov's principal–agent problem?*

Dylan Possamaï[†] Nizar Touzi[‡]

July 14, 2020

Abstract

This paper provides a complete review of the continuous–time optimal contracting problem introduced by Sannikov [54], in the extended context allowing for possibly different discount rates of both parties. The agent's problem is to seek for optimal effort, given the compensation scheme proposed by the principal over a random horizon. Then, given the optimal agent's response, the principal determines the best compensation scheme in terms of running payment, retirement, and lump–sum payment at retirement.

A Golden Parachute is a situation where the agent ceases any effort at some positive stopping time, and receives a payment afterwards, possibly under the form of a lump sum payment, or of a continuous stream of payments. We show that a Golden Parachute only exists in certain specific circumstances. This is in contrast with the results claimed by Sannikov [54], where the only requirement is a positive agent's marginal cost of effort at zero. Namely, we show that there is no Golden Parachute if this parameter is too large. Similarly, in the context of a concave marginal utility, there is no Golden Parachute if the agent's utility function has a too negative curvature at zero.

In the general case, we provide a rigorous analysis of this problem, and we prove that an agent with positive reservation utility is either never retired by the principal, or retired above some given threshold (as in Sannikov's solution). In particular, different discount factors induce naturally a *face–lifted utility function*, which allows to reduce the whole analysis to a setting similar to the equal–discount rates one. Finally, we also confirm that an agent with small reservation utility does have an informational rent, meaning that the principal optimally offers him a contract with strictly higher utility value.

Key words: continuous–time principal–agent, optimal control and stopping, face–lifting.

1 Introduction

Principal–agent problems naturally stem from questions of optimal contracting between two parties – a principal ('she') and an agent ('he'), when the agent's effort cannot be observed or contracted upon. Mathematically, they are formulated as a Stackelberg non–zero sum game, and can also be identified to bi–level optimisation problems in the operations research literature. The number of articles related to this topic is staggering, mainly due to the wide spectrum of concrete problems where this theory is able to provide relevant results, for instance for moral hazard problem in microeconomics with applications to corporate governance, portfolio management, and many other areas of economics and finance.

The first, and seminal, paper on principal–agent problems in continuous–time is by Holmström and Milgrom [34], who show that the optimal contract is linear in the output process, in a finite horizon setting with CARA utility functions for both parties, and when the agent's effort impacts solely the drift of the output process. This paper is the first to highlight that optimal contracting problems tend to be easier to

*We are grateful to Yuliy Sannikov for his insightful comments on the first version of this paper. This work benefited from support of the ANR project PACMAN ANR–16–CE05–0027.

[†]Columbia University, IEOR department, USA, dp2917@columbia.edu

[‡]CMAP, École Polytechnique, 91128 Palaiseau Cedex, France, nizar.touzi@polytechnique.edu. This authors is also grateful for the financial support from the Chaires FiME-FDD and Financial Risks of the Louis Bachelier Institute.

address in continuous–time, an observation which has been confirmed by the large continuous–time literature in this area. Holmström and Milgrom's work was extended by Schättler and Sung [56], Sung [66; 67], Müller [43; 44], and Hellwig and Schmidt [32; 31]. While the aforementioned papers use continuous–time extensions of the celebrated first–order approach from the contract theory literature in static cases, see for instance Rogerson [52], the papers by Williams [70; 71; 72] and Cvitanić, Wan, and Zhang [15; 16; 17] use the stochastic maximum principle and forward–backward stochastic differential equations to characterise the optimal compensation for more general utility functions, see also the excellent monograph by Cvitanić and Zhang [14].[1]

The seminal work of Sannikov [54], see also Sannikov [55], represents a genuine breakthrough in this vast literature from various perspectives. First, from the methodological viewpoint, Sannikov introduced the idea to focus on the dynamic continuation value of the agent as a state variable for the principal's problem. Although this idea was already acknowledged throughout the discrete–time literature on this problem, an illuminating example being Spear and Srivastava [63], its systematic implementation in continuous–time offers an elegant solution approach by means of a representation result of the dynamic value function. Second, the infinite horizon setting considered by Sannikov revealed remarkable economic implications. Indeed, Sannikov's main conclusions are that the principal optimally retires the agent, offering him a Golden Parachute, that is to say a lifetime constant continuous stream of consumption, when his continuation utility reaches a sufficiently high level, and that an agent with small reservation utility possesses an informational rent, in the sense that he is offered a contract with strictly higher value.

The main objective of our paper is twofold. First, we revisit Sannikov's seminal work, but putting a stronger weight on technical rigour, which is unfortunately lacking in some key parts of [54]. We would like to emphasise that this should not be seen in any case as a reason to underestimate the importance of this paper, given the groundbreaking novelties recalled above. In contrast, our first aim is to try and contribute even more to the success of [54] by making it more accessible to a wider community of mathematicians and economists, whose overall understanding of the model may be hindered by the technical gaps in [54]. Notice that we are not the first to try and obtain rigorously the results claimed in [54]. For instance, Strulovici and Szydlowski [65, Section 4.3] offers a more rigorous take on the existence of optimal contracts in the model. However, the authors take for granted the fact that [54] proves that the HJB equation for the principal's problem has a smooth solution, while we will argue that the proof has important gaps. Similarly, the unpublished PhD thesis of Choi [12] aims at putting the problem on rigorous foundations. Nonetheless, existence of optimal contracts is not addressed there, and the results rely on the assumption that it is never optimal to retire the agent temporarily, while our approach actually proves that this is the case. We also would like to refer to the recent work of Décamps and Villeneuve [20], where the authors study a related, but different, contracting problem, and where again the heart of the analysis is technical clarity: this should be an additional illustration that actually proving rigorously results in this literature is a challenging task.

Our second goal is to prove that our analysis extends beyond the case where the principal and the agent have the same discount rates. It is an important feature, as most models[2] in the discrete– or continuous–time literature either allow for risk–averse agents who are as patient as the principal, as in Sannikov [54], Fong [26], Myerson [45], and Hajjej, Hillairet, Mnif, and Pontier [28], or for more impatient, but risk–neutral agents, as in DeMarzo and Sannikov [21], Biais, Mariotti, Plantin, and Rochet [4], Biais, Mariotti, Rochet, and Villeneuve [5], Biais, Mariotti, and Rochet [6], He [30], Piskorski and Tchistyi [50], Piskorski and Westerfield [51], DeMarzo, Fishman, He, and Wang [22], Pagès and Possamaï [48], or Williams [72]. Even more surprisingly, our analysis can also accommodate the case where the principal is actually strictly more impatient than the agent. More precisely, when the principal is more impatient, but not too much (the actual bound depends on the level of risk–aversion of the agent), the solution exhibits no fundamental

[1] Other early continuous–time contract theory models were proposed by Adrian and Westerfield [1], Biais, Mariotti, Plantin, and Rochet [4], Biais, Mariotti, Rochet, and Villeneuve [5], Biais, Mariotti, and Rochet [6], Capponi and Frei [9], DeMarzo and Sannikov [21], DeMarzo, Fishman, He, and Wang [22], Fong [26], He [30], Hoffmann and Pfeil [33], Ju and Wan [35], Keiber [37], Leung [39], Mirrlees and Raimondo [42], Myerson [45], Ou-Yang [46], Pagès [47], Pagès and Possamaï [48], Piskorski and Tchistyi [50], Piskorski and Westerfield [51], Sannikov [53], Schroder, Sinha, and Levental [58], Van Long and Sorger [68], Westerfield [69], Zhang [73], Zhou [74], or Zhu [75].

[2] Exceptions are the recent work by Hajjej, Hillairet, and Mnif [29], where the agent is risk–averse and more impatient than the principal. However, they do not obtain clear results saying that the hypotheses of their verification result [29, Theorem 4.3] can be verified in practice, as well as the work of Lin, Ren, Touzi, and Yang [41], but there the emphasis is more on obtaining general methods to attack infinite horizon moral hazard problems.

difference compared to the case where the principal is more patient. However, when the discount rate of the principal falls below a critical threshold, the problem degenerates, optimal contracts cease to exist, and the principal can achieve her first–best value with appropriately defined sequences of incentive–compatible contracts. As far as we know, our paper is the first offering such a comprehensive analysis.

Our main findings are the following. First, in contrast with the overall message from [54], we show that a Golden Parachute only exists in some specific situations. It never happens if the agent's marginal cost of effort at zero is zero, or is sufficiently large. And it never happens if the agent's marginal utility is also concave, and his utility function has sufficiently large negative curvature at zero, with a level depending on the marginal cost of effort at zero. We also confirm rigorously the existence of informational rent for an agent with small reservation utility. Under our set of assumptions, our Theorem 3.6 provides a necessary and sufficient condition for this important economic effect to occur. The condition combines the curvature of the agent's utility function at zero, and his marginal cost of effort at zero. We emphasise that our rigorous presentation involves advanced tools from stochastic control theory and partial differential equations. In particular, the justification of the solution claimed by Sannikov in [54] requires the use of Perron's existence approach, combined with the theory of viscosity solutions, and it is unclear to us how the proof could be significantly simplified.

Finally, from the methodological and theoretical point of view, we have highlighted a novel phenomenon in (properly renormalised) moral hazard problems with risk–aversion and different discount rates, where the principal's problem has an optimal stopping component, in the sense that she can terminate the contract. Indeed, we proved that the problem could actually be treated as in the case with similar discount rates, but provided that in the principal's optimisation, the certainty equivalent of the agent's continuation utility paid upon retirement is not computed using the inverse utility function of the agent, but using instead a 'face–lifted', or a 'shadow utility' function, obtained as the solution to a specific deterministic control problem. In words, this deterministic control problem assesses whether upon termination of the contract, it could actually be profitable for the principal to instead retire the agent by providing him, for a certain amount of time, a deterministic rent while he exerts no effort. Though we present this method in the specific context of the model in [54], it applies to generic moral hazard problems with early retirement possibilities. To the best of our knowledge, such a phenomenon has not been observed before, neither in the contract theory literature, nor in the optimal stopping literature.[3]

The paper is organised as follows. Section 2 provides a rigorous formulation of the continuous–time contracting problem, with a clear description of the set of contracts, and introduces the face–lifted utility $\overline{F}$. Our main results are given in Section 3. Thus, Section 3.1 provides some conditions under which no Golden Parachute can exist, which can all be recovered by the more abstract sufficient condition $\mathfrak{I}_0(\overline{F}', \overline{F}'') > 0$ at infinity, where δ is the ratio of the discount rate of the agent and the principal, and $\mathfrak{I}_0$ is defined in (3.2). Next, Section 3.2 identifies the value function of the principal and describes optimal contracts, while Section 3.3 presents our numerical illustrations, and Section 3.4 discusses the gaps in [54]. Subsequently, for completeness, we provide a review of the first–best version of [54]'s model in Section 4, thus highlighting the very different nature of the first–best optimal contract, for which a Golden Parachute never exists. Then, Section 5 uses the result of Lin, Ren, Touzi, and Yang [41], itself an extension of earlier results by Cvitanić, Possamaï, and Touzi [18; 19], which justify rigorously Sannikov's [54] remarkable reduction of the principal's Stackelberg game problem into a standard control–and–stopping problem. Such a reduction opens the door for the use of standard tools of stochastic control theory. In particular, we treat the case of a very impatient principal in Section 6, which can be addressed directly by exhibiting a sequence of contracts inducing a degenerate situation where both parties achieve as large a payment as possible. The alternative case of reasonably impatient principal is analysed by means of the corresponding dynamic programming equation introduced in Section 7, where we also provide a verification result following the standard theory. In Section 9, we provide a rigorous analysis of the dynamic programming equation, and we isolate a set of

[3]The 'face–lifting' phenomenon corresponds to the so–called boundary layer effect in singular optimal control problems, and appeared naturally in various pricing problems in finance, either with hedging constraints or market frictions, see for instance Broadie, Cvitanić, and Soner [8], Bouchard and Touzi [7], Chassagneux, Élie, and Kharroubi [10], Guasoni, Rásonyi, and Schachermayer [27], Soner and Touzi [59; 60; 61; 62], Cheridito, Soner, and Touzi [11], and Schmock, Shreve, and Wystup [57], or for utility maximisation problems, see Larsen, Soner, and Žitković [38]. However all these references consider either 'pure' optimal control or stochastic target problems, while in our context, the face–lifting phenomenon occurs because of an optimal stopping problem, and is therefore of a different nature.

conditions which guarantee that the solution is of the form claimed in [54]. Finally, Section 10 complements our results and examines the possibility of existence of a Golden Parachute in the context of the finite horizon Holmström and Milgrom [34], where both parties are now allowed to be risk–averse.

2 Sannikov's contracting problem

2.1 Output process, agent's effort, and contract

This section reports our understanding of the continuous time contracting model in Sannikov [54]. Let $(\Omega, \mathcal{F}, \mathbb{P}^0)$ be a probability space carrying a one–dimensional $\mathbb{P}^0$–Brownian motion W^0. For fixed parameters $\sigma > 0$ and $X_0 \in \mathbb{R}$, the **output process** is defined by

$$X_t := X_0 + \sigma W_t^0,\ t \geq 0.$$

We denote by $\mathbb{F}$ the $\mathbb{P}^0$–augmentation of the natural filtration of X (or equivalently, of W^0), which is well–known to satisfy the usual conditions. We next introduce distributions $\mathbb{P}^\alpha$ of the output process under effort α, so as to induce the dynamics $\mathrm{d}X_t = \alpha_t \mathrm{d}t + \sigma \mathrm{d}W_t^\alpha$, for some $\mathbb{P}^\alpha$–Brownian motion W^α. This is naturally accomplished by means of the following argument based on the Girsanov transformation.

Let $\mathcal{A}$ be the collection of all $\mathbb{F}$–predictable processes α with values in a compact subset A of $[0, \infty)$, containing 0. For all $\alpha \in \mathcal{A}$, we may introduce an equivalent probability measure $\mathbb{P}^\alpha$ so that the process $W^\alpha := W^0 - \int_0^\cdot \frac{\alpha_s}{\sigma} \mathrm{d}s$ is a $\mathbb{P}^\alpha$–Brownian motion, and the process X can be written in terms of W^α as

$$X_t = X_0 + \int_0^t \alpha_s \mathrm{d}s + \sigma W_t^\alpha,\ t \geq 0.$$

Any $\alpha \in \mathcal{A}$ is called an **effort** process, and is interpreted as an action exerted in order to affect the distribution of the output process from $\mathbb{P}^0$ to $\mathbb{P}^\alpha$.

A **contract** is a triple $\mathbf{C} := (\tau, \pi, \xi)$, where $\tau \in \mathcal{T}$, the set of all $\mathbb{F}$–stopping times, ξ is a non–negative $\mathcal{F}_\tau$–measurable random variable, and $\pi \in \Pi$, the set of $\mathbb{F}$–predictable non–negative processes. Here, τ represents a retirement time, π is a process of continuous payment rate until retirement, and ξ is a lump–sum payment at retirement, which may be interpreted as a Golden Parachute in the terminology of Sannikov [54], see Definition 2.4 below.

We shall introduce later in Section 2.5 the collection $\mathfrak{C}^0$ of admissible contracts, by imposing some integrability requirements. These contracts allow to formulate the contracting problem which sets the terms of the delegation by the principal (she) of the output process to the agent (he). Namely, the principal seeks to design the optimal contract so as to guarantee that the agent best serves her objectives, while optimising his own interest.

2.2 The agent's problem

The agent preferences are defined by

- a utility function $u : [0, \infty) \longrightarrow [0, \infty)$ which is increasing, strictly concave, twice continuously differentiable on $(0, \infty)$, satisfies $u(0) = 0$ together with the (one–sided) Inada condition $\lim_{x\to\infty} u'(x) = 0$ and the growth condition

$$c_0\big(-1 + \pi^{\frac{1}{\gamma}}\big) \leq u(\pi) \leq c_1\big(1 + \pi^{\frac{1}{\gamma}}\big),\ \pi \geq 0,\ \text{for some } (c_0, c_1) \in (0, \infty)^2,\ \text{and some } \gamma > 1, \tag{2.1}$$

  which implies that $u(\infty) = \infty$, and $u^{-1}(y) \leq C\big(1 + y^\gamma\big)$, for any $(y, \pi) \in [0, \infty)$, and for some $C > 0$;
- a cost function $h : [0, \infty) \longrightarrow [0, \infty)$ assumed to be increasing, strictly convex, continuously differentiable, with $h(0) = 0$;
- a fixed discount rate $r > 0$.

Given a contract $\mathbf{C} := (\tau, \pi, \xi) \in \mathfrak{C}^0$ and $\alpha \in \mathcal{A}$, the utility obtained by the agent is defined by the problem

$$V^{\mathrm{A}}(\mathbf{C}) := \sup_{\alpha\in\mathcal{A}} J^{\mathrm{A}}(\mathbf{C}, \alpha),\ \text{where } J^{\mathrm{A}}(\mathbf{C}, \alpha) := \mathbb{E}^{\mathbb{P}^\alpha}\bigg[\mathrm{e}^{-r\tau} u(\xi) + \int_0^\tau r\mathrm{e}^{-rs}\big(u(\pi_s) - h(\alpha_s)\big)\mathrm{d}s\bigg]. \tag{2.2}$$

As $u \geq 0$, and A is bounded, notice that $J^{\mathrm{A}}(\mathbf{C},\alpha) \in \mathbb{R} \cup \{\infty\}$ is well–defined. Moreover, as the agent is allowed to choose zero effort, inducing $J^{\mathrm{A}}(\mathbf{C},0) \geq 0$, it follows that $V^{\mathrm{A}}(\mathbf{C}) \geq 0$ for any proposed contract $\mathbf{C} \in \mathfrak{C}^0$. We denote by

$$\mathcal{A}^\star(\mathbf{C}) := \big\{\alpha \in \mathcal{A} : V^{\mathrm{A}}(\mathbf{C}) = J^{\mathrm{A}}(\mathbf{C},\alpha)\big\},$$

the (possibly empty) set of all optimal responses of the agent.

In addition, the agent only accepts contracts which provide him with a utility above some fixed threshold $u(R)$, where $R \geq 0$, called participation level. Thus he is only willing to consider contracts in the subset

$$\mathfrak{C}^0_R := \big\{\mathbf{C} \in \mathfrak{C}^0 : V^{\mathrm{A}}(\mathbf{C}) \geq u(R)\big\}.$$

Observe that the final lump–sum utility for the agent can be written as $u(\xi) = \int_\tau^\infty r\mathrm{e}^{-rt}u(\xi)\mathrm{d}t$, so that it can be equivalently implemented by the payment of the lifetime consumption ξ after retirement at time τ. We shall comment further on this normalisation in Remark 2.3 below.

2.3 The principal's problem

The principal is risk–neutral with the objective of maximising her overall revenue induced by the agent's effort and the promised compensation

$$J^{\mathrm{P}}(\mathbf{C},\alpha) := \mathbb{E}^{\mathbb{P}^\alpha}\bigg[-\mathrm{e}^{-\rho\tau}\xi + \int_0^\tau \rho\mathrm{e}^{-\rho s}(\mathrm{d}X_s - \pi_s\mathrm{d}s)\bigg],$$

where we consider here an extension of Sannikov [54], allowing the principal to have a possibly different discount rate $\rho > 0$ from that of the agent.

Observe that for any $\alpha \in \mathcal{A}$, $\mathbb{E}^{\mathbb{P}^\alpha}\big[\int_0^\tau \mathrm{e}^{-2\rho s}\mathrm{d}s\big] \leq \int_0^\infty \mathrm{e}^{-2\rho s}\mathrm{d}s = \frac{1}{2\rho} < \infty$. Then, by standard Itō integration theory, we have $\mathbb{E}^{\mathbb{P}^\alpha}\big[\int_0^\tau \mathrm{e}^{-\rho s}\sigma\mathrm{d}W^\alpha_s\big] = 0$ for all stopping time $\tau \in \mathcal{T}$, implying that

$$J^{\mathrm{P}}(\mathbf{C},\alpha) = \mathbb{E}^{\mathbb{P}^\alpha}\bigg[-\mathrm{e}^{-\rho\tau}\xi + \int_0^\tau \rho\mathrm{e}^{-\rho s}(\alpha_s - \pi_s)\mathrm{d}s\bigg],$$

which is well–defined in $\{-\infty\} \cup \mathbb{R}$, due to the boundedness of A and the non–negativity of ξ and π.

We also notice again that the lump–sum payment ξ at time τ can be rewritten as $\xi = \int_\tau^\infty \rho\mathrm{e}^{-\rho t}\xi\mathrm{d}t$, and so it can be implemented by the lifetime payment at rate ξ after τ, in agreement with the corresponding interpretation in the agent's problem.

The principal's problem is defined as follows: anticipating the agent's optimal response, she chooses the contract which best serves her objective under the participation constraint

$$V^{\mathrm{P}} := \sup_{\mathbf{C} \in \mathfrak{C}^0_R}\ \sup_{\alpha \in \mathcal{A}^\star(\mathbf{C})} J^{\mathrm{P}}(\mathbf{C},\alpha). \tag{2.3}$$

2.4 Reformulation and face–lifted utility

We next re–write our contracting problem equivalently by using the opposite of the inverse of the agent's utility

$$F := -u^{-1}\mathbf{1}_{[0,u(\infty))} - \infty\mathbf{1}_{\mathbb{R}_+\setminus[0,u(\infty))}.$$

Then, denoting $\zeta := u(\xi)$ and $\eta := u(\pi)$, the criterion of the agent becomes

$$J^{\mathrm{A}}(\mathbf{C},\alpha) = \mathbb{E}^{\mathbb{P}^\alpha}\bigg[\mathrm{e}^{-r\tau}\zeta + \int_0^\tau r\mathrm{e}^{-rt}\big(\eta_t - h(\alpha_t)\big)\mathrm{d}t\bigg],\ (\mathbf{C},\alpha) \in \mathfrak{C}^0 \times \mathcal{A}, \tag{2.4}$$

where we abuse notations and indifferently refer as contract the triplet (τ,π,ξ), or the triplet (τ,η,ζ). We will use this identification implicitly throughout the paper. As for the principal, we have

$$J^{\mathrm{P}}(\mathbf{C},\alpha) = \mathbb{E}^{\mathbb{P}^\alpha}\bigg[\mathrm{e}^{-\rho\tau}F(\zeta) + \int_0^\tau \rho\mathrm{e}^{-\rho t}\big(\alpha_t + F(\eta_t)\big)\mathrm{d}t\bigg],\ (\mathbf{C},\alpha) \in \mathfrak{C}^0_R \times \mathcal{A}.$$

Here, the (negative) reward of the principal by stopping at τ is $F(\zeta)$.

Our first result shows that in general, the principal may be able to improve her reward by not ending the contract at some time τ with a lump–sum payment to the agent, but by instead discouraging the agent from exerting any efforts (which can be understood as an alternative way of ending the contract), and offering him a continuous consumption. The improved (or face–lifted, hereafter) reward is naturally defined by means of the following deterministic control problem

$$\overline{F}(y_0) := \sup_{p\in\mathcal{B}_{\mathbb{R}_+}} \sup_{T\in[0,T_0^{y_0,p}]} \bigg\{ \mathrm{e}^{-\rho T} F\big(y^{y_0,p}(T)\big) + \int_0^T \rho \mathrm{e}^{-\rho t} F\big(p(t)\big)\mathrm{d}t \bigg\},\ y_0 \geq 0, \tag{2.5}$$

where $\mathcal{B}_{\mathbb{R}_+}$ is the set of Borel measurable maps from $\mathbb{R}_+$ to $\mathbb{R}_+$, and for all $(y_0, p) \in \mathbb{R}_+ \times \mathcal{B}_{\mathbb{R}_+}$,

$$T_0^{y_0,p} := \inf\big\{t \geq 0 : y^{y_0,p}(t) \leq 0\big\} \in [0,\infty],$$

and the state process $y^{y_0,p}$ is defined by the controlled first–order ODE

$$y^{y_0,p}(0) = y_0,\ \dot{y}^{y_0,p}(t) = r\big(y^{y_0,p}(t) - p(t)\big),\ t > 0. \tag{2.6}$$

To better understand the expression (2.5) for the improved payment, notice that for any $p \in \mathcal{B}_{\mathbb{R}_+}$, direct integration of this linear ODE leads to

$$y_0 = \mathrm{e}^{-rT} y^{y_0,p}(T) + \int_0^T \mathrm{e}^{-rt} p(t)\mathrm{d}t,\ \text{for all } y_0 \geq 0,\ \text{and } T \leq T_0^{y_0,p},$$

meaning that for a given state of the world ω, the agent is indifferent between a lump–sum payment $\xi(\omega)$ at some retirement time $\tau(\omega)$, and a continuous payment $p(t)$ on the time interval $[\tau(\omega), \tau(\omega) + T]$, with zero effort on this time interval, and a retirement deferred to $\tau(\omega) + T$, where the lump–sum payment is now $\xi'(\omega) := u^{-1}\big(y^{\zeta(\omega),p}(T)\big)$. The restriction $T \leq T_0^{y_0,p}$ on such deferral policies is induced by the fact that the agent is protected by limited liability, and therefore can only receive non–negative payments. The idea is that while the agent is indifferent between these two alternatives, the discrepancy between the discount rates may be such that the principal can actually benefit from postponing retirement.

An immediate consequence of this is that the value function of the principal can be expressed in its relaxed formulation as

$$\overline{V}^{\mathrm{P}} := \sup_{\mathbf{C}\in\mathfrak{C}_R} \sup_{\alpha\in\mathcal{A}^\star(\mathbf{C})} \bar{J}^{\mathrm{P}}(\mathbf{C},\alpha),\ \text{where } \bar{J}^{\mathrm{P}}(\mathbf{C},\alpha) := \mathbb{E}^{\mathbb{P}^\alpha}\bigg[\mathrm{e}^{-\rho\tau}\overline{F}(\zeta) + \int_0^\tau \rho\mathrm{e}^{-\rho t}\big(\alpha_t + F(\eta_t)\big)\mathrm{d}t\bigg], \tag{2.7}$$

where $\mathfrak{C}_R := \{\mathbf{C} \in \mathfrak{C} : V^{\mathrm{A}}(\mathbf{C}) \geq u(R)$, for some subset $\mathfrak{C} \subset \mathfrak{C}^0$ defined in Section 2.5 below.

