Relational Contracts: Public versus Private Savings

Francesc Dilmé* **Daniel Garrett†**

June 9, 2019

PRELIMINARY VERSION

Abstract

We study relational contracting with a risk-averse agent who thus has preferences for smoothing consumption over time. The agent has 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. 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. We suggest that modeling informal/relational incentives on consumption/savings decisions is a pertinent alternative to the approach in existing literature on formal contracts in dynamic moral hazard.

JEL Classification: D82

Keywords: relational contracts, risk-averse agents

*University of Bonn.

†Toulouse School of Economics.

1 Introduction

The literature on repeated moral hazard (such as Rogerson, 1985, and Fudenberg, Holmstrom and Milgrom, 1990) highlights the value to the principal of controlling the agent's consumption/savings decisions. Given an optimal dynamic contract in the setting of Rogerson, the agent is required 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 most 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. 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)). Also, in the US, many employers have access to employee credit information (a 2012 survey by the Society for Human Resource Management found that 47% of US employers access this information when making hiring decisions).

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. To this end, we compare the evolution of optimal relational contracts in a simple deterministic environment. We assume that (i) effort and hence output are jointly observed, (ii) workers are risk-averse and hence have a preference for smooth consumption, and (iii) consumption is either fully observable to the principal or fully unobservable.

When consumption is fully observable, an optimal relational contract calls for a specified level of consumption expenditure; failure to consume at this level implies equilibrium punishment in the form of termination of productive employment. One interpretation is that an employee with an insufficiently frivolous lifestyle (in terms of car, dress, leisure activities and perquisites) is let go (perhaps dismissed after being deemed a poor fit with the corporate cul-

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

ture). When the principal fails to observe consumption, the agent is naturally free to smooth consumption optimally without putting the future of the relationship at risk. Although reality might often lie between these two extremes – e.g., some consumption or savings decisions are observable or at least deemed relevant by the employer, while others are not – studying the extremes of observed and unobserved consumption is illuminative and simplifies the analysis.

As in a number of other relational contracting settings, there are two relevant incentive constraints in each period. 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 willingness of the principal to pay the agreed compensation to the agent, which (given the absence of formal contracts) is entirely at the employer's discretion. The interesting cases occur when the players are insufficiently patient, so that the principal's constraints bind, at least in some periods.

For the cases of interest, when consumption is unobserved, we find that the profitability of the relationship deteriorates over time. This deterioration has its source in the value of the agent's outside option of quitting the relationship (after deviating from the required effort levels), which improves over time. In particular, an agent who intends to quit in a given period can optimally smooth his consumption by choosing the same consumption inside and outside of the relationship, from the beginning. The later he quits, i.e. the longer he chooses to work, the greater his lifetime earnings, and the smaller the value for additional earnings obtainable by obediently choosing the effort prescribed by the contract. Hence, at later dates, lower effort can be induced from the agent for the same payments, reducing the profitability of the contract. This effect feeds back on itself, for the payments the principal can credibly make to the agent become smaller as the future profitability of the relationship shrinks with time.

We document that, with unobservable consumption, effort and output are either decreasing over time, or they are initially constant over time and then (once the credibility constraint binds) decreasing. In the first case, wages fall over time (since the credibility constraint binds and tightens over time as explained above), while in the second case wages may initially rise before declining. When consumption is instead observed, the optimal contract is designed so the value to the agent of quitting falls rather than increases over time. This is achieved by requiring the agent, in the informal relationship, to consume more than he would like, driving down the balance on his savings account. As a result, the longer the agent works, the smaller his total savings, which, other things equal, makes remaining obediently in the relationship more desirable. Hence, the remuneration that must be paid to compensate a given level of effort falls over time, the relationship becomes more profitable, and the principal's credibility constraints are relaxed over time. As a result, payments and effort increase with time.

We believe our work is relevant to research on wage-tenure profiles. The common empirical finding is increasing wage-tenure profiles. A force that one might expect to push in this direction is that wealthier workers are harder to motivate; hence workers with long tenure should be paid more. We show, however, that if incentives are relational, the implications of wealth accumulation (or at least the possibility for the worker to accumulate wealth over time) tends to work in the opposite direction, because it harms the profitability of the relationship after long tenure. In turn, this reduces the compensation that the principal can credibly promise to pay. Conversely, requiring high consumption spending by the worker can help to keep the worker poor. Rather than implying a reduction in compensation payments over time, this can instead make the relationship grow more profitable, so the principal can credibly promise higher compensation.

1.1 Other literature

This paper contributes to the literature on relational contracts, reviewed in MacLeod (2007) and Malcomson (2013). Much of this literature focusses on moral-hazard settings without savings. Some papers introduce non-persistent information asymmetry, like Fuchs (2007, about the worker's evaluation) and Li and Matouschek (2013, about the principal's opportunity cost of funds). In Halac (2012), the principal holds a permanent private information about her outside option.

Our analysis introduces savings into a model where the agent is risk averse and thus has preferences for smoothing consumption, and it studies how the contract dynamics depend on the observability of the consumption decisions by the agent. Both savings and risk aversion have received little attention in the 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 (compensation is equal to consumption).

Our paper is related the literature on dynamic contracts with private savings, like Edmans et al. (2009), He (2012), and Di Tella and Sannikov (2018). In this literature, the principal has full commitment power. As a consequence, a common result is that optimal contracts feature no private savings by the agent. A common result is that optimal contracts, instead of featuring front-loaded consumption patterns as in the case where saving is observed (see Rogerson (1985)), feature backloaded payments to the agent which discourage private savings. Unlike these papers, in the version of our model where consumption/savings are privately observed, the level of the agent's consumption typically deviates from the level of his pay.

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.[2]We further assume that v, when evaluated on positive consumption values, is twice continuously differentiable, strictly increasing, strictly concave, 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 and 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 our attention on *deterministic* relational contracts, so that a relational contract is simply an agreement on the sequence $(e_t, w_t, c_t, b_t)_{t \geq 1}$. We understand this as simply reflecting the players' inability to randomize.[3] We then consider contracts

[2]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).

[3]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.

satisfying the following feasibility constraint.

Definition 2.1. A *feasible relational contract* is a sequence $(e_t, w_t, c_t, b_t)_{t\geq 1}$ with the following feasibility conditions:

1. **Non-negativity:** $e_t, w_t \geq 0$ and $c_t > 0$ for all t.

2. **Balance dynamics and constraint:** Conditions (1) and (2) hold.

3. **Bounded sequences:** The sequences $(e_t)_{t\geq 1}$, $(c_t)_{t\geq 1}$ and $(w_t)_{t\geq 1}$ are bounded.

Our approach, below, is to provide a natural set of conditions under which we can view feasible relational contracts as "self-enforcing". While our arguments reference fixed equilibrium concepts – Perfect Bayesian Equilibrium in the unobservable-consumption case, and Subgame Perfect Nash Equilibrium in the observable-consumption case – we proceed somewhat informally and reach the definition of "self-enforcing" directly without a full definition of strategies and equilibria. We thus skirt complications in defining strategies that relate, for instance, to the long-run condition on the agent's balance (which depends on both the agent's consumption and principal's payments). We also avoid difficulties in defining Perfect Bayesian Equilibrium (as have been reognized elsewhere). However, in Appendix B, we do provide strategies, and beliefs in the case of unobservable consumption, that coincide (on path) with self-enforcing agreements as we define below, and which would satisfy reasonable equilibrium definitions (and restrictions on strategies).

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 $(e_t, c_t, w_t, b_t)_{t\geq 1}$ such that the payoff of the agent

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

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 $(e_t, c_t, w_t, b_t)_{t\geq 1}$, 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 typically 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 some version of *Perfect Bayesian Equilibrium* (PBE); our task, to find an equilibrium with outcomes $(e_t, c_t, w_t, b_t)_{t\geq 1}$ that is optimal for the principal.

Note that both principal and agent have the option, at any date, to deviate from the agreement. In this case, the principal and agent can be held to payoffs no smaller than the ones obtained when the principal makes no payments and the agent exerts no effort from then on; the agent, in addition, perfectly smoothing the balance of his account over the infinite future (such continuation play might be described as "autarky"). More explicitly, if the agent is paid zero at all future dates, the best he can do is to exert zero effort and perfectly smooth his balance, attaining the autarky payoff. If the agent exerts zero effort at all future dates, the best the principal can do is to make no payments, again attaining the autarky payoff.

We now consider, without loss of optimality, agreements $(e_t, c_t, w_t, b_t)_{t\geq 1}$ that can be sustained with any deviation in effort or payment being punished by autarky. In this section, 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). We here provide a definition of a self-enforcing agreement as one that is sustainable by autarky punishments, and then exhibit in Appendix B strategies and beliefs that would satisfy our view of a PBE, and whose equilibrium outcomes are the ones specified in a self-enforcing agreement.[4]

Suppose that the agent deviates in his effort choice for the first time at time t. That is, suppose the agent exerts effort equal to e_s for all $s < t$, and an effort different from 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

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

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 $c_t = \bar{c}_\infty$ for all t.

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

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

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

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

[4]We skirt the difficulty of defining PBE more generally, although for any sensible definition the minmax payoffs for the players would be those determined by autarky.

The principal remains willing to continue abiding by the agreement if, at each time t, the payment w_t that is due is less than her continuation payoff in the agreement. Thus we require that, for all $t \geq 1$,

$$w_t \leq \sum_{s=t+1}^{\infty} \delta^{s-t}(e_s - w_s). \tag{PC$_t$}$$

Our main objective will be to determine relational contracts that maximize the principal's payoff subject to the above incentive constraints. Formally, we refer to *self-enforcing (relational) contracts*, being feasible relational contracts $(e_t, c_t, b_t, w_t)_{t\geq 1}$ that satisfy the following conditions:

1. **Agent's incentive constraint:** $c_t = \bar{c}_\infty$ for all t, and condition ($\mathrm{AC}_t^{\mathrm{un}}$) holds for all $t \geq 1$.
2. **Principal's constraint:** Condition (PC_t) holds for all $t \geq 1$.

An *optimal contract* is then a self-enforcing contract that maximizes the payoff of the principal.

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". Our formal result is as follows.

Lemma 4.1. *If there is an optimal self-enforcing relational contract, then there is one with the same sequence of efforts such that, for all $t \geq 1$,*

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

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

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

If condition (4) does not hold, 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

(AC_t^{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, 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 (FP_t^{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, $e_t > 0$ implies

$$w_t \in \left(\frac{\psi(e_t)}{v'(\bar{c}_{t-1})}, \frac{\psi(e_t)}{v'(\bar{c}_t)} \right). \tag{5}$$

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 $\frac{w_t}{\psi(e_t)}$ increases with t. In other words, the payments needed to keep the agent 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 $(e_t)_{t\geq 1}$.[5] We next discuss the implementation of first-best contracts (Section 4.1), before moving to consider optimal contracts when the first-best is not achievable (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 solution, 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. According to the condition (5), payments to the agent increase over time. In the long run, payments approach

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

[5]Note that from $(e_t)_{t\geq 1}$ we can obtain $(\bar{c}_t)_{t\geq 1}$ using (FP_t^{un}) (so the corresponding optimal consumption $c_t = \bar{c}_\infty$ is also pinned down). Then $(w_t)_{t\geq 1}$ is obtained from Equation (3), and $(b_t)_{t\geq 1}$ from Equation (1).

