CESifo WORKING PAPERS

6590 2017

July 2017

Relational Contracts with Private Information on the Future Value of the Relationship: The Upside of Implicit Downsizing Costs

Matthias Fahn, Nicolas Klein

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6590*
Category 11: Industrial Organisation

Relational Contracts with Private Information on the Future Value of the Relationship: The Upside of Implicit Downsizing Costs

Abstract

We analyze a relational contracting problem, in which the principal has private information about the future value of the relationship. In order to reduce bonus payments, the principal is tempted to claim that the value of the future relationship is lower than it actually is. To induce truth-telling, the optimal relational contract may introduce distortions after a bad report. For some levels of the discount factor, output is reduced by more than would be sequentially optimal. This distortion is attenuated over time even if prospects remain bad. Our model thus provides an alternative explanation for indirect short-run costs of downsizing.

JEL-Codes: C730, D860.

Matthias Fahn
LMU Munich
Department of Economics
Munich / Germany
matthias.fahn@econ.lmu.de

Nicolas Klein
Université de Montréal
Department of Economics
Montréal / Québec / Canada
kleinnic@yahoo.com

July 21, 2017
We thank Daniel Barron, Catherine Bobtcheff, Sylvain Chassang, Florian Englmaier, Willie Fuchs, Tahmina Hadjer, Marina Halac, Martin Hellwig, Richard Holden, Johannes Hörner, Sebastian Kranz, Nicolas Lambert, Steve Leider, Jin Li, Thomas Mariotti, Niko Matouschek, Aniko Öry, Mike Powell, Sven Rady, Markus Reisinger, Andreas Roider, Larry Samuelson, Klaus Schmidt, Marco Schwarz, Robert Ulbricht, as well as attendees at seminars at the Universtitat de Barcelona, the MPI Bonn, the ETH Zurich, McGill, Guelph, the Université Laval, the Université de Montréal, Munich, Paris (Séminaire Roy), Regensburg, Toulouse, the UNSW, and at various conferences, for helpful comments. Nicolas Klein gratefully acknowledges financial assistance from the Fonds de Recherche du Québec Société et Culture, and the Social Sciences and Humanities Research Council of Canada.

1 Introduction

In many instances, organizations face difficulties in providing the proper incentives to their members because performance cannot be verified, i.e. enforced by a court. As noted by the literature on relational contracts, however, the mutual dependence that repeated interaction between the same parties fosters may allow contracting parties to overcome these difficulties through the building of mutual trust. This engenders an implicit, or "relational," contract between them, whereby the principal "voluntarily" rewards the agent for his effort. As the worst the agent can do to the principal is to break off the relationship entirely, the most the principal can credibly promise as a reward is the value of the entire future relationship to her.

Our goal here is to analyze the workings of such relational contracts when, at the time of deciding on rewards, the principal knows more about the future development, and hence the value, of the relationship. Indeed, management may e.g. be better informed about the likely evolution of demand for a firm's product than workers. In such a context, workers must trust management not to use its informational advantage against them, e.g. by fraudulently claiming a threat of future demand contraction to cut their bonus payments or even let go of them.

We show that an optimal relational contract in such a setting can lead to a dynamic that has been discussed in the strategic management literature, which has noted that downsizing often seems less effective than originally anticipated.[1] The prevailing explanation for these *implicit downsizing costs* seems to be that surviving employees tend to consider downsizing as a breach of a "psychological contract" (Love and Kraatz (2009)), and thus switch to a kind of 'punishment mode' in response. As Cascio (1993) writes: "Study after study shows that following a downsizing, surviving employees become narrow-minded, self-absorbed, and risk averse. Morale sinks, productivity drops, and survivors distrust management." Love and Kraatz (2009) write: "Though downsizing was perfectly legal and widely advocated as an efficient business practice, it connoted opportunism and signaled that a firm was an untrustworthy actor that might not be counted on to meet its commitments in the future. Employees clearly interpreted downsizing as a betrayal and characterized downsizers as untrustworthy."

[1] See e.g. Cascio (1993) and Datta, Guthrie, Basuil, and Pandey (2010).

Yet, there is some evidence to suggest that this ‘punishment mode’ only lasts for a limited period of time. Indeed, Goesaert, Heinz, and Vanormelingen (2015) show that firm performance tends to drop at the downsizing event, recovering at best to pre-downsizing levels afterwards. The survey paper by Datta, Guthrie, Basuil, and Pandey (2010) quotes several studies showing that the benefits of downsizing, if any, will materialize only 2-3 years after the downsizing event (see bottom of their page 355, and the references cited there). Thus, punishment does not seem to be of the grim-trigger form, which would constitute the harshest, and hence optimal, penal code.[2]

We do not find the prevailing explanations, according to which implicit downsizing costs are a phenomenon that is supposed to occur off the equilibrium path of play, very convincing. Indeed, the very fact that it seems to occur so regularly that it has been noted in the management literature would in our view militate against its being an off-path phenomenon. By contrast, our paper will provide a story generating implicit downsizing costs as an *on-path* phenomenon. They will indeed arise from a relational, or “psychological,” contract, yet not from its breakdown but rather as *part of* the path of play in an optimal relational contract, where they will act as a commitment device only to downsize when it is necessary to do so.

More specifically, our model starts from the standard relational-contracting framework, in which a principal and an agent interact repeatedly over time. The agent has to exert effort to produce output, which translates into a profit for the principal. Effort is costly to the agent. By assuming that the agent is risk neutral and effort costs are linear in the level of effort exerted, we can interpret our agent as representing the firm’s total workforce, which is made up of homogeneous workers.[3] Only one-period contracts are possible; these cannot condition on the agent’s effort choice, which, although observable, contains subjective aspects and is hence not contractible. As effort is perfectly observable by both parties to the relationship, however, continuation play can depend on the level of effort observed. In particular, the principal can pay the agent a discretionary bonus for choosing the right level of effort; this bonus can be enforced by the agent’s threat to leave the relationship if a bonus pay-

[2]See Abreu (1988).

[3]This interpretation presupposes multilateral relational contracts, by which a deviation in the relationship with one agent is punished by a complete loss of trust in all other relationships, see Levin (2002).

ment that was due to him was reneged on. Therefore, the principal can only credibly commit to a bonus that is at most as high as the expected value of the continuation of the relationship to her. The principal's profits, which are generated by the (publicly observable) output in a given period, depend on the binary state of the world in that period, which is only privately observed by the principal. The effort level the principal wants to induce may thus well depend on the current state of the world. The state of the world for the next period is privately observed by the principal before she decides on the current period's bonus. She thus has some private information on the value of the continuation of the relationship when she decides whether to pay out the bonus, or to renege, and thus to end the relationship.

Our analysis shows that, even though there is only one-sided private information, some surplus may optimally be destroyed along the path of play, leading to *implicit downsizing costs*. The goal of this arrangement is to deter the principal from mulcting the agent of the bonus due to him by understating the value of the continuation of the relationship. Indeed, lest the principal be tempted by such a deviation, continuation play following a pessimistic announcement must be rendered sufficiently unattractive. One way of achieving this goal would be to force the principal to pay the agent a transfer whenever the continuation value is low. This, however, turns out not to be optimal in our setting, the reason being that this penalty would hurt a truthful on-path principal and a lying off-path principal alike. By contrast, a distortion in the agent's effort hits an off-path principal, who has falsely claimed that effort is less productive, more than an on-path principal, who has been honest in invoking a low productivity of effort.

In most of the paper, we focus on the case in which the principal's type is iid across periods. In this case, only a distortion in the next period hits an off-path principal more severely than an on-path principal, and consequently, implicit downsizing costs only last for one period. Indeed, the management literature has noted that, at the occurrence of downsizing events, there is often some overshooting in the reduction of labor input, as evidenced by the fact that firms tend to increase labor input again shortly after downsizing, while the firm's environment has not changed.[4]

In Section 5, we extend our analysis to persistent shocks. In this case,

[4]See Cascio (1993).

distortions gradually attenuate over time but only ever vanish in the limit. The reason is that, with persistent shocks, an off-path principal is hit more severely by a distortion in any future period, but the difference in on-path vs. off-path costs diminishes with distance in time.

The idea that repeated interaction endogenously creates some scope for commitment via implicit contracts has been applied to labor markets by Bull (1987), as well as MacLeod and Malcomson (1989).[5] These early papers abstracted from informational asymmetries, focusing instead on the question of how incentives can be governed by non-contractual agreements. Levin (2003) augmented the analysis by introducing informational asymmetries, analysing the cases in which the employee privately knows his effort costs (adverse selection), his effort level can only be imperfectly observed (moral hazard), as well as the case in which the employer privately observes a performance measure, while not observing the agent's effort choice directly. Malcomson (2016) introduces persistent types into Levin's (2003) adverse-selection model, and finds that a full separation of types is not possible when continuation payoffs are on the Pareto frontier. Malcomson (2015) augments Levin's (2003) adverse-selection model by the introduction of different principal-types denoting the productivity of the agent's effort in the *current* period. At the time the principal decides on her bonus payment, however, she does not have any additional information concerning *future* productivity, in contrast to our setting. Halac (2012) analyzes the case of a principal who privately knows the value of her outside option while not being able to observe the agent's effort level directly. In Halac (2012), there is no direct productive distortion in the agent's not knowing the principal's private information; in our setting, by contrast, the first-best level of effort depends on its productivity. In Li and Matouschek (2013), the principal has one-sided private information as well. In contrast to our setting, this information pertains to the cost of *transferring* surplus to the agent, rather than producing surplus. Furthermore, the private information pertains to the *current* period; information about the future is symmetrically held. This allows Li and Matouschek (2013) to apply recursive techniques. In contrast to the implicit downsizing costs in our setting, they find that every optimal equilibrium has the property of being *sequentially* optimal as well.

The rest of the paper is set up as follows: Section 2 introduces the model;

[5] See Malcomson (2012) for an overview of the literature on relational contracts.

Section 3 reviews some benchmarks; Section 4 presents the main results, while Section 5 discusses an extension to persistent shocks. Section 6 concludes. Proofs not given in the text can be found in the Appendix.

2 The Model

Players. There is one principal ("she") and one agent ("he"), who are both risk neutral and who interact repeatedly in periods $t = 1, 2, \cdots$.

Actions. At the beginning of every period t, the principal makes an employment offer to the agent, consisting of a fixed wage $w_t \in [-\bar{w}, \bar{w}]$, where $\bar{w} > 0$ is assumed to be large enough. The agent then accepts ($d_t = 1$) or rejects ($d_t = 0$) the employment offer. If he accepts, the wage w_t is immediately paid. (If $w_t < 0$, the agent pays $-w_t$ to the principal.) He subsequently chooses his effort level $n_t \in \mathbb{R}_+$. At the end of the period, the principal can pay the agent a non-contractible, non-negative, bonus $b_t \in [0, \bar{b}]$, where $\bar{b} > 0$ is assumed to be large enough. Furthermore, she can send a non-verifiable cheap-talk message $\hat{\theta}_t \in \{\theta^l, \theta^h\}$ at this time.

Information. The public events (i.e. those that can be observed by both the principal and the agent) in period t are given by $h_t = \left\{w_t,\, d_t,\, y_t,\, b_t,\, \hat{\theta}_t\right\}$, where $y_t = g(n_t)$. The production function $g : \mathbb{R}_+ \to \mathbb{R}_+$ is C^2, satisfies $g' > 0 > g''$ and $\lim_{n \downarrow 0} g'(n) = \infty$, $\lim_{n \to \infty} g'(n) = 0$. It is commonly known by the players. A *public history* of length $t-1$, h^{t-1} (for $t \geq 2$) collects the public events up to, and including, time $t-1$, i.e. $h^{t-1} := (h_\tau)_{\tau=1}^{t-1}$. We denote the set of public histories of length $t-1$ by $\mathcal{H}^{t-1}$. (We set $\mathcal{H}^0 = \{\emptyset\}$.) In each period, a strategy for the agent specifies what wage offers to accept as a function of the previous public history, and what level of effort to exert if he accepts employment as a function of the previous public history and current-period wages. Formally, it is a sequence of mappings $\left\{\sigma_t^A\right\}_{t=1}^{\infty}$, where, for each $t \in \mathbb{N}$, $\sigma_t^A = (d_t, n_t)$, and $d_t : \mathcal{H}^{t-1} \times [-\bar{w}, \bar{w}] \to \{0, 1\}$, $(h^{t-1}, w_t) \mapsto d_t(h^{t-1}, w_t)$ and $n_t : \mathcal{H}^{t-1} \times [-\bar{w}, \bar{w}] \times \{0, 1\} \to \mathbb{R}_+$, $(h^{t-1}, w_t, d_t) \mapsto n_t(h^{t-1}, w_t, d_t)$.