The following result states the equivalence of our original contracting problem[4] with $\overline{V}^{\mathrm{P}}$, and characterises the face–lifted reward $\overline{F}$ in closed form in terms of the concave conjugate functions

$$F^\star(p) := \inf_{y\geq 0}\big\{yp - F(y)\big\},\ \text{and } \overline{F}^\star(p) := \inf_{y\geq 0}\big\{yp - \overline{F}(y)\big\},\ p \in \mathbb{R}.$$

Notice that $F^\star = 0$ on $\mathbb{R}_-$, and that our condition (2.1) on the agent's utility function is immediately converted for $F^\star$ into

$$-c_0^\star\big(1 + |p|^{\gamma^\star}\big) \leq F^\star(p) \leq c_1^\star\big(1 - |p|^{\gamma^\star}\big),\ p \leq 0,\ \text{with } \frac{1}{\gamma} - \frac{1}{\gamma^\star} = 1,\ \text{for some } (c_0^\star, c_1^\star) \in (0,\infty)^2. \tag{2.8}$$

Proposition 2.1. *We have $V^{\mathrm{P}} = \overline{V}^{\mathrm{P}}$, and the face–lifted reward function satisfies*

(i) $\overline{F} = 0$, *if* $\rho \geq \gamma r$;

(ii) $\overline{F} = F$, *if* $\rho = r$;

[4] As observed by Yuliy Sannikov in private communication with us, the principal problem may actually be directly defined by the relaxed formulation (2.7).

(iii) *if either $\rho \in (r, \gamma r)$, or $\rho \in (0, r)$ and $\lim_{y\to\infty} \frac{F'(y)}{yF''(y)}$ exists, we have $\overline{F} = (\overline{F}^\star)^\star$ where*

$$\overline{F}^\star(p) = \frac{1}{1-\delta}\bigg(\frac{|p|}{\delta}\bigg)^{\frac{1}{1-\delta}} \int_b^{\delta p} |x|^{-1-\frac{1}{1-\delta}} F^\star(x)\mathrm{d}x, \text{ with } \delta := \frac{\rho}{r},\ b := -\infty \mathbf{1}_{\{r<\rho\}} + F'(0)\mathbf{1}_{\{r>\rho\}}. \tag{2.9}$$

In particular $\overline{F}$ is decreasing, strictly concave, $r\overline{F}'(0) = \rho F'(0)\mathbf{1}_{\{r\geq\rho\}}$, and $\overline{F}^\star$ satisfies similar bounds to (2.8), with appropriate positive constants $\bar{c}_0^\star$, and $\bar{c}_1^\star$, which translate into bounds on $\overline{F}$ similar to those in (2.1), with appropriate positive constants $\bar{c}_0$, and $\bar{c}_1$. Moreover, the supremum over T in (2.5) is attained at $T_0^{y_0,p}$, meaning that

$$\overline{F}(y_0) = \sup_{p\in\mathcal{B}_{\mathbb{R}_+}} \bigg\{ \int_0^{T^{y_0,p}} \rho \mathrm{e}^{-\rho t} F\big(p(t)\big)\mathrm{d}t \bigg\},\ y_0 \geq 0. \tag{2.10}$$

The equality $V^{\mathrm{P}} = \overline{V}^{\mathrm{P}}$ in Proposition 2.1 is a direct consequence of our definition of admissible contracts in Section 2.5 below. The remaining claims are proved in Appendix A, and provide the following significant results. In the case $\rho = r$ considered by Sannikov [54], the principal never gains by postponing retirement and allowing the agent to produce zero effort for a while. On the other hand, when $\rho \neq r$, and ρ is not too large, it is always optimal to postpone retirement and $\overline{F}$ is a non–trivial majorant of F. Finally, when the principal becomes a lot more impatient than the agent, we actually have $\overline{F} = 0$, meaning that she can bring back the cost of permanently retiring the agent to 0.

Example 2.2. *Let $u(\pi) := \pi^{1/\gamma}$, and $\rho \neq r$ with $\rho < \gamma r$, then $F(y) = -y^\gamma$, and we compute directly*

$$F^\star(p) = -(\gamma-1)\bigg(\frac{|p|}{\gamma}\bigg)^{\gamma/(\gamma-1)},\ \overline{F}^\star(p) = -\frac{\rho(\gamma-1)^2}{r\gamma-\rho}\bigg(\frac{r|p|}{\rho\gamma}\bigg)^{\frac{\gamma}{\gamma-1}},\ p \leq 0,$$

and it follows from Proposition 2.1 *that*

$$\overline{F}(y) = -\bigg(\frac{r\gamma-\rho}{\rho(\gamma-1)}\bigg)^{\gamma-1}\bigg(\frac{ry}{\rho}\bigg)^{\gamma},\ y \geq 0.$$

Remark 2.3. *The normalisation of the running rewards of the principal and the agent by their corresponding discount rates in* Equation (2.2) *and* Equation (2.3), *is not fundamental,* per se. *However, the face–lifted principal's benefit function plays a crucial role to relate equivalent formulations of the problem. Consider for instance the agent's criterion*

$$\overline{J}_0^{\mathrm{A}}(\mathbf{C}, \alpha) := \mathbb{E}^{\mathbb{P}^\alpha}\bigg[\mathrm{e}^{-r\tau}u(\xi) + \int_0^\tau \mathrm{e}^{-rs}\big(u(\pi_s) - h(\alpha_s)\big)\mathrm{d}s\bigg],$$

which differs from J^{A} in (2.2) by the form of discount factor e^{-rt} instead of $r\mathrm{e}^{-rt}$. Similarly, change the principal's criterion to

$$\overline{J}_0^{\mathrm{P}}(\mathbf{C}, \alpha) := \mathbb{E}^{\mathbb{P}^\alpha}\bigg[-\mathrm{e}^{-\rho\tau}\xi + \int_0^\tau \mathrm{e}^{-\rho t}\big(\alpha_t - \pi_t\big)\mathrm{d}t\bigg].$$

Then, following the same argument, the corresponding face–lifted utility function is

$$\overline{F}_0(y_0) := \sup_{p\in\mathcal{B}_{\mathbb{R}_+}} \sup_{T\in[0,T_0^{y_0,p}]} \bigg\{\mathrm{e}^{-\rho T}F\big(y^{y_0,p}(T)\big) + \int_0^T \mathrm{e}^{-\rho t}F\big(p(t)\big)\mathrm{d}t\bigg\},\ y_0 \geq 0,$$

with controlled state satisfying for any $p \in \mathcal{B}_{\mathbb{R}_+}$, $y^{y_0,p}(0) = y_0$, and $\dot{y}^{y_0,p}(t) = \big(ry^{y_0,p}(t) - p(t)\big)$, $t > 0$. The corresponding Hamilton–Jacobi equation is

$$\min\Big\{\overline{F}_0 - F, \rho\overline{F}_0 - ry\overline{F}_0' + F^\star(\overline{F}_0')\Big\} = 0.$$

In particular, in the case $\rho = r$ of equal discount rates, we see immediately that $\overline{F}_0(y) := \frac{1}{r}F(ry)$, $y \geq 0$, is a solution of this equation. Consequently the decision of retiring the agent should be discussed by comparing the principal's value function to $\overline{F}_0$ instead of F in this case, see Definition 2.4 *below. In this sense, the setting of [54] is the only parametrisation of the problem with $\rho = r$ for which the face–lifted retirement reward function $\overline{F}$ coincides with F.*

2.5 Admissible contracts and Golden Parachute

For technical reasons, we introduce further integrability conditions which guarantee that both criteria of the agent and the principal are finite, and more importantly, allow to apply the reduction result of Lin, Ren, Touzi, and Yang [41]. We denote by $\mathfrak{C}$ the collection of all contracts $\mathbf{C} := (\tau, \pi, \xi)$, satisfying in addition the following integrability condition

$$\lim_{n\to\infty} \sup_{\alpha\in\mathcal{A}} \mathbb{P}^\alpha[\tau \geq n] = 0, \text{ and } \sup_{\alpha\in\mathcal{A}} \mathbb{E}^{\mathbb{P}^\alpha}\bigg[\big(\mathrm{e}^{-r'\tau}|\xi|\big)^\gamma + \int_0^\tau \big(\mathrm{e}^{-r's}|\pi_s|\big)^\gamma \mathrm{d}s\bigg] < \infty, \tag{2.11}$$

for some $r' \in (0, r \wedge \frac{\rho}{\gamma})$.

In order to guarantee that the equality $V^{\mathrm{P}} = \overline{V}^{\mathrm{P}}$ of Proposition 2.1 holds, we define the set $\mathfrak{C}^0$ as the collection of all triples (τ^0, π^0, ξ^0) such that

$$\tau^0 = \tau + T,\ \pi^0 = \pi \mathbf{1}_{[0,\tau)} + p\mathbf{1}_{[\tau,\tau^0)},\ \text{and } u(\xi^0) = y^{u(\xi),p}(T),$$

for some $(\tau, \pi, \xi) \in \mathfrak{C}$, and $\mathcal{F}_\tau$–measurable p with values in $\mathcal{B}_{\mathbb{R}_+}$, and T with values in $\big[0, T_0^{u(\xi),p}\big]$.

We can now introduce the notion of Golden Parachute which may have two different meanings in our relaxed formulation (2.7)

(i) in Sannikov's formulation, the retirement time τ is not explicitly involved in the model formulation. Instead, a Golden Parachute is defined as a stopping time τ such that the agent exerts no effort while receiving a constant consumption on $[\tau, \infty)$;

(ii) our definition of contracts includes a retirement time τ, and we may naturally define a situation of Golden Parachute by $\tau > 0$ and $\xi > 0$, $\mathbb{P}^0$–a.s.

Definition 2.4. *We say that the contracting model exhibits a* **Golden Parachute***, if there exists an optimal contract* $(\tau^\star, \pi^\star, \xi^\star) \in \mathfrak{C}_R$ *for the relaxed formulation of the principal problem* (2.7) *such that* $\tau^\star > 0$*, and* $\mathbb{P}^0[\xi^\star > 0] > 0$.

In other words, a Golden Parachute corresponds to a situation where there is a high–retirement point for the agent, with either lump–sum payment at retirement or continuous payment after retirement, where retirement means that the agent ceases to exert any effort forever.

Remark 2.5. *We shall provide in* Section 4 *a complete characterisation of the first–best version of our contracting problem*

$$V^{\mathrm{P,FB}} := \sup\Big\{J^{\mathrm{P}}(\mathbf{C}, \alpha) : \mathbf{C} \in \mathfrak{C}^{\mathrm{FB}},\ \alpha \in \mathcal{A},\ \text{and } J^{\mathrm{A}}(\mathbf{C}, \alpha) \geq u(R)\Big\},$$

where $\mathfrak{C}^{\mathrm{FB}}$ *is an appropriate extension of our* $\mathfrak{C}$*. In particular,* Theorem 4.1 *shows that the first–best optimal contract exhibits no Golden Parachute.*

3 Main results

3.1 Some cases of non–existence of a Golden Parachute

Our first main result provides a necessary condition for the potential optimality of a Golden Parachute, and then deduces some sufficient conditions which exclude the existence of a Golden Parachute, thus contrasting with the results claimed in Sannikov [54]. Our statement requires to introduce the convex conjugate of the cost of effort function

$$h^\star(z) := \sup_{a\in A} \{za - h(a)\},\ z \in \mathbb{R}. \tag{3.1}$$

We also introduce the corresponding subgradient $\partial h^\star(z) := \{a \in A : h^\star(z) = za - h(a)\}$, together with the second order differential operator

$$\mathfrak{I}_0(v', v'') := \sup_{z\in\mathbb{R},\ \hat a\in\partial h^\star(z)} \big\{\hat a + h(\hat a)\delta v' + \eta z^2 \delta v''\big\},\ \text{for all } C^2 \text{ function } v,\ \text{where } \delta := \frac{r}{\rho},\ \eta := \frac{1}{2} r\sigma^2. \tag{3.2}$$

Proposition 3.1. *If a* Golden Parachute *in the sense of* Definition 2.4 *exists, then*

$$\sup_{y\geq\bar y}\Big\{\mathfrak{I}_0(\overline{F}',\overline{F}'')(y)\Big\}\leq 0,\ \textit{for some } \bar y>0,\ \text{or equivalently}\ \sup_{p\leq\bar p}\bigg\{\mathfrak{I}_0\bigg(p,\frac{1}{(\overline{F}^\star)''(p)}\bigg)\bigg\}\leq 0,\ \textit{for some } \bar p<0.$$

In particular, there is no Golden Parachute *whenever either*

(NGP1) $h'(0)=0$;

(NGP2) *or* $h'(0)>0$, $\overline{F}''$ *is non–increasing, and* $\mathfrak{I}_0\big(\overline{F}'(0),\overline{F}''(0)\big)>0$;

(NGP3) *or* $h'(0)>0$, A *is an interval, and* $h\in C^3$ *with*

$$\inf_{a\in A}\bigg\{\frac{\big((h')^2\big)''(a)}{h''(a)}\bigg\}\geq\frac{1}{\eta}\sup_{y\geq 0}\bigg\{-\frac{\overline{F}'(y)}{\overline{F}''(y)}\bigg\},\ \text{and}\ \sup_{y\geq 0}\Big\{\overline{F}'(y)+2\eta\overline{F}''(y)h''(0)\Big\}\leq-\frac{1}{\delta h'(0)}.$$

Remark 3.2. *Assume for simplicity that* $F'(0)=0$, *then* $\overline{F}'(0)=0$ *by* Proposition 2.1. *Under our condition that* $\bar a:=\max A<\infty$, *we have*

$$\sup_{z\geq h'(0)}\big\{z^{-2}\max\hat A(z)\big\}\leq\frac{\bar a}{(h'(0))^2}.$$

Then, when $\overline{F}''$ *is non–increasing, the existence of a* Golden Parachute *implies that* $(h'(0))^2<\bar a/\big(-\eta\overline{F}''(0)\big)$. *In other words, the second alternative* (NGP2) *of* Proposition 3.1 *states that there is no* Golden Parachute *for sufficiently large* $h'(0)$. *As for the third alternative* (NGP3), *notice that the first condition is automatically satisfied whenever* $(h')^2$ *is convex, while the second one again requires* $h'(0)$ *to be large enough.*

Example 3.3. Sannikov [54, Figure 1] *considers the situation* $\delta=1$, *(so that* $\overline{F}=F$*), and*

$$F(y)=-y^2,\ y\geq 0,\ h(a):=\frac{1}{2}ha^2+\beta a,\ a\in A=\mathbb{R}_+,\ \text{for some positive constants } h \text{ and } \beta.$$

Notice that, given Sannikov*'s conclusion that a* Golden Parachute *exists, the unboundedness of* A *is not problematic, as the optimal effort remains bounded, so that the problem is unchanged by restricting to the corresponding compact subset of* A. *Under the present specification, we have*

$$F''(y)=-2,\ \text{and}\ \sup_{z\geq\beta}\big\{z^{-2}\max\hat A(z)\big\}=\sup_{a\geq 0}\bigg\{\frac{a}{h(a+\beta)^2}\bigg\}=\frac{1}{4h\beta}.$$

Then, since $F'(0)=0$ *in this case, the second alternative in* Proposition 3.1 *can be reformulated as*

$$\text{(NGP2) if and only if } 8\beta\eta h\geq 1.$$

3.2 Complete solution of the contracting problem

Recall from Proposition 2.1 that $\overline{F}=0$ when $\delta\gamma\leq 1$. Our first result shows that the solution of the contracting problem is degenerate in this case. Indeed, we shall exhibit a sequence of admissible contracts which induces a utility as large as we want for the agent, and reaches the highest possible level for the principal, namely $\bar a$. Roughly speaking, these contracts make small intermediate payments, enforce the highest possible effort for the agent at all times, and promise to pay him an extremely high value after an extremely long time. By exploiting the large discrepancy between the discount rates of the agent and the principal, we show that the continuation utilities of both parties reach their maximum.

We emphasise that this result is in line with the solution of the first–best contracting problem of Section 4 below, where we also exhibit a sequence of contracts which induce arbitrarily large level of utility for the agent, while providing the principal with a value as close as we want to her universal maximal utility of $\bar a$. We refer the reader to Section 4 and Section 6 for more intuitions on the contracts we construct.

Theorem 3.4. *Let* $\rho\geq\gamma r$. *Then* $V^{\mathrm{P}}=\bar a$, *there is no optimal contract achieving this value, and the second–best value of the principal coincides with her first–best value.*

The proof of this result is reported in Section 6. We next focus on the more interesting case $\rho < \gamma r$. Similar to Sannikov [54], the solution of the contracting problem is characterised by means of the second–order differential equation

$$v(0) = 0, \text{ and } v - \delta y v' + F^\star(\delta v') - \mathfrak{I}_0(v, v')^+ = 0, \text{ on } [0, \infty). \tag{3.3}$$

Our main results hold under the following assumption.

Assumption 3.5. *Either $\beta := h'(0) > 0$, or $A \supset [0, \bar{a}_0]$ for some $\bar{a}_0 > 0$. Moreover, if $\rho \in (0, r)$ then $\lim_{y\to\infty} F'(y)/\big(yF''(y)\big)$ exists.*

Theorem 3.6. *Let* Assumption 3.5 *hold true, and let $\mathcal{S} := \big\{v = \overline{F}\big\}$. Then*

(i) *there exists a unique solution $v \in C^2(\mathbb{R}_+)$ of (3.3), such that $0 \leq (v - \overline{F})(y) \leq C \log(1 + \log(1 + y))$, $y \geq 0$, for some $C > 0$;*

(ii) *v is strictly concave, ultimately decreasing, $v'(0) \geq 0$, and whenever $F'(0) = 0$, we have $v'(0) > 0$ if and only if $\mathfrak{I}_0\big(0, \overline{F}''(0)\big) > 0$;*

(iii) *if $\beta = 0$, then $\mathcal{S} = \{0\}$, and if $\beta > 0$, and in addition the maps F and $\mathfrak{I}_0$ of (7.2) are analytic, then $\mathcal{S} = \{0\} \cup [y_{\rm gp}, \infty)$ for some $y_{\rm gp} \in [0, \infty]$;*

(iv) *if $\mathcal{S} = \{0\} \cup [y_{\rm gp}, \infty)$ for some $y_{\rm gp} < \infty$, then*

$$\overline{V}^{\rm P} = \sup_{y \geq u(R)} v(y),$$

and the supremum is attained at some $\hat{y} \geq u(R)$. Defining $\hat{z} : [0, \infty) \longrightarrow \mathbb{R}$ to be a (measurable) maximiser of $\mathfrak{I}_0(v', v'')$, $\hat{\pi} : [0, \infty) \longrightarrow \mathbb{R}$ to be a (measurable) minimiser of $F^\star(\delta v')$, there exists a unique weak solution to the SDE *corresponding to $\widehat{Y} := Y^{\hat{y}, \hat{z}(\widehat{Y}), \hat{\pi}(\widehat{Y})}$. In particular, the contract $\big(\hat{\tau}, \hat{\pi}(\widehat{Y}), u^{-1}(\widehat{Y}_{\hat{\tau}})\big)$, where*

$$\hat{\tau} := \inf\Big\{t \geq 0 : \widehat{Y}_t \notin (0, y_{\rm gp})\Big\},$$

is an optimal contract for the relaxed principal problem (2.7).

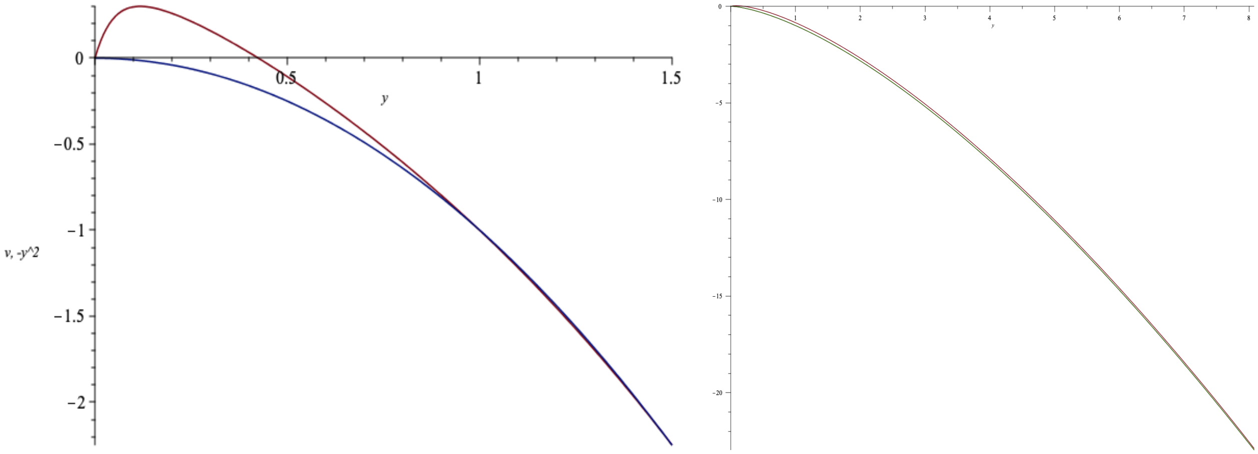
Remark 3.7. Sannikov *mentions that if* 'the agent had a higher discount rate than the principal, then with time the principal's benefit from output outweighs the cost of the agent's effort,' *and that* 'it is sensible to avoid permanent retirement by allowing the agent to suspend effort temporarily.' ([54, pp. 959])*. Our result shows that this statement is not correct: having $\delta > 1$ does not change the nature of the solution to the problem.*

Remark 3.8. *The case $\mathcal{S} = \{0\}$ is not covered by* Theorem 3.6.*(iv), due to the fact that in this case, the optimal retirement time $\tau^\star$ may be infinite with positive probability, and therefore cannot satisfy the integrability requirement in* Equation (2.11)*. This is however not a critical issue. Indeed, the integrability condition on admissible stopping times in* Equation (2.11) *is taken from the general result in* [41]*. But a detailed reading of their arguments shows that they only require it in order to be able to treat moral hazard problems where the agent is allowed to control the volatility of the output process, for which they need a theory for second–order* backward SDEs *with random horizon, which is obtained in* Lin, Ren, Touzi, and Yang [40], *but does not allow for infinite horizon. In our problem of interest, the agent only controls the drift of X, meaning that the classical theory of* backward SDEs *is sufficient, and these objects are known to be well–posed even with infinite horizon, see for instance* Papapantoleon, Possamaï, and Saplaouras [49]*. With these results in hand, we can straightforwardly extend the general reduction result of* Section 5 *to include possibly infinite retirement times, and then obtain a verification result general enough to cover these situations. As this is not central to our message, we refrain to go to this level of generality.*

3.3 Numerical illustration

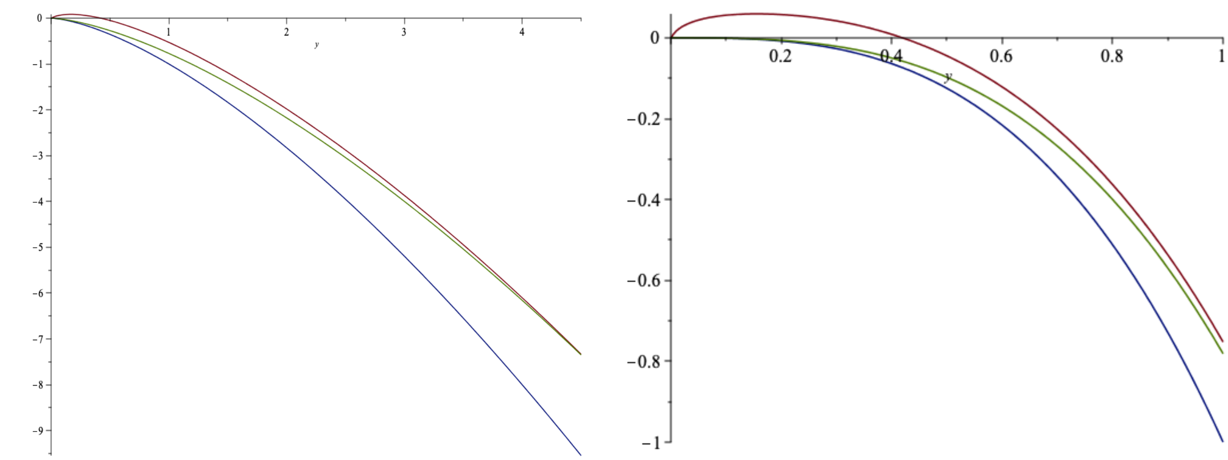
We next provide some numerical results with the cost of effort function from Example 3.3, and utility function $u(\pi) := \pi^\gamma$, $\gamma > 1$. We of course choose the model parameters so that neither (NGP2) nor (NGP3) are satisfied, since in those cases the solution is $\overline{F}$ everywhere.

Figure 1a takes the parameters in [54] (with $\gamma = 2$, $\eta = 0.05$, $h = 0.5$, $\beta = 0.4$, and $\delta = 1$), and shows the archetypical case where a Golden Parachute exists, as in [54]. Figure 1b however (with $\gamma = 3/2$, $\eta = h = 1$, $\beta = 0.01$, and $\delta = 1$) suggests strongly that v remains always strictly above F but becomes asymptotically close to it, a case for which a Golden Parachute would not exist.