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 4.1. *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 the 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 is not implementable.

Proposition 4.2. *Suppose the principal cannot attain the first-best payoff in a self-enforcing relational 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 $e_\infty > 0$ 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.*

An intuition for this result is as follows. The relevant deviations for the agent involve obediently choosing the effort specified by the contract for some time, and then quitting the relationship. The longer the agent obediently follows the effort choices dictated by the agreement, the more income he accumulates. Therefore, to make deviations in later periods unprofitable, the payments made in return for a fixed level of effort increase. As a consequence, the relationship becomes less profitable for the principal. The principal's constraint (PC_t) is then more difficult to satisfy in later periods; in particular, high payments are less able to be credibly promised. This effect in fact feeds back on itself, since a more constrained relationship features lower effort and is even less profitable.

Certain elements of the proof shed light on the forces underlying this result. For instance, one step in the proof (Lemma A.7) is to rule out the possibility that a constant effort policy is optimal. Suppose, for example, that there is a self-enforcing contract with constant effort e_∞, and let the payments and the equilibrium consumption $\bar{c}_\infty$ be determined by Equation ($\mathrm{FP}_t^{\mathrm{un}}$). Then payments increase over time, and converge to $\frac{\psi(e_\infty)}{v'(\bar{c}_\infty)}$. The principal's enforcement constraint (PC_t) is then satisfied if and only if

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

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 best choice of a constant effort e_∞^*, 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 $\psi'\left(e_\infty^*\right) < v'\left(\bar{c}_\infty\right)$, and so any 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 that increases the NPV of effort, but (assuming that payments continue to satisfy ($\mathrm{FP}_t^{\mathrm{un}}$)) leaves the principal's constraints (PC_t) intact. Concretely, we consider increasing effort 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 ($\mathrm{FP}_t^{\mathrm{un}}$), the principal's constraint (PC_t) is eventually violated (since v is strictly concave, it is more costly to compensate the agent for his effort, and payments 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.

The fact the optimal effort policy is not constant turns out to imply that the principal's constraint (PC_t) must hold as an equality in some period (this follows as a consequence of Lemma A.3). We also 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 $e_{\hat{t}+1} < e_{\hat{t}}$ and the constraint holds with equality also at $\hat{t}+1$. Hence effort strictly decreases over time.

The argument can be summarized as follows. By assumption, the principal's constraint

(PC$_t$) at date $\hat{t}$ holds as an equality, i.e.

$$w_{\hat{t}} = \sum_{s=\hat{t}+1}^{\infty} \delta^{s-\hat{t}} \left(e_s - w_s\right).$$

We establish (in Lemma A.1) that $\psi'(e_t) \leq v'(\bar{c}_\infty)$ for all t, which means that (under Condition (FP$_t^{\mathrm{un}}$)) 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), we then have $e_{\hat{t}+1} - w_{\hat{t}+1} > e_s - w_s$ for all $s > \hat{t}+1$. Therefore,

$$w_{\hat{t}} > \sum_{s=\hat{t}+2}^{\infty} \delta^{s-\hat{t}-1} \left(e_s - w_s\right) \geq w_{\hat{t}+1}$$

where the second equality is the principal's constraint (PC$_t$) at date $\hat{t}+1$. This shows why a binding constraint for the principal at $\hat{t}$ implies a strictly lower payment and (using (5)) effort in the following period. In particular, it confirms that the declining profitability of the relationship implies that, when the principal is constrained, payments and hence effort must decrease with time.

In order to translate the finding of Proposition 4.2 into predictions for payments and the agent's balance, it is important to 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$_t^{\mathrm{un}}$).

Proposition 4.3. *Suppose the principal cannot attain the first-best payoff in a self-enforcing relational contract. Fix an optimal policy and let $\bar{t}$ be the date from which effort is strictly decreasing (see Proposition 4.2). Then, Condition (FP$_t^{un}$) holds for all $t > \bar{t}$.*

Proposition 4.3 implies that, given an optimal effort policy $(e_t)_{t\geq 1}$, payments w_t are uniquely determined from date $\bar{t}+1$ onwards. Our arguments, which took payments to be determined by Condition (FP$_t^{\mathrm{un}}$), then imply immediately the dynamics.

Corollary 4.1. *Suppose the principal cannot attain the first-best payoff in a self-enforcing relational contract. Fix an optimal effort policy and let $\bar{t}$ be the date from which effort is strictly decreasing (see Proposition 4.2). 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 effort policy, the principal's constraint (PC$_t$) holds with equality at any $\hat{t} > \bar{t}$. As explained above, this permits us to conclude that $w_{\hat{t}+1} < w_{\hat{t}}$.

The fact that the agent's balance increases over time follows straightforwardly from Equation (1) and Equation (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

$$b_t = \frac{\bar{c}_\infty}{1-\delta} - \sum_{\tau \geq t} \delta^{\tau - t} w_\tau$$

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 $\bar{t}$.

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)).

As in Section 4, our objective is to obtain relational contracts that maximize the principal's payoff. Analogous to the arguments made in the previous section, we consider contracts that are self-enforcing when deviations 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. The relevant deviations are then those in which the principal, at any given date t, makes a payment equal to zero, and where the agent, at any date t, chooses effort equal to zero and consumes $b_t(1-\delta)$ where b_t is the publicly observed balance at the beginning of date t.

The above specification means that the agent's payoff, if complying until date $t-1$ and optimally failing to comply from t onwards, is now

$$\sum_{s=1}^{t-1} \delta^{s-1} \left(v\left(c_s\right) - e_s\right) + \delta^{t-1} \frac{v((1-\delta)b_t)}{1-\delta}.$$

Thus, the agent is willing to exert effort e_t and consume c_t at date t if and only if

$$\frac{v\left(b_t\left(1-\delta\right)\right)}{1-\delta} \leq \sum_{s=t}^{\infty} \delta^{s-t} \left(v\left(c_s\right) - e_s\right). \tag{$\mathrm{AC}_t^{\mathrm{ob}}$}$$

The key difference to Condition (AC_t^{un}) is that continuing to publicly honour the agreement up to date $t-1$ ensures that the agent begins period t with the specified balance b_t, which in turn determines the wealth he has available to spend in autarky. Condition (AC_t^{un}), on the other hand, takes into account that the agent quitting the relationship at date t can save in advance for this event, because consumption is not observed.

The principal's constraint, nonetheless, is the same as when consumption is unobservable. Therefore, we have that for the observable-consumption case, a *self-enforcing (relational) contract* is a feasible relational contract $(e_t, c_t, b_t, w_t)_{t\geq 1}$ satisfying the following conditions:

1. **Agent's incentive constraint:** Condition (AC_t^{ob}) holds for all $t \geq 1$.

2. **Principal's constraint:** Condition (PC_t) holds for all $t \geq 1$.

We begin with 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. *Suppose $(e_t, c_t, w_t, b_t)_{t\geq 1}$ is an optimal relational contract. Then there exists another optimal contract 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 $t \geq 1$,*

$$\frac{v\left(b_t\left(1-\delta\right)\right)}{1-\delta} = \sum_{s=t}^{\infty} \delta^{s-t}\left(v\left(c_s\right) - \psi(e_s)\right). \tag{7}$$

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

$$v(c_t) - \psi(e_t) + \tfrac{\delta}{1-\delta} v((1-\delta)b_{t+1}) = \tfrac{1}{1-\delta} v((1-\delta)b_t). \tag{FP_t^{ob}}$$

This says that the agent is indifferent between quitting at date t and smoothing the balance b_t optimally over the infinite future, and working one more period, exerting effort e_t and consuming c_t, before quitting at date $t+1$ and smoothing the balance 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 $w_t = w^{FB}\left(b_1\right) \equiv c^{FB}(b_1) - (1-\delta)b_1$ 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. Assuming the agent is willing to abide by the contract, the principal's payoff can 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. Importantly, note that the agent is thus paid earlier than for the "fastest payments" of the unobservable-consumption case (in the unobservable-consumption case, the NPV of payments corresponding to constant first-best effort is the same, but payments determined as "fastest" strictly increase over time).

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 forever. The condition for implementing the first-best is then, analogous to Condition (6),

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

This condition is more easily satisfied than in the unobservable 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

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

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

5.2 Optimal contract with observed consumption

The principal's continuation payoff from date t onwards is given by

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

We can restrict attention to payments that keep the agent indifferent to quitting the agreement, as given in Lemma 5.1. At each date t, given the balance b_t, the continuation contract must then maximize the principal's continuation payoff subject to the indifference condition in Equation (7). This problem is entirely forward-looking. It depends only on the agent's balance b_t, as determined by the contract, at the beginning of period t.

Given the above, we let $V(b_t)$ denote the maximum of the principal's date-t continuation payoff given the agent's date-t balance b_t (we show in Lemma A.15 that this maximum is attained). The principal's problem can be stated recursively as

$$V(b_t) = \max_{e_t, b_{t+1}, c_t} \big(e_t - (\delta b_{t+1} - b_t + c_t) + \delta V(b_{t+1})\big) \tag{9}$$

subject to agent's indifference condition ($\mathrm{FP}_t^{\mathrm{ob}}$) and to the principal's constraint

$$\delta b_{t+1} - b_t + c_t \leq \delta V(b_{t+1}). \tag{10}$$

Here, the interpretation is that the agent's payment w_t can be divided into date-t consumption $c_t \in \mathbb{R}$ and savings $\delta b_{t+1} - b_t \in \mathbb{R}$. Additionally, the effort and the payment must be non-negative.

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

Proposition 5.1. *Suppose that, given the balance b_1, the principal's optimal contract fails to obtain the first-best payoff $V^{FB}(b_1)$. Then the agent's balance b_t and consumption c_t decline strictly over time, with $b_t \to b_\infty$ for some $b_\infty > 0$. Effort e_t and the payments w_t determined by the Conditions (FP_t^{ob}) increase strictly over time. We have $V(b_t) \to V^{FB}(b_\infty)$ as $t \to \infty$, and hence the contract converges to the first-best contract for balance b_∞.*

A heuristic account of the forces behind this result is as follows. When the constraint in ($\mathrm{FP}_t^{\mathrm{ob}}$) 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} in the subsequent period, increases the principal's continuation payoff and relaxes the date-t constraint. 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 balance, the contract is close to first-best, and so the value of continuing to distort consumption vanishes.

Central to our analysis is an Euler equation

$$1 - \frac{v'((1-\delta)b_{t+1})}{v'(c_t)} = \frac{v'(c_{t+1})}{\psi'(e_{t+1})}\left(1 - \frac{v'((1-\delta)b_{t+1})}{v'(c_{t+1})}\right)$$

which we use to derive several key properties. This equation captures the relationship between a dynamic distortion — when the first-best is not implementable $c_{t+1} < c_t$ — and a static distortion — $\psi'(e_{t+1}) < v'(c_{t+1})$ (recall, in the first-best policy, the marginal disutility of effort is *equal* to the marginal utility of consumption). For instance, as $b_t \to b_\infty$, consumption falls to its lowest bound, becoming almost constant, so $\frac{v'(c_{t+1})}{\psi'(e_{t+1})} \to 1$, which accords with convergence to the first-best policy.

