Relational Contracts: Public versus Private Savings*

Francesc Dilmé[†] **Daniel Garrett**[‡]

October 7, 2019

PRELIMINARY VERSION

Abstract

We study relational contracting with an agent who has consumption-smoothing preferences as well as the ability to save to defer consumption (or to borrow). We compare principal-optimal relational contracts in two settings. The first where the agent's consumption and savings decisions are private, and the second where these decisions are publicly observed. In the first case, the agent smooths his consumption over time, the agent's effort and payments eventually decrease over time, and the balances on his savings account eventually increase. In essence, the relationship eventually deteriorates with time. In the second case, the relational contract can specify the level of consumption by the agent in each period. The optimal contract calls for the agent to consume earlier than he would like, consumption and balances on the account fall over time, and effort and payments to the agent increase. Our results suggest a possible explanation for low savings rates in certain industries where incentive pay plays an important role.

JEL Classification: D82

Keywords: relational contracts, consumption smoothing preferences, private savings

*We are grateful for helpful comments from Nageeb Ali, Daniel Bird, Mathias Fahn, William Fuchs, Drew Fudenberg, Chris Shannon, Andy Skrzypacz, Rani Spiegler, Philipp Strack, Balazs Szentes, Jean Tirole, Marta Troya Martinez, Mark Voorneveld, and seminar participants at Barcelona Graduate School of Economics, Berkeley, Boston University, Hebrew University of Jerusalem, Higher School of Economics, Stockholm School of Economics, Tel Aviv University, the University of Bonn, University Pompeu Fabra, the Stony Brook International Conference on Game Theory, the 2nd Japanese-German Workshop on Contracts and Incentives, and the 5th Workshop on Relational Contracts in Madrid. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 714147).

[†]University of Bonn.

[‡]Toulouse School of Economics.

1 Introduction

Early literature on repeated moral hazard with a risk-averse agent — for instance Rogerson (1985) and Fudenberg, Holmstrom and Milgrom (1990) — highlighted the value to the principal of controlling the agent's consumption or savings. Optimal dynamic contracts in the setting of Rogerson force the agent to consume more than he would like early in the relationship (the agent would gain by secretly saving and deferring some consumption to a later date). In the framework of Fudenberg, Holmstrom and Milgrom, a sequence of short-term contracts can often implement the outcome of a long-term contract if consumption/savings decisions can be stipulated by contract.[1] Yet, in modern employment relationships, workers' consumption expenditures remain at their discretion and are rarely subject to formal agreement.

Nonetheless, employers in some settings *are* able to monitor, at least to a degree, the consumption and savings decisions of workers. "Conspicuous consumption" decisions include choices of clothes, car, or leisure activities.[2] Some savings decisions (such as pension contributions, student loan repayments or repayments of employer-offered mortgages) are observed as direct deductions from employee paychecks. Consumption expenditures are sometimes deducted from pay (for instance, executive compensation whose monetary component might, in principle, depend on the extent of discretionary perquisites, with the two acting as substitute rewards for good performance (see Bennardo, Chiappori and Song, 2010)).

In this paper, we ask how the evolution of employment relationships can be expected to depend on the observability of consumption and savings decisions when formal agreements about these decisions are ruled out, but where *relational* incentives might nonetheless exist. The idea that many productive relationships rest on goodwill and implicit agreements, together with the threat of future punishments for deviations, has been developed following work such as Bull (1987), MacLeod and Malcomson (1989) and Levin (2003). While such agreements often concern both the level of output delivered to the principal, as well as agent pay (or bonuses), we introduce the possibility that they also concern the level of consumption, to the extent this is jointly observed. Implicit understandings on appropriate consumption levels could reflect part of a workplace's culture, sustained through repeated interaction with the employer. For instance, one interpretation is that an employee with an insufficiently frivolous lifestyle (in terms of car, dress, leisure activities and perquisites) is dismissed, having been deemed a "poor fit" with the firm's culture.

We consider a simple relational contracting model in which it is possible to study two

[1] Other examples where the principal controls the level of agent consumption include Sannikov (2008) and Garrett and Pavan (2015).

[2] The term "conspicuous consumption" was introduced by Veblen (1899).

polar opposite cases: one where consumption is perfectly jointly observed, and the other where consumption is unobservable to the principal. Although reality might often lie between these two extremes, studying each separately is illuminative and simplifies the analysis.

Our simple setting is deterministic, with the output enjoyed by the principal equal to the agent's effort. A contract can then be understood as an agreement on streams of effort, (bonus) pay and agent consumption. The initial balance on the agent's savings account is common knowledge, and the evolution of these savings are determined by pay and consumption. The agent has concave utility over money and hence a preference for smooth consumption.

When consumption is observable, our problem is to determine the sequence of effort, pay and consumption that maximizes the principal's payoff among all sequences that can arise as outcomes in subgame perfect equilibrium. When consumption is unobservable, the relevant concept is instead perfect-Bayesian equilibrium. In either case, there is no loss in considering equilibria where public deviations are punished by "autarky". This means not only that effort and payments to the agent stop, but also that the agent optimally smooths his available savings over the infinite future. Hence, the agent's payoff in autarky depends on the level of the balance on his savings account.

As in a number of other relational contracting settings, the conditions for equilibria correspond to two key sets of constraints. One set of constraints ensures willingness of the agent never to quit the relationship by deviating from the agreed effort and consumption, if observed. The other constraints ensure credibility of the principal's payments to the agent. That is, they ensure willingness of the principal to pay the agreed compensation, which (given the absence of formal contracts) is entirely at the principal's discretion.

We begin by studying the problem with unobservable consumption, which we view as an important benchmark. In this case, any relational contract must involve constant consumption, since the agent finds it optimal to perfectly smooth his lifetime income and faces no incentives to do otherwise. Similarly, any public deviation by the agent — i.e., shirking on effort which effectively ends the relationship — must optimally be accompanied by perfectly smoothed consumption. This means that, if the agent deviates in effort at a given date, he optimally reduces his consumption from the initial date relative to what he would consume in equilibrium. We argue that there is no loss in considering payments to the agent that keep him indifferent to such optimal deviations. Since lifetime earnings accumulate the longer the agent obediently exerts effort, and since agent utility is concave, the pay needed to keep the agent from deviating from a given level of effort increases with time. Provided such payments are credible, the contract can then implement the first-best outcome; effort is constant and pay increases over time.

For lower levels of patience, the first-best outcome cannot be sustained in equilibrium. Recall that, relative to a given level of effort, pay to the agent must increase with time; in a sense, the agent becomes more costly to compensate to keep in the relationship. This means that the contract becomes less profitable with time, making it more difficult to sustain agreement at later dates. In particular, we show that the payments that the principal can credibly promise eventually decline, and so effort decreases as well.

The dynamics of the relationship are remarkably different when consumption is instead observed. Note that the possibility for the informal agreement to condition on consumption enlarges the set of sustainable outcomes for two reasons. One is that non-constant consumption paths become possible, due to implicit incentives on the level of consumption. The other is that, if the agent deviates from the specified effort at a given date, he cannot lower his consumption at earlier dates without this being detected, triggering premature termination of his pay. This effectively reduces the profitability for the agent of deviations from equilibrium effort (compared to what he could obtain were he able to reduce consumption at earlier dates and continue being paid). That deviations in effort are less profitable permits the agent to be paid earlier in the relationship (i.e., it permits pay to be more front-loaded). In relation to the first-best policy, this flattens the profile of pay which helps to relax the principal's credibility constraints. First-best outcomes are thus sustainable for a broader range of model parameters.

Now consider the case where the first-best outcome is not sustainable. In this case, the optimal contract stipulates high consumption for the agent early in the relationship, which drives down the balance on his account. Given a lower balance, lower pay is needed to keep the agent obedient to the contract at a fixed level of effort. For this reason, the agent's compensation relative to effort declines. Since the relationship grows more profitable with time, the pay the principal can credibly offer increases; agent effort and the principal's profits therefore increase with time. The key to understanding these dynamics lies in recognizing the value of increasing the profitability of the relationship at later dates, which in turn relaxes the principal's credibility constraint and permits high pay early on. The channel for increasing future profits is to induce higher consumption by the agent than is privately optimal, reducing the balance on his account and making him more willing to work for low pay. In fact, we show that this process continues indefinitely, with the balance approaching a level at which a first-best contract can be sustained. Hence, in the long-run, the contract with the agent is approximately first best. While these dynamics might be thought of as a form of "immiseration" (the agent's payoff falls towards a lower bound), they are borne out of the principal's inability to commit, and are not related to immiseration results documented in papers such as Thomas and Worrall (1990) and Atkeson and Lucas (1992).[3]

As mentioned above, in real-world relationships, not all of an agent's consumption expenditures are likely to be observable to the principal. Nonetheless, our analysis of fully observable consumption suggests a force that could, at least potentially, be relevant in some settings. In particular, our results suggest a theory of high consumption and low savings that favors the profits of the principal. The idea could be relevant in certain industries such as banking, where high remuneration (especially through discretionary bonuses) is accompanied by a propensity for high spending. Such a propensity seems widely acknowledged: Former British banker Geraint Anderson has commented that bankers are essentially trapped in a culture of high consumption, adding that life as a banker is "like a gilded cage".[4] Sarah Butcher, at *efinancialcareers*, quotes Gary Goldstein, co-founder of executive search firm Whitney Partners, suggesting that high-earning bankers often have low savings:[5]

> "It's really not that unusual to find Wall Street bankers who are close to declaring themselves bankrupt.... Some people are really struggling."

Mark Gongloff, writing for *Huffpost Business*, explains that bankers "are under constant social pressure to spend and spend some more".[6] This kind of behavior would increase the willingness of bankers to continue devoting long hours, which seems advantagenous to firms and might even be encouraged by them. The idea that the behavior reflects cultural norms, that may be sustained through repeated interaction, offers an alternative to the possibility that high consumption is simply about signaling.

A further source of evidence on pressure to spend may come from supply arrangements, especially those managed by large firms. One possibility is that larger firms push for expenditures by suppliers that harm their financial positions, perhaps making them more financially dependent. A possible example comes from the US poultry industry, where chicken farmers supply to a few large firms that dominate the industry. Alison Moodie at *The Guardian* documented the situation of a farmer that contracted with the chicken producer Tyson Foods.[7]

[3] Williams (2011), however, shows that the immiseration result of Thomas and Worrall is sensitive to the process for income shocks.

[4] 'The perils of earning a £100,000 salary' by Jon Kelly, *BBC News Magazine*, 22 September 2010, https://www.bbc.com/news/magazine-11382591.

[5] 'When a million isn't enough: why top bankers are struggling to get by' by Sarah Butcher, *efinancialcareers,* 29 April 2013, https://news.efinancialcareers.com/uk-en/140070/when-a-million-isnt-enough-why-top-bankers-are-struggling-to-get-by/.

[6] 'Bankers Explain How They Cannot Possibly Live On $1 Million Pay' by Mark Gongloff, *Huffpost Business*, 1 May 2013, https://www.huffpost.com/entry/bankers-1-million-pay_n_3188177.

[7] 'Fowl play: the chicken farmers being bullied by big poultry', by Alison Moodie, *The Guardian*, 22 April 2017, https://www.theguardian.com/sustainable-business/2017/apr/22/chicken-farmers-big-poultry-rules.

Farmer Alton Terry entered an exclusive agreement with Tyson. After the first couple of years, Tyson began demanding additional expenditures on equipment, things like extra feed bins and chicken houses that Terry believed unnecessary. Terry commented: "If we are independent contractors, then why does the company have the right to tell us what equipment to use?" When Terry failed to comply, the relationship deteriorated and, in the end, it was terminated. Our theory seems consistent with this kind of story, although in our model the agent's expenditures are on consumption rather than improving productivity.

The rest of this paper is as follows. Next, in Section 1.1, we provide a summary of relevant literature. Section 2 introduces a model, Section 3 considers the first-best contract, Section 4 considers relational contracts when consumption is unobservable to the principal, and Section 5 considers relational contracts where consumption is observable. While the results in Sections 3 to 5 suppose the agent can save, Section 6 determines the optimal relational contract when the agent cannot save and so consumes what he is paid in every period. Section 7 concludes.

1.1 Other literature

This paper contributes to the literature on relational contracts, reviewed in MacLeod (2007) and Malcomson (2013). While much of this literature is interested in settings with moral hazard, the players most often have linear preferences over money and savings plays no role. Some examples include Levin (2003), Fuchs (2007), Chassang (2010), Halac (2012), Li and Matouschek (2013), Yang (2013), Malcomson (2016) and Fong and Li (2017). Also unlike our paper, much of the interest in these works lies in the role of exogenous uncertainty, which is often a source of private information for one of the parties.

Our model features concave agent utility over payments, giving the agent a preference for smooth consumption. We also allow the agent to save and dissave. The role of consumption-smoothing preferences and saving has been given little attention in the relational contracting literature to date. Pearce and Stacchetti (1998) consider a relational contracting model with a risk-averse agent, but there is no scope for the agent to save. Bull (1987) considers a model with overlapping generations of risk-averse agents and a long-lived principal where the agents can privately save. Yet his concern is only with the conditions under which the first-best outcome can be sustained.

Apart from models with productive effort, there are settings with lending and insurance. Thomas and Worrall (1990) consider a model where the principal insures a risk-averse agent. The paper devotes some attention to self-enforcing insurance arrangements (thus relaxing insurer commitment in the spirit of relational contracting); however, the agent is not permitted to save by himself. Thomas and Worrall (1994) study the problem of an investor in an agent (country) that can steal the invested capital and walk away from the relationship. The setting is relational (neither party can commit), and they consider separately cases where the capital of the agent can accumulate (arguably akin to saving), and where the agent is risk averse.

An important driver of dynamics in our model is evolution of the agent's outside option of ceasing productive effort and "living off" the balance on his account. The role of the value of the agent's outside option in repeated relationships has been emphasized in work such as Baker, Gibbons and Murphy (1994) and McAdams (2011). In Baker, Gibbons and Murphy, the outside option is endogenously determined by the possibility of contracting on objective performance measures, while in McAdams (2011) it is endogenously determined by the opportunities of partners to a relationship to rematch. In general, the higher the outside options of the parties to a relationship, the harder it is for a productive relationship to be sustained. In some papers, as in ours, the outside option evolves dynamically. One example is Thomas and Worrall's (1994) model of capital accumulation mentioned above, where the value to the agent of departing with the accumulated capital increases with time. Another is Garicano and Rayo (2017), where the agent is paid by increments in productive knowledge which increases the value of his outside option.[8]

Our paper is also related to work on dynamic moral hazard where the agent can privately save. Examples include Edmans et al. (2009), Abraham, Koehne and Pavoni (2011), He (2012), and Di Tella and Sannikov (2018). In this work, the principal has full commitment power. As a consequence, there is no loss in considering contracts that feature no private savings by the agent in equilibrium. A common result is that optimal contracts, instead of featuring front-loaded payment patterns as in the case where saving is observed (see Rogerson, 1985), feature more backloaded payments which discourage private savings. However, the restriction to contracts where the agent consumes what he is paid is only one possibility; given that the principal commits, there is a multiplicity of optimal contracts where payments are spread differently across time. This is a key difference to our work where the precise timing of payments is more restricted, and often uniquely pinned down. The reason is that the level of payments plays a critical role in determining whether the principal is willing to continue abiding by the contract (that is, the payments stipulated by the contract need to be credible).

[8] Also related is Fudenberg and Rayo (2017); the agent's outside option improves over time also due to the accumulation of knowledge, but the paper focuses on the case where the principal can commit to a long-term contract.

2 Setting

Environment and preferences. A principal and agent meet in discrete time $t = 1, 2, \ldots$. The common discount factor is $\delta \in (0,1)$. In every period t, first the agent exerts an effort $e_t \geq 0$ and consumes an amount $c_t \in \mathbb{R}$. Then, the principal makes a discretionary payment $w_t \geq 0$ to the agent.

The agent has initial savings balance $b_1 > 0$ as well as access to a savings technology, which accumulates interest at rate $\frac{1-\delta}{\delta}$. The agent's balance at time $t \geq 1$ then satisfies

$$b_{t+1} = \frac{b_t + w_t - c_t}{\delta} = b_1 \delta^{-t} + \sum_{s=1}^{t} \delta^{s-t-1} (w_s - c_s) . \tag{1}$$

Balances can, in principle, be negative (i.e., the agent can borrow), but we will impose the following standard feasibility constraint:

$$\lim_{t \to \infty} \delta^{t-1} b_t \geq 0. \tag{2}$$

The agent's felicity from consumption c_t in any period t is denoted $v(c_t)$, where $v : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$. We assume that $v(c)$ is real-valued for $c > 0$, and assumes value $-\infty$ otherwise. This will mean that the agent prefers any stream of consumption which remains strictly positive ahead of any stream such that consumption is non-positive in at least one period.[9] We further assume that v, when evaluated on positive consumption values, is twice continuously differentiable, strictly increasing and strictly concave. We also assume that it is onto all of $\mathbb{R}$, and satisfies $\lim_{c \searrow 0} v(c) = -\infty$.

In every period, the agent decides his effort level $e_t \in \mathbb{R}_+$. The agent's disutility of (non-negative) effort e_t is $\psi(e_t)$. We assume that ψ is continuously differentiable, strictly increasing, strictly convex, and such that $\psi(0) = \psi'(0) = 0$, and that $\lim_{e \to \infty} \psi'(e) = \infty$.

The agent's date-t payoff is then $v(c_t) - \psi(e_t)$, while the principal's is $e_t - w_t$ (hence, we interpret effort as equal to the output enjoyed by the principal).

Relational contracts. We focus on *deterministic* relational contracts. We understand this as simply reflecting the players' inability to randomize.[10] Somewhat abusively, we identify

[9] Consumption will be strictly positive in any self-enforcing agreement. However, we do not rule out that balances on the agent's account may become negative, in which case (in the event of a deviation) negative consumption may be needed to satisfy the constraint (2).

[10] More generally, in examining contracts that are optimal for the principal, whether random contracts can improve on deterministic ones might be expected to depend on the nature of risk aversion (e.g., whether v exhibits increasing or decreasing risk aversion). Our results below, however, will hold irrespective of how risk preferences change with the level of consumption.

relational contracts with their equilibrium outcomes, denote them $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$. We restrict these outcomes to satisfy the following feasibility constraints.

Definition 2.1. A *feasible relational contract* is a sequence $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$ satisfying the following feasibility conditions:

1. **Non-negativity:** $\tilde{e}_t, \tilde{w}_t \geq 0$ and $\tilde{c}_t > 0$ for all t.
2. **Balance dynamics and constraint:** Conditions (1) and (2) hold.
3. **Bounded sequences:** The sequences $(\tilde{e}_t)_{t\geq 1}$, $(\tilde{c}_t)_{t\geq 1}$ and $(\tilde{w}_t)_{t\geq 1}$ are bounded.

3 First best and principal full commitment

Consider the problem of maximizing the principal's payoff subject only to the constraint that the agent is initially willing to participate. More specifically, we look for profit-maximizing, feasible relational contracts $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$ such that the payoff of the agent

$$\sum_{t=1}^{\infty} \delta^{t-1} \left(v(\tilde{c}_t) - \psi\left(\tilde{e}_t\right)\right)$$

is no lower than his autarky value, $\frac{1}{1-\delta} v((1-\delta)b_1)$.

Proposition 3.1. *Consider maximizing the principal's discounted payoff by choice of feasible contracts $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$, subject to ensuring the agent a payoff at least his autarky value $\frac{1}{1-\delta} v((1-\delta)b_1)$. Optimal effort and consumption are constant at $e^{FB}(b_1)$ and $c^{FB}(b_1)$, respectively, being the unique solutions to:*

1. ***First order condition:*** *$\psi'(e^{FB}(b_1)) = v'(c^{FB}(b_1))$, and*
2. ***Agent's indifference condition:*** *$v(c^{FB}(b_1)) - \psi(e^{FB}(b_1)) = v((1-\delta)b_1)$.*

Furthermore, the payoff of the principal is $V^{FB}(b_1) \equiv \frac{1}{1-\delta}(e^{FB}(b_1) - (c^{FB}(b_1) - (1-\delta)b_1))$, which is a decreasing function of b_1.

The results in the proposition are easily anticipated. Given that v is concave, it is optimal to prescribe constant consumption. Similarly, the convexity of the disutility of effort implies the optimality of constant effort. At an optimum, the agent is indifferent between participating in the contract and autarky.

It is worth observing that the first-best policies depend on both b_1 and δ, although we reduce the notational burden by making dependence only on b_1 explicit. As the second condition in Proposition 3.1 indicates, the first-best policies are uniquely determined by $(1-\delta)b_1$, which is the agent's consumption in autarky.

The first-best problem corresponds to one in which both principal and agent can fully commit at date 1 to contractual terms over the infinite future. Such a contract can stipulate a constant payment $c^{FB}(b_1) - b_1(1-\delta)$ to the agent for delivering effort $e^{FB}(b_1)$ in each period (and zero payments if the agent ever deviates). Whether the agent's consumption is also agreed is immaterial, since, given the contractual payments, the agent optimally sets consumption equal to $c^{FB}(b_1)$ every period. We discuss below the implementation of the first best effort also when the agent cannot commit, and when neither the principal nor agent can commit.

4 Unobservable consumption

This section studies the case where, at each time t, the principal can observe the previous and current effort choices of the agent $(e_s)_{s=1}^t$, but not the consumption choices (or the agent's balance). In settings where the commitment of the players cannot be taken for granted, the relational contracting literature looks for contracts that are "self-enforcing". This means that each party to the agreement is willing to adhere to it in each period, given that continuation play depends on such adherence. In our setting with private consumption, the natural equilibrium concept is then *Perfect Bayesian Equilibrium* (PBE); our task, to find an equilibrium with outcomes $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$ that are optimal for the principal.

To define equilibrium, let us first define the admissible strategies. For $t \geq 0$, a t-history for the agent is $h_t^A = (e_s, w_s, c_s)_{1\leq s<t}$, which specifies the observed effort, payments and consumption up until time $t-1$. The set of such histories at date $t \geq 1$ is $\mathcal{H}_t^A = \mathbb{R}_+^{2(t-1)} \times \mathbb{R}^{t-1}$ (with the convention that $\mathbb{R}^0 = \emptyset$). Note that, given h_t^A and the agent's initial balance b_1, we can completely determine the evolution of the balance up to b_t at date t using (1). Somewhat abusively, we denote the date-t balance by $b\left(h_t^A\right)$. A t-history for the principal is $h_t^P = (e_s, w_s)_{1\leq s<t}$. The set of such histories at date $t \geq 1$ is $\mathcal{H}_t^P = \mathbb{R}_+^{2(t-1)}$.

A strategy for the agent is then a collection of functions

$$\alpha_t : \mathcal{H}_t^A \to \mathbb{R}_+ \times \mathbb{R},\ t \geq 1,$$

and a strategy for the principal is a collection of functions

$$\sigma_t : \mathcal{H}_t^P \times \mathbb{R}_+ \to \mathbb{R}_+,\ t \geq 1.$$

Here, α_t maps the t-history of outcomes (as are fully observed by the agent) to a pair (e_t, c_t) of effort and consumption. Also, σ_t maps the past history of jointly observable outcomes (efforts and payments) up to $t-1$, together with the observed effort choice e_t of the agent, to a payment w_t.

We restrict strategies for the agent by requiring that, for any strategy $(\alpha_t)_{t=1}^{\infty}$ employed by the agent, there is $l \in \mathbb{R}$ such that $b\left(h_t^A\right) - c_t$ is bounded above l, over all t and histories h_t^A. Here, for every t and h_t^A, c_t is determined using α_t. Note that the choice of l can depend on the strategy; we can view it as a choice of the agent.

A PBE then comprises agent and principal strategies as set out above, together with beliefs of the principal at every $\left(h_t^P, e_t\right) \in \mathcal{H}_t^P \times \mathbb{R}_+$, for all t. We require strategies to be sequentially optimal, subject to the above restriction for agent strategies (i.e., any deviation by the agent must respect the restriction on his strategy). Beliefs must be updated according to Bayes' rule wherever possible.

The rationale for the restriction on agent strategies can be understood as follows. Consider a date t when the agent's balance is b_t, and any $T > t$. If payments to the agent are $w_t, w_{t+1}, \ldots, w_{T-1}$ and the agent consumes $c_t, c_{t+1}, \ldots, c_{T-1}$, then the date-$T$ balance is

$$b_T = \delta^{t-T} b_t + \sum_{s=t}^{T-1} \delta^{s-T} \left(w_s - c_s\right).$$

Therefore, if the agent's date-T consumption is determined by an admissible strategy, we have

$$b_T - c_T = \delta^{t-T} b_t + \sum_{s=t}^{T-1} \delta^{s-T} \left(w_s - c_s\right) - c_T \geq l.$$

Since the inequality holds uniformly across T, and taking limits assuming that these exist,

$$\sum_{s=t}^{\infty} \delta^{s-t} c_s \leq b_t + \sum_{s=t}^{\infty} \delta^{s-t} w_s.$$

This shows that the NPV of consumption from t onwards must be no greater than the agent's date-t balance plus the NPV of future earnings. Our restriction on strategies thus disciplines the agent's consumption not only on path, but also in case of deviations.

We are then interested in which relational contracts $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$ are the outcomes of some PBE. Without loss of generality, we suppose publicly observed deviations from these contractual outcomes are punished by "autarky". This means that, if the agent deviates from the agreed effort $\tilde{e}_t$, or if the principal deviates from the agreed payment $\tilde{w}_t$, the principal makes no payments and the agent exerts no effort from then on; the agent perfectly smoothing the balance of his account over the infinite future. Deviations by the agent from the specified consumption, provided they are not accompanied by any deviation in effort, go unpunished (i.e., the principal continues to adhere to the payments specified by the agreement).

To characterize which contracts $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$ can be outcomes of a PBE, we argue as follows. Suppose that the agent deviates in his effort choice for the first time at date t. That is, suppose the agent exerts effort equal to $\tilde{e}_s$ for all $s < t$, and an effort different from $\tilde{e}_t$ at time t (while the principal chooses to continue with the agreement until observing a deviation). Given the first public deviation is at time t, the agent optimally sets consumption in every period equal to

$$\bar{c}_{t-1} \equiv (1-\delta)\left(b_1 + \sum_{s=1}^{t-1} \delta^{s-1}\tilde{w}_s\right) \tag{3}$$

so as to completely smooth (and exhaust) lifetime earnings. When the agent instead continues to choose effort obediently in every period, optimal consumption is $\bar{c}_\infty$, determined by taking $t = \infty$ in Equation (3). Clearly, any contract in which the agent behaves obediently must then specify $\tilde{c}_t = \bar{c}_\infty$ for all t.

Given the above, the maximum payoff the agent achieves when deviating in choice of effort at date t is

$$\frac{1}{1-\delta}v(\bar{c}_{t-1}) - \sum_{s=1}^{t-1}\delta^{s-1}\psi(\tilde{e}_s).$$

Hence, the agent does not want to deviate from the agreement if, for all $t \geq 1$,

$$\frac{1}{1-\delta}v(\bar{c}_{t-1}) - \sum_{s=1}^{t-1}\delta^{s-1}\psi(\tilde{e}_s) \leq \frac{1}{1-\delta}v(\bar{c}_\infty) - \sum_{s=1}^{\infty}\delta^{s-1}\psi(\tilde{e}_s). \tag{AC$_t^{\text{un}}$}$$

The principal remains willing to continue abiding by the agreement if, at each time t, the payment $\tilde{w}_t$ that is due is less than her continuation payoff in the agreement. The exact requirement is that, for all $t \geq 1$,

$$\tilde{w}_t \leq \sum_{s=t+1}^{\infty}\delta^{s-t}(\tilde{e}_s - \tilde{w}_s). \tag{PC$_t$}$$

We are thus able to show the following.

Proposition 4.1. *A feasible sequence $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$ is the outcome of a PBE in our environment if and only if, for all $t \geq 1$, Conditions (AC_t^{un}) and (PC_t) are satisfied, and $\tilde{c}_t = \bar{c}_\infty$.*

.

The result simplifies our task of finding a PBE that maximizes the payoff of the principal. Inspired by the terminology of the literature, we refer to a contract $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$ that satisfies the conditions of Proposition 4.1 as *self-enforceable*.[11] Our task then reduces to finding a feasible sequence $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$ that maximizes the principal's payoff subject to it being self-enforceable. We term such a contract *optimal*.

To determine the properties of optimal contracts, we first show that we can restrict attention to contracts with a particular pattern of payments over time. This pattern involves paying the agent as early as possible, subject to satisfying the agent's incentive constraints. This requires that the agent's obedience constraints in Condition ($\mathrm{AC}_t^{\mathrm{un}}$) hold with equality for all $t \geq 1$. Inspired by the terminology of Board (2011), we refer to this condition as "fastest payments". We show the following.