(a) v (red), F (blue) (b) v (red), F (green)

The next two sets of figures show what happens when $\delta \neq 1$. More precisely, Figure 2a (with $\gamma = 3/2$, $\eta = h = 1$, $\beta = 0.01$, and $\delta = 3/4$) shows a case where v becomes equal to $\overline{F}$ after a while and a Golden parachute does exist, while, at least numerically, Figure 2b (with $\gamma = 3$, $\eta = h = 1$, $\beta = 0.01$, and $\delta = 2$), seems to show that v remains always above $\overline{F}$, and that no Golden Parachute exists.



(a) v (red), $\overline{F}$ (green), F (blue) (b) v (red), $\overline{F}$ (green), F (blue)

3.4 Sannikov's solution

In this subsection, we specialise the discussion to the case $\delta = 1$ to better compare with [54]. Notice that the HJB equation considered by Sannikov in [54, Equation (5)] is the same as our Equation (3.3) when restricted to the continuation region

$$v - yv' + F^\star(v') - \mathfrak{I}_0(v, v')^+ = 0,\ y \in [0, y_{\rm gp}],\ v(0) = F(0),\ v(y_{\rm gp}) = F(y_{\rm gp}) \text{ and } v'(y_{\rm gp}) = F'(y_{\rm gp}),$$

which corresponds to the natural guess that the stopping region $\mathcal{S} = \{v = F\}$ is of the form $\{0\} \cup [y_{\rm gp}, \infty)$, with some free boundary point $y_{\rm gp} < \infty$ to be determined so as to guarantee that the smooth–fit condition

$v'(y_{\rm gp}) = F'(y_{\rm gp})$ holds. Such a guess is more naturally justified by the optimal stopping component of the principal's problem in our formulation. We shall also see that it is necessary in order to apply the verification argument of Proposition 7.2 below (which in fact requires C^2 regularity).

A few pages later, namely in [54, Equation (6)], the author rewrites this ODE with $\mathfrak{I}_0$ instead of $\mathfrak{I}_0^+$

$$v - yv' + F^\star(v') - \mathfrak{I}_0(v, v') = 0,\ y \in [0, y_{\rm gp}],\ v(0) = F(0),\ v(y_{\rm gp}) = F(y_{\rm gp}) \text{ and } v'(y_{\rm gp}) = F'(y_{\rm gp}). \quad (3.4)$$

This is motivated by the natural guess that the principal is expected to induce a positive effort for the agent on the continuation region. More importantly, direct manipulations allow to reformulate the last equation equivalently as

$$v'' = \inf_{z \geq h'(0),\ \hat{a} \in \hat{A}(z)} \bigg\{ \frac{v - yv' + F^\star(v') - \hat{a} - h(\hat{a})v'}{\eta z^2} \bigg\}, \quad (3.5)$$

thus reducing the equation to an explicit non–linear second–order ODE under the additional restriction to a positive marginal cost of effort, that is to say when $h'(0) > 0$.

Next, assuming that $y_{\rm gp} < \infty$, the potential explosion of the solution due to the superlinear feature of $F^\star$ is bypassed, as the concavity of v implies that v' is bounded in $[v'(y_{\rm gp}), v'(0)]$. Although this assumption is not always true, see Proposition 3.1, we continue along the line of Sannikov. Then, it follows from the standard Cauchy–Lipschitz theorem that the last ODE, with initial data $v(0) = 0$ and $v'(0) = b$, has a unique classical solution for any choice of b, say v_b. Then, Sannikov argues that it is possible to choose b so that this solution v_b indeed solves Equation (3.4). Although, Sannikov's proof of this claim is not rigorous, we show in the subsequent analysis that this result may be correct for sufficiently small β. However, notice that our main results given in Section 3.2 and Section 3.1, show that

- for $\beta = 0$, there is no $y_{\rm gp} \geq 0$ such that the solution of the dynamic programming equation (3.3) agrees with F on $[y_{\rm gp}, \infty)$, see (NGP1) of Proposition 3.1;
- for $\beta > 0$ sufficiently small, we prove under additional conditions that the solution of (3.3) may exhibit the behaviour claimed by Sannikov. In fact, we shall prove that the stopping region $\mathcal{S}$ is either reduced to $\{0\}$, or is of the form $\{0\} \cup [y_{\rm gp}, \infty)$ for some $y_{\rm gp} \geq 0$. This requires some involved technical arguments which are displayed in Section 9 below;
- when the curvature at zero $u''(0)$ of the agent's utility is sufficiently large negative, the stopping region $\mathcal{S}$ is always reduced to $\{0\}$ for whatever value of $\beta > 0$. See (NGP2) of Proposition 3.1.

Finally, we observe that in [54, Figure 6], the value function v is tangent to F at the point $y_{\rm gp}$, but seems to be strictly above F on $(y_{\rm gp}, \infty)$! We believe that the function plotted in this figure is the solution of (3.4). Although this solution coincides with the solution of the dynamic programming equation (3.3) on the continuation region $[0, y_{\rm gp}]$, this figure shows that it lies strictly above it on the stopping region $(y_{\rm gp}, \infty)$ where the principal optimally retires the agent. Hence, this seems to be a concrete numerical evidence that the dynamic programming equation is not equivalent to (3.4).

4 The first–best contracting problem

This section reports for completeness the solution of the first–best version of the contracting problem

$$V^{\rm P,FB} := \sup \Big\{ J^{\rm P}(\mathbf{C}, \alpha) : \mathbf{C} \in \mathfrak{C}^{\rm FB},\ \alpha \in \mathcal{A},\ \text{and } J^{\rm A}(\mathbf{C}, \alpha) \geq u(R) \Big\},$$

where $\mathfrak{C}^{\rm FB}$ consists of all contracts (τ, π, ξ) where $\tau \in \mathcal{T}$ is a stopping time with values in $[0, \infty]$, and (π, ξ) satisfy the integrability condition of (2.11). In particular, we shall see that the first–best optimal contract exhibits no Golden Parachute.

We first consider the case where $\delta\gamma \leq 1$, which is somewhat degenerate. Indeed, as mentioned earlier, we can find a sequence of admissible contracts which ensure a utility as large as we want for the agent, and reaches the highest possible level for the principal, namely $\bar{a}$. The idea is to offer no intermediate payments, to ask the agent to exert maximal effort at all times, and to retire him after a very long time, at which we offer him a very large lump–sum payment. The difficulty is then in how to calibrate the speed at which the retirement time and the final payment explode, so as to ensure that the principal's utility still increases.

Theorem 4.1. *Assume that $\delta\gamma \leq 1$. Then, we have $V^{\mathrm{P,FB}} = \bar{a}$, and there does not exist an optimal contract.*

Proof. Notice first that the limited liability constraints on the payments made to the agent, and the fact that A is bounded by $\bar{a}$ imply immediately that for any $(\mathbf{C}, \alpha) \in \mathfrak{C}^{\mathrm{FB}} \times \mathcal{A}$, we have

$$J^{\mathrm{P}}(\mathbf{C}, \alpha) \leq \bar{a}.$$

Moreover, the only way this can be an equality is to choose $\alpha = \bar{a}$, and $\mathbf{C} := (\tau, \pi, \xi)$ such that $\pi = 0$ and $F(\xi)\mathrm{e}^{\rho\tau} = 0$, which means that either $\tau = \infty$, or $\xi = 0$. However, such contracts do not satisfy the participation constraint of the agent, and therefore there cannot exist an admissible contract attaining the upper bound $\bar{a}$ for the principal. We will however show that one can find a sequence of admissible contracts which allows to approach $\bar{a}$ as close as we want.

For any $\varepsilon > 0$, let us consider the following contract: $\tau^\varepsilon := -\log(\varepsilon)/\varepsilon$, $\pi^\varepsilon := 0$, $\xi^\varepsilon := \varepsilon^{-1}\mathrm{e}^{\gamma(r-\varepsilon)\tau^\varepsilon}$, with the level of effort $\alpha^\varepsilon := \bar{a}$. Since these contracts are defined by deterministic components, they automatically satisfy the integrability condition of (2.11). Notice also that when ε goes to 0, both τ^ε and ξ^ε converge to ∞. Therefore, we can choose ε small enough and find a constant $C > 0$, independent of ε, such that $u(\xi^\varepsilon) \geq C(\xi^\varepsilon)^{1/\gamma}$. The utility received by the agent is then

$$\mathrm{e}^{-r\tau^\varepsilon} u\big(\xi^\varepsilon\big) - h(\bar{a})\big(1 - \mathrm{e}^{-r\tau^\varepsilon}\big) \geq \frac{C}{\varepsilon^{\gamma-1}} - h(\bar{a}) \underset{\varepsilon \to 0}{\longrightarrow} \infty,$$

so that the agent's participation constraint is satisfied for ε small enough. The principal's utility is

$$-\mathrm{e}^{-\rho\tau^\varepsilon}\xi^\varepsilon + \bar{a}\big(1 - \mathrm{e}^{-\rho\tau^\varepsilon}\big) = \mathrm{e}^{\rho(1-\delta\gamma)\frac{\log(\varepsilon)}{\varepsilon}}\varepsilon^{\gamma-1} + \bar{a}\big(1 - \mathrm{e}^{-\rho\tau^\varepsilon}\big) \underset{\varepsilon \to 0}{\longrightarrow} \bar{a},$$

since $\delta\gamma \leq 1$, which ends the proof in this case. □

When $\delta\gamma > 1$, the problem does not degenerate any longer, unless the reservation utility of the agent is too low and either $\bar{a}$ is too small, or $(F^\star)'(0)$ is too large. The solution is expressed in terms of the function

$$G^\star(p) := \sup_{a \in A}\big\{a + ph(a)\big\},\ p \in \mathbb{R}.$$

Theorem 4.2. *Let $\delta\gamma > 1$. Then*

(i) *if $u(R) \leq -h(\bar{a}) + \big(F^\star\big)'(0)$, the value function of the first–best problem is $V^{\mathrm{P,FB}} = \bar{a}$, and there is no optimal contract which achieves this value*;

(ii) *otherwise, $V^{\mathrm{P,FB}} = -\lambda^\star\delta u(R) + \big(G^\star - F^\star\big)\big(-\delta\lambda^\star\big)$, where $\lambda^\star$ is the unique positive solution of*

$$-u(R) - \int_0^\infty r\mathrm{e}^{-rt}\big(G^\star - F^\star\big)'\big(-\delta\lambda^\star\mathrm{e}^{(\rho-r)t}\big)\mathrm{d}t = 0.$$

Moreover, the agent's participation constraint is saturated, with first–best optimal contract

$$\tau^\star = \infty,\ \text{and}\ \pi_t^\star \in \hat{U}\bigg(\frac{\mathrm{e}^{(r-\rho)t}}{\delta\lambda^\star}\bigg),\ a_t^\star \in \hat{A}\bigg(\frac{\mathrm{e}^{(r-\rho)t}}{\delta\lambda^\star}\bigg),\ t \geq 0,$$

where for any $z \in \mathbb{R}$, $\hat{U}(z) := \mathrm{argmin}_{p\geq 0}\big\{zp - u(p)\big\}$.

Proof. By the standard Karush–Kuhn–Tucker method, we rewrite the first best problem as

$$\begin{aligned}
&\inf_{\lambda\geq 0}\bigg\{-\lambda u(R) + \sup_{(\mathbf{C},\alpha)\in\mathfrak{C}^{\mathrm{FB}}\times\mathcal{A}} \mathbb{E}^{\mathbb{P}^\alpha}\bigg[-\big(\mathrm{e}^{-\rho\tau}\xi - \mathrm{e}^{-r\tau}\lambda u(\xi)\big) - \int_0^\tau \Big(\rho\mathrm{e}^{-\rho t}\pi_t - r\mathrm{e}^{-rt}\lambda u(\pi_t)\Big)\mathrm{d}t \\
&\qquad\qquad + \int_0^\tau \Big(\rho\mathrm{e}^{-\rho t}\alpha_t - r\mathrm{e}^{-rt}\lambda h(\alpha_t)\Big)\mathrm{d}t\bigg]\bigg\} \\
&= \inf_{\lambda\geq 0}\bigg\{-\lambda u(R) + \sup_{(\tau,\alpha)\in\mathcal{T}\times\mathcal{A}} \mathbb{E}^{\mathbb{P}^\alpha}\bigg[-\mathrm{e}^{-\rho\tau}F^\star\big(-\lambda\mathrm{e}^{(\rho-r)\tau}\big) + \int_0^\tau \rho\mathrm{e}^{-\rho t}\big(G^\star - F^\star\big)\big(-\delta\lambda\mathrm{e}^{(\rho-r)t}\big)\mathrm{d}t\bigg]\bigg\} \\
&= \inf_{\lambda\geq 0}\Big\{-\lambda u(R) + \sup_{T\geq 0} f(T)\Big\},\quad f(T) := -\mathrm{e}^{-\rho T}F^\star\big(-\lambda\mathrm{e}^{(\rho-r)T}\big) + \int_0^T \rho\mathrm{e}^{-\rho t}\big(G^\star - F^\star\big)\big(-\delta\lambda\mathrm{e}^{(\rho-r)t}\big)\mathrm{d}t.
\end{aligned}$$

As $G^\star \geq 0$, and $F^\star$ is concave, we have

$$f'(T) \geq \rho \mathrm{e}^{-\rho T}\Big(F^\star\big(-\lambda \mathrm{e}^{(\rho-r)T}\big) - F^\star\big(-\lambda\delta \mathrm{e}^{(\rho-r)T}\big) + \lambda(1-\delta)\mathrm{e}^{(\rho-r)T}(F^\star)'\big(-\lambda \mathrm{e}^{(\rho-r)T}\big)\Big) \geq 0,\ T \geq 0.$$

Then, the supremum over $T \geq 0$ is attained at infinity, with $\lim_{T\to\infty} f(T) = \phi(\lambda) < \infty$, as $\delta\gamma > 1$, where

$$\phi(\lambda) := \int_0^\infty \rho \mathrm{e}^{-\rho t}\big(G^\star - F^\star\big)\big(-\delta\lambda \mathrm{e}^{(\rho-r)t}\big)\mathrm{d}t < \infty,\ \text{and}\ V_0^{\mathrm{P,FB}} = \inf_{\lambda\geq 0}\big\{-\lambda u(R) + \phi(\lambda)\big\}.$$

Notice that ϕ is strictly convex, with $\phi(0) = (G^\star - F^\star)(0) = G^\star(0) = \bar{a}$, and $\lim_{\lambda\to\infty}\phi(\lambda) = \infty$, since $G^\star$ has linear growth (recall that A is compact), and $F^\star$ grows as $(-p)^{\gamma/(\gamma-1)}$ at $-\infty$. We also compute directly that

$$\phi'(0) = -\big(G^\star - F^\star\big)'(0) \leq -\big(G^\star - F^\star\big)'(0) = -h(\bar{a}) + \big(F^\star\big)'(0),$$

where the last equality follows from the differentiability of $G^\star$ and the observation that $G^\star(p) = \bar{a} + ph(\bar{a})$ for $p \geq 0$. Consequently, the minimum in the last expression of $V^{\mathrm{P,FB}}$ is attained

- either at $\lambda^\star = 0$, if $-u(R) - h(\bar{a}) + \big(F^\star\big)'(0) \geq 0$, inducing the value $V^{\mathrm{P,FB}} = (G^\star - F^\star)(0) = G^\star(0) = \bar{a}$,
- or at the unique solution $\lambda^\star$ of $-u(R) - \phi'(\lambda^\star) = 0$. Then it follows from a direct integration by parts that $-\lambda^\star(1-\delta)u(R) = \lambda^\star(1-\delta)\phi'(\lambda^\star) = (G^\star - F^\star)(-\delta\lambda^\star) - \phi(\lambda^\star)$, and therefore

$$V^{\mathrm{P,FB}} = -\lambda^\star u(R) + \phi(\lambda^\star) = -\lambda^\star\delta u(R) + \big(G^\star - F^\star\big)\big(-\delta\lambda^\star\big).$$

Notice finally that when $\lambda^\star = 0$, similar to the proof of Theorem 4.1, there is no optimal contract. □

5 Reduction to a mixed control–and–stopping problem

In order to prove our main results reported in Section xsect:mainresults, we use the general approach of Lin, Ren, Touzi, and Yang [41][5] which justifies the remarkable solution approach introduced by Sannikov [54], reducing the Stackelberg game problem of the principal (2.7) into a standard stochastic control one.

To do this, observe that the Hamiltonian of the agent's problem is given by convex conjugate function $h^\star$ introduced in (3.1), and that the corresponding sub–gradient contains all possible optimal agent responses

$$\hat{A}(z) := \partial h^\star(z) = \{a \in A : h^\star(z) = za - h(a)\}.$$

As A is closed and h is strictly convex[6], notice that

$$\hat{A}(z) \neq \emptyset,\ \text{whenever}\ h^\star(z) < \infty,\ \text{and}\ \hat{A}(z) = \{0\},\ \text{for}\ z \leq h'(0), \tag{5.1}$$

because $a \longmapsto za - h(a)$ is decreasing whenever $z \leq h'(0)$. We also abuse notations slightly, and for any $\mathbb{F}$–predictable, real–valued process Z and any $\alpha \in \mathcal{A}$, we write $\alpha \in \hat{A}(Z)$ whenever $\alpha_t \in \hat{A}(Z_t)$, $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}^0$–a.e.

Then, the lump–sum payment $\xi = u^{-1}(\zeta)$ promised by the principal at τ takes the form

$$\zeta = Y_\tau^{Y_0,Z,\pi} = Y_0 + r\int_0^\tau Z_t \mathrm{d}X_t + \big(Y_t^{Y_0,Z,\pi} - h^\star(Z_t) - \eta_t\big)\mathrm{d}t, \tag{5.2}$$

where $Y^{Y_0,Z,\pi}$ represents the continuation utility of the agent given a continuous consumption stream $\pi = u^{-1}(\eta)$ and Z satisfies the integrability condition

$$\sup_{\alpha\in\mathcal{A}}\mathbb{E}^{\mathbb{P}^\alpha}\bigg[\sup_{0\leq t\leq\tau}\big(\mathrm{e}^{-r't}|Y_t|\big)^p\bigg] < \infty,\ \text{and}\ \sup_{\alpha\in\mathcal{A}}\mathbb{E}^{\mathbb{P}^\alpha}\bigg[\bigg(\int_0^\tau(\mathrm{e}^{-r't}|Z_t|)^2\mathrm{d}t\bigg)^{\frac{p}{2}}\bigg] < \infty. \tag{5.3}$$

[5] See Footnote 7. The methodology developed in [41] extends the finite maturity setting of Cvitanić, Possamaï, and Touzi [19] and is largely inspired by the method developed in Sannikov [54].

[6] If A is an interval, then the strict convexity of h guarantees that $\hat{A}(z)$ is a singleton. However, for a general closed subset A, the maximiser may not be unique.

Remark 5.1. *As observed by Sannikov [54], notice that the non–negativity condition on u and h implies that the so–called limited liability condition $Y^{Y_0,Z,\pi} \geq 0$ is satisfied. Indeed, as the dynamics of the process $Y^{Y_0,Z,\pi}$ are given by $\mathrm{d}Y_t^{Y_0,Z,\pi} = r\big(Y_t^{Y_0,Z,\pi} + h^\star(Z_t) - \eta_t\big)\mathrm{d}t + \sigma r Z_t \mathrm{d}W_t^0$, under the agent's optimal response, we see that 0 is an absorption point for the continuation utility with optimal effort 0.*

By the main reduction result of [41],[7] we may rewrite the principal's problem (2.7) as

$$V^{\mathrm{P}} = \sup_{Y_0 \geq u(R)} V(Y_0), \text{ where } V(Y_0) := \sup_{\substack{(\tau,Z,\pi)\in\mathcal{Z}(Y_0)\\ \hat{a}\in\hat{A}(Z)}} J(\tau,\pi,Z,\hat{a}), \tag{5.4}$$

and

$$J(\tau,\pi,Z,\hat{a}) := \mathbb{E}^{\mathbb{P}^{\hat{a}}}\bigg[\mathrm{e}^{-\rho\tau}\overline{F}\big(Y_\tau^{Y_0,Z,\pi}\big) + \int_0^\tau \rho \mathrm{e}^{-\rho t}\big(\hat{a}_t + F(\eta_t)\big)\mathrm{d}t\bigg]. \tag{5.5}$$

Here $\mathcal{Z}(Y_0)$ is the collection of all triples (τ, Z, π) such that τ, $\hat{a}$ and $\xi = -\overline{F}(Y_\tau^{Y_0,Z,\pi})$ satisfy the integrability conditions (2.11), for some $\hat{a} \in \hat{A}(Z)$, and therefore also (5.3), together with the limited liability condition $Y^{Y_0,Z,\pi} \geq 0$ of Remark 5.1.

The last control problem only involves the dynamics of $Y^{Y_0,Z,\pi}$ under the optimal response of the agent (due to the principal's criterion which does not involve anymore the state variable X)

$$\mathrm{d}Y_t^{Y_0,Z,\pi} = r\big(Y_t^{Y_0,Z,\pi} + h\big(\hat{a}_t\big) - \eta_t\big)\big)\mathrm{d}t + rZ_t\sigma \mathrm{d}W_t^{\hat{a}},\ \mathbb{P}^{\hat{a}}\text{–a.s., for all } \hat{a} \in \hat{A}(Z). \tag{5.6}$$

6 Second–best value for a (very) impatient principal

We now provide the proof of Theorem 3.4 by using the problem reduction from the previous section. Notice first that whenever $V^{\mathrm{P}} = \bar{a}$, then the result of Theorem 4.1 shows that there cannot exist an optimal contract, and that the first–best and second–best value coincide. Our proof is based on an explicit construction of a sequence of contracts following the idea used in the proof of Theorem 4.1: we want to have a retirement time going to ∞, associated with a large lump–sum payment. However, because we are now in the second–best case, we need to offer the agent contracts which are incentive–compatible with the level of effort $\bar{a}$, meaning that these contracts cannot be deterministic. This can however be achieved by choosing a large enough and constant control process Z in Equation (5.6). The price to pay now with such contracts is that the continuation utility of the agent may reach 0 in finite time with positive probability, thus preventing the principal from offering a large lump–sum payment. This thus requires to carefully control the probability of early termination of the contract, and we show that by offering the agent a sufficiently large utility, this probability can be made arbitrarily small.

Proof of Theorem 3.4. Let us fix some $y_0 > 0$, $z > h'(\bar{a})$. It is immediate that in this case $\hat{A}(z) = \{\bar{a}\}$. For arbitrary $\varepsilon \in (0, r\wedge 1)$, consider the continuous payment $\pi_t^\varepsilon := u^{-1}\big(\varepsilon Y_t^\varepsilon\big)$, $t \geq 0$, where $Y^\varepsilon := Y^{y_0/\sqrt{\varepsilon},z,\pi^\varepsilon}$ is the corresponding continuation utility of the agent, which is given by

$$Y_t^\varepsilon = \frac{y_0}{\sqrt{\varepsilon}} + \int_0^t \big((r-\varepsilon)Y_s^\varepsilon + rh(\bar{a})\big)\mathrm{d}s + rz\sigma W_t^{\bar{a}},\ t \geq 0.$$

Notice that Y^ε an Ornstein–Uhlenbeck process under $\mathbb{P}^{\bar{a}}$, whose SDE can be solved explicitly:

$$Y_t^\varepsilon = \mathrm{e}^{(r-\varepsilon)t}\frac{y_0}{\sqrt{\varepsilon}} + \frac{r}{r-\varepsilon}h(\bar{a})\big(\mathrm{e}^{r(t-\varepsilon)} - 1\big) + rz\sigma\int_0^t \mathrm{e}^{(r-\varepsilon)(t-s)}\mathrm{d}W_s^{\bar{a}},\ t \geq 0.$$

Let now $T_0^\varepsilon := \inf\big\{t > 0 : Y_t^\varepsilon = 0\big\}$, and consider the contract $\mathbf{C}_\varepsilon$ with retirement time $\tau^\varepsilon := \big(-\frac{\log(\varepsilon)}{\varepsilon}\big)\wedge T_0^\varepsilon$, continuous payments π^ε, and terminal payment $\xi^\varepsilon := u^{-1}\big(Y_{\tau^\varepsilon}^\varepsilon\big)$. We know from the general results in Section 5 that such a contract provides the agent with utility $-\log(\varepsilon)y_0$, which he will accept for ε small

[7] By the growth condition (2.1) on u, the integrability condition (2.11) implies that

$$\sup_{\alpha\in\mathcal{A}} \mathbb{E}^{\mathbb{P}^\alpha}\bigg[\big(\mathrm{e}^{-r'\tau}u(\xi)\big)^\gamma + \int_0^\tau \big(\mathrm{e}^{-r's}u(\pi_s)\big)^\gamma \mathrm{d}s\bigg] < \infty,$$

which is precisely the integrability condition required by Lin, Ren, Touzi, and Yang [41].

enough, regardless of the level of his participation constraint. Indeed, all the integrability requirements are obviously satisfied here, since z is deterministic, τ^ε is bounded, and from the explicit formula for Y^ε.