The principal additionally knows her *type* in period $t+1$, $\theta_{t+1} \in \{\theta^l, \theta^h\}$, before deciding on the bonus payment b_t in period t; the agent never learns the realizations of the principal's types. The values satisfy $\theta^h > \theta^l > 0$ and are commonly known. We write $\theta^t = \{\theta_\tau\}_{\tau=1}^{t}$ for the sequence of realizations of the principal's types up to, and including, period t. The *principal events* in period t are given by $\mathfrak{h}_t = \left\{w_t,\ d_t,\ y_t,\ \theta_{t+1}, b_t,\ \hat{\theta}_t\right\}$; that is, the principal learns about her period-$t+1$ type already in period t, before paying the bonus in the respective period. A *principal history* of length $t-1$, $\mathfrak{h}^{t-1}$ (for $t \geq 2$) collects the principal events up to, and including, time $t-1$, i.e. $\mathfrak{h}^{t-1} := (\mathfrak{h}_\tau)_{\tau=1}^{t-1}$. We denote the set of principal-histories of length $t-1$ by $\mathfrak{H}^{t-1}$. We assume that the principal's type in period $t=1$ is commonly known to be $\theta_1 = \theta^h$ and thus set $\mathfrak{H}^0 = \{\theta^h\}$. In each period, a pure strategy for the principal specifies his wage offers as a function of the previous principal history, as well as his bonus payment and report as a function of the previous history, current-period wages and output, as well as his type in the next period. Formally, it is a sequence of mappings $\left\{\sigma_t^P\right\}_{t=1}^{\infty}$, where, for each $t \in \mathbb{N}$, $\sigma_t^P = (w_t, b_t, \hat{\theta}_t)$, and $w_t : \mathfrak{H}^{t-1} \to [-\bar{w}, \bar{w}]$, $\mathfrak{h}^{t-1} \mapsto w_t(\mathfrak{h}^{t-1})$, $b_t : \mathfrak{H}^{t-1} \times [-\bar{w}, \bar{w}] \times \{0,1\} \times \mathbb{R}_+ \times \{\theta^l, \theta^h\} \to [0, \bar{b}]$, $(\mathfrak{h}^{t-1}, w_t, d_t, y_t, \theta_{t+1}) \mapsto b_t(\mathfrak{h}^{t-1}, w_t, d_t, y_t, \theta_{t+1})$, with the restriction that $d_t = 0 \Rightarrow b_t(\mathfrak{h}^{t-1}, w_t, d_t, y_t, \theta_{t+1}) = 0$, and $\hat{\theta}_t : \mathfrak{H}^{t-1} \times [-\bar{w}, \bar{w}] \times \{0,1\} \times \mathbb{R}_+ \times \{\theta^l, \theta^h\} \to \{\theta^l, \theta^h\}$, $(\mathfrak{h}^{t-1}, w_t, d_t, y_t, \theta_{t+1}) \mapsto \hat{\theta}_t(\mathfrak{h}^{t-1}, w_t, d_t, y_t, \theta_{t+1})$. A pure public strategy by the principal is a pure strategy which does not condition on her *past* private information; formally, a strategy $\left\{\sigma_t^P\right\}_{t=1}^{\infty}$ is said to be a public strategy if, for each period $t \in \mathbb{N}$, it can be written $\sigma_t^P = (\tilde{w}_t, \tilde{b}_t, \tilde{\hat{\theta}}_t)$, where $\tilde{w}_t : \mathcal{H}^{t-1} \to [-\bar{w}, \bar{w}]$, $\tilde{b}_t : \mathcal{H}^{t-1} \times [-\bar{w}, \bar{w}] \times \{0,1\} \times \mathbb{R}_+ \times \{\theta^l, \theta^h\} \to [0, \bar{b}]$ and $\tilde{\hat{\theta}}_t : \mathcal{H}^{t-1} \times [-\bar{w}, \bar{w}] \times \{0,1\} \times \mathbb{R}_+ \times \{\theta^l, \theta^h\} \to \{\theta^l, \theta^h\}$.

While $\theta_1 = \theta^h$, the principal's types $\{\theta_t\}_{t=2}^{\infty}$ are i.i.d. across periods (except in Section 5); for all $t = 2, 3, \cdots$, the probability that $\theta_t = \theta^h$ is $q \in (0,1)$. The probability q, as well as the principal's type in the first period, are common knowledge.

Payoffs. The principal's period payoff in period t is given by $d_t(\theta_t y_t - w_t) - b_t$; the agent's is given by $d_t(w_t - n_t c) + b_t$, where $c > 0$ is his marginal cost of effort. Both players discount future payoffs with the discount factor $\delta \in (0,1)$.

Our solution concept is perfect Bayesian equilibrium in (pure) public strategies (PPE), as defined above. There are no long-term contracts or other means for the principal or the agent to commit to a certain course of action. In particular, the output y_t is assumed to be non-verifiable.

Our objective is to find a PPE that maximizes the principal's ex ante expected profit Π_1 among all PPE. As expected surplus can be transferred freely through w_1, the fixed wage in the first period, any equilibrium maximizing Π_1

also maximizes the players' joint surplus given the constraints, as shown by the following proposition, which parallels Levin's (2003) Theorem 1.

Proposition 1 *Suppose there exists a PPE leading to a joint surplus of $s \geq 0$. Then, there exists a PPE giving the principal an expected payoff of π and the agent an expected payoff of u, for any $(\pi, u) \in \{(x, y) \in \mathbb{R}_+ : x + y = s\}$.*

Proof.

The proof follows that of Theorem 1 in Levin (2003) and is therefore omitted. ■

As on-path equilibrium actions are completely determined by past type realizations, we shall replace histories as defined above with the history of previously reported types, which, on the equilibrium path, coincide with the history of past type realizations. By our choice of equilibrium concept, there is no loss to focussing on truth-telling equilibria; i.e., on the equilibrium path, reported types will coincide with the history of past type realizations, $\theta^t = \{\theta_\tau\}_{\tau=1}^t$. In a slight abuse of notation, we will thus write $w(\theta^t)$ for $w_t(\mathfrak{h}^{t-1})$, and $n(\theta^t)$ for $n_t(h^{t-1}, w(\theta^t), 1)$, the agent's effort choice on the equilibrium path in period t given history θ^t. In addition, we shall use superscripts h or l to indicate the type in period $t+1$, given history θ^t, writing, for instance, $b^h(\theta^t)$ for $b_t(\mathfrak{h}^{t-1}, w(\theta^t), 1, y_t, \theta^h)$, the principal's on-path bonus payment after history θ^t, given that $\theta_{t+1} = \theta^h$. By the same token, we write $\Pi(\theta^t) = \Pi^i(\theta^{t-1})$ for the principal's expected on-path profit at the beginning of period t, given the history of type realizations θ^t and given that $\theta_t = \theta^i$ ($i \in \{h, l\}$).

Thus, we can write

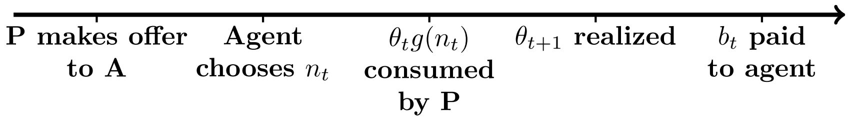
$$\begin{aligned}\Pi(\theta^t) =& d(\theta^t)\left[\theta_t g(n(\theta^t)) - w(\theta^t)\right]\\ &+ q\left(-b^h(\theta^t) + \delta\Pi^h(\theta^t)\right) + (1-q)\left(-b^l(\theta^t) + \delta\Pi^l(\theta^t)\right).\end{aligned}$$

for the principal's expected on-path profits for a given history of types θ^t, and

$$\begin{aligned}U(\theta^t) =& d(\theta^t)\left[w(\theta^t) - n(\theta^t)c\right] + q\left(b^h(\theta^t) + \delta U^h(\theta^t)\right)\\ &+ (1-q)\left(b^l(\theta^t) + \delta U^l(\theta^t)\right).\end{aligned}$$

for the agent's expected on-path utility in period t.

The following figure summarizes the timing within each period:



3 Some Benchmarks

In this section, we analyze a few natural benchmarks against which to measure our equilibrium. This will help us better understand the role of the non-contractibility of effort as well as that of the principal's private information.

3.1 The First Best

Suppose the principal and the agent acted cooperatively so as to maximize their joint surplus. Our assumptions on the production function g immediately imply that, in all periods t, the effort chosen would be equal to $n^{FB}(\theta_t)$, with $n^{FB}(\theta_t)$ being defined by the first-order condition

$$\theta_t g'(n^{FB}(\theta_t)) = c.$$

For the remainder of the paper, we define $n_h^{FB} \equiv n^{FB}(\theta^h)$ and $n_l^{FB} \equiv n^{FB}(\theta^l)$.

3.2 Verifiable Effort

Now, suppose that the agent's effort choice was not just observable but also verifiable, while the principal's type was her private information and both the principal and the agent maximized their own respective payoffs. Since the agent's effort is verifiable, the principal and the agent can write a binding contract specifying, in each period t and given any history θ^t, that $w_t = 0$, as well as $b_t = n_t c$ if $n_t = n^{FB}(\theta_t)$ and $b_t = 0$ otherwise. This sequence of contracts implements first-best effort levels, and, since the principal collects the entire surplus, there is no sequence of contracts generating higher profits.

3.3 Static Subgame-Perfect Equilibrium

If the game is played only once, the principal will never pay a positive bonus, whatever the agent's effort level may have been. Anticipating this, the agent chooses $n_1 = 0$, implying $y_1 = 0$. In any equilibrium of the repeated game, either party's can always guarantee itself this static SPE payoff, which constitutes its minmax-payoff. As we are interested in the best possible PPE for the principal, it is without loss for us to focus on equilibria in which any observable deviation triggers this harshest possible punishment.[6]

3.4 Public Types

Now, let us suppose that the principal's type is public information, while the agent's effort is non-contractible. Thus, we assume that the agent observes next period's type at the same time as the principal does, implying that we here allow the agent to condition his strategy on the principal-histories rather than only the coarser public histories. In this case, there is no informational asymmetry; agency problems arise merely on account of the non-contractibility of effort.

The agent always has the option of rejecting the principal's offers forever, guaranteeing him a utility of 0. Therefore, after any history in any equilibrium, his expected utility will be at least 0, i.e., the following Individual Rationality constraint must hold, for all histories θ^t:

$$U(\theta^t) \geq 0. \tag{IR}$$

Furthermore, after pocketing the fixed wages $w(\theta^t)$, the agent must find it optimal to exert the level of effort he is supposed to exert in equilibrium, namely $n(\theta^t)$. Thus, his utility when exerting $n(\theta^t)$ must be at least as high as his utility from exerting any other level of effort. As effort levels are observable, it is without loss for us to focus on equilibria in which any deviation by the agent is punished in the harshest possible way, by giving him a continuation utility of 0; in such an equilibrium, therefore, any possible deviation is dominated by a deviation to an effort level of 0. Thus, the agent's Incentive

[6] See Abreu (1988) on the optimality of such simple penal codes.

Compatibility Constraint is given by

$$-n(\theta^t)c + q\left(b^h(\theta^t) + \delta U^h(\theta^t)\right) + (1-q)\left(b^l(\theta^t) + \delta U^l(\theta^t)\right) \geq 0. \tag{IC}$$

It must also be optimal, after any history θ^t, for the principal to make the bonus payments she is supposed to make in equilibrium. Indeed, as effort levels and bonus payments are not contractible, these must be self-enforcing. Again, we can focus without loss of generality on equilibria in which the principal is punished with a continuation profit of 0 whenever she does not pay out the bonus she is supposed to pay out; her best deviation in this case is to paying a bonus of 0. This yields the following dynamic enforcement constraints

$$-b^h(\theta^t) + \delta\Pi^h(\theta^t) \geq 0 \tag{DEh}$$

$$-b^l(\theta^t) + \delta\Pi^l(\theta^t) \geq 0. \tag{DEl}$$

It is standard to verify that (DEh) and (DEl) can equivalently be combined into a single constraint,

$$-\left(qb^h(\theta^t) + (1-q)b^l(\theta^t)\right) + \delta\left(q\Pi^h(\theta^t) + (1-q)\Pi^l(\theta^t)\right) \geq 0. \tag{DE}$$

The (DE) constraint states that the future benefits of honoring the relational contract must be sufficiently large for the principal that she is willing to bear today's costs. Whereas these costs manifest themselves in (expected) bonus payments, the benefits are provided by the discounted difference between on- and off-path future profits. Since off-path profits, i.e., profits after a deviation, are zero, the benefits are identical to expected future profits. These determine the principal's commitment in the relationship and consequently her credibility in the relational contract. This *credibility problem,* which asks "should one party believe another's promise?" (Gibbons and Henderson (2012)), is at the heart of most relational contracting models.

Finally, it must be optimal for the principal to offer the equilibrium contract to the agent, i.e., $\Pi(\theta^t) \geq 0$. This, however, is already implied by the (DE) constraint and our assumption that bonus payments are positive.