Finally, analogous to Lemma 4.3, we would like to show also that, when there is no self-enforcing first-best contract, the timing of payments is uniquely determined by the Conditions ($\mathrm{FP}_t^{\mathrm{ob}}$).

Lemma 5.2. *Suppose the principal cannot attain the first-best payoff in a self-enforcing 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.*

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

A Proofs

A.1 Proofs of the results in Section 3

Proof of Proposition 3.1. Consider the problem of maximizing the principal's payoff

$$\sum_{t=1}^{\infty} \delta^{t-1} \left(e_t - w_t\right)$$

subject to the constraint that

$$\sum_{t=1}^{\infty} \delta^{t-1} \left(v\left(c_t\right) - \psi\left(e_t\right)\right) \geq \frac{v\left(b_1\left(1-\delta\right)\right)}{1-\delta}$$

together with (2), that is,

$$b_1 + \sum_{t=1}^{\infty} \delta^{t-1} \left(w_t - c_t\right) \geq 0.$$

The Lagrangian for this problem is

$$\sum_{t=1}^{\infty} \delta^{t-1} \left(e_t - w_t\right) + \lambda \left(\sum_{t=1}^{\infty} \delta^{t-1} \left(v\left(c_t\right) - \psi\left(e_t\right)\right) - \frac{v\left(b_1\left(1-\delta\right)\right)}{1-\delta} \right) + \mu \left(b_1 + \sum_{t=1}^{\infty} \delta^{t-1} \left(w_t - c_t\right) \right)$$

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

$$\psi'\left(e_t\right) = 1/\lambda$$

and

$$v'\left(c_t\right) = \frac{\mu}{\lambda}$$

for all t. That is, effort e_t must be constant at some e and consumption c_t constant at some c. The two constraints binding is necessary for an optimum. The second constraint yields

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

The first constraint may be written as

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

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

$$\frac{e}{1-\delta} - \frac{v^{-1}\left(\psi\left(e\right) + v\left(b_1\left(1-\delta\right)\right)\right)}{1-\delta} + b_1. \tag{11}$$

The first-order condition for a maximum is

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

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

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

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

Now consider why the first-best payoff $V^{FB}(b_1)$ is decreasing in b_1. It is clear that the first-best effort $e^{FB}(b_1)$ is continuous in b_1. Standard arguments (see Milgrom and Segal, 2002) can then be used to establish, using (11), that $V^{FB}(b_1)$ is absolutely continuous and differentiable a.e., with derivative

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

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

A.2 Proofs of the results in Section 4

Proof of Lemma 4.1

Proof. Fix an optimal (and hence self-enforcing) relational contract $(e_t, c_t, w_t, b_t)_{t\geq 1}$. We show first that

$$\frac{v(\bar{c}_\infty)}{1-\delta} - \sum_{s=1}^{\infty} \delta^{s-1}\psi(e_s) \tag{12}$$

is equal to $\frac{v(b_1(1-\delta))}{1-\delta}$. Clearly the only way this can fail in a self-enforcing relational contract is if (12) strictly exceeds $\frac{v(b_1(1-\delta))}{1-\delta}$. However, in this case, there is a more profitable contract in which e_1 is strictly increased by a small amount. This leaves the agent's constraints (AC_t^{un}) unchanged at all periods except the initial period, where it continues to hold (provided the increase in e_1 is small, the agent prefers to continue in the contract forever by choosing the specified effort levels, rather than shirk at date 1 and consume $b_1(1-\delta)$ in every period). The principal's constraints (PC_t) are unaffected.

Next observe that, if

$$\frac{v(\bar{c}_{t-1})}{1-\delta} - \sum_{s=1}^{t-1} \delta^{s-1}\psi(e_s) \tag{13}$$

exceeds $\frac{v(b_1(1-\delta))}{1-\delta}$ at any date t, then the agent prefers to work up to date $t-1$ and then shirk (choose effort zero) at date t, as opposed to working at the prescribed level forever. Hence, the contract is not self-enforcing.

Finally, suppose that the expression (13) 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

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

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 (13) 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 w_t in the new relational contract are non-negative. At every date t, the agent is indifferent between working at the agreed level and shirking (putting $e_t = 0$). 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 4.1 and 4.2

Proof. The proof of the Proposition 4.1 is divided into eight 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 A.3 shows that if the principal's constraint (PC_t) is slack at date t, then effort is higher at date t than in adjacent periods.

4. Lemma A.4 uses the previous lemma to show that the contract becomes stationary in the long run.

5. Lemma A.5 shows that the 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 the first-best policy to be achievable, 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: $\psi'(e_t) \leq v'(\bar{c}_\infty) \leq v'(\bar{c}_t)$ (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. *If $(e_t, c_t, w_t, b_t)_{t\geq 1}$ is a self-enforcing relational contract and $\psi'(e_t) > v'(\bar{c}_\infty)$ for some t, then there is a contract achieving a higher payoff. 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-enforcing relational contract satisfies $\psi'(e_t) \leq v'(\bar{c}_\infty)$ for all t, we have $\bar{c}_\infty \leq \hat{c}$.

Proof. Take a contract satisfying condition ($\mathrm{FP}_t^{\mathrm{un}}$) for all t, and such that $\psi'(e_{t^*}) > v'(\bar{c}_\infty)$ for some $t^* \in \mathbb{N}$. Consider a new contract satisfying condition ($\mathrm{FP}_t^{\mathrm{un}}$), with efforts coinciding with the original contract at all times except at t^*, where we specify $e'_{t^*} = e_{t^*} - \varepsilon$, for some small $\varepsilon > 0$. Consider a decrease in the date-t^* payment, denoted Δw, satisfying

$$\frac{v(\bar{c}_\infty) - v\left(\bar{c}_\infty - (1-\delta)\,\delta^{t^*-1}\Delta w\right)}{1-\delta} = \delta^{t^*-1}\big(\psi(e_{t^*}) - \psi(e_{t^*} - \varepsilon)\big).$$

We have that $v'(\bar{c}_\infty)\,\Delta w = \psi'(e_t)\,\varepsilon + o(\varepsilon)$; hence the wage saving is $\frac{\psi'(e_{t^*})\varepsilon}{v'(\bar{c}_\infty)} + o(\varepsilon)$.[6] The gain to the principal of the change in the contract (in date-t^* terms) is equal to the savings in payments minus the change in the effort, i.e. $\frac{\psi'(e_{t^*})\varepsilon}{v'(\bar{c}_\infty)} - \varepsilon + o(\varepsilon)$, which is strictly positive since $\psi'(e_{t^*}) > v'(\bar{c}_\infty)$.

The change in the contract is such that the equilibrium payoff of the agent is unchanged, and hence equal to $\frac{v(b_1(1-\delta))}{1-\delta}$. Now suppose that the agent deviates from the terms of the contract for the first time at any date $t > t^*$ by exerting zero effort from then on and consuming optimally. Since v is strictly concave, the effect of the change of contract on the agent's payoff under such deviations is strictly negative (the agent consumes less by $(1-\delta)\,\delta^{t^*-1}\Delta w$ in each period, and the reduction in agent consumption utility due to this change is greater than

[6] Here, we use that feasibility of the sequences requires that payments in particular are bounded; hence $\bar{c}_\infty$ is finite and $v'(\bar{c}_\infty) > 0$.

the equilibrium effect, i.e. where he remains obedient). Hence, under the new contract, the agent strictly prefers to be obedient from date $t^* + 1$ onwards. At dates $t \leq t^*$, the agent is still indifferent between quitting the relationship (by exerting zero effort and consuming optimally) and choosing effort obediently for all time (and optimally setting consumption equal to $\bar{c}_\infty - (1-\delta)\,\delta^{t^*-1}\Delta w$ in each period).

It is also easy to see that the principal's constraints (PC_t) continue to be satisfied: They are relaxed at dates t^* and earlier, and unchanged after date t^*.

We have thus established that we can restrict attention to effort policies such that $z\left(v'\left(\bar{c}_\infty\right)\right)$ is an upper bound on effort (with $\bar{c}_\infty$ pinned down by the effort policy, as described in Footnote 3). Because (FP_t^{un}) holds at $t = \infty$, we must then have that

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

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

The above result establishes that the marginal disutility of effort $\psi'\left(e_t\right)$ in an optimal contract is bounded by $v'\left(\bar{c}_\infty\right)$, which is certainly no greater than $v'\left(b_1\left(1-\delta\right)\right)$, given our restriction that payments w_t are non-negative. In turn, because $v'\left(\bar{c}_\infty\right)$ is no less than $v'\left(\hat{c}\right)$, the condition (FP_t^{un}) implies that payments to the agent are bounded by some $\bar{w} > 0$ (see Condition (5)). We now prove existence of an optimal contract.

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

Proof. Note that, under the condition "fastest payments" given in (FP_t^{un}), the relational contract is determined solely by the effort policy $\left(e_t\right)_{t\geq 1}$ (see Footnote 3). Hence, the payoff obtained by the principal can be written

$$W\left(\left(e_t\right)_{t=1}^{\infty}\right) = \sum_{t=1}^{\infty} \delta^{t-1}\left(e_t - w_t\right)$$

where $\sum_{t=1}^{\infty} \delta^{t-1} w_t$ is determined by

$$\frac{v\left(\left(1-\delta\right)b_1 + \left(1-\delta\right)\sum_{t=1}^{\infty}\delta^{t-1}w_t\right)}{1-\delta} = \sum_{t=1}^{\infty}\delta^{t-1}\psi(e_t) + \frac{v\left(\left(1-\delta\right)b_1\right)}{1-\delta}. \tag{14}$$

Given that, from Lemma A.1, we can restrict attention to effort policies in $\left[0, z\left(v'\left(b_1\left(1-\delta\right)\right)\right)\right]^{\infty}$, $\sum_{t=1}^{\infty}\delta^{t-1}w_t$ 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 in $[0, z(v'(b_1(1-\delta)))]$ such that the relational contract is self-enforcing. Consider then a sequence of policies $((e_t^n)_{t=1}^{\infty})_{n=1}^{\infty}$ with

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

for all n, and for which the contract defined by each effort policy (see Footnote 3) is self-enforcing. There then exists a sequence $(e_t^{\infty})_{t\geq 1} \in [0, z(v'(b_1(1-\delta)))]^{\infty}$ and a subsequence $\left((e_t^{n_k})_{t\geq 1}\right)_{k\geq 1}$, with $(e_t^{n_k})_{t\geq 1}$ convergent pointwise to $(e_t^{\infty})_{t\geq 1}$. Since

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

as $k \to \infty$ (by continuity of ψ and discounting) we have (by continuity of v and Equation (14))

$$\sum_{t=1}^{\infty} \delta^{t-1} w_t^{n_k} \to \sum_{t=1}^{\infty} \delta^{t-1} w_t^{\infty}$$