We now compute the principal's utility induced by this contract

$$J^{\mathrm{P}}(\mathbf{C}_\varepsilon,\bar a)=\mathbb{E}^{\mathbb{P}^{\bar a}}\big[\mathrm{e}^{-\rho\tau^\varepsilon}F\big(Y^\varepsilon_{\tau^\varepsilon}\big)\big]+\mathbb{E}^{\mathbb{P}^{\bar a}}\bigg[\int_0^{\tau^\varepsilon}\rho\mathrm{e}^{-\rho t}F\big(\varepsilon Y^\varepsilon_t\big)\mathrm{d}t\bigg]+\bar a\Big(1-\mathbb{E}^{\mathbb{P}^{\bar a}}\big[\mathrm{e}^{-\rho\tau^\varepsilon}\big]\Big).$$

Step 1. For $\varepsilon<r$, we have $T^\varepsilon_0>\bar T^\varepsilon_0:=\inf\big\{t>0:\bar Y^\varepsilon_t=0\big\}$, where $\bar Y^\varepsilon_t:=y_0/\sqrt{\varepsilon}+rh(\bar a)t+rz\sigma W^{\bar a}_t$, $t\geq 0$. The law of $\bar T^\varepsilon_0$ is well–known (see for instance Karatzas and Shreve [36, Equation (5.13)]), and we have

$$\mathbb{P}^{\bar a}\big[T^\varepsilon_0<\infty\big]\leq\mathbb{P}^{\bar a}\big[\bar T^\varepsilon_0<\infty\big]=\exp\bigg(-\frac{2h(\bar a)y_0}{r\sigma^2z^2\sqrt{\varepsilon}}\bigg)\underset{\varepsilon\to 0}{\longrightarrow}0. \tag{6.1}$$

This implies that

$$\mathbb{E}^{\mathbb{P}^{\bar a}}\big[\mathrm{e}^{-\rho\tau^\varepsilon}\big]=\mathrm{e}^{\rho\frac{\log(\varepsilon)}{\varepsilon}}\mathbb{P}^{\bar a}\big[T^\varepsilon_0=\infty\big]+\mathbb{E}^{\mathbb{P}^{\bar a}}\Big[\mathrm{e}^{-\rho\tau^\varepsilon}\mathbf{1}_{\{T^\varepsilon_0<\infty\}}\Big]\leq\mathrm{e}^{\rho\frac{\log(\varepsilon)}{\varepsilon}}\mathbb{P}^{\bar a}\big[T^\varepsilon_0=\infty\big]+\mathbb{P}^{\bar a}\big[T^\varepsilon_0<\infty\big]\underset{\varepsilon\to 0}{\longrightarrow}0.$$

Step 2. Next, we have that there exists some $C>0$, which may change value from line to line, but is independent of ε, such that for any $t\in\big[0,T^\varepsilon_0\big]$

$$\begin{aligned}0\leq-\mathrm{e}^{-\rho t}F\big(Y^\varepsilon_t\big)&\leq C\mathrm{e}^{-\rho t}\big(1+\big|Y^\varepsilon_t\big|^\gamma\big)\\&\leq C\mathrm{e}^{-\rho t}\bigg(1+\big(1+\varepsilon^{-\gamma/2}\big)\mathrm{e}^{(r-\varepsilon)\gamma t}+\mathrm{e}^{\gamma(r-\varepsilon)t}\bigg|\int_0^t\mathrm{e}^{-(r-\varepsilon)s}\mathrm{d}W^{\bar a}_s\bigg|^\gamma\bigg).\end{aligned}$$

Then, as the last stochastic integral is a Gaussian random variable, and $\delta\gamma\leq 1$, we see that

$$\begin{aligned}0&\leq-\mathbb{E}^{\mathbb{P}^{\bar a}}\big[\mathrm{e}^{-\rho\tau^\varepsilon}F\big(Y^\varepsilon_{\tau^\varepsilon}\big)\big]\\&\leq C\mathbb{E}^{\mathbb{P}^{\bar a}}\bigg[\mathrm{e}^{-\rho\tau^\varepsilon}\bigg(1+\big(1+|\log(\varepsilon)|^\gamma\big)\mathrm{e}^{(r-\varepsilon)\gamma\tau^\varepsilon}+\mathrm{e}^{\gamma(r-\varepsilon)\tau^\varepsilon}\bigg|\int_0^{\tau^\varepsilon}\mathrm{e}^{-(r-\varepsilon)s}\mathrm{d}W^{\bar a}_s\bigg|^\gamma\bigg)\bigg]\\&\leq C\mathrm{e}^{\rho\frac{\log(\varepsilon)}{\varepsilon}}\bigg(1+\big(1+\varepsilon^{-\gamma/2}\big)\mathrm{e}^{(r-\varepsilon)\gamma\frac{-\log(\varepsilon)}{\varepsilon}}+\mathrm{e}^{-\gamma(r-\varepsilon)\frac{\log(\varepsilon)}{\varepsilon}}\mathbb{E}^{\mathbb{P}^{\bar a}}\bigg[\bigg|\int_0^{-\frac{\log(\varepsilon)}{\varepsilon}}\mathrm{e}^{-(r-\varepsilon)s}\mathrm{d}W^{\bar a}_s\bigg|^\gamma\bigg]\bigg)\\&\quad+C\mathbb{E}^{\mathbb{P}^{\bar a}}\bigg[\mathbf{1}_{\{T^\varepsilon_0<\infty\}}\mathrm{e}^{-\rho\tau^\varepsilon}\bigg(1+\big(1+\varepsilon^{-\gamma/2}\big)\mathrm{e}^{(r-\varepsilon)\gamma\tau^\varepsilon}+\mathrm{e}^{\gamma(r-\varepsilon)\tau^\varepsilon}\bigg|\int_0^{\tau^\varepsilon}\mathrm{e}^{-(r-\varepsilon)s}\mathrm{d}W^{\bar a}_s\bigg|^\gamma\bigg)\bigg]\\&\leq C\mathrm{e}^{\rho\frac{\log(\varepsilon)}{\varepsilon}}\bigg(1+\big(1+\varepsilon^{-\gamma/2}\big)\mathrm{e}^{-(r-\varepsilon)\gamma\frac{\log(\varepsilon)}{\varepsilon}}+\mathrm{e}^{-\gamma(r-\varepsilon)\frac{\log(\varepsilon)}{\varepsilon}}\Big(1-\mathrm{e}^{2(r-\varepsilon)\frac{\log(\varepsilon)}{\varepsilon}}\Big)^{\frac{\gamma}{2}}\bigg]\bigg)\\&\quad+C\big(1+\varepsilon^{-\gamma/2}\big)\mathbb{P}^{\bar a}\big[T^\varepsilon_0<\infty\big]+C\mathbb{E}^{\mathbb{P}^{\bar a}}\bigg[\mathbf{1}_{\{T^\varepsilon_0<\infty\}}\bigg|\int_0^{\tau^\varepsilon}\mathrm{e}^{-(r-\varepsilon)s}\mathrm{d}W^{\bar a}_s\bigg|^\gamma\bigg].\end{aligned} \tag{6.2}$$

It can be checked directly that since $\delta\gamma\leq 1$, the first term on the right–hand side of Equation (6.2) goes to 0 as ε go to 0. By (6.1), the second term also goes to 0 as ε goes to 0. Finally, for the third one, we have using Cauchy–Schwarz inequality and Burkholder–Davis–Gundy's inequality

$$\begin{aligned}\mathbb{E}^{\mathbb{P}^{\bar a}}\bigg[\mathbf{1}_{\{T^\varepsilon_0<\infty\}}\bigg|\int_0^{\tau^\varepsilon}\mathrm{e}^{-(r-\varepsilon)s}\mathrm{d}W^{\bar a}_s\bigg|^\gamma\bigg]&\leq\Big(\mathbb{P}^{\bar a}\big[T^\varepsilon_0<\infty\big]\Big)^{\frac12}\mathbb{E}^{\mathbb{P}^{\bar a}}\bigg[\bigg|\int_0^{\tau^\varepsilon}\mathrm{e}^{-(r-\varepsilon)s}\mathrm{d}W^{\bar a}_s\bigg|^{2\gamma}\bigg]^{\frac12}\\&\leq C\Big(\mathbb{P}^{\bar a}\big[T^\varepsilon_0<\infty\big]\Big)^{\frac12}\bigg(\int_0^\infty\mathrm{e}^{-2(r-\varepsilon)s}\mathrm{d}s\bigg)^{\frac{\gamma}{2}}\\&\leq C\Big(\mathbb{P}^{\bar a}\big[T^\varepsilon_0<\infty\big]\Big)^{\frac12}\underset{\varepsilon\to 0}{\longrightarrow}0,\end{aligned}$$

Step 3. Notice also at this point that when $\delta\gamma<1$, we can follow all the steps above but take instead $\pi^\varepsilon=0$ in the contract. Then, all the terms appearing still converge to 0 when ε goes to 0, and it is enough in this case to conclude that $\lim_{\varepsilon\to 0}J^{\mathrm{P}}(\mathbf{C}_\varepsilon,\bar a)=\bar a$.

In the case $\delta\gamma = 1$, it remains to control the continuous payment term

$$\begin{aligned}
0 &\leq -\mathbb{E}^{\mathbb{P}^{\bar{a}}}\bigg[\int_0^{\tau^\varepsilon}\rho\mathrm{e}^{-\rho t}F\big(\varepsilon Y_t^\varepsilon\big)\mathrm{d}t\bigg]\\
&\leq -\int_0^\infty \rho\mathrm{e}^{-\rho t}\mathbb{E}^{\mathbb{P}^{\bar{a}}}\Big[\mathbf{1}_{\{\varepsilon|Y_t^\varepsilon|\leq 1\}}F\big(\varepsilon|Y_t^\varepsilon|\big)\Big]\mathrm{d}t + C\int_0^\infty \rho\mathrm{e}^{-\rho t}\mathbb{E}^{\mathbb{P}^{\bar{a}}}\Big[\mathbf{1}_{\{\varepsilon|Y_t^\varepsilon|>1\}}\big(1+\varepsilon^\gamma|Y_t^\varepsilon|^\gamma\big)\Big]\mathrm{d}t\\
&\leq -\int_0^\infty \rho\mathrm{e}^{-\rho t}\mathbb{E}^{\mathbb{P}^{\bar{a}}}\Big[F\big(1\wedge|\varepsilon Y_t^\varepsilon|\big)\Big]\mathrm{d}t + C\varepsilon^\gamma\int_0^\infty \rho\mathrm{e}^{-\rho t}\mathbb{E}^{\mathbb{P}^{\bar{a}}}\big[|Y_t^\varepsilon|^\gamma\big]\mathrm{d}t + C\int_0^\infty \rho\mathrm{e}^{-\rho t}\mathbb{P}^{\bar{a}}\big[\varepsilon|Y_t^\varepsilon|\geq 1\big]\mathrm{d}t.
\end{aligned}$$

Notice next that we have that for any $t \geq 0$

$$\varepsilon^\gamma|Y_t^\varepsilon|^\gamma \leq C\bigg(\varepsilon^{\gamma/2} + \varepsilon^\gamma\mathrm{e}^{\gamma(r-\varepsilon)t} + \varepsilon^\gamma\bigg|\int_0^t \mathrm{e}^{(r-\varepsilon)(t-s)}\mathrm{d}W_s^{\bar{a}}\bigg|^\gamma\bigg).$$

Therefore, we have since $\gamma > 1$

$$\begin{aligned}
0 \leq \varepsilon^\gamma\int_0^\infty \rho\mathrm{e}^{-\rho t}\mathbb{E}^{\mathbb{P}^{\bar{a}}}\big[|Y_t^\varepsilon|^\gamma\big]\mathrm{d}t &\leq \varepsilon^{\gamma/2} + C\rho\frac{\varepsilon^{\gamma-1}}{\gamma} + \rho\varepsilon^\gamma\int_0^\infty \mathrm{e}^{-\gamma\varepsilon t}\Big(1-\mathrm{e}^{-2(r-\varepsilon)t}\Big)^{\frac{\gamma}{2}}\mathrm{d}t\\
&\leq \varepsilon^\gamma(-\log(\varepsilon))^\gamma + 2\rho\frac{\varepsilon^{\gamma-1}}{\gamma} \underset{\varepsilon\to 0}{\longrightarrow} 0.
\end{aligned}$$

Finally, since for any $t \geq 0$, $\varepsilon|Y^\varepsilon|$ converges $\mathbb{P}^{\bar{a}}$–a.s. to 0, it is immediate by dominated convergence that

$$\int_0^\infty \rho\mathrm{e}^{-\rho t}\mathbb{P}^{\bar{a}}\big[\varepsilon|Y_t^\varepsilon|\geq 1\big]\mathrm{d}t \underset{\varepsilon\to 0}{\longrightarrow} 0, \text{ and } \int_0^\infty \rho\mathrm{e}^{-\rho t}\mathbb{E}^{\mathbb{P}^{\bar{a}}}\Big[F\big(1\wedge|\varepsilon Y_t^\varepsilon|\big)\Big]\mathrm{d}t \underset{\varepsilon\to 0}{\longrightarrow} 0,$$

which concludes the proof.

7 Dynamic programming equation

This section prepares for the proof of the remaining main results of Section 3 by applying the dynamic programming approach to solve the mixed control–and–stopping problem (5.4)–(5.5).

Notice that this problem is stationary in time due to the infinite horizon feature, and the time homogeneity of the dynamics of Y. By standard stochastic control theory, together with Remark 5.1, the corresponding HJB equation is

$$v(0) = 0, \text{ and } \min\big\{v - \overline{F}, \mathbf{L}v\big\} = 0, \text{ on } (0,\infty), \tag{7.1}$$

where for any $y > 0$

$$\mathbf{L}v(y) := v - \delta yv'(y) + F^\star\big(\delta v'(y)\big) - \mathfrak{I}_0\big(v(y), v'(y)\big)^+ = v(y) - F(y) - \mathbf{T}F\big(y, \delta v'(y)\big) - \mathfrak{I}_0\big(v'(y), v''(y)\big)^+,$$

and the second order differential operator $\mathfrak{I}_0$ is as introduced in Equation (3.2), and can be rewritten thanks to (5.1) as

$$\mathfrak{I}_0(p,q) = \infty\mathbf{1}_{\{q>0\}} + \mathbf{1}_{\{q\leq 0\}}\sup_{z\geq h'(0),\, \hat{a}\in\hat{A}(z)}\big\{\hat{a} + h(\hat{a})p + \eta z^2 q\big\},\ (p,q)\in\mathbb{R}^2, \tag{7.2}$$

where

$$\mathbf{T}F(y,p) := yp - F(y) - F^\star(p),\ y\geq 0,\ p\in\mathbb{R}.$$

Observe by definition that

$$\mathbf{T}F(y,p) \geq 0, \text{ and } \mathbf{T}F\big(y, F'(y)\big) = 0, \text{ for all } y \geq 0. \tag{7.3}$$

Moreover, the face–lifted principal reward function $\overline{F}$ introduced in (2.5) satisfies

$$\overline{F} - F - \mathbf{T}F\big(\cdot, \delta\overline{F}'\big) = 0, \text{ on } \mathbb{R}_+, \tag{7.4}$$

see the proof of Proposition 2.1 in Appendix A.

Remark 7.1. (i) *Notice that* $\mathbf{L}\overline{F} \le 0$ *on* $\mathbb{R}_+$. *Indeed* $\mathbf{L}\overline{F} = \overline{F} - F - \mathbf{T}\overline{F}(y, \delta\overline{F}') - \mathfrak{I}_0(\overline{F}', \overline{F}'')^+ = -\mathfrak{I}_0(\delta\overline{F}', \delta\overline{F}'')^+ \le 0$ *by* (7.4).

(ii) Equation (7.1) *is equivalent to*

$$v(0) = 0, \text{ and } \mathbf{L}v = 0, \text{ on } (0, \infty), \tag{7.5}$$

which agrees exactly with Equation (3.3) *used in the statement of* Theorem 3.4. *Indeed, if v is a solution of* (7.1), *then* $\mathbf{L}v = 0$ *on* $\mathcal{S}^c$, *where* $\mathcal{S} := \{v = \overline{F}\}$ *is the so–called stopping region, and* $\mathbf{L}v = \mathbf{L}\overline{F} \ge 0$ *on* $\mathcal{S}$, *which implies that* $\mathbf{L}v = 0$ *on* $\mathbb{R}_+$ *by part* (i) *of the present remark.*

Conversely, assuming that $\mathbf{L}v = 0$, *we see that* $v = F + \mathbf{T}F(\cdot, \delta v') + \mathfrak{I}_0(v', v'')^+ \ge F + \mathbf{T}F(\cdot, \delta v')$, *and therefore v is a supersolution of* (7.4). *By* Lemma A.2, *this implies that* $v \ge \overline{F}$, *and we conclude that v solves* Equation (7.1).

We next provide a verification argument which is the standard justification of the importance of the dynamic programming equation (7.1), and which guides the subsequent technical analysis to solve the contracting problem.

Proposition 7.2. (i) *Let* $v \in C^2(\mathbb{R}_+)$ *be a super–solution of* (7.1), *i.e.* $v(0) \ge 0$, *and* $\mathbf{L}v \ge 0$. *Then* $v \ge V$ *on* $\mathbb{R}_+$.

(ii) *Assume further that* $v(0) = 0$, $\mathbf{L}v = 0$ *on the continuation region* $\mathcal{S}^c := \{v > \overline{F}\}$, *and that*

- *for any* $y > 0$, *there exists a maximiser* $\hat{z}(y)$ *of* $I(\delta v', \delta v'')(y)$ *such that the* SDE (5.6), *with, for any* $t \ge 0$, $u(\pi^\star_t) := (F^\star)'\big(\delta v'(Y_t)\big)$, $Z^\star_t := \hat{z}(Y_t)$, *and* $\hat{a}^\star_t \in \hat{A}(Z^\star_t)$, *has a weak solution*;
- *defining* $\tau^\star := \inf\Big\{t : Y_t^{y,Z^\star,\pi^\star} \notin \mathcal{S}^c\Big\}$, *the triplet* $(\tau^\star, Z^\star, \pi^\star)$ *belongs to* $\mathcal{Z}(Y_0)$.

Then $v(Y_0) = V(Y_0)$.

(iii) *If in addition v is ultimately decreasing, then the value function of the principal is* $V^{\mathrm{P}} = v(Y_0^\star)$, *for some* $Y_0^\star \ge u(R)$ *with optimal contract* $\xi^\star$ *given by*

$$u(\xi^\star) := Y_0^\star + r\int_0^{\tau^\star} Z_t^\star \mathrm{d}X_t + r\int_0^{\tau^\star} \big(Y_t - h^\star(Z_t^\star) - u(\pi^\star)\big)\mathrm{d}t.$$

Proof. (i) We first prove that $v \ge V$. For an arbitrary $Y_0 \ge 0$, and $(\tau, Z, \pi) \in \mathcal{Z}(Y_0)$ with corresponding $\hat{a} \in \hat{A}(Z)$, we introduce $\tau_n := \tau \wedge \inf\{t \ge 0 : Y_t \ge n\}$, and we directly compute by Itō's formula that

$$\begin{aligned} v(Y_0) &= \mathrm{e}^{-\rho\tau_n} v(Y_{\tau_n}) - \int_0^{\tau_n} \mathrm{e}^{-\rho t}\Big(-\rho v + \partial_t v + (y + h(\hat{a}_t) - u(\pi_t)) r v_y + \frac{1}{2}\sigma^2 r^2 Z_t^2 v_{yy}\Big)(Y_t)\mathrm{d}t \\ &\quad - \int_0^{\tau_n} \mathrm{e}^{-\rho t} v_y(Y_t) r Z_t \sigma \mathrm{d}W_t^{\hat{a}} \\ &\ge \mathrm{e}^{-\rho\tau_n}\overline{F}(Y_{\tau_n}) + \int_0^{\tau_n} \mathrm{e}^{-\rho t}\big(\mathbf{L}v(Y_t) + \hat{a}_t - \pi_t\big)\mathrm{d}t - \int_0^{\tau_n} \mathrm{e}^{-\rho t} v_y(Y_t) r Z_t \sigma \mathrm{d}W_t^{\hat{a}} \\ &\ge \mathrm{e}^{-\rho\tau_n}\overline{F}(Y_{\tau_n}) + \int_0^{\tau_n} \mathrm{e}^{-\rho t}\big(\hat{a}_t - \pi_t\big)\mathrm{d}t - \int_0^{\tau_n} \mathrm{e}^{-\rho t} v_y(Y_t) r Z_t \sigma \mathrm{d}W_t^{\hat{a}}. \end{aligned}$$

Since v_y is bounded on $[0, \tau_n]$ and Z satisfies (5.3), this implies that

$$v(Y_0) \ge \mathbb{E}^{\mathbb{P}^{\hat{a}}}\bigg[\mathrm{e}^{-\rho\tau_n}\overline{F}(Y_{\tau_n}) + \int_0^{\tau_n} \mathrm{e}^{-\rho t}\big(\hat{a}_t - \pi_t\big)\mathrm{d}t\bigg] \longrightarrow \mathbb{E}^{\mathbb{P}^{\hat{a}}}\bigg[\mathrm{e}^{-\rho\tau}\overline{F}(Y_\tau) + \int_0^{\tau} \mathrm{e}^{-\rho t}\big(\hat{a}_t - \pi_t\big)\mathrm{d}t\bigg], \text{ as } n \longrightarrow \infty,$$

where the last convergence follows from the fact that

$$\big|\mathrm{e}^{-\rho\tau_n}\overline{F}(Y_{\tau_n})\big| \le C\big(1 + \mathrm{e}^{-\rho\tau_n} Y_{\tau_n}^\gamma\big) \le C\bigg(1 + \sup_{0 \le t \le \tau} \big(\mathrm{e}^{-\frac{\rho}{\gamma}t} Y_t\big)^\gamma\bigg),$$

by the estimate stated in Proposition 2.1, together with the integrability conditions on π in (2.11) and on Y in (5.3). By the arbitrariness of $(\tau, Z, \pi) \in \mathcal{Z}(Y_0)$, this shows that $v(Y_0) \ge V(Y_0)$.

To prove (ii), we now repeat the previous argument starting from the control $(\tau^\star, Z^\star, \pi^\star)$ introduced in the statement, and denoting $Y^\star$ the induced controlled state process. As $Z^\star_t$ and $u(\pi^\star_t)$ are maximisers of $I(\delta v', \delta v'')(Y^\star_t)$ and $F^\star(\delta v'(Y^\star_t))$, respectively, we see that for any $\hat{a}^\star \in \hat{A}(Z^\star)$

$$v(Y_0) = \mathbb{E}^{\mathbb{P}^{\hat{a}^\star}}\bigg[\mathrm{e}^{-\rho\tau_n} v(Y^\star_{\tau^\star_n}) + \int_0^{\tau^\star_n} \mathrm{e}^{-\rho t}\big(\hat{a}^\star_t - \pi^\star_t\big)\mathrm{d}t\bigg] \underset{n\to\infty}{\longrightarrow} \mathbb{E}^{\mathbb{P}^{\hat{a}^\star}}\bigg[\mathrm{e}^{-\rho\tau^\star} v\big(Y^\star_{\tau^\star}\big) + \int_0^{\tau^\star} \mathrm{e}^{-\rho t}\big(\hat{a}^\star_t - \pi^\star_t\big)\mathrm{d}t\bigg]$$
$$= \mathbb{E}^{\mathbb{P}^{\hat{a}^\star}}\bigg[\mathrm{e}^{-\rho\tau^\star}\overline{F}\big(Y^\star_{\tau^\star}\big) + \int_0^{\tau^\star} \mathrm{e}^{-\rho t}\big(\hat{a}^\star_t - \pi^\star_t\big)\mathrm{d}t\bigg],$$

since $v = \overline{F}$ on the boundary of $\mathcal{S}$.

(iii) Finally, v is concave by Remark 7.1 (iii). As it is assumed to be ultimately decreasing, the existence of a maximiser $Y^\star_0$ of $v(y)$ on $[u(R), \infty)$ follows, and we obtain that $V^{\mathrm{P}} = \sup_{Y_0 \geq u(R)} v(y) = v(Y^\star_0)$. □

8 On the existence of a Golden Parachute

This section reports the proof of Proposition 3.1 by analysing the action of the operator $\mathbf{L}$ on the face–lifted principal's reward $\overline{F}$. Indeed, if there is a Golden Parachute, then the value function of the principal coincides with $\overline{F}$ on $[y_{\mathrm{gp}}, \infty)$, and we must therefore have $\mathbf{L}\overline{F} \geq 0$ on that interval. In view of Remark 7.1.(i), we must have in fact $\mathbf{L}\overline{F} = 0$ on $[y_{\mathrm{gp}}, \infty)$. By definition of $\overline{F}$, this means that we must have $\mathfrak{I}_0\big(\overline{F}'(y), \overline{F}''(y)\big)^+ = 0$ for any large enough y, and that $\mathfrak{I}_0\big(\overline{F}'(y), \overline{F}''(y)\big) > 0$ for y in a set of non–empty interior. Hence the first part of the statement. The equivalence with the condition written in terms of $\overline{F}^\star$ can be obtained by evaluating $\mathfrak{I}_0\big(\overline{F}'(y), \overline{F}''(y)\big)^+$ at the point $y = \big(\overline{F}'\big)^{-1}(p)$, and by computing that $\overline{F}''(y) = 1/(\overline{F}^\star)''(p)$.

Consequently, we now justify the sufficient conditions of the proposition by verifying some cases where $\overline{F}$ either solves Equation (7.1) on the whole $\mathbb{R}_+$, or nowhere.

Lemma 8.1. *Let $\beta := h'(0)$. We have*

(i) $\mathbf{L}\overline{F}(y) = 0$ *for some* $y > 0$*, if and only if* $\mathfrak{I}_0(\overline{F}', \overline{F}'')(y)^+ = 0$*;*

(ii) *if* $\beta > 0$*, then* $\mathbf{L}\overline{F} = 0$ *on* $[y_1, \infty)$*, for some* $y_1 \leq \big(\overline{F}'\big)^{-1}\Big(\overline{F}'(0) \wedge \frac{-1}{\beta\delta}\Big)$*;*

(iii) *if* $\beta = 0$*, and* $A \supset [0, \bar{a}]$ *for some* $\bar{a} > 0$*, then* $\mathbf{L}\overline{F} < 0$ *on* $(0, \infty)$*.*

Proof. (i) follows immediately from the definition of $\overline{F}$. To prove (ii), recall from Proposition 2.1 that $\overline{F}$ is decreasing and strictly concave on $[0, \infty)$, implying that

$$0 \leq \mathfrak{I}_0\big(\delta\overline{F}', \delta\overline{F}''\big)^+ \leq \sup_{z\in\mathbb{R},\ \hat{a}\in\hat{A}(z)} \big\{\hat{a} + h(\hat{a})\delta\overline{F}'\big\} \leq \sup_{z\in\mathbb{R},\ \hat{a}\in\hat{A}(z)} \big\{\hat{a}(1 + \beta\delta\overline{F}')\big\},$$

where the last inequality is a consequence of the convexity of h, which implies that $h(\hat{a}) \geq h(0) + h'(0)\hat{a} = \beta\hat{a}$. Now, observe that $y_0 := (\overline{F}')^{-1}\big(\overline{F}'(0) \wedge \frac{-1}{\beta\delta}\big)$ is such that $1 + \beta\delta\overline{F}' \leq 0$ on $[y_0, \infty)$. Then, since $\hat{A}(0) = \{0\}$, we deduce that $\mathfrak{I}_0(\overline{F}', \overline{F}'')^+ = 0$ on $[y_0, \infty)$.