Thus, our problem is to maximize $\Pi^h(\emptyset)$, subject to (IR), (IC), and (DE), through our choice of effort levels $n(\theta^t)$, wage and bonus payments $w(\theta^t)$, $b^l(\theta^t)$ and $b^h(\theta^t)$, for all histories θ^t. The following lemma details some characteristics of an optimal solution.

Lemma 1 Assume that the firm's type is publicly observable. Then, there exists a profit-maximizing equilibrium in which the agent never gets a rent, that is,

- $qb^h(\theta^t) + (1-q)b^l(\theta^t) = n(\theta^t)c$ and
- $w(\theta^t) = 0$ for every history θ^t.

Furthermore, equilibrium effort only depends on the current state, that is, $n(\theta^t) = n(\theta_t)$.

The lemma shows that there exists an optimal equilibrium in which the agent does not get a rent after any history. Indeed, suppose there was a history θ^t at which the agent optimally received a rent. In this case, it is possible to lower the wages $w(\theta^t)$ and to increase the corresponding bonus payment in the previous period in such a way that profits and constraints remain unaffected. This operation can be repeated until we reach the first period, when the principal can expropriate the agent's entire rent.

Furthermore, the (IC) constraint will bind after any history. Indeed, suppose to the contrary that it was slack after some history θ^t. Then, it is possible to reduce $b^h(\theta^t)$ or $b^l(\theta^t)$ while increasing $w(\theta^t)$ by the corresponding amount. This leaves the previous constraints unaffected but relaxes the (DE) constraint in the current period. As (IC) constraints bind and the agent does not get a rent after any history, we have $w(\theta^t) = 0$ after all histories θ^t.

The lemma also shows that the equilibrium is stationary. Hence, we can write $n(\theta^h)$ and $n(\theta^l)$ for the respective equilibrium effort levels in any period t. The reason for this is that, in the case of observable types, every deviation is observable; there is therefore no reason to burn any surplus on the equilibrium path of play. Players therefore endeavor to be as close as possible to the first-best effort level in any period. What is possible, in turn, does not depend on the history by virtue of our assumption that types are iid across periods. The first-best target, meanwhile, depends on the history only via the current type.

Note that, as is also the case e.g. in Levin (2003) or MacLeod and Malcomson (1989), *enforceable* effort in any given period does not depend on the current type but only on the principal's credibility in the relational contract, that is, her expected future profits. Indeed, current output has already been realized and thus is sunk once the principal decides on the bonus payment. Optimal effort, on the other hand, depends on today's type. This tension delivers the intuition for the following proposition, which summarizes a profit-maximizing equilibrium with public types. There, note that fixing expected per-period profits, the principal's credibility is solely determined by the discount factor δ. Therefore, if δ is high enough, the first best is achievable and the credibility problem resolved. For intermediate levels of the discount factor, n_h^{FB} is no longer enforceable, while n_l^{FB} still is. In this case, the highest enforceable effort level is chosen in all periods t in which $\theta_t = \theta^h$, while n_l^{FB} is enforced in all periods τ in which $\theta_\tau = \theta^l$. If the discount factor is so low that even n_l^{FB} can no longer be enforced, the highest enforceable effort level is enacted in all periods.

Proposition 2 Assume the firm's type is publicly observable. Then, there are levels of the discount factor, $\overline{\delta}$ and $\underline{\delta}$, with $0 < \underline{\delta} < \overline{\delta} < 1$, such that

- $n(\theta^h) = n_h^{FB}$ and $n(\theta^l) = n_l^{FB}$ for $\delta \geq \overline{\delta}$;
- $n(\theta^l) = n_l^{FB} < n(\theta^h) < n_h^{FB}$ for $\underline{\delta} < \delta < \overline{\delta}$;
- $n(\theta^h) = n(\theta^l) \leq n_l^{FB}$ for $\delta \leq \underline{\delta}$.

Note that the principal's credibility today depends on next period's type. Thus, she can credibly commit to a higher bonus payment if tomorrow's type is high. If (DE) binds, it is indeed (strictly) optimal to have $b^h(\theta^t) > b^l(\theta^t)$ because of the agent's risk neutrality. In this case, the agent's compensation is not only a function of his own performance, but also depends on random events that are out of his control. This stands in contrast to Holmstrom (1979)'s informativeness principle, which states that only measures of performance that reveal information about the effort level chosen by the agent should be included in the compensation contract.

4 Private Types

In this section, we assume that the principal's type is her private information. Thus, she has to be given incentives not to misrepresent her true type. A straightforward response would be to make the bonus payment independent of next period's type; however, while feasible, such an approach is generally not optimal. In the following, we will explore how asymmetric information on future profits affects the properties of a profit-maximizing relational contract.

First, given our equilibrium concept, it is without loss for us to focus on truth-telling equilibria. Indeed, suppose to the contrary that, in an optimal equilibrium, the principal did not reveal his private information after some history θ^t. As we are looking for an equilibrium in public strategies, this implies that continuation play after (θ^t, θ^h) and (θ^t, θ^l) must be identical. Yet, as only future prospects matter for enforceable actions, this continuation play would also be available if the principal disclosed her private information.

In truth-telling equilibrium, the principal needs sufficient incentives to reveal her type in every period. Specifically, after any history θ^t, it must be optimal for the principal to pay out $b^h(\theta^t)$ (rather than $b^l(\theta^t)$) if tomorrow's state is high, and $b^l(\theta^t)$ (rather than $b^h(\theta^t)$) if tomorrow's state is low; other bonus payments never occur on the path of play and can therefore be deterred by threatening the principal with a continuation profit of 0. Lest punishment be triggered, once the principal has paid out $b^l(\theta^t)$ at the end of period t, she can only induce effort $n^l(\theta^t)$ in period $t+1$.[7]

Because, for any strategy choice by the agent, the principal always has a best response which is a public strategy, we only need to check the principal's incentives to deviate to another public strategy. Furthermore, thanks to discounting, the One-Deviation principle applies in our setting. Therefore, if tomorrow's state is high but the principal pays out the low-type bonus (or reports $\hat{\theta}_{t+1} = \theta^l$, in case they are equal) instead, her continuation payoff in period $t+1$ can be written as

$$\begin{aligned}\tilde{\Pi}^l(\theta^t) =& \theta^h g(n^l(\theta^t)) - w^l(\theta^t) \\ & + q\left(-b^{lh}(\theta^t) + \delta\Pi^{lh}(\theta^t)\right) + (1-q)\left(-b^{ll}(\theta^t) + \delta\Pi^{ll}(\theta^t)\right),\end{aligned}$$

where the second superscript describes the type in period $t+2$.

By the same token, if tomorrow's state is low but the principal pays out

[7] Note that a formal mechanism to transmit messages would not be required, whenever the size of the bonus depends on tomorrow's type, i.e. $b^h(\theta^t) \neq b^l(\theta^t)$. In this case, bonus payments serve as a message and also determine next period's equilibrium effort. In our equilibrium, whenever the principal's report in period $t+1$ does not correspond to the bonus having been paid in period t, punishment is triggered. When $b^h(\theta^t) = b^l(\theta^t)$ while $n^h(\theta^t) \neq n^l(\theta^t)$, a message is needed to tell the agent which level of effort to choose in period $t+1$.

the high-type bonus instead, her continuation payoff in period $t+1$ is

$$\begin{aligned}\tilde{\Pi}^h(\theta^t) =& \theta^l g(n^h(\theta^t)) - w^h(\theta^t) \\ &+ q\left(-b^{hh}(\theta^t) + \delta\Pi^{hh}(\theta^t)\right) + (1-q)\left(-b^{hl}(\theta^t) + \delta\Pi^{hl}(\theta^t)\right).\end{aligned}$$

Therefore, the principal is willing to tell the truth in equilibrium following history θ^t if and only if

$$-b^h(\theta^t) + \delta\Pi^h(\theta^t) \geq -b^l(\theta^t) + \delta\tilde{\Pi}^l(\theta^t) \quad \text{(TTh)}$$

$$-b^l(\theta^t) + \delta\Pi^l(\theta^t) \geq -b^h(\theta^t) + \delta\tilde{\Pi}^h(\theta^t). \quad \text{(TTl)}$$

As $\tilde{\Pi}^l(\theta^t) = \Pi^l(\theta^t) + \theta^h g(n^l(\theta^t)) - \theta^l g(n^l(\theta^t))$ and $\tilde{\Pi}^h(\theta^t) = \Pi^h(\theta^t) - \theta^h g(n^h(\theta^t)) + \theta^l g(n^h(\theta^t))$, we can rewrite these constraints as follows:

$$-b^h(\theta^t) + \delta\Pi^h(\theta^t) \geq -b^l(\theta^t) + \delta\Pi^l(\theta^t) + \delta g(n^l(\theta^t))\left(\theta^h - \theta^l\right) \quad \text{(TTh)}$$

$$-b^l(\theta^t) + \delta\Pi^l(\theta^t) \geq -b^h(\theta^t) + \delta\Pi^h(\theta^t) - \delta g(n^h(\theta^t))\left(\theta^h - \theta^l\right). \quad \text{(TTl)}$$

Thus, the principal's objective is to maximize $\Pi(\theta^1) = \theta^h g(n(\theta^1)) - w(\theta^1) + q\left(-b^h(\theta^1) + \delta\Pi^h(\theta^1)\right) + (1-q)\left(-b^l(\theta^1) + \delta\Pi^l(\theta^1)\right)$, where $\theta^1 = \theta_1 = \theta^h$, subject to (DEh), (DEl), (TTh), (TTl), (IR) and (IC) at each history θ^t. In the following subsection, we shall give some preliminary results concerning the structure of an optimal equilibrium, before we turn to the presentation of our main results.

4.1 Preliminaries

The object of this subsection is to simplify the problem by eliminating some of the constraints while deriving some structural properties of an optimal equilibrium. We begin with the simple observation that (DEh) can be omitted.

Lemma 2 For any history θ^t, the (DEh) constraint can be omitted.

Proof. Adding (DEl) and (TTh) gives $-b^h(\theta^t) + \delta\Pi^h(\theta^t) \geq \delta g(n^l(\theta^t))\left(\theta^h - \theta^l\right)$. Since the right hand side is positive, this implies (DEh). ■

The following lemma summarizes some structural properties of an optimal equilibrium.

Lemma 3 There exists an optimal equilibrium with the properties that, for every history θ^t,

- $U(\theta^t) = 0$,
- $\Pi^h(\theta^t) \geq \tilde{\Pi}^l(\theta^t)$,
- $b^h(\theta^t) \geq b^l(\theta^t)$,
- the (TTl) constraint can be omitted,
- $n(\theta^t)c = qb^h(\theta^t) + (1-q)b^l(\theta^t)$ and $w(\theta^t) = 0$.

Thus, as in the case of public types, the agent does not get a rent after any history. The intuition for this result remains the same: Front-loading rent payments can only relax the (DE)-constraint, and once the first period is reached, the agent's rent can be expropriated via the fixed wages $w(\theta^1)$. Moreover, the principal's bonus payment is weakly higher if the agent's productivity tomorrow, and hence the value of continuing the relationship for the principal, is higher. Consequently, she will never want to claim that the agent's productivity tomorrow is higher than it actually is; i.e., the (TTl) constraint can be omitted. Thus, on the principal's side, we are left with only the (DEl) and (TTh) constraints. A rather similar argument to the case of public types then establishes that the (IC)-constraint will bind and that $w(\theta^t) = 0$ after all histories θ^t.

The following lemma shows that the (DEl) and (TTh) constraints can be combined into one.

Lemma 4 Maximum profits in the problem in which (TTh) and (DEl) are replaced by the following constraint (EC) equal maximum equilibrium profits:

$$-n(\theta^t)c + \delta\left(q\Pi^h(\theta^t) + (1-q)\Pi^l(\theta^t)\right) \geq \delta q g(n^l(\theta^t))\left(\theta^h - \theta^l\right). \qquad \text{(EC)}$$

Optimal bonus payments are given by $b^h(\theta^t) = b^l(\theta^t) = n(\theta^t)c$ if $n(\theta^t)c \leq \delta\Pi^l(\theta^t)$, and $b^h(\theta^t) = \frac{1}{q}\left[n(\theta^t)c - \delta(1-q)\Pi^l(\theta^t)\right] > \delta\Pi^l(\theta^t) = b^l(\theta^t)$ otherwise.