Lemma 4.1. *If there is an optimal relational contract, then there is another optimal contract* $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$ *with the same sequence of efforts such that, for all* $t \geq 1$,

$$\frac{v\left(\bar{c}_{t-1}\right)}{1-\delta} - \sum_{s=1}^{t-1} \delta^{s-1}\psi(\tilde{e}_s) = \frac{v\left(\left(1-\delta\right)b_1\right)}{1-\delta}. \qquad (\mathrm{FP}_t^{\mathrm{un}})$$

An explanation for the result is as follows. First, it is optimal to hold the agent to his outside option, and hence

$$\frac{v\left(\bar{c}_{\infty}\right)}{1-\delta} - \sum_{t=1}^{\infty} \delta^{t-1}\psi(\tilde{e}_t) = \frac{v\left(\left(1-\delta\right)b_1\right)}{1-\delta}. \qquad (4)$$

If Condition (4) does not hold, $\tilde{e}_1$ can be slightly increased (keeping the rest of the contract the same) so that the constraints ($\mathrm{AC}_t^{\mathrm{un}}$) and ($\mathrm{PC}_t$) continue to hold for all t. Second, when ($\mathrm{FP}_t^{\mathrm{un}}$) holds for all t, the agent is paid as early as possible while preserving the constraints ($\mathrm{AC}_t^{\mathrm{un}}$). The agent cannot be paid earlier, otherwise he will prefer to work obediently for a certain number of periods, save his income at a higher rate than specified in the agreement, and then quit by exerting no effort. It is easily seen that moving payments earlier in time only relaxes the "principal's constraints" (PC_t).

When Condition ($\mathrm{FP}_t^{\mathrm{un}}$) is satisfied for all dates, we observe that the payment the agent receives, relative to the disutility of effort incurred, increases over time. In particular, we observe that, for any t, $\tilde{e}_t > 0$ implies

$$\tilde{w}_t \in \left(\frac{\psi\left(\tilde{e}_t\right)}{v'\left(\bar{c}_{t-1}\right)}, \frac{\psi\left(\tilde{e}_t\right)}{v'\left(\bar{c}_t\right)}\right). \qquad (5)$$

[11] We think this terminology is appropriate since we identify a contract with equilibrium outcomes, rather than the complete specification of strategies (and beliefs). Outcomes being self-enforceable simply means that they can be supported in equilibrium.

Note here that $\bar{c}_t$, as defined by (3), increases with t. This observation turns out to be important for understanding the dynamics of optimal relational contracts, particularly because it implies that the ratio $\frac{\tilde{w}_t}{\psi(\tilde{e}_t)}$ increases with t. In other words, the payments needed to keep the agent obediently in the relationship, relative to the disutility of effort incurred, increase with time.

The usefulness of Lemma 4.1 is that it permits the design of the relational contract to be reduced to the choice of an effort sequence $(\tilde{e}_t)_{t\geq 1}$. From $(\tilde{e}_t)_{t\geq 1}$ we can obtain $(\bar{c}_t)_{t\geq 1}$ using (FP_t^{un}) (so the corresponding consumption $\tilde{c}_t = \bar{c}_\infty$ is also pinned down). Then $(\tilde{w}_t)_{t\geq 1}$ is obtained from Equation (3), and $(\tilde{b}_t)_{t\geq 1}$ from Equation (1). We next discuss the implementation of first-best contracts (Section 4.1), before moving to consider optimal contracts when there is no first-best contract that is self-enforceable (Section 4.2).

4.1 Implementation of the first-best contract

Lemma 4.1 is useful for understanding the conditions under which the first-best contract characterized in Proposition 3.1 can be implemented. For instance, we can observe that the first-best outcome, which involves effort and consumption $(e^{FB}(b_1), c^{FB}(b_1))$ in each period, can be implemented when the principal can commit to the agreement, but the agent cannot commit. For this, we simply suppose the principal agrees to payments satisfying the conditions in Equation (FP_t^{un}), provided the agent chooses effort obediently. Any deviation by the agent from the required effort is met with zero payments from then on.

Now consider whether the principal can attain the first-best payoff when neither the principal nor agent can commit; i.e., whether there is a first-best contract that is self-enforceable. According to Condition (5), payments to the agent increase over time. In the long run, payments approach

$$\frac{\psi\left(e^{FB}(b_1)\right)}{v'\left(c^{FB}(b_1)\right)}.$$

Therefore, verifying the principal's constraint (PC_t) is satisfied amounts to verifying that

$$\frac{\psi\left(e^{FB}(b_1)\right)}{v'\left(c^{FB}(b_1)\right)} \leq \frac{\delta}{1-\delta}\left(e^{FB}(b_1) - \frac{\psi\left(e^{FB}(b_1)\right)}{v'\left(c^{FB}(b_1)\right)}\right). \tag{6}$$

The right-hand side is the limiting value of the principal's future discounted profits in the agreement, while the left-hand side is the limiting value of the payment to the agent. Because there is no loss in restricting attention to "fastest payments" (due to Lemma 4.1), this condition is also necessary, and so we have the following result.

Proposition 4.2. *Suppose that neither the principal nor agent can commit to the terms of the agreement and that consumption is unobservable. Then the principal attains the first-best payoff in an optimal contract if and only if Condition* (6) *is satisfied.*

Note that an optimal contract for the principal coincides with the first best whenever Condition (6) is satisfied. While understanding the parameter range for which this condition holds is therefore important for understanding the optimal contract, this is complicated by the dependence of the first-best policy on both b_1 and δ. For instance, we were unable to establish in general monotonicity in δ. However, if we vary δ while allowing b_1 to adjust, holding $b_1(1-\delta)$ constant, then the first-best policies remain constant; there is then a threshold value of δ above which Condition (6) is satisfied, and below which it fails.

4.2 Main characterization for unobservable consumption

We now state our main result for the unobservable consumption case, which is a characterization of optimal effort when the first-best effort cannot be sustained.

Proposition 4.3. *An optimal contract exists. Suppose the principal cannot attain the first-best payoff in a self-enforceable contract. Then, in any optimal contract, effort is constant up to some date* $\bar{t} \geq 1$*, and is subsequently strictly decreasing, converging to a value* $\tilde{e}_\infty > 0$ *in the long run. There exist parameters for which* $\bar{t} > 1$ *in any optimal contract; in particular, effort can indeed be constant in the initial periods.*

The dynamics of the optimal effort policy can be explained as follows. There may be a sequence of consecutive initial periods in which the principal's constraint (PC_t) is slack. In this case, as we explain below, effort is constant over the initial periods. Considering then the "fastest payments" determined in Lemma 4.1, the payments $\tilde{w}_t$ rise over the initial periods, until the principal's constraint binds. The reason for the rising payments is that, as the agent obediently chooses the effort specified by the agreement, he can save his income over time. Given that the agent has a concave utility of consumption, rising payments are therefore needed to make deviating from the agreement unprofitable. Once the principal's constraint binds, payments must be reduced, which is only possible by reducing the level of effort. Note that how much effort can be asked without violating the principal's constraint depends on the future profitability of the relationship. Profitability declines over time, both because higher payments must be made relative to the agent's disutility of effort, and because the effort that can be requested is less. The fact that profitability declines contributes to the decline in effort, which creates a feedback loop.

In terms of our approach to proving Proposition 4.3, a couple of initial observations are worth making. First, unlike the case of observable consumption studied below, we do not appear to have a convenient recursive representation of the optimal contracting problem; at least not one with a single state variable. In the observable consumption case, what takes the place of the agent's constraint (AC_t^{un}) (namely, the constraint (AC_t^{ob})) at any date t determines feasible continuation contracts as a function of the date-t balance $\tilde{b}_t$. Put simply, the jointly observed balance is a convenient state variable when consumption is observed. However, in the present case with unobservable consumption, the constraint (AC_t^{un}) determines feasible continuation contracts as a function of both past pay, as summarized by $\bar{c}_{t-1}$, and the date t itself. Any recursive formulation of the problem then seems to require two state variables and we have not found such a formulation especially useful.

Second, our main mode of attack relies on variational arguments. For contracts that fail to exhibit the dynamics anticipated by Proposition 4.3, we are able to construct more profitable contracts that satisfy all the constraints in Proposition 4.1. We demonstrate some of these kinds of arguments in the text below, focusing on arguments that shed most light on the dynamics in the result.

We begin with the fact that optimal effort is constant over the initial periods when the principal's constraint (PC_t) is slack, which can be seen as a consequence of the following result.

Lemma 4.2. *Suppose that $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ is an optimal relational contract. Suppose that the principal's constraint is slack at t^*, i.e. $\tilde{w}_{t^*} < \sum_{s=t^*+1}^{\infty} \delta^{s-t^*} (\tilde{e}_s - \tilde{w}_s)$. Then, $\tilde{e}_{t^*-1}, \tilde{e}_{t^*+1} \leq \tilde{e}_{t^*}$.*

Proof. **Proof that $\tilde{e}_{t^*+1} \leq \tilde{e}_{t^*}$.** Suppose, for the sake of contradiction, that $\tilde{e}_{t^*+1} > \tilde{e}_{t^*}$. We can choose a new contract with efforts $(\tilde{e}'_t)_{t\geq 1}$, and payments $(\tilde{w}'_t)_{t\geq 1}$ chosen to satisfy Equation (FP_t^{un}), such that they coincide with the original policy except in periods t^* and t^*+1. In these periods, $\tilde{e}'_{t^*}$ and $\tilde{e}'_{t^*+1}$ are such that $\tilde{e}_{t^*} < \tilde{e}'_{t^*} \leq \tilde{e}'_{t^*+1} < \tilde{e}_{t^*+1}$ and

$$\psi\left(\tilde{e}'_{t^*}\right) + \delta\psi\left(\tilde{e}'_{t^*+1}\right) = \psi\left(\tilde{e}_{t^*}\right) + \delta\psi\left(\tilde{e}_{t^*+1}\right),$$

which implies (by convexity of ψ) that $\tilde{e}'_{t^*} + \delta\tilde{e}'_{t^*+1} > \tilde{e}_{t^*} + \delta\tilde{e}_{t^*+1}$. We then have also that $\tilde{w}_{t^*} < \tilde{w}'_{t^*}$ and $\tilde{w}'_{t^*} + \delta\tilde{w}'_{t^*+1} = \tilde{w}_{t^*} + \delta\tilde{w}_{t^*+1}$ (since the NPV of payments does not change, equilibrium consumption $\tilde{c}_t$ does not change in any period t; so the balance $\tilde{b}_{t^*+1}$ increases). Provided the changes are small, the principal's constraint (PC_t) at t^* remains satisfied. Since the payment at time t^* is higher in the new contract than under the original one (because $\tilde{e}'_{t^*} > \tilde{e}_{t^*}$), the payment $\tilde{w}_{t^*+1}$ at time t^*+1 is lower, so the principal's constraint is relaxed at date t^*+1. Since the NPV of output goes up, the principal's constraint is relaxed at all periods before t^*. The contract after date t^*+1 is unaffected. The modified contract is thus self-enforceable, and it is strictly more profitable than the original, establishing a contradiction.

Proof that $\tilde{e}_{t^*-1} \leq \tilde{e}_{t^*}$. Suppose now, for the sake of contradiction, that $\tilde{e}_{t^*-1} > \tilde{e}_{t^*}$. We can choose again a new contract with efforts $(\tilde{e}'_t)_{t\geq 1}$ and payments $(\tilde{w}'_t)_{t\geq 1}$ satisfying Equation (FP_t^{un}) that coincides with $(\tilde{e}_t)_{t\geq 1}$ except in periods t^*-1 and t^*, specifying $\tilde{e}'_{t^*-1}$ and $\tilde{e}'_{t^*}$ so that $\tilde{e}_{t^*} < \tilde{e}'_{t^*} \leq \tilde{e}'_{t^*-1} < \tilde{e}_{t^*-1}$ and

$$\psi\left(\tilde{e}'_{t^*-1}\right) + \delta\psi\left(\tilde{e}'_{t^*}\right) = \psi\left(\tilde{e}_{t^*-1}\right) + \delta\psi\left(\tilde{e}_{t^*}\right),$$

which implies that $\tilde{e}'_{t^*-1} + \delta\tilde{e}'_{t^*} > \tilde{e}_{t^*-1} + \delta\tilde{e}_{t^*}$. Note that

$$\tilde{w}'_{t^*-1} + \delta\tilde{w}'_{t^*} = \tilde{w}_{t^*-1} + \delta\tilde{w}_{t^*}.$$

Also, $\tilde{w}'_{t^*-1} < \tilde{w}_{t^*-1}$ and $\tilde{w}'_{t^*} > \tilde{w}_{t^*}$ Provided the changes are small, the principal's constraint (PC_t) at t^* remains satisfied. Moreover, the principal's constraints are relaxed at date t^*-1, and because the NPV of effort increases, also at all earlier dates. Therefore, the principal's constraints are satisfied at all dates and the principal's payoff strictly increases. Again, this establishes a contradiction. □

The implication of constant effort is then straightforward. Suppose that the principal's constraint (PC_t) is slack on adjacent dates t^* and t^*+1 say. Because the constraint is slack at t^*, $\tilde{e}_{t^*+1} \leq \tilde{e}_{t^*}$. Because the constraint is slack at t^*+1, $\tilde{e}_{t^*} \leq \tilde{e}_{t^*+1}$. Therefore, $\tilde{e}_{t^*} = \tilde{e}_{t^*+1}$. The explanation for constant effort over the initial periods is thus as follows. If the principal's constraints (PC_t) are slack in the initial periods, and effort is *not* constant over these periods, then effort can be smoothed over time, yielding a more profitable contract that still satisfies the constraints in Proposition 4.1. Such smoothing should be expected to be profitable given that the disutility of effort is strictly convex (i.e., differences in effort across periods are inefficient).

Given Lemma 4.2, Proposition 4.3 reveals that the principal's constraints (PC_t) hold with equality from some date onwards. One step in arriving at Proposition 4.3 is therefore establishing that it cannot be optimal for the principal's constraints to be slack at all dates, in which case Lemma 4.2 would imply that optimal effort is constant over all dates. Suppose then that effort is constant at some value, say $\tilde{e}_\infty$ (using the notation of the proposition). Let the payments and the equilibrium consumption $\bar{c}_\infty$ be determined by Equation (FP_t^{un}). Then payments increase over time, and converge to $\frac{\psi(\tilde{e}_\infty)}{v'(\bar{c}_\infty)}$. The principal's constraint (PC_t) is then satisfied if and only if

$$\frac{\psi\left(\tilde{e}_\infty\right)}{v'\left(\bar{c}_\infty\right)} \leq \frac{\delta}{1-\delta}\left(\tilde{e}_\infty - \frac{\psi\left(\tilde{e}_\infty\right)}{v'\left(\bar{c}_\infty\right)}\right),$$

where the left-hand side can be read as the limiting payment to the agent, while the right-hand side is the limiting NPV of future profits to the principal. For the most profitable choice of a constant effort $\tilde{e}^*_\infty$, this inequality holds as equality. The principal's constraints (PC$_t$) tighten over time, but never hold with equality.

Because effort is below the first-best level, we have $\psi'(\tilde{e}^*_\infty) < v'(\bar{c}_\infty)$, and so any (small) increase in effort, together with a change in payments that leaves the agent's payoff in the contract unchanged, raises profits. We therefore suggest a perturbation to the constant-effort contract (see Lemma A.7 in the Appendix) that increases the NPV of effort, but (assuming that payments continue to satisfy (FP$_t^{\text{un}}$)) leaves the principal's constraints (PC$_t$) intact. Concretely, we consider increasing effort by a little at date one and lowering it by a constant amount in future periods. If we only raise effort at date one, leaving other effort values unchanged and assuming that payments are adjusted to satisfy (FP$_t^{\text{un}}$), the principal's constraint (PC$_t$) is eventually violated (since v is strictly concave and total pay increases, it becomes more costly to compensate the agent for his effort; in particular, payments must increase in all periods). Therefore the reduction in effort at future dates is a "correction" intended to relax the principal's constraint (PC$_t$) when it is tightest. This part of the proof is illuminative, for it highlights the value in reducing effort at later dates when the principal's constraint is tightest and increasing effort early on when the principal's constraint is most slack.

We have thus established that the optimal policy is not constant and the principal's constraint (PC$_t$) holds as an equality in some period. We are also able to show (in Lemma A.5) that effort is weakly decreasing with time. Lemma A.6 then establishes that, if the principal's constraint (PC$_t$) holds with equality at some date $\hat{t}$, then $\tilde{e}_{\hat{t}+1} < \tilde{e}_{\hat{t}}$ and the constraint holds with equality also at $\hat{t}+1$. Hence effort strictly decreases over time.

The argument of Lemma A.6 can be summarized as follows. By assumption, the principal's constraint (PC$_t$) at date $\hat{t}$ holds as an equality, i.e.

$$\tilde{w}_{\hat{t}} = \sum_{s=\hat{t}+1}^{\infty} \delta^{s-\hat{t}} \left(\tilde{e}_s - \tilde{w}_s\right).$$

We establish (in Lemma A.1) that $\psi'(\tilde{e}_t) \le v'(\bar{c}_\infty)$ for all t, which means that (under Condition (FP$_t^{\text{un}}$)) reductions in effort *reduce* per-period profits to the principal. Because effort decreases weakly over time, and using the consequence of "fastest payments" in Condition (5) (which, roughly stated, says that payments relative to effort must increase over time), we then have $\tilde{e}_{\hat{t}+1} - \tilde{w}_{\hat{t}+1} > \tilde{e}_s - \tilde{w}_s$ for all $s > \hat{t}+1$. Therefore,

$$\tilde{w}_{\hat{t}} = \sum_{s=\hat{t}+1}^{\infty} \delta^{s-\hat{t}} \left(\tilde{e}_s - \tilde{w}_s\right) > \sum_{s=\hat{t}+2}^{\infty} \delta^{s-\hat{t}-1} \left(\tilde{e}_s - \tilde{w}_s\right) \ge \tilde{w}_{\hat{t}+1}$$

where the second inequality is the principal's constraint (PC_t) at date $\hat{t}+1$. This shows why the principal's constraint holding with equality at $\hat{t}$ implies a strictly lower payment and (using (5)) effort in the following period $\hat{t}+1$. In turn, using Lemma 4.2, the principal's constraint must hold again with equality at $\hat{t}+1$. That is, if the principal's constraint holds with equality at a given date, it must hold with equality from then on. This shows how the declining profitability of the relationship translates into payments and hence effort that decrease with time.

The above argument highlights that the principal's constraint eventually binds in each period, which means the principal is indifferent between paying the agent and reneging (we show below that this must be true in *any* optimal contract, not only those satisfying Condition (FP_t^{un})). This observation is also made in Ray (2002) as a step to providing, in a quite general (though distinct) relational contracting environment, a sense in which agent payoffs are backloaded.[12]

We would like to translate the findings of Proposition 4.3 into predictions for payments and the agent's balance. This is complicated by a potential multiplicity of optimal payment paths $(\tilde{w}_t)_{t\geq 1}$. That is, while Lemma 4.1 tells us that it is optimal for Condition (FP_t^{un}) to hold at all dates, this is not the only possibility. We therefore obtain a partial converse for Lemma 4.1. In particular, we show that, once the principal's constraints begin to bind after date $\bar{t}$, payments to the agent are uniquely determined by Condition (FP_t^{un}).

Proposition 4.4. *Suppose the principal cannot attain the first-best payoff in a self-enforceable contract. Fix any optimal policy $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ and let $\bar{t}$ be the date from which effort is strictly decreasing (see Proposition 4.3). Then, Condition (FP_t^{un}) holds for all $t > \bar{t}$.*

Proposition 4.4 implies that, given an optimal effort policy $(\tilde{e}_t)_{t\geq 1}$, payments $\tilde{w}_t$ are uniquely determined from date $\bar{t}+1$ onwards. Our arguments, which took payments to be determined by Condition (FP_t^{un}), then imply immediately the dynamics.

Corollary 4.1. *Suppose the principal cannot attain the first-best payoff in a self-enforceable relational contract. Fix any optimal contract $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ and let $\bar{t}$ be the date from which effort is strictly decreasing (see Proposition 4.3). Then, payments to the agent strictly decrease from date $\bar{t}+1$ onwards, while the agent's balances strictly increase.*

The reason for the result is the one explained above. Fixing an optimal contract, the principal's constraint (PC_t) holds with equality at any $\hat{t} > \bar{t}$. As explained above, this permits us to conclude that $\tilde{w}_{\hat{t}+1} < \tilde{w}_{\hat{t}}$.

[12] Ray considers settings where enforceability constraints prevent the implementation of an efficient contract. His main result is that, after enough time, continuation contracts maximize agent payoffs over constrained efficient continuation contracts that satisfy enforceability constraints.

The fact that the agent's balance increases over time follows straightforwardly from Equations (1) and (2), taken to hold with equality. These observations, together with the fact that the agent consumes a constant $\bar{c}_\infty$ per period, yield in particular that

$$\tilde{b}_t = \frac{\bar{c}_\infty}{1-\delta} - \sum_{\tau \geq t} \delta^{\tau - t} \tilde{w}_\tau$$

which strictly increases with t when payments to the agent fall over time.

Note that, when $\bar{t} > 1$, the principal's constraint (PC_t) is initially slack. In this case, Condition ($\mathrm{FP}_t^{\mathrm{un}}$) need not hold at $t < \bar{t}$, and so payments before date $\bar{t}$ are not uniquely determined. When this "fastest payments" condition is nonetheless taken to hold, payments in fact increase over time up to date $\bar{t}$ (as was mentioned above).

5 Observed consumption

We now study the case where, at each time t, before making the payment w_t, the principal can observe the agent's past and present-period effort choices $(e_s)_{s=1}^t$ as well as past and present-period consumption choices $(c_s)_{s=1}^t$. Since payments and consumption are commonly observed, the balance b_t at the beginning of each period t is also commonly known (using Equation (1)).

The game is now one of complete information, and the relevant solution concept *subgame perfect Nash* equilibrium (SPNE). Both players observe at date t the public history $h_t = (e_s, w_s, c_s)_{1 \leq s < t}$. The set of such histories at date $t \geq 1$ is $\mathcal{H}_t = \mathbb{R}_+^{2(t-1)} \times \mathbb{R}^{t-1}$ (with the convention that $\mathbb{R}^0 = \emptyset$). The date-t balance is written $b(h_t)$. Re-using notation from Section 4 introduces no confusion, so we describe a strategy for the agent as a collection of functions

$$\alpha_t : \mathcal{H}_t \to \mathbb{R}_+ \times \mathbb{R},\ t \geq 1,$$

and a strategy for the principal as a collection of functions

$$\sigma_t : \mathcal{H}_t \times \mathbb{R}_+ \to \mathbb{R}_+,\ t \geq 1.$$

Here, α_t maps the t-history of outcomes to a pair (e_t, c_t) of effort and consumption. Also, σ_t maps these outcomes, together with the observed effort choice e_t of the agent, to a payment w_t. As in Section 4, we restrict strategies for the agent by requiring that, for any strategy $(\alpha_t)_{t=1}^\infty$ employed by the agent, there is $l \in \mathbb{R}$ such that $b\left(h_t^A\right) - c_t$ is bounded above l, over all t and histories h_t^A.

Again we identify a relational contract with the equilibrium outcomes. Also as before, our objective is to obtain relational contracts that maximize the principal's payoff. A first step is then to determine the (public) outcomes $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ that can be sustained in equilibrium (outcomes that we then also label "self-enforceable"). Analogous to the arguments made in the previous section, we suppose deviations from the agreed contract are punished by "autarky". That is, when either player deviates from the contract, all future effort and payments cease, and the agent perfectly smooths his balance over time (consuming $b_t(1-\delta)$ when his balance is b_t). Of course, now autarky follows not only deviations in effort and payments, but also in consumption.

Suppose that the agreed contract is $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$, and deviations are punished by autarky. The agent's payoff, if complying until date $t-1$ and optimally failing to comply from t onwards, is now

$$\sum_{s=1}^{t-1} \delta^{s-1}\left(v\left(\tilde{c}_s\right) - \psi\left(\tilde{e}_s\right)\right) + \delta^{t-1}\frac{v((1-\delta)\tilde{b}_t)}{1-\delta}.$$

This takes into account that the agent who fails to comply from date t onwards optimally exerts no effort and consumes $(1-\delta)\tilde{b}_t$ per period; the consumption choice is optimal given the restriction on agent strategies, which recall implies any sequence of agent consumptions $(c_s)_{s\geq t}$ satisfies $\sum_{s=t}^{\infty}\delta^{s-t}c_s \leq \tilde{b}_t$. Thus, the agent is willing to follow the prescription of the contract if and only if, at all dates t,

$$\frac{v\left(\tilde{b}_t\left(1-\delta\right)\right)}{1-\delta} \leq \sum_{s=t}^{\infty}\delta^{s-t}\left(v\left(\tilde{c}_s\right) - \psi\left(\tilde{e}_s\right)\right). \qquad (\mathrm{AC}_t^{\mathrm{ob}})$$

The key difference to Condition ($\mathrm{AC}_t^{\mathrm{un}}$) is that continuing to publicly honor the agreement up to date $t-1$ ensures that the agent begins period t with the specified balance $\tilde{b}_t$, which in turn determines the wealth he has available to spend in autarky. Condition ($\mathrm{AC}_t^{\mathrm{un}}$), on the other hand, takes into account that the agent who publicly deviates at date t (by shirking on effort) can save in advance for this event, because consumption is not observed.

We can characterize the sustainable outcomes as follows.

Proposition 5.1. *A feasible sequence $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$ is the outcome of a SPNE in the environment where consumption is observed if and only if, for all $t \geq 1$, Conditions (AC_t^{ob}) and (PC_t) are satisfied.*

Notice here that the principal's constraint (PC_t) is the same as in Section 4. A feasible sequence $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$ satisfying the conditions in the proposition is again *self-enforceable* and such a contract chosen to maximize the principal's payoff is *optimal.* We can now state a result similar to Lemma 4.1: it is without loss of generality to focus on relational contracts where the agent is indifferent to quitting at any period.

Lemma 5.1. *For any optimal contract, there exists another optimal contract* $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ *with the same effort and consumption where the timing of payments ensures that agent constraints hold with equality in all periods; that is, for all* $t \geq 1$,

$$\frac{v\left(\tilde{b}_t\,(1-\delta)\right)}{1-\delta} = \sum_{s=t}^{\infty} \delta^{s-t}\left(v\left(\tilde{c}_s\right) - \psi\left(\tilde{e}_s\right)\right). \tag{7}$$

Lemma 5.1 implies that we can focus on relational contracts where, for all $t \geq 1$,

$$v(\tilde{c}_t) - \psi(\tilde{e}_t) + \tfrac{\delta}{1-\delta} v((1-\delta)\tilde{b}_{t+1}) = \tfrac{1}{1-\delta} v((1-\delta)\tilde{b}_t). \tag{FP$_t^{\text{ob}}$}$$

This says that the agent is indifferent between quitting at date t and smoothing the balance $\tilde{b}_t$ optimally over the infinite future, and working one more period, exerting effort $\tilde{e}_t$ and consuming $\tilde{c}_t$, before quitting at date $t+1$ (i.e., ceasing to exert effort) and smoothing the balance $\tilde{b}_{t+1}$ over the infinite future.

5.1 Implementing the first best

If consumption is observable, contracts that condition on the observed consumption permit additional flexibility in the timing of payments. For example, suppose that the principal fully commits to a contract that insists on consumption $c^{FB}(b_1)$ in every period and makes a constant payment $\tilde{w}_t = w^{FB}(b_1) \equiv c^{FB}(b_1) - (1-\delta)b_1$ at the end of each period t, provided that effort $e^{FB}(b_1)$ and consumption $c^{FB}(b_1)$ have always been chosen. After any deviation from this effort and consumption, payments to the agent are zero forever after. The agent is willing to abide by the contract, and the principal's payoff can then be written

$$V^{FB}(b_1) = \frac{e^{FB}(b_1) - w^{FB}(b_1)}{1-\delta}.$$

Note then that, having been compliant in the contract, the agent reaches any date with a balance b_1, and is then indifferent between quitting the contract and continuing to abide by it forever. In fact, these payments are the "fastest", ensuring that Equation (7) is always satisfied. An important observation will be that the agent is thus paid earlier than for the "fastest payments" of the unobservable-consumption case (i.e., than for the payments determined to satisfy Equation (FP$_t^{\text{un}}$) given first-best effort; recall that these payments increase strictly over time, although the NPV of the payments is the same as for the observable case).

Now consider when the first-best policy is implementable given the principal cannot commit. By Lemma 5.1, we can restrict attention to the same contracts, where the agent is indifferent in every period between quitting and continuing obediently. We then have the following result.