(iii) As A contains an interval, h is strictly convex, $\overline{F}$ is concave, and $h'(0) = 0$, we have that

$$\begin{aligned}
\mathfrak{I}_0\big(\overline{F}'(y), \overline{F}''(y)\big)^+ &= \sup_{z\geq 0,\ \hat{a}\in\hat{A}(z)} \big\{\hat{a} + h(\hat{a})\delta\overline{F}'(y) + \eta z^2\delta\overline{F}''(y)\big\} \\
&\geq \sup_{z\geq 0,\ \hat{a}\in\hat{A}(z)\subset[0,\bar{a}]} \big\{\hat{a} + h(\hat{a})\delta\overline{F}'(y) + \eta z^2\delta\overline{F}''(y)\big\} \\
&= \sup_{z\geq 0} \Big\{(h')^{-1}(z) \wedge \bar{a} + h\big((h')^{-1}(z) \wedge \bar{a}\big)\delta\overline{F}'(y) + \eta z^2\delta\overline{F}''(y)\Big\} \\
&= \sup_{a\in[0,\bar{a}]} \big\{a + h(a)\delta\overline{F}'(y) + \eta\big(h'(a)\big)^2\delta\overline{F}''(y)\big\},\ y > 0.
\end{aligned}$$

Now notice that since $h'(0) = 0$, the derivative at $a = 0$ of the map inside the supremum above is equal to $1 > 0$. Therefore, this map is increasing on a right–neighbourhood of 0, and thus for any $y > 0$, we have $\mathfrak{I}_0\big(\overline{F}'(y), \overline{F}''(y)\big)^+ > 0$. □

The next result complements Lemma 8.1.(ii) by exploring the regions where $\mathbf{L}\overline{F} = 0$ may hold under additional conditions on either $\overline{F}$ or h. Let

$$\phi(y) := \sup_{z\geq\beta,\ \hat{a}\in\hat{A}(z)} \big\{\hat{a} + h(\hat{a})\delta\overline{F}'(y) + \eta z^2\delta\overline{F}''(y)\big\},\ y \geq 0,\ \text{so that } \mathfrak{I}_0(\overline{F}', \overline{F}'')^+ = \phi^+,$$

and notice that ϕ is non–increasing, whenever $\overline{F}''$ is.

Lemma 8.2. *Let $\beta := h'(0) > 0$. Then the following holds*

(i) *if $\overline{F}''$ is non–increasing, then*

$$\big\{\mathbf{L}\overline{F} = 0\big\} = \{\mathfrak{I}_0(\overline{F}', \overline{F}'')^+ = 0\} = [y_1, \infty),\ \text{where } y_1 := \inf\big\{y \geq 0 : \phi(y) \leq 0\big\} < \infty.$$

In particular, $\mathbf{L}\overline{F} = 0$ on $\mathbb{R}_+$ if and only if $\mathfrak{I}_0\big(\overline{F}'(0), \overline{F}''(0)\big)^+ = 0$;

(ii) *if $A = [0, \bar{a}]$ for some $\bar{a} > 0$, and $h \in C^3$ with*

$$\inf_{a\in A}\left\{\frac{\big((h')^2\big)''(a)}{h''(a)}\right\} > 0,\ \text{and } \overline{F}'(0) + 2\eta\overline{F}''(0)h''(0) < \frac{-1}{\beta\delta}, \tag{8.1}$$

then $\mathfrak{I}_0(\overline{F}', \overline{F}'')^+ = 0$, on $[0, y_0] \cup [y_1, \infty)$, for some $y_0 > 0$, and $y_1 \leq (F')^{-1}\Big(\overline{F}'(0) \wedge \frac{-1}{\beta\delta}\Big)$;

(iii) *In the context of (NGP3) we have $I\big(\delta\overline{F}', \delta\overline{F}''\big) = 0$,* on $(0, \infty)$.

Proof. (i) Since $\mathfrak{I}_0^+ \geq 0$ and $\mathbf{L}\overline{F} = \mathfrak{I}_0\big(\overline{F}', \overline{F}''\big)^+$ as in the previous proof, the first part of (i) follows immediately, and we see that $y_1 < \infty$ by Lemma 8.1.(iii). Next, we just observe that $\phi(0) \leq 0$ if and only if $\hat{a} + h(\hat{a})\delta\overline{F}'(0) + \eta z^2\delta\overline{F}''(0) \leq 0$, for all $z \geq \beta$ and $\hat{a} \in \hat{A}(z)$, which provides the required condition given that $\overline{F}'(0) \leq 0$.

(ii) The existence of y_1 is direct as in (i). Next, under our assumption on A, we have directly that

$$\mathfrak{I}_0\big(\overline{F}'(y), \overline{F}''(y)\big)^+ = \sup_{a\in A}\Big\{\underbrace{a + h(a)\delta\overline{F}'(y) + \eta\big(h'(a)\big)^2\delta\overline{F}''(y)}_{=:\psi(a,y)}\Big\}^+,\ y > 0.$$

Notice that

$$\partial_{aa}\psi(a, y) = h''(a)\delta\overline{F}'(y) + \eta\big((h')^2\big)''(a)\delta\overline{F}''(y),\ (a, y) \in A \times [0, \infty),$$

so that the first condition in Equation (8.1) implies that $\sup_{a\in A}\partial_{aa}\psi(a, 0) < 0$, and therefore by continuity, $\sup_{a\in A}\partial_{aa}\phi(a, y) \leq 0$ on some small interval $[0, y_0]$, $y_0 > 0$. This shows that $\psi(\cdot, y)$ is concave in a, for y in this interval. We next compute that, reducing y_0 if necessary

$$\partial_a\psi(0, y) = 1 + \beta\delta\Big(\overline{F}'(y) + 2\eta\overline{F}''(y)h''(0)\Big) \leq 0,\ y \in [0, y_0],$$

by the second condition in Equation (8.1) and the continuity of $\overline{F}'(0)''$. Hence, for any $y \in [0, y_0]$, the function $a \longmapsto \psi(a, y)$ is non–increasing, concave, and thus attains its maximum at $a = 0$, implying that $\mathfrak{I}_0\big(\overline{F}'(y), \overline{F}''(y)\big)^+ = \psi(0, y)^+ = 0$.

(iii) This is very similar to (ii). Simply notice that now the first condition in (NGP3) implies that for any $y \geq 0$, $A \ni a \longmapsto \psi(a, y)$ is concave, while the second condition in (NGP3) implies that for any $y \geq 0$, $\partial_a\psi(0, y) \leq 0$, and thus that ϕ is non–increasing in a for any $y \geq 0$, and therefore the desired result. ☐

9 Analysis of the dynamic programming equation

Throughout this section, Assumption 3.5 is in force. We start with proving the strict concavity of v.

Lemma 9.1. *Any continuous solution of* Equation (7.1) *is strictly concave.*

Proof. To prove concavity, suppose to the contrary that v is strictly convex on some non–empty open interval $(y_0, y_1) \subset \mathbb{R}_+$, then we would have that $-v'' < 0$ in the viscosity sense on (y_0, y_1), and thus that $-\mathfrak{I}_0(v', v'')^+ = -\infty$ on (y_0, y_1) (still in the viscosity sense), contradicting the fact that v is a continuous viscosity solution of Equation (7.1).

The strict concavity follows the same line of argument as in [54]. Suppose to the contrary that $v(y) = b_0 + by$ for y in some interval $[y_0, y_1] \subset \mathbb{R}_+$, then

$$b_0 + (1-\delta)by + F^\star(\delta b) - \mathfrak{I}_0(b, 0)^+ = 0,\ y \in [y_0, y_1].$$

If $\delta \neq 1$, this implies that $b = 0$, and therefore $b_0 = \mathfrak{I}_0(0, 0)^+ = \bar{a} > 0$. In particular $\mathfrak{I}_0^+ = \mathfrak{I}_0$.

We next argue that this ODE is in addition uniformly elliptic. This is immediate when $\beta > 0$. For $\beta = 0$, we have $[0, \bar{a}_0] \subset A$ by Assumption 3.5, and

$$\mathfrak{I}_0\big(\delta v', \delta v''\big)^+ \geq \sup_{z\geq 0,\ \hat{a}\in\hat{A}(z)\subset[0,\bar{a}_0]} \big\{\hat{a} + \delta h(\hat{a})v' + \delta\eta z^2 v''(y)\big\} = \sup_{a\in[0,\bar{a}_0]} \big\{ \underbrace{a + \delta h(a)v'(y) + \delta\eta\big(h'(a)\big)^2 v''(y)}_{=:\Phi(a,y)} \big\},$$

and $\Phi(0, y) = 0$, $\partial_a\Phi(0, y) = 1$, for any $y > 0$. Then, for any compact subset of $(0, \infty)$, the supremum in $\mathfrak{I}_0\big(\delta v', \delta v''\big)^+$ is attained on $[\varepsilon, \bar{a}_0]$ for some $\varepsilon > 0$, independent of y (but of course depending on the chosen compact set). Hence, the ODE can always be written in explicit form on any compact subset of $(0, \infty)$.

Consequently, the standard Cauchy–Lipschitz existence and uniqueness theory applies. By uniqueness of the solution of Equation (7.1) with boundary condition $v(y_0) = b_0$ and $v'(y_0) = 0$, we deduce that $v = b_0$ on $[0, y_0]$, contradicting the boundary condition $v(0) = 0$. If $\delta = 1$, we also see by the same argument that $v(y) = b_0 + by$ on $[0, y_0]$, so that $v(0) = 0$ implies that $b_0 = 0$, and we get $F^\star(\delta b) - \mathfrak{I}_0(\delta b, 0)^+ = 0$ and therefore $F^\star(\delta b) = \mathfrak{I}_0(\delta b, 0)^+ = 0$, which again cannot happen. ☐

Remark 9.2. *By* Lemma 9.1, *it is natural to introduce the concave dual function* $v^\star(p) := \inf_{y\geq 0} \{yp - v(y)\}$, $p \in \mathbb{R}$. *Then, if in addition* v *is a* C^2 *solution of the dynamic programming equation,* $v^\star$ *solves the equation*

$$\mathbf{L}^\star v^\star(p) := v^\star(p) - F^\star(\delta p) + (\delta - 1)p(v^\star)'(p) + \mathfrak{I}_0\bigg(p, \frac{1}{(v^\star)''(p)}\bigg)^+ = 0,\ p \in \mathbb{R}. \tag{9.1}$$

This follows by evaluating (7.5) *at the point* $y = (v')^{-1}(p)$ *and by computing that* $v''(y) = 1/(v^\star)''(p)$.[8]

Lemma 9.3. *There is a unique solution* v *of* Equation (7.1), *such that* $0 \leq (v - \overline{F})(y) \leq C\log(1 + \log(1 + y))$, $y \geq 0$, *for some* $C > 0$. *Besides,* v *is strictly concave, ultimately decreasing, and belongs to* $C^2(\mathbb{R}_+)$.

Proof. By Remark 7.1 and Lemma B.1 below, $\overline{F}$ and $\overline{F}_b$ are respectively sub–solution and super–solution of (7.1), for b large enough. Lemma B.3 below shows that this equation satisfies comparison between sub–solutions and super–solutions lying between $\overline{F}$ and $\overline{F}_b$. Then, since $\overline{F}(0) = \overline{F}_b(0) = 0$, we deduce from Perron's existence result, see Crandall, Ishii, and Lions [13, Theorem 4.1], that Equation (7.1) has a viscosity solution v_b lying between $\overline{F}$ and $\overline{F}_b$. Using Lemma B.3, this solution must be unique, and thus does not depend on b, so that we can denote it by v.

Recall that v is strictly concave by Lemma 9.1. Since it is below $\overline{F}_b$, it has to be ultimately decreasing, as $\overline{F}_b$ is. In addition, v is differentiable Lebesgue–a.e., and we may define the measurable set

$$I := \big\{y \geq 0 : v(y) - F(y) - \mathbf{T}F\big(y, \delta v'(y)\big) = 0\big\}.$$

For almost–every $y \in I$, we have by using the inequality $v - \overline{F} \geq 0$ (which must hold for any y since it holds in the viscosity sense and both v and $\overline{F}$ are continuous) and the definitions of $F^\star$ and $\mathbf{T}$

$$F^\star\big(\delta v'(y)\big) = \delta y v'(y) - v(y) \leq \delta y v'(y) - F(y) \leq F^\star\big(\delta v'(y)\big).$$

[8] Such a transformation can also be conducted if the solution is expressed in the sense of viscosity solutions (as it will be needed later), but one has to be careful as strict convexity is not sufficient, see Alvarez, Lasry, and Lions [3, Proposition 5] and the remark after its proof.

Therefore, the above inequalities must be equalities almost–everywhere on I, which in particular implies that $v = \overline{F}$, a.e. on I, and therefore everywhere on I since v and $\overline{F}$ are continuous. We also automatically have that v is C^2 on I.

On the other hand, we have on $\mathbb{R}_+ \setminus I$ that $\mathfrak{I}_0\big(v', v''\big)^+ = \mathfrak{I}_0\big(v', v''\big) > 0$, almost–everywhere. Arguing as in the proof of Lemma 9.1, we see that v is a viscosity solution of a (locally) uniformly elliptic ODE

$$v'' = L(y, v, v'), \text{ on } \mathbb{R}_+ \setminus I,$$

for some locally Lipschitz nonlinearity L. Consequently, since $\mathbb{R}_+ \setminus I = \big(v - \overline{F}\big)^{-1}((0,\infty))$ is an open set by continuity of v and $\overline{F}$, the standard Cauchy–Lipschitz theorem then shows that v must also be smooth on $\mathbb{R}_+ \setminus I$. Finally, any point y_0 on the boundary $\partial I \cap (0,\infty)$ is a minimiser of the difference $v - \overline{F}$. As such

- we have by the first order condition that $v'(y_0-) - \overline{F}'(y_0) \leq 0 \leq v'(y_0+) - \overline{F}'(y_0)$, implying by concavity that v is differentiable at y_0 and $v'(y_0) = \overline{F}'(y_0)$;
- we also have $v''(y_0) \geq \overline{F}''(y_0)$, implying by continuity that

$$0 = \mathbf{L}v(y_0) = -\mathfrak{I}_0\big(\overline{F}'(y_0), v''(y_0)\big) \leq -\mathfrak{I}_0\big(\overline{F}'(y_0), \overline{F}''(y_0)\big) \leq 0,$$

which implies that $v''(y_0) = \overline{F}''(0)$, and $\mathfrak{I}_0\big(\overline{F}'(y_0), \overline{F}''(y_0)\big)^+ = 0$. Hence, $y_0 \in \big\{\mathfrak{I}_0\big(\overline{F}', \overline{F}''\big)^+ = 0\big\}$. ☐

Lemma 9.4. *Let $\beta > 0$, assume that F and $\mathfrak{I}_0$ are analytic, and let v be the solution from* Lemma 9.3. *Then $\big\{v = \overline{F}\big\} \cap (0,\infty)$ is a possibly empty interval unbounded to the right.*

Proof. (i) Suppose $v = \overline{F}$ on some interval $[y_0, y_1]$, and let us show that we must have $v = \overline{F}$ on $[y_0, \infty)$, which implies in particular that $\mathfrak{I}_0(\overline{F}', \overline{F}'')^+ = 0$ on $[y_0, \infty)$, and consequently $I \cap (0,\infty)$ has the claimed form, where the set I is introduced in the proof of Lemma 9.3.

To see this, suppose to the contrary that $\mathbf{L}\overline{F} < 0$ on some neighbourhood (y_1, y_1') to the right of y_1, and denote $\phi(y) := \big(v - F\big)(y) - \mathbf{T}F(y, \delta v'(y))$. Clearly, $\phi \geq 0$ on $\mathbb{R}_+$ and $\phi = 0$ on $[y_0, y_1]$, so that

$$0 = \big(v - \overline{F}\big)(y_1) = \min_{y \in (y_0, y_1')} \big(v - \overline{F}\big)(y), \text{ and } 0 = \phi(y_1) = \min_{y \in (y_0, y_1')} \phi(y). \tag{9.2}$$

Next, reducing $y_1' > y_1$ if necessary, we have that on (y_1, y_1'), both v and $\overline{F}$ are strictly concave and decreasing, and we claim that the supremum in $\mathfrak{I}_0(\delta v', \delta v'')^+$ cannot be attained on a right–neighbourhood of β. Indeed, otherwise $\mathfrak{I}_0(\delta v', \delta v'')^+$ would be equal to 0 on some interval at the right of y_1, since we have

$$\mathfrak{I}_0(v', v'') = \sup_{\hat{a} \in \hat{A}(\beta)} \big\{\hat{a} + \delta v' h(\hat{a}) + \delta\eta v''\beta^2\big\} = \delta\eta v''\beta^2 < 0,$$

so that, by continuity, this supremum remains negative on a right–neighbourhood of y_1, and $\mathfrak{I}_0(v', v'')^+ = \mathfrak{I}_0(v', v'')^+ = 0$. Consequently, v and $\overline{F}$ are both solutions of the ODE $w - \delta y w' + F^\star(\delta w') = 0$ with same boundary condition at y_1, implying that $v = \overline{F}$ on some interval at the right of y_1, and therefore contradicting the definition of y_1 as a (right) extreme point of $\big\{v = \overline{F}\big\}$.

Moreover, since A is bounded, the supremum over z must be attained on a compact set, meaning that for some $\beta_{\min} > \beta$ and some finite $\beta_{\max} > \beta_{\min}$

$$\mathfrak{I}_0(v', v'')^+ = \mathfrak{I}_0(v', v'') = \max_{z \in [\beta_{\min}, \beta_{\max}],\ \hat{a} \in \hat{A}(z)} \big\{\hat{a} + \delta v' h(\hat{a}) + \delta\eta z^2 v''\big\}.$$

By our assumptions on F and $\mathfrak{I}_0$, the function v solves on $\mathbb{R}_+ \setminus I$ an explicit ODE with analytic non–linearity. Indeed, $\mathfrak{I}_0$ is invertible and does not take the value 0, implying that its reciprocal function is still analytic. By Cauchy–Kowaleski's theorem, we deduce that v is also analytic on $[y_1, y_1')$, and therefore C^∞. Using all the above results, we have that on (y_1, y_1')

$$\begin{aligned}
\mathfrak{I}_0\big(\overline{F}', \overline{F}''\big)^+ > 0 = -\mathbf{L}v &= \mathfrak{I}_0(v', v'')^+ - \phi \\
&\geq \mathfrak{I}_0(F', F'')^+ - \phi + \delta\Big(\eta\beta^2\big(v'' - \overline{F}''\big)^+ - \eta\beta_{\max}^2\big(v'' - \overline{F}''\big)^- + h(\underline{a})\big(v' - \overline{F}'\big)^+\Big) \\
&\quad - \delta h(\overline{a})\big(v' - \overline{F}'\big)^-.
\end{aligned}$$

Denoting $c_1 := (\eta\beta^2) \wedge h(\underline{a}) > 0$ and $c_2 := (\eta\beta^2_{\max}) \vee h(\overline{a}) > 0$, we have then

$$\delta c_1\big[(v'' - \overline{F}'')^+ + \big(v' - \overline{F}'\big)^+\big] - \delta c_2\big[(v'' - \overline{F}'')^- + \big(v' - \overline{F}'\big)^-\big] < \phi, \text{ on } (y_1, y_1'). \tag{9.3}$$

As $v'(y_1) = \overline{F}'(y_1)$, $v''(y_1+) \geq \overline{F}''(y_1)$ and $\phi(y_1) = 0$, by (9.2), these inequalities imply by sending $y \searrow y_1$, that $(v'' - \overline{F}'')(y_1) = 0$. This implies in turn that ϕ is differentiable at y_1, and, since v and $\overline{F}$ are C^∞ on the right of y_1, and y_1 is a local minimum of $v - \overline{F}$

$$\big(v - \overline{F}\big)'''(y_1+) \geq 0, \text{ and } \phi'(y_1) = -v''(y_1)\big(y_1 - (F')^{-1} \circ v'(y_1)\big) = 0,$$

Then, dividing (9.3) by $(y - y_1)$ and sending $y \searrow y_1$, we get $(v''' - \overline{F}''')(y_1+) = 0$, and we deduce that ϕ is twice differentiable at y_1, and

$$\big(v - \overline{F}\big)''''(y_1+) \geq 0, \text{ and } \phi''(y_1) = 0.$$

Direct iteration of this argument shows that $\big(v - \overline{F}\big)$ is infinitely differentiable at the point y_1, with 0 derivative of any order. Since it is an analytic function on a right–neighbourhood on y_1, we deduce that $v - \overline{F} = 0$ on a right–neighbourhood of y_1, which contradicts the definition of y_1. □

Together with the previous lemmatas, the following result concludes the proof of Theorem 3.6.(i)–(ii)–(iii).

Lemma 9.5. *Let v be the solution constructed in* Lemma 9.3, *and define $\mathcal{S} := \{v = \overline{F}\}$. Then*

(i) *we have $v'(0) \geq 0$, and whenever $F'(0) = 0$, we have $v'(0) > 0$ if and only if $I\big(0, \delta\overline{F}''(0)\big) > 0$;*

(ii) *if $\beta = 0$, then $\mathcal{S} = \{0\}$.*

Proof. (i) By continuity, we have

$$0 = \mathbf{L}v(0) = F^\star\big(\delta v'(0)\big) - I\big(\delta v'(0), \delta v''(0)\big).$$

Since $F^\star \leq 0$ and $I \geq 0$, it follows that $F^\star\big(\delta v'(0)\big) = 0$, and consequently $v'(0) \geq 0$, since $F^\star < 0$ on $(-\infty, 0)$. Because $I(p, q)$ is non–decreasing in p, we deduce that $0 = I\big(\delta v'(0), \delta v''(0)\big) \geq I\big(0, \delta v''(0)\big)$. However, under our assumptions, $I\big(0, \delta\overline{F}''(0)\big) > 0$. Then it follows from the non–decrease of $I(p, q)$ in q that $v''(0) < \overline{F}''(0)$. Consequently, we have $v \geq \overline{F}$, $v(0) = \overline{F}(0)$, $v'(0) \geq \overline{F}'(0) = 0$ and $v''(0) < \overline{F}''(0) \leq 0$, implying that $v'(0) > 0$, as required.

(ii) By Lemma 8.1.(iii), we know that $\overline{F}$ never solves the ODE, and we claim that this implies that $v > \overline{F}$ on $(0, \infty)$. Indeed, notice that any contact point y_0 of v and $\overline{F}$ is a local minimiser of the difference $v - \overline{F}$, so that $v' = \overline{F}'$ at such a point. Then, as $I \geq 0$, it follows from Equation (7.1) that $\mathbf{T}F\big(y_0, \delta\overline{F}'(y_0)\big) = 0$ which cannot happen unless $y_0 = 0$. □

We finally prove Theorem 3.6.(iv) by using the verification result provided in Proposition 7.2, in order to show that one can identify the value function of the principal with the function v constructed in Theorem 3.6.

Proof of Theorem 3.6.(iv). The existence of $\hat{y}$ is immediate by the strict concavity of v, and the fact that it is ultimately decreasing. Then, the rest of the proof simply requires to check that the assumptions in Proposition 7.2 are satisfied here. First of all, notice that the map $\hat{z}$ is bounded from above since A is compact, and from below by β, and it is continuous on $\mathcal{S}^c$ because v is C^2 there. Similarly, the map $\hat{\pi}$ is bounded on $\mathcal{S}^c$, from below by 0 and from above as well because $\mathcal{S}^c$ is a bounded set under our assumptions. The existence of a unique weak solution for $\widehat{Y}$ is then direct from Stroock and Varadhan [64, Corollary 6.4.4].[9] Notice in addition that $\widehat{Y}$ has moments of any order under $\mathbb{P}^0$ (and thus under any $\mathbb{P}^\alpha$, $\alpha \in \mathcal{A}$, recall that A is compact). It remains to verify that $\hat{\tau}$ satisfies (2.11). However, $\widehat{Y}$ is a one–dimensional Markov process for which the boundaries 0 and $y_{\rm gp}$ are regular and accessible, it is therefore well–known that $\hat{\tau}$ is finite with probability 1. Since A is compact, the densities $\mathrm{d}\mathbb{P}^\alpha/\mathrm{d}\mathbb{P}^0$ all have moments of any order, uniformly in $\alpha \in \mathcal{A}$, from which it is immediate that (2.11) holds. □

[9] The drift of $\widehat{Y}$ is not bounded as required in [64, Corollary 6.4.4], because of the term $r\widehat{Y}$. However, it suffices to apply the result to $\big(\mathrm{e}^{rt}\widehat{Y}_t\big)_{t\geq 0}$.