The left-hand side of the (EC)-constraint is identical to the left-hand side of the (DE)-constraint with public types. It weighs the cost of compensating the agent for his effort costs, $n(\theta^t)c$, against discounted expected future profits, $\delta\left(q\Pi^h(\theta^t)+(1-q)\Pi^l(\theta^t)\right)$, which determine the principal's stakes in the relationship. With public types, this left-hand side had to exceed 0 for the principal to be willing to incur the cost of compensating the agent for his effort costs. With private types, by contrast, this has to be weakly greater than $\delta q g(n^l(\theta^t))\left(\theta^h-\theta^l\right)\geq 0$, which is an expression for the principal's information rent. Indeed, if (DE) constraints bind, the principal would like to transfer her entire future profits to the agent. But this is not feasible if the principal's type tomorrow is θ^h (which happens with probability q), because she always has the option of falsely claiming that the type is θ^l. If she does so, she will get $\theta^h g(n^l(\theta^t))$ in the next period, rather than just $\theta^l g(n^l(\theta^t))$, which determines the bonus the principal is supposed to pay. As (EC) shows, it is on account of this information rent that a given level of effort is harder to implement with private types.

Put differently, the agency problem here consists not only in the non-verifiability of the agent's performance measures, but also in the necessity of preventing the principal from claiming her type to be lower than it actually is in order to save on her bonus payments. Lying generally does not come for free, though, because only the respective low-type effort can be implemented in the subsequent period. Thus, for the same reason as in the case of public types, it can still be optimal to have $b^h(\theta^t)>b^l(\theta^t)$, despite the principal's temptation to lie. Now, the principal's tradeoff boils down to a comparison of today's benefits of a deviation (a lower bonus payment) with tomorrow's costs (a lower output). This aspects adds another dimension to the credibility problem typical for relational contracts. Indeed, the principal's credibility is reduced by the information rent she can always secure herself because of her private information. As we shall see below, tweaking tomorrow's costs of lying, by adjusting the output level given tomorrow's type is low, can be a way of boosting the principal's credibility today.

While effort dynamics in the case of public types are completely stationary (see Lemma 1), this is no longer the case with private types, as the following lemma shows. In order to state this lemma, we define, for every history $\theta^t := (\theta^h, \theta_2, \theta_3, \cdots, \theta_t)$, the function

$$i(\theta^t) := \begin{cases} 0 & \text{if } \theta_t = \theta^h \\ \max\left\{\iota \in \mathbb{N} : \theta_{t-\iota+n} = \theta^l \,\forall\, n \in \{0, \cdots, \iota\}\right\} + 1 & \text{if } \theta_t = \theta^l \end{cases},$$

which indicates the number of consecutive low periods immediately preceding, and including, period t along a given history θ^t.

Lemma 5 There exists an optimal equilibrium with the property that, for every two histories θ^t and $\tilde{\theta}^{\tilde{t}}$, $n^h(\theta^t) = n^h(\tilde{\theta}^{\tilde{t}})$. Furthermore, for every history θ^t, $n^l(\theta^t) = n^l_{i(\theta^t)}$.

Proof. Consider an optimum satisfying the properties of Lemmas 3 and 4. Suppose that there exists a history θ^t such that $\Pi^h(\theta^t) < \max_{\hat{\theta}^\tau} \Pi^h(\hat{\theta}^\tau)$. Replace the continuation play following $\left(\theta^t, \theta^h\right)$ by the continuation play following $\left(\tilde{\theta}, \theta^h\right)$, where $\tilde{\theta} \in \operatorname{argmax}_{\hat{\theta}^\tau} \Pi^h(\hat{\theta}^\tau)$. By virtue of our iid assumption, this is feasible. This increases profits and relaxes some (EC) constraints without tightening any previous ones. This establishes that $\Pi^h(\theta^t) = \overline{\Pi}^h$ for all θ^t (if two different continuation plays lead to $\operatorname{argmax}_{\hat{\theta}^\tau} \Pi^h(\hat{\theta}^\tau)$, we select one to be played after all histories $\left(\theta^t, \theta^h\right)$). Therefore, there exists an optimum in which for any history θ^t, $n^h(\theta^t) = \overline{n}^h$ and $n^l(\theta^t) = n^l_{i(\theta^t)}$. ■

The lemma shows the optimal effort level continues to be the same in all high periods. The reason is that there is no trade-off with respect to effort levels in high periods. Choosing them closer to the first-best benchmark both increases the objective and relaxes the constraint; indeed, making a high period more attractive makes the principal less inclined falsely to claim to be in a low period. The effort level in a low period, by contrast, depends on the history, albeit only via the distance of the current period to the last previous high period. Indeed, there is a trade-off with respect to the effort level in a low period. Making a low period less attractive lowers the objective but relaxes the constraint as it makes it less enticing for the principal falsely to claim to be in a low period; by making it less attractive to be in a low period tomorrow, one can thus enhance the principal's credibility today. Thus, the optimal effort level in a given low period depends on the optimal effort level today, leading to the dynamics described in the lemma. In the following, we shall write $n^h := n(\theta^t)$ for all θ^t such that $\theta_t = \theta^h$; we shall write $n^l_i = n^l_{i(\theta^t)} = n(\theta^{t+1})$ for all $\theta^{t+1} = (\theta^t, \theta^l)$. By the same token, we shall write Π^h and Π^l_i for the corresponding optimal profits.

Armed with these results, we can rewrite our problem. The objective is to choose $\left(n^h, n_i^l\right)_{i\in\mathbb{N}}$ so as to maximize

$$\Pi^h = \frac{1-\delta(1-q)}{1-\delta}\left(\theta^h g(n^h) - n^h c\right) + \frac{1-\delta(1-q)}{1-\delta}\delta(1-q)\sum_{i=0}^{\infty}\left(\delta(1-q)\right)^i\left(\theta^l g(n_i^l) - n_i^l c\right),$$

subject to

$$-n^h c + \delta\left(q\Pi^h + (1-q)\Pi_0^l\right) \geq \delta q g(n_0^l)\left(\theta^h - \theta^l\right). \qquad \text{(ECh)}$$

and

$$-n_i^l c + \delta\left(q\Pi^h + (1-q)\Pi_{i+1}^l\right) \geq \delta q g(n_{i+1}^l)\left(\theta^h - \theta^l\right) \qquad \text{(ECli)}$$

for all $i \in \mathbb{N}$.

The following two lemmas summarize further aspects of an optimal equilibrium: Effort levels are always weakly below first-best levels, and profits are weakly increasing in the discount factor δ.

Lemma 6 There exists an optimal equilibrium with the property that $n_i^l \leq n_l^{FB}$ and $n^h \leq n_h^{FB}$.

Lemma 7 For every history θ^t, maximal profits $\Pi(\theta^t)$ are weakly increasing in δ. Furthermore, a higher δ relaxes (EC) constraints.

4.2 The Optimality of Implicit Downsizing Costs

We are now ready to state our main results. Firstly, if the discount factor is close enough to 1, the first best can be achieved.

Proposition 3 There exists a $\overline{\delta} \in (0,1)$ such that for all $\delta \geq \overline{\delta}$, the unique optimal equilibrium implies first-best effort levels n_h^{FB}/n_l^{FB}.

To get an intuition for the forces at play, recall that the (EC)-constraints in fact capture two distinct effects. On the one hand, there is the classical effect coming from the dynamic-enforcement constraints that the principal would never be willing to make a bonus payment exceeding the discounted expected value of the continuation of the relationship to her. As we have seen above, this constraint can only ever bind in our setting if the principal observes the next period to be low (see Lemma 2). On the other hand, there is the need to incentivize the principal to tell the truth. Indeed, a higher enforceable bonus when next period is high may tempt the principal to lie in order to reduce her bonus payments in the current period. A straightforward response to this temptation is a reduction of $b^h(\theta^t)$, accompanied by an appropriate increase of $b^l(\theta^t)$ to leave incentives for the agent unaffected. This, however, is restricted by $\delta\Pi^l$, which is the most the principal would be willing to pay given that tomorrow's type is low. Yet, as δ, and hence $\delta\Pi^l$, increase, it becomes possible to increase b^l without violating (DEl); this in turn reduces the principal's incentives to lie. The proposition now shows that, when δ is close enough to 1, the (EC) constraint will hold, and hence the principal will not have any incentives to lie or to renege on her bonus payment.

Our next proposition characterizes an optimal outcome, given that the discount factor is too low to implement n_h^{FB} but high enough to implement n_l^{FB}.

Proposition 4 There exist discount factors $\underline{\delta}$ and $\overline{\delta}$, with $0 < \underline{\delta} < \overline{\delta} < 1$, such that, in an optimal equilibrium, for $\delta \in (\underline{\delta}, \overline{\delta})$, n^h and n_0^l are inefficiently low: $n_0^l < n_l^{FB} < n^h < n_h^{FB}$, and, for all $i \geq 1$, $n_i^l = n_l^{FB}$.

Note that, for the first-best solution, the (ECh) and (ECli) constraints are identical but for the first term, which is n_h^{FB} and n_l^{FB}, respectively. Thus, as δ decreases, (ECh) starts binding before the (ECli) constraints do. When this happens, n_h^{FB} is no longer implementable and n^h is hence reduced below first-best levels. Yet, as Proposition 4 shows, n_0^l is reduced below n_l^{FB} as well, even though (ECl0) does not bind. This "overshooting" relaxes (ECh) and thus allows for a smaller reduction in n^h than would otherwise be necessary.

Because the principal needs to be dissuaded from claiming that next period's type is low when it is in fact high, low periods need to be rendered less attractive, and, in particular, those low periods that follow periods in which the principal needs a lot of credibility, i.e., high periods. A natural, surplus-neutral, way of achieving this goal would be to force the principal to make a transfer to the agent if he claims next period's type to be low. However, as Proposition 4 shows, this would not be optimal in our setting, as such a transfer would hit an on-path principal who truthfully claimed next period's type to be low in the same way as it would hit a lying principal. Put differently, such a transfer would relax (TTh), but tighten (DEl) to the same extent. Therefore, (EC) constraints, which are combinations of the respective (TTh) and (DEl) constraints, would not be relaxed.

The distortion of effort levels as proposed by Proposition 4, which can be interpreted as *implicit downsizing costs*, hits a lying off-path principal harder than a truthful principal. Indeed, when lying, the principal's profits are decreased by $\theta^h g' - c$ on the margin, while they are only decreased by $\theta^l g' - c$ when she is telling truth. Thus, the game exhibits memory, and the equilibrium is not sequentially optimal, in that $n_l^{FB} \left(> n_0^l\right)$ would be implemented if the game newly started with a low state. This contrasts with the finding in Li and Matouschek (2013), where every optimal equilibrium was sequentially optimal. In our iid model, this distortion in effort levels only lasts a single period, and $n_i^l = n_l^{FB}$ for $i > 0$. This is due to two reasons. First, reducing n_i^l for $i > 0$ would not allow to further increase n^h because the resulting distortions in later periods would hit on-path and off-path principals alike. Second, for discount factors above $\underline{\delta}$, (ECl) constraints do not bind and first-best effort levels are feasible. Thus, implicit downsizing costs indeed optimally arise on the equilibrium path.

5 Persistent Shocks

So far, we have assumed that the principal's types across periods are iid. In this section, we show that implicit downsizing costs may also obtain if shocks are permanent. Specifically, let us assume that, as before, the principal starts out with a high type, but that the type remains high for another period with time-invariant probability q. With probability $1-q$, the type switches to low and remains low forever. In all other aspects, the setup is as before.

As the problem conditional on the type still being high is stationary, it is without loss for us to restrict attention to equilibria in which actions do not depend on calendar time. Therefore, equilibrium high-type effort is constant, whereas low-type effort depends on the distance in time to the (now permanent) switch from high to low. Thus, equilibrium profits can be written

$$\Pi^h = \theta^h g(n^h) - qb^h - (1-q)b_0^l - w^h + \delta q \Pi^h + \delta(1-q)\Pi_0^l$$

$$\Pi_i^l = \theta^l g(n_i^l) - b_{i+1}^l - w_i^l + \delta \Pi_{i+1}^l,$$

where w^h and w_i^l are defined analogously to n^h and n_i^l.

The objective is to maximize Π^h, subject to the following constraints. First, the dynamic enforcement (DE) constraints must be satisfied for b^h and all b_i^l:

$$-b^h + \delta\Pi^h \geq 0 \tag{DEh}$$

$$-b_i^l + \delta\Pi_i^l \geq 0 \forall i \geq 0. \tag{DEli}$$

As long as the principal has not announced a switch to the low state, the following truth-telling constraints must hold:

$$-b^h + \delta\Pi^h \geq -b_0^l + \delta\tilde{\Pi}_0^l \tag{TTh}$$

$$-b_0^l + \delta\Pi_0^l \geq -b^h + \delta\tilde{\Pi}^h, \tag{TTl}$$

where

$$\tilde{\Pi}_i^l = \theta^h g(n_i^l) - b_{i+1}^l - w_i^l + \delta\left[q\tilde{\Pi}_{i+1}^l + (1-q)\Pi_{i+1}^l\right] = \Pi_i^l + \sum_{\tau=0}^{\infty} (\delta q)^\tau g(n_\tau^l)(\theta^h - \theta^l)$$

and

$$\tilde{\Pi}^h = \theta^l g(n^h) - qb^h - (1-q)b_0^l - w^h + \delta\Pi_0^l = \Pi^h - (\theta^h - \theta^l)\frac{g(n^h)}{1-\delta q}.$$

Note that our formulation of $\tilde{\Pi}_i^l$ takes into account that the principal does not renege after having falsely announced a switch to state θ^l in the past. This requires $-b_i^l + \delta\tilde{\Pi}_i^l \geq 0$, which holds given the (DEli) constraints and, as can be shown, given $\Pi_i^l < \tilde{\Pi}_i^l$.