Proposition 5.2. *Suppose that neither the principal nor agent can commit to the terms of the agreement and that consumption is observable. Then the principal attains the first-best payoff in an optimal contract if and only if*

$$w^{FB}(b_1) \leq \frac{\delta}{1-\delta}\big(e^{FB}(b_1) - w^{FB}(b_1)\big). \tag{8}$$

Condition (8) is more easily satisfied than Condition (6) (the condition for the unobervable consumption case); i.e., if the first-best effort and consumption is sustained in the unobservable consumption case, then it is sustained when consumption is observed. To see this, note that by concavity of v, and because $c^{FB}(b_1) > (1-\delta)b_1$, we have

$$v\big(c^{FB}(b_1)\big) - v\left((1-\delta)b_1\right) > v'\big(c^{FB}(b_1)\big)\big(c^{FB}(b_1) - (1-\delta)b_1\big) = v'\big(c^{FB}(b_1)\big)w^{FB}(b_1).$$

Because the agent is kept indifferent to quitting, $v\big(c^{FB}(b_1)\big) - v\left((1-\delta)b_1\right) = \psi\big(e^{FB}(b_1)\big)$. Therefore,

$$w^{FB}(b_1) < \frac{\psi\left(e^{FB}(b_1)\right)}{v'\left(c^{FB}(b_1)\right)}.$$

We have thus observed that the constant payment to the agent in the observed-consumption case, namely $w^{FB}(b_1)$, is below the limiting payment in the unobserved-consumption case, namely $\frac{\psi\left(e^{FB}(b_1)\right)}{v'\left(c^{FB}(b_1)\right)}$, implying that the contract is more easily self-enforceable in the former case.

5.2 Optimal contract with observed consumption

Now consider the problem of deriving an optimal contract. We can restrict attention to payments that keep the agent indifferent to quitting the agreement, as given in Lemma 5.1. The agent's problem can now be written recursively, with the balance $\tilde{b}_t$ a state variable for the relationship, applying the principle of optimality. Indeed, suppose for any date t that the continuation contract $(\tilde{e}_s, \tilde{w}_s, \tilde{c}_s, \tilde{b}_s)_{s\geq t}$ does not maximize the continuation value of the relationship to the principal $\sum_{s=t}^{\infty} \delta^{s-t}(e_s - w_s)$, subject to the satisfaction of constraints (PC_t) from t onwards; i.e., there is some more profitable continuation contract $(\tilde{e}'_s, \tilde{w}'_s, \tilde{c}'_s, \tilde{b}'_s)_{s\geq t}$. Then this contract can be substituted, increasing the continuation value $\sum_{s=t}^{\infty} \delta^{s-t}(e_s - w_s)$, maintaining the agent indifference conditions (7) at all dates and continuing to satisfy the principal's constraints (PC_t), and increasing the principal's payoff from the contract overall.

Since the optimal contract maximizes the principal's continuation profits, an optimal contract $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$ must solve a sequence of sub-problems with value $V(\tilde{b}_t)$ given by

$$V(\tilde{b}_t) = \max_{e_t\in\mathbb{R}_+;\ b_{t+1}, c_t\in\mathbb{R}} \left(e_t - \Big(\delta b_{t+1} - \tilde{b}_t + c_t\Big) + \delta V\left(b_{t+1}\right)\right) \tag{9}$$

subject to the agent's indifference condition

$$v(c_t) - \psi(e_t) + \tfrac{\delta}{1-\delta} v((1-\delta)b_{t+1}) = \tfrac{1}{1-\delta} v((1-\delta)\tilde{b}_t). \tag{10}$$

and to the principal's constraint

$$\delta b_{t+1} - \tilde{b}_t + c_t \leq \delta V\left(b_{t+1}\right). \tag{11}$$

The left-hand side of (11) can be understood as the agent's payment w_t, which is divided into date-t consumption $c_t \in \mathbb{R}$ and savings $\delta b_{t+1} - b_t \in \mathbb{R}$. Note that the effort e_t must be non-negative; non-negativity of the payment $\delta b_{t+1} - \tilde{b}_t + c_t$ is assured by the equality (10).

We show that any optimal policy for the principal can be characterized as follows.

Proposition 5.3. *An optimal contract exists. Suppose that, given the balance b_1, an optimal contract $(\tilde{e}_t, \tilde{w}_t, \tilde{c}_t, \tilde{b}_t)_{t\geq 1}$ fails to obtain the first-best payoff $V^{FB}\left(b_1\right)$. Then the agent's balance $\tilde{b}_t$ and consumption $\tilde{c}_t$ decline strictly over time, with $\tilde{b}_t \to \tilde{b}_\infty$ for some $\tilde{b}_\infty > 0$. Effort $\tilde{e}_t$ and the payments $\tilde{w}_t$ determined by the Conditions (FP_t^{ob}) increase strictly over time. We have $V\left(\tilde{b}_t\right) \to V^{FB}\left(\tilde{b}_\infty\right)$ as $t \to \infty$, and hence effort and consumption converge to first-best levels for the balance $\tilde{b}_\infty$.*

A heuristic account of the forces behind this result is as follows. When the principal's constraint ((PC_t) or equivalently (11)) binds, effort is suppressed. That is, if the principal could increase effort and credibly increase payments to keep the agent as well off, she would gain by doing so. However, the principal's value function $V(\cdot)$ is strictly decreasing; intuitively, because a lower balance makes the agent cheaper to compensate to keep him in the agreement. Therefore, for any date t, reducing the balance b_{t+1} increases the principal's continuation payoff $V\left(\tilde{b}_{t+1}\right)$ and relaxes the principal's date-t constraint (11). Therefore, the principal asks the agent to consume earlier than he would like, driving the balance down over time. This continues to a point where, given the revised balance, the contract is close to first best, and so the value of continuing to distort consumption vanishes.

It is worth pointing out here that the dynamics of $V\left(\tilde{b}_t\right)$ are determinative of both the dynamics of effort and payments. When there is no self-enforceable first-best contract, $V\left(\tilde{b}_t\right)$ strictly increases with t, and moreover the principal's constraint (11) binds. Therefore, we must have both $\tilde{w}_t = \delta V\left(\tilde{b}_t\right)$ and $V\left(\tilde{b}_t\right) = \tilde{e}_t - \tilde{w}_t + \delta V\left(\tilde{b}_{t+1}\right) = \tilde{e}_t$. (Note that the conclusion the the principal's constraint (11) binds is obtained under the assumption that Conditions (FP_t^{ob}) hold at all dates; but we establish in Proposition 5.4 below that, when the first-best payoff is not attainable, this is not only consistent with optimality but also necessary.)

It is also worth mentioning how we show $V\left(\tilde{b}_t\right)$ is strictly increasing with t. As noted above, that $V(\cdot)$ is a strictly decreasing function seems intuitive given that the agent should be easier to motivate when his balance is low. To see this formally (following the logic of Lemma A.13 in the Appendix), notice that when the agent's indifference condition ($\mathrm{FP}_t^{\mathrm{ob}}$) holds, we have

$$\tilde{e}_t = \psi^{-1}\left(v(\tilde{c}_t) - \tfrac{1}{1-\delta}v((1-\delta)\tilde{b}_t) + \tfrac{\delta}{1-\delta}v((1-\delta)\tilde{b}_{t+1})\right).$$

Consider a small decrease in the date-t balance $\tilde{b}_t$. If date-t consumption is reduced by the same amount, then we can keep the date-$t+1$ balance $\tilde{b}_{t+1}$ unchanged, we can keep the date-$t+1$ continuation contract unchanged, and by implication payments to the agent do not change. Effort is determined by the above equation, i.e. to maintain agent indifference. Note then that (as we establish in Lemma A.11), $\tilde{c}_t > (1-\delta)\tilde{b}_t$, so that $v'(\tilde{c}_t) < v'((1-\delta)\tilde{b}_t)$, showing that effort in fact *increases*. This shows that profits to the principal increase. This argument can then be used to show that $V(\cdot)$ is increasing. Finally, we have to show that $\tilde{b}_t$ strictly decreases with t. The argument is more involved and involves a variational argument: if $\tilde{b}_t$ fails to be strictly decreasing in t, it is possible to specify a strictly more profitable contract (this is established in Lemma A.12 in the Appendix).

Another part of our analysis that is worth mentioning is an Euler equation

$$1 - \frac{v'\left((1-\delta)\,\tilde{b}_{t+1}\right)}{v'(\tilde{c}_t)} = \frac{v'(\tilde{c}_{t+1})}{\psi'(\tilde{e}_{t+1})}\left(1 - \frac{v'\left((1-\delta)\,\tilde{b}_{t+1}\right)}{v'(\tilde{c}_{t+1})}\right)$$

which must hold for an optimal contract at all dates t, and which we use to derive several key properties. Such a condition is derived (in Lemma A.11) by fixing the contract at and before $t-1$, and from date $t+2$ onwards, and then requiring the contractual variables at t and $t+1$ to be chosen optimally. The equation captures the relationship between a static distortion in effort and a dynamic distortion in consumption. In particular, when the first best is not sustainable, we are able to show that $\psi'(\tilde{e}_{t+1}) < v'(\tilde{c}_{t+1})$ (reflecting a static (downward) distortion in effort), and correspondingly $(1-\delta)\,\tilde{b}_{t+1} < \tilde{c}_{t+1} < \tilde{c}_t$ (i.e., consumption strictly decreases over time, which is a dynamic distortion). Such a trade-off should be anticipated, since declining consumption (which is an inefficient departure from constant consumption) increases the agent's marginal utility of consumption, which makes him easier to motivate and permits higher effort and profits at later dates, relaxing the principal's credibility constraint (PC_t) and permitting higher effort also early on. As $\tilde{b}_t \to \tilde{b}_\infty$, consumption falls to its lower bound, becoming almost constant, so $\frac{v'(\tilde{c}_{t+1})}{\psi'(\tilde{e}_{t+1})} \to 1$, which accords with convergence to the first-best policy.

Finally, analogous to Proposition 4.4, we would like to show also that, when there is no self-enforceable first-best contract, the timing of payments is uniquely determined (given the effort and consumption policies) by the Conditions ($\mathrm{FP}_t^{\mathrm{ob}}$).

Proposition 5.4. *Suppose the principal cannot attain the first-best payoff in a self-enforceable relational contract. Then, in any policy that is optimal for the principal, Condition (FP_t^{ob}) holds at all dates. Hence payments to the agent strictly increase over time.*

The logic of this result is that if the Condition ($\mathrm{FP}_t^{\mathrm{ob}}$) fails, then payments can be made earlier in time, while maintaining the agent constraints ($\mathrm{AC}_t^{\mathrm{ob}}$). This induces slack in the principal's constraint (PC_t), which can then be exploited by increasing payments and effort, increasing profits. As noted for the case of unobservable consumption, such an observation is related to arguments in Ray (2002).

6 Optimal relational contract in the absence of savings

Having studied optimal relational contracts when the agent can save, it is natural to ask: what happens if savings are ruled out? Dynamics in the above scenarios were driven by the evolution, or potential evolution, of the balance on the savings account. We might then conjecture that an optimal contract in the absence of savings should be stationary, mirroring what occurs when the agent has linear preferences over consumption (see Levin, 2003, and the discussion by Malcomson, 2013). Below we provide a framework where this conjecture is correct.

Consistent with the analysis above, let us restrict attention to deterministic relational contracts, which are sequences of effort and consumption $(e_t, c_t)_{t\geq 1}$, with the interpretation that the agent is paid and consumes c_t at date t. We require $e_t \geq 0$ for all t. In addition, to avoid the implication of infinite profits for the principal, we assert that the principal must pay $c_t \geq \underline{c}$ in all periods, for $\underline{c}$ satisfying $v(\underline{c}) = 0$. The payoff zero from the minimal payment $\underline{c}$ is a normalization. One interpretation is that the relationship is governed by an established employment agreement which stipulates payment at least $\underline{c}$ irrespective of the agent's performance (think of this as a base contract). See Fong and Li (2017) for a similar wage floor in a model of moral hazard with stochastic output and agent limited liability. (Note that we could give the environments with savings above a similar interpretation if the agent is taken to have initial balance zero on the account, and the principal is obliged to pay at least $(1-\delta)\, b_1$ per period irrespective of the agent's performance. In this case, the payments considered in the earlier analysis are in addition to the amount $(1-\delta)\, b_1$.)

As in the previous section, we are in a setting of complete information, the relevant solution concept is SPNE, and we can consider any deviation in effort e_t or payment c_t to be punished by autarky. In this case, autarky involves effort set to zero and payments set to $\underline{c}$.

A relational contract $(\tilde{e}_t, \tilde{c}_t)_{t\geq 1}$ is now self-enforceable if the principal's constraints hold, which can be written

$$\sum_{s=t}^{\infty} \delta^{s-t} \tilde{c}_s \leq \sum_{s=t+1}^{\infty} \delta^{s-t} \tilde{e}_s \tag{12}$$

for all $t \geq 1$, and if the agent's constraints holds, i.e.

$$U_t \geq 0 \tag{13}$$

for all $t \geq 1$, where $U_t \equiv \sum_{s=t}^{\infty} \delta^{s-t}(v(\tilde{c}_s) - \psi(\tilde{e}_s))$ is the continuation utility of the agent at time t. The contract design problem is to maximize the principal's payoff

$$\sum_{t=1}^{\infty} \delta^{t-1} (e_t - c_t)$$

by choice of a self-enforceable contract $(e_t, c_t)_{t\geq 1}$.

Proposition 6.1. *There is a unique optimal contract, and it is stationary. Letting* (e^{FB}, c^{FB}) *be the unique pair satisfying* $\psi'(e^{FB}) = v'(c^{FB})$ *and* $\psi(e^{FB}) = v(c^{FB})$*, we have*

1. *If* $c^{FB} \leq \delta e^{FB}$*, then the optimal contract satisfies* $(\tilde{e}_t, \tilde{c}_t) = (e^{FB}, c^{FB})$ *for all* $t \geq 1$.
2. *If* $c^{FB} > \delta e^{FB}$*, then the optimal contract satisfies* $(\tilde{e}_t, \tilde{c}_t) = (e^*, \delta e^*)$ *for all* $t \geq 1$*, where* e^* *is the unique value satisfying* $\psi(e^*) = v(\delta e^*)$.

Proof. **Step 1.** Let $(\tilde{e}_t, \tilde{c}_t)_{t\geq 1}$ be an optimal contract. We first note that $U_1 = 0$. Otherwise, the effort at time 1 can be increased, so the payoff of the principal increases and, if the effort increase is small enough, all constraints are satisfied.

Step 2. We now prove that, for all $t \geq 1$, if $U_{t+1} > 0$ then $\tilde{c}_{t+1} \leq \tilde{c}_t$ and $\tilde{e}_{t+1} \geq \tilde{e}_t$. To see the former we assume, for the sake of contradiction, that $\tilde{c}_{t'+1} > \tilde{c}_{t'}$ for some $t' \geq 1$. Consider a contract $(\tilde{e}'_t, \tilde{c}'_t)_{t\geq 1}$ where payments at times different from t' and $t'+1$ remain the same, and all efforts remain the same. Slightly increase the date-t' payment so that $\tilde{c}'_{t'} > \tilde{c}_{t'}$ and slightly reduce date-$t'+1$ consumption so that $\tilde{c}'_{t'+1} < \tilde{c}_{t'+1}$, maintaining $\tilde{c}'_{t'+1} > \tilde{c}'_{t'}$, as well as

$$v(\tilde{c}'_{t'}) + \delta v(\tilde{c}'_{t'+1}) = v(\tilde{c}_{t'}) + \delta v(\tilde{c}_{t'+1}) \ .$$

The agent's constraint (13) is affected only at date $t'+1$: $U_{t'+1}$ decreases, but if the modification to the contract is small enough, we have $U_{t'+1} > 0$ and so the agent's constraint is satisfied.

Since $\tilde{c}'_{t'+1} < \tilde{c}_{t'+1}$ (while all consumption and efforts after $t'+1$ remain the same), the principal's constraint (12) is relaxed at time $t'+1$. Given that v is strictly concave, the time-t' present value of the payments decreases, so the principal's constraint is relaxed at all dates t' and earlier. Hence, the contract $(\tilde{e}'_t, \tilde{c}'_t)_{t\geq 1}$ satisfies the relevant constraints and it is strictly more profitable than the original, contradicting the optimality of the original.

We now prove that, for all $t \geq 1$, if $U_{t+1} > 0$ then $\tilde{e}_{t+1} \geq \tilde{e}_t$. To see this we assume, for the sake of contradiction, that $\tilde{e}_{t'+1} < \tilde{e}_{t'}$ for some $t' \geq 1$. Consider a contract $(\tilde{e}'_t, \tilde{c}'_t)_{t\geq 1}$ where payments remain the same at all times, and where efforts differ only at t' and $t'+1$. In particular, we set $\tilde{e}'_{t'+1}$ slightly higher than $\tilde{e}_{t'+1}$, and choose $\tilde{e}'_{t'}$ so that

$$\psi(\tilde{e}'_{t'}) + \delta\psi(\tilde{e}'_{t'+1}) = \psi(\tilde{e}_{t'}) + \delta\psi(\tilde{e}_{t'+1}) \ .$$

The agent's constraints (13) are then unaffected, except at $t'+1$ where the constraint is tightened, but for a small enough adjustment to the contract we still have $U_{t'+1} > 0$. Principal constraints (12) remain unaffected from $t'+1$ onwards. Since $\tilde{e}'_{t'+1} > \tilde{e}_{t'+1}$ (while all consumption and efforts after $t'+1$ remain the same), the principal's constraint is relaxed at time t'. Given that ψ is convex, the date-t' present value of the effort increases (provided the changes in effort are small enough), so the principal's constraint is relaxed at all times before t' as well. Again, the modified contract is more profitable, which gives our contradiction.

Step 3. We now prove that $U_t = 0$ for all t. We already established in Step 1 that $U_1 = 0$. Now suppose we have established that $U_s = 0$ for all $s \in \{1, \ldots, t\}$; we want to show that $U_{t+1} = 0$. The result will then follow by induction.

Suppose then for a contradiction that $U_{t+1} > 0$. Let then $t+T$ be the first date after t at which $U_{t+T} = 0$ ($T = +\infty$ means that $U_s > 0$ for all $s > t$), so we have assumed $T > 1$. We have

$$U_t = \sum_{s=t}^{t+T-1} \delta^{s-t}(v(\tilde{c}_s) - \psi(\tilde{e}_s)) \geq \sum_{s=t}^{t+T-2} \delta^{s-t}(v(\tilde{c}_s) - \psi(\tilde{e}_s)) \ ,$$

where the inequality follows from the fact that (1) when $T < +\infty$, since $U_{t+T-1} > 0$, we have $v(\tilde{c}_{t+T-1}) - \psi(\tilde{e}_{t+T-1}) > 0$, and (2) when $T = +\infty$, the equation holds with equality. Note then that, since $\tilde{c}_{s+1} \leq \tilde{c}_s$ and $\tilde{e}_{s+1} \geq \tilde{e}_s$ for all $s \in \{t, ..., t+T-2\}$ we have

$$U_t \geq \sum_{s=t}^{t+T-2} \delta^{s-t}(v(\tilde{c}_{s+1}) - \psi(\tilde{e}_{s+1})) = \sum_{s=t+1}^{t+T-1} \delta^{s-t-1}(v(\tilde{c}_s) - \psi(\tilde{e}_s)) = U_{t+1} > 0 \ .$$

This is a contradiction because $U_t = 0$. This establishes that, indeed, $U_{t+1} = 0$.

Note that the conclusion $U_t = 0$ for all t implies that the optimal contract is stationary. For any date t the continuation of an optimal contract $(\tilde{e}_s, \tilde{c}_s)_{s\geq t}$ must maximize continuation profits

$$\sum_{s=t}^{\infty} \delta^{s-t} (e_s - c_s)$$

subject to $v(c_s) = \psi(e_s)$ and to principal constraints (12)

$$\sum_{s'=s}^{\infty} \delta^{s'-s} \tilde{c}_{s'} \leq \sum_{s'=s+1}^{\infty} \delta^{s'-s} \tilde{e}_{s'}$$

for all $s \geq t$. In particular, note that maximizing continuation profits at date t only relaxes the principal's constraints (12) at dates $t-1$ and earlier.

Step 4. Given that any optimal relational contract is stationary, we can now consider the problem of maximizing by choice of effort e and payment c the per period profit $e-c$ subject to the agent earning zero payoff (that this must hold is established above), i.e. $v(c) - \psi(e) = 0$, and subject to the principal's constraint (12) which takes the form $c \leq \delta e$. Case 1 in the proposition then corresponds to the case where the principal's constraint does not bind, while Case 2 corresponds to the one where it binds. □

The result confirms a sense in which interesting dynamics arise in our models with consumption-smoothing preferences due to the possibility of agent savings. It turns out that, when the agent cannot save, the agent is compensated for effort only through the current-period payment; in particular, the agent's continuation payoff is always zero. As shown in the proof above, this insight markedly simplifies the search for an optimal contract.

7 Conclusions

This paper has studied optimal relational contracts in a simple deterministic setting where the agent has consumption-smoothing preferences and is able to save. We contrasted the case where the agent's consumption is unobservable to the principal (the case of "private savings") and where consumption is observed. In the case where consumption is unobservable, we found that the relationship eventually becomes less profitable with time, meaning that the payments the principal can credibly offer must decline. Hence effort eventually declines with time. When consumption is instead observable, the agent consumes inefficiently early (i.e., saves too little), the balance on his savings account gradually declines, the relationship becomes more profitable as the agent grows easier to incentivize, payments to the agent gradually increase, and the agent's effort increases. In either case, the agent's payoff is equal to his autarky payoff; the fact that the principal earns higher profits when consumption is observable can therefore be entirely attributed to improvements in efficiency. While the agent consumes inefficiently early when consumption is observable, the loss in surplus due to a failure to smooth consumption is more than offset by an amelioration of the moral hazard problem: i.e., the agent can be induced to exert higher effort.

We conclude with some open questions. First, while we studied the extremes of observable and unobservable consumption, we think there would be value in considering a model where only part of an agent's consumption expenditures are observable. For instance, the agent might gain utility from both "conspicuous" and "inconspicuous" consumption. An optimal relational contract might call for inefficient amounts of the former to impoverish the agent over time, similar to the case of observable consumption studied in Section 5. Unfortunately, such a model seems less tractable than the cases of fully observable or fully unobservable consumption.

Another question we left unexplored is the possible role of exogenous uncertainty. This could take a range of forms. For instance, the agent's initial balance could be random and the agent's private information, the agent could receive taste or income shocks over time, there could be random exogenous separations, or there could be shocks to the principal's willingness or ability to pay compensation. These possibilities again seem to substantially complicate the analysis. Further questions concern both testable implications and implications for policy. In terms of testable implications, one can possibly compare the wage structure and productivity of firms with older and younger workers. In terms of policy, one could examine in the model the role of unemployment insurance as well as incentives/subsidies to save.

References

[1] Abraham, Arpad, Sebastian Koehne, and Nicola Pavoni, 2011, 'On the first-order approach in principal–agent models with hidden borrowing and lending,' *Journal of Economic Theory*, 146, 1331-1361.

[2] Atkeson, Andrew and Robert E. Lucas, 1992, 'On Efficient Distribution with Private Information,' *Review of Economic Studies*, 59, 427-453.

[3] Baker, George, Robert Gibbons and Kevin J. Murphy, 1994, 'Subjective Performance Measures in Optimal Incentive Contracts,' *Quarterly Journal of Economics*, 109, 1125-1156.

[4] Bennardo, Alberto, Pierre-Andre Chiappori and Joon Song, 2010, 'Perks as Second Best Optimal Compensations,' CSEF Working Paper No. 244.

[5] Board, Simon, 'Relational Contracts and the Value of Loyalty,' *American Economic Review*, 101, 3349-3367.

[6] Bull, Clive, 1987, 'The Existence of Self-Enforcing Implicit Contracts,' *Quarterly Journal of Economics*, 147-159.

[7] Chassang, Sylvain, 2010, 'Building Routines: Learning, Cooperation, and the Dynamics of Incomplete Relational Contracts,' *American Economic Review*, 100, 448-465.

[8] Di Tella, Sebastian and Yuliy Sannikov, 2018, 'Optimal Asset Management Contracts with Hidden Savings,' mimeo Stanford University.

[9] Edmans, Alex, Xavier Gabaix, Tomasz Sadzik, and Yuliy Sannikov, 2012, 'Dynamic CEO Compensation,' *Journal of Finance*, 67, 1603-1647.

[10] Fong, Yuk-fai and Jin Li, 2017, 'Relational contracts, limited liability, and employment dynamics,' *Journal of Economic Theory*, 169, 270-293.

[11] Fuchs, William, 2007, 'Contracting with Repeated Moral Hazard and Private Evaluations,' *American Economic Review*, 97, 1432-1448.

[12] Fudenberg, Drew, Bengt Holmstrom and Paul Milgrom, 1990, 'Short-term contracts and long-term agency relationships,' *Journal of Economic Theory*, 51, 1-31.

[13] Fudenberg, Drew and Luis Rayo, 2019, 'Training and Effort Dynamics in Apprenticeship,' mimeo MIT and Northwestern U..

[14] Garicano, Luis and Luis Rayo, 2017, 'Relational Knowledge Transfers,' *American Economic Review*, 107, 2695-2730.

[15] Garrett, Daniel and Alessandro Pavan, 2015, 'Dynamic Managerial Compensation: A Variational Approach,' *Journal of Economic Theory*, 159, 775-818.

[16] Halac, Marina, 'Relational Contracts and the Value of Relationships,' *American Economic Review*, 102, 750-79.

[17] He, Zhiguo, 2012, 'Dynamic Compensation Contracts with Private Savings,' *Review of Financial Studies*, 25, 1494-1549.

[18] Levin, Jonathan, 2003, 'Relational Incentive Contracts,' *American Economic Review*, 93, 835-857.

[19] Li, Jin and Niko Matouschek, 2013, 'Managing Conflicts in Relational Contracts,' *American Economic Review*, 103, 2328-51.

[20] MacLeod, W. Bentley, 2007, 'Reputations, Relationships, and Contract Enforcement,' *Journal of Economic Literature*, 45, 595-628.

[21] Macleod, W. Bentley, and James M. Malcomson, 1989, 'Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment,' *Econometrica*, 57, 447-480.

[22] Malcomson, James M., 2013, 'Relational incentive contracts' In Robert Gibbons and John Roberts (eds.), *Handbook of Organizational Economics*. Princeton University Press.

[23] Malcomson, James M., 2016, 'Relational incentive contracts with persistent private information,' *Econometrica*, 84, 317-346.

[24] McAdams, David, 2011, 'Performance and Turnover in a Stochastic Partnership,' *American Economic Journal: Microeconomics*, 3, 107-142.

[25] Pearce, David G. and Ennio Stacchetti, 1993, 'The Interaction of Implicit and Explicit Contracts in Repeated Agency,' *Games and Economic Behavior*, 23, 75-96.

[26] Ray, Debraj, 2002, 'The Time Structure of Self-Enforcing Agreements,' *Econometrica*, 70, 547-582.

[27] Rogerson, William P., 1985, 'Repeated Moral Hazard,' *Econometrica*, 53, 69-76.

[28] Sannikov, Yuliy, 2008, 'A Continuous-Time Version of the Principal-Agent Problem,' *Review of Economic Studies*, 75, 957-984.

[29] Thomas, Jonathan and Tim Worrall, 1990, 'Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem,' *Journal of Economic Theory*, 51, 367-390.

[30] Thomas, Jonathan and Tim Worrall, 1994, 'Foreign Direct Investment and the Risk of Appropriation,' *Review of Economic Studies*, 61, 81-108.

[31] Veblen, Thorstein, 1899, *The Theory of the Leisure Class*, New York: McMillan Co..

[32] Williams, Noah, 2011, 'Persistent Private Information,' *Econometrica,* 79, 1233-1275.

A Proofs

A.1 Proofs of the results in Section 3

Proof of Proposition 3.1.

Proof. Consider the problem of maximizing the principal's payoff

$$\sum_{t=1}^{\infty} \delta^{t-1} \left(\tilde{e}_t - \tilde{w}_t \right)$$

subject to the constraint that

$$\sum_{t=1}^{\infty} \delta^{t-1} \left(v\left(\tilde{c}_t \right) - \psi\left(\tilde{e}_t \right) \right) \geq \frac{v\left(b_1 \left(1-\delta \right) \right)}{1-\delta}$$

together with (2), that is,

$$b_1 + \sum_{t=1}^{\infty} \delta^{t-1} \left(\tilde{w}_t - \tilde{c}_t \right) \geq 0.$$

The Lagrangian for this problem is

$$\sum_{t=1}^{\infty} \delta^{t-1} \left(\tilde{e}_t - \tilde{w}_t \right) + \lambda \left(\sum_{t=1}^{\infty} \delta^{t-1} \left(v\left(\tilde{c}_t \right) - \psi\left(\tilde{e}_t \right) \right) - \frac{v\left(b_1 \left(1-\delta \right) \right)}{1-\delta} \right) + \mu \left(b_1 + \sum_{t=1}^{\infty} \delta^{t-1} \left(\tilde{w}_t - \tilde{c}_t \right) \right)$$

where λ and μ are multipliers on the constraints. First-order conditions are

$$\psi'\left(\tilde{e}_t \right) = 1/\lambda$$

and

$$v'\left(\tilde{c}_t \right) = \frac{\mu}{\lambda}$$

for all t. That is, effort $\tilde{e}_t$ must be constant at some $\tilde{e}$ and consumption $\tilde{c}_t$ constant at some $\tilde{c}$. The two constraints binding is necessary for an optimum. The second constraint yields

$$\sum_{t=1}^{\infty} \delta^{t-1} \tilde{w}_t = \frac{\tilde{c}}{1-\delta} - b_1.$$

The first constraint may be written as

$$\tilde{c} = v^{-1} \left(\psi\left(\tilde{e} \right) + v\left(b_1 \left(1-\delta \right) \right) \right),$$

which is the second condition in the proposition. Incorporating the two constraints, the principal then maximizes

$$\frac{\tilde{e}}{1-\delta}-\frac{v^{-1}\left(\psi\left(\tilde{e}\right)+v\left(b_1\left(1-\delta\right)\right)\right)}{1-\delta}+b_1. \tag{14}$$

The first-order condition for a maximum is

$$1-\frac{\psi'\left(\tilde{e}\right)}{v'\left(\tilde{c}\right)}=0,$$

which gives the first condition in the proposition. Optimal effort solves

$$\psi'\left(\tilde{e}\right)=v'\left(v^{-1}\left(\psi\left(\tilde{e}\right)+v\left(b_1\left(1-\delta\right)\right)\right)\right),$$

where the left-hand side is increasing from zero to $+\infty$ in $\tilde{e}$, and the right-hand side is positive and decreasing in $\tilde{e}$. Thus optimal effort $e^{FB}\left(b_1\right)$ and hence consumption $c^{FB}\left(b_1\right)$ are uniquely determined.