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

Our result will then follow if we can show that the contract defined by $(e_t^{\infty})_{t\geq 1}$ (see Footnote 3) is self-enforcing. Because the payment policy paired with $(e_t^{\infty})_{t\geq 1}$ satisfies ($\mathrm{FP}_t^{\mathrm{un}}$), the agent's incentive constraints ($\mathrm{AC}_t^{\mathrm{un}}$) are satisfied, so it remains to check the principal's incentive constraints (PC_t) are satisfied. Suppose with a view to contradiction that there is some t^* at which the principal's constraint does not hold, and so

$$w_{t^*}^{\infty} > \sum_{s=t^*+1}^{\infty} \delta^{s-t^*}(e_s^{\infty} - w_s^{\infty}).$$

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

$$w_{t^*}^{n_k} > \sum_{s=t^*+1}^{\infty} \delta^{s-t^{n_k}}(e_s^{n_k} - w_s^{n_k})$$

contradicting that $(e_t^{n_k})_{t\geq 1}$ determines a self-enforcing agreement. □

We next establish an important property of periods where an optimal relational contract is such that the principal's constraint is slack:

Lemma A.3. *Suppose that $(e_t, c_t, w_t, b_t)_{t\geq 1}$ is an optimal relational contract. Suppose that the principal's constraint is slack at t^*, i.e. $w_{t^*} < \sum_{s=t^*+1}^{\infty} \delta^{s-t^*}(e_s - w_s)$. Then, $e_{t^*-1}, e_{t^*+1} \leq e_{t^*}$.*

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

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

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

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

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

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

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

Also, $w'_{t^*-1} < w_{t^*-1}$ and $w'_{t^*} > w_{t^*}$ 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. □

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

Lemma A.4. *Suppose that $(e_t, c_t, w_t, b_t)_{t\geq 1}$ is an optimal relational contract. Then, there exists an effort/payment pair (e_∞, w_∞) such that $\lim_{t\to\infty}(e_t, w_t) = (e_\infty, w_\infty)$.*

Proof. **Step 0.** We first prove that, if $(e_t, c_t, w_t, b_t)_{t\geq 1}$ is an optimal relational contract satisfying ($\mathrm{FP}_t^{\mathrm{un}}$), then

$$\lim_{t\to\infty}\left(w_t - \frac{\psi(e_t)}{v'(\bar{c}_\infty)}\right) = 0.$$

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

$$\frac{v\left(\bar{c}_{t-1} + (1-\delta)\,\delta^{t-1} w_t\right) - v\left(\bar{c}_{t-1}\right)}{1-\delta} = \delta^{t-1}\psi(e_t)\,.$$

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

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

which proves the result.

Step 1. Define $\bar{e} \equiv \limsup_{t\to\infty} e_t$, 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 $e \in [0, \bar{e}]$,

$$\frac{\psi(e)}{v'(\bar{c}_\infty)} \leq \frac{\delta}{1-\delta}\left(e - \frac{\psi(e)}{v'(\bar{c}_\infty)}\right). \tag{15}$$

That is, the principal's constraints 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 (15) is satisfied at $\bar{e}$, then it is satisfied for all $e \in [0, \bar{e}]$; this follows because the left-hand side of (15) 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 (15) is not satisfied for some $e \in [0, \bar{e}]$. Then we must have

$$\frac{\psi(\bar{e})}{v'(\bar{c}_\infty)} > \frac{\delta}{1-\delta}\left(\bar{e} - \frac{\psi(\bar{e})}{v'(\bar{c}_\infty)}\right). \tag{16}$$

Observe then that

$$e_t - w_t \leq \bar{e} - \frac{\psi(\bar{e})}{v'(\bar{c}_\infty)} + \varepsilon_t$$

for some sequence $(\varepsilon_t)_{t=1}^{\infty}$ convergent to zero. This follows because $w_t - \frac{\psi(e_t)}{v'(\bar{c}_\infty)} \to 0$ 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 $\limsup_{t\to\infty} e_t$.
We therefore have that

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

where the last inequality holds by equation (16). However, Step 0 implies that the superior limit of wages 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 $\limsup_{t\to\infty} e_t$ (with $(e_t)_{t\geq 1}$ the effort profile in an enforceable relational contract).

Step 2. Assume, for the sake of contradiction, that $\underline{e} \equiv \liminf_{t\to\infty} e_t < \bar{e}$. In this case, there exists some $t' > 1$ such that $e_{t'} < \min\{\bar{e}, e_{t'+1}\}$.

Step 2a. We have

$$w_{t'} \leq \frac{\delta}{1-\delta}\left(e_{t'+1} - \frac{\psi\left(e_{t'+1}\right)}{v'\left(\bar{c}_\infty\right)}\right). \tag{17}$$

This follows because (i) $w_{t'} \leq \frac{\psi(e_{t'})}{v'(\bar{c}_\infty)}$ by assumption that payments satisfy condition (FP_t^{un});[7] (ii) $\frac{\psi(e_{t'})}{v'(\bar{c}_\infty)} \leq \frac{\delta}{1-\delta}\left(e_{t'} - \frac{\psi(e_{t'})}{v'(\bar{c}_\infty)}\right)$, by assumption that $e_{t'} < \bar{e}$ and by Step 1, and (iii) $e_{t'} - \frac{\psi(e_{t'})}{v'(\bar{c}_\infty)} \leq e_{t'+1} - \frac{\psi\left(e_{t'+1}\right)}{v'(\bar{c}_\infty)}$ because $z\left(v'\left(\bar{c}_\infty\right)\right) \geq e_{t'+1} > e_{t'}$ with the first inequality following from Lemma A.1.

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

$$\begin{aligned} w_{t+1} - w_t &= \frac{\bar{c}_{t+1} - \bar{c}_t}{\delta^t(1-\delta)} - \frac{\bar{c}_t - \bar{c}_{t-1}}{\delta^{t-1}(1-\delta)} \\ &\geq \frac{v(\bar{c}_{t+1}) - v(\bar{c}_t)}{\delta^t(1-\delta)v'(\bar{c}_t)} - \frac{v(\bar{c}_t) - v(\bar{c}_{t-1})}{\delta^{t-1}(1-\delta)v'(\bar{c}_t)} \\ &= \frac{\psi(e_{t+1}) - \psi(e_t)}{v'(\bar{c}_t)}, \end{aligned}$$

where we used that v is concave. Hence, we have that $e_{t+1} > e_t$ implies $w_{t+1} > w_t$.

[7] This follows from

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

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

Since t' was chosen so that $e_{t'+1} > e_{t'}$, we have $w_{t'+1} > w_{t'}$. Hence,

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

where the second inequality uses (i) Equation (17) from Step 2a, and (ii) the principal's constraint (PC_t) in period $t'+1$. The third inequality uses that $w_{t'+1} \leq \frac{\psi(e_{t'+1})}{v'(\bar{c}_\infty)}$, which again follows from Equation ($\mathrm{FP}_t^{\mathrm{un}}$).

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 A.3, 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 $(e_t)_{t\geq 1}$ is a weakly decreasing sequence. Therefore, for all t, $e_t \geq e_\infty \equiv \lim_{s\to\infty} e_s$.*

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

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

Because payments satisfy ($\mathrm{FP}_t^{\mathrm{un}}$), we have for all t,

$$w_t \in \left(\frac{\psi(e_t)}{v'(\bar{c}_{t-1})}, \frac{\psi(e_t)}{v'(\bar{c}_t)} \right). \tag{18}$$

Thus, for any $s > t^*(t')$,

$$e_{t^*(t')} - w_{t^*(t')} > e_{t^*(t')} - \frac{\psi(e_{t^*(t')})}{v'(\bar{c}_{t^*(t')})} \geq e_{t^*(t')} - \frac{\psi(e_{t^*(t')})}{v'(\bar{c}_{s-1})} \geq e_s - \frac{\psi(e_s)}{v'(\bar{c}_{s-1})} > e_s - w_s. \tag{19}$$

The first inequality follows from Equation (18); 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 $e_s \leq e_{t^*(t')}$ for $s > t^*(t')$ by definition of $t^*(t')$. The fourth inequality follows because $w_s > \frac{\psi(e_s)}{v'(\bar{c}_{s-1})}$ by our previous observation, i.e. Equation (18).

Equation (19) implies that

$$e_{t^*(t')} - w_{t^*(t')} > (1-\delta) \sum_{s=t^*(t')+1}^{\infty} \delta^{s-t^*(t')-1} (e_s - w_s),$$

so that

$$\begin{aligned}
\sum_{s=t^*(t')}^{\infty} \delta^{s-t^*(t')} (e_s - w_s) &= e_{t^*(t')} - w_{t^*(t')} + \delta \sum_{s=t^*(t')+1}^{\infty} \delta^{s-t^*(t')-1} (e_s - w_s) \\
&> (1-\delta) \sum_{s=t^*(t')+1}^{\infty} \delta^{s-t^*(t')-1} (e_s - w_s) + \delta \sum_{s=t^*(t')+1}^{\infty} \delta^{s-t^*(t')-1} (e_s - w_s) \\
&= \sum_{s=t^*(t')+1}^{\infty} \delta^{s-t^*(t')-1} (e_s - w_s). \qquad (20)
\end{aligned}$$

Recall from Lemma A.3 that the principal's constraint must bind at $t^*(t') - 1$ (since $e_{t^*(t')} > e_{t^*(t')-1}$ by the definition of $t^*(t')$). The inequality (20), then implies that $w_{t^*(t')-1} > w_{t^*(t')}$. But then, recalling Equation (18), we have

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

Hence, $e_{t^*(t')-1} > e_{t^*(t')}$, 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 $e_{\bar{t}} > e_{\bar{t}+1}$. Hence, by Lemma A.3, 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 (19) imply that $e_{\bar{t}+1} - w_{\bar{t}+1} > e_s - w_s$ 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 $w_{\bar{t}} > w_{\bar{t}+1}$. Indeed, because the principal's constraint binds at $\bar{t}$,

$$\begin{aligned}
w_{\bar{t}} &= \delta\left(e_{\bar{t}+1} - w_{\bar{t}+1} + \delta \sum_{s=\bar{t}+2}^{\infty} \delta^{s-\bar{t}-2}(e_s - w_s)\right) \\
&> \delta\left((1-\delta) \sum_{s=\bar{t}+2}^{\infty} \delta^{s-\bar{t}-2}(e_s - w_s) + \delta \sum_{s=\bar{t}+2}^{\infty} \delta^{s-\bar{t}-2}(e_s - w_s)\right) \\
&= \sum_{s=\bar{t}+2}^{\infty} \delta^{s-\bar{t}-1}(e_s - w_s) \\
&\geq w_{\bar{t}+1}.
\end{aligned}$$

The final inequality follows from the principal's constraint (PC_t) at date $\bar{t}+1$. Using (18), we have $e_{\bar{t}+1} < e_{\bar{t}}$. Hence, by Lemma A.3, 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. $\square$

The above lemma 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 ($\mathrm{FP}_t^{\mathrm{un}}$) holds at all dates t, and which has the same effort profile. Hence any optimal contract (whether or not condition ($\mathrm{FP}_t^{\mathrm{un}}$) 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 binds under the payment profile satisfying ($\mathrm{FP}_t^{\mathrm{un}}$) 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-enforcing 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* $\frac{\psi(e^{FB}(b_1))}{v'(c^{FB}(b_1))} \leq \frac{\delta}{1-\delta}\left(e^{FB}(b_1) - \frac{\psi(e^{FB}(b_1))}{v'(c^{FB}(b_1))}\right)$. *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 ($\mathrm{FP}_t^{\mathrm{un}}$), 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 ($\mathrm{FP}_t^{\mathrm{un}}$) 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 ealier. In this case, effort is constant from the initial date up to $\bar{t}-1$ (by Lemma A.3). Second, we might have that the principal's constraint (PC_t) is slack at all dates. Effort is then constant (by Lemma A.3), but not first-best. Letting e_∞ 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 $v'(\bar{c}_\infty) > \psi'(e_\infty)$.