Finally, the (IC) and (IR) constraints are as before.

The proofs of Lemmas 2 and 3 go through essentially unchanged. This implies, inter alia, that $b^h \geq b_0^l$, $n^h c = qb^h + (1-q)b_0^l$, $n_i^l c = b_{i+1}^l$ and that (TTh) and (DEli) are the relevant constraints.

The proof of Lemma 4 goes through essentially unchanged as well. Therefore, we can equivalently replace (TTh) and (DEl0) by the following (ECh) constraint

$$-n^h c + \delta q \Pi^h + \delta (1-q) \Pi_0^l \geq \left(\theta^h - \theta^l\right) \sum_{i=0}^{\infty} (\delta q)^{i+1} g(n_i^l). \tag{ECh}$$

We furthermore need to keep track of

$$-n_i^l c + \delta \Pi_{i+1}^l \geq 0 \tag{DEli}$$

for all $i \geq 0$. The proof of Lemma 6 goes through unchanged as well.

As before, the right hand side of (ECh) expresses the information rent the principal can secure herself by falsely claiming that the state is low. With iid shocks, the principal gets $\theta^h g(n_0^l)$ after a lie whereas the agent believes she gets $\theta^l g(n_0^l)$, the principal's informational advantage extending but to the next period. Here, by contrast, her informational advantage extends to the first (random) period after her lie in which the state indeed switches to low. As the principal maintains her informational advantage from one period to the next with probability q, the expression for the information rent is now $\delta q \left(\theta^h - \theta^l\right) \sum_{i=0}^{\infty} (\delta q)^i g(n_i^l)$, while it was $\delta q \left(\theta^h - \theta^l\right) g(n_0^l)$ before.

To show that, as before, constraints are tightened for lower values of δ, we prove two new lemmas. Our first lemma shows that the agent's effort level is weakly higher before a shock.

Lemma 8 *The effort levels satisfy* $n^h \geq \sup_{i \in \mathbb{N}} n_i^l$.

We can now prove the following:

Lemma 9 *Maximal profits* Π^h *and* Π_i^l *(*$i \in \mathbb{N}$*) are weakly increasing in* δ*. Furthermore, a higher* δ *relaxes the (ECh) and (DEli)-constraints.*

As $\delta \to 1$, the left-hand sides of the (ECh) and (DEli)-constraints for first-best effort levels diverge to infinity, while the right-hand side of (ECh) converges to $\frac{q}{1-q}(\theta^h - \theta^l) g(n_l^{FB}) < \infty$. Thus, if δ is sufficiently large, the first-best effort levels n_h^{FB} and n_l^{FB} can be implemented. As δ leaves this range, it is of interest whether (ECh) or (DEli) constraints start binding first. Let $\delta^l := \frac{n_l^{FB} c}{\theta^l g(n_l^{FB})}$ denote the discount factor at which (DEli)-constraints start binding for first-best effort levels, and δ^h the corresponding discount factor for the (ECh)-constraint. It can be shown that (ECh) binds first if

$q < \frac{\theta^l g(n_l^{FB})\left(n_h^{FB} - n_l^{FB}\right)}{\theta^h n_l^{FB}\left(g(n_h^{FB}) - g(n_l^{FB})\right)}$; i.e. in this case, $\delta^l < \delta^h$. For this case, the following proposition shows that overshooting of the effort reduction may arise with persistent shocks as well.

Proposition 5 *Assume* $q < \frac{\theta^l g(n_l^{FB})\left(n_h^{FB} - n_l^{FB}\right)}{\theta^h n_l^{FB}\left(g(n_h^{FB}) - g(n_l^{FB})\right)}$ *and* $\delta \in [\delta^l, \delta^h)$. *Then,* $n^h < n_h^{FB}$. *Furthermore, for all* $i \in \mathbb{N}$, $n_i^l < n_{i+1}^l < n_l^{FB}$, *with* $\lim_{i \to \infty} n_i^l = n_l^{FB}$.

Whereas we still observe overshooting, the recovery is gradual and never complete. Recall that in the case of iid shocks, having a distortion is optimal one period after the announcement of a low state because the off-path costs (i.e. if the state is in fact high) are larger than the on-path costs (i.e. if the state is indeed low). Because states are iid, though, costs are the same on path and off path in subsequent periods; there is thus no gain to imposing further distortions, as the agent reverts to telling the truth after one lie by the One-deviation principle.

With persistent shocks, however, falsely claiming that the type is low forces the principal to stick to announcing the low state forever thereafter. As, in expectation, the costs imposed by a distortion in effort in *any* future period are higher off path than on path as there always is some chance that the type is still high after T periods (for any T), it is optimal to keep distorting in all future periods, as, on account of the concave production function g, it is better to smooth out distortions. The further in the past the first announcement of the low state lies, though, the more likely it becomes that the state will indeed have switched to low in the meantime; i.e., the difference in off-path vs. on-path costs imposed by the distortion decreases. It is therefore optimal to distort the less the further past the announcement of the switch to the low state one is. As the expected cost difference becomes negligible over time, the distortion eventually vanishes at a rate that is decreasing in the probability of remaining in the high state, $q = \frac{\theta^l g'(n_{i+1}^l) - c}{\theta^l g'(n_i^l) - c}$. As with iid shocks, our optimal self-enforcing contract is thus not sequentially optimal since n_l^{FB} would satisfy all (DEli) constraints.

6 Conclusion

In this paper, we have shown that the phenomenon of implicit downsizing costs can be explained as an optimal commitment device for a principal not opportunistically to misrepresent her private information. In order to prevent downsizing when it is not necessary, an optimal relational contract imposes a cost on the principal whenever she announces bad news. This in turn enhances the principal's credibility and profits.

Appendix

Proof of Lemma 1

We shall first show that there exists an optimal equilibrium such that $U(\theta^t) = 0$ for all histories θ^t. If $U(\theta^1) > 0$, reduce $w(\theta^1)$ by $U(\theta^1)$. For $t > 1$, assume to the contrary that, in an optimal equilibrium, $U^i(\theta^t) > 0$ for some history θ^t and $i \in \{h, l\}$. Now, reduce $w^i(\theta^t)$ by $U^i(\theta^t)$ and increase the respective bonus in the previous period, $b^i(\theta^t)$, by $\delta U^i(\theta^t)$. Since $-b^i(\theta^t) + \delta\Pi^i(\theta^t)$ and $b^i(\theta^t) + \delta U^i(\theta^t)$ remain unchanged, this change leaves the agent's (IC) constraints as well as all of the principal's constraints at history θ^t and all predecessor histories unaffected. Furthermore, the principal's profits at history θ^t as well as in all predecessor histories remain unchanged. We can thus without loss focus on equilibria such that $U(\theta^t) = 0$ for all histories θ^t.

Now, suppose that there exists a history θ^τ after which the (IC) constraint does not bind. Note that a non-binding (IC) constraint implies that either $b^h(\theta^\tau) > 0$ or $b^l(\theta^\tau) > 0$. Thus, there exists an $\varepsilon > 0$ such that, if either $b^h(\theta^\tau)$ is reduced by $\frac{\varepsilon}{q}$ or $b^l(\theta^\tau)$ by $\frac{\varepsilon}{1-q}$, the (IC) constraint is still satisfied. If $w(\theta^\tau)$ is at the same time increased by ε, the (DE) constraint for history θ^τ is relaxed, and all constraints for all other histories θ^t are unaffected by this change. This adjustment potentially increases profits if (DE) for history θ^τ binds, and leaves profits unaffected if (DE) for history θ^τ is slack, hence is optimal. Thus, we have shown that there exists an optimal equilibrium with the property that $w(\theta^t) = 0$, $U(\theta^t) = 0$, and $qb^h(\theta^t) + (1-q)b^l(\theta^t) = n(\theta^t)c$ for all histories θ^t.

To prove the final part of the Lemma, we first rewrite the (DE) constraint:

$$-n(\theta^t)c + \delta\left(q\Pi^h(\theta^t) + (1-q)\Pi^l(\theta^t)\right) \geq 0. \tag{DE}$$

In addition, note that effort levels will never exceed the first best (otherwise, a reduction would increase profits without violating any of the constraints). Now, assume that there are histories $\theta^{\tilde{\tau}}$ and $\theta^{\overline{\tau}}$, with $n^h(\theta^{\tilde{\tau}}) > n^h(\theta^{\overline{\tau}})$. If the profits being produced in the continuation play following $\left(\theta^{\overline{\tau}}, \theta^h\right)$ are higher, it is possible to implement $n^h(\theta^{\tilde{\tau}})$ with the continuation play following $\left(\theta^{\overline{\tau}}, \theta^h\right)$. In this case, the principal can therefore increase her profits following history $\left(\theta^{\overline{\tau}}, \theta^h\right)$ by increasing the current period's effort level to $n^h(\theta^{\tilde{\tau}})$, while leaving the continuation play unchanged. Now, suppose that it is not possible to implement $n^h(\theta^{\tilde{\tau}})$ with the continuation play following $\left(\theta^{\overline{\tau}}, \theta^h\right)$. This implies that the profits created by the continuation play following $\left(\theta^{\overline{\tau}}, \theta^h\right)$ are lower than the continuation play following $\left(\theta^{\tilde{\tau}}, \theta^h\right)$. Furthermore, because $n^h(\theta^{\tilde{\tau}})$ is enforceable, it is possible to replace the continuation play following $\left(\theta^{\overline{\tau}}, \theta^h\right)$ with the continuation play following $\left(\theta^{\tilde{\tau}}, \theta^h\right)$, thereby relaxing the (DE) constraint in $\overline{\tau}$. It thus becomes possible to increase $n^h(\theta^{\overline{\tau}})$ to $n^h(\theta^{\tilde{\tau}})$. This increases both the principal's current and future profits. A similar argument applies to the low state. Hence, equilibrium effort only depends on the current state. ■

Proof of Proposition 2

To ease the notational burden, we write $n^h \equiv n(\theta^h)$ and $n^l \equiv n(\theta^l)$. The Lagrangian for the firm's problem can be written as

$$\begin{aligned}\mathcal{L} = &\left(\theta^h g(n^h) - n^h c\right)\left(1 + \frac{\delta q}{1-\delta}\right) + \left(\theta^l g(n^l) - n^l c\right)\frac{\delta(1-q)}{1-\delta} \\ &+ \lambda_{DE_h}\left[-n^h c + \frac{\delta}{1-\delta}\left[q(\theta^h g(n^h) - n^h c) + (1-q)(\theta^l g(n^l) - n^l c)\right]\right] \\ &+ \lambda_{DE_l}\left[-n^l c + \frac{\delta}{1-\delta}\left[q(\theta^h g(n^h) - n^h c) + (1-q)(\theta^l g(n^l) - n^l c)\right]\right],\end{aligned}$$

where λ_{DE_i} denotes the Lagrange multiplier associated with the constraint (DE_i), for $i \in \{l, h\}$.

By strict concavity of g, the first-order conditions are both necessary and sufficient for an optimum. By the Inada Conditions on g, optimal effort levels are interior, and hence characterized by $\frac{\partial \mathcal{L}}{\partial n^i} = 0$, as well as $\lambda_{DE_i}\frac{\partial \mathcal{L}}{\partial \lambda_{DE_i}} = 0$, for both $i \in \{h, l\}$. One computes

$$\frac{\partial \mathcal{L}}{\partial n^h} = \left(\theta^h g'(n^h) - c\right)\left[1 + \frac{\delta}{1-\delta} q(1 + \lambda_{DE_h} + \lambda_{DE_l})\right] - \lambda_{DE_h} c;$$

$$\frac{\partial \mathcal{L}}{\partial n^l} = \left(\theta^l g'(n^l) - c\right) \frac{\delta}{1-\delta}(1-q)(1+\lambda_{DE_h}+\lambda_{DE_l}) - \lambda_{DE_l} c.$$

As $n^h \geq n^l$ at an optimum, we know that $\lambda_{DE_h} = 0$ implies $\lambda_{DE_l} = 0$. As our system of equations characterizing the solution $\left(n^h, n^l, \lambda_{DE_h}, \lambda_{DE_l}\right)$ is (jointly) continuous in $\left(n^h, n^l, \lambda_{DE_h}, \lambda_{DE_l}, \delta\right)$, the solutions $\left(n^h, n^l, \lambda_{DE_h}, \lambda_{DE_l}\right)$ can be written as continuous functions of δ. Thus, profits Π^h and Π^l are continuous in δ.

