

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

by

Matthias FAHN  
Nicolas KLEIN

Working Paper No. 1714  
July 2017

---

Johannes Kepler University of Linz  
Department of Economics  
Altenberger Strasse 69  
A-4040 Linz - Auhof, Austria  
[www.econ.jku.at](http://www.econ.jku.at)

[matthias.fahn@jku.at](mailto:matthias.fahn@jku.at)  
(ab 1.10.2017)

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

Matthias Fahn<sup>†</sup> and Nicolas Klein<sup>‡</sup>

October 24, 2018

## 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.

---

\*We thank the editor, John Asker, and three anonymous referees, for their suggestions and guidance, which have greatly improved the paper. We are also grateful to 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 Universitat de Barcelona, the MPI Bonn, McGill, Guelph, the Université Laval, the JKU Linz, the Université de Montréal, Munich, Paris (Séminaire Roy), Regensburg, Toulouse, the UNSW, the ETH Zurich, and at various conferences, for helpful comments. Matthias Fahn gratefully acknowledges financial support from the German Science Foundation (DFG) through collaborative research center CRC TRR 190. 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.

<sup>†</sup>JKU Linz and CESifo, matthias.fahn@jku.at.

<sup>‡</sup>Université de Montreal and CIREQ, kleinnic@yahoo.com.

# 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. 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.<sup>1</sup> 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), p. 100, 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.” Moreover, there is evidence to suggest that this “punishment mode” only lasts for a limited period of time.<sup>2</sup> Our paper provides an alternative explanation for the temporary lack of effectiveness of downsizing. In this view, implicit downsizing costs do

---

<sup>1</sup>See e.g. Cascio (1993) and Datta, Guthrie, Basuil, and Pandey (2010).

<sup>2</sup>See e.g. Goesaert, Heinz, and Vanormelingen (2015); Meuse, Bergmann, Vanderheiden, and Roraff (2004); Meuse and Dai (2013)

not indicate a lack of trustworthiness, nor do they result from punishment for a broken promise. Instead, they will arise as part of the path of play in an optimal relational contract, acting as a commitment device only to downsize when it is necessary to do so. As a consequence, these implicit downsizing costs allow for increased productivity in good times.

More specifically, our model starts from a standard relational-contracting framework, in which a principal and an agent interact repeatedly over time. The agent has to exert costly effort to produce output, which translates into a profit for the principal. Effort, while observable to both parties, is not contractible. Thus, bonuses paid to the agent for choosing the right level of effort can only be enforced via relational contracts. There, the principal can credibly commit to pay at most 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 ("type") 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 type of 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. Such an effort distortion reduces output and profits below levels that would be feasible at this point in time – which

however is optimal because it allows for higher output and profits in earlier, high-state, periods.

Thus, the negative consequences of downsizing that have been observed<sup>3</sup> might indeed serve as a commitment device for firms. In addition, a number of studies have shown that these negative consequences may only last for a limited period of time. Cascio (1993) e.g. notes that firms tend to increase labor input again shortly after downsizing, while the firm’s environment has not changed. Meuse, Bergmann, Vanderheiden, and Roraff (2004) and Meuse and Dai (2013) demonstrate that, while downsizing firms perform significantly worse than other firms by several financial measures, this difference gradually vanishes, eventually becoming insignificant. Conducting a survey of employees of a large high-tech firm, Amabile and Conti (1999) find that productivity significantly declined during and immediately after the downsizing process, recovering again after a while. 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.

This anecdotal evidence for our mechanism is further supported by studies pointing out that implicit downsizing costs are particularly prevalent in cases where incentives are arguably provided by relational contracts. For example, Guthrie and Datta (2008) find that the negative effect of downsizing is almost entirely driven by industries with high research and development (R&D) intensities, high growth, and low capital intensities. Furthermore, Guthrie and Datta (2008) state that employee effort components such as “creativity and initiative” (p. 112) are important in these industries, and that those are “associated with organizational variability and enhanced discretion, increasing the relative benefits derived from employee initiatives and contributions” (p.112). Since these aspects cannot be easily measured and hence specified in formal, court-enforceable, contracts, we would argue that they indicate a larger importance of relational contracts in the respective industries.