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

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

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

$$w_t > \sum_{s=t+1}^{\infty} \delta^{s-t}(e_\infty - w_t),$$

so the principal's constraint is violated at t. If instead $\frac{\psi(e_\infty)}{v'(\bar{c}_\infty)} < \frac{\delta}{1-\delta}\left(e_\infty - \frac{\psi(e_\infty)}{v'(\bar{c}_\infty)}\right)$, we have w_t remains bounded below $\sum_{s=t+1}^{\infty} \delta^{s-t}(e_\infty - w_t)$. 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 the condition (21) can be written as

$$\frac{\psi(e_\infty)}{v'(\bar{c}_\infty)} = \delta e_\infty.$$

Because ψ is strictly convex, we have

$$\frac{\psi'(e_\infty)}{v'(\bar{c}_\infty)} > \delta.$$

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 $\psi(e_\infty + \varepsilon) - \psi(e_\infty)$. 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

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

or

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

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

$$\frac{\psi(e_\infty - \kappa(\varepsilon))}{v'(\bar{c}_\infty(\varepsilon,\kappa(\varepsilon)))} - \delta(e_\infty - \kappa(\varepsilon)) = 0. \tag{22}$$

The derivative of the left-hand side of (22) with respect to $\kappa(\varepsilon)$, evaluated at $(\varepsilon,\kappa(\varepsilon)) = (0,0)$, is

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

This is strictly negative, using that $\frac{\psi'(e_\infty)}{v'(\bar{c}_\infty(0,0))} > \delta$. The derivative of the left-hand side of (22) instead with respect to ε, evaluated at $(\varepsilon,\kappa(\varepsilon)) = (0,0)$, is

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

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

$$\begin{aligned}\kappa'(0) &= \frac{v''(\bar{c}_\infty(0,0))\left(\frac{(1-\delta)\psi'(e_\infty)}{v'(\bar{c}_\infty(0,0))^3}\right)\psi(e_\infty)}{\delta - \frac{\psi'(e_\infty)}{v'(\bar{c}_\infty(0,0))} + v''(\bar{c}_\infty(0,0))\left(\frac{\delta^{T-1}\psi'(e_\infty)}{v'(\bar{c}_\infty(0,0))^3}\right)\psi(e_\infty)} \\ &< \frac{1-\delta}{\delta^{T-1}}\end{aligned} \tag{23}$$

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

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 (23) 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. The principal's payoff increases by

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

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

Note next that, when there is no self-enforcing 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-enforcing 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 policy, effort is constant over at least the first two dates (i.e., $e_1 = e_2$).*

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-enforcing 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^*. For any $\varepsilon > 0$ and period T, there exists $\hat{\delta}(T, \varepsilon)$ such that, for $\delta \in \left(\hat{\delta}(T, \varepsilon), \delta^*\right)$, $\max_{t \leq T} |e_t - e^*| \leq \varepsilon$, where $(e_t)_{t\geq 1}$ is the optimal effort policy for parameter δ.

By Lemma A.1, considering $\delta \leq \delta^*$, any optimal effort policy is contained in $[0, z(v'(\gamma))]^\infty$. The principal's payoff under a self-enforcing relational contract with arbitrary effort policy $(e_t)_{t=1}^\infty$ (and satisfying Condition (FP_t^{un})) is

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

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

Fix $\delta = \delta^*$, and fix any $\varepsilon > 0$ and any period T. We then have that, for an effort policy $(e_t)_{t\geq 1}$ satisfying $\max_{t \leq T} \left|e_t - e^{FB}(b_1)\right| \geq \varepsilon$, 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 $(e_t)_{t\geq 1}$. 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 $\delta \in \left(\hat{\delta}(T, \varepsilon), \delta^*\right)$. $\square$

(End of the proof of Proposition 4.2.) $\square$

Proof of Proposition 4.3.

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

$$\frac{v(\bar{c}_{t'-1})}{1-\delta} - \sum_{s=1}^{t'-1} \delta^{s-1}\psi(e_s) < \frac{v(\bar{c}_\infty)}{1-\delta} - \sum_{t=1}^{\infty} \delta^{t-1}\psi(e_t) = \frac{v((1-\delta)\, b_1)}{1-\delta}.$$

We can then increase $w_{t'-1}$ by $\varepsilon > 0$ and decrease $w_{t'}$ 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-enforcing 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 A.3. That is, given Proposition 4.2 and the assumption that $t' > \bar{t}$ implies that $e_{t'} < e_{t'-1}$, 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 $(e'_t)_{t\geq 1}$ and payments $(w'_t)_{t\geq 1}$ satisfying Equation (FP_t^{un}) that coincides with $(e_t)_{t\geq 1}$ except in periods $t'-1$ and t', so we have $e_{t'} < e'_{t'} \leq e'_{t'-1} < e_{t'-1}$ and

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

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

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

Also, $w'_{t'-1} < w_{t'-1}$ and $w'_{t'} > w_{t'}$. 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 Lemma 5.1

Proof. Fix an optimal contract $(e_t, c_t, w_t, b_t)_{t\geq 1}$ and suppose that Condition (7) is not satisfied for all t. Since the contract is incentive compatible, we have, for all t, Equation ($\mathrm{AC}_t^{\mathrm{ob}}$) 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 b_t by ε and, for appropriately chosen ε, the constraint ($\mathrm{AC}_t^{\mathrm{ob}}$) 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.[8] Constraints are also unaffected from $t+1$ onwards, since the contract remains unchanged at these dates (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 ($\mathrm{AC}_t^{\mathrm{ob}}$) holds as a strict inequality, yielding a contract for which ($\mathrm{AC}_t^{\mathrm{ob}}$) holds at all dates.

If the inequality is strict at $t = 1$, then both c_1 and w_1 can be reduced by the same small amount $\varepsilon > 0$, leaving 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 $t > 1$, 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. □

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

Proof of Proposition 5.1

Proof. Lemma 5.1 implies that, when constructing an optimal contract (satisfying fastest payments), we only need to keep track of the balance b_t (note that the agent's continuation value is $\frac{1}{1-\delta}v((1-\delta)b_t)$). As a result, we can use dynamic programming techniques and maximize the principal's payoff "period by period". We continue the proof by rewriting the principal's problem in the following useful, recursive way. Define first the function $\tilde{e}$ as follows:

$$e_t = \tilde{e}(c_t, b_t, b_{t+1}) \equiv \psi^{-1}\Big(v(c_t) + \tfrac{\delta}{1-\delta}v\left((1-\delta)\,b_{t+1}\right) - \tfrac{1}{1-\delta}v((1-\delta)b_t)\Big). \tag{24}$$

So we may substitute the agent's indifference condition, writing

$$V(b_t) = \max_{c_t, b_{t+1}} \left(\tilde{e}(c_t, b_t, b_{t+1}) - (\delta b_{t+1} - b_t + c_t) + \delta V(b_{t+1})\right) \tag{25}$$

subject to the principal's constraint

$$\delta b_{t+1} - b_t + c_t \leq \delta V\left(b_{t+1}\right) \tag{26}$$

and to the requirement that

$$v(c_t) + \tfrac{\delta}{1-\delta}v\left((1-\delta)\,b_{t+1}\right) - \tfrac{1}{1-\delta}v((1-\delta)b_t) \geq 0. \tag{27}$$

The continuation of the proof of the Proposition 5.1 focusses on contracts satisfying the fastest payments condition ($\mathrm{FP}_t^{\mathrm{ob}}$) for all t. It consists of seven lemmas. Lemmas A.10 to A.14 take existence of an optimal policy as given, while Lemma A.15 establishes this existence.

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 bounds the marginal disutility of effort as a function of the marginal utility of consumption.
3. Lemma A.11 proves the validity of the Euler equation and shows that consumption is weakly decreasing in time.
4. Lemma A.12 shows that either the optimal contract gives the principal his first-best payoff and the balance is constant over time, or the balance is strictly decreasing towards some $b_\infty > 0$.
5. Lemma A.13 shows that if an optimal contract does not achieve the first best, then the specified sequences of effort and payments strictly increase over time.
6. Lemma A.14 shows that if an optimal contract does not achieve the first best, then the specified sequence of consumption strictly decreases over time.

7. Lemma A.15 shows that an optimal contract exists.

Lemma A.9. *Fix an optimal relational contract* $(e_t, c_t, w_t, b_t)_{t\geq 1}$*. For all* t*,* $V(b_t) \in (0, V^{FB}(b_t)]$*. If* $V(b_t) = V^{FB}(b_t)$*, 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. If the first best is attainable at a given date t and balance b_t, then (using Proposition 3.1) it is attained by constant consumption equal to $c^{FB}(b_t)$ and constant effort equal to $e^{FB}(b_t)$. 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 b_t. Hence, the contract is stationary.

We now show that $V(b_t) > 0$ irrespective of the value $b_t > 0$. For this, we need to state the conditions for a stationary contract (i.e., $(e_\tau, c_\tau, w_\tau, b_\tau)_{\tau\geq t}$ with $(e_\tau, c_\tau, w_\tau, b_\tau) = (e, c, w, b)$ for all $\tau \geq t$) to be self-enforcing. The principal's constraint (PC_t) at any date may be written

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

The agent is willing to remain obedient to the agreement if and only if

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

Hence, the contract is self-enforcing if and only if $e \in [\frac{1}{\delta}(c - (1-\delta)b), \psi^{-1}(v(c) - v((1-\delta)b))]$.

Now, fixing b_t, let $\hat{e}(c) \equiv \psi^{-1}(v(c) - v((1-\delta)b_t))$. Note that $\hat{e}((1-\delta)b_t) = 0$ and $\hat{e}'((1-\delta)b_t) = +\infty$. Therefore, it is clear that a self-enforcing stationary contract giving the principal a strictly positive payoff exists, so $V(b_t) > 0$. □

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

Proof. **Proof that** $e_t, c_t, w_t, b_t > 0$ **for all** t**.** We first prove that $w_t > 0$ for all t. We do this assuming, for the sake of contradiction, that $w_t = 0$ for some t. This implies that $e_t = 0$, $c_t = (1-\delta)b_t$ and $b_{t+1} = b_t$ (this is the only possibility for Condition (27) to be satisfied). This implies $V(b_t) = \delta V(b_t)$, that is, $V(b_t) = 0$, but this contradicts Lemma A.9.