The left-hand sides of the (DE_i) constraints are increasing in δ,[8] hence maximum enforceable effort increases in δ as well.

For $\delta \to 1$, (DE_i) are satisfied for first-best effort levels, since $\theta g(n^{FB}(\theta)) - n^{FB}(\theta)c > 0$ for both $\theta \in \{\theta^h, \theta^l\}$. Thus, there exists a $\bar{\delta} \in [0,1)$ such that $\lambda_{DE_h} = \lambda_{DE_l} = 0$ for all $\delta > \bar{\delta}$. For $\delta = 0$, no positive effort can be enforced. Thus, $\bar{\delta} > 0$. Moreover, by continuity of the (DE_i)-constraints in δ, for every pair of effort levels (n^h, n^l) between zero and the respective first-best effort levels n_l^{FB} and n_h^{FB}, there exists a discount factor $\delta(n^h, n^l)$ such that the constraint (DE_h) holds for $\delta \geq \delta(n^h, n^l)$ and is violated for $\delta < \delta(n^h, n^l)$. Set $\bar{\delta} = \delta(n_h^{FB}, n_l^{FB})$. Since $n_l^{FB} < n_h^{FB}$, (DE_l) holds with slackness at $n^l = n_l^{FB}$ for $\delta = \bar{\delta}$. Let $\underline{n}^h(\delta)$ be defined by $\theta^h g'(\underline{n}^h(\delta)) = c\frac{1-\delta(1-q)}{\delta q}$; as g' is continuous, strictly decreasing and takes on all values in $(0, \infty)$, $\underline{n}^h(\delta)$ exists and is unique; furthermore, the Inverse Function Theorem implies that it is a continuous function of δ. As the partial derivative of (DE_h) with respect to n^h is always strictly negative at $n^h = n_h^{FB}$, we have that $\underline{n}^h(\delta) < n_h^{FB}$. Clearly, the solution $\hat{n}^h$ to the optimization problem in which only (DE_h) is imposed entails $\hat{n}^h \in [\underline{n}^h(\delta), n_h^{FB}]$. Direct computation shows the partial derivative of (DE_h) with respect to n^h to be strictly negative on $(\underline{n}^h(\delta), n_h^{FB})$, while its partial derivative with respect to δ is strictly positive and, since $\delta \leq \bar{\delta} < 1$, bounded. Therefore $\hat{n}^h$ is a continuous function of δ, and thus, by continuity of (DE_l) in (n^h, δ), there exists a $\underline{\delta} \in (0, \bar{\delta})$ such that (DE_l) continues to hold with slackness for all $\delta \in (\underline{\delta}, \bar{\delta}]$. This implies $n^l = n_l^{FB} < n^h < n_h^{FB}$. For $\delta \leq \underline{\delta}$, both (DE) constraints bind, and hence $n^h = n^l$. ■

Before we prove Lemma 3, we first prove the following auxiliary lemma.

Lemma 10 For any history θ^t with $n^h(\theta^t) \neq n^l(\theta^t)$, (TTh) and (TTl) are not both binding.

[8] This can be shown formally by an argument analogous to the one underlying the proof of Lemma 7.

Proof of Lemma 10.

Assume there is a history θ^τ where both constraints bind simultaneously even though $n^h(\theta^\tau) \neq n^l(\theta^\tau)$. Then, (TTh) implies $b^h(\theta^\tau) = b^l(\theta^\tau) + \delta\Pi^h(\theta^\tau) - \delta\tilde{\Pi}^l(\theta^\tau)$. Plugging this into the binding (TTl) constraint yields $g(n^l(\theta^\tau))\left(\theta^h - \theta^l\right) = g(n^h(\theta^\tau))\left(\theta^h - \theta^l\right)$. Since $\theta^h - \theta^l > 0$ and g is strictly increasing, this contradicts the claim that both constraints bind for $n^h(\theta^\tau) \neq n^l(\theta^\tau)$. ■

Proof of Lemma 3

We start with proving the first two parts. Suppose to the contrary that there exists a history θ^t of length $t \geq 1$ and an equilibrium such that, following history θ^t, the principal is strictly better off in this equilibrium than in any equilibrium satisfying points 1.-2. We show by construction that this cannot be the case.

1. Assume that, in an optimal equilibrium, $U^i(\theta^t) > 0,\ i \in \{h, l\}$ for some history θ^t of length t. Reduce $w^i(\theta^t)$ by $U^i(\theta^t)$ and increase the respective bonus in the previous period, $b^i(\theta^t)$, by $\delta U^i(\theta^t)$. Since $-b^i(\theta^t) + \delta\Pi^i(\theta^t)$ and $b^i(\theta^t) + \delta U^i(\theta^t)$ remain unchanged, this change leaves the agent's (IC) constraints as well all of the principal's constraints at history θ^t and all predecessor histories unaffected. Furthermore, the principal's profits at history θ^t as well as in all predecessor histories remain unchanged.

   Repeat this step for all histories of length t and of length $t + 1$.

2. Assume that $\Pi^h(\theta^t) < \tilde{\Pi}^l(\theta^t)$. Replace play after (θ^t, θ^h) by play after (θ^t, θ^l). This leads to on-path profits of $\hat{\Pi}^h(\theta^t) = \tilde{\Pi}^l(\theta^t)$. Set $b^h_{new}(\theta^t) = b^l_{new}(\theta^t) = n(\theta^t)c$, while increasing $w(\theta^t)$ by $\delta q\left(\hat{\Pi}^h(\theta^t) - \Pi^h(\theta^t)\right) + q\left(b^h_{old}(\theta^t) - b^h_{new}(\theta^t)\right) + (1-q)\left(b^l_{old}(\theta^t) - b^l_{new}(\theta^t)\right)$. (By Step 1. and the fact that (IC) at history θ^t holds, this increase is weakly larger than $q\delta\left(\hat{\Pi}^h(\theta^t) - \Pi^h(\theta^t)\right)$.) (TTh), (TTl) and (IC) at history θ^t now hold with equality. Previous constraints remain unchanged, with the exception of previous (IC)-constraints, which are relaxed. It remains to be shown that the (DEl)-constraint at history θ^t continues to hold. As the proof of Lemma 2 shows, the fact that (DEl) and (TTh) previously held

at history θ^t, together with Step 1, implies

$$-n(\theta^t)c + \delta\left\{q\left[\Pi^h(\theta^t) - (\theta^h - \theta^l)g(n^l(\theta^t))\right] + (1-q)\Pi^l(\theta^t)\right\} \geq 0.$$

As $\Pi^h(\theta^t) < \Pi^l(\theta^t) + (\theta^h - \theta^l)g(n^l(\theta^t)) = \tilde{\Pi}^l(\theta^t)$, this implies $-n(\theta^t)c + \delta\Pi^l(\theta^t) \geq 0$, which was to be shown.

Furthermore, we can show (for later use) that, for histories θ^t such that $n^h(\theta^t) \leq n^l(\theta^t)$, $\Pi^l(\theta^t) \geq \tilde{\Pi}^h(\theta^t)$. To the contrary, assume that $\Pi^l(\theta^t) < \tilde{\Pi}^h(\theta^t)$. Replace play after (θ^t, θ^l) by play after (θ^t, θ^h). This leads to on-path profits of $\hat{\Pi}^l(\theta^t) = \tilde{\Pi}^h(\theta^t)$. Set $b^h_{new}(\theta^t) = b^l_{new}(\theta^t) = n(\theta^t)c$, while increasing $w(\theta^t)$ by $\delta(1-q)\left(\hat{\Pi}^l(\theta^t) - \Pi^l(\theta^t)\right) + q\left(b^h_{old}(\theta^t) - b^h_{new}(\theta^t)\right) + (1-q)\left(b^l_{old}(\theta^t) - b^l_{new}(\theta^t)\right)$. (TTh), (TTl) and (IC) at history θ^t now hold with equality. Previous constraints remain unchanged, with the exception of previous (IC)-constraints, which are relaxed. It remains to be shown that (DEl)-constraint at history θ^t continues to hold. As the proof of Lemma 2 shows, the fact that (DEl) and (TTh) previously held at history θ^t, together with Step 1, implies

$$-n(\theta^t)c + \delta\left\{q\left[\Pi^h(\theta^t) - (\theta^h - \theta^l)g(n^l(\theta^t))\right] + (1-q)\Pi^l(\theta^t)\right\} \geq 0.$$

As $\Pi^l(\theta^t) < \Pi^h(\theta^t) - (\theta^h - \theta^l)g(n^h(\theta^t)) = \tilde{\Pi}^h(\theta^t)$, this implies $-n(\theta^t)c + \delta\Pi^h(\theta^t) \geq \delta(\theta^h - \theta^l)\left(qg(n^l(\theta^t) + (1-q)g(n^h(\theta^t))\right)$. As $n^h(\theta^t) \leq n^l(\theta^t)$, this implies $-n(\theta^t)c + \delta\Pi^h(\theta^t) \geq \delta(\theta^h - \theta^l)g(n^h(\theta^t))$, or $-n(\theta^t)c + \delta\hat{\Pi}^l(\theta^t) \geq 0$, which was to be shown.

After Operation 2., we have to repeat Operations 1. As Operations 1. leave profits and effort levels unchanged, there is no need to repeat Operation 2. after that. Furthermore, we can repeat these operations for all histories of length t and after that for all histories of length $t-1$, $t-2$, $\cdots$. Finally, assume $U(\theta^1) > 0$. Reducing $w(\theta^1)$ by $U(\theta^1)$ increases $\Pi(\theta^1)$ and only affects the agent's first-period (IR) constraint, which continues to hold.

To show that $b^h(\theta^t) \geq b^l(\theta^t)$ for all histories θ^t, assume to the contrary that there exists a history θ^t such that $b^h(\theta^t) < b^l(\theta^t)$. Because of part 2, this implies that (TTh) is slack. Increase $b^h(\theta^t)$ by a small $\varepsilon > 0$ and reduce $b^l(\theta^t)$ by $\frac{q}{1-q}\varepsilon$. This leaves all (IC) constraints unaffected and relaxes the (DEl) and (TTl) constraints at history θ^t. (TTh) is tightened, while nonetheless remaining slack as long as $b^h(\theta^t) < b^l(\theta^t)$. Finally, all constraints and profits at predecessor histories remain unchanged.

We now show that the (TTl) constraint can be omitted. If $n^h(\theta^t) \leq n^l(\theta^t)$, this follows immediately from the fact that $b^h(\theta^t) \geq b^l(\theta^t)$ and $\Pi^l(\theta^t) \geq \tilde{\Pi}^h(\theta^t)$. So suppose that $n^h(\theta^t) > n^l(\theta^t)$, and suppose that the (TTl) constraint binds. By Lemma 10, this implies that the (TTh) constraint is slack. We can therefore increase $b^h(\theta^t)$ by a small $\varepsilon > 0$ while decreasing $w(\theta^t)$ by $q\varepsilon$ (or, if $b^l(\theta^t) > 0$, we can decrease $b^l(\theta^t) > 0$ by $\frac{q}{1-q}\varepsilon$ instead). This leaves all previous constraints and profits unaffected yet relaxes the current (IC) constraint (or leaves the current (IC) constraint unchanged and relaxes the current (DEl) constraint).

Finally, assume to the contrary that there exists a history θ^t such that $-n(\theta^t)c + qb^h(\theta^t) + (1-q)b^l(\theta^t) > 0$. If $b^h(\theta^t) > b^l(\theta^t)$, reduce $b^h(\theta^t)$ by a small $\varepsilon > 0$ and increase $w(\theta^t)$ by $q\varepsilon$. If $b^h(\theta^t) = b^l(\theta^t)$, reduce $b^h(\theta^t)$ and $b^l(\theta^t)$ by a small $\varepsilon > 0$ and increase $w(\theta^t)$ by ε. In the first case, this relaxes the (TTh) constraint at history θ^t; in the second case, it relaxes the (DEl) constraint at history θ^t. Profits, the agent's utility and all other constraints at history θ^t or its predecessor histories are unaffected. Because $U(\theta^t) = w(\theta^t) - n(\theta^t)c + qb^h(\theta^t) + (1-q)b^l(\theta^t) = 0$, a binding (IC) constraint implies that $w(\theta^t) = 0$ for all histories θ^t. ■

Proof of Lemma 4

By Lemma 3, we can without loss focus on equilibria in which

$$n(\theta^t)c = qb^h(\theta^t) + (1-q)b^l(\theta^t) \tag{1}$$

at every history θ^t. Using (1) and multiplying (TTh) with q and adding it to (DEl) yields (EC).