Now consider why the first-best payoff $V^{FB}\left(b_1\right)$ is decreasing in b_1. Standard arguments (see Milgrom and Segal, 2002) can then be used to establish, using (14), that $V^{FB}\left(b_1\right)$ is absolutely continuous and differentiable a.e., with derivative

$$1-\frac{v'\left(b_1\left(1-\delta\right)\right)}{v'\left(c^{FB}\left(b_1\right)\right)}$$

which is strictly negative because $b_1\left(1-\delta\right)<c^{FB}\left(b_1\right)$. □

A.2 Proofs of the results in Section 4

Proof of Proposition 4.1

Proof. **Necessity.** Fix a PBE with outcomes $(\tilde{e}_t,\tilde{c}_t,\tilde{w}_t,\tilde{b}_t)_{t\geq 1}$, say. Suppose first it is not true that $\tilde{c}_t=\bar{c}_\infty$ for all t. Then the agent can choose a strategy which specifies obedient (on-path) effort and specifies consumption given in each period by $\bar{c}_\infty$ (as long as the public history stays on path, otherwise he follows the original strategy). Note that

$$b_1+\sum_{s=1}^{\infty}\delta^{s-1}\left(\tilde{w}_s-\bar{c}_\infty\right)=0.$$

Hence,

$$\begin{aligned}
b_1+\sum_{s=1}^{t-1}\delta^{s-1}\left(\tilde{w}_s-\bar{c}_\infty\right)&=-\sum_{s=t}^{\infty}\delta^{s-1}\left(\tilde{w}_s-\bar{c}_\infty\right)\\
&\geq-\sum_{s=t}^{\infty}\delta^{s-1}\left(\bar{w}-\bar{c}_\infty\right)\\
&=-\frac{\delta^{t-1}}{1-\delta}\left(\bar{w}-\bar{c}_\infty\right),
\end{aligned}$$

where $\bar{w} = \sup_t \{\tilde{w}_t\}$. Hence, letting b_t be the balance implied by the agent's deviation at any date t,

$$b_t - \bar{c}_\infty = b_1 \delta^{-t+1} + \sum_{s=1}^{t-1} \delta^{s-t} \left(\tilde{w}_s - \bar{c}_\infty\right) - \bar{c}_\infty \geq \frac{\bar{w} - \bar{c}_\infty}{1-\delta} - \bar{c}_\infty$$

which shows that the deviation is available to the agent (i.e., $b_t - \bar{c}_\infty$ remains bounded below). Moreover, it is clear that the agent obtains a strictly higher payoff.

If ($\mathrm{AC}_t^{\mathrm{un}}$) is not satisfied, then the agent can consume $\bar{c}_{t-1} \leq \bar{c}_\infty$ in every period and exert the specified effort $e_s = \tilde{e}_s$ up to date $t-1$ (as long as the public history stays on path, otherwise he follows the original strategy), and zero effort from then on, attaining a higher payoff. Because this strategy generates an outcome in which the agent consumes less in every period, the balance b_s is above $\tilde{b}_s$ in all periods up to date t, and it remains strictly positive from then on. Therefore again the deviation is available to the agent (i.e., certainly $b_s - c_s$ is bounded below by $-B$ over all periods s).

If (PC_t) is not satisfied at date t, the principal can pay zero from then on. The principal then attains a non-negative continuation payoff, while by continuing to make the payments $\tilde{w}_t$, the principal obtains a negative continuation payoff

$$\sum_{s=t+1}^{\infty} \delta^{s-t}(\tilde{e}_s - \tilde{w}_s) - \tilde{w}_t.$$

Therefore, the principal has a profitable deviation at date t.

Sufficiency. We provide strategies and beliefs that constitute a Perfect Bayesian Equilibrium with outcome $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ satisfying the conditions of the proposition.

Recall feasibility stipulates that the sequence $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ is bounded. Let

$$A = \sup_t \sum_{\tau=t}^{\infty} \delta^{\tau-t} \tilde{w}_t,$$

and let $B = A/(1-\delta) \geq 0$. The value $-B$ will provide a lower bound on $b\left(h_t^A\right) - c_t$ across all agent histories $h_t^A \in \mathcal{H}_t^A$, where c_t is determined using α_t, so the constraint on agent strategies is respected. (Observe that if the agent's balance at date t, $b\left(h_t^A\right)$, takes a large negative value, then the agent can choose negative consumption c_t at date t to ensure $b\left(h_t^A\right) - c_t \geq -B$.)

We specify a PBE as follows. On the principal's side, put $\sigma_t\left(h_t^P, e_t\right) = \tilde{w}_t$ if $(e_s, w_s) = (\tilde{e}_s, \tilde{w}_s)$ for all $s \leq t-1$ and $e_t = \tilde{e}_t$, and put $\sigma_t\left(h_t^P, \tilde{e}_t\right) = 0$ otherwise.

For the agent's strategy, put $\alpha_t\left(h_t^A\right) = (\tilde{e}_t, \tilde{c}_t)$ if $(e_s, c_s, w_s) = (\tilde{e}_s, \tilde{c}_s, \tilde{w}_s)$ for all $s < t$. Let $C(b) = b + B$ if $b + B \leq 0$ and put $C(b) = b(1-\delta)$ otherwise. Put $\alpha_t\left(h_t^A\right) = (0, C(b_t))$ whenever $(e_s, w_s) \neq (\tilde{e}_s, \tilde{w}_s)$ for some $s \leq t-1$. Then if the agent's balance b_t at such a history h_t^A is such that $b_t + B > 0$, we have $b_t - C(b_t) = \delta b_t > -B$. Conversely, if $b_t + B \leq 0$, $b_t - C(b_t) = -B$. Note that this specification of the agent's strategy will guarantee its sequential optimality at date-t histories where $(e_s, w_s) \neq (\tilde{e}_s, \tilde{w}_s)$ for some $s < t$. This is because, due to our restriction on the agent's strategies, consumption paths $(c_s)_{s=t}^{\infty}$ must satisfy

$$\sum_{s=t}^{\infty} \delta^{s-t} c_s \leq b_t,$$

where b_t is the date-t balance.[13]

Determining the agent's equilibrium strategy for the remaining possible histories, where $(e_s, w_s) = (\tilde{e}_s, \tilde{w}_s)$ for all $s \leq t-1$, and yet $c_s \neq \tilde{c}_s$ for some values $s \leq t-1$, is then more involved. Consider a history h_t^A for the agent with $(e_s, w_s) = (\tilde{e}_s, \tilde{w}_s)$ for all $s \leq t-1$, and yet $c_s \neq \tilde{c}_s$ for some values. First note that, if

$$b\left(h_t^A\right) + \sum_{\tau=t}^{\infty} \delta^{\tau-t} \tilde{w}_\tau \tag{15}$$

is non-positive, then at history h_t^A the payoff of the agent is $-\infty$ under any feasible continuation strategy. In this case we might as well specify that $e_t = \tilde{e}_t$ and $c_t = C\left(b\left(h_t^A\right)\right)$, with C defined above (this ensures that $b\left(h_t^A\right) - c_t \geq -B$, so the restriction on agent strategies is respected). Assume then that the expression (15) is strictly positive. In this case, an optimal continuation strategy for the agent, given the principal's strategy, should induce a continuation outcome of the following form. There should be some $t' \geq t$ (possibly $+\infty$) so that effort is $e_s = \tilde{e}_s$ for all $s \in \{t, t+1, \ldots, t'-1\}$, and so that effort is $e_s = 0$ for $s \geq t'$. Consumption should be specified optimally. Because of our restriction on the agent's strategies, the continuation consumption outcome $(c_s)_{s \geq t}$ must satisfy

$$\sum_{s=t}^{\infty} \delta^{s-t} c_s \leq b\left(h_t^A\right) + \sum_{\tau=t}^{t'-1} \delta^{\tau-t} \tilde{w}_\tau$$

Given the concavity of v, the highest continuation payoff the agent can achieve at date t, given that he works obediently until date $t'-1$, is that obtained by putting

$$c_s = (1-\delta)\left(b\left(h_t^A\right) + \sum_{\tau=t}^{t'-1} \delta^{\tau-t} \tilde{w}_\tau\right)$$

for all $s \geq t$. (Note that, when this expression is negative, the agent's payoff is necessarily $-\infty$, and hence the specified consumption is optimal.)

[13] Note here that, if $b_t \leq 0$, the agent's continuation payoff is $-\infty$ independently of his feasible continuation strategy, and therefore any feasible continuation play is optimal. If $b_t > 0$, then it is optimal to perfectly smooth consumption, setting it to $b_t(1-\delta)$ in each period.

We can now consider the problem of choosing the optimal "public deviation" time t', given optimal consumption as specified above. The existence of a solution to this problem follows from "continuity at infinity" of the agent's payoff in the public deviation date t'; i.e., because, for all t, all histories h_t^A for which there is, as yet, no public deviation ($e_s = \tilde{e}_s$ and $w_s = \tilde{w}_s$ for all $s < t$),

$$\frac{v\left((1-\delta)\left(b\left(h_t^A\right)+\sum_{\tau=t}^{t'-1}\delta^{\tau-t}\tilde{w}_\tau\right)\right)}{1-\delta}-\sum_{\tau=t}^{t'-1}\delta^{\tau-t}\psi\left(\tilde{e}_\tau\right)$$
$$\longrightarrow \frac{v\left((1-\delta)\left(b\left(h_t^A\right)+\sum_{\tau=t}^{\infty}\delta^{\tau-t}\tilde{w}_\tau\right)\right)}{1-\delta}-\sum_{\tau=t}^{\infty}\delta^{\tau-t}\psi\left(\tilde{e}_\tau\right)$$

as $t' \to \infty$. This follows by continuity of v and because $\sum_{\tau=t}^{t'-1}\delta^{\tau-t}\tilde{w}_\tau$ is convergent to some finite value. In determining the agent's continuation strategy at date t and private history h_t^A, we take t' to be the largest value that attains the optimal payoff for the agent (again, continuity at infinity implies that such a largest value exists, and it could be $+\infty$). Hence, the strategy specifies that, at private history h_t^A for the agent, effort is $e_t = \tilde{e}_t$ if $t' > t$ and $e_t = 0$ if $t' = t$, and consumption is $c_t = (1-\delta)\left(b\left(h_t^A\right)+\sum_{\tau=t}^{t'-1}\delta^{\tau-t}\tilde{w}_\tau\right)$. Note that, given the definition of B and that expression (15) is positive, we have $b\left(h_t^A\right) - c_t \geq -B$.

Finally, let us specify the principal's beliefs at date t about the agent's previous consumption choices. Let these beliefs be degenerate sequences, and denote the believed consumption up to date t by $(\hat{c}_s)_{s=1}^t$. If date t is such that there has not been a public deviation (i.e., $(e_s, w_s) = (\tilde{e}_s, \tilde{w}_s)$ for all $s \leq t-1$ and if $e_t = \tilde{e}_t$), then the principal believes that the agent has consumed as the strategy specifies; that is, $\hat{c}_s = \tilde{c}_s$ at each date $s \leq t$. If instead $(e_s, w_s) \neq (\tilde{e}_s, \tilde{w}_s)$ for some $s \leq t-1$, or if $e_t \neq \tilde{e}_t$, let $t' \leq t$ be the first date s at which $(e_s, w_s) \neq (\tilde{e}_s, \tilde{w}_s)$, or if there is no such date, let $t' = t$. If $e_{t'} \neq \tilde{e}_{t'}$ (so the agent is first to publicly deviate), we let the principal's belief be given by

$$\hat{c}_s = (1-\delta)\left(b_1 + \sum_{\tau=1}^{t'-1}\delta^{\tau-1}\tilde{w}_\tau\right)$$

for all $s \in \{1, \ldots, t'-1\}$ (i.e., the principal believes that the agent optimally consumed from time 1 given the deviation and the principal's specified strategy), while, for all $s \in \{t', \ldots, t\}$, the principal believes that the agent consumes

$$\hat{c}_s = C\left(\hat{b}_s\right),$$

where $\hat{b}_s$ are beliefs on the agent's balance determined recursively from the principal's payments and the agent's believed consumption (i.e., $\hat{b}_s = (\hat{b}_{s-1} + w_{s-1} - \hat{c}_{s-1})/\delta$, with $\hat{b}_1 = b_1$).

If $e_{t'} = \tilde{e}_{t'}$ (so the principal is first to publicly deviate), then the principal believes that the agent consumes $\hat{c}_s = \tilde{c}_s$ for all $s \leq t'$ and $\hat{c}_s = C\left(\hat{b}_s\right)$ for all $s = t'+1, ..., t$ (again, the values of $\hat{b}_s$ are determined recursively by $\hat{b}_s = (\hat{b}_{s-1} + w_{s-1} - \hat{c}_{s-1})/\delta$, with $\hat{b}_1 = b_1$). These beliefs are consistent with updating of the principal's prior beliefs according to the specified strategy of the agent whenever there is no public evidence the agent's strategy has not been followed.

We now want to verify the sequential optimality of the above strategies, given the beliefs. First, note that at any information set at which the principal has not yet observed a deviation, the fact that $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ satisfies Condition (PC_t) implies that the principal optimally sets $w_t = \tilde{w}_t$ (if $w_t \neq \tilde{w}_t$, then the principal's continuation profits are no greater than zero). If instead the principal has observed a deviation, then the principal can obtain at most zero, since the agent exerts no effort, and hence paying $w_t = 0$ is optimal.

Next, at any date-t history, the agent's continuation strategy maximizes his payoff given the principal's strategy and the requirement that the sequence of consumption generated, namely $(c_s)_{s\geq t}$, satisfies

$$\sum_{s=t}^{\infty} \delta^{s-t} c_s \leq b_t + \sum_{s=t}^{\infty} \delta^{s-t} w_s, \tag{16}$$

where b_t is the date-t balance and $(w_s)_{s\geq t}$ is the induced payment stream. Then, for instance, at date 1, the agent prefers to follow the specified strategy and obtain discounted payoff

$$\frac{1}{1-\delta} v(\bar{c}_\infty) - \sum_{s=t}^{\infty} \delta^{s-t} \psi(\tilde{e}_s)$$

rather than shirk for the first time at date t and choose the optimal consumption path given (16), i.e. consume $\bar{c}_{t-1}$ in each period. The payoff obtained by the latter strategy is

$$\frac{1}{1-\delta} v(\bar{c}_{t-1}) - \sum_{s=1}^{t-1} \delta^{s-1} \psi(\tilde{e}_s).$$

At dates when some deviation has occurred, we specified the agent's continuation play to be sequentially optimal, again given that the constraint (16) must be satisfied. □

Proof of Lemma 4.1

Proof. Fix an optimal relational contract $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$; that is, a contract that maximizes the principal's discounted payoff subject to the conditions of Proposition 4.1. We show first that

$$\frac{v\left(\bar{c}_\infty\right)}{1-\delta} - \sum_{s=1}^{\infty} \delta^{s-1} \psi(\tilde{e}_s) \tag{17}$$

is equal to $\frac{v(b_1(1-\delta))}{1-\delta}$. Clearly the only way this can fail in a self-enforceable relational contract is if (17) strictly exceeds $\frac{v(b_1(1-\delta))}{1-\delta}$. However, in this case, there is a more profitable contract for which the conditions of Proposition 4.1 still hold, and in which $\tilde{e}_1$ increases by a small amount. In particular, the inequalities (AC_t^{un}) are unchanged at all periods except the first; the inequality that relates to the first period continues to hold because the change in $\tilde{e}_1$ is small. The inequalities (PC_t) are unaffected.

Next, the previous observation implies that, if

$$\frac{v\left(\bar{c}_{t-1}\right)}{1-\delta} - \sum_{s=1}^{t-1} \delta^{s-1} \psi(\tilde{e}_s) \tag{18}$$

exceeds $\frac{v(b_1(1-\delta))}{1-\delta}$ at any date t, then the inequality (AC_t^{un}) must not be satisfied; i.e., the conditions of Proposition 4.1 are not satisfied (there can be no PBE with contractual outcomes $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$).

Finally, suppose that the expression (18) is strictly less than $\frac{v(b_1(1-\delta))}{1-\delta}$ at some increasing sequence of dates $(t_n)_{n=1}^{N}$, where N may be finite or infinite. For each n, there is $\varepsilon_n > 0$ such that

$$\frac{1}{1-\delta} v(\bar{c}_{t_n-1} + \delta^{t_n-2}\varepsilon_n(1-\delta)) - \sum_{s=1}^{t_n-1} \delta^{s-1}\psi(\tilde{e}_s) = \frac{v\left(b_1\left(1-\delta\right)\right)}{1-\delta}.$$

Increase w_{t_n-1} by ε_n, and reduce w_{t_n} by $\frac{\varepsilon_n}{\delta}$; note that this leads to a change in $\bar{c}_{t_n-1}$, but does not affect $\bar{c}_t$ for $t \neq t_n$. After this adjustment has been made for each n, we have a relational contract for which the expression (18) is equal to $\frac{v(b_1(1-\delta))}{1-\delta}$ at all dates t. Also, because ψ is non-negative, $\bar{c}_t$ must be a non-decreasing sequence, and hence all payments $\tilde{w}_t$ in the new relational contract are non-negative. Hence the "agent's constraints" (AC_t^{un}) are satisfied. Also, the principal's constraints (PC_t) are satisfied. To see the latter, note that these constraints are affected by the adjustments to the original contract only at dates satisfying $t = t_n$ for some n. At such dates the principal's constraint is *slackened* by the amount $\frac{\varepsilon_n}{\delta}$. □

Proof of Propositions 4.2 and 4.3

Proof. The remaining steps in the proof of the Proposition 4.3 are divided into seven lemmas. The proof of Proposition 4.2 is provided in the process, in Lemma A.7. Throughout, we restrict attention to payments determined under the restriction to "fastest payments", i.e. satisfying Condition (FP_t^{un}).

1. Lemma A.1 bounds effort and hence payments.

2. Lemma A.2 shows that an optimal relational contract exists.

3. Lemma POSITIVE shows that the principal achieves a strictly positive payoff and that effort remains strictly positive in any optimal relational contract.

4. Lemma A.4 shows that the contract becomes stationary in the long run.

5. Lemma A.5 shows that effort is weakly decreasing.

6. Lemma A.6 establishes that, if the principal's constraint (PC_t) binds at date t, then it continues to bind at all future dates. Also effort strictly decreases over these dates.

7. Lemma A.7 establishes the condition for a first-best policy to be self-enforceable, and that when this condition is not satisfied there exists a date $\bar{t}$ satisfying the properties in the proposition (i.e., effort is constant up to date $\bar{t}$, and subsequently strictly decreasing).

8. Lemma A.8 establishes that the date $\bar{t}$ can be strictly greater than one.

The following lemma argues that we can restrict attention to contracts such that the marginal disutility of effort is bounded by the marginal utility of consumption: $\psi'(\tilde{e}_t) \leq v'(\bar{c}_\infty) \leq v'(\bar{c}_t)$ (the inequality $v'(\bar{c}_\infty) \leq v'(\bar{c}_t)$ comes from the fact that $\bar{c}_t$ is increasing in time, so $\bar{c}_t \leq \bar{c}_\infty$).

Lemma A.1. *There is no loss of optimality in restricting to self-enforceable contracts such that $\psi'(\tilde{e}_t) \leq v'(\bar{c}_\infty)$ for all t. Additionally, let z be the inverse of ψ' and let $\hat{c}$ be the highest value of c such that*

$$v(c) - \psi(z(v'(c))) \leq v(b_1(1-\delta)),$$

which exists because ψ and z are increasing, v' is decreasing and v is onto all of $\mathbb{R}$. Then, if a self-enforceable relational contract satisfies $\psi'(\tilde{e}_t) \leq v'(\bar{c}_\infty)$ for all t, we have $\bar{c}_\infty \leq \hat{c}$.

Proof. Take a contract satisfying Condition (FP_t^{un}) for all t, and let t^* be the first date at which $\psi'(\tilde{e}_{t^*}) > v'(\bar{c}_\infty)$. We can adjust such a contract by reducing date t^* effort by $\eta > 0$ (holding effort at other dates fixed). This determines a new contract, with adjusted consumption and payments, again satisfying (FP_t^{un}) for all t. Let us index the revised effort policy by the date-t^* adjustment η, writing $\tilde{e}_t(\eta)$ for all t. Correspondingly, write

$$\bar{c}_\infty(\eta) \equiv (1-\delta)\left(b_1 + \sum_{s=1}^{\infty} \delta^{s-1} \tilde{w}_s(\eta)\right)$$

where $(\tilde{w}_s(\eta))_{s \geq 1}$ are the payments determined from the adjusted effort policy. Then

$$\frac{v(\bar{c}_\infty(0)) - v(\bar{c}_\infty(\eta))}{1-\delta} = \delta^{t^*-1}\big(\psi(\tilde{e}_{t^*}(0)) - \psi(\tilde{e}_{t^*}(\eta))\big).$$

Differentiating with respect to η,

$$\frac{\bar{c}'_{\infty}(\eta)}{1-\delta} = \frac{-\delta^{t^*-1}\psi'(\tilde{e}_{t^*}(\eta))}{v'(\bar{c}_{\infty}(\eta))}.$$

This expression coincides with the derivative of the NPV of payments to the agent with respect to η. The derivative of the principal's profits is therefore

$$-\delta^{t^*-1} + \frac{\delta^{t^*-1}\psi'(\tilde{e}_{t^*}(\eta))}{v'(\bar{c}_{\infty}(\eta))}$$

which is strictly positive for $\eta \in [0, \bar{\eta})$, with $\bar{\eta}$ satisfying $\psi'(\tilde{e}_{t^*}(\bar{\eta})) = v'(\bar{c}_{\infty}(\bar{\eta}))$. The effect on profit from reducing date t^* effort by $\bar{\eta}$ is therefore to increase it by

$$\int_0^{\bar{\eta}} \left(-\delta^{t^*-1} + \frac{\delta^{t^*-1}\psi'(\tilde{e}_{t^*}(\eta))}{v'(\bar{c}_{\infty}(\eta))} \right) d\eta > 0.$$

Note (from (FP_t^{un})) that payments $\tilde{w}_t(\bar{\eta})$ are reduced for all $t \geq t^*$, with the implication that the principal's constraints (PC_t) are relaxed. Hence, the new contract is self-enforceable. Note also that $v'(\bar{c}_{\infty}(\bar{\eta})) > \psi'(\tilde{e}_t(\bar{\eta}))$ for all $t < t^*$. We can therefore continue iteratively, by proceeding to the next date $t > t^*$ at which $v'(\bar{c}_{\infty}(\bar{\eta})) < \psi'(\tilde{e}_t(\bar{\eta}))$, if any, and reducing effort precisely as for at t^*. Proceeding sequentially, we obtain a self-enforceable contract for which $\psi'(\tilde{e}_t) \leq v'(\bar{c}_{\infty})$ at all dates t, and which is strictly more profitable than the original.

The above establishes that we can restrict attention to effort policies such that $\tilde{e}_t \leq z(v'(\bar{c}_{\infty}))$ (with $\bar{c}_{\infty}$ determined by the effort policy). Because (FP_t^{un}) holds at $t = \infty$, we must then have that

$$v(\bar{c}_{\infty}) - \psi(z(v'(\bar{c}_{\infty}))) \leq v(b_1(1-\delta)),$$

which implies $\bar{c}_{\infty} \leq \hat{c}$ as required. □

The above result establishes that the marginal disutility of effort $\psi'(\tilde{e}_t)$ in an optimal contract is bounded by $v'(\bar{c}_{\infty})$, which is certainly no greater than $v'(b_1(1-\delta))$, given feasibility dictates payments $\tilde{w}_t$ are non-negative. Also, we established that $\bar{c}_{\infty} \leq \hat{c}$, and so $\bar{c}_{t-1} \leq \hat{c}$ for all t as well. Therefore, given an effort policy $(\tilde{e}_t)_{t\geq 1}$, we can recover the payment policy $(\tilde{w}_t)_{t\geq 1}$ from (FP_t^{un}), and these payments must be bounded by some $\bar{w} > 0$ (see Condition (5)).

Lemma A.2. *An optimal relational contract exists.*

Proof. As we already observed, under the condition "fastest payments" given in ($\mathrm{FP}_t^{\mathrm{un}}$), the relational contract is determined solely by the effort policy $(\tilde{e}_t)_{t\geq 1}$. Hence, the payoff obtained by the principal can be written

$$W\left((\tilde{e}_t)_{t=1}^{\infty}\right) = \sum_{t=1}^{\infty} \delta^{t-1}\left(\tilde{e}_t - \tilde{w}_t\right)$$

where $\sum_{t=1}^{\infty} \delta^{t-1}\tilde{w}_t$ satisfies

$$\frac{v\left((1-\delta)\, b_1 + (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1}\tilde{w}_t\right)}{1-\delta} = \sum_{t=1}^{\infty} \delta^{t-1}\psi(\tilde{e}_t) + \frac{v\left((1-\delta)\, b_1\right)}{1-\delta}. \tag{19}$$

Note that, from Lemma A.1, we can restrict attention to effort policies in $[0, z\left(v'\left(b_1\left(1-\delta\right)\right)\right)]^{\infty}$, and that $\sum_{t=1}^{\infty} \delta^{t-1}\tilde{w}_t$ remains bounded below $\frac{\hat{c}}{1-\delta} - b_1$ over such policies (where $\hat{c}$ is defined in Lemma A.1).

Now, let $W^{\sup}$ be the supremum of $W(\cdot)$ over effort policies $(\tilde{e}_t)_{t\geq 1}$ in $[0, z\left(v'\left(b_1\left(1-\delta\right)\right)\right)]$ for which the implied contract $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ (i.e., the one implied by ($\mathrm{FP}_t^{\mathrm{un}}$)) satisfies the principal's constraints (PC_t) (such contracts thus satisfy all the conditions of Proposition 4.1). Consider then a sequence of policies $((\tilde{e}_t^n)_{t=1}^{\infty})_{n=1}^{\infty}$ with values in $[0, z\left(v'\left(b_1\left(1-\delta\right)\right)\right)]$ and with

$$W\left((\tilde{e}_t^n)_{t=1}^{\infty}\right) > W^{\sup} - 1/n$$

for all n, and for which the contract defined by each effort policy (using ($\mathrm{FP}_t^{\mathrm{un}}$)) satisfies the principal's constraints (PC_t). There then exists a sequence $(\tilde{e}_t^{\infty})_{t\geq 1} \in [0, z\left(v'\left(b_1\left(1-\delta\right)\right)\right)]^{\infty}$ and a subsequence $\left((\tilde{e}_t^{n_k})_{t\geq 1}\right)_{k\geq 1}$ convergent pointwise to $(\tilde{e}_t^{\infty})_{t\geq 1}$. Since

$$\sum_{t=1}^{\infty} \delta^{t-1}\psi\left(\tilde{e}_t^{n_k}\right) \to \sum_{t=1}^{\infty} \delta^{t-1}\psi\left(\tilde{e}_t^{\infty}\right)$$

as $k \to \infty$ (by continuity of ψ and discounting) we have (by continuity of v and Equation (19))

$$\sum_{t=1}^{\infty} \delta^{t-1}\tilde{w}_t^{n_k} \to \sum_{t=1}^{\infty} \delta^{t-1}\tilde{w}_t^{\infty}$$

where $\tilde{w}_t^{n_k}$ and $\tilde{w}_t^{\infty}$ denote the corresponding payments derived through ($\mathrm{FP}_t^{\mathrm{un}}$). Hence, $W\left((\tilde{e}_t^{\infty})_{t\geq 1}\right) = W^{\sup}$.