10 Holmström–Milgrom's model with early retirement

In this section, we explore whether adding risk–aversion for the principal fundamentally modifies the results for the existence of a Golden Parachute. As such, we consider a variation which mixes both Sannikov's model [54] and Holmström and Milgrom's [34]. There is a finite horizon $T > 0$, the contracts only stipulate a lump–sum payment at τ (meaning that π is always 0). Moreover, the agent also has a CARA utility and, abusing notations slightly, for any admissible contract $\mathbf{C} := (\tau, \xi)$, we have

$$V^{\mathrm{A}}(\mathbf{C}) := \sup_{\alpha\in\mathcal{A}} J^{\mathrm{A}}(\mathbf{C},\alpha), \text{ where } J^{\mathrm{A}}(\mathbf{C},\alpha) := \mathbb{E}^{\mathbb{P}^\alpha}\bigg[\mathrm{U}_{\mathrm{A}}\bigg(\xi - \int_0^\tau h(\alpha_s)\mathrm{d}s\bigg)\bigg],$$

where $\mathrm{U}_{\mathrm{A}}(x) := -\mathrm{e}^{-\psi x}$, $x \in \mathbb{R}$, for some $\psi > 0$. Using again the general approaches from Cvitanić, Possamaï, and Touzi [19] and Lin, Ren, Touzi, and Yang [41][10], we can show that the lump–sum payment ξ at time τ takes the form

$$\xi = Y_\tau^{Y_0,Z} = Y_0 + \int_0^\tau Z_t \mathrm{d}X_t - \int_0^\tau \bigg(h^\star(Z_t) - \frac{1}{2}\psi\sigma^2 Z_t^2\bigg)\mathrm{d}t,$$

where this time $Y^{Y_0,Z}$ should be interpreted as the certainty equivalent of the agent. In turn, the principal's problem boils now down to

$$V^{\mathrm{P}} = \sup_{Y_0\geq R} \sup_{(\tau,Z,\hat{a})\in\mathfrak{Z}(Y_0)\times\hat{A}(Z)} \mathbb{E}^{\mathbb{P}^{\hat{a}}}\Big[\mathrm{U}_{\mathrm{P}}\Big(X_{\tau\wedge T} - Y_{\tau\wedge T}^{Y_0,Z}\Big)\Big], \tag{10.1}$$

where $\mathrm{U}_{\mathrm{P}}(x) := -\mathrm{e}^{-\eta x}$, $x \in \mathbb{R}$, for some $\eta > 0$, and $\mathfrak{Z}(Y_0)$ is a proper reformulation of $\mathcal{Z}(Y_0)$ in this context.

In the present context, a Golden Parachute is a situation where the optimal retirement time chosen by the principal lies in $(0, T)$ with positive probability. As $V^{\mathrm{P}}(t,x,y) = U_{\mathrm{P}}(x-y)$ upon retirement, the existence of a Golden Parachute is reduced to the non–emptiness of the stopping region before maturity $\big\{(t,x,y) : t < T \text{ and } V^{\mathrm{P}}(t,x,y) = U_{\mathrm{P}}(x-y)\big\}$.

Similar to Section 8, we shall explore the potential existence of a Golden Parachute by analysing the action of the dynamic programming operator on the obstacle $U_{\mathrm{P}}(x-y)$. In the present context, the dynamic programming equation corresponding to the reduced principal problem (10.1) is given by

$$\min\bigg\{v - \mathrm{U}_{\mathrm{P}}(x-y); -\partial_t v - \frac{\sigma^2}{2}v_{xx} - M(v_x, v_y, v_{yy}, v_{xy})\bigg\} = 0,\ v(t,x,0) = \mathrm{U}_{\mathrm{P}}(x),\ \text{on } [0,T)\times\mathbb{R}\times(0,\infty),$$

$$v(T,x,y) = \mathrm{U}_{\mathrm{P}}\big(x-y\big),\ \text{on } (x,y)\in\mathbb{R}\times[0,\infty),$$

where

$$M(q_1,q_2,\gamma_1,\gamma_2) := \sup_{(z,\hat{a})\in\mathbb{R}\times\hat{A}(z)} \bigg\{\hat{a}(z)q_1 + \bigg(\frac{\sigma^2\psi}{2}z^2 + h\big(\hat{a}(z)\big)\bigg)q_2 + \frac{\sigma^2}{2}z^2\gamma_1 + \sigma^2 z\gamma_2\bigg\}.$$

For the sake of clarity, the following result focuses on the case where the agent's cost of effort is quadratic.

Lemma 10.1. *In the present setting, assume that $A = [0,\infty)$, and $h(a) = ha^2/2 + \beta a$, $a \geq 0$, for some $h > 0$ and $\beta \geq 0$. A necessary condition for a* Golden Parachute *to exist is*

$$\beta \geq \underline{\beta} := \left(1 - \sqrt{\frac{\sigma^2 h\psi(1+\sigma^2 h\eta)}{1+\sigma^2 h(\psi+\eta)}}\right)^+.$$

Proof. With the choice of A and h in the statement of the lemma, we have that when $\sigma^2 h(\psi q_2 + \gamma_1) + q_2 < 0$, and $\psi q_2 + \gamma_1 < 0$

$$M(q_1,q_2,\gamma_1,\gamma_2) = \frac{1}{2h}\max\bigg\{\sup_{z\geq\beta}\Big\{\big(\sigma^2 h(\psi q_2+\gamma_1)+q_2\big)z^2 + 2\big(\sigma^2 h\gamma_2 + q_1\big)z - \beta\big(2q_1+\beta q_2\big)\Big\},$$

$$\sup_{z<\beta}\Big\{\sigma^2 h\big(\psi q_2+\gamma_1\big)z^2 + 2\sigma^2 h\gamma_2 z\Big\}\bigg\}.$$

[10] See also Cvitanić, Possamaï, and Touzi [18], Aïd, Possamaï, and Touzi [2], Élie and Possamaï [23], Élie, Mastrolia, and Possamaï [25], or Élie, Hubert, Mastrolia, and Possamaï [24] for models with CARA utilities using this approach.

We now evaluate M along the appropriate derivatives of the map $(x,y)\longmapsto \mathrm{U_P}(x-y)$, which corresponds to the substitutions

$$q_1\longleftrightarrow -\eta\mathrm{U_P}(x-y),\ q_2\longleftrightarrow \eta\mathrm{U_P}(x-y),\ \gamma_1\longleftrightarrow \eta^2\mathrm{U_P}(x-y),\ \gamma_2\longleftrightarrow -\eta^2\mathrm{U_P}(x-y).$$

Defining $\mathcal{M}(x,y):=M\big(-\eta\mathrm{U_P}(x-y),\eta\mathrm{U_P}(x-y),\eta^2\mathrm{U_P}(x-y),-\eta^2\mathrm{U_P}(x-y)\big)$, we compute directly that for $\psi>0$, $\frac{\eta}{\psi+\eta}<\frac{\sigma^2h\eta+1}{\sigma^2h(\psi+\eta)+1}$, from which we have

$$\mathcal{M}(x,y)=-\frac{\eta\mathrm{U_P}(x-y)}{2h}\begin{cases}\dfrac{(\sigma^2h\eta+1)^2}{\sigma^2h(\psi+\eta)+1}+\beta\big(\beta-2\big),\ \text{if}\ \dfrac{\eta}{\psi+\eta}>\beta,\\ \dfrac{\sigma^2h\eta^2}{\psi+\eta},\ \text{if}\ \dfrac{\sigma^2h\eta+1}{\sigma^2h(\psi+\eta)+1}<\beta.\end{cases}$$

In the intermediary case $\frac{\eta}{\psi+\eta}\leq\beta\leq\frac{\sigma^2h\eta+1}{\sigma^2h(\psi+\eta)+1}$, to find the maximum in the expression of $\mathcal{M}$, we need to solve the inequality

$$\frac{(\sigma^2h\eta+1)^2}{\sigma^2h(\psi+\eta)+1}+\beta\big(\beta-2\big)\geq\frac{\sigma^2h\eta^2}{\psi+\eta}\Longleftrightarrow \beta^2-2\beta+\frac{\sigma^2h\eta^2+2\sigma^2h\eta\psi+\eta+\psi}{(\psi+\eta)(1+\sigma^2h(\psi+\eta))}\geq 0.$$

It is straightforward to check that the second–order polynomial in β above has two positive roots, the smallest one only belonging to $\big[\frac{\eta}{\psi+\eta},\frac{\sigma^2h\eta+1}{\sigma^2h(\psi+\eta)+1}\big]$, and therefore that for $\frac{\eta}{\psi+\eta}\leq\beta\leq\frac{\sigma^2h\eta+1}{\sigma^2h(\psi+\eta)+1}$

$$\beta^2-2\beta+\frac{\sigma^2h\eta^2+2\sigma^2h\eta\psi+\eta+\psi}{(\psi+\eta)(1+\sigma^2h(\psi+\eta))}\geq 0\Longleftrightarrow \beta\leq 1-\psi\sigma\sqrt{\frac{h}{(\psi+\eta)(1+\sigma^2h(\psi+\eta))}}=:\overline{\beta}.$$

Overall, we thus have

$$\mathcal{M}(x,y)=-\frac{\eta\mathrm{U_P}(x-y)}{2h}\begin{cases}\dfrac{(\sigma^2h\eta+1)^2}{\sigma^2h(\psi+\eta)+1}+\beta\big(\beta-2\big),\ \text{if}\ \overline{\beta}>\beta,\\ \dfrac{\sigma^2h\eta^2}{\psi+\eta},\ \text{if}\ \overline{\beta}\leq\beta.\end{cases}$$

Hence, the diffusion operator in the PDE applied to $(x,y)\longmapsto\mathrm{U_P}(x-y)$ is exactly equal to

$$-\frac{\eta\mathrm{U_P}(x-y)}{2h}\begin{cases}\eta\sigma^2h-\dfrac{(\sigma^2h\eta+1)^2}{\sigma^2h(\psi+\eta)+1}-\beta\big(\beta-2\big),\ \text{if}\ \overline{\beta}>\beta,\\ \eta\sigma^2h-\dfrac{\sigma^2h\eta^2}{\psi+\eta},\ \text{if}\ \overline{\beta}\leq\beta.\end{cases}$$

When $\beta\geq\bar{\beta}$, the above quantity is always non–negative, and when $\beta<\overline{\beta}$, one can check that the second–order polynomial in β we get always has real roots, that the largest one is above $\overline{\beta}$, the lowest one is below $\overline{\beta}$, and therefore that it will be non–negative if and only if $\beta\geq\underline{\beta}$. □

The message from the previous lemma is that whenever the lower bound for β given in the statement is positive, which is equivalent to having

$$1>\sigma^2h\eta\big(\sigma^2h\psi-1\big), \tag{10.2}$$

a Golden Parachute cannot exist for small values of β. Since the classical Holmström and Milgrom's model has exactly $\beta=0$, Golden Parachute cannot exist as soon as Equation (10.2) holds, which happens for small risk aversions for either the principal or the agent. This means that golden parachutes can only arise in situations where we have either high risk–aversions, or high marginal costs for the agent, or high uncertainty on the returns of the output X. Though one should keep in mind that the setting is now somewhat different from [54], this is in stark contrast with the statement that 'if we allow the principal to be explicitly risk averse we can expect the qualitative features of the optimal contract (including retirement) to be the same as with risk neutrality.' ([54, Remark 3]).

Appendices

A Face–lifted principal's reward

This section is dedicated to the proof of Proposition 2.1.

A.1 Very impatient principal may reduce her loss to zero

We first consider the case $\rho \geq \gamma r$ of Proposition 2.1.(i). Notice that we always have $\overline{F}(0) = 0$, and that since F is non–positive, we have $\overline{F} \leq 0$. Besides, by our assumptions on u, there exists $M > 0$ and $C > 0$, such that for any $y \geq M$, $F(y) \geq -Cy^\gamma$. Fix some some $y_0 > 0$ and some $\varepsilon > 0$, and consider then the following control

$$p(t) := \mathbf{1}_{[t^\star,\infty)}(t)\rho\varepsilon y(t),\ t \geq 0,$$

where $t^\star$ is the first instant at which $y^{y_0,0}$ reaches the value M. We immediately have that

$$y^{y_0,p}(t) = y_0 \mathrm{e}^{rt}\mathbf{1}_{[0,t^\star)}(t) + M\mathrm{e}^{(r-\rho\varepsilon)(t-t^\star)}\mathbf{1}_{[t^\star,\infty)}(t),\ t \geq 0.$$

In particular, $T_0^{y_0,p} = \infty$, and for $T > t^\star$

$$\begin{aligned}
\overline{F}(y_0) &\geq \mathrm{e}^{-\rho T}F\big(y^{y_0,p}(T)\big) + \int_0^T \rho\mathrm{e}^{-\rho t}F\big(p(t)\big)\mathrm{d}t \\
&= \mathrm{e}^{-\rho T}F\big(M\mathrm{e}^{(r-\rho\varepsilon)(T-t^\star)}\big) + \int_{t^\star}^T \rho\mathrm{e}^{-\rho t}F\big(\rho\varepsilon M\mathrm{e}^{(r-\rho\varepsilon)(t-t^\star)}\big)\mathrm{d}t \\
&\geq -CM^\gamma \mathrm{e}^{\gamma(\rho\varepsilon-r)t^\star-\rho T(1-\gamma\frac{\rho}{r}+\varepsilon\gamma)} - \frac{C\varepsilon^\gamma\rho^\gamma M^\gamma\mathrm{e}^{-\rho t^\star}}{1-\gamma\frac{\rho}{r}+\gamma\varepsilon}\Big(1-\mathrm{e}^{-\rho(T-t^\star)(1-\gamma\frac{\rho}{r}+\gamma\varepsilon)}\Big) \underset{T\to\infty}{\longrightarrow} \frac{-C\varepsilon^\gamma\rho^\gamma M^\gamma\mathrm{e}^{-\rho t^\star}}{1-\gamma\delta+\gamma\varepsilon},
\end{aligned}$$

by the condition $\rho \geq \gamma r$. As $\gamma > 1$, the last limit converges to 0 as $\varepsilon \searrow 0$.

A.2 Non–degenerate face–lifted utility

By standard control theory, the Hamilton–Jacobi equation corresponding to the mixed control–stopping problem defining $\overline{F}$ is

$$\min\Big\{w - F, w - \delta y w' + F^\star\big(\delta w'\big)\Big\} = 0,\ \text{on } (0,\infty),\ w(0) = 0.$$

Notice first that, similar to Remark 7.1.(ii), this ODE is equivalent to

$$w - \delta y w' + F^\star\big(\delta w'\big) = 0,\ \text{on } (0,\infty),\ w(0) = 0, \tag{A.1}$$

as the last equation implies that $w - F = \delta y w' - F - F^\star\big(\delta w'\big) \geq 0$.

Now, this ODE has 0 as a trivial solution, and when $\delta = 1$, it has a unique strictly concave solution given by F. The following lemma addresses the general case.

Lemma A.1. *Let $\delta \neq 1$ and $\gamma\delta > 1$, Denote by $w^\star$ be the function introduced in (2.9), and let $w := (w^\star)^\star$ be its concave conjugate. Then w is a solution of (A.1) satisfying $\bar{c}_0(-1+y^\gamma) \leq w(y) \leq \bar{c}_0(-1+y^\gamma)$. Moreover, $w'(0) = F'(0)\frac{1}{\delta}\mathbf{1}_{\{\delta\geq 1\}}$.*

Proof. Notice first that if w solves Equation (A.1), then, whenever $w'(0)$ is finite, by letting y go to 0, we get $F^\star\big(\delta w'(0)\big) = 0$. This implies that $\delta w'(0) \geq F'(0)$, and as $w \geq F$ we deduce that $w'(0) \geq \frac{F'(0)}{1\vee\delta}$, an inequality obviously satisfied when $w'(0) = \infty$, which is the only possible infinite value by concavity of w. We now show that

$$w'(0) = F'(0)\frac{1}{\delta}\mathbf{1}_{\{\delta\geq 1\}}. \tag{A.2}$$

To see this, we consider the following alternative cases.

• $\delta \geq 1$. Assume to the contrary that $\delta w'(0) > F'(0)$, then $\delta w' > F'(0)$ on $[0, \varepsilon)$ for some $\varepsilon > 0$, by continuity. This in turn implies that $F^\star(\delta w') = 0$ on $[0, \varepsilon)$, and equation (A.1) reduces to $w(y) - \delta y w'(y) = 0$, on $[0, \varepsilon)$, and we get

$$w(y) = \text{Const.} \frac{w(y_0)}{y_0^{\frac{1}{\delta}}} y^{\frac{1}{\delta}}, 0 < y_0 \leq y < \varepsilon. \tag{A.3}$$

As $w(0) = 0$ and $w \leq 0$, we see that $w(y_0)/y_0^{\frac{1}{\delta}} \longrightarrow 0$, as $y_0 \searrow 0$, and therefore $w = 0$ on $[0, \varepsilon)$, contradicting the strict concavity of w.

• $\delta < 1$. Since $w'(0) \geq F'(0)$, we have again $F^\star(\delta w') = 0$ on $[0, \varepsilon)$, for some $\varepsilon > 0$, and by arguying as in the previous case, we arrive to again to the same conclusion (A.3). However, as $\delta < 1$, the only way to avoid explosion of $w(y_0)/y_0^{\frac{1}{\delta}}$ as $y \searrow 0$ is that $w'(0) = 0$.

In particular, notice that (A.2) implies that any strictly concave solution of (A.1) is decreasing. Next, we can use convex duality and consider the dual of w, given by $w^\star(p) := \inf_{y \geq 0} \{py - w(y)\}$. Notice that since we proved that we needed to have $w'(0) = F'(0)\frac{1}{\delta}\mathbf{1}_{\{\delta > 1\}} =: f_\delta$, the domain over which $w^\star$ is naturally defined is $(-\infty, f_\delta]$. As such the ODE satisfied by $w^\star$ is

$$-w^\star(p) + (1 - \delta)p\big(w^\star\big)'(p) + F^\star(\delta p) = 0,\ p < f_\delta,\ w^\star(f_\delta) = 0. \tag{A.4}$$

This linear ODE has the generic solution, for any $C \in \mathbb{R}$ and $\varepsilon > 0$

$$w^\star(p) = (-p)^{-\frac{1}{\delta - 1}} \left(C - \frac{1}{1 - \delta} \int_p^{f_\delta - \varepsilon} \frac{F^\star(\delta x)}{(-x)^{1 + \frac{1}{1 - \delta}}} \mathrm{d}x \right),\ p < f_\delta. \tag{A.5}$$

We need to study the behaviour of the solution when p goes to $f_\delta-$. We will thus consider three cases.

Case 1: $f_\delta < 0$**.** In this case, we can take directly $\varepsilon = 0$ in Equation (A.5), as the integrand has no singularity, and the boundary condition at f_δ imposes that $C = 0$, so that the solution is uniquely determined by

$$w^\star(p) = \frac{(-p)^{-\frac{1}{\delta - 1}}}{\delta - 1} \int_p^{f_\delta} \frac{F^\star(\delta x)}{(-x)^{1 + \frac{1}{1 - \delta}}} \mathrm{d}x,\ p \leq f_\delta.$$

In this case, we necessarily have $\delta > 1$, therefore, by Lemma A.4, $w^\star$ is strictly concave, increasing, and $w^\star \leq F^\star$. This immediately proves that w is unique, strictly concave, decreasing, and above F. Besides, the explicit formula we obtained shows by direct integration and using Equation (2.8) that $w^\star$ satisfies also Equation (2.8) with appropriate constants, which directly implies the required inequalities for w.

Case 2: $f_\delta = 0$ **and** $\delta > 1$**.** Under this condition, we can take $\varepsilon = 0$ in Equation (A.5), leading to

$$\frac{(-p)^{-\frac{1}{\delta - 1}}}{1 - \delta} \int_p^0 \frac{F^\star(\delta x)}{(-x)^{1 + \frac{1}{1 - \delta}}} \mathrm{d}x \underset{p \to 0-}{=} o(1),$$

and thus converges to 0 as p goes to $0-$. Therefore, when $\delta > 1$, we need to take again $C = 0$ in Equation (A.5) to satisfy the boundary condition $w^\star(0) = 0$, and our solution is uniquely determined. Moreover, by Lemma A.4, $w^\star$ is strictly concave, increasing, and $w^\star \leq F^\star$. This immediately proves that w is unique, strictly concave, decreasing, and above F. We deduce that w satisfies the required inequalities as in the previous case.

Case 3: $\delta < 1$**.** In this case, it can be checked that for any $C \in \mathbb{R}$ and $\varepsilon > 0$, we have

$$\lim_{p \to 0-} (-p)^{-\frac{1}{\delta - 1}} \left(C - \frac{1}{1 - \delta} \int_p^{-\varepsilon} \frac{F^\star(\delta x)}{(-x)^{1 + \frac{1}{1 - \delta}}} \mathrm{d}x \right) = 0.$$

We therefore have, *a priori*, infinitely many possible solutions to the ODE. However, notice that the growth imposed on w translates into $\bar{c}_0^*(-1 + |p|^{\frac{\gamma}{\gamma - 1}}) \leq w^\star(p) \leq \bar{c}_1^*(1 + |p|^{\frac{\gamma}{\gamma - 1}})$, and this implies that

$$|p|^{-\frac{1}{1 - \delta}} w^\star(p) \leq \bar{c}_1^* \Big(|p|^{-\frac{1}{1 - \delta}} + |p|^{\frac{1 - \gamma\delta}{(1 - \delta)(\gamma - 1)}} \Big) \underset{p \to -\infty}{\longrightarrow} 0, \tag{A.6}$$

as $\gamma\delta > 1$. Then, C and ε must be such that $C = \frac{1}{1-\delta}\int_{-\infty}^{-\varepsilon} \frac{F^\star(\delta x)}{(-x)^{1+\frac{1}{1-\delta}}}\mathrm{d}x$, where the finiteness of the last integral is satisfied in our setting again by $\gamma\delta > 1$. Consequently, $w^\star$ is uniquely determined and given by

$$w^\star(p) = \frac{(-p)^{-\frac{1}{\delta-1}}}{(1-\delta)}\int_{-\infty}^{p} \frac{F^\star(\delta x)}{(-x)^{1+\frac{1}{1-\delta}}}\mathrm{d}x,\ p \le 0, \tag{A.7}$$

and we can again use Lemma A.4 to conclude. □

The following is the easy inequality in a generic verification theorem for $\overline{F}$.

Lemma A.2. *Let w be a C^2 super–solution of $w - \delta y w' + F^\star(\delta w') \ge 0$ on $\mathbb{R}_+$. Then $w \ge \overline{F}$.*

Proof. We first observe that $w - F \ge \mathbf{T}F(y, \delta w') \ge 0$. We next compute for all $p \in \mathcal{B}_{\mathbb{R}_+}$ and $T \le T_0^{y_0,p}$ that

$$\begin{aligned} w(y_0) &= \mathrm{e}^{-\rho T} w\big(y^{y_0,p}(T)\big) + \int_0^T \rho\mathrm{e}^{-\rho t}\Big(w(y^{y_0,p}(t)) - \delta\big(y^{y_0,p}(t) - p(t)w'(y^{y_0,p}(t))\big)\Big)\mathrm{d}t \\ &\ge \mathrm{e}^{-\rho T} F\big(y^{y_0,p}(T)\big) + \int_0^T \rho\mathrm{e}^{-\rho t}F\big(p(t)\big)\mathrm{d}t, \end{aligned}$$

by the super–solution property of w. The arbitrariness of $p \in \mathcal{B}_{\mathbb{R}_+}$ and $T \le T_0^{y_0,p}$ implies that $w \ge \overline{F}$. □

Lemma A.3. *Let $\delta \ne 1$ and $\delta\gamma > 1$. Assume further that* Assumption 3.5 *holds. Then $\overline{F} = \big(\overline{F}^\star\big)^\star$, where $\overline{F}^\star$ is given explicitly in (2.9), is the unique solution of* Equation (A.1) *in the class of functions satisfying $\bar{c}_0(-1 + y^\gamma) \le \overline{F}(y) \le \bar{c}_0(-1 + y^\gamma)$. Moreover, $\overline{F}$ is a strictly concave decreasing majorant of F, with $\overline{F}'(0) = \delta^{-1}F'(0)\mathbf{1}_{\{\delta>1\}}$.*

Proof. We show by a standard verification argument that $w = \overline{F}$ where w is the solution of (A.1) whose concave dual was derived explicitly in the Lemma A.1. By Lemma A.2, w is an upper bound for $\overline{F}$, i.e. $\overline{F} \le w$.

On the other hand, consider for any $y > 0$ the maximiser in the definition of $F^\star(\delta w'(y))$, that is to say $(F^\star)'\big(\delta w'(y)\big)$ as a feedback control

$$\dot{y}_t^\star = r\big(y_t^\star - p_t^\star\big),\ \text{where } p_t^\star := (F^\star)'\big(\delta w'(y_t^\star)\big).$$

Direct differentiation of (A.1) provides that for any $y > 0$, $(1-\delta)w'(y) = \big(y - (F^\star)'(\delta w'(y))\big)\delta w''(y)$, so that

$$\dot{y}_t^\star = r\Big(\frac{1}{\delta} - 1\Big)y_t^\star h(y_t^\star),\ t \ge 0,\ \text{where } h(y) := \frac{w'(y)}{yw''(y)},\ y > 0.$$

Since $h \ge 0$, we see that $y^\star$ is decreasing when $\delta > 1$, and is therefore well–defined at least until the hitting time of zero $T^\star := T_0^{y_0,p^\star} < \infty$. In contrast, when $\delta < 1$, $y^\star$ is increasing until some explosion time $\bar{T}$, and $T^\star = \infty$.

Following the same calculation as in the first step of the present proof, we see that under the control $p^\star$, all inequalities are turned into equalities, leading for any $T \in [0, \bar{T}]$ to

$$w(y_0) = \mathrm{e}^{-\rho T\wedge T^\star} w(y_{T\wedge T^\star}^\star) + \int_0^{T\wedge T^\star} \rho\mathrm{e}^{-\rho t}F(p_t^\star)\mathrm{d}t. \tag{A.8}$$

First, by the previous step, when $\delta > 1$, we have $T^\star < \infty$, and we obtain by sending T to ∞ and using the boundary condition $w(0) = 0$ that $(T^\star, p^\star)$ attains the upper bound $w(y_0)$, and is therefore an optimal control for the problem $\overline{F}$.