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

$$\hat{c}_t = v^{-1}\big(\psi(\varepsilon) - \tfrac{\delta}{1-\delta}v((1-\delta)b_{t+1}) + \tfrac{1}{1-\delta}v((1-\delta)b_t)\big);$$

and raise the principal's date-t payment to $\hat{w}_t = w_t + \hat{c}_t - c_t$. 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 $\hat{w}_t < \delta V(b_{t+1})$, and the principal's payoff strictly increases. Also, the agent remains willing to be obedient at all periods. If instead $w_t = \delta V(b_{t+1})$, we have $V(b_t) = 0$, but this contradicts Lemma A.9.

That $c_t, b_t > 0$ 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 $\psi'(e_t) \leq v'(c_t)$ for all t, and $\psi'(e_t) < v'(c_t)$ only if $w_t = \delta V(b_{t+1})$. Define

$$\underline{c}(b_t, b_{t+1}) \equiv v^{-1}\left(\tfrac{1}{1-\delta} v((1-\delta)b_t) - \tfrac{\delta}{1-\delta} v\left(\left(1-\delta\right) b_{t+1}\right)\right),$$

interpreted as the lowest consumption level that permits the agent's constraint to be satisfied, for fixed values of b_t and b_{t+1}. Consider the problem of maximizing

$$\tilde{e}(\hat{c}_t, b_t, b_{t+1}) - (\delta b_{t+1} - b_t + \hat{c}_t) + \delta V(b_{t+1}) \tag{28}$$

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

$$\psi'(\tilde{e}(c^*(b_t, b_{t+1}), b_t, b_{t+1})) = v'(c^*(b_t, b_{t+1})).$$

If we have $\delta b_{t+1} - b_t + c^*(b_t, b_{t+1}) \leq \delta V(b_{t+1})$ then it is clear that optimality requires that $c_t = c^*(b_t, b_{t+1})$. Otherwise, given the concavity of (28), we must have

$$c_t = \delta V(b_{t+1}) - \delta b_{t+1} + b_t < c^*(b_t, b_{t+1})$$

and hence $w_t = \delta V(b_{t+1})$. In this case,

$$e_t = \tilde{e}(c_t, b_t, b_{t+1}) < \tilde{e}(c^*(b_t, b_{t+1}), b_t, b_{t+1}),$$

and so we have $\psi'\left(e_t\right) < v'(c^*(b_t, b_{t+1})) < v'\left(c_t\right)$. □

The following result establishes the montonicity of the consumption plan.

Lemma A.11. *Any optimal contract $(e_t, c_t, w_t, b_t)_{t \geq 1}$ satisfies the Euler equation*

$$1 - \frac{v'\left(\left(1-\delta\right) b_{t+1}\right)}{v'\left(c_t\right)} = \frac{v'\left(c_{t+1}\right)}{\psi'\left(e_{t+1}\right)} \left(1 - \frac{v'\left(\left(1-\delta\right) b_{t+1}\right)}{v'\left(c_{t+1}\right)}\right) \tag{29}$$

in all periods. Furthermore, $c_t \geq c_{t+1} > (1-\delta) b_{t+1}$ for all t.

Proof. We divide the proof in 3 steps:

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

$$v(\hat{c}_t) - \psi(\hat{e}_t) + \tfrac{\delta}{1-\delta} v\big(\tfrac{1-\delta}{\delta}(b_t + \hat{w}_t - \hat{c}_t)\big) = \tfrac{1}{1-\delta} v((1-\delta)b_t), \tag{30}$$

$$v\big(\tfrac{1}{\delta}(b_t + \hat{w}_t - \hat{c}_t) + \hat{w}_{t+1} - \delta b_{t+2}\big) - \psi(\hat{e}_{t+1}) + \tfrac{\delta}{1-\delta} v((1-\delta)b_{t+2}) = \tfrac{1}{1-\delta} v\big(\tfrac{1-\delta}{\delta}(b_t + \hat{w}_t - \hat{c}_t)\big), \tag{31}$$

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

Fix $\hat{e}_t = e_t$ and $\hat{w}_{t+1} = w_{t+1}$. Equations (30) and (31) implicitly define $\hat{e}_{t+1}$ and $\hat{w}_t$ as functions of $\hat{c}_t$. Let these functions be denoted $\tilde{e}_{t+1}(\cdot)$ and $\tilde{w}_t(\cdot)$, respectively. We can use the implicit function theorem to compute the derivatives at $\hat{c}_t = c_t$:

$$\tilde{e}'_{t+1}(c_t) = \frac{v'(c_t)\big(v'((1-\delta)b_{t+1}) - v'(c_{t+1})\big)}{\delta \psi'\left(\tilde{e}_{t+1}(c_t)\right) v'((1-\delta)b_{t+1})} \quad \text{and} \quad \tilde{w}'_t(c_t) = 1 - \frac{v'(c_t)}{v'((1-\delta)b_{t+1})}.$$

If $\hat{c}_t$ is chosen to be equal to $c_t + \varepsilon$, for some (positive or negative) ε small, the total effect on the continuation payoff of the principal at time t is $(-\tilde{w}'_t(c_t) + \delta \tilde{e}'_{t+1}(c_t))\varepsilon + o(\varepsilon)$ (where $o(\varepsilon)$ represents terms that vanish faster than ε as $\varepsilon \to 0$). Hence, a necessary condition for optimality is that $-\tilde{w}'_t(c_t) + \delta \tilde{e}'_{t+1}(c_t) = 0$, which is equivalent the Euler equation (29).

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

1. If both sides of the Euler equation are strictly positive, then $c_t < c_{t+1} < (1-\delta)b_{t+1}$. In this case, since $e_{t+1} > 0$ (from Lemma A.10), we have $b_{t+2} > b_{t+1}$ (from the agent's constraint (24)). If the constraint does not bind at $t+2$, then $c_{t+2} = c_{t+1}$, and if it binds, $c_{t+1} < c_{t+2} < (1-\delta)b_{t+2}$.

2. If both sides of the Euler equation are zero, then $c_t = c_{t+1} = (1-\delta)b_{t+1}$.

3. If both sides of the Euler equation are strictly negative, then $c_t > c_{t+1} > (1-\delta)b_{t+1}$.

Step 2: We now prove that if $c_t \leq c_{t+1} \leq (1-\delta)b_{t+1}$ then $c_s \leq c_{s+1} < (1-\delta)b_{s+1}$ for all $s > t$. Assume first that there is a period t such that $c_{t+1} \leq (1-\delta)b_{t+1}$. Hence, since $e_{t+1} > 0$, we have $b_{t+2} > b_{t+1}$. This shows that the Euler satisfies

$$1 - \frac{v'\left((1-\delta)\, b_{t+2}\right)}{v'\left(c_{t+1}\right)} = \frac{v'\left(c_{t+2}\right)}{\psi'\left(e_{t+2}\right)} \left(1 - \frac{v'\left((1-\delta)\, b_{t+2}\right)}{v'\left(c_{t+2}\right)}\right) > 0.$$

Since $v'(c_{t+2})/\psi'(e_{t+2}) \geq 1$, $(1-\delta)b_{t+2} > c_{t+2} \geq c_{t+1}$. The result then follows by induction.

Step 3: We prove that $c_t > (1-\delta)b_t$ for all $t > 1$, and therefore it is (weakly) decreasing in t. Assume then, for the sake of contradiction, that there is a $t' > 1$ such that $c_{t'} \leq (1-\delta)b_{t'}$. Then $c_{t'-1} \leq c_{t'}$ (either $v'(c_{t'}) = \psi'(e_{t'})$ and so $c_{t'-1} = c_{t'}$, or $v'(c_{t'}) > \psi'(e_{t'})$ and Case 1 or Case 2 from above applies). Hence, from Step 2, we have that $(1-\delta)b_{s+1} > c_{s+1} \geq c_s$ for all $s \geq t'$.

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

$$\sum_{s=t'}^{\infty} \delta^{s-t'} v(c_s) > \tfrac{1}{1-\delta} v((1-\delta)b_{t'}),$$

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

Consider the original contract, and note that, for all $s \geq 1$, $b_s + \sum_{\tau=s}^{\infty} \delta^{\tau-s} w_\tau = \sum_{\tau=s}^{\infty} \delta^{\tau-s} c_\tau$. Note that since (by Step 2) consumption is weakly increasing from t' onwards these quantities 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

$$\frac{\bar{c}}{1-\delta} = \sum_{\tau=t''}^{\infty} \delta^{\tau-t''} c_\tau$$

which implies that $\frac{\bar{c}}{1-\delta} < \sum_{\tau=s}^{\infty} \delta^{\tau-s} c_\tau$ 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 b_s for $s \leq t''$, but differ for $s > t''$. We have for all $s > t''$

$$b_s^{new} + \sum_{\tau=s}^{\infty} \delta^{\tau-s} w_\tau = \frac{\bar{c}}{1-\delta} < \sum_{\tau=s}^{\infty} \delta^{\tau-s} c_\tau = b_s + \sum_{\tau=s}^{\infty} \delta^{\tau-s} w_s.$$

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

Now, we want to investigate the agent's constraint in each period $s \geq 1$. We wish to examine the agent's willingness to continue in the contract, rather than "quit", i.e. to optimally deviate by exerting no effort and smoothing consumption. Note first that, for all $s \leq t''$, the agent anticipates a strictly higher continuation payoff under the new contract, but the value of quitting optimally at date s remains the same. Hence, the agent strictly prefers to continue under the contract than quit at all such dates.

To understand how the agent's payoff changes at $s > t''$, we consider an "intermediate" adjustment to the contract. In particular, we consider adjusting the original contract from date s onwards by letting the agent optimally/efficiently smooth consumption from date s onwards. We denote the smoothed consumption $\bar{c}_s(b_s)$. This is equal to $(1-\delta)\sum_{\tau=s}^{\infty}\delta^{\tau-s}c_\tau = (1-\delta)\left(b_s + \sum_{\tau=s}^{\infty}\delta^{\tau-s}w_\tau\right)$.

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

$$\sum_{\tau=s}^{\infty}\delta^{\tau-s}\left(v\left(\bar{c}_s(b_s)\right)-\psi(e_\tau)\right) \geq \sum_{\tau=s}^{\infty}\delta^{\tau-s}\left(v(c_\tau)-\psi(e_\tau)\right) \geq \frac{v\left(b_s(1-\delta)\right)}{1-\delta}.$$

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

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

Note that $\bar{c} = \bar{c}_s(b_s)-(1-\delta)\left(b_s-b_s^{new}\right)$. Therefore, we can now write, for $s > t''$,

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

This shows that, under the new contract, where the consumption is smoothed from date t'' onwards (with the agent consuming $\bar{c}$), the agent prefers to continue in the contract from date s onwards, rather than to quit at date s and optimally smooth consumption.

We have thus defined a new policy with consumption given by $c_s^{new} = c_s$ 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 w_s and e_s for all s. Finally, the agent's balances are given by $b_s^{new} = b_s$ 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 does not bind for $s \leq t''$ (this follows because $\sum_{\tau=t''}^{\infty}\delta^{\tau-t''}v\left(c_\tau^{new}\right) = \sum_{\tau=t''}^{\infty}\delta^{\tau-t''}v(\bar{c}) > \sum_{\tau=t''}^{\infty}\delta^{s'-t}v(c_\tau)$ by strict concavity of v).