Our result will then follow if we can show that the contract defined by $(\tilde{e}_t^{\infty})_{t\geq 1}$ satisfies the principal's constraints (PC_t). Suppose with a view to contradiction that there is some t^* at which the principal's constraint does not hold, and so

$$\tilde{w}_{t^*}^{\infty} > \sum_{s=t^*+1}^{\infty} \delta^{s-t^*}\left(\tilde{e}_s^{\infty} - \tilde{w}_s^{\infty}\right).$$

It is easily verified from (FP_t^{un}) and the pointwise convergence of $(\tilde{e}_t^{n_k})_{t\geq 1}$ to $(\tilde{e}_t^{\infty})_{t\geq 1}$ that $\tilde{w}_t^{n_k} \to \tilde{w}_t^{\infty}$ for each t. Therefore, for large enough k,

$$\tilde{w}_{t^*}^{n_k} > \sum_{s=t^*+1}^{\infty} \delta^{s-t^{n_k}} \left(\tilde{e}_s^{n_k} - \tilde{w}_s^{n_k}\right)$$

contradicting that the contract determined by $(\tilde{e}_t^{n_k})_{t\geq 1}$ satisfies the principal's constraints (PC_t). □

We now establish the following regarding the non-degeneracy of optimal contracts.

Lemma A.3. *The principal obtains a strictly positive payoff in any optimal contract* $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$. *Moreover, effort* $\tilde{e}_t$ *is strictly positive at all dates.*

Proof. Consider effort set constant to some level $\tilde{e} > 0$. Let $z(\tilde{e}) = \sum_{t=1}^{\infty} \delta^{t-1} \tilde{w}_t$ be the NPV of payments that must be made to the agent when satisfying the indifference conditions (FP_t^{un}). Recall that this satisfies

$$\frac{v\left((1-\delta) b_1 + (1-\delta) z(\tilde{e})\right)}{1-\delta} = \frac{\psi(\tilde{e})}{1-\delta} + \frac{v\left((1-\delta) b_1\right)}{1-\delta}.$$

Differentiating with respect to $\tilde{e}$ yields

$$z'(\tilde{e}) = \frac{\psi'(\tilde{e})}{(1-\delta)\, v'\left((1-\delta) b_1 + (1-\delta) z(\tilde{e})\right)}$$

showing that the principal's payoff is strictly positive for small positive $\tilde{e}$. Moreover, payments determined by the conditions (FP_t^{un}) rise over time approaching

$$\frac{\psi(\tilde{e})}{v'\left((1-\delta) b_1 + (1-\delta) z(\tilde{e})\right)}$$

which is $o(\tilde{e})$ (i.e., vanishes much faster than $\tilde{e}$). It is then easy to see that, when $\tilde{e}$ is small enough, all the principal constraints (PC_t) are satisfied. Hence the contract determined from specifying constant effort $\tilde{e}$, for small $\tilde{e}$, is self-enforceable and generates a strictly positive payoff.

Now we show that, in an optimal contract, effort is strictly positive in every period. Suppose without loss of generality that payments are determined from effort using (FP_t^{un}). First note that the principal's continuation profits

$$\sum_{s=t}^{\infty} \delta^{s-t} \left(\tilde{e}_s - \tilde{w}_s\right)$$

must be strictly positive at all dates. Otherwise, if this expression is zero at some date t, then $\tilde{w}_{t-1} = 0$. Condition (FP_t^{un}) then implies that $\tilde{e}_{t-1} = 0$. Iterating backwards, we find the optimal profit is zero in contradiction with the previous claim. Suppose then that effort is zero at some date, and consider a date t such that effort is zero at this date but strictly positive at the subsequent date. Then $\tilde{w}_t = 0$ and so the principal's constraint (PC_t) is slack at date t. However, this contradicts Lemma 4.2.

□

We now establish an important property of relational contracts: they become (approximately) stationary in the long run.

Lemma A.4. *Suppose that $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ is an optimal relational contract. Then, there exists an effort/payment pair $(\tilde{e}_\infty, \tilde{w}_\infty)$ such that $\lim_{t\to\infty} (\tilde{e}_t, \tilde{w}_t) = (\tilde{e}_\infty, \tilde{w}_\infty)$.*

Proof. **Step 0.** We first prove that, if $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ is an optimal relational contract satisfying (FP_t^{un}), then

$$\lim_{t\to\infty} \left(\tilde{w}_t - \frac{\psi(\tilde{e}_t)}{v'(\bar{c}_\infty)} \right) = 0.$$

Condition (FP_t^{un}) implies that, for any $t \geq 1$,

$$\frac{v(\bar{c}_{t-1} + (1-\delta)\,\delta^{t-1}\tilde{w}_t) - v(\bar{c}_{t-1})}{1-\delta} = \delta^{t-1}\psi(\tilde{e}_t).$$

(Recall that $\bar{c}_0 = (1-\delta)b_1$.) Recall from Lemma A.1 that effort and hence payments remain bounded. Thus, as $t \to \infty$, $\tilde{w}_t \delta^{t-1} \to 0$, and

$$v'(\bar{c}_\infty)\,\tilde{w}_t\delta^{t-1} + o\left(\tilde{w}_t\delta^{t-1}\right) = \delta^{t-1}\psi(\tilde{e}_t),$$

which proves the result.

Step 1. Define $\bar{e} \equiv \limsup_{t\to\infty} \tilde{e}_t$, which we know from Lemma A.1 is no greater than $z(v'(\bar{c}_\infty))$ (recall that z is the inverse of ψ'). We now show that, for any $e \in [0, \bar{e}]$,

$$\frac{\psi(e)}{v'(\bar{c}_\infty)} \leq \frac{\delta}{1-\delta}\left(e - \frac{\psi(e)}{v'(\bar{c}_\infty)}\right). \tag{20}$$

That is, the principal's constraints (PC_t) would be satisfied if paying a constant wage $\frac{\psi(e)}{v'(\bar{c}_\infty)}$ per period, in return for effort $e \leq \bar{e}$. Note that, if the inequality (20) is satisfied at $\bar{e}$, then it is satisfied for all $e \in [0, \bar{e}]$; this follows because the left-hand side of (20) is convex, and equal to zero at zero, while the right hand side is concave, and also equal to zero at zero.

Assume now for the sake of contradiction that the inequality (20) is not satisfied for some $e \in [0, \bar{e}]$. Then we must have

$$\frac{\psi(\bar{e})}{v'(\bar{c}_\infty)} > \frac{\delta}{1-\delta}\left(\bar{e} - \frac{\psi(\bar{e})}{v'(\bar{c}_\infty)}\right). \tag{21}$$

Observe then that

$$\tilde{e}_t - \tilde{w}_t \leq \bar{e} - \frac{\psi(\bar{e})}{v'(\bar{c}_\infty)} + \varepsilon_t$$

for some sequence $(\varepsilon_t)_{t=1}^{\infty}$ convergent to zero. This follows because $\tilde{w}_t - \frac{\psi(\tilde{e}_t)}{v'(\bar{c}_\infty)} \to 0$ as $t \to \infty$ (by Step 0), because $e - \frac{\psi(e)}{v'(\bar{c}_\infty)}$ increases over effort levels in $[0, \bar{e}]$ (since $\psi'(\bar{e}) \leq v'(\bar{c}_\infty)$), and by definition of $\bar{e}$ as $\limsup_{t\to\infty} \tilde{e}_t$.
We therefore have that

$$\limsup_{t\to\infty} \sum_{s=t+1}^{\infty} \delta^{s-t}(\tilde{e}_s - \tilde{w}_s) \leq \frac{\delta}{1-\delta}\left(\bar{e} - \frac{\psi(\bar{e})}{v'(\bar{c}_\infty)}\right) < \frac{\psi(\bar{e})}{v'(\bar{c}_\infty)},$$

where the last inequality holds by equation (21). However, Step 0 implies that the superior limit of payments to the agent must be $\frac{\psi(\bar{e})}{v'(\bar{c}_\infty)}$, which means that the principal's constraint (PC_t) is not satisfied at some time. This contradicts the definition of $\bar{e}$ as $\limsup_{t\to\infty} \tilde{e}_t$ (with $(\tilde{e}_t)_{t\geq 1}$ the effort profile in a self-enforceable relational contract).

Step 2. We complete the proof by showing that $\liminf_{t\to\infty} \tilde{e}_t = \bar{e}$. Assume, for the sake of contradiction, that $\liminf_{t\to\infty} \tilde{e}_t < \bar{e}$. In this case, there exists some $t' > 1$ such that $\tilde{e}_{t'} < \min\{\bar{e}, \tilde{e}_{t'+1}\}$.

Step 2a. We have

$$\tilde{w}_{t'} \leq \frac{\delta}{1-\delta}\left(\tilde{e}_{t'+1} - \frac{\psi(\tilde{e}_{t'+1})}{v'(\bar{c}_\infty)}\right). \tag{22}$$

This follows because (i) $\tilde{w}_{t'} \leq \frac{\psi(\tilde{e}_{t'})}{v'(\bar{c}_\infty)}$ by assumption that payments satisfy condition (FP_t^{un});[14] (ii) $\frac{\psi(\tilde{e}_{t'})}{v'(\bar{c}_\infty)} \leq \frac{\delta}{1-\delta}\left(\tilde{e}_{t'} - \frac{\psi(\tilde{e}_{t'})}{v'(\bar{c}_\infty)}\right)$, by assumption that $\tilde{e}_{t'} < \bar{e}$ and by Step 1, and (iii) $\tilde{e}_{t'} - \frac{\psi(\tilde{e}_{t'})}{v'(\bar{c}_\infty)} \leq e_{t'+1} - \frac{\psi(\tilde{e}_{t'+1})}{v'(\bar{c}_\infty)}$ because $z(v'(\bar{c}_\infty)) \geq \tilde{e}_{t'+1} > \tilde{e}_{t'}$ with the first inequality following from Lemma A.1.

[14] This follows from

$$\psi(\tilde{e}_{t'}) = \frac{v(\bar{c}_{t'}) - v(\bar{c}_{t'} - (1-\delta)\delta^{t'-1}\tilde{w}_{t'})}{(1-\delta)\delta^{t'-1}} \geq \tilde{w}_{t'} v'(\bar{c}_{t'}) \geq \tilde{w}_{t'} v'(\bar{c}_\infty).$$

Intuitively, the payment $\tilde{w}_{t'}$ makes the agent indifferent between working at date t' (and collecting $\tilde{w}_{t'}$ and then quitting), and instead quitting at $t'-1$, saving on the disutility of effort $\psi(\tilde{e}_{t'})$; the payment $\tilde{w}_{t'}$ required for this indifference is less than $\frac{\psi(\tilde{e}_{t'})}{v'(\bar{c}_\infty)}$ because the agent's marginal utility of money associated with the payment $\tilde{w}_{t'}$, conditional on quitting the relationship after t', is higher than $v'(\bar{c}_{t'})$.

Step 2b. We now show that the principal's constraint (PC_t) is slack at t'. Note first that, for any $t \geq 1$, we have

$$\begin{aligned}
\tilde{w}_{t+1} - \tilde{w}_t &= \frac{\bar{c}_{t+1} - \bar{c}_t}{\delta^t(1-\delta)} - \frac{\bar{c}_t - \bar{c}_{t-1}}{\delta^{t-1}(1-\delta)} \\
&\geq \frac{v(\bar{c}_{t+1}) - v(\bar{c}_t)}{\delta^t(1-\delta)v'(\bar{c}_t)} - \frac{v(\bar{c}_t) - v(\bar{c}_{t-1})}{\delta^{t-1}(1-\delta)v'(\bar{c}_t)} \\
&= \frac{\psi(\tilde{e}_{t+1}) - \psi(\tilde{e}_t)}{v'(\bar{c}_t)},
\end{aligned}$$

where we used that v is concave. Hence, we have that $\tilde{e}_{t+1} > \tilde{e}_t$ implies $\tilde{w}_{t+1} > \tilde{w}_t$.

Since t' was chosen so that $\tilde{e}_{t'+1} > \tilde{e}_{t'}$, we have $\tilde{w}_{t'+1} > \tilde{w}_{t'}$. Hence,

$$\begin{aligned}
\tilde{w}_{t'} &< (1-\delta)\,\tilde{w}_{t'} + \delta \tilde{w}_{t'+1} \\
&\leq \delta \left(\tilde{e}_{t'+1} - \frac{\psi\left(\tilde{e}_{t'+1}\right)}{v'\left(\bar{c}_\infty\right)} \right) + \delta \sum_{s=t'+2}^{\infty} \delta^{s-t'-1} \left(\tilde{e}_s - \tilde{w}_s\right) \\
&\leq \sum_{s=t'+1}^{\infty} \delta^{s-t'} \left(\tilde{e}_s - \tilde{w}_s\right),
\end{aligned}$$

where the second inequality uses (i) Equation (22) from Step 2a, and (ii) the principal's constraint (PC_t) in period $t'+1$. The third inequality uses that $\tilde{w}_{t'+1} \leq \frac{\psi(\tilde{e}_{t'+1})}{v'(\bar{c}_\infty)}$, which again follows from Equation (FP_t^{un}).

Step 2c. We finish the proof with the following observation. The fact the principal's constraint (PC_t) is slack at time t' (proven in Step 2b) contradicts Lemma 4.2, since effort is strictly higher at $t'+1$ than at t'. □

The following lemma determines that, in an optimal contract, effort is weakly decreasing.

Lemma A.5. *In an optimal contract, the effort policy $(\tilde{e}_t)_{t\geq 1}$ is a weakly decreasing sequence. Therefore, for all t, $\tilde{e}_t \geq \tilde{e}_\infty \equiv \lim_{s\to\infty} \tilde{e}_s$.*

Proof. By Lemma A.4, $(\tilde{e}_t)_{t=1}^{\infty}$ is a convergent sequence, so using the notation in its proof, we have $\tilde{e}_\infty = \bar{e}$. Step 2 in the proof of Lemma A.4 proves that there is no time t' such that $\tilde{e}_{t'} < \min\{\bar{e}, \tilde{e}_{t'+1}\}$. Hence, there is no t' such that $\tilde{e}_{t'} < \tilde{e}_\infty$.

Now suppose, for the sake of contradiction, that $(\tilde{e}_t)_{t=1}^{\infty}$ is not a weakly decreasing sequence. Thus, there exists a date t' where $\max_{t>t'} \tilde{e}_t > \tilde{e}_{t'}$ (the maximum exists by the first part of this proof, as well as the existence of $\lim_{t\to\infty} \tilde{e}_t = \tilde{e}_\infty$ by Lemma A.4). Let $t^*(t')$ be the smallest value $t > t'$ where the maximum is attained, that is, $\tilde{e}_{t^*(t')} = \max_{t>t'} \tilde{e}_t$.

Because payments satisfy (FP_t^{un}), we have for all t, that (5) holds. Thus, for any $s > t^*(t')$,

$$\tilde{e}_{t^*(t')} - \tilde{w}_{t^*(t')} > \tilde{e}_{t^*(t')} - \frac{\psi(\tilde{e}_{t^*(t')})}{v'(\bar{c}_{t^*(t')})} \geq \tilde{e}_{t^*(t')} - \frac{\psi(\tilde{e}_{t^*(t')})}{v'(\bar{c}_{s-1})} \geq \tilde{e}_s - \frac{\psi(\tilde{e}_s)}{v'(\bar{c}_{s-1})} > \tilde{e}_s - \tilde{w}_s. \tag{23}$$

The first inequality follows from Equation (5) and because $\tilde{e}_{t^*(t')}$ is strictly positive (by Lemma A.3); the second inequality follows because $\bar{c}_{s-1} \geq \bar{c}_{t^*(t')}$. The third inequality follows because $e - \frac{\psi(e)}{v'(\bar{c}_{s-1})}$ is increasing in e over $[0, z(v'(\bar{c}_\infty))]$, and because $\tilde{e}_s \leq \tilde{e}_{t^*(t')}$ for $s > t^*(t')$ by definition of $t^*(t')$. The fourth inequality follows because $\tilde{w}_s > \frac{\psi(\tilde{e}_s)}{v'(\bar{c}_{s-1})}$ by Equation (5), and because $\tilde{e}_s$ is strictly positive (again by Lemma A.3).

Equation (23) implies that

$$\tilde{e}_{t^*(t')} - \tilde{w}_{t^*(t')} > (1-\delta) \sum_{s=t^*(t')+1}^{\infty} \delta^{s-t^*(t')-1} (\tilde{e}_s - \tilde{w}_s),$$

so that

$$\begin{aligned}
\sum_{s=t^*(t')}^{\infty} \delta^{s-t^*(t')} (\tilde{e}_s - \tilde{w}_s) &= \tilde{e}_{t^*(t')} - \tilde{w}_{t^*(t')} + \delta \sum_{s=t^*(t')+1}^{\infty} \delta^{s-t^*(t')-1} (\tilde{e}_s - \tilde{w}_s) \\
&> (1-\delta) \sum_{s=t^*(t')+1}^{\infty} \delta^{s-t^*(t')-1} (\tilde{e}_s - \tilde{w}_s) + \delta \sum_{s=t^*(t')+1}^{\infty} \delta^{s-t^*(t')-1} (\tilde{e}_s - \tilde{w}_s) \\
&= \sum_{s=t^*(t')+1}^{\infty} \delta^{s-t^*(t')-1} (\tilde{e}_s - \tilde{w}_s).
\end{aligned} \tag{24}$$

Recall from Lemma 4.2 that the principal's constraint must bind at $t^*(t') - 1$ (since $\tilde{e}_{t^*(t')} > \tilde{e}_{t^*(t')-1}$ by the definition of $t^*(t')$). The inequality (24), then implies that $\tilde{w}_{t^*(t')-1} > \tilde{w}_{t^*(t')}$. But then, recalling Equation (5), we have

$$\frac{\psi\left(\tilde{e}_{t^*(t')-1}\right)}{v'\left(\bar{c}_{t^*(t')-1}\right)} > \tilde{w}_{t^*(t')-1} > \tilde{w}_{t^*(t')} > \frac{\psi(\tilde{e}_{t^*(t')})}{v'\left(\bar{c}_{t^*(t')-1}\right)}.$$

Hence, $\tilde{e}_{t^*(t')-1} > \tilde{e}_{t^*(t')}$, contradicting the definition of $t^*(t')$. □

Having shown that the effort is weakly decreasing in an optimal relational contract (Lemma A.5) we now show that, in fact, it is strictly decreasing when the principal's constraint holds with equality.

Lemma A.6. *If the principal's constraint* (PC_t) *holds with equality at some date* $\bar{t}$, *then* $\tilde{e}_{\bar{t}} > \tilde{e}_{\bar{t}+1}$. *Hence, by Lemma 4.2, the principal's constraint also holds with equality at* $\bar{t}+1$.

Proof. The same arguments we used in Lemma A.5 to establish the inequalities in (23) imply that $\tilde{e}_{\bar{t}+1} - \tilde{w}_{\bar{t}+1} > \tilde{e}_s - \tilde{w}_s$ for all $s > \bar{t}+1$. In turn, this means that if the principal's constraint (PC$_t$) holds with equality at $\bar{t}$, then $\tilde{w}_{\bar{t}} > \tilde{w}_{\bar{t}+1}$. Indeed, because the principal's constraint holds with equality at $\bar{t}$,

$$\begin{aligned}
\tilde{w}_{\bar{t}} &= \delta\Bigg(\tilde{e}_{\bar{t}+1} - \tilde{w}_{\bar{t}+1} + \delta \sum_{s=\bar{t}+2}^{\infty} \delta^{s-\bar{t}-2}(\tilde{e}_s - \tilde{w}_s)\Bigg) \\
&> \delta\Bigg((1-\delta) \sum_{s=\bar{t}+2}^{\infty} \delta^{s-\bar{t}-2}(\tilde{e}_s - \tilde{w}_s) + \delta \sum_{s=\bar{t}+2}^{\infty} \delta^{s-\bar{t}-2}(\tilde{e}_s - \tilde{w}_s)\Bigg) \\
&= \sum_{s=\bar{t}+2}^{\infty} \delta^{s-\bar{t}-1}(\tilde{e}_s - \tilde{w}_s) \\
&\geq \tilde{w}_{\bar{t}+1}.
\end{aligned}$$

The final inequality follows from the principal's constraint (PC$_t$) at date $\bar{t}+1$. Using Equation (5), we have $\tilde{e}_{\bar{t}+1} < \tilde{e}_{\bar{t}}$. Hence, by Lemma 4.2, the principal's constraint holds with equality at $\bar{t}+1$. Therefore, by induction, the principal's constraint fails to be slack at all future dates and effort strictly decreases from $\bar{t}$ onwards. □

Lemma (A.6) implies that if the principal's constraint holds with equality at some $\bar{t}$, then effort is strictly decreasing forever after (and the principal's constraints hold with equality forever after). Recall that, for any optimal contract, Lemma 4.1 establishes that we can obtain an optimal contract in which condition (FP$_t^{\text{un}}$) holds at all dates t, and which has the same effort profile. Hence any optimal contract (whether or not condition (FP$_t^{\text{un}}$) holds – i.e., whether or not the agent is indifferent to quitting at all dates) satisfies the pattern implied by the above lemmas. In particular, we have established that either: (a) the principal's effort is always constant and the principal's constraint never holds with equality under the payment profile satisfying (FP$_t^{\text{un}}$) for all t, or (b) effort is constant up to some date, and strictly decreasing thereafter. The purpose of the following result is to establish that, when the first best cannot be sustained in a self-enforceable contract, the effort policy is necessarily the one satisfying Case (b).

Lemma A.7. *An optimal contract achieves the first-best payoff of the principal if and only if Condition (6) holds. If this condition is not satisfied, then there is a time $\bar{t} \in \mathbb{N}$ such that the principal's constraint is slack if and only if $t < \bar{t}$.*

Proof. Consider payments satisfying (FP$_t^{\text{un}}$), for all t, and determined given the first-best effort. From Proposition 3.1, the first-best effort is $e^{FB}(b_1)$. Step 0 in the proof of Lemma A.4

shows that the payments tend to $\frac{\psi(e^{FB}(b_1))}{v'(c^{FB}(b_1))}$. Furthermore, given the concavity of v, Equations (3) and (FP_t^{un}) imply that equilibrium payments to the agent increase over time. Hence, the upper limit of payments is given by $\frac{\psi(e^{FB}(b_1))}{v'(c^{FB}(b_1))}$ while the lower limit of per-period profits is given by $e^{FB}(b_1) - \frac{\psi(e^{FB}(b_1))}{v'(c^{FB}(b_1))}$, which establishes the condition for implementation of the first best in the lemma.

Consider now an optimal contract that is not first best. Lemma A.6 established that there are two possibilities. First, we might have a finite date $\bar{t} \in \mathbb{N}$, with the principal's constraint (PC_t) holding with equality at $\bar{t}$, and every subsequent date, but slack at dates $\bar{t}-1$ and earlier. In this case, effort is constant from the initial date up to $\bar{t}-1$ (by Lemma 4.2). Second, we might have that the principal's constraint (PC_t) is slack at all dates. Effort is then constant over all periods (by Lemma 4.2), but not first-best. Letting $\tilde{e}_\infty$ be the constant effort level and $\bar{c}_\infty$ equilibrium consumption, Proposition 3.1 then implies that $v'(\bar{c}_\infty) \neq \psi'(e_\infty)$. By Lemma A.1, we have $v'(\bar{c}_\infty) > \psi'(\tilde{e}_\infty)$.

Assuming that payments to the agent satisfy Condition (FP_t^{un}) for all t, we have $\tilde{w}_t$ increasing over time and converging to $\frac{\psi(\tilde{e}_\infty)}{v'(\bar{c}_\infty)}$ from below. We claim then that

$$\frac{\psi\left(\tilde{e}_\infty\right)}{v'\left(\bar{c}_\infty\right)} = \frac{\delta}{1-\delta}\left(\tilde{e}_\infty - \frac{\psi\left(\tilde{e}_\infty\right)}{v'\left(\bar{c}_\infty\right)}\right). \tag{25}$$

If instead $\frac{\psi(\tilde{e}_\infty)}{v'(\bar{c}_\infty)} > \frac{\delta}{1-\delta}\left(\tilde{e}_\infty - \frac{\psi(\tilde{e}_\infty)}{v'(\bar{c}_\infty)}\right)$, then, for large enough t we must have

$$\tilde{w}_t > \sum_{s=t+1}^{\infty} \delta^{s-t}\left(\tilde{e}_\infty - \tilde{w}_t\right),$$

so the principal's constraint is violated at t. If instead $\frac{\psi(\tilde{e}_\infty)}{v'(\bar{c}_\infty)} < \frac{\delta}{1-\delta}\left(\tilde{e}_\infty - \frac{\psi(\tilde{e}_\infty)}{v'(\bar{c}_\infty)}\right)$, we have $\tilde{w}_t$ remains bounded below $\sum_{s=t+1}^{\infty} \delta^{s-t}\left(\tilde{e}_\infty - \tilde{w}_t\right)$. Without violating ($\text{PC}_t$), effort can then be increased by a small constant amount across all periods, with payments determined via condition (FP_t^{un}). This increases profits.

Note then that Condition (25) can be written as

$$\frac{\psi\left(\tilde{e}_\infty\right)}{v'\left(\bar{c}_\infty\right)} = \delta \tilde{e}_\infty.$$

Because ψ is strictly convex, we have

$$\frac{\psi'\left(\tilde{e}_\infty\right)}{v'\left(\bar{c}_\infty\right)} > \delta.$$

We will now consider an adjusted contract in which effort increases at date 1 by $\varepsilon > 0$, raising the disutility of effort at date 1 by $\psi\left(\tilde{e}_\infty + \varepsilon\right) - \psi\left(\tilde{e}_\infty\right)$. Because payments to the agent increase at all dates under condition (FP_t^{un}), the new policy will not satisfy the principal's constraint (PC_t) if this is the only adjustment. We therefore simultaneously reduce effort from any fixed date $T \geq 2$ onwards by $\kappa(\varepsilon) > 0$ (i.e., $e_t = e_\infty - \kappa(\varepsilon)$ for $t \geq T$).

We let $\bar{c}_\infty(\varepsilon, \kappa(\varepsilon))$ denote equilibrium consumption under the new plan (naturally, $\bar{c}_\infty(0,0)$ denotes payments and consumption under the original plan). The new payments satisfy

$$\begin{aligned}\frac{v(\bar{c}_\infty(\varepsilon,\kappa(\varepsilon)))}{1-\delta} - \frac{v(\bar{c}_\infty(0,0))}{1-\delta} =& \psi(\tilde{e}_\infty + \varepsilon) - \psi(\tilde{e}_\infty) \\ &- \frac{\delta^{T-1}}{1-\delta}(\psi(\tilde{e}_\infty) - \psi(\tilde{e}_\infty - \kappa(\varepsilon)))\end{aligned}$$

or

$$\bar{c}_\infty(\varepsilon,\kappa(\varepsilon)) = v^{-1}\left(\begin{array}{c}(1-\delta)(\psi(\tilde{e}_\infty+\varepsilon) - \psi(\tilde{e}_\infty)) \\ -\delta^{T-1}(\psi(\tilde{e}_\infty) - \psi(\tilde{e}_\infty - \kappa(\varepsilon))) + v(\bar{c}_\infty(0,0))\end{array}\right)$$

We will determine the value for $\kappa(\varepsilon)$ by the equality

$$\frac{\psi(\tilde{e}_\infty - \kappa(\varepsilon))}{v'(\bar{c}_\infty(\varepsilon,\kappa(\varepsilon)))} - \delta(\tilde{e}_\infty - \kappa(\varepsilon)) = 0. \tag{26}$$

Then, provided ε is small enough, the principal's constraints (PC_t) continue to hold. The derivative of the left-hand side of (26) with respect to $\kappa(\varepsilon)$, evaluated at $(\varepsilon,\kappa(\varepsilon)) = (0,0)$, is

$$\delta - \frac{\psi'(\tilde{e}_\infty)}{v'(\bar{c}_\infty(0,0))} + v''(\bar{c}_\infty(0,0))\left(\frac{\delta^{T-1}\psi'(\tilde{e}_\infty)}{v'(\bar{c}_\infty(0,0))^3}\right)\psi(\tilde{e}_\infty).$$

This is strictly negative, using that $\frac{\psi'(\tilde{e}_\infty)}{v'(\bar{c}_\infty(0,0))} > \delta$. The derivative of the left-hand side of (26) instead with respect to ε, evaluated at $(\varepsilon,\kappa(\varepsilon)) = (0,0)$, is

$$-v''(\bar{c}_\infty(0,0))\left(\frac{(1-\delta)\psi'(\tilde{e}_\infty)}{v'(\bar{c}_\infty(0,0))^3}\right)\psi(\tilde{e}_\infty).$$

The implicit function theorem then gives us that κ is locally well-defined, unique and continuously differentiable, with derivative approaching

$$\begin{aligned}\kappa'(0) =& \frac{v''(\bar{c}_\infty(0,0))\left(\frac{(1-\delta)\psi'(\tilde{e}_\infty)}{v'(\bar{c}_\infty(0,0))^3}\right)\psi(\tilde{e}_\infty)}{\delta - \frac{\psi'(\tilde{e}_\infty)}{v'(\bar{c}_\infty(0,0))} + v''(\bar{c}_\infty(0,0))\left(\frac{\delta^{T-1}\psi'(\tilde{e}_\infty)}{v'(\bar{c}_\infty(0,0))^3}\right)\psi(\tilde{e}_\infty)} \\ &< \frac{1-\delta}{\delta^{T-1}}\end{aligned} \tag{27}$$

as $\varepsilon \to 0$ (that $\kappa'(0) < \frac{1-\delta}{\delta^{T-1}}$ follows because $\frac{\psi'(\tilde{e}_\infty)}{v'(\bar{c}_\infty(0,0))} > \delta$).