In the alternative case $\delta < 1$, we have $T^\star = \infty$. In the rest of this proof, we show that

$$\bar{h} := \sup_{y \ge \hat{y}}\{h(y)\} < \frac{1}{\gamma(1-\delta)},\ \text{for some } \hat{y} > 0. \tag{A.9}$$

Then, $y^\star$ is defined on $\mathbb{R}_+$, i.e. $\bar{T}=\infty$, and since $\hat{T}:=\inf\{t\geq 0: y^\star_t\geq \hat{y}\}<\infty$, we deduce from the growth of w that for some $C>0$, whose value may change from line to line, and any $t\geq\hat{T}$

$$\mathrm{e}^{-\rho(t-\hat{T})}|w(y_t^\star)|\leq C\mathrm{e}^{-\rho(t-\hat{T})}\big(1+|y_t^\star|^\gamma\big)\leq C\mathrm{e}^{-\rho(t-\hat{T})}\Big(1+\mathrm{e}^{\rho\gamma(1-\delta)\int_{\hat{T}}^t h(y_s^\star)\mathrm{d}s}\Big)$$
$$\leq C\mathrm{e}^{-\rho(t-\hat{T})}\Big(1+\mathrm{e}^{\rho\gamma(1-\delta)\bar{h}(t-\hat{T})}\Big)\underset{t\to\infty}{\longrightarrow}0,$$

as $1-\gamma(1-\delta)\bar{h}>0$. Sending T to ∞ in (A.8) this provides again that $(T^\star,p^\star)=(\infty,p^\star)$ attains the upper bound $w(y_0)$.

In order to verify (A.9), we prove equivalently that the concave dual $w^\star$ satisfies

$$\sup_{p\leq\hat{p}}\left\{\frac{p(w^\star)''(p)}{(w^\star)'(p)}\right\}<\frac{1}{\gamma(1-\delta)},\ \text{for some } \hat{p}<0. \tag{A.10}$$

Differentiating the ODE (A.4) satisfied by $w^\star$, and using the expression of $w^\star$ from Equation (A.7) in the present case, we see that

$$\frac{p(w^\star)''}{(w^\star)'}=\frac{\delta}{1-\delta}-\Psi(\delta p),\ \text{with } \Psi(p):=\frac{-p\psi'(p)}{\psi(p)},\ \text{and } \psi(p):=(-p)^{-\frac{1}{1-\delta}}F^\star(p)-\int_{-\infty}^p\frac{F^\star(u)}{(1-\delta)(-u)^{1+\frac{1}{1-\delta}}}\mathrm{d}u.$$

Notice that $\lim_{p\to-\infty}\psi(p)=0$, and that ψ is easily shown to be non–negative and non–decreasing. Therefore, if $-p\psi'(p)$ does not go to 0 as p goes to $-\infty$, we have that $\lim_{p\to-\infty}\Psi(\delta p)=\infty$, and Equation (A.10) automatically holds. Now if $-p\psi'(p)\underset{p\to-\infty}{\longrightarrow}0$, it follows from l'Hôpital's rule and our assumptions that

$$\lim_{y\to\infty}\left\{\frac{F'(y)}{yF''(y)}\right\}=\lim_{p\to-\infty}\left\{\frac{p(F^\star)''(p)}{(F^\star)'(p)}\right\}=\frac{\delta}{1-\delta}-\lim_{p\to-\infty}\left\{\frac{-\psi'(p)-p\psi''}{\psi'(p)}\right\}.$$

Then, using again l'Hôpital's rule, we deduce

$$\limsup_{p\to-\infty}\left\{\frac{p(w^\star)''(p)}{(w^\star)'(p)}\right\}=\frac{\delta}{1-\delta}-\liminf_{p\to-\infty}\left\{\frac{-p\psi'(p)}{\psi(p)}\right\}$$
$$=\frac{\delta}{1-\delta}-\liminf_{p\to-\infty}\left\{\frac{-\psi'(p)-p\psi''}{\psi'(p)}\right\}=\lim_{y\to\infty}\left\{\frac{F'(y)}{yF''(y)}\right\}.$$

Then, assuming to the contrary that (A.10) does not hold means that, for fixed $\gamma_0\in(1+\gamma(1-\delta),\gamma)$, we may find $y_0>0$ such that $\frac{F''(y)}{F'(y)}\leq(\gamma_0-1)\frac{1}{y}$, for $y\geq y_0$. Integrating twice and recalling that $F\leq 0$, this implies that

$$F(y)\geq F(y_0)+\frac{y_0F'(y_0)}{\gamma_0}\left(\left(\frac{y}{y_0}\right)^{\gamma_0}-1\right),\ \text{for all } y\geq y_0,$$

which in turn leads to the following contradiction $\frac{y_0F'(y_0)}{\gamma_0}\leq\lim\frac{F(y)}{y^{\gamma_0}}=-\infty$ by our assumption on the growth of F together with the fact that $\gamma_0<\gamma$. □

We end this section with the result used in the proof of Lemma A.1.

Lemma A.4. *Let $\delta\neq 1$, and let $\overline{F}^\star$ be a solution of*

$$-\overline{F}^\star+(1-\delta)p\big(\overline{F}^\star\big)'+F^\star(\delta p)=0,\ p<\frac{F'(0)}{\delta}\mathbf{1}_{\{\delta>1\}},\ \overline{F}^\star\left(\frac{F'(0)}{\delta}\mathbf{1}_{\{\delta>1\}}\right)=0. \tag{A.11}$$

Then $\overline{F}^\star\leq F^\star$, $\overline{F}^\star$ is strictly concave and increasing.

Proof. Denote $\phi:=F^\star-\overline{F}^\star$, and notice that Equation (A.11) says that for any $p<F'(0)/\delta\mathbf{1}_{\{\delta>1\}}=:f_\delta$

$$\phi(p)=F^\star(p)-F^\star(\delta p)+(\delta-1)p\big(\overline{F}^\star\big)'(p)\geq(1-\delta)p\phi'(p),$$

by the concavity of $F^\star$. In other words, if we define for $p<f_\delta$, $\psi(p):=(-p)^{\frac{1}{\delta-1}}$, we have

$$\psi'(p)=\frac{(-p)^{\frac{1}{\delta-1}}}{1-\delta}\big(\phi(p)-(1-\delta)p\phi'(p)\big).$$

We need to distinguish two cases, depending on whether $\delta > 1$ or $\delta < 1$. First, if $\delta > 1$, we have that ψ is non–increasing, and thus for any $p < f_\delta$

$$(-p)^{\frac{1}{\delta-1}}\phi(p) \geq \lim_{p\to f_\delta}\Big\{(-p)^{\frac{1}{\delta-1}}\phi(p)\Big\} = 0,$$

since $\overline{F}^\star(f_\delta) = F^\star(f_\delta) = 0$ (recall that $F^\star$ is 0 above $F'(0)$, which is itself below f_δ, since $\delta > 1$). Similarly, when $\delta < 1$, by arguying as in (A.6), we arrive at the conclusion $\phi \geq 0$, and thus that $\overline{F}^\star \leq F^\star$, as desired.

Next, by direct differentiation of Equation (A.11), then substituting the expression of $(\overline{F}^\star)'$ from Equation (A.11), and finally using the strict concavity of $F^\star$, we deduce that for any $p < f_\delta$

$$\begin{aligned}(\delta-1)^2p^2(\overline{F}^\star)''(p) = \delta(\delta-1)p\big((F^\star)'(\delta p) - (\overline{F}^\star)'(p)\big) &= \delta(\delta-1)p(F^\star)'(\delta p) - \delta\big(F^\star(\delta p) - \overline{F}^\star(p)\big)\\ &< \delta\big(\overline{F}^\star(p) - F^\star(p)\big) \leq 0,\end{aligned}$$

thus proving the strict concavity of $\overline{F}^\star$.

Finally, since $\overline{F}^\star$ is strictly concave, remains below $F^\star$ which is increasing on $(-\infty, F'(0)]$, and 0 on $[F'(0), f_\delta \wedge F'(0)]$, then $\overline{F}^\star$ must also be increasing on its domain. ☐

B Ingredients for Perron's existence method

This section provides two main technical results which were needed to justify the existence of a solution of the dynamic programming equation in Lemma 9.3. We first prove the existence of a super–solution for the dynamic programming equation with appropriate growth. Then, we show that, despite the exploding feature of $F^\star$, the dynamic programming equation satisfies a comparison result.

Lemma B.1. (Super–solution) *Let $g : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ be a C^2, increasing, strictly concave function, with at most logarithmic growth at infinity, such that $g(0) = 0$ and the following possibly infinite limit exists*

$$c_o := \lim_{y\to\infty}\frac{-yg''(y)}{g'(y)} \;\geq\; (1-\delta^{-1})^+.$$

For $b > 0$, let $\overline{F}_b(y) := \overline{F}(y) + bg(y)$, $y \geq 0$. Then, for b sufficiently large, we may find a super–solution $\bar{v}$ of (7.1) with growth at infinity controlled by $\overline{F}_b$.

Proof. We proceed in five steps.

Step 1. As $\mathfrak{I}_0$ is Lipschitz in (p, q), it follows from the standard Cauchy–Lipschitz theorem that we may consider the maximal solution v_b on $[0, \bar{y})$, for some $\bar{y}$, of

$$v_b - \delta y v_b' + F^\star\big(\delta v_b'\big) - \mathfrak{I}_0\big(\delta v_b', \delta v_b''\big) = 0,\ v_b(0) = 0,\ v_b'(0) = b,$$

by writing this equation in its explicit form (3.5). Indeed, this is immediate when $\beta > 0$, and when $\beta = 0$, we can argue as in the proof of Lemma 9.1.

Next, as long as v_b is non–decreasing, we necessarily have $v_b \geq F$ (recall that $F(0) = 0$ and F is decreasing), and therefore $\mathfrak{I}_0(v_b', v_b'') = v_b - \delta y v_b' + F^\star\big(\delta v_b'\big) \geq F - \delta y v_b' + F^\star\big(\delta v_b'\big) \geq 0$. Consequently $\mathfrak{I}_0(\delta v_b', \delta v_b'') = \mathfrak{I}_0(\delta v_b', \delta v_b'')^+$ as long as v_b is non–decreasing, and v_b is a solution of the required equation (7.1) on this region.

Step 2. We first consider the case where v_b remains increasing on $[0, \bar{y})$. Then, v_b solves the ODE on $[0, \infty)$, i.e. $\bar{y} = \infty$, and we claim that

$$0 \leq v_b \leq \bar{a},\ \text{ and } \bar{v}_b := v_b + \overline{F} \text{ is a super–solution of (7.1).}$$

The last statement follows from the fact that

$$\begin{aligned}\mathbf{L}\bar{v}_b(y) &= v_b(y) + \overline{F}(y) - \delta y\big(v_b'(y) + \overline{F}'(y)\big) + F^\star\big(\delta v_b'(y) + \delta\overline{F}'(y)\big) - \mathfrak{I}_0\big(\delta v_b'(y) + \delta\overline{F}'(y), \delta v_b''(y) + \delta\overline{F}''(y)\big)^+ \\ &\geq v_b(y) + \overline{F}(y) - \delta y\big(v_b'(y) + F'(y)\big) + F^\star\big(\delta v_b'(y) + \delta\overline{F}'(y)\big) - \mathfrak{I}_0\big(\delta v_b'(y), \delta v_b''(y)\big)^+ \\ &\quad - \sup_{z\in\mathbb{R},\ \hat{a}\in\hat{A}(z)} \big\{h(\hat{a})\delta\overline{F}'(y) + \eta z^2\delta\overline{F}''(y)\big\} \\ &= F^\star\big(\delta v_b'(y) + \delta\overline{F}'(y)\big) - F^\star\big(\delta\overline{F}'(y)\big) - \sup_{z\in\mathbb{R},\ \hat{a}\in\hat{A}(z)} \big\{h(\hat{a})\delta\overline{F}'(y) + \eta z^2\delta\overline{F}''(y)\big\} \\ &\geq - \sup_{z\in\mathbb{R},\ \hat{a}\in\hat{A}(z)} \big\{h(\hat{a})\delta\overline{F}'(y) + \eta z^2\delta\overline{F}''(y)\big\} \geq 0,\end{aligned}$$

by the non–decrease of $F^\star$, together with direct manipulation of the supremum and the negativity of $\overline{F}'$ and $\overline{F}''$. To verify the claim that $v_b \leq \bar{a}$, we consider two separate cases.

(i) *Case* $\delta \leq 1$: by the concavity of v_b and the increase of $\mathfrak{I}_0$ in q, we have

$$0 = v_b(y) - \delta y v_b'(y) - \mathfrak{I}_0\big(\delta v_b'(y), \delta v_b''(y)\big) \geq v_b(y) - \delta y v_b'(y) - \bar{a} - h(\bar{a})\delta v_b'(y).$$

Assume to the contrary that $v_b(y_0) > \bar{a}$, for some $y_0 > 0$. Then, $v_b > \bar{a}$ on $[y_0, y_1)$ for some $y_1 > y_0$, and it follows from the last inequality that

$$\frac{v_b'(y)}{v_b(y) - \bar{a}} \geq \frac{1}{\delta y + h(\bar{a})}, \text{ and therefore } v_b(y) - \bar{a} \geq (v_b(y_0) - \bar{a})\left(\frac{\delta y + h(\bar{a})}{\delta y_0 + h(\bar{a})}\right)^{\frac{1}{\delta}},\ y \in [y_0, y_1).$$

This shows that we may take $y_1 = \infty$. Moreover, this completes the proof in the case $\delta < 1$ as it contradicts the concavity of v_b. In the remaining case $\delta = 1$, this shows that v_b is affine to the right of y_0, and as this affine function solves the ODE on $\mathbb{R}_+$, we deduce that $v_b(y) = b_0 + by$ is affine on $\mathbb{R}_+$. But this cannot happen as the equation imposes that the constant $b_0 = \mathfrak{I}_0(b, 0) = \bar{a} + bh(\bar{a})$, while the boundary condition imposes that $v_b(0) = b_0 = 0$.

(ii) *Case* $\delta > 1$: as v_b is increasing, we have $0 = v_b - \delta y v_b' - \mathfrak{I}_0(\delta v_b', \delta v_b'')^+ \leq v_b - \delta y v_b'$, which implies that $v_b(y) \leq y^{\frac{1}{\delta}}$. Indeed, if we define $w_b(y) := y^{1/\delta} - v_b(y)$, then the previous inequality implies directly that $\big(w_b(y)y^{-1/\delta}\big)' \geq 0$. Since for $\delta > 1$, we have $\lim_{y\to 0+} w_b(y)y^{-1/\delta} = 1 > 0$, we obtain the desired result.

Next, let ψ be a non–negative continuous function defined on a neighbourhood of the origin with $\psi(0) = 0$. We shall specify this function later, and we denote $\psi_\varepsilon := \psi(\varepsilon)$ for all small $\varepsilon > 0$. Assuming to the contrary that $2\eta := v_b(y_o) - \bar{a} > 0$, for some $y_o \in (0, \bar{y})$, we may find for all $\varepsilon > 0$ a maximiser $y_\varepsilon > 0$ of

$$M_\varepsilon := \max_{y\geq 0} \big\{v_b(y) - \bar{a} - \varepsilon y^{\frac{1}{\delta}+\psi_\varepsilon}\big\} = v_b(y_\varepsilon) - \bar{a} - \varepsilon y_\varepsilon^{\frac{1}{\delta}+\psi_\varepsilon}.$$

Besides, for ε small enough, we have

$$M_\varepsilon \geq v_b(y_o) - \bar{a} - \varepsilon y_o^{\frac{1}{\delta}+\psi_\varepsilon} \geq \eta > 0,$$

showing that y_ε must be an interior maximiser for ε small enough. In particular, for such small values of ε, we have $v_b'(y_\varepsilon) = \varepsilon\big(\frac{1}{\delta} + \psi_\varepsilon\big)y_\varepsilon^{\frac{1}{\delta}+\psi_\varepsilon-1}$, as well as $v_b''(y_\varepsilon) \leq 0$, and it follows from the definition of v_b that

$$\begin{aligned}0 = v_b(y_\varepsilon) - \delta y_\varepsilon v_b'(y_\varepsilon) - \mathfrak{I}_0\big(\delta v_b'(y_\varepsilon), \delta v_b''(y_\varepsilon)\big)^+ &\geq v_b(y_\varepsilon) - \delta y_\varepsilon v_b'(y_\varepsilon) - \mathfrak{I}_0\big(\delta v_b'(y_\varepsilon), 0\big)^+ \\ &= v_b(y_\varepsilon) - \delta y_\varepsilon v'(y_\varepsilon) - \bar{a} - h(\bar{a})v'(y_\varepsilon) \\ &\geq \eta - \delta\varepsilon\psi_\varepsilon y_\varepsilon^{\frac{1}{\delta}+\psi_\varepsilon} - h(\bar{a})\delta\left(\frac{1}{\delta} + \psi_\varepsilon\right)\varepsilon y_\varepsilon^{\frac{1}{\delta}+\psi_\varepsilon-1}. \qquad \text{(B.1)}\end{aligned}$$

On the other hand, we have $\eta \leq v_b(y_\varepsilon) - \bar{a} - \varepsilon y_\varepsilon^{\frac{1}{\delta}+\psi_\varepsilon} \leq y_\varepsilon^{\frac{1}{\delta}} - \bar{a} - \varepsilon y_\varepsilon^{\frac{1}{\delta}+\psi_\varepsilon}$, which implies that as ε goes to 0

- either $(y_\varepsilon)_{\varepsilon>0}$ remains bounded (and is in any case bounded away from 0, since v_b is non–decreasing) and, by sending $\varepsilon \searrow 0$, (B.1) implies that $0 \geq \eta$, contradicting the positivity of η;

• or $y_\varepsilon \longrightarrow_{\varepsilon\to 0} \infty$ along some sub–sequence, with $\varepsilon y_\varepsilon^{\psi_\varepsilon} < 1$. We claim that we may choose the function ψ so that for ε small enough

$$\psi_\varepsilon = \varepsilon^{1+\frac{1}{\delta\psi_\varepsilon}}. \tag{B.2}$$

In this case, $\varepsilon\psi_\varepsilon y_\varepsilon^{\frac{1}{\delta}+\psi_\varepsilon} = \varepsilon^2 y_\varepsilon^{\psi_\varepsilon} \frac{\psi_\varepsilon y_\varepsilon^{\frac{1}{\delta}}}{\varepsilon} \leq \varepsilon^2\psi_\varepsilon$, and (B.1) provides again a contradiction by sending $\varepsilon \searrow 0$.

We finally justify the existence of ψ_ε satisfying (B.2) by verifying directly that for ε small enough, the function $f_\varepsilon(\psi) := \psi\varepsilon^{-1-\frac{1}{\delta\psi}}$ is decreasing on $[\varepsilon,\infty)$, with $f(\varepsilon) > 1$ and fro any $\psi > 0$, $f(\varepsilon^{2+\frac{1}{\delta\psi}}) < 1$. in particular, for any $\psi > 0$, $\varepsilon^{2+\frac{1}{\delta\psi}} < \psi_\varepsilon < \varepsilon$ implying that $\psi_\varepsilon \longrightarrow 0$ as $\varepsilon \searrow 0$.

Step 3: otherwise, if v_b is ultimately decreasing, then we may find a point of maximum $\hat{y}_b > 0$, and we shall justify in **Step** 4 below that

$$v_b'(\hat{y}_b) = 0, \text{ and } \hat{y}_b \nearrow \infty, \text{ as } b \nearrow \infty. \tag{B.3}$$

Let $\hat{b} := -\overline{F}'(\hat{y}_b)$, and set

$$\bar{v}_b(y) := v_b(y)\mathbf{1}_{[0,\hat{y}_b]}(y) + \big(v_b(\hat{y}_b) + \overline{F}_{\hat{b}}(y) - \overline{F}_{\hat{b}}(\hat{y}_b)\big)\mathbf{1}_{(\hat{y}_b,\infty)}(y), \ y \geq 0.$$

By definition $\mathbf{L}\bar{v}_b = 0$ on $[0,\hat{y}_b]$. On $[\hat{y}_b,\infty)$, we compute directly that

$$\mathbf{L}\bar{v}_b(y) = v_b(\hat{y}_b) - \overline{F}(\hat{y}_b) + \mathbf{L}\overline{F}_{\hat{b}}(y),$$

Since $v_b(\hat{y}_b) > 0 > \overline{F}(\hat{y}_b)$, this implies that

$$\mathbf{L}\bar{v}_b(y) > \mathbf{L}\overline{F}_{\hat{b}}(y) \geq 0, \text{ for large } \hat{b},$$

by Step 5 below, together with (B.3) which implies that $\hat{b} = -F'(\hat{y}_b) \nearrow \infty$ as $b \nearrow \infty$.

Step 4: we next justify (B.3). By Remark 9.2, the concave dual $v_b^\star$ of v_b solves the ODE (9.1), which reduces on $[0,b]$ (that is to say the domain where $y \leq \hat{y}_b$) to

$$-(v_b^\star)''(p) = \delta\eta \inf_{z\geq\beta,\ \hat{a}\in\hat{A}(z)} \left\{\frac{z^2}{v^\star(p) + \hat{a} + \delta h(\hat{a})p + (\delta-1)p(v_b^\star)'(p)}\right\}, \text{ on } [0,b],\ v_b^\star(b) = 0,\ (v_b^\star)'(0) = \hat{y}_b,$$

by the identity $(v_b^\star)' = (v_b')^{-1}$, and the fact that on the domain where v_b is increasing, we can replace $\mathfrak{I}_0^+$ by $\mathfrak{I}_0$ in the ODE.

Assume first that $\delta \leq 1$. As $0 = v(0) = -\sup_{p\in\mathbb{R}} v^\star(p)$, and since $v_b^\star$ is concave and increasing on $[0,b]$, we deduce from the previous PDE that

$$-(v_b^\star)'' \geq \eta\delta\frac{\beta^2}{\bar{a} + \delta h(\bar{a})p},$$

which provides by direct integration between 0 and b that

$$\hat{y}_b = (v_b^\star)'(0) \geq -\big((v_b^\star)'(b) - (v_b^\star)'(0)\big) \geq \eta\delta\beta^2\big(\log\big(\bar{a} + \delta h(\bar{a})b\big) - \log(\bar{a})\big) \longrightarrow \infty, \text{ as } b \nearrow \infty.$$

Similarly, when $\delta > 1$, we have

$$-(v_b^\star)'' \geq \eta\delta\frac{\beta^2}{\bar{a} + (\delta h(\bar{a}) + (\delta-1)\hat{y}_b)p},$$

which provides by direct integration between 0 and b that

$$\hat{y}_b \geq \eta\delta\beta^2\big(\log\big(\bar{a} + (\delta h(\bar{a}) + (\delta-1)\hat{y}_b)b\big) - \log(\bar{a})\big).$$

Now, if $\hat{y}_b$ remained bounded as b goes to ∞, the above inequality would lead to a contradiction for b large. We can then let b go to ∞ and deduce again that $\hat{y}_b$ goes to ∞ as b goes to ∞.

Step 5: It remains to prove that $\overline{F}_b$ is a super–solution of (7.1) on $\{\overline{F}_b' \leq 0\}$ for sufficiently large $b > 0$. We first observe that

$$\{\overline{F}_b' \leq 0\} = [y_0(b), \infty), \text{ where } y_0(b) \text{ is the unique solution of } \overline{F}'\big(y_0(b)\big) + bg'\big(y_0(b)\big) = 0,$$

which exists as $\overline{F}_b$ is strictly concave and ultimately decreasing. Moreover, we get by direct differentiation, and using the fact that g is increasing, and both g and $\overline{F}$ are strictly concave

$$y_0'(b) = -\frac{g'(y_0(b))}{bg''(y_0(b)) + \overline{F}''(y_0(b))} > 0, \ b \geq 0.$$

Hence $b \longmapsto y_0(b)$ is increasing, and it is immediate from the equation defining $y_0(b)$ that it must go to ∞ as b goes to ∞. Next, by the non–decrease of $F^\star$ and $\mathfrak{I}_0(p,.)^+$, we directly compute that

$$\mathbf{L}\overline{F}_b \geq bf(y) - I(\delta\overline{F}_b', \delta\overline{F}'') \geq bf(y) - \bar{a}, \text{ where } f(y) := g(y) - \delta y g'(y), \ y \geq 0.$$

Direct calculations show that

$$f'(y) = g'(y)\bigg(1 - \delta - \delta\frac{yg''(y)}{g'(y)}\bigg), \ y \geq 0.$$

If $c_o = \infty$, then for y large enough f must be increasing. Otherwise, we have

$$f'(y) \underset{y\to\infty}{\sim} g'(y)\big(1 - \delta + \delta c_o\big),$$

and f is still increasing for large values of y. We can then choose b large enough so that f is increasing on $[y_0(b), \infty)$, and we can then deduce that $\mathbf{L}\overline{F}_b \geq bf\big(y_0(b)\big) - \bar{a}$ on $[y_0(b), \infty)$. We now complete the proof by showing that $f\big(y_0(b)\big) - \bar{a} \geq 0$ for sufficiently large b. However this is immediate since we have that f is increasing on $[y_0(b), \infty)$ for b large, that it converges to ∞ as f goes to ∞, and we have already proved that $y_0(b)$ increases to ∞ as b goes to ∞. □

Remark B.2. *In* Lemma B.1*, there are several possible choices for the function g. For instance, $g(y) = \log(1+y)$, or $g(y) = \log\big(1 + \log(1+y)\big)$, both verify the required properties with $c_o = 1$. This actually extends to arbitrary many iterations of the logarithm. In particular, the upper bound for v_b in* Lemma 9.3 *below can be improved, but the one we give is enough for our purpose here.*

We conclude this section by reporting the comparison result used in the proof of Lemma 9.3.