Finally, we can define a policy 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. For dates $t''+1$ onwards, the continuation contract remains unchanged (i.e., given by $\left(e_s^{new}, c_s^{new}, w_s^{new}, b_s^{new}\right)_{s>t''}$). $\square$

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

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. Suppose from now on that $V(b_1) < V^{FB}(b_1)$.

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

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

$$\begin{aligned} c'_t = \bar{c} &\equiv (1-\delta)\sum_{\tau\geq\hat{t}} \delta^{\tau-\hat{t}} w_\tau + (1-\delta) b_{\hat{t}} \\ &= (1-\delta)\sum_{\tau\geq\hat{t}} \delta^{\tau-\hat{t}} c_\tau. \end{aligned}$$

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

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

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

Step 1b. We now want to show that the agent's constraint ($\mathrm{AC}_t^{\mathrm{ob}}$) 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 $c_{\hat{t}} \geq \bar{c}$. Therefore,

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

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 ($\mathrm{AC}_t^{\mathrm{ob}}$) at date $\hat{t}+1$, plus the observation that $\bar{b} < b_{\hat{t}+1}$, then implies

$$\sum_{\tau\geq\hat{t}+1} \delta^{\tau-\hat{t}-1} v(\bar{c}) - \sum_{\tau\geq\hat{t}+1} \delta^{\tau-\hat{t}-1}\psi(\bar{e}) > \frac{v\left((1-\delta)\,\bar{b}\right)}{1-\delta}, \tag{32}$$

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

Note then that

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

(with a strict inequality if the consumption levels c_τ 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 $\bar{e} \geq (1-\delta)\sum_{\tau\geq\hat{t}+1}\delta^{\tau-\hat{t}-1}e_\tau$. Therefore,

$$\begin{aligned}\sum_{\tau\geq\hat{t}+1}\delta^{\tau-\hat{t}}e'_\tau-\sum_{\tau\geq\hat{t}}\delta^{\tau-\hat{t}}w'_\tau &= \frac{\delta\bar{e}}{1-\delta}-\sum_{\tau\geq\hat{t}}\delta^{\tau-\hat{t}}w_\tau\\ &\geq \sum_{\tau\geq\hat{t}+1}\delta^{\tau-\hat{t}}e_\tau-\sum_{\tau\geq\hat{t}}\delta^{\tau-\hat{t}}w_\tau\\ &\geq 0 \end{aligned}\tag{33}$$

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 e'_t is constant for $t \geq \hat{t}+1$, and because 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

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

which is satisfied because (i) $e'_\tau = e_\tau$ for $\tau \leq \hat{t}$, and $w'_\tau = w_\tau$ for $\tau < \hat{t}$, (ii) the first inequality in Equation (33) 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

$$v\left(c_{\hat{t}}\right)-\psi\left(e_{\hat{t}}\right)=\frac{v\left(\left(1-\delta\right)\bar{b}\right)-\delta v\left(\left(1-\delta\right)b_{\hat{t}+1}\right)}{1-\delta}<v\left(\left(1-\delta\right)\bar{b}\right).$$

Hence,

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

where the final inequality follows from (32). Hence $e_{\hat{t}} > \bar{e}$. Recall that $e'_{\hat{t}} = e_{\hat{t}}$, and $e'_\tau = \bar{e}$ for $\tau > \hat{t}$; so we have $e'_{\hat{t}} > e'_\tau$ for all $\tau > \hat{t}$.

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

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

and such that

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

Substitute $e''_{\hat{t}}$ for $e'_{\hat{t}}$ and $e''_{\hat{t}+1}$ for e'_τ, 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 ($\mathrm{AC}_t^{\mathrm{ob}}$) remains satisfied at $\hat{t}$, and at all earlier dates. Note that, due to (32), the agent's constraints ($\mathrm{AC}_t^{\mathrm{ob}}$) 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 $V(b_1) < V^{FB}(b_1)$ then $(b_t)_{t\geq 1}$ is a strictly decreasing sequence.

Step 2a. We first prove that if $b_t = b_{t+1}$ then $V(b_t) = V^{FB}(b_t)$. To do this, note that if $b_t = b_{t+1}$, then it is optimal to specify $c_\tau = c_t$, $w_\tau = w_t$, and $e_\tau = e_t$ for all $\tau > t$; that is, it must be optimal for the contract to be stationary from period t onwards. The Euler equation (29) then requires that $\psi'(e_\tau) = v'(c_\tau)$ for all $\tau \geq t+1$,[9] and by stationarity also $\psi'(e_t) = v'(c_t)$. Then, e_τ and c_τ satisfy, for all $\tau \geq t$, the first-order and agent's indifference conditions in Proposition 3.1, given initial balance b_t. Therefore they are the first-best effort and consumption given balance b_t. This shows that $V(b_t) = V^{FB}(b_t)$, as desired.

Step 2b. We now prove that if $V(b_1) < V^{FB}(b_1)$ then $V(b_t) < V^{FB}(b_t)$ for all $t \geq 1$ and, in addition, $(b_t)_{t\geq 1}$ is strictly decreasing. Suppose that $V(b_t) < V^{FB}(b_t)$, which by Step 1 and Step 2a implies $b_{t+1} < b_t$. 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 b_{t+1}. This implies that $e_\tau = e^{FB}(b_{t+1})$ and $c_\tau = c^{FB}(b_{t+1})$ for all $\tau > t$. By assumption that the agent's constraint ($\mathrm{AC}_t^{\mathrm{ob}}$) is satisfied with equality in all periods, we then have $b_\tau = b_{t+1}$ for all $\tau > t+1$. Hence, the contract is stationary from $t+1$ onwards; in particular, the payment is constant at $w_\tau = \bar{w}$ for $\tau \geq t+1$, for some value $\bar{w}$.

From the Euler equation (29) and the fact that $v'(c_{t+1}) = \psi'(e_{t+1})$, we have $c_t = c_{t+1}$.

[9] To see this, recall from Lemma A.11 that $c_\tau > b_\tau$.

Hence, using $b_{t+2} = b_{t+1} < b_t$, we have (using (FP$_t^{\text{ob}}$))

$$\begin{aligned}\psi(e_t) &= v(c_t) + \tfrac{\delta}{1-\delta}v((1-\delta)b_{t+1}) - \tfrac{1}{1-\delta}v((1-\delta)b_t) \\ &< v(c_{t+1}) + \tfrac{\delta}{1-\delta}v((1-\delta)b_{t+2}) - \tfrac{1}{1-\delta}v((1-\delta)b_{t+1}) = \psi(e_{t+1}).\end{aligned}$$

Consequently, $e_t < e_{t+1}$, and so $\frac{\psi'(e_t)}{v'(c_t)} < \frac{\psi'(e_{t+1})}{v'(c_{t+1})} = 1$. We then know that the principal's constraint (PC$_t$) binds at t, and so

$$\sum_{s=t+1}^{\infty} \delta^{s-t} e_s = \sum_{s=t}^{\infty} \delta^{s-t} w_s = \sum_{s=t}^{\infty} \delta^{s-t} c_s - b_t.$$

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

$$\delta e^{FB}(b_{t+1}) = c^{FB}(b_{t+1}) - (1-\delta)b_t < c^{FB}(b_{t+1}) - (1-\delta)b_{t+1} = \bar{w}.$$

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 $b_\infty > 0$ and $V(b_\infty) = V^{FB}(b_\infty)$.

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

Step 3b. We now prove that $b_\infty > 0$. We show that $\lim_{b \searrow 0} \frac{c^{FB}(b) - (1-\delta)b}{e^{FB}(b)} = 0$ and so, by Equation (8), there exists some $\bar{b}$ such that an optimal contract achieves the first-best payoff of the principal for all $b \leq \bar{b}$. Since $v(c^{FB}(b)) - v((1-\delta)b) = \psi(e^{FB}(b)) > 0$, 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 $\lim_{b\searrow 0} e^{FB}(b) = +\infty$. Thus we have established that, given $V(b_1) < V^{FB}(b_1)$, the sequence $(b_t)_{t\geq 1}$ is strictly decreasing and convergent to some positive value b_∞.

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

$$V(b_\infty) = \tilde{e}(c_\infty, b_\infty, b_\infty) = \psi^{-1}(v(c_\infty) - v((1-\delta)b_\infty)).$$

Since $V(b_\infty) > 0$ (recall Lemma (A.9)), $c_\infty > (1-\delta)b_\infty$. Therefore, the Euler equation (29) implies that, necessarily, $\lim_{t\to\infty} \frac{v'(c_{t+1})}{\psi'(e_{t+1})} = 1$, and therefore $e_\infty \equiv \lim_{t\to\infty} e_t$ exists. It is then clear that both Conditions 1 and 2 of Proposition 3.1 hold for e_∞, c_∞, and b_∞ (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 $(e_t)_{t\geq 1}$ and $(w_t)_{t\geq 1}$ are strictly increasing.*

Proof. Recall from Lemma A.12 we have that, if $V(b_1) < V^{FB}(b_1)$, then $(b_t)_{t\geq 1}$ 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 $\tilde{b} \in (b^* - \varepsilon, b^*)$ satisfying $V(\tilde{b}) \leq V(b^*)$.

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 $b_t = b^*$, and consider a change to $b_t = b^* - \nu$ for ν arbitrarily small and such that $V(b^* - \nu) \leq V(b^*)$. Then we can reduce c_t by the same amount ν, holding b_{t+1} and w_t, as well as all other variables, constant (that a reduction in date-t consumption is possible follows from Lemma A.11, since $c_t > (1-\delta)b_t$). Note then that, provided ν is small enough,

$$v(c_t - \nu) - \tfrac{1}{1-\delta} v((1-\delta)(b_t - \nu)) > v(c_t) - \tfrac{1}{1-\delta} v((1-\delta)b_t),$$

which follows again because $c_t > (1-\delta)\, b_t$ (by Lemma A.11) and by concavity of v. Hence, we have

$$\tilde{e}(c_t - \nu, b_t - \nu, b_{t+1}) > \tilde{e}(c_t, b_t, b_{t+1}).$$

By construction, the agent remains indifferent to continuing in the contract at all dates (we have that ($\mathrm{AC}_t^{\mathrm{ob}}$) 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 w_t is unchanged). Hence,

$$V(b_t) = e_t - w_t + \delta V(b_{t+1}) < \tilde{e}(c_t - \nu, b_t - \nu, b_{t+1}) - w_t + \delta V(b_{t+1}) \leq V(b_t - \nu).$$

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

Lemma A.14. *If the first-best contract is not implementable given b_1, then consumption strictly declines over time.*

Proof. Consider an optimal relational contract $(e_t, c_t, w_t, b_t)_{t\geq 1}$ not implementing the first best. By Lemma A.11, we know that $c_{t-1} \geq c_t$ for all $t \geq 2$. Hence, if the result fails, we must have $c_{t-1} = c_t$ for some t. We then have, by Equation (29) (and noting that $c_t > (1-\delta)\, b_t$, also by Lemma A.11), that $\psi'(e_t) = v'(c_t)$. Also, by the previous result, $e_{t+1} > e_t$.