The NPV of effort increases by

$$\varepsilon - \frac{\delta^{T-1}}{1-\delta}\kappa(\varepsilon) = \left(1 - \frac{\delta^{T-1}}{1-\delta}\kappa'(0)\right)\varepsilon + o(\varepsilon).$$

From the inequality (27) we have $1-\frac{\delta^{T-1}}{1-\delta}\kappa'(0)>0$, and so the increase in effort is strictly positive for ε small enough. Using that (FP_t^{un}) continues to hold under the modified effort policy at $t=\infty$, a marginal increase in the NPV of effort is compensated by an increase in the NPV of payments to the agent by $\frac{\psi'(\tilde{e}_\infty)}{v'(\bar{c}_\infty(0,0))}$. Therefore, the principal's payoff increases by

$$\left(1-\frac{\psi'(\tilde{e}_\infty)}{v'(\bar{c}_\infty(0,0))}\right)\left(1-\frac{\delta^{T-1}}{1-\delta}\kappa'(0)\right)\varepsilon+o(\varepsilon)$$

which is strictly positive for small enough ε, recalling that $1-\frac{\psi'(\tilde{e}_\infty)}{v'(\bar{c}_\infty(0,0))}>0$. □

Note next that, when there is no self-enforceable contract implementing the first-best, $\bar{t}$ could equal one, so that the principal's constraint is not slack in any period. In this case effort decreases strictly over time. However, one may be interested to determine whether it is possible instead that the principal's constraints (PC_t) are initially slack for several periods, so that effort is initially constant (before beginning to strictly decrease at some future date). This can be guaranteed for appropriate values of the discount factor δ.

Lemma A.8. *For any v and ψ, there exists a discount factor δ and initial balance b_1 such that (i) the first-best is not sustainable in a self-enforceable contract, and (ii) for any optimal contract, the principal's constraint (PC_t) is slack for at least $t=1,2$, and hence, in an optimal contract, effort is constant over at least the first two dates (i.e., $\tilde{e}_1=\tilde{e}_2$).*

Proof. Fix v and ψ satisfying the properties in the model set-up, and fix a scalar $\gamma>0$. Define the function $b_1(\delta)=\frac{\gamma}{1-\delta}$. As explained in the main text, there is then a threshold value δ^* such that $\delta\geq\delta^*$ and $b_1=b_1(\delta)$ implies the first-best policy is part of a self-enforceable contract, while $\delta<\delta^*$ and $b_1=b_1(\delta)$ implies this is not the case. We therefore aim to show that the principal's constraint (PC_t) is slack over some initial periods when δ is below, but close enough to, δ^*, and with $b_1=b_1(\delta)$. We do so in three steps. In these steps, we let δ parameterize the environment, leaving $b_1=b_1(\delta)$ implicit.

Step 1. First, by considering constant effort policies, it is easily seen that the principal's payoff in an optimal contract approaches that for parameters δ^* and $b_1^*=b_1(\delta^*)$ as $\delta\to\delta^*$ from below.

Step 2. Next, let e^* be the first-best effort for parameters δ^* and b_1^*. We show that, for any $\varepsilon>0$ and period T, there exists $\hat{\delta}(T,\varepsilon)$ such that, for $\delta\in\left(\hat{\delta}(T,\varepsilon),\delta^*\right)$, $\max_{t\leq T}\left|\tilde{e}_t-e^*\right|\leq\varepsilon$, where $(\tilde{e}_t)_{t\geq1}$ is the optimal effort policy for parameter δ.

By Lemma A.1, for $\delta\leq\delta^*$, any optimal effort policy is contained in $[0,z(v'(\gamma))]^\infty$. The principal's payoff under a self-enforceable relational contract with arbitrary effort policy $(\tilde{e}_t)_{t=1}^\infty$

(and satisfying Condition (FP_t^{un})) is

$$\sum_{t=1}^{\infty}\delta^{t-1}\tilde{e}_t-\frac{v^{-1}\left((1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\psi\left(\tilde{e}_t\right)+v\left(\gamma\right)\right)}{1-\delta}+b_1\left(\delta\right)$$

which varies continuously in δ, with the continuity uniform over effort policies contained in $[0,z\left(v'\left(\gamma\right)\right)]^{\infty}$.

Fix $\delta=\delta^*$, and fix any $\varepsilon>0$ and any period T. We then have that, for an effort policy $(\tilde{e}_t)_{t\geq 1}$ contained in $[0,z\left(v'\left(\gamma\right)\right)]^{\infty}$ and satisfying $\max_{t\leq T}\left|\tilde{e}_t-e^{FB}(b_1)\right|\geq\varepsilon$, and for payments satisfying Condition (FP_t^{un}), the principal's payoff is less than that sustained by the first-best contract by some amount $\nu>0$. This follows from uniqueness of the first-best policy and continuity of the principal's objective in the effort policy $(\tilde{e}_t)_{t\geq 1}$. However, the aforementioned continuity of the principal's payoff in δ, together with Step 1, implies that, when δ is close enough to (but below) δ^*, such an effort policy cannot be optimal.

Step 3. Notice that, for $\delta=\delta^*$, under the first-best policy, the principal's constraint (PC_t) is slack at both dates $t=1$ and $t=2$. It is then easily verified that, provided ε is taken small enough, and T large enough, these constraints must also be slack under an optimal policy when $\delta\in\left(\hat{\delta}\left(T,\varepsilon\right),\delta^*\right)$. □

(End of the proof of Proposition 4.3.) □

Proof of Proposition 4.4.

Proof. **Step 1.** Note first that since the contract is optimal, the condition in Equation (4) holds. Therefore, if the contract is self-enforceable and Condition (FP_t^{un}) fails for some $t'>\bar{t}$, we must have

$$\frac{v\left(\bar{c}_{t'-1}\right)}{1-\delta}-\sum_{s=1}^{t'-1}\delta^{s-1}\psi(\tilde{e}_s)<\frac{v\left(\bar{c}_{\infty}\right)}{1-\delta}-\sum_{t=1}^{\infty}\delta^{t-1}\psi(\tilde{e}_t)=\frac{v\left((1-\delta)\,b_1\right)}{1-\delta}.$$

We can then increase $\tilde{w}_{t'-1}$ by $\varepsilon>0$ and decrease $\tilde{w}_{t'}$ by ε/δ. For an appropriate choice of ε, the agent's constraint (AC_t^{un}) holds as an equality at date t', and it is satisfied at all other dates. The principal's constraint (PC_t) is slackened at t', and continues to hold at all other dates. Crucially, we arrive at a new self-enforceable contract that obtains the same payoff for the principal, and for which (i) the agent's constraint (AC_t^{un}) holds as an equality at t', and (ii) the principal's constraint (PC_t) is slack at this date.

Step 2. We now follow an analogous argument to the second part of the proof of Lemma 4.2. That is, given Proposition 4.3 and the assumption that $t'>\bar{t}$ implies that $\tilde{e}_{t'}<\tilde{e}_{t'-1}$, we show there is a contract for the principal that is strictly more profitable. Take as a starting point the adjusted contract, where the principal's constraint is slack at date t'. We can choose a new contract with efforts $(\tilde{e}'_t)_{t\geq 1}$ and payments $(\tilde{w}'_t)_{t\geq 1}$ satisfying Equation ($\mathrm{FP}_t^{\mathrm{un}}$) that coincides with $(\tilde{e}_t)_{t\geq 1}$ except in periods $t'-1$ and t', so we have $\tilde{e}_{t'} < \tilde{e}'_{t'} \leq \tilde{e}'_{t'-1} < \tilde{e}_{t'-1}$ and

$$\psi\left(\tilde{e}'_{t'-1}\right) + \delta\psi\left(\tilde{e}'_{t'}\right) = \psi\left(\tilde{e}_{t'-1}\right) + \delta\psi\left(\tilde{e}_{t'}\right),$$

which implies that $\tilde{e}'_{t'-1} + \delta\tilde{e}'_{t'} > \tilde{e}_{t'-1} + \delta\tilde{e}_{t'}$. Note that

$$\tilde{w}'_{t'-1} + \delta\tilde{w}'_{t'} = \tilde{w}_{t'-1} + \delta\tilde{w}_{t'}.$$

Also, $\tilde{w}'_{t'-1} < \tilde{w}_{t'-1}$ and $\tilde{w}'_{t'} > \tilde{w}_{t'}$. Provided the changes are small, the principal's constraint (PC_t) at t' remains satisfied. Moreover, the principal's constraints are relaxed at date $t'-1$, and because the NPV of effort increases, also at all earlier dates. Therefore, the principal's constraints are satisfied at all dates and the principal's payoff strictly increases.

□

A.3 Proofs of the results in Section 5

Proof of Proposition 5.1

Proof. **Necessity**. Consider supposed equilibrium strategies $(\alpha_t)_{t\geq 1}$ and $(\sigma_t)_{t\geq 1}$. Suppose Condition ($\mathrm{AC}_t^{\mathrm{ob}}$) fails at some date, and let t be the earliest such date. Then the agent obtains a strictly higher payoff by abiding by the contract to date $t-1$ and then consuming $\tilde{b}_t(1-\delta)$ from then on, as opposed to following the putative equilibrium strategy. That is, the agent has a profitable deviation. If instead (PC_t) fails at some date, then there is a first date t at which it is violated. The principal has a profitable deviation by ceasing all payments from then on.

Sufficiency. Let $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ be a feasible contract satisfying conditions ($\mathrm{AC}_t^{\mathrm{ob}}$) and ($\mathrm{PC}_t$). Specify strategies $(\alpha_t)_{t\geq 1}$ and $(\sigma_t)_{t\geq 1}$ for the agent and principal as follows. Provided that $(e_s, c_s, w_s) = (\tilde{e}_s, \tilde{c}_s, \tilde{w}_s)$ for all $s < t$, the agent consumes $\tilde{c}_t$ and chooses effort $\tilde{e}_t$. Otherwise, the agent consumes $(1-\delta)\, b_t$ and chooses effort zero. Provided that $(e_s, c_s, w_s) = (\tilde{e}_s, \tilde{c}_s, \tilde{w}_s)$ for all $s < t$, and $(e_t, c_t) = (\tilde{e}_t, \tilde{c}_t)$, the principal pays $\tilde{w}_t$. Otherwise, she pays zero.

Now, we can check that the players do not want to deviate at any history. Consider a date-t history at which $(e_s, c_s, w_s) = (\tilde{e}_s, \tilde{c}_s, \tilde{w}_s)$ for all $s < t$. Given the principal's strategy, the agent obtains continuation payoff $\sum_{s=t}^{\infty} \delta^{s-t}\left(v\left(\tilde{c}_s\right) - \tilde{e}_s\right)$ by remaining faithful to the specified strategy. If the agent deviates, then he is paid zero from t onwards. Recall from the discussion in Section 4 that the restriction on the agent's strategy implies, for any stream of consumption $(c_s)_{s\geq t}$ chosen by the agent,

$$\sum_{s=t}^{\infty} \delta^{s-t} c_s \leq \tilde{b}_t,$$

and so optimal consumption is set at $(1-\delta)\,\tilde{b}_t$ for all dates. Therefore, the highest payoff the agent can obtain under a deviation is $\frac{v\left(\tilde{b}_t(1-\delta)\right)}{1-\delta}$. The inequality ($\text{AC}_t^{\text{ob}}$) states that this is less than $\sum_{s=t}^{\infty} \delta^{s-t} \left(v\left(\tilde{c}_s\right) - \tilde{e}_s\right)$. At any other history (i.e., if $(e_s, c_s, w_s) \neq (\tilde{e}_s, \tilde{c}_s, \tilde{w}_s)$ for some $s \neq t$), the agent will never be paid again and so consumption $(c_s)_{s\geq t}$ must satisfy $\sum_{s=t}^{\infty} \delta^{s-t} c_s \leq b_t$. Hence, consuming $(1-\delta)\, b_t$ and putting no effort is optimal.

On the principal's side, for a date-t history at which $(e_s, c_s, w_s) = (\tilde{e}_s, \tilde{c}_s, \tilde{w}_s)$ for all $s < t$, and given $(e_t, c_t) = (\tilde{e}_t, \tilde{c}_t)$, the principal's continuation payoff from following the specified strategy is $\sum_{s=t+1}^{\infty} \delta^{s-t}(\tilde{e}_s - \tilde{w}_s)$ which is larger than the date-t payment $\tilde{w}_t$ by Condition (PC_t). Hence, the principal prefers to follow the specified strategy. Following any deviation, given the agent's strategy, the principal clearly finds paying zero optimal.

□

Proof of Lemma 5.1

Proof. Fix an optimal contract $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ and suppose that Condition 7 is not satisfied for all t. Since the contract is self-enforceable, we have, for all t, Equation (AC_t^{ob}) holds. Suppose then that the inequality is strict at some $t > 1$. Then consider a new relational contract with payment reduced at date t by $\varepsilon > 0$, and with payment increased at date $t-1$ by $\delta\varepsilon$; hold the consumption and effort profile the same. This change increases $\tilde{b}_t$ by ε and, for appropriately chosen ε, the constraint (AC_t^{ob}) holds with equality. The principal's constraint (PC_t) is slackened at date t (payments are smaller in that period and the same from then on), and its constraints are unaffected in all earlier periods.[15] Constraints are also unaffected from $t+1$ onwards, since the contract remains unchanged at these dates ($\tilde{b}_{t+1}$ is unchanged). The principal's payoff remains unchanged.

The adjustment to the contract therefore still yields an optimal contract. The adjustments can be applied sequentially at the dates for which (AC_t^{ob}) holds as a strict inequality, yielding a contract for which (AC_t^{ob}) holds at all dates.

If the inequality is strict at $t = 1$, then both $\tilde{c}_1$ and $\tilde{w}_1$ can be reduced by the same small amount $\varepsilon > 0$, leaving $\tilde{b}_2$ unchanged, and keeping the rest of the relational contract the same. This adjustment leaves unchanged the constraints of the principal in all periods $t > 1$,

[15]Note that the principal's constraint in period $t-1$ can be written as $\sum_{s=t-1}^{\infty} \delta^{s-t+1} \tilde{w}_s \leq \sum_{s=t}^{\infty} \delta^{s-t+1} \tilde{e}_s$. Since the left hand side remains the same after the suggested change, the constraint of the principal still holds.

and slackens the principal's constraint at date 1. It also leaves the constraints of the agent unaffected in all periods $t > 1$, and if $\varepsilon > 0$ is small enough, the agent's date-1 constraint is still satisfied. This increases the principal's payoff.

Finally, note that when Condition (7) is satisfied for all t, all payments to the agent are non-negative given that the disutility of effort is non-negative. This ensures that the above adjustments also yield a contract that is feasible. □

Proof of Proposition 5.2.

Proof. Suppose that condition (8) is satisfied. As in the main text, consider the contract where the agent puts effort $e^{FB}(b_1)$ per period, is paid $w^{FB}(b_1)$ in each period, and where the agent consumes $c^{FB}(b_1)$ in each period. The agent's balance remains constant at b_1, and by the agent's indifference condition in Proposition 3.1, the agent's constraint (AC_t^{ob}) is satisfied. Condition (8) is simply the principal's constraint (PC_t). Hence, the first-best contract is self-enforceable. □

Proof of Proposition 5.3

Proof. It will be useful to write the recursive problem in the main text by substituting out agent effort. To this end, define a function $\hat{e}$ by:

$$\hat{e}(c_t, b_t, b_{t+1}) \equiv \psi^{-1}\Big(v(c_t) + \tfrac{\delta}{1-\delta}v\left((1-\delta)\,b_{t+1}\right) - \tfrac{1}{1-\delta}v((1-\delta)b_t)\Big). \tag{28}$$

We can then write the principal's optimal payoff given balance $\tilde{b}_t > 0$ (which we establish below can be attained by a self-enforceable contract) as follows:

$$V(\tilde{b}_t) = \max_{c_t, b_{t+1}} \left(\hat{e}(c_t, \tilde{b}_t, b_{t+1}) - (\delta b_{t+1} - \tilde{b}_t + c_t) + \delta V(b_{t+1})\right) \tag{29}$$

subject to the principal's constraint

$$\delta b_{t+1} - \tilde{b}_t + c_t \leq \delta V\left(b_{t+1}\right) \tag{30}$$

and to the requirement that the implied effort is non-negative, i.e.

$$v(c_t) + \tfrac{\delta}{1-\delta}v\left((1-\delta)\,b_{t+1}\right) - \tfrac{1}{1-\delta}v((1-\delta)\tilde{b}_t) \geq 0. \tag{31}$$

The proof of Proposition 5.3 will now consist of eight lemmas.

1. Lemma A.9 shows that the principal's payoff is strictly positive for all b_1.

2. Lemma A.10 is analogous to Lemma A.1 in that it bounds the marginal disutility of effort by the marginal utility of consumption in any optimal contract.

3. Lemma A.11 proves the validity of the Euler equation for any optimal contract and shows that consumption in an optimal contract is weakly decreasing in time.

4. Lemma A.12 shows that any optimal contract either gives the principal his first-best payoff and the balance is constant over time, or the balance is strictly decreasing towards some $\tilde{b}_\infty > 0$.

5. Lemma A.13 shows that if an optimal contract does not achieve the first best, then the continuation payoff of the principal strictly increases over time.

6. Lemma A.14 shows that if an optimal contract does not achieve the first best, then the principal's constraint (PC_t) holds with equality in every period.

7. Lemma A.15 shows that, in any optimal contract, effort and payments strictly increase over time, while consumption strictly decreases.

8. Lemma A.16 shows that an optimal contract exists.

Lemma A.9. *Fix an optimal relational contract* $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$. *For all* t *and* $\tilde{b}_t > 0$, $V(\tilde{b}_t) \in (0, V^{FB}(\tilde{b}_t)]$. *If* $V(\tilde{b}_t) = V^{FB}(\tilde{b}_t)$, *the contract satisfying the conditions* (FP_t^{ob}) *is stationary.*

Proof. The principal can never do better than offering the first-best contract: i.e., $V(b) \leq V^{FB}(b)$ for all $b > 0$. If the first best is attainable at a given date t and balance $\tilde{b}_t$, then (using Proposition 3.1) it is attained by constant consumption equal to $c^{FB}(\tilde{b}_t)$ and constant effort equal to $e^{FB}(\tilde{b}_t)$. If the contract satisfies "fastest payments", then, in particular, Equation (7) holds at all $\tau \geq t$, and hence the balance must remain constant at $\tilde{b}_t$, as explained in the main text. Hence, the contract is stationary.

Now let us show that $V\left(\tilde{b}_t\right) > 0$ irrespective of the value $\tilde{b}_t > 0$. For this, let us state the conditions for a stationary contract (i.e., $(\tilde{e}_\tau, \tilde{c}_\tau, \tilde{w}_\tau, \tilde{b}_\tau)_{\tau\geq t}$ with $(\tilde{e}_\tau, \tilde{c}_\tau, \tilde{w}_\tau, \tilde{b}_\tau) = (\tilde{e}, \tilde{c}, \tilde{w}, \tilde{b})$ for all $\tau \geq t$) to be self-enforceable. The principal's constraint (PC_t) at any date may be written

$$\delta\tilde{e} \geq \tilde{w} = \tilde{c} - (1-\delta)\tilde{b}.$$

The agent's constraint (AC_t^{ob}) is

$$v(\tilde{c}) - \psi(\tilde{e}) \geq v((1-\delta)\tilde{b}).$$

Taking the latter to hold with equality, we have $\tilde{e} = \breve{e}(\tilde{c}) \equiv \psi^{-1}(v(\tilde{c}) - v((1-\delta)\tilde{b}_t))$. That $\breve{e}((1-\delta)\tilde{b}_t) = 0$ and $\breve{e}'((1-\delta)\tilde{b}_t) = +\infty$ shows that the the former inequality (the principal's constraint) is satisfied when $\tilde{c}$ is above but close enough to $(1-\delta)\tilde{b}$. That $\delta\tilde{e} \geq \tilde{w}$ and $\tilde{e} > 0$ confirms that the principal's payoff $\frac{\tilde{e}-\tilde{w}}{1-\delta}$ is strictly positive. That is, $V(\tilde{b}_t) > 0$.

□

Lemma A.10. *In any optimal contract* $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ *that satisfies the conditions* (FP_t^{ob})*, we have* $\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t > 0$ *for all* t*. Furthermore,* $\psi'(\tilde{e}_t) \leq v'(\tilde{c}_t)$ *for all* t*, and* $\psi'(\tilde{e}_t) < v'(\tilde{c}_t)$ *only if* $\tilde{w}_t = \delta V(\tilde{b}_{t+1})$*.*

Proof. **Proof that** $\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t > 0$ **for all** t**.** We first prove that $\tilde{w}_t > 0$ for all t. We do this assuming, for the sake of contradiction, that $\tilde{w}_t = 0$ for some t. This implies that $\tilde{e}_t = 0$, $\tilde{c}_t = (1-\delta)\tilde{b}_t$ and $\tilde{b}_{t+1} = \tilde{b}_t$ (this is the only possibility for Condition (31) to be satisfied). This implies $V(\tilde{b}_t) = \delta V(\tilde{b}_t)$, that is, $V(\tilde{b}_t) = 0$, but this contradicts Lemma A.9.

To prove that $\tilde{e}_t > 0$ for all t, suppose to the contrary that $\tilde{e}_t = 0$ for some t. If $\tilde{w}_t < \delta V(\tilde{b}_{t+1})$, we can raise effort to $\check{e}_t = \varepsilon$ at date t for $\varepsilon > 0$; raise date-t consumption to

$$\check{c}_t = v^{-1}\big(\psi(\varepsilon) - \tfrac{\delta}{1-\delta}v((1-\delta)\tilde{b}_{t+1}) + \tfrac{1}{1-\delta}v((1-\delta)\tilde{b}_t)\big);$$

and raise the principal's date-t payment to $\check{w}_t = \tilde{w}_t + \check{c}_t - \tilde{c}_t$. Thus, the agent's balance at $t+1$ remains unchanged, and the only adjustments to the contract are at date t. For ε sufficiently small, we have $\check{w}_t < \delta V(\tilde{b}_{t+1})$, and the principal's payoff strictly increases. Also, the agent remains willing to be obedient at all periods. If instead $\tilde{w}_t = \delta V(\tilde{b}_{t+1})$, we have $V(\tilde{b}_t) = 0$, but this contradicts Lemma A.9.

That $\tilde{c}_t, \tilde{b}_t > 0$ for all t follows immediately from our assumption that the Conditions (FP_t^{ob}) hold at all dates t, and because $b_1 > 0$.

Proof that $\psi'(\tilde{e}_t) \leq v'(\tilde{c}_t)$ **for all** t**, and** $\psi'(\tilde{e}_t) < v'(\tilde{c}_t)$ **only if** $\tilde{w}_t = \delta V(\tilde{b}_{t+1})$**.** Define

$$\underline{c}(\tilde{b}_t, \tilde{b}_{t+1}) \equiv v^{-1}\big(\tfrac{1}{1-\delta}v((1-\delta)\tilde{b}_t) - \tfrac{\delta}{1-\delta}v\left((1-\delta)\,\tilde{b}_{t+1}\right)\big),$$

interpreted as the lowest consumption level that permits the constraint (31) to be satisfied, for fixed values of $\tilde{b}_t$ and $\tilde{b}_{t+1}$. Consider the problem of maximizing

$$\hat{e}(c_t, \tilde{b}_t, \tilde{b}_{t+1}) - (\delta\tilde{b}_{t+1} - \tilde{b}_t + c_t) + \delta V(\tilde{b}_{t+1}) \tag{32}$$

with respect to c_t on $[\underline{c}(\tilde{b}_t, \tilde{b}_{t+1}), +\infty)$. Given that $\hat{e}(\cdot, \tilde{b}_t, \tilde{b}_{t+1})$ is a continuous and strictly concave function, and that $\lim_{c\to+\infty} \frac{\partial \hat{e}(c,\tilde{b}_t,\tilde{b}_{t+1})}{\partial c} = 0$, there is a unique solution of the maximization problem, denoted $c^*(\tilde{b}_t, \tilde{b}_{t+1})$. Furthermore, since $\psi'(0) = 0$, we have that $c^*(\tilde{b}_t, \tilde{b}_{t+1}) > \underline{c}(\tilde{b}_t, \tilde{b}_{t+1})$, and the first-order condition establishes

$$\psi'(\hat{e}(c^*(\tilde{b}_t, \tilde{b}_{t+1}), \tilde{b}_t, \tilde{b}_{t+1})) = v'(c^*(\tilde{b}_t, \tilde{b}_{t+1})).$$

If we have $\delta\tilde{b}_{t+1} - \tilde{b}_t + c^*(\tilde{b}_t, \tilde{b}_{t+1}) \leq \delta V(\tilde{b}_{t+1})$ then it is clear that optimality requires that $\tilde{c}_t = c^*(\tilde{b}_t, \tilde{b}_{t+1})$. Otherwise, given the concavity of (32), we must have

$$\tilde{c}_t = \delta V(\tilde{b}_{t+1}) - \delta\tilde{b}_{t+1} + \tilde{b}_t < c^*(\tilde{b}_t, \tilde{b}_{t+1})$$

and hence $\tilde{w}_t = \delta V(\tilde{b}_{t+1})$. In this case,

$$\tilde{e}_t = \hat{e}(\tilde{c}_t, \tilde{b}_t, \tilde{b}_{t+1}) < \hat{e}(c^*(\tilde{b}_t, \tilde{b}_{t+1}), \tilde{b}_t, \tilde{b}_{t+1}),$$

and so we have $\psi'(\tilde{e}_t) < v'(c^*(\tilde{b}_t, \tilde{b}_{t+1})) < v'(\tilde{c}_t)$. □

The following result establishes the Euler equation given in the main text as well as the monotonicity of the consumption plan.

Lemma A.11. *Any optimal contract* $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ *satisfies the Euler equation*

$$1 - \frac{v'\left((1-\delta)\,\tilde{b}_{t+1}\right)}{v'(\tilde{c}_t)} = \frac{v'(\tilde{c}_{t+1})}{\psi'(\tilde{e}_{t+1})}\left(1 - \frac{v'\left((1-\delta)\,\tilde{b}_{t+1}\right)}{v'(\tilde{c}_{t+1})}\right) \tag{33}$$

in all periods. Furthermore, $\tilde{c}_t \geq \tilde{c}_{t+1} > (1-\delta)\tilde{b}_{t+1}$ *for all* t.

Proof. We divide the proof in 3 steps:

Step 1: Fix an optimal contract $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$. Consider a contract $(\check{e}_t, \check{c}_t, \check{w}_t, \check{b}_t)_{t\geq 1}$, coinciding with the original contract in all periods except for periods t and $t+1$ (so, also, $\check{b}_t = \tilde{b}_t$). We specify that the new contract keeps the agent indifferent between being obedient and optimally deviating in all periods. This requires

$$v(\check{c}_t) - \psi(\check{e}_t) + \tfrac{\delta}{1-\delta} v\big(\tfrac{1-\delta}{\delta}(\tilde{b}_t + \check{w}_t - \check{c}_t)\big) = \tfrac{1}{1-\delta} v((1-\delta)\tilde{b}_t), \tag{34}$$

$$v\big(\tfrac{1}{\delta}(\tilde{b}_t + \check{w}_t - \check{c}_t) + \check{w}_{t+1} - \delta\tilde{b}_{t+2}\big) - \psi(\check{e}_{t+1}) + \tfrac{\delta}{1-\delta} v((1-\delta)\tilde{b}_{t+2}) = \tfrac{1}{1-\delta} v\big(\tfrac{1-\delta}{\delta}(\tilde{b}_t + \check{w}_t - \check{c}_t)\big), \tag{35}$$

which uses that consumption in period $t+1$ under the new contract is $\check{c}_{t+1} = \frac{1}{\delta}(\tilde{b}_t + \check{w}_t - \check{c}_t) + \check{w}_{t+1} - \delta\tilde{b}_{t+2}$ (guaranteeing the agent has savings $\tilde{b}_{t+2}$ at date $t+2$).

Fix $\check{e}_t = \tilde{e}_t$ and $\check{w}_{t+1} = \tilde{w}_{t+1}$. Equations (34) and (35) implicitly define $\check{e}_{t+1}$ and $\check{w}_t$ as functions of $\check{c}_t$. Let these functions be denoted $\hat{e}_{t+1}(\cdot)$ and $\hat{w}_t(\cdot)$, respectively. We can use the implicit function theorem to compute the derivatives at $\check{c}_t = \tilde{c}_t$:

$$\hat{e}'_{t+1}(\tilde{c}_t) = \frac{v'(\tilde{c}_t)\big(v'((1-\delta)\tilde{b}_{t+1}) - v'(\tilde{c}_{t+1})\big)}{\delta\psi'(\hat{e}_{t+1}(\tilde{c}_t))\, v'((1-\delta)\tilde{b}_{t+1})} \quad \text{and} \quad \hat{w}'_t(\tilde{c}_t) = 1 - \frac{v'(\tilde{c}_t)}{v'((1-\delta)\tilde{b}_{t+1})}.$$

If $\check{c}_t$ is chosen to be equal to $\tilde{c}_t + \varepsilon$, for some (positive or negative) ε small, the total effect on the continuation payoff of the principal at time t is $(-\hat{w}_t'(\tilde{c}_t) + \delta\hat{e}_{t+1}'(\tilde{c}_t))\varepsilon + o(\varepsilon)$ (where $o(\varepsilon)$ represents terms that vanish faster than ε as $\varepsilon \to 0$). Hence, a necessary condition for optimality is that $-\hat{w}_t'(\tilde{c}_t) + \delta\hat{e}_{t+1}'(\tilde{c}_t) = 0$, which is equivalent the Euler equation (33).