In a next step, we explore in Section 5.1 how the precise timing of in-

---

<sup>3</sup>Dozens of studies have analyzed the consequences of downsizing. While some have found downsizing to have the intended effect, such as e.g. a reduction of organizational slack, streamlined operations, or enhanced effectiveness, others find negative effects on firm performance (see Guthrie and Datta (2008) or Datta, Guthrie, Basuil, and Pandey (2010) for summaries of these studies). In the economics literature, by contrast, the consequences of downsizing have so far received less attention. Exceptions include e.g. Goesaert, Heinz, and Vanormelingen (2015) or Ahammer, Grübl, Lapinski, and Winter-Ebmer (2018).

formation revelation affects our outcomes. First, we assume that the state of the world is revealed later than in our main model, at the beginning of each respective period. In this case, private information is not costly, and the principal can credibly promise the full expected continuation value as a means of providing incentives. This implies that it is feasible and optimal to make the agent's compensation in a period independent of next period's type. In a second step, we explore the effects of the type being revealed earlier than in our main setting, assuming that next period's state of the world is observed by the principal already at the beginning of the current period, before the agent has chosen current-period effort. Now, the principal's private information is again costly and truth-telling conditions constrain profits, albeit due to an issue that has been absent before: Because a high state of the world in the next period potentially allows for higher effort – and consequently higher profits – in the current period, the principal has an incentive to misrepresent tomorrow's type as high, and subsequently to renege on the promised payment. It turns out that, on account of this constraint, it is not possible to generate higher profits if tomorrow's type is high. In contrast to before, truth-telling can now also be achieved via fixed payments made to the agent at the beginning of a period. Put differently, either effort will be independent of tomorrow's state of the world, or a high state tomorrow also triggers higher effort, albeit with a fixed payment made to the agent *before* effort is delivered. This payment has to fully make up for the increased value of production.

Thus, when the principal strives to motivate the agent to exert effort, she is tempted to claim that the future looks bright and that hence the agent will be compensated for his hard work. Yet when the principal is supposed to compensate the agent, she is tempted to claim that the future looks grim – and that the agent will consequently have to accept lower compensation.

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. Consequently, implicit downsizing costs will only last for one period in this setting, after which effort increases to its undistorted level even if the firm's prospects remain bad. In Section 5.2, we extend our analysis to (fully or partially) persistent shocks. In these cases, 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.

Finally, note that, in our model, we explore the employment of one agent who exerts effort. We choose this approach for reasons of tractability, and in order better to relate our approach to “standard” relational contracting models. However, one could also interpret our effort levels as the aggregate labor input by a firm’s workforce.<sup>4</sup> Thus, a reduction of effort levels in our model can be construed as either a diminution of the size of the workforce or a reduction in individual workers’ effort levels. While these two very distinct real-world phenomena map into a reduction of effort in our model, our analysis would suggest that, *given* a firm engages in downsizing, this should optimally be accompanied by some sort of temporary decline in productivity. Whereas the former interpretation is supported by Cascio (1993), who reports that following downsizing, some “managers ended up replacing some of the very people they had dismissed” (p. 98), the latter seems to correspond to the turmoil discussed in most of the management literature. To align our results with this observation, note that, in reality, employment is also affected by many aspects outside the scope of our model. For example, keeping a worker on causes fixed per-period costs of employment, while firing and/or replacing a worker is costly as well.<sup>5</sup> Whereas fixed per-period costs of employment accrue in every period, separation and replacement costs accrue only once for a given employment relationship. This would suggest that temporary variations of total labor input would rather be effected by changes in individual effort levels, while the workforce is kept stable in order to economize on replacement costs. Long-term changes in desired labor input, on the other hand, would tend to involve a larger focus on adjusting the size of the firm’s workforce. In this view, our theoretical results for the case of (at least partially) persistent shocks are in line with the observation made by the management literature that permanent downsizing decisions are often followed by a phase of turmoil.

---

<sup>4</sup>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).

<sup>5</sup>For example, Dube, Freeman, and Reich (2010) report that costs to replace a worker in California (including recruitment, selection and screening, as well as the costs of learning on the job and separation costs) amount to between \$ 2,000 and \$ 7,000. Muehleman and Leiser (2018) show that in Switzerland, the average hiring costs amount to about 16 weeks of wages.

## Related Literature

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).<sup>6</sup> 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, analyzing 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