Consider then reducing the payment and consumption at date t by a small amount $\varepsilon > 0$, while reducing effort at date $t+1$ by $\nu(\varepsilon)$ so as to keep the agent's payoff in the contract unchanged. This requires

$$v(c_t) - v(c_t - \varepsilon) = \delta\left(\psi(e_{t+1}) - \psi(e_{t+1} - \nu(\varepsilon))\right).$$

We have $\nu(\varepsilon) = \frac{v'(c_t)}{\delta\psi'(e_{t+1})}\varepsilon + o(\varepsilon)$, and hence $\varepsilon - \delta\nu(\varepsilon) = \varepsilon\left(1 - \frac{v'(c_t)}{\psi'(e_{t+1})}\right) + o(\varepsilon)$, which is strictly positive for ε sufficiently small, because $v'(c_t) < \psi'(e_{t+1})$. This shows that the principal's continuation payoff at the moment of making the date-t payment strictly increases. Hence, the principal's date-t constraint (PC_t), and all earlier principal constraints, are relaxed, and profits strictly increase. Since the agent is asked for less effort at date $t+1$, the agent then strictly prefers to continue in the contract at date $t+1$ (the balance at that date is b_{t+1}, as in the original contract). The agent's constraints at all other dates are unaffected. This shows that the original contract cannot have been optimal. $\square$

Lemma A.15. *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 the first best is implementable (this follows from Equation (8)), and so existence is established. The remainder of the proof is needed for the values b_1 such that the first-best is not implementable as part of a self-enforcing agreement.

We denote by $\Pi(b_t)$ the sequences $(c_s, b_{s+1})_{s=t}^{\infty}$ that satisfy, for all $s \geq t$,[10]

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

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

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

We can write the functional equation for the problem as

$$TW(b_t) = \sup_{c_t > 0, b_{t+1} > 0} \left(\tilde{e}(c_t, b_t, b_{t+1}) - (\delta b_{t+1} - b_t + c_t) + \delta W(b_{t+1})\right) \tag{34}$$

subject to the principal's constraint

$$\delta b_{t+1} - b_t + c_t \leq \delta W(b_{t+1}) \tag{35}$$

and to

$$v(c_t) + \tfrac{\delta}{1-\delta} v\left((1-\delta)\, b_{t+1}\right) - \tfrac{1}{1-\delta} v((1-\delta)b_t) \geq 0. \tag{36}$$

Outline of Proof. Note also that the operator T is monotone: if $W_1 \geq W_2$, then $TW_1 \geq TW_2$. Also, we have $TV^{FB} \leq V^{FB}$. Proceeding iteratively, we have that $(T^n V^{FB}(b_t))$ is a decreasing sequence for all $b_t > 0$. Therefore there is some pointwise limit, 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 that respects the principal and agent constraints. 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 $V(b_t) \geq$

[10] We do not impose the third feasibility requirement "boundedness" of the sequences. The arguments above imply that boundedness is satisfied at an optimal contract.

$\bar{V}(b_t)$. Suppose that V is a fixed point of T (which we prove in 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 (34) to (36) 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 $I(b_t) = [l(b_t), u(b_t)]$ with $l(b_t) > 0$, and (b) consumption can be taken to be bounded above. The latter will follow from establishing Part (a) and by our assumption that $\lim_{e\to\infty}\psi'(e) = \infty$.

Observe then that, irrespective of b_t,

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

This implies that, if b_{t+1} is large enough, the principal's constraint (35), is violated. Hence, we can b_{t+1} to be bounded above by some $u(b_t)$.

We now show that, given b_t, satisfaction of the constraints in Equations (35) and (36) implies that b_{t+1} must be no less than some $l(b_t) > 0$. 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

$$V^{FB}(b_{t+1}) \geq \tilde{V}(b_{t+1}) \equiv \frac{v^{-1}\left(\frac{1}{1-\delta} v((1-\delta)b_t) - \frac{\delta}{1-\delta} v((1-\delta)b_{t+1})\right) - b_t}{\delta} + b_{t+1}. \tag{37}$$

Now, notice that the right-hand side of Equation (37) 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

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

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

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

Therefore, by the envelope theorem,

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

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

$$1-\frac{v'((1-\delta)b_{t+1})}{v'\Big(\underbrace{v^{-1}\Big(\frac{1}{1-\delta}v((1-\delta)b_t)-\frac{\delta}{1-\delta}v((1-\delta)b_{t+1})\Big)}\Big)}. \tag{38}$$

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.

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

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

for some $\left(\hat{c}_t,\hat{b}_{t+1}\right)$ satisfying the constraints in the functional equation, i.e. Equations (35) and (36) 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 $\left(\hat{c}_s, \hat{b}_{s+1}\right)_{s=t}^{\infty} \in \Pi(b_t)$. But (for ν small enough) this sequence generates a payoff to the principal higher than $V(b_t)$, contradicting the definition of the latter. Hence, $V(b_t)$ is an upper bound on

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

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

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

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

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

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

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

□

(End of the proof of Proposition 5.1.) □

Proof of Lemma 5.2.

Proof. If this is not the case, then there is a date t such that

$$\frac{v(b_t(1-\delta))}{1-\delta} < \sum_{s=t}^{\infty} \delta^{s-t}\left(v(c_s) - \psi(e_s)\right).$$

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 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 e'_t, and the date-$t+1$ effort to e'_{t+1}, with $e_t < e'_t < e'_{t+1} < e_{t+1}$, and with

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

All other variables remain unchanged. This affects the agent constraints 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. $\square$

B Appendix providing "equilibria" corresponding to self-enforcing agreements

Equilibrium for Section 4 (unobserved consumption). Below, for any "self-enforcing agreement" in the environment of Section 4, we provide strategies and beliefs that we would understand as constituting a Perfect Bayesian Equilibrium according to an appropriate definition of such concept for the environment with private savings.

For $t \geq 0$, a t-history for the agent is $h_t^A = (e_s, w_s, c_s, b_{s+1})_{1 \leq s < t}$, which specifies the observed effort, payment and consumption up until time $t-1$, and balances up to date t (taking the initial balance b_1 as a parameter of the model, rather than an outcome). The set of such histories at date $t \geq 1$ is $\mathcal{H}_t^A = \mathbb{R}_+^{2(t-1)} \times \mathbb{R}^{2(t-1)}$ (with the convention that $\mathbb{R}^0 = \emptyset$). 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 past 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.

Departing from the notation in the main text, let $\left(e_s^*, w_s^*, c_s^*, b_{s+1}^*\right)_{s \geq 1}$ be a self-enforcing agreement. We specify a PBE as follows. On the principal's side, put $\sigma_t = w_t^*$ if $(e_s, w_s) = (e_s^*, w_s^*)$ for all $s \leq t-1$ and $e_t = e_t^*$, and put $\sigma_t = 0$ otherwise.

For the agent's strategy, put $\alpha_t = (e_t^*, c_t^*)$ whenever $(e_s, w_s, c_s, b_{s+1}) = \left(e_s^*, w_s^*, c_s^*, b_{s+1}^*\right)$ for all $s < t$. Put $\alpha_t = (0, b_t (1-\delta))$ whenever $(e_s, w_s) \neq (e_s^*, w_s^*)$ for some $s \leq t-1$. Determining the agent's equilibrium strategy for the remaining possible events, where $(e_s, w_s) = (e_s^*, w_s^*)$ for all $s \leq t-1$, and yet $(c_s, b_{s+1}) \neq \left(c_s^*, b_{s+1}^*\right)$ for some values $s \leq t-1$, is then more involved. It requires determining an optimal continuation strategy for the agent given that there is as yet no public deviation, and yet b_t may be different from b_t^*. Such an optimal strategy involves setting for some $t' \geq t$, $e_s = e_s^*$ for all $t \leq s < t'$, putting $e_{t'} = 0$, and putting for all $s \in \{t, \ldots, t'\}$

$$c_s = (1-\delta)\left(b_t^* + \sum_{\tau=t}^{t'-1} \delta^{\tau-t} w_\tau^*\right).$$

(After date t', the deviation $(e_{t'}, w_{t'}) \neq (e_{t'}^*, w_{t'}^*)$ is publicly observed, so continuation play is determined as above.) Allowing that $t' = \infty$ – i.e., the agent chooses $e_s = e_s^*$ at all $s \geq t$ – the problem of choosing the optimal "public deviation" time t' has a solution. This follows from "continuity at infinity" of the agent's payoff in the public deviation date t'; i.e., because

$$\frac{v\left((1-\delta)\left(b_t^* + \sum_{\tau=t}^{t'-1} \delta^{\tau-t} w_\tau^*\right)\right)}{1-\delta} - \sum_{\tau=t}^{t'-1} \delta^{\tau-t} \psi\left(e_\tau^*\right)$$
$$\to \frac{v\left((1-\delta)\left(b_t^* + \sum_{\tau=t}^{\infty} \delta^{\tau-t} w_\tau^*\right)\right)}{1-\delta} - \sum_{\tau=t}^{\infty} \delta^{\tau-t} \psi\left(e_\tau^*\right)$$

as $t' \to \infty$, which follows by continuity of v and because $\sum_{\tau=t}^{t'-1} \delta^{\tau-t} w_\tau^*$ is convergent to some finite value (since the payments w_τ^* are bounded by assumption).

Finally, the principal's beliefs at each information set, at each date t, may be specified by putting probability one on the agent having consumed c_s^* at each date $s \leq t-1$ in case $(e_s, w_s) = (e_s^*, w_s^*)$ for all $s \leq t-1$, and if $e_t = e_t^*$; the principal then believes $b_{s+1} = b_{s+1}^*$ for all $s \leq t-1$. If instead $(e_s, w_s) \neq (e_s^*, w_s^*)$ for some $s \leq t-1$, or if $e_t \neq e_t^*$, let $t' \leq t$ be the first date at which $(e_s, w_s) \neq (e_s^*, w_s^*)$ for $s \leq t-1$, or if there is no such date, let $t' = t$. Then, if $e_{t'} \neq e_{t'}^*$ (so the agent is first to publicly deviate), let the principal believe that

$$c_s = (1-\delta)\left(b_1 + \sum_{\tau=1}^{t'-1} \delta^{\tau-1} w_\tau^*\right)$$

for all $s \in \{1, \ldots, t'\}$, while, for all $s > t'$, the principal believes that the agent consumes

$$c_s = (1-\delta)\, b_s$$

(with the balance b_s determined mechanically by Equation (1)). If $e_{t'} = e_{t'}^*$ (so the principal is first to publicly deviate), then the principal believes that the agent consumes $c_s = c_s^*$ for all $s \leq t'$ and $c_s = (1-\delta)\, b_s$ for all $s > t'$. 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.

It is then clear that, given $\left(e_s^*, w_s^*, c_s^*, b_{s+1}^*\right)_{s\geq 1}$ is "self-enforcing" as defined in the main text, the strategies specified for the principal and agent are optimal at histories such that no deviation is yet observed. We specified also an optimal continuation strategy for the agent at histories where effort and payments are as specified in the agreement, but only the agent's consumption differed. Finally, at dates when the agent has deviated in choice of effort, or principal has deviated in choice of payment, the agent finds it optimal to choose zero effort and perfectly smooth consumption, while the principal finds it optimal to make no payments. Hence, the above strategies are sequentially optimal.