The Euler equation implies that if $v'(\tilde{c}_{t+1}) = \psi'(\tilde{e}_{t+1})$ we have $\tilde{c}_t = \tilde{c}_{t+1}$. From Lemma A.10 we have that, if $v'(\tilde{c}_{t+1}) \neq \psi'(\tilde{e}_{t+1})$, there are three possibilities:

1. If both sides of the Euler equation are strictly positive, then $\tilde{c}_t < \tilde{c}_{t+1} < (1-\delta)\tilde{b}_{t+1}$. In this case, since $\tilde{e}_{t+1} > 0$ (from Lemma A.10), we have $\tilde{b}_{t+2} > \tilde{b}_{t+1}$ (from the agent's constraint (28)). If the constraint does not bind at $t+2$, then $\tilde{c}_{t+2} = \tilde{c}_{t+1}$, and if it binds, $\tilde{c}_{t+1} < \tilde{c}_{t+2} < (1-\delta)\tilde{b}_{t+2}$.

2. If both sides of the Euler equation are zero, then $\tilde{c}_t = \tilde{c}_{t+1} = (1-\delta)\tilde{b}_{t+1}$.

3. If both sides of the Euler equation are strictly negative, then $\tilde{c}_t > \tilde{c}_{t+1} > (1-\delta)\tilde{b}_{t+1}$.

Step 2: We now prove that if $\tilde{c}_t \leq \tilde{c}_{t+1} \leq (1-\delta)\tilde{b}_{t+1}$ then $\tilde{c}_s \leq \tilde{c}_{s+1} < (1-\delta)\tilde{b}_{s+1}$ for all $s > t$. Assume first that there is a period t such that $\tilde{c}_{t+1} \leq (1-\delta)\tilde{b}_{t+1}$. Hence, since $\tilde{e}_{t+1} > 0$, we have $\tilde{b}_{t+2} > \tilde{b}_{t+1}$. This shows that the Euler satisfies

$$1 - \frac{v'\left(\left(1-\delta\right)\tilde{b}_{t+2}\right)}{v'\left(\tilde{c}_{t+1}\right)} = \frac{v'\left(\tilde{c}_{t+2}\right)}{\psi'\left(\tilde{e}_{t+2}\right)}\left(1 - \frac{v'\left(\left(1-\delta\right)\tilde{b}_{t+2}\right)}{v'\left(\tilde{c}_{t+2}\right)}\right) > 0.$$

Since $v'\left(\tilde{c}_{t+2}\right)/\psi'\left(\tilde{e}_{t+2}\right) \geq 1$, $(1-\delta)\tilde{b}_{t+2} > \tilde{c}_{t+2} \geq \tilde{c}_{t+1}$. The result then follows by induction.

Step 3: We prove that $\tilde{c}_t > (1-\delta)\tilde{b}_t$ for all $t > 1$; it then follows immediately from Step 1 that consumption is (weakly) decreasing in t. Assume then, for the sake of contradiction, that there is a $t' > 1$ such that $\tilde{c}_{t'} \leq (1-\delta)\tilde{b}_{t'}$. Then $\tilde{c}_{t'-1} \leq \tilde{c}_{t'}$ (either $v'(\tilde{c}_{t'}) = \psi'(\tilde{e}_{t'})$ and so $\tilde{c}_{t'-1} = \tilde{c}_{t'}$, or $v'(\tilde{c}_{t'}) > \psi'(\tilde{e}_{t'})$ and Case 1 or Case 2 from above applies). Hence, from Step 2, we have that $(1-\delta)\tilde{b}_{s+1} > \tilde{c}_{s+1} \geq \tilde{c}_s$ for all $s \geq t'$.

Also, since effort is strictly positive at all times, we have

$$\sum_{s=t'}^{\infty} \delta^{s-t'} v(\tilde{c}_s) > \tfrac{1}{1-\delta} v((1-\delta)\tilde{b}_{t'}),$$

and so there must be a period $s \geq t'$ where $\tilde{c}_{s+1} > \tilde{c}_{t'}$ (recall we assumed $\tilde{c}_{t'} \leq (1-\delta)\tilde{b}_{t'}$). Let t'' be the earliest such period, and note that it satisfies $\tilde{c}_{t''} < \tilde{c}_{t''+1}$. We now want to show that the principal can offer a strictly more profitable contract satisfying the conditions (PC_t) and (AC_t^{ob}).

We will use that, for all $s \geq 1$, $\tilde{b}_s + \sum_{\tau=s}^{\infty} \delta^{\tau-s}\tilde{w}_\tau = \sum_{\tau=s}^{\infty} \delta^{\tau-s}\tilde{c}_\tau$ (were this not true, feasibility of $(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t)_{t\geq 1}$ implies that $\lim_{t\to\infty} \delta^{t-1}b_t > 0$; but then the payments $\tilde{w}_t$ could be reduced, say by a small constant amount in every period, yielding a feasible contract satisfying the constraints (PC$_t$) and (AC$_t^{\text{ob}}$), and giving a higher payoff to the principal). Note that since (by Step 2) consumption is weakly increasing from t' onwards, so these expressions are weakly increasing with s from t' onwards, and increase strictly between $s = t''$ and $s = t'' + 1$.

The key step is to consider a "new" policy, with the same equilibrium payments to the agent, but a different agreed consumption sequence. In particular, we specify consumption $\bar{c}$ in each period from t'' onwards, where

$$\bar{c} = (1-\delta) \sum_{\tau=t''}^{\infty} \delta^{\tau-t''} \tilde{c}_\tau$$

which note implies that $\bar{c} < (1-\delta) \sum_{\tau=s}^{\infty} \delta^{\tau-s}\tilde{c}_\tau$ for all $s > t''$. Consumption before t'' remains as under the original policy.

Since we keep the timing of payments to the agent the same, the agent's balance evolves differently under the new contract. We denote these balances by b_s^{new} for all $s \geq 1$. These equal $\tilde{b}_s$ for $s \leq t''$, but differ for $s > t''$. We have for all $s > t''$

$$b_s^{new} + \sum_{\tau=s}^{\infty} \delta^{\tau-s}\tilde{w}_\tau = \frac{\bar{c}}{1-\delta} < \sum_{\tau=s}^{\infty} \delta^{\tau-s}\tilde{c}_\tau = \tilde{b}_s + \sum_{\tau=s}^{\infty} \delta^{\tau-s}\tilde{w}_s.$$

Hence, $b_s^{new} < \tilde{b}_s$ for all $s > t''$.

Now, we want to investigate the agent's constraint (AC$_t^{\text{ob}}$) in each period $s \geq 1$. For all $s \leq t''$, the agent anticipates a strictly higher continuation payoff under the new contract, i.e.

$$\sum_{r=s}^{\infty} \delta^{r-s} \left(v\left(\bar{c}\right) - \psi\left(\tilde{e}_r\right)\right) \geq \sum_{r=s}^{\infty} \delta^{r-s} \left(v\left(\tilde{c}_r\right) - \psi\left(\tilde{e}_r\right)\right),$$

while $v\left(b_s^{new}(1-\delta)\right) = v\left(\tilde{b}_s(1-\delta)\right)$, and hence the agent's constraints continue to be satisfied at these dates.

To understand how the agent's constraints change at $s > t''$, we consider an "intermediate" adjustment to the contract. In particular, we consider adjusting the original contract from date s onwards by specifying optimal/efficient smoothing of consumption from date s onwards. We denote the smoothed consumption $\bar{c}_s\left(\tilde{b}_s\right)$. This is equal to $(1-\delta)\sum_{\tau=s}^{\infty} \delta^{\tau-s}\tilde{c}_\tau = (1-\delta)\left(\tilde{b}_s + \sum_{\tau=s}^{\infty} \delta^{\tau-s}\tilde{w}_\tau\right)$.

Note that, because the agent consuming $\bar{c}_s\left(\tilde{b}_s\right)$ in each period from s onwards improves the agent's payoff relative to the original agreement, we have

$$\sum_{\tau=s}^{\infty} \delta^{\tau-s} \left(v \left(\bar{c}_s \left(\tilde{b}_s \right) \right) - \psi \left(\tilde{e}_\tau \right) \right) \geq \sum_{\tau=s}^{\infty} \delta^{\tau-s} \left(v \left(\tilde{c}_\tau \right) - \psi \left(\tilde{e}_\tau \right) \right) \geq \frac{v \left(\tilde{b}_s \left(1-\delta \right) \right)}{1-\delta}, \tag{36}$$

where the second inequality follows because the agent's constraints (AC_t^{ob}) are satisfied in the original contract.

Because ψ is non-negative, the above inequalities imply $\bar{c}_s \left(\tilde{b}_s \right) \geq \tilde{b}_s \left(1-\delta \right)$. Therefore, since v is concave, we have

$$v \left(\bar{c}_s \left(\tilde{b}_s \right) \right) - v \left(\bar{c}_s \left(\tilde{b}_s \right) - \left(1-\delta \right) \left(\tilde{b}_s - b_s^{new} \right) \right) \leq v \left(\tilde{b}_s \left(1-\delta \right) \right) - v \left(\tilde{b}_s \left(1-\delta \right) - \left(1-\delta \right) \left(\tilde{b}_s - b_s^{new} \right) \right). \tag{37}$$

Note that $\bar{c} = \bar{c}_s \left(\tilde{b}_s \right) - \left(1-\delta \right) \left(\tilde{b}_s - b_s^{new} \right)$. Combining Equations (36) and (37), we have that for $s > t''$,

$$\sum_{\tau=s}^{\infty} \delta^{\tau-s} \left(v \left(\bar{c} \right) - \psi \left(\tilde{e}_\tau \right) \right) \geq \frac{v \left(b_s^{new} \left(1-\delta \right) \right)}{1-\delta}.$$

This shows that, under the new contract, where the consumption is smoothed from date t'' onwards (with the agent consuming $\bar{c}$ per period) and where the agent balances are given by b_s^{new}, the agent's constraints (AC_t^{ob}) are satisfied also at dates following t''.

We have thus defined a new policy with consumption given by $c_s^{new} = \tilde{c}_s$ for $s < t''$ and by $c_s^{new} = \bar{c}$ for $s \geq t''$. Also the payments w_s^{new} and efforts e_s^{new} are defined to equal $\tilde{w}_s$ and $\tilde{e}_s$ for all s. Finally, the agent's balances are given by $b_s^{new} = \tilde{b}_s$ for $s \leq t''$ and by b_s^{new} being implied by the payment schedule and consumption for $s > t''$ as noted above. We have, for all s, the agent's constraint is satisfied. Furthermore, the agent's constraint is slack for $s \leq t''$ (this follows because $\sum_{\tau=t''}^{\infty} \delta^{\tau-t''} v \left(c_\tau^{new} \right) = \sum_{\tau=t''}^{\infty} \delta^{\tau-t''} v \left(\bar{c} \right) > \sum_{\tau=t''}^{\infty} \delta^{s'-t} v \left(\tilde{c}_\tau \right)$ by strict concavity of v).

Finally, we can define a contract that further adjusts the "new" one defined by $\left(e_s^{new}, c_s^{new}, w_s^{new}, b_s^{new} \right)_{s \geq 1}$ by reducing $c_{t''}^{new}$ and $w_{t''}^{new}$ by a small enough ε that the agent's constraint at date t'', and all earlier dates, continues to be satisfied. At other dates, the contract remains unchanged. The principal obtains a strictly higher payoff in this contract than in the original, contradicting the optimality of the original. □

Lemma A.12. *$(\tilde{b}_t)_{t \geq 1}$ is a weakly decreasing sequence. It is constant when the first-best payoff is achievable at b_1, and strictly decreasing towards some $\tilde{b}_\infty > 0$ otherwise. Also, $V(\tilde{b}_\infty) = V^{FB}(\tilde{b}_\infty)$.*

Proof. **Step 0.** If the first-best payoff is achievable at b_1, then equilibrium consumption and effort is uniquely determined by the conditions in Proposition 3.1. Because we restrict attention to payments timed to satisfy Equation (7), the balance is constant as claimed in the lemma (and explained in the main text). Suppose from now on that $V(b_1) < V^{FB}(b_1)$.

Step 1. Proof that $(\tilde{b}_t)_{t\geq 1}$ is weakly decreasing. Suppose that $\tilde{b}_{\hat{t}+1} > \tilde{b}_{\hat{t}}$ for some date $\hat{t}$. We construct a self-enforceable contract that achieves strictly higher profits for the principal.

Step 1a. First, denote a new contract by $(\tilde{e}'_t, \tilde{c}'_t, \tilde{w}'_t, \tilde{b}'_t)_{t\geq 1}$, which we will choose to coincide with the original contract until $\hat{t}-1$, and with $\tilde{e}'_{\hat{t}} = \tilde{e}_{\hat{t}}$. For dates $t \geq \hat{t}$, let

$$\begin{aligned}\tilde{c}'_t = \bar{c} &\equiv (1-\delta)\sum_{\tau\geq\hat{t}} \delta^{\tau-\hat{t}}\tilde{w}_\tau + (1-\delta)\tilde{b}_{\hat{t}} \\ &= (1-\delta)\sum_{\tau\geq\hat{t}} \delta^{\tau-\hat{t}}\tilde{c}_\tau.\end{aligned}$$

For dates $t \geq \hat{t}+1$, let $\tilde{e}'_t = \bar{e}$, where $\bar{e}$ is defined by

$$\psi(\bar{e}) = (1-\delta)\sum_{\tau\geq\hat{t}+1} \delta^{\tau-\hat{t}-1}\psi(\tilde{e}_\tau).$$

Let also, for all $t \geq \hat{t}$, $\tilde{w}'_t = \bar{w}$, where $\bar{w} = (1-\delta)\sum_{\tau\geq\hat{t}} \delta^{\tau-\hat{t}}\tilde{w}_\tau$. Thus, we must have $\tilde{b}'_t = \bar{b} \equiv \tilde{b}_{\hat{t}}$ for all $t \geq \hat{t}$.

Step 1b. We now want to show that, for the new contract, the agent's constraint (AC_t^{ob}) is satisfied at all dates. Note that the new contract is stationary from date $\hat{t}+1$ onwards. Let's then consider the agent's constraint for these dates. Note first that, by the previous lemma, we must have $\tilde{c}_{\hat{t}} \geq \bar{c}$. Therefore,

$$\sum_{\tau\geq\hat{t}+1} \delta^{\tau-\hat{t}-1}\bar{c} \geq \sum_{\tau\geq\hat{t}+1} \delta^{\tau-\hat{t}-1}\tilde{c}_\tau.$$

Also, the NPV of disutility of effort from date $\hat{t}+1$ onwards is the same for both policies. The fact that the original policy satisfies the agent's constraint (AC_t^{ob}) at date $\hat{t}+1$, plus the observation that $\bar{b} < b_{\hat{t}+1}$, then implies

$$\sum_{\tau\geq\hat{t}+1} \delta^{\tau-\hat{t}-1}v(\bar{c}) - \sum_{\tau\geq\hat{t}+1} \delta^{\tau-\hat{t}-1}\psi(\bar{e}) > \frac{v\left((1-\delta)\,\bar{b}\right)}{1-\delta}, \tag{38}$$

which means that the agent's constraint is satisfied as a *strict inequality* from $\hat{t}+1$ onwards.

Note then that

$$\sum_{\tau\geq\hat{t}} \delta^{\tau-\hat{t}}v(\bar{c}) \geq \sum_{\tau\geq\hat{t}} \delta^{\tau-\hat{t}}v(\tilde{c}_\tau)$$

(with a strict inequality if the consumption levels $\tilde{c}_\tau$ for $\tau \geq \hat{t}$ are non-constant). Also, the NPV of the disutility of effort is the same from $\hat{t}$ onwards under both policies. Therefore, the agent's constraint continues to be satisfied at $\hat{t}$, and by the same logic all earlier periods.

Step 1c. Now we show that the principal's constraint (PC_t) is satisfied in all periods. Because the NPV of disutility of effort from date $\hat{t}+1$ onwards is the same under both policies; and because ψ is convex, we have $\bar{e} \geq (1-\delta)\sum_{\tau \geq \hat{t}+1} \delta^{\tau-\hat{t}-1}\tilde{e}_\tau$. Therefore,

$$\begin{aligned}
\sum_{\tau \geq \hat{t}+1} \delta^{\tau-\hat{t}}\tilde{e}'_\tau - \sum_{\tau \geq \hat{t}} \delta^{\tau-\hat{t}}\tilde{w}'_\tau &= \frac{\delta \bar{e}}{1-\delta} - \sum_{\tau \geq \hat{t}} \delta^{\tau-\hat{t}}\tilde{w}_\tau \\
&\geq \sum_{\tau \geq \hat{t}+1} \delta^{\tau-\hat{t}}\tilde{e}_\tau - \sum_{\tau \geq \hat{t}} \delta^{\tau-\hat{t}}\tilde{w}_\tau \\
&\geq 0 \qquad (39)
\end{aligned}$$

where the second inequality holds because the principal's constraint is satisfied at date $\hat{t}$ under the original policy. Hence the principal's constraint is satisfied under the new policy at date $\hat{t}$. Because $\tilde{e}'_t$ is constant for $t \geq \hat{t}+1$, and because $\tilde{w}'_t$ is constant for $t \geq \hat{t}$, the same inequality implies the satisfaction of the principal's constraint also from $\hat{t}+1$ onwards. Checking that the principal's constraint is satisfied also at dates before $\hat{t}$ follows the same logic. For $t < \hat{t}$, the principal's constraint is

$$\sum_{\tau=t+1}^{\hat{t}} \delta^{\tau-t}\tilde{e}'_\tau - \sum_{\tau=t}^{\hat{t}-1} \delta^{\tau-t}\tilde{w}'_\tau + \delta^{\hat{t}-t}\left(\sum_{\tau \geq \hat{t}+1} \delta^{\tau-\hat{t}}\tilde{e}'_\tau - \sum_{\tau \geq \hat{t}} \delta^{\tau-\hat{t}}\tilde{w}'_\tau\right) \geq 0,$$

which is satisfied because (i) $\tilde{e}'_\tau = \tilde{e}_\tau$ for $\tau \leq \hat{t}$, and $\tilde{w}'_\tau = \tilde{w}_\tau$ for $\tau < \hat{t}$, (ii) the first inequality in Equation (39) holds, and (iii) the principal's constraint is satisfied at date t under the original policy.

Step 1d. Finally, we show that the contract can be further (slightly) adjusted to a self-enforcing contract with a strictly higher payoff for the principal. The original contract was taken to satisfy

$$v\left(\tilde{c}_{\hat{t}}\right) - \psi\left(\tilde{e}_{\hat{t}}\right) = \frac{v\left((1-\delta)\bar{b}\right) - \delta v\left((1-\delta)\tilde{b}_{\hat{t}+1}\right)}{1-\delta} < v\left((1-\delta)\bar{b}\right).$$

Hence,

$$\psi\left(\tilde{e}_{\hat{t}}\right) > v\left(\tilde{c}_{\hat{t}}\right) - v\left((1-\delta)\bar{b}\right) \geq v\left(\bar{c}\right) - v\left((1-\delta)\bar{b}\right) > \psi\left(\bar{e}\right)$$

where the final inequality follows from (38). Hence $\tilde{e}_{\hat{t}} > \bar{e}$. Recall that $\tilde{e}'_{\hat{t}} = \tilde{e}_{\hat{t}}$, and $\tilde{e}'_\tau = \bar{e}$ for $\tau > \hat{t}$; so we have $\tilde{e}'_{\hat{t}} > \tilde{e}'_\tau$ for all $\tau > \hat{t}$.

Now, pick $\tilde{e}''_{\hat{t}}$ and $\tilde{e}''_{\hat{t}+1}$, with

$$\tilde{e}'_{\hat{t}+1} < \tilde{e}''_{\hat{t}+1} < \tilde{e}''_{\hat{t}} < \tilde{e}'_{\hat{t}}$$

and such that

$$\psi(\tilde{e}''_{\hat{t}}) + \frac{\delta}{1-\delta}\psi(\tilde{e}''_{\hat{t}+1}) = \psi(\tilde{e}'_{\hat{t}}) + \frac{\delta}{1-\delta}\psi(\tilde{e}'_{\hat{t}+1}).$$

Substitute $\tilde{e}''_{\hat{t}}$ for $\tilde{e}'_{\hat{t}}$ and $\tilde{e}''_{\hat{t}+1}$ for $\tilde{e}'_{\tau}$, for all $\tau \geq \hat{t}+1$, in the contract defined in Step 1a. The agent's value from remaining in the contract from $\hat{t}$ onwards remains unchanged, so the agent's constraint (AC_t^{ob}) remains satisfied at $\hat{t}$, and at all earlier dates. Note that, due to (38), the agent's constraints (AC_t^{ob}) at dates $\hat{t}+1$ onwards are slack under the contract defined in Step 1a, and hence continue to be satisfied under the contract with the further modification, provided the adjustment in effort is small. Moreover, because ψ is strictly convex, the NPV of effort from date $\hat{t}$ onwards increases; so the principal's payoff strictly increases. Also, the principal's constraints (PC_t) clearly continue to be satisfied.

Step 2. Proof that if $V(\tilde{b}_1) < V^{FB}(\tilde{b}_1)$ then $(\tilde{b}_t)_{t\geq 1}$ is a strictly decreasing sequence.

Step 2a. We first prove that if $\tilde{b}_t = \tilde{b}_{t+1}$ then $V(\tilde{b}_t) = V^{FB}(\tilde{b}_t)$. To do this, note that if $\tilde{b}_t = \tilde{b}_{t+1}$, then it is optimal to specify $\tilde{c}_\tau = \tilde{c}_t$, $\tilde{w}_\tau = \tilde{w}_t$, and $\tilde{e}_\tau = \tilde{e}_t$ for all $\tau > t$; that is, it must be optimal for the contract to be stationary from period t onwards. The Euler equation (33) then requires that $\psi'(\tilde{e}_\tau) = v'(\tilde{c}_\tau)$ for all $\tau \geq t+1$,[16] and by stationarity also $\psi'(\tilde{e}_t) = v'(\tilde{c}_t)$. Then, $\tilde{e}_\tau$ and $\tilde{c}_\tau$ satisfy, for all $\tau \geq t$, the first-order and agent's indifference conditions in Proposition 3.1, given initial balance $\tilde{b}_t$. Therefore they are the first-best effort and consumption given balance $\tilde{b}_t$. This shows that $V(\tilde{b}_t) = V^{FB}(\tilde{b}_t)$, as desired.

Step 2b. We now prove that if $V(\tilde{b}_1) < V^{FB}(\tilde{b}_1)$ then $V(\tilde{b}_t) < V^{FB}(\tilde{b}_t)$ for all $t \geq 1$ and, in addition, $(\tilde{b}_t)_{t\geq 1}$ is strictly decreasing. Suppose that $V(\tilde{b}_t) < V^{FB}(\tilde{b}_t)$, which by Step 1 and Step 2a implies $\tilde{b}_{t+1} < \tilde{b}_t$. Suppose for a contradiction that an optimal relational contract achieves the first-best continuation payoff for the principal at date $t+1$, when the balance is $\tilde{b}_{t+1}$. This implies that $\tilde{e}_\tau = e^{FB}(\tilde{b}_{t+1})$ and $\tilde{c}_\tau = c^{FB}(\tilde{b}_{t+1})$ for all $\tau > t$. By assumption that the agent's constraint (AC_t^{ob}) is satisfied with equality in all periods, we then have $\tilde{b}_\tau = \tilde{b}_{t+1}$ for all $\tau > t+1$. Hence, the contract is stationary from $t+1$ onwards; in particular, the payment is constant at $\tilde{w}_\tau = \bar{w}$ for $\tau \geq t+1$, for some value $\bar{w}$.

From the Euler equation (33) and the fact that $v'(\tilde{c}_{t+1}) = \psi'(\tilde{e}_{t+1})$, we have $\tilde{c}_t = \tilde{c}_{t+1}$. Hence, using $\tilde{b}_{t+2} = \tilde{b}_{t+1} < \tilde{b}_t$, we have (using ($\text{FP}_t^{\text{ob}}$))

$$\begin{aligned}\psi(\tilde{e}_t) &= v(\tilde{c}_t) + \tfrac{\delta}{1-\delta}v((1-\delta)\tilde{b}_{t+1}) - \tfrac{1}{1-\delta}v((1-\delta)\tilde{b}_t)\\ &< v(\tilde{c}_{t+1}) + \tfrac{\delta}{1-\delta}v((1-\delta)\tilde{b}_{t+2}) - \tfrac{1}{1-\delta}v((1-\delta)\tilde{b}_{t+1}) = \psi(\tilde{e}_{t+1}).\end{aligned}$$

[16] To see this, recall from Lemma A.11 that $\tilde{c}_\tau > \tilde{b}_\tau$.

Consequently, $\tilde{e}_t < \tilde{e}_{t+1}$, and so $\frac{\psi'(\tilde{e}_t)}{v'(\tilde{c}_t)} < \frac{\psi'(\tilde{e}_{t+1})}{v'(\tilde{c}_{t+1})} = 1$. We then know (from Lemma A.10) that the principal's constraint (PC$_t$) binds at t, and so

$$\sum_{s=t+1}^{\infty} \delta^{s-t} \tilde{e}_s = \sum_{s=t}^{\infty} \delta^{s-t} \tilde{w}_s = \sum_{s=t}^{\infty} \delta^{s-t} \tilde{c}_s - \tilde{b}_t.$$

Using that $\tilde{e}_\tau = e^{FB}(\tilde{b}_{t+1})$ for all $\tau \geq t+1$, and $\tilde{c}_\tau = c^{FB}(\tilde{b}_{t+1})$ for all $\tau \geq t$, we can write this condition as

$$\delta e^{FB}(\tilde{b}_{t+1}) = c^{FB}(\tilde{b}_{t+1}) - (1-\delta)\tilde{b}_t < c^{FB}(\tilde{b}_{t+1}) - (1-\delta)\tilde{b}_{t+1} = \bar{w}.$$

This equation implies that the principal's constraint (PC$_t$) in period $t+1$ (as well as at future dates) is violated, so we reach a contradiction.

Step 3. Proof that $\tilde{b}_\infty > 0$ and $V(\tilde{b}_\infty) = V^{FB}(\tilde{b}_\infty)$.

Since $\left(\tilde{b}_t\right)_{t\geq 1}$ is a decreasing sequence, the limit $\lim_{t\to\infty} \tilde{b}_t$ exists. We want to show this limit, call it $\tilde{b}_\infty$, is strictly positive and $V(\tilde{b}_\infty) = V^{FB}(\tilde{b}_\infty)$.

Step 3a. We first show that the function V is continuous at any $b > 0$. Suppose there is a point of discontinuity $\check{b} > 0$. Then there is $\varepsilon > 0$ and a sequence $(b_n)_{n=1}^{\infty}$ convergent to $\check{b}$ with $|V(b_n) - V(\check{b})| \geq \varepsilon$ for all n. Let $\check{c}$ be an optimal consumption choice when the balance is $\check{b}$, let $\check{b}'$ be a corresponding optimal next-period balance, and let $\hat{e}(\check{c}, \check{b}, \check{b}')$ be the corresponding effort. Let c_n be an optimal consumption when the balance is b_n, let b_n' be a corresponding next-period balance, and let $\hat{e}(c_n, b_n, b_n')$ be the corresponding effort. Then, for n large enough, if $V(b_n) \leq V(\check{b}) - \varepsilon$, we reach a contradiction because the principal's payoff at balance b_n is at least that obtained by specifying consumption $\check{c}$, next-period balance $\check{b}'$ and effort $\hat{e}(\check{c}, b_n, \check{b}')$ (which is strictly positive since $\hat{e}(\check{c}, \check{b}, \check{b}') > 0$ and provided b_n is sufficiently close to $\check{b}$), and then specifying an optimal continuation policy given the next period balance is $\check{b}'$ rather than b_n'. If instead $V(b_n) \geq V(\check{b}) + \varepsilon$, we reach a contradiction because the principal's payoff at balance $\check{b}$ is at least that obtained by specifying consumption c_n, next-period balance b_n', effort $\hat{e}(c_n, \check{b}, b_n')$ (which is strictly positive since $\hat{e}(c_n, b_n, b_n') > 0$ and provided $\check{b}$ is close enough to b_n), and then specifying a continuation policy that is optimal given the next period balance is b_n' rather than $\check{b}'$.