To prove that (EC) implies (TTh) and (DEl) given (1), assume that we are at an optimum satisfying the properties of Lemma 3 and that (EC) holds. We shall now show that it is always possible to find non-negative bonus payments $b^h(\theta^t)$ and $b^l(\theta^t)$ such that (1) holds, and that (DEl) and (TTh) are both satisfied. Toward this purpose, we set $b^l(\theta^t) = \min\left\{\delta\Pi^l(\theta^t), n(\theta^t)c\right\} \geq 0$. First suppose that $n(\theta^t)c \leq \delta\Pi^l(\theta^t)$. In this case, we set $b^h(\theta^t) = n(\theta^t)c$. Now, (DEl) will trivially hold (with slackness if $n(\theta^t)c < \delta\Pi^l(\theta^t)$). Using

$b^h(\theta^t) = n(\theta^t)c$ in (TTh) yields $\delta\Pi^h(\theta^t) \geq \delta g(n^l(\theta^t))\left(\theta^h - \theta^l\right) + \delta\Pi^l(\theta^t)$, which is implied by the second part of Lemma 3. Now suppose that $n(\theta^t)c > \delta\Pi^l(\theta^t)$. In this case, we set $b^h(\theta^t) = \frac{1}{q}\left[n(\theta^t)c - \delta(1-q)\Pi^l(\theta^t)\right] > 0$. Clearly, (DEl) will trivially hold with equality (because $b^l(\theta^t) = \delta\Pi^l(\theta^t)$). Substituting $b^h(\theta^t)$ into (TTh) yields $\frac{1}{q}$ times (EC). ■

Proof of Lemma 6

Consider an optimum satisfying the properties of Lemmas 3, 4 and 5. Suppose there exists a history θ^t such that $n(\theta^t) > n^{FB}(\theta_t)$. Reduce $n(\theta^t)$ by a small $\varepsilon > 0$. This increases profits and relaxes the (EC) constraints at all predecessor histories. ■

Proof of Lemma 7

Consider a given discount factor $\hat{\delta}$ and the associated sequence of optimal actions $\left(n^h(\hat{\delta}), n_i^l(\hat{\delta})\right)_{i\in\mathbb{N}}$. We first show that a higher δ relaxes (EC) constraints; i.e., for any discount factor $\tilde{\delta} > \hat{\delta}$, previously optimal actions $n^h(\hat{\delta})$ and $n_i^l(\hat{\delta})$ continue to satisfy the (EC) constraints. We show this by induction over the number of periods, starting from the first period, in which the discount factor rises from $\hat{\delta}$ to $\tilde{\delta}$. First, suppose only the discount factor between the first and the second period rises. The (EC) constraint in the first period can be written as $-n^h c + \delta q\left[\Pi^h - g(n_0^l)\left(\theta^h - \theta^l\right)\right] + \delta(1-q)\Pi_0^l \geq 0$. In Lemma 3 we showed that, at our optimum, $\Pi^h(\theta^t) \geq \Pi^l(\theta^t) + g(n^l(\theta^t))\left(\theta^h - \theta^l\right)$ for all histories θ^t. Since $\Pi^l(\theta^t) \geq 0$, the term in square brackets is non-negative. Hence, (EC) in period 1 becomes slacker, and the actions that were optimal at the discount factor $\hat{\delta}$ can still be enforced at the higher discount factor $\tilde{\delta}$. By Lemma 6, these actions lead to (weakly) higher profits. The argument for the induction step is analogous. ■

Proof of Proposition 3

The (EC) constraint to enforce first-best effort levels is given by

$$-n^{FB}(\theta_t)c + \delta\left(q\Pi^{h,FB} + (1-q)\Pi_0^{l,FB}\right) - \delta q g(n_l^{FB})\left(\theta^h - \theta^l\right) \geq 0.$$

The left-hand side can be bounded from below by

$$
\begin{aligned}
&- n^{FB}(\theta_t)c + \delta q \Pi^{h,FB} - \delta q g(n_l^{FB}) \left(\theta^h - \theta^l\right) \\
\geq &- n^{FB}(\theta_t)c + \delta q \left(\theta^h g(n_h^{FB}) - n_h^{FB} c\right) \left(\frac{1-\delta\left(1-q\right)}{1-\delta}\right) - \delta q g(n_l^{FB}) \left(\theta^h - \theta^l\right).
\end{aligned}
$$

Since $\theta^h g(n_h^{FB}) - n_h^{FB} c > 0$ by assumption and because $g(n_l^{FB})$ is finite, this expression diverges to infinity as $\delta \to 1$. Since, by Lemma 7, (EC) constraints are relaxed by larger values of δ, the claim follows. ■

Proof of Proposition 4

Define $\overline{\delta} \in (0,1)$ as the smallest discount factor such that (ECh) holds as an equality for first-best effort levels $n^h = n_h^{FB}$ and $n_i^l = n_l^{FB}$, for all $i \in \mathbb{N}$; i.e., $\overline{\delta}$ is the smallest discount factor such that

$$
-n_h^{FB} c + \overline{\delta}\left(q \Pi^{h,FB} + (1-q)\Pi^{l,FB}\right) = \overline{\delta} q g(n_l^{FB}) \left(\theta^h - \theta^l\right).
$$

Note that given first-best effort levels, (ECh) is continuous in δ. Furthermore, $\overline{\delta} > 0$ follows from no effort being enforceable for $\delta = 0$. Because $n_h^{FB} > n_l^{FB}$, all (ECl) constraints are slack at $\overline{\delta}$ for first-best effort levels.

Now, consider the relaxed problem of maximizing Π^h subject only to (ECh). The Lagrange function for this problem is given by

$$
\begin{aligned}
\mathcal{L} = \Pi^h + \lambda_{ECh} \Bigg[&- n^h c + \frac{\delta q}{1-\delta(1-q)} \Pi^h + \delta\left(\left(\theta^l - q\theta^h\right) g(n_0^l) - (1-q) n_0^l c\right) \\
&+ \sum_{\tau=1}^{\infty} \left(\delta(1-q)\right)^{\tau+1} \left(\theta^l g(n_\tau^l) - n_\tau^l c\right) \Bigg]
\end{aligned}
$$

where $\Pi^h = \frac{1-\delta(1-q)}{1-\delta}\left(\theta^h g(n^h) - n^h c\right) + \frac{1-\delta(1-q)}{1-\delta}\delta(1-q)\left[\sum_{i=0}^{\infty}\left(\delta(1-q)\right)^i \left(\theta^l g(n_i^l) - n_i^l c\right)\right]$. By our assumptions on g, the objective function and the constraint are twice continuously differentiable in the choice variables $\left(n^h, n_i^l\right)_{i\in\mathbb{N}}$. If $\theta^l \geq q\theta^h$, the Lagrangian is strictly concave in the choice variables, and the first-order conditions are necessary and sufficient for an optimum. If $\theta^l < q\theta^h$, the first-order conditions are necessary for a global optimum.[9]

[9] In this case, one can show that a global optimum exists and that it entails $n^h \in (0, n_{FB}^h)$ by substituting the binding (ECh) constraint into the objective. Indeed, considering n_0^l as

The first-order conditions for our reduced problem are given by

$$\frac{\partial \mathcal{L}}{\partial n^h} = \left(\theta^h g'(n^h) - c\right)\left(\frac{1-\delta(1-q)}{1-\delta} + \lambda_{ECh}\frac{\delta q}{1-\delta(1-q)}\right) - c\lambda_{ECh} = 0;$$

$$\frac{\partial \mathcal{L}}{\partial n_0^l} = \delta(1-q)\left(\theta^l g'(n_0^l) - c\right)\frac{1-\delta(1-q)}{1-\delta}(1+\lambda_{ECh}) - \lambda_{ECh}\delta q g'(n_0^l)\left(\theta^h - \theta^l\right) = 0 \text{ if } n_0^l > 0;$$

$$\lambda_{ECh}[-n^h c + \frac{\delta q}{1-\delta(1-q)}\Pi^h$$
$$+\delta\left(\left(\theta^l - q\theta^h\right)g(n_0^l) - (1-q)n_0^l c\right) + \sum_{\tau=1}^{\infty}(\delta(1-q))^{\tau+1}\left(\theta^l g(n_\tau^l) - n_\tau^l c\right)] = 0.$$

Furthermore, optimality requires $\frac{\partial L}{\partial n_i^l} = 0$, implying $\theta^l g'(n_i^l) = c$, for all $i \geq 1$.

Thus, once (ECh) binds and hence $\lambda_{ECh} > 0$, $\theta^h g'(n^h) - c$ must be positive for the respective first-order condition to hold; n^h will thus be below its first-best level. In addition, if $n_0^l > 0$, $\theta^l g'(n_0^l) - c$ must be positive for the first-order condition to hold, so that n_0^l will be below its first-best level as well. Effort levels n_i^l are at their efficient level n_l^{FB} for all $i \geq 1$.

Let $\underline{n}^h(\delta)$ be defined by $\theta^h g'(\underline{n}^h(\delta)) = c\frac{1-\delta(1-q)}{\delta q}$. As g' is continuous, strictly decreasing and takes on all values in $(0,\infty)$, $\underline{n}^h(\delta)$ exists and is unique; furthermore, the Inverse Function Theorem implies that it is a continuous function of δ. Moreover, define $\underline{\tilde{n}}^l(\delta)$ by $g'(\underline{\tilde{n}}^l(\delta)) = c(1-q)\frac{1-\delta(1-q)}{1-\delta}\left[\frac{1-\delta(1-q(1-q))}{1-\delta}\theta^l - q\theta^h\right]^{-1}$ and $\underline{n}^l(\delta)$ by

$$\underline{n}^l(\delta) = \begin{cases} \underline{\tilde{n}}^l(\delta) & \text{if } \frac{1-\delta(1-q(1-q))}{1-\delta}\theta^l - q\theta^h > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Again, as g' is continuous, strictly decreasing and takes on all values in $(0,\infty)$, $\underline{n}^l(\delta)$ exists and is unique; furthermore, the Inverse Function Theorem implies that it is a continuous function of δ. Clearly, the solution $(\hat{n}^h, \hat{n}_0^l)(\delta)$ to the

a function of n^h, one shows that this objective function is strictly concave in n^h, strictly increasing for n^h close to 0, and, given that we can impose without loss that $n_0^l \leq n_{FB}^l$ by Lemma 6, decreasing at $n^h = n_{FB}^h$. Of course, as the global optimum satisfies the first-order conditions, the properties we derive from them apply to the optimum in this case as well.

problem in which only (ECh) is imposed entails $(\hat{n}^h, \hat{n}_0^l)(\delta) \in \mathcal{I}$, where $\mathcal{I} := [\underline{n}^h(\delta), n_h^{FB}] \times [\underline{n}^l(\delta), n_l^{FB}]$.[10] Direct computation shows the partial derivatives of (ECh) with respect to n^h and n_0^l respectively to be strictly negative a.e. on $\mathcal{I}$, while, because $\delta \leq \bar{\delta} < 1$, its partial derivative with respect to δ is bounded. Hence, it is feasible to have a policy $(\hat{n}^h, \hat{n}_0^l)$ that is continuous in δ, implying that the optimal profits $\hat{\Pi}^h$ in this problem are a continuous function of δ. As (n^h, n_0^l) impacts the (ECli) constraints only via the profits Π^h, and since these constraints are continuous in Π^h, all (ECli) constraints hold for the solutions of this reduced problem in a neighborhood of $\overline{\delta}$.[11] By the argument underlying the proof of Lemma 7, the (ECh) constraint becomes tighter as the discount factor δ decreases. Thus, $\hat{\Pi}^h(\delta)$ is (weakly) increasing. We can thus take $\underline{\delta}$ as low as the discount factor at which the (ECli) constraints, $i \geq 1$, just hold as an equality for $n_i^l = n_l^{FB}$, and $n^h = \hat{n}^h$ and $n_0^l = \hat{n}_0^l$, as characterized by the Kuhn-Tucker system above.

It remains to show that $n^h > n_l^{FB}$. Suppose to the contrary that $n^h \leq n_l^{FB}$. Yet this solution is dominated by $\hat{n}^h = \hat{n}_0^l = n_i^l = n_l^{FB}$, which leads to higher profits and is feasible since all (ECli)-constraints (for $i \geq 1$) hold for $n_i^l = n_l^{FB}$ even for the initial n^h and n_0^l. ■

Proof of Lemma 8

Suppose to the contrary that a policy $\sigma = \left(n^h, n_i^l\right)_{i\in\mathbb{N}}$ such that $n^h < \sup_{i\in\mathbb{N}} n_i^l =: \bar{n}^l$ is optimal. Then, the policy $\hat{\sigma} = \left(\hat{n}^h, \hat{n}_i^l\right)_{i\in\mathbb{N}}$ given by $\hat{n}^h = \hat{n}_i^l = \bar{n}^l$ for all $i \in \mathbb{N}$ leads to higher profits $\hat{\Pi}^h > \Pi^h$ and $\hat{\Pi}^l \geq \Pi_i^l$ $(i \in \mathbb{N})$, where $\hat{\Pi}^h$ (Π^h) and $\hat{\Pi}^l$ (Π_i^l) are the profits associated with policy $\hat{\sigma}$ (σ), respectively. As policy σ satisfies all (DEli)-constraints, we have that $-n_i^l c + \delta\hat{\Pi}^l \geq -n_i^l c + \delta\Pi_i^l \geq 0$. This implies $-\bar{n}^l c + \delta\hat{\Pi}^l \geq 0$, i.e., the policy $\hat{\sigma}$ satisfies all (DEli)-constraints. As, under the policy $\hat{\sigma}$, we have that $\hat{\Pi}^h = \hat{\Pi}^l + \delta q(\theta^h - \theta^l) \sum_{i=0}^{\infty} (\delta q)^i g(\hat{n}_i^l)$, the (ECh)-constraint simplifies to $-\bar{n}^l c + \delta\hat{\Pi}^l \geq 0$, which holds by our previous step. This is a contradiction to policy σ being optimal. ■

[10] One shows that $\underline{n}^l < n_l^{FB}$ $(\underline{n}^h < n_h^{FB})$ by showing that the partial derivative of (ECh) with respect to n_0^l (n^h) is always strictly negative at $n_0^l = n_l^{FB}$ $(n^h = n_h^{FB})$.