Lemma B.3. (Comparison) *Let u and v be respectively a viscosity sub–solution and a viscosity super–solution of* (7.1)*, such that for $\varphi \in \{u, v\}$ and for some $b > 0$*

$$\overline{F}(y) \leq \varphi(y) \leq \overline{F}(y) + b\log\big(1 + \log(1+y)\big), \ y \geq 0.$$

Then $u \leq v$ on $\mathbb{R}_+$.

Remark B.4. *The specific upper bound in the statement of* Lemma B.3 *with an iterated logarithm is not important* per se*. Indeed, the proof goes through as long as the upper bound is of the form $bg(y)$ for some positive, increasing, strictly concave map g, null at 0, growing strictly slower at ∞ than $\log(y)$. And we have already seen in* Remark B.2 *that we could find infinitely many such functions such that $\overline{F} + bg$ is a super–solution of* Equation (7.1) *for b large enough.*

Proof of Lemma B.3*.* Notice that u and v are respectively viscosity sub–solution and super–solution of the equation

$$w + G(y, w', w'') = 0, \text{ on } \mathbb{R}_+, \tag{B.4}$$

where the nonlinearity G is given, for any $(y, p, q) \in \mathbb{R}_+ \times \mathbb{R}^2$, by

$$G(y,p,q) := -\delta y p + F^\star(\delta p) - \mathfrak{I}_0\big(\delta p, \delta q\big)^+,$$

Our objective is to follow Crandall, Ishii, and Lions [13, Section 3], and adapt the arguments there to our context. Fix some $\nu > 0$ and define $\mu := 2 \vee (\gamma + \nu)$. We consider for any $\alpha > 0$ and $\varepsilon > 0$ the map

$$\psi_{\alpha,\varepsilon}(x,y) := u(x) - v(y) - \frac{\alpha}{\mu}|x-y|^\mu - \varepsilon\log(1+y), \ (x,y) \in \mathbb{R}_+^2.$$

We let for any $\alpha > 0$ and $\varepsilon > 0$

$$M_{\alpha,\varepsilon} := \sup_{(x,y)\in\mathbb{R}_+^2} \psi_{\alpha,\varepsilon}(x,y).$$

By the growth assumptions on u and v, the supremum in the definition of $M_{\alpha,\varepsilon}$ is attained, and we can define an $\mathbb{R}_+^2$–valued sequence $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})_{\alpha>0}$ such that for any $\alpha > 0$ and $\varepsilon > 0$

$$M_{\alpha,\varepsilon} = \psi_{\alpha,\varepsilon}(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}).$$

Since the supremum is attained on a compact set, we can find for any $\varepsilon > 0$ a further subsequence, denoted by $(x_n^\varepsilon, y_n^\varepsilon)_{n\in\mathbb{N}} := (x_{\alpha_n,\varepsilon}, y_{\alpha_n,\varepsilon})_{n\in\mathbb{N}}$, converging to some $(\hat{x}^\varepsilon, \hat{y}^\varepsilon)$. Moreover, by standard arguments from viscosity solution theory (see for instance Crandall, Ishii, and Lions [13, Proposition 3.7]), we have

$$\hat{x}^\varepsilon = \hat{y}^\varepsilon,\ \lim_{n\to\infty} \alpha_n \big|x_n^\varepsilon - y_n^\varepsilon\big|^\mu = 0,\ M_\varepsilon := \lim_{n\to\infty} M_{\alpha_n,\varepsilon} = \sup_{y\geq 0}\, (u-v)(y) - \varepsilon \log\big(1+\hat{y}^\varepsilon\big).$$

Let us now assume that there is some $y_o > 0$ such that $\eta := (u-v)(y_o) > 0$. Then, we have for any $n \in \mathbb{N}$ and $\varepsilon > 0$

$$\eta - \varepsilon \log(1+y_o) \leq M_{\alpha_n,\varepsilon} = \psi_{\alpha,\varepsilon}(x_{\alpha_n}^\varepsilon, y_{\alpha_n}^\varepsilon).$$

In particular, for n sufficiently large, we have that x_n^ε and y_n^ε are both positive, and we assume for notational simplicity that we took the appropriate subsequence. Using Crandall–Ishii's lemma (see Crandall, Ishii, and Lions [13, Theorem 3.2]), we can find for each integer n, an $\mathbb{R}_+^2$–valued sequence $(X_n^\varepsilon, Y_n^\varepsilon)_{n\in\mathbb{N}}$ such that

$$\big(\alpha_n(x_n^\varepsilon - y_n^\varepsilon)^{\mu-1}, X_n^\varepsilon\big) \in \overline{J}^{2,+}u(x_n^\varepsilon),\ \bigg(\alpha_n(x_n^\varepsilon - y_n^\varepsilon)^{\mu-1} - \frac{\varepsilon}{1+y_n^\varepsilon}, Y_n^\varepsilon\bigg) \in \overline{J}^{2,-}v(y_n^\varepsilon),$$

with the notation $x^\phi := \mathrm{sgn}(x)|x|^\phi$ for all $\phi > 0$ and $x \in \mathbb{R}$, and

$$-\bigg(\frac{1}{\lambda} + \|C_n^\varepsilon\|\bigg) \leq \begin{pmatrix} X_n^\varepsilon & 0 \\ 0 & -Y_n^\varepsilon \end{pmatrix} \leq C_n^\varepsilon\Big(I_2 + \lambda C_n^\varepsilon\Big),\quad \text{for all}\quad \lambda > 0,$$

where I_2 is the two–dimensional identity matrix, and

$$C_n^\varepsilon := a_n^\varepsilon A + b_n^\varepsilon B,\ a_n^\varepsilon := \alpha_n(\mu-1)|x_n^\varepsilon - y_n^\varepsilon|^{\mu-2},\ b_n^\varepsilon := \frac{\varepsilon}{(1+y_n^\varepsilon)^2},\ A := \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},\ B := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

and where we use the spectral norm for symmetric matrices. Take $\lambda = \|C_n^\varepsilon\|^{-1}$, we get

$$-2\|C_n^\varepsilon\| I_2 \leq \begin{pmatrix} X_n^\varepsilon & 0 \\ 0 & -Y_n^\varepsilon \end{pmatrix} \leq C_n^\varepsilon\bigg(I_2 + \frac{C_n^\varepsilon}{\|C_n^\varepsilon\|}\bigg).$$

This implies in particular (simply multiply the above inequality by $(1,1)$ to the left and $(1,1)^\top$ to the right) that for any $n \in \mathbb{N}$

$$X_n^\varepsilon - Y_n^\varepsilon \leq \frac{(b_n^\varepsilon)^2}{\|C_n^\varepsilon\|} + b_n^\varepsilon.$$

By the sub–solution and super–solution properties of u and v, we have for any $n \in \mathbb{N}$

$$u(x_n^\varepsilon) + G\big(x_n^\varepsilon, \alpha_n(x_n^\varepsilon - y_n^\varepsilon)^{\mu-1}, X_n^\varepsilon\big) \leq 0 \leq v(y_n^\varepsilon) + G\bigg(y_n^\varepsilon, \alpha_n(x_n^\varepsilon - y_n^\varepsilon)^{\mu-1} - \frac{\varepsilon}{1+y_n^\varepsilon}, Y_n^\varepsilon\bigg).$$

We deduce that $\eta - \varepsilon\log(1+y_o) \leq u(x_n^\varepsilon) - v(y_n^\varepsilon) \leq F_{n,\varepsilon}$, where

$$F_{n,\varepsilon} := G\bigg(y_n^\varepsilon, \alpha_n(x_n^\varepsilon - y_n^\varepsilon)^{\mu-1} - \frac{\varepsilon}{1+y_n^\varepsilon}, Y_n^\varepsilon\bigg) - G\big(x_n^\varepsilon, \alpha_n(x_n^\varepsilon - y_n^\varepsilon)^{\mu-1}, X_n^\varepsilon\big).$$

Now notice that since $\mathfrak{I}_0^+$ is Lipschitz continuous and non–decreasing with respect to its second variable, and since $F^\star$ is non–decreasing, we have for some $c_o > 0$

$$\begin{aligned}
F_{n,\varepsilon} &= \delta\alpha_n|x_n^\varepsilon - y_n^\varepsilon|^\mu + \delta\varepsilon\frac{y_n^\varepsilon}{1+y_n^\varepsilon} + F^\star\bigg(\delta\alpha_n(x_n^\varepsilon - y_n^\varepsilon)^{\mu-1} - \frac{\delta\varepsilon}{1+y_n^\varepsilon}\bigg) - F^\star\big(\delta\alpha_n(x_n^\varepsilon - y_n^\varepsilon)^{\mu-1}\big)\\
&\quad + \mathfrak{I}_0\big(\delta\alpha_n(x_n^\varepsilon - y_n^\varepsilon)^{\mu-1}, X_n^\varepsilon\big)^+ - \mathfrak{I}_0\bigg(\delta\alpha_n(x_n^\varepsilon - y_n^\varepsilon)^{\mu-1} - \frac{\delta\varepsilon}{1+y_n^\varepsilon}, Y_n^\varepsilon\bigg)^+\\
&\leq \delta\alpha_n|x_n^\varepsilon - y_n^\varepsilon|^\mu + \delta\varepsilon + \mathfrak{I}_0\bigg(\delta\alpha_n(x_n^\varepsilon - y_n^\varepsilon)^{\mu-1}, Y_n^\varepsilon + \frac{(b_n^\varepsilon)^2}{\|C_n^\varepsilon\|} + b_n^\varepsilon\bigg)^+\\
&\quad - \mathfrak{I}_0\bigg(\delta\alpha_n(x_n^\varepsilon - y_n^\varepsilon)^{\mu-1} - \frac{\delta\varepsilon}{1+y_n^\varepsilon}, Y_n^\varepsilon\bigg)^+\\
&\leq \delta\alpha_n|x_n^\varepsilon - y_n^\varepsilon|^\mu + \delta\varepsilon + c_o\bigg(\frac{(b_n^\varepsilon)^2}{\|C_n^\varepsilon\|} + b_n^\varepsilon\bigg) + c_o\frac{\delta\varepsilon}{1+y_n^\varepsilon}\\
&\leq \delta\alpha_n|x_n^\varepsilon - y_n^\varepsilon|^\mu + (1+c_o)\delta\varepsilon + c_o b_n^\varepsilon + c_o\frac{(b_n^\varepsilon)^2}{\|C_n^\varepsilon\|}.
\end{aligned}$$

We now want to let n go to ∞, and will distinguish two cases. First, if $\big(\|C_n^\varepsilon\|\big)_{n\in\mathbb{N}}$ is unbounded, taking a subsequence if necessary, we deduce by letting n go to ∞ that

$$\eta - \varepsilon\log(1+y_o) \leq (1+c_o)\delta\varepsilon + c_o\frac{\varepsilon}{(1+\hat{y}^\varepsilon)^2},$$

which gives a contradiction when ε goes to 0.

If now $\big(\|C_n^\varepsilon\|\big)_{n\in\mathbb{N}}$ remains bounded, we take a converging subsequence, and notice that we then have for some $a^\varepsilon \in \mathbb{R}$

$$\|C_n^\varepsilon\| \underset{n\to\infty}{\longrightarrow} \|C^\varepsilon\| := \bigg\|a^\varepsilon A + \underbrace{\frac{\varepsilon}{(1+\hat{y}^\varepsilon)^2}}_{=:b^\varepsilon}B\bigg\|.$$

If the sequence $(a^\varepsilon)_{\varepsilon>0}$ is unbounded, we take again a subsequence and get a contradiction by letting ε go to 0 in

$$\eta - \varepsilon\log(1+y_o) \leq (1+c_o)\delta\varepsilon + c_o b^\varepsilon + c_o\frac{(b^\varepsilon)^2}{\|C^\varepsilon\|}. \tag{B.5}$$

If now $(a^\varepsilon)_{\varepsilon>0}$ remains bounded and is such that $a := \text{limsup}_{\varepsilon\to0}a^\varepsilon \neq 0$, or $a := \text{liminf}_{\varepsilon\to0}a^\varepsilon \neq 0$, then taking another subsequence, we obtain a contradiction by letting ε go to 0 in Equation (B.5). Finally, if $\lim_{\varepsilon\to0} a^\varepsilon = 0$, then three cases can occur

(i) first, if $a^\varepsilon \underset{\varepsilon\to0}{=} o\big(b^\varepsilon\big)$, then $\frac{(b^\varepsilon)^2}{\|C^\varepsilon\|} \underset{\varepsilon\to0}{\sim} b^\varepsilon \underset{\varepsilon\to0}{\longrightarrow} 0$, and we conclude again by letting ε go to 0 in Equation (B.5);

(ii) if instead $b^\varepsilon \underset{\varepsilon\to0}{=} o\big(a^\varepsilon\big)$, then $\frac{(b^\varepsilon)^2}{\|C^\varepsilon\|} \underset{\varepsilon\to0}{\sim} \frac{b^\varepsilon}{a^\varepsilon}b^\varepsilon \underset{\varepsilon\to0}{\longrightarrow} 0$, and we conclude similarly;

(iii) finally, if $a^\varepsilon \underset{\varepsilon\to0}{\sim} cb^\varepsilon$ for some $c \neq 0$, then $\frac{(b^\varepsilon)^2}{\|C^\varepsilon\|} \underset{\varepsilon\to0}{\sim} \frac{b^\varepsilon}{\|cA+B\|} \underset{\varepsilon\to0}{\longrightarrow} 0$, and we get once more a contradiction.

□

References

[1] T. Adrian and M.M. Westerfield. Disagreement and learning in a dynamic contracting model. *Review of Financial Studies*, 22(10):3873–3906, 2009.

[2] R. Aïd, D. Possamaï, and N. Touzi. Optimal electricity demand response contracting with responsiveness incentives. *arXiv preprint arXiv:1810.09063*, 2018.

[3] O. Alvarez, J.-M. Lasry, and P.-L. Lions. Convex viscosity solutions and state constraints. *Journal de Mathématiques Pures et Appliquées*, 76(3):265–288, 1997.

[4] B. Biais, T. Mariotti, G. Plantin, and J.-C. Rochet. Dynamic security design: convergence to continuous time and asset pricing implications. *The Review of Economic Studies*, 74(2):345–390, 2007.

[5] B. Biais, T. Mariotti, J.-C. Rochet, and S. Villeneuve. Large risks, limited liability, and dynamic moral hazard. *Econometrica*, 78(1):73–118, 2010.

[6] B. Biais, T. Mariotti, and J.-C. Rochet. Dynamic financial contracting. In D. Acemoglu, M. Arellano, and E. Dekel, editors, *Advances in economics and econometrics, 10th world congress of the Econometric Society, volume 1, economic theory*, number 49 in Econometric Society Monographs, pages 125–171. Cambridge University Press, 2013.

[7] B. Bouchard and N. Touzi. Explicit solution to the multivariate super–replication problem under transaction costs. *The Annals of Applied Probability*, 10(3):685–708, 2000.

[8] M. Broadie, J. Cvitanić, and H.M. Soner. Optimal replication of contingent claims under portfolio constraints. *The Review of Financial Studies*, 11(1):59–79, 1998.

[9] A. Capponi and C. Frei. Dynamic contracting: accidents lead to nonlinear contracts. *SIAM Journal on Financial Mathematics*, 6(1):959–983, 2015.

[10] J.-F. Chassagneux, R. Élie, and I. Kharroubi. When terminal facelift enforces Delta constraints. *Finance and Stochastics*, 19(2):329–362, 2015.

[11] P. Cheridito, H.M. Soner, and N. Touzi. The multi–dimensional super–replication problem under gamma constraints. *Annales de l'institut Henri Poincaré, analyse non linéaire* (C), 22(5):633–666, 2005.

[12] S.M. Choi. *On Sannikov's continuous–time principal–agent problem*. PhD thesis, University of California Berkeley, 2014.

[13] M.G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, 27(1):1–67, 1992.

[14] J. Cvitanić and J. Zhang. *Contract theory in continuous–time models*. Springer, 2012.

[15] J. Cvitanić, X. Wan, and J. Zhang. Optimal contracts in continuous–time models. *International Journal of Stochastic Analysis*, 2006(095203), 2006.

[16] J. Cvitanić, X. Wan, and J. Zhang. Principal–agent problems with exit options. *The B.E. Journal of Theoretical Economics*, 8(1):23, 2008.

[17] J. Cvitanić, X. Wan, and J. Zhang. Optimal compensation with hidden action and lump–sum payment in a continuous–time model. *Applied Mathematics and Optimization*, 59(1):99–146, 2009.

[18] J. Cvitanić, D. Possamaï, and N. Touzi. Moral hazard in dynamic risk management. *Management Science*, 63(10):3328–3346, 2017.

[19] J. Cvitanić, D. Possamaï, and N. Touzi. Dynamic programming approach to principal–agent problems. *Finance and Stochastics*, 22(1):1–37, 2018.

[20] J.-P. Décamps and S. Villeneuve. A two–dimensional control problem arising from dynamic contracting theory. *Finance and Stochastics*, 23(1):1–28, 2019.

[21] P.M. DeMarzo and Y. Sannikov. Optimal security design and dynamic capital structure in a continuous–time agency model. *The Journal of Finance*, 61(6):2681–2724, 2006.

[22] P.M. DeMarzo, M.J. Fishman, Z. He, and N. Wang. Dynamic agency and the q theory of investment. *The Journal of Finance*, 67(6):2295–2340, 2012.

[23] R. Élie and D. Possamaï. Contracting theory with competitive interacting agents. *SIAM Journal on Control and Optimization*, 57(2):1157–1188, 2019.

[24] R. Élie, E. Hubert, T. Mastrolia, and D. Possamaï. Mean–field moral hazard for optimal energy demand response management. *arXiv preprint arXiv:1902.10405*, 2019.

[25] R. Élie, T. Mastrolia, and D. Possamaï. A tale of a principal and many many agents. *Mathematics of Operations Research*, 44(2):440–467, 2019.

[26] K. Fong. Evaluating skilled experts: optimal scoring rules for surgeons. Stanford university, 2009.

[27] P. Guasoni, M. Rásonyi, and W. Schachermayer. Consistent price systems and face–lifting pricing under transaction costs. *The Annals of Applied Probability*, 18(2):491–520, 2008.

[28] I. Hajjej, C. Hillairet, M. Mnif, and M. Pontier. Optimal contract with moral hazard for public private partnerships. *Stochastics: An International Journal of Probability and Stochastic Processes*, 89(6–7): 1015–1038, 2017.

[29] I. Hajjej, C. Hillairet, and M. Mnif. Optimal stopping contract for public private partnerships under moral hazard. *arXiv preprint arXiv:1910.05538*, 2019.

[30] Z. He. Optimal executive compensation when firm size follows geometric brownian motion. *Review of Financial Studies*, 22(2):859–892, 2009.

[31] M.F. Hellwig. The role of boundary solutions in principal–agent problems of the Holmström–Milgrom type. *Journal of Economic Theory*, 136(1):446–475, 2007.

[32] M.F. Hellwig and K.M. Schmidt. Discrete–time approximations of the Holmström–Milgrom Brownian–motion model of intertemporal incentive provision. *Econometrica*, 70(6):2225–2264, 2002.

[33] F. Hoffmann and S. Pfeil. Reward for luck in a dynamic agency model. *The Review of Financial Studies*, 23(9):3329–3345, 2010.

[34] B. Holmström and P. Milgrom. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica*, 55(2):303–328, 1987.

[35] N. Ju and X. Wan. Optimal compensation and pay–performance sensitivity in a continuous–time principal–agent model. *Management Science*, 58(3):641–657, 2012.

[36] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate texts in mathematics*. Springer–Verlag New York, 2nd edition, 1998.

[37] K.L. Keiber. Overconfidence in the continuous–time principal–agent problem. Technical report, WHU Otto Beisheim Graduate School of Management, 2003.

[38] K. Larsen, H.M. Soner, and G. Žitković. Facelifting in utility maximization. *Finance and Stochastics*, 20(1):99–121, 2016.

[39] R.C.W. Leung. Continuous–time principal–agent problem with drift and stochastic volatility control: with applications to delegated portfolio management. Technical report, Haas School of Business, University of California Berkeley, 2014.

[40] Y. Lin, Z. Ren, N. Touzi, and J. Yang. Second order backward SDE with random terminal time. *arXiv preprint arXiv:1802.02260*, 2018.

[41] Y. Lin, Z. Ren, N. Touzi, and J. Yang. Random horizon principal–agent problem. *arXiv preprint arXiv:2002.10982*, 2020.

[42] J.A. Mirrlees and R.C. Raimondo. Strategies in the principal–agent model. *Economic Theory*, 53(3): 605–656, 2013.

[43] H.M. Müller. The first–best sharing rule in the continuous–time principal–agent problem with exponential utility. *Journal of Economic Theory*, 79(2):276–280, 1998.

[44] H.M. Müller. Asymptotic efficiency in dynamic principal–agent problems. *Journal of Economic Theory*, 91(2):292–301, 2000.

[45] R.B. Myerson. Leadership, trust, and power: dynamic moral hazard in high office. University of Chicago, 2008.

[46] H. Ou-Yang. Optimal contracts in a continuous–time delegated portfolio management problem. *Review of Financial Studies*, 16(1):173–208, 2003.

[47] H. Pagès. Bank monitoring incentives and optimal ABS. *Journal of Financial Intermediation*, 22(1): 30–54, 2013.

[48] H. Pagès and D. Possamaï. A mathematical treatment of bank monitoring incentives. *Finance and Stochastics*, 18(1):39–73, 2014.

[49] A. Papapantoleon, D. Possamaï, and A. Saplaouras. Existence and uniqueness for BSDEs with jumps: the whole nine yards. *Electronic Journal of Probability*, 23(121):1–68, 2018.

[50] T. Piskorski and A. Tchistyi. Optimal mortgage design. *Review of Financial Studies*, 23(8):3098–3140, 2010.

[51] T. Piskorski and M.M. Westerfield. Optimal dynamic contracts with moral hazard and costly monitoring. *Journal of Economic Theory*, 166:242–281, 2016.

[52] W.P. Rogerson. The first–order approach to principal–agent problems. *Econometrica*, 53(6):1357–1368, 1985.

[53] Y. Sannikov. Agency problems, screening and increasing credit lines. Princeton university, 2007.

[54] Y. Sannikov. A continuous–time version of the principal–agent problem. *The Review of Economic Studies*, 75(3):957–984, 2008.

[55] Y. Sannikov. Contracts: the theory of dynamic principal–agent relationships and the continuous–time approach. In D. Acemoglu, M. Arellano, and E. Dekel, editors, *Advances in economics and econometrics, 10th world congress of the Econometric Society, volume 1, economic theory*, number 49 in Econometric Society Monographs, pages 89–124. Cambridge University Press, 2013.

[56] H. Schättler and J. Sung. The first–order approach to the continuous–time principal–agent problem with exponential utility. *Journal of Economic Theory*, 61(2):331–371, 1993.

[57] U. Schmock, S.E. Shreve, and U. Wystup. Valuation of exotic options under shortselling constraints. *Finance and Stochastics*, 6(2):143–172, 2002.

[58] M.D. Schroder, S. Sinha, and S. Levental. The continuous–time principal–agent problem with moral hazard and recursive preferences. Technical report, Michigan State University, 2010.

[59] H.M. Soner and N. Touzi. Superreplication under gamma constraints. *SIAM Journal on Control and Optimization*, 39(1):73–96, 2000.

[60] H.M. Soner and N. Touzi. Dynamic programming for stochastic target problems and geometric flows. *Journal of the European Mathematical Society*, 4(3):201–236, 2002.

[61] H.M. Soner and N. Touzi. The problem of super–replication under constraints. In P. Bank, E. Baudoin, H. Föllmer, L.C.G. Rogers, H.M. Soner, and N. Touzi, editors, *Paris–Princeton lectures on mathematical finance 2002*, volume 1814 of *Lecture notes in mathematics*, pages 133–172. Springer, 2003.

[62] H.M. Soner and N. Touzi. Hedging under gamma constraints by optimal stopping and face–lifting. *Mathematical Finance*, 17(1):59–79, 2007.

[63] S.E. Spear and S. Srivastava. On repeated moral hazard with discounting. *The Review of Economic Studies*, 54(4):599–617, 1987.

[64] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der mathematischen Wissenschaften*. Springer–Verlag Berlin Heidelberg, 1997.

[65] B. Strulovici and M. Szydlowski. On the smoothness of value functions and the existence of optimal strategies in diffusion models. *Journal of Economic Theory*, 159:1016–1055, 2015.

[66] J. Sung. Linearity with project selection and controllable diffusion rate in continuous–time principal–agent problems. *The RAND Journal of Economics*, 26(4):720–743, 1995.

[67] J. Sung. Corporate insurance and managerial incentives. *Journal of Economic Theory*, 74(2):297–332, 1997.

[68] N. Van Long and G. Sorger. A dynamic principal–agent problem as a feedback Stackelberg differential game. *Central European Journal of Operations Research*, 18(4):491–509, 2010.

[69] M.M. Westerfield. Optimal dynamic contracts with hidden actions in continuous time. University of Southern California, 2006.

[70] N. Williams. On dynamic principal–agent problems in continuous time. University of Wisconsin, Madison, 2008.

[71] N. Williams. Persistent private information. *Econometrica*, 79(4):1233–1275, 2011.

[72] N. Williams. A solvable continuous time dynamic principal–agent model. *Journal of Economic Theory*, 159(part B):989–1015, 2015.

[73] Y. Zhang. Dynamic contracting with persistent shocks. *Journal of Economic Theory*, 144(2):635–675, 2009.

[74] Y. Zhou. Principal–agent analysis in continuous–time. Technical report, The Chinese University of Hong Kong, 2006.

[75] J.Y. Zhu. *Sticky incentives and dynamic agency*. PhD thesis, University of California Berkeley, 2011.