Step 3b. We now prove that $\tilde{b}_\infty > 0$. For this, we first show $\lim_{b\searrow 0} \frac{c^{FB}(b)-(1-\delta)b}{e^{FB}(b)} = 0$ and so, by Equation (8), there exists some $\bar{b} > 0$ such that an optimal contract achieves the first-best payoff of the principal for all $b \leq \bar{b}$. This follows after noting that $v(c^{FB}(b)) - v((1-\delta)b) = \psi(e^{FB}(b)) > 0$, so we have that either $\lim_{b\searrow 0} c^{FB}(b) = 0$ or $\lim_{b\searrow 0} e^{FB}(b) = +\infty$. Since $\psi'(e^{FB}(b)) = v'(c^{FB}(b))$ we have, in fact, that both $\lim_{b\searrow 0} c^{FB}(b) = 0$ and $\lim_{b\searrow 0} e^{FB}(b) =$

$+\infty$, which establishes the result. Next, recall from Step 2 that, given $V(b_1) < V^{FB}(b_1)$, the sequence $\left(\tilde{b}_t\right)_{t\geq 1}$ of balances in the optimal contract is strictly decreasing and such that $V\left(\tilde{b}_t\right) < V^{FB}\left(\tilde{b}_t\right)$ for all t. That is, $\tilde{b}_t$ remains above $\bar{b}$, and so converges to some value $\tilde{b}_\infty \geq \bar{b}$.

Step 3c. We finally prove that $V(\tilde{b}_\infty) = V^{FB}(\tilde{b}_\infty)$. Recall we assumed that $V(b_1) < V^{FB}(b_1)$. By the continuity of V established in Step 3a, we have that $\lim_{t\to\infty} V\left(\tilde{b}_t\right) = V\left(\tilde{b}_\infty\right)$. Because the principal's constraint (PC_t) binds for all t, we have $V\left(\tilde{b}_t\right) = \hat{e}(\tilde{c}_t, \tilde{b}_t, \tilde{b}_{t+1})$ for all t. By continuity of $\hat{e}(\cdot,\cdot,\cdot)$, we have $\lim_{t\to\infty} \hat{e}(\tilde{c}_t, \tilde{b}_t, \tilde{b}_{t+1}) = \hat{e}(\tilde{c}_\infty, \tilde{b}_\infty, \tilde{b}_\infty)$, where $\tilde{c}_\infty \equiv \lim_{t\to\infty} \tilde{c}_t$, which exists because $\tilde{c}_t$ is decreasing and remains above $\tilde{b}_\infty$ by Lemma A.11. Therefore,

$$V\left(\tilde{b}_\infty\right) = \hat{e}(\tilde{c}_\infty, \tilde{b}_\infty, \tilde{b}_\infty) = \psi^{-1}(v(\tilde{c}_\infty) - v((1-\delta)\tilde{b}_\infty)).$$

Since $V(\tilde{b}_\infty) > 0$ (recall Lemma (A.9)), $\tilde{c}_\infty > (1-\delta)\tilde{b}_\infty$. Therefore, the Euler equation (33) implies that, necessarily, $\lim_{t\to\infty} \frac{v'(\tilde{c}_{t+1})}{\psi'(\tilde{e}_{t+1})} = 1$, and therefore $\tilde{e}_\infty \equiv \lim_{t\to\infty} \tilde{e}_t$ exists. It is then clear that both Conditions 1 and 2 of Proposition 3.1 hold for $\tilde{e}_\infty$, $\tilde{c}_\infty$, and $\tilde{b}_\infty$ (instead of $e^{FB}(b_1)$, $c^{FB}(b_1)$, and b_1). This establishes the result. □

Lemma A.13. *Assume $V(b_1) < V^{FB}(b_1)$. Then $\left(V\left(\tilde{b}_t\right)\right)_{t\geq 1}$ is a strictly increasing sequence.*

Proof. Recall from Lemma A.12 we have that, if $V(b_1) < V^{FB}(b_1)$, then $(\tilde{b}_t)_{t\geq 1}$ is strictly decreasing. Therefore, the result will follow if we can show $V(\cdot)$ is strictly decreasing.

Step 1. We show that if $V(\cdot)$ fails to be strictly decreasing, then there exists a value $b^* > 0$ such that, for every $\varepsilon > 0$, there is a $\check{b} \in (b^* - \varepsilon, b^*)$ satisfying $V(\check{b}) \leq V(b^*)$.

First, by Step 3a of the proof of the previous lemma, $V(\cdot)$ is continuous on strictly positive values. Suppose $V(\cdot)$ fails to be strictly decreasing, which means that there are values b', b'' with $0 < b' < b''$, and with $V(b') \leq V(b'')$. Consider maximizing V on $[b', b'']$. If the maximum (which exists by continuity of V) is $V(b'')$, then we may take $b^* = b''$. If the maximum is greater than $V(b'')$, then we may take any maximizer in $(b', b'']$ to be b^*.

Step 2. Consider the optimal continuation contract when $\tilde{b}_t = b^*$, and consider a change to $\tilde{b}_t = b^* - \nu$ for ν arbitrarily small and such that $V(b^* - \nu) \leq V(b^*)$. Then we can reduce $\tilde{c}_t$ by the same amount ν, holding $\tilde{b}_{t+1}$ and $\tilde{w}_t$, as well as all other variables, constant. Note then that, provided ν is small enough,

$$v(\tilde{c}_t - \nu) - \tfrac{1}{1-\delta} v((1-\delta)(\tilde{b}_t - \nu)) > v(\tilde{c}_t) - \tfrac{1}{1-\delta} v((1-\delta)\tilde{b}_t),$$

which follows again because $\tilde{c}_t > (1-\delta)\,\tilde{b}_t$ (by Lemma A.11) and by concavity of v. Hence, we have

$$\hat{e}(\tilde{c}_t - \nu, \tilde{b}_t - \nu, \tilde{b}_{t+1}) > \tilde{e}(\tilde{c}_t, \tilde{b}_t, \tilde{b}_{t+1}).$$

By construction, the agent remains indifferent to continuing in the contract at all dates (we have that (AC_t^{ob}) holds as an equality at all dates). The continuation of the relationship from $t+1$ onwards is precisely as before, and therefore the principal's constraint at date t is satisfied (since $\tilde{w}_t$ is unchanged). Hence,

$$V(\tilde{b}_t) = \tilde{e}_t - \tilde{w}_t + \delta V(\tilde{b}_{t+1}) < \hat{e}(\tilde{c}_t - \nu, \tilde{b}_t - \nu, \tilde{b}_{t+1}) - \tilde{w}_t + \delta V(\tilde{b}_{t+1}) \leq V(\tilde{b}_t - \nu).$$

However, this contradicts $V\left(b^* - \nu\right) \leq V\left(b^*\right)$. □

Lemma A.14. *Assume $V(b_1) < V^{FB}(b_1)$. Then, in an optimal contract $\left(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t\right)_{t\geq 1}$, $v'\left(\tilde{c}_t\right) > \psi'\left(\tilde{e}_t\right)$ for all t.*

Proof. We show that $v'\left(\tilde{c}_{\check{t}}\right) = \psi'\left(\tilde{e}_{\check{t}}\right)$ at some $\check{t}$ implies $v'\left(\tilde{c}_{\check{t}+1}\right) = \psi'\left(\tilde{e}_{\check{t}+1}\right)$. Otherwise, we have by Lemma A.10 that $v'\left(\tilde{c}_{\check{t}}\right) = \psi'\left(\tilde{e}_{\check{t}}\right)$ and $v'\left(\tilde{c}_{\check{t}+1}\right) > \psi'\left(\tilde{e}_{\check{t}+1}\right)$, and the latter implies $\tilde{w}_{\check{t}+1} = \delta V(\tilde{b}_{\check{t}+2})$. In turn, this implies

$$\begin{aligned}
\tilde{e}_{\check{t}+1} &= \tilde{e}_{\check{t}+1} - \tilde{w}_{\check{t}+1} + \delta V(\tilde{b}_{\check{t}+2}) \\
&= V\left(\tilde{b}_{\check{t}+1}\right) \\
&> V\left(\tilde{b}_{\check{t}}\right) \\
&= \tilde{e}_{\check{t}} - \tilde{w}_{\check{t}} + \delta V(\tilde{b}_{\check{t}+1}) \\
&\geq \tilde{e}_{\check{t}}.
\end{aligned}$$

There are two cases: either $\tilde{w}_{\check{t}} < \sum_{s=\check{t}+1}^{\infty} \delta^{s-\check{t}}\left(\tilde{e}_s - \tilde{w}_s\right)$ or $\tilde{w}_{\check{t}} = \sum_{s=\check{t}+1}^{\infty} \delta^{s-\check{t}}\left(\tilde{e}_s - \tilde{w}_s\right)$. Consider the first. Define a new contract $\left(\tilde{e}'_t, \tilde{c}'_t, \tilde{w}'_t, \tilde{b}'_t\right)_{t\geq 1}$, which is identical to the original, except that $\tilde{e}'_{\check{t}} = \tilde{e}_{\check{t}} + \varepsilon$ and $\tilde{e}'_{\check{t}+1} = \tilde{e}_{\check{t}+1} - \nu\left(\varepsilon\right)$, with $\nu\left(\varepsilon\right)$ defined by

$$\psi\left(\tilde{e}_{\check{t}} + \varepsilon\right) + \delta\psi\left(\tilde{e}_{\check{t}+1} - \nu\left(\varepsilon\right)\right) = \psi\left(\tilde{e}_{\check{t}}\right) + \delta\psi\left(\tilde{e}_{\check{t}+1}\right).$$

Thus

$$\nu'\left(\varepsilon\right) = \frac{\psi'\left(\tilde{e}_t\right)}{\delta\psi'\left(\tilde{e}_{t+1}\right)}$$

and so the change in the NPV of effort is

$$\varepsilon - \delta\nu(\varepsilon) = \left(1 - \frac{\psi'(\tilde{e}_t)}{\psi'(\tilde{e}_{t+1})}\right)\varepsilon + o(\varepsilon)$$

which is strictly positive for ε sufficiently small. It is easy to see that the agent's constraint (AC$_t^{ob}$) is unchanged at all dates except $\check{t}' + 1$, when the constraint is relaxed. The principal's constraint (PC$_t$) is unchanged from date $\check{t} + 1$ onwards, relaxed at date $\check{t} - 1$ and earlier (because the NPV of effort increases), but is tightened at date $\check{t}$. Provided ε is small enough, the date-$\check{t}$ constraint remains intact. Profits increase, contradicting the optimality of the original contract.

Now suppose that $\tilde{w}_{\check{t}} = \sum_{s=\check{t}+1}^{\infty} \delta^{s-\check{t}} (\tilde{e}_s - \tilde{w}_s)$, and note the above adjustment now leads to a violation of the principal's constraint at date $\check{t}$. To ensure the constraint is satisfied, we further adjust the modified contract by subtracting $\eta(\varepsilon)$ from date $\check{t}$ effort, and reducing the date $\check{t}$ payment and consumption by an amount $\gamma(\varepsilon)$ to leave agent payoffs unchanged. The date-$\check{t}$ principal constraint (PC$_t$) then holds as equality if we set $\gamma(\varepsilon) = \delta\nu(\varepsilon)$. We have

$$v(\tilde{c}'_{\check{t}} - \gamma(\varepsilon)) - \psi(\tilde{e}'_{\check{t}} - \eta(\varepsilon)) = v(\tilde{c}'_{\check{t}}) - \psi(\tilde{e}'_{\check{t}})$$

and so

$$v'(\tilde{c}'_{\check{t}})\gamma'(\varepsilon) = \psi'(\tilde{e}'_{\check{t}})\eta'(\varepsilon).$$

The effect on date-$\check{t}$ profits is to reduce them by

$$\begin{aligned} \eta(\varepsilon) - \gamma(\varepsilon) &= (\eta'(\varepsilon) - \gamma'(\varepsilon))\varepsilon + o(\varepsilon) \\ &= \eta'(\varepsilon)\left(1 - \frac{\psi'(\tilde{e}'_{\check{t}})}{v'(\tilde{c}'_{\check{t}})}\right)\varepsilon + o(\varepsilon) \\ &= o(\varepsilon). \end{aligned}$$

Hence, for small enough ε, the overall effect on profits relative to the original contract $\left(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t\right)_{t\geq 1}$, is positive, with the continuation profits from date $\check{t}$ increasing. The principal's constraint (PC$_t$) is relaxed at dates $\check{t} - 1$ and earlier, it is satisfied by construction at $\check{t}$, and it is unchanged from date $\check{t} + 1$ onwards. Again, the fact profits strictly increase contradicts the optimality of $\left(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t\right)_{t\geq 1}$.

We have shown then that $v'(\tilde{c}_{\check{t}}) = \psi'(\tilde{e}_{\check{t}})$ at some $\check{t}$ implies $v'(\tilde{c}_t) = \psi'(\tilde{e}_t)$ for all $t \geq \check{t}$. By the Euler equation (33), consumption and effort remain constant from $\check{t}$ onwards. However, the fact that the agent's balance must strictly decrease over time contradicts that Equation (FP$_t^{ob}$) is presumed to hold at all dates. Hence, we do not have $v'(\tilde{c}_t) = \psi'(\tilde{e}_t)$ at any t, establishing the result.

□

Lemma A.14 implies by Lemma A.10 that, if $V(b_1) < V^{FB}(b_1)$, the principal's constraint (PC_t) holds with equality in every period. We use this to show the following.

Lemma A.15. *If $V(b_1) < V^{FB}(b_1)$, then, in any optimal contract $\left(\tilde{e}_t, \tilde{c}_t, \tilde{w}_t, \tilde{b}_t\right)_{t\geq 1}$, effort $\tilde{e}_t$ and payments $\tilde{w}_t$ strictly increase over time, while consumption $\tilde{c}_t$ strictly declines over time.*

Proof. The fact that the principal's constraint binds at every date, as argued above, can be stated as $\tilde{w}_t = \delta V\left(\tilde{b}_{t+1}\right)$ for all t; hence payments are strictly increasing in t by Lemma A.13. We also have $V\left(\tilde{b}_t\right) = \tilde{e}_t$ for all t, so effort is strictly increasing as well.

Now consider consumption. By Lemma A.11, we know that $\tilde{c}_{t-1} \geq \tilde{c}_t$ for all $t \geq 2$. Hence, if consumption fails to be strictly decreasing, we must have $\tilde{c}_{t-1} = \tilde{c}_t$ for some t. We then have, by Equation (33) (and noting that $\tilde{c}_t > (1-\delta)\,\tilde{b}_t$, also by Lemma A.11), that $\psi'(\tilde{e}_t) = v'(\tilde{c}_t)$. However, this contradicts Lemma A.14. $\square$

Lemma A.16. *An optimal contract exists.*

Proof. If $\delta \geq \frac{c^{FB}(b_t)-(1-\delta)b_t}{e^{FB}(b_t)} \in (0,1)$ then there is a self-enforceable efficient contract (by Proposition 5.2), and so existence is established. The remainder of the proof is needed for the values b_1 such that there is no self-enforceable first-best contract.

We denote by $\Pi(b_t)$ the sequences $(c_s, b_{s+1})_{s=t}^{\infty}$ that satisfy, for all $s \geq t$,

$$\delta b_{s+1} - b_s + c_s \leq \sum_{\tau=s+1}^{\infty} \delta^{\tau-s}\left(\hat{e}(c_\tau, b_\tau, b_{\tau+1}) - (\delta b_{\tau+1} - b_\tau + c_\tau)\right)$$

as well as

$$v(c_s) + \tfrac{\delta}{1-\delta} v\left((1-\delta)\, b_{s+1}\right) - \tfrac{1}{1-\delta} v((1-\delta) b_s) \geq 0.$$

Given any $b_t > 0$, let

$$V(b_t) = \sup_{(c_s, b_{s+1})_{s=t}^{\infty} \in \Pi(b_t)} \sum_{s=t}^{\infty} \delta^{s-t}\left(\hat{e}(c_s, b_s, b_{s+1}) - (\delta b_{s+1} - b_s + c_s)\right).$$

We can write the functional equation for the problem as

$$TW(b_t) = \sup_{c_t > 0,\, b_{t+1} > 0} \left(\hat{e}(c_t, b_t, b_{t+1}) - (\delta b_{t+1} - b_t + c_t) + \delta W(b_{t+1})\right) \tag{40}$$

subject to the principal's constraint

$$\delta b_{t+1} - b_t + c_t \leq \delta W(b_{t+1}) \tag{41}$$

and to

$$v(c_t) + \tfrac{\delta}{1-\delta} v\left((1-\delta)\, b_{t+1}\right) - \tfrac{1}{1-\delta} v((1-\delta) b_t) \geq 0. \tag{42}$$

Outline of Proof. Note that the operator T is monotone: if $W_1 \geq W_2$, then $TW_1 \geq TW_2$. Also, we have $TV^{FB} \leq V^{FB}$. Applying T to both sides, we have that $(T^n V^{FB}(b_t))_{n\geq 1}$ is a decreasing sequence for all $b_t > 0$. Therefore there is some pointwise limit of $T^n V^{FB}$, call it $\bar{V}$. Straightforward continuity arguments show that $\bar{V}$ is a fixed point of T.

We want to show that $\bar{V}(b_t) = V(b_t)$ and that this payoff is attained by a feasible policy $(\tilde{e}_s, \tilde{c}_s, \tilde{w}_s, \tilde{b}_s)_{s\geq t}$, that respects the principal and agent constraints, (PC_t) and (AC_t^{ob}). We first establish (in Step 1) the existence of a feasible policy that attains the payoff $\bar{V}(b_t)$ for the principal; this establishes that $V(b_t) \geq \bar{V}(b_t)$. We then show that V is a fixed point of T (see Step 2 below). Then the fact that $V \leq V^{FB}$, together with the fact that T is monotone, implies

$$V(b_t) = \lim_{n\to\infty} T^n V(b_t) \leq \lim_{n\to\infty} T^n V^{FB}(b_t) = \bar{V}(b_t)$$

which completes the proof.

Step 1. Determining a policy from $\bar{V}$: We want to show that the supremum in the problem defined by Equations (40) to (42) for $W = \bar{V}$ is attained by some values c_t and b_{t+1} at each $b_t > 0$. By analogous arguments to Step 3a of the proof of Lemma A.12, we have that $\bar{V}$ is continuous. Therefore, our supremum will be attained if (a) the values of b_{t+1} that satisfy the constraints of the functional equation are contained in a bounded interval $\left[l^b(b_t), u^b(b_t)\right]$ with $l^b(b_t) > 0$, and (b) consumption is also constrained to come from a bounded interval $[l^c(b_t), u^c(b_t)]$ with $l^c(b_t) > 0$.

We begin with Part (a). Observe that

$$\lim_{b_{t+1}\to\infty} \delta(\bar{V}(b_{t+1}) - b_{t+1}) \leq \lim_{b_{t+1}\to\infty} \delta(V^{FB}(b_{t+1}) - b_{t+1}) = -\infty\ .$$

Hence, the satisfaction of the principal's constraint (41) implies b_{t+1} must be bounded above by some $u^b(b_t)$.

We now show that, given b_t, satisfaction of the constraints in Equations (41) and (42) implies that b_{t+1} must be no less than some $l^b(b_t) > 0$. Assume, for the sake of contradiction, that b_{t+1} can be taken arbitrarily close to 0, given b_t, without violating either of these constraints.

In particular, consider $b_{t+1} < \bar{b}$, where $\bar{b} > 0$ is such that $\bar{V}(b) = V^{FB}(b)$ for all $b \in (0, \bar{b}]$ (note that it exists by Step 3b of the proof of Lemma A.12). These constraints may be written

$$v(c_t) \geq \tfrac{1}{1-\delta} v((1-\delta)b_t) - \tfrac{\delta}{1-\delta} v((1-\delta)b_{t+1}) \quad \text{and} \quad c_t \leq b_t + \delta(V^{FB}(b_{t+1}) - b_{t+1}).$$

Combining these two equations we have

$$V^{FB}(b_{t+1}) \geq \tilde{V}(b_{t+1}) \equiv \frac{v^{-1}\left(\tfrac{1}{1-\delta} v((1-\delta)b_t) - \tfrac{\delta}{1-\delta} v((1-\delta)b_{t+1})\right) - b_t}{\delta} + b_{t+1}. \tag{43}$$

Now, notice that the right-hand side of Equation (43) tends to $+\infty$ as $b_{t+1} \to 0$. Hence, if the constraints are satisfied, we must have $\lim_{b_{t+1}\to 0} V^{FB}(b_{t+1}) = +\infty$ and

$$\lim_{b_{t+1}\to 0} \frac{\tilde{V}(b_{t+1})}{V^{FB}(b_{t+1})} \leq 1.$$

However, we now show that the value of this limit is instead $+\infty$.

First, notice that

$$V^{FB}(b_{t+1}) = \tfrac{1}{1-\delta} \max_{w} \left\{ \psi^{-1}\left(v\left(b_{t+1}(1-\delta) + w\right) - v\left(b_{t+1}(1-\delta)\right)\right) - w \right\}.$$

At the optimal choice of w, we have $c^{FB}(b_{t+1}) = b_{t+1}(1-\delta) + w$, and

$$e^{FB}(b_{t+1}) = \psi^{-1}\left(v\left(b_{t+1}(1-\delta) + w\right) - v\left(b_{t+1}(1-\delta)\right)\right).$$

Therefore, by the envelope theorem,

$$\frac{d}{db_{t+1}} V^{FB}(b_{t+1}) = \frac{v'\left(c^{FB}(b_{t+1})\right) - v'\left(b_{t+1}(1-\delta)\right)}{\psi'\left(e^{FB}(b_{t+1})\right)} = 1 - \frac{v'\left(b_{t+1}(1-\delta)\right)}{\psi'\left(e^{FB}(b_{t+1})\right)}.$$

On the other hand, the derivative of $\tilde{V}(b_{t+1})$ is given by

$$1 - \frac{v'((1-\delta)b_{t+1})}{v'\left(v^{-1}\left(\tfrac{1}{1-\delta} v((1-\delta)b_t) - \tfrac{\delta}{1-\delta} v((1-\delta)b_{t+1})\right)\right)}. \tag{44}$$

From l'Hôpital's rule, we have that

$$\begin{aligned}
\lim_{b_{t+1}\to 0} \frac{\tilde{V}(b_{t+1})}{V^{FB}(b_{t+1})} &= \lim_{b_{t+1}\to 0} \frac{\frac{\mathrm{d}}{\mathrm{d}b_{t+1}} \tilde{V}(b_{t+1})}{\frac{\mathrm{d}}{\mathrm{d}b_{t+1}} V^{FB}(b_{t+1})} \\
&= \lim_{b_{t+1}\to 0} \frac{\frac{\mathrm{d}}{\mathrm{d}b_{t+1}} \tilde{V}(b_{t+1}) - 1}{\frac{\mathrm{d}}{\mathrm{d}b_{t+1}} V^{FB}(b_{t+1}) - 1} \\
&= \lim_{b_{t+1}\to 0} \frac{-\frac{v'((1-\delta)b_{t+1})}{v'\left(v^{-1}\left(\frac{1}{1-\delta} v((1-\delta)b_t) - \frac{\delta}{1-\delta} v((1-\delta)b_{t+1})\right)\right)}}{-\frac{v'((1-\delta)b_{t+1})}{v'(c^{FB}(b_{t+1}))}} \\
&= \lim_{b_{t+1}\to 0} \frac{v'(c^{FB}(b_{t+1}))}{v'\left(v^{-1}\left(\frac{1}{1-\delta} v((1-\delta)b_t) - \frac{\delta}{1-\delta} v((1-\delta)b_{t+1})\right)\right)} \\
&= +\infty.
\end{aligned}$$

The second equality holds because both the numerator and the denominator tend to $-\infty$. The final equality holds because $c^{FB}(b_{t+1}) \to 0$ as $b_{t+1} \to 0$, by Step 3b of Lemma A.12.

We have therefore shown that, given a date-t balance b_t, the choices of b_{t+1} that are available in the above program while satisfying the constraints Equations (41) and (42) come from some bounded set $\left[l^b\left(b_t\right), u^b\left(b_t\right)\right]$ with $l^b\left(b_t\right) > 0$. It is then immediate that consumption c_t must be chosen from some bounded interval $\left[l^c\left(b_t\right), u^c\left(b_t\right)\right]$ as well. Hence, given the continuity of $\bar{V}$, the problem defined by Equations (40) to (42) and $W = \bar{V}$ has a solution. We can then solve iteratively to determine an optimal (date-t) continuation policy $(c_s, b_{s+1})_{s=t}^{\infty}$. As noted above, this establishes $V\left(b_t\right) \geq \bar{V}\left(b_t\right)$.

Step 2. Showing that the supremum function V satisfies the functional equation (and hence $V\left(b_t\right) \leq \bar{V}\left(b_t\right)$) . Consider any $b_t > 0$ and first suppose $V\left(b_t\right)$ is strictly less than

$$\hat{e}(c_t, b_t, b_{t+1}) - (\delta b_{t+1} - b_t + c_t) + \delta V\left(b_{t+1}\right)$$

for some $\left(\hat{c}_t, \hat{b}_{t+1}\right)$ satisfying the constraints in the functional equation, i.e. Equations (41) and (42) for $W = V$. Then we can take a policy $\left(\hat{c}_s, \hat{b}_{s+1}\right)_{s=t+1}^{\infty} \in \Pi\left(\hat{b}_{t+1}\right)$ and generating payoff within $\nu > 0$ of $V\left(\hat{b}_{t+1}\right)$. Then, after reducing $\hat{c}_t$ by an amount that can be taken arbitrarily close to zero as $\nu \to 0$ (to ensure the principal's constraint is satisfied at date t), we have $\left(\hat{c}_s, \hat{b}_{s+1}\right)_{s=t}^{\infty} \in \Pi\left(b_t\right)$. But (for ν small enough) this sequence generates a payoff to the principal higher than $V\left(b_t\right)$, contradicting the definition of the latter. Hence, $V\left(b_t\right)$ is an upper bound on

$$\hat{e}(c_t, b_t, b_{t+1}) - (\delta b_{t+1} - b_t + c_t) + \delta V\left(b_{t+1}\right)$$

over policies (c_t, b_{t+1}) satisfying the constraints of the functional equation (41) and (42).

Let $\left(c'_s, b'_{s+1}\right)_{s=t}^{\infty} \in \Pi\left(b_t\right)$ be a policy that generates payoff within ν of $V\left(b_t\right)$. Then $\left(c'_t, b'_{t+1}\right)$ satisfies the FE constraints (41) and (42). Also, $\left(c'_s, b'_{s+1}\right)_{s=t+1}^{\infty} \in \Pi\left(b'_{t+1}\right)$. Hence, by definition of V,

$$\begin{aligned}
&\hat{e}(c'_t, b_t, b'_{t+1}) - (\delta b'_{t+1} - b_t + c'_t) + \delta V\left(b'_{t+1}\right) \\
\geq &\hat{e}(c'_t, b_t, b'_{t+1}) - (\delta b'_{t+1} - b_t + c'_t) + \sum_{\tau=t+1}^{\infty} \delta^{\tau-t}\left(\hat{e}(c'_\tau, b'_\tau, b'_{\tau+1}) - (\delta b'_{\tau+1} - b'_\tau + c'_\tau)\right) \\
\geq &V\left(b_t\right) - \nu
\end{aligned}$$

This shows that $V\left(b_t\right)$ is the least upper bound for

$$\hat{e}(c_t, b_t, b_{t+1}) - (\delta b_{t+1} - b_t + c_t) + \delta V\left(b_{t+1}\right)$$

over policies (c_t, b_{t+1}) satisfying the constraints (41) and (42) of the functional equation.

□

(End of the proof of Proposition 5.3.) □

Proof of Lemma 5.4.

Proof. If the result does not hold, then there is a date t such that

$$\frac{v\left(\tilde{b}_t(1-\delta)\right)}{1-\delta} < \sum_{s=t}^{\infty} \delta^{s-t}\left(v\left(\tilde{c}_s\right) - \psi(\tilde{e}_s)\right).$$

We can increase the payment to the agent at date $t-1$ by $\varepsilon\delta$ for $\varepsilon > 0$, and reduce the date-t payment by ε. All other variables are unchanged. Provided ε is small enough, all constraints are preserved. Because the date-t payment is reduced, the principal's constraint (PC_t) is then slack at date t.

Because the contract is optimal, but not first best, we have that effort strictly increases over time. We can then change the date-t effort to a value $\tilde{e}'_t$, and the date-$t+1$ effort to $\tilde{e}'_{t+1}$, with $\tilde{e}_t < \tilde{e}'_t < \tilde{e}'_{t+1} < \tilde{e}_{t+1}$, and with

$$\psi\left(\tilde{e}'_t\right) + \delta\psi\left(\tilde{e}'_{t+1}\right) = \psi\left(\tilde{e}_t\right) + \delta\psi\left(\tilde{e}_{t+1}\right).$$

All other variables remain unchanged. This affects the agent constraints (AC_t^{ob}) by increasing the profitability of remaining in the contract from date $t+1$ onwards (i.e., the date-$t+1$ constraint is slackened). It relaxes the principal's constraint at date $t-1$ and earlier, because the NPV of effort increases (by convexity of ψ). It tightens the principal's constraint at date t, but provided the changes are small, it remains slack. The principal's constraints are unaffected from date $t+1$ onwards. Because the NPV of effort increases, profits strictly increase. □