[11] As the only exception, there is a direct impact of $\hat{n}_0^l$ in (ECl0). Yet, as $\hat{n}_0^l \leq n_l^{FB}$, (ECl0) is slacker than the other (ECli) constraints, and thus continues to hold as well.

Proof of Lemma 9

Suppose the discount factor rises from $\hat{\delta}$ to $\tilde{\delta} > \hat{\delta}$. The actions that were optimal at $\hat{\delta}$ continue to satisfy all (DEli) for $\tilde{\delta}$. By Lemma 6, these actions lead to weakly higher profits. It thus only remains to show that (ECh) is relaxed as δ increases. For this, we compute the derivative $\mathcal{D}$ of (ECh) with respect to δ, which works out as

$$\mathcal{D} = q \left[\Pi^h + \delta \Pi^{h'} - (\theta^h - \theta^l) \sum_{i=0}^{\infty} (1+i)(\delta q)^i g(n_i^l) \right] + (1-q) \left[\Pi_0^l + \delta \Pi_0^{l'} \right].$$

As

$$\Pi^h = \frac{1}{1-\delta q} \left[\theta^h g(n^h) - n^h c + \delta(1-q)\Pi_0^l \right],$$

we have

$$\Pi^{h'} = \frac{1-q}{1-\delta q} [\Pi_0^l + \delta \Pi_0^{l'}] + \frac{q}{(1-\delta q)^2} \left[\theta^h g(n^h) - n^h c + \delta(1-q)\Pi_0^l \right].$$

Furthermore, as

$$\Pi_0^l = \sum_{i=0}^{\infty} \delta^i \left(\theta^l g(n_i^l) - n_i^l c \right),$$

we have

$$\Pi_0^l + \delta(1-\delta q)\Pi_0^{l'} = \sum_{i=0}^{\infty} (1 + (1-\delta q)i)\delta^i \left(\theta^l g(n_i^l) - n_i^l c \right).$$

Inserting this gives us

$$\begin{aligned} (1-\delta q)^2 \mathcal{D} = q(\theta^h g(n^h) - n^h c) + (1-q) \sum_{i=0}^{\infty} (1 + (1-\delta q)i)\, \delta^i \left(\theta^l g(n_i^l) - n_i^l c \right) \\ - q(\theta^h - \theta^l)(1-\delta q)^2 \sum_{i=0}^{\infty} (1+i)(\delta q)^i g(n_i^l). \end{aligned}$$

To show that $\mathcal{D} \geq 0$, it is sufficient to show that

$$q(\theta^h g(n^h) - n^h c) + (1-q) \sum_{i=0}^{\infty} (1 + (1-\delta q)i)\, \delta^i \left(\theta^l g(n_i^l) - n_i^l c \right) - q(\theta^h - \theta^l) g(\bar{n}^l) \geq 0,$$

where we have used that $\sum_{i=0}^{\infty}(1+i)(\delta q)^i = \frac{1}{(1-\delta q)^2}$ and $\sup_{i\in\mathbb{N}} n_i^l =: \bar{n}^l$. We can rewrite this as

$$q\left[\theta^h(g(n^h)-g(\bar{n}^l)) - \left(n^h - \sum_{i=0}^{\infty}(1+(1-\delta q)i)\delta^i n_i^l\right)c \right.$$
$$\left. +\theta^l\left(g(\bar{n}^l) - \sum_{i=0}^{\infty}(1+(1-\delta q)i)\delta^i g(n_i^l)\right)\right]$$
$$+\sum_{i=0}^{\infty}(1+(1-\delta q)i)\delta^i(\theta^l g(n_i^l) - n_i^l c) \geq 0.$$

By Lemma 8, we know that $n^h \geq \bar{n}^l$; by Lemma 6, this implies that $\theta^h g(n^h) - n^h c \geq \theta^h g(\bar{n}^l) - \bar{n}^l c$. Thus, it is sufficient for $\mathcal{D} \geq 0$ that

$$q\left[\theta^l g(\bar{n}^l) - \bar{n}^l c - \sum_{i=0}^{\infty}(1+(1-\delta q)i)\delta^i(\theta^l g(n_i^l) - n_i^l c)\right] + \sum_{i=0}^{\infty}(1+(1-\delta q)i)\delta^i(\theta^l g(n_i^l) - n_i^l c) \geq 0,$$

which was to be shown. ■

Proof of Proposition 5

We first omit (DEl) constraints and show ex post that they hold at the solutions of the relaxed problem. Denoting by λ the Lagrange parameter associated with the (ECh) constraint, the Lagrange function equals

$$\mathcal{L} = \frac{\theta^h g(n^h) - n^h c + \delta(1-q)\sum_{i=0}^{\infty}\delta^i\left(\theta^l g(n_i^l) - n_i^l c\right)}{1-\delta q}(1+\delta q\lambda)$$
$$+\lambda\left[-n^h c + \sum_{i=0}^{\infty}\delta^{i+1}\left[\left((1-q)\,\theta^l - \left(\theta^h - \theta^l\right)q^{i+1}\right)g(n_i^l) - (1-q)\,n_i^l c\right]\right],$$

yielding first-order conditions

$$\frac{\partial \mathcal{L}}{\partial n^h} = \frac{\theta^h g'(n^h) - c}{1-\delta q}(1+\delta q\lambda) - \lambda c = 0 \tag{2}$$

$$\frac{\partial \mathcal{L}}{\partial n_i^l} = \delta^{i+1}\Bigg\{ \left(\theta^l g'(n_i^l) - c\right)\left(\frac{(1-q)}{1-\delta q}(1+\delta q\lambda) + \lambda(1-q)\right) \\ -\lambda q^{i+1}\left(\theta^h - \theta^l\right) g'(n_i^l)\Bigg\} = 0 \tag{3}$$

$\delta < \delta^h$ implies $\lambda > 0$. Hence, condition (2) gives $n^h < n_h^{FB}$, whereas (3) gives $n_i^l < n_l^{FB}$. Condition (3) also implies that $\lim_{i\to\infty} n_i^l = n_l^{FB}$: Since $q < 1$, $\lim_{i\to\infty} q^{i+1} = 0$, hence $\theta^l g'(n_i^l) - c = 0$.

To show that $n_i^l < n_{i+1}^l$, rewrite conditions (3) for n_i^l and for n_{i+1}^l as
$\left(\theta^l g'(n_i^l) - c\right)\frac{(1-q)}{1-\delta q} = -\lambda\left[\frac{(1-q)}{1-\delta q}\left(\theta^l g'(n_i^l) - c\right) - q^{i+1}\left(\theta^h - \theta^l\right)g'(n_i^l)\right]$
$\left(\theta^l g'(n_{i+1}^l) - c\right)\frac{(1-q)}{1-\delta q} = -\lambda\left[\frac{(1-q)}{1-\delta q}\left(\theta^l g'(n_{i+1}^l) - c\right) - q^{i+2}\left(\theta^h - \theta^l\right)g'(n_i^l)\right]$.
Dividing the first by the second equality yields the necessary condition

$$\frac{\theta^l g'(n_i^l) - c}{\theta^l g'(n_{i+1}^l) - c} = \frac{\frac{(1-q)}{1-\delta q}\left(\theta^l g'(n_i^l) - c\right) - q^{i+1}\left(\theta^h - \theta^l\right)g'(n_i^l)}{\frac{(1-q)}{1-\delta q}\left(\theta^l g'(n_{i+1}^l) - c\right) - q^{i+2}\left(\theta^h - \theta^l\right)g'(n_i^l)},$$

which becomes

$$q^{i+1}\left(\theta^h - \theta^l\right)g'(n_i^l)\frac{\left(\theta^l g'(n_{i+1}^l) - c\right) - q\left(\theta^l g'(n_i^l) - c\right)}{\left(\theta^l g'(n_{i+1}^l) - c\right)\left[\frac{(1-q)}{1-\delta q}\left(\theta^l g'(n_{i+1}^l) - c\right) - q^{i+2}\left(\theta^h - \theta^l\right)g'(n_i^l)\right]} = 0.$$

The denominator of this expression must be different from zero: $\left(\theta^l g'(n_{i+1}^l) - c\right) > 0$ because $n_{i+1}^l < n_l^{FB}$. The term in squared brackets must be strictly negative: It captures the partial derivative of the left hand side of the (ECh) constraint with respect to n_{i+1}^l. If it were positive, a larger value of n_{i+1}^l (which is feasible) would relax the (ECh) constraint, contradicting that it binds. Therefore, the term is zero if and only if its numerator is zero, yielding

$$\frac{\theta^l g'(n_{i+1}^l) - c}{\theta^l g'(n_i^l) - c} = q.$$

Because $q < 1$ and $g(\cdot)$ is strictly concave, $n_{i+1}^l > n_i^l$.

Finally, note that the derived n_i^l satisfy all (DEli) constraints, $-n_i^l c + \delta\Pi_{i+1}^l \geq 0$. Since $n_{i+1}^l > n_i^l \forall i$, $\Pi_{i+1}^l > \frac{\theta^l g(n_i^l) - n_i^l c}{1-\delta}$, hence it is sufficient to show that

$-n_i^l c + \delta\frac{\theta^l g(n_i^l) - n_i^l c}{1-\delta} \geq 0$, that is $-n_i^l c + \delta\theta^l g(n_i^l) \geq 0$, holds. Because $\delta \geq \delta^l$, this condition would hold for $n_i^l = n_l^{FB}$. Because $g(\cdot)$ is strictly increasing and concave, and because $g(0) = 0$, $-n_l^{FB}c + \delta\theta^l g(n_l^{FB}) \geq 0$ implies that this also holds for all $n_i^l < n_l^{FB}$. ■

References

ABREU, D. (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56(6), 1713–1733.

BULL, C. (1987): "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, 102(1), 147–159.

CASCIO, W. F. (1993): "Downsizing: What Do We Know? What Have We Learned?," *Academy of Management Executive*, 7(1), 95–104.

DATTA, D. K., J. P. GUTHRIE, D. BASUIL, AND A. PANDEY (2010): "Causes and Effects of Employee Downsizing: A Review and Synthesis," *Journal of Management*, 36(1), 281–348.

GIBBONS, R., AND R. HENDERSON (2012): *What Do Managers Do?: Exploring Persistent Performance Differences among Seemingly Similar Enterprises*pp. 680–731. Princeton University Press.

GOESAERT, T., M. HEINZ, AND S. VANORMELINGEN (2015): "Downsizing and firm performance: evidence from German firm data," *Industrial and Corporate Change*, 24(6), 1443–1472.

HALAC, M. (2012): "Relational Contracts and the Value of Relationships," *American Economic Review*, 102(2), 750–779.

HOLMSTROM, B. (1979): "Moral Hazard and Observability," *The Bell Journal of Economics*, 10(1), 74–91.

LEVIN, J. (2002): "Multilateral Contracting and the Employment Relationship," *The Quarterly Journal of Economics*, 117(3), 1075–1103.

LEVIN, J. (2003): "Relational incentive contracts," *American Economic Review*, 93(3), 835–857.

LI, J., AND N. MATOUSCHEK (2013): "Managing Conflicts in Relational Contracts," *The American Economic Review*, 103(6), 2328–2351.

LOVE, E. G., AND M. KRAATZ (2009): "Character, Conformity, or the Bottom Line? How and Why Downsizing Affected Corporate Reputation," *Academy of Management Executive*, 52(2), 314–335.

MacLeod, W. B., and J. M. Malcomson (1989): "Implicit contracts, incentive compatibility, and involuntary unemployment," *Econometrica*, 57(2), 447–480.

Malcomson, J. M. (2012): *Relational Incentive Contracts*pp. 1014–1065. Princeton University Press.

——— (2015): "Relational incentive contracts with productivity shocks," *Games and Economic Behavior*, 92, 122 – 137.

——— (2016): "Relational Incentive Contracts With Persistent Private Information," *Econometrica*, 84(1), 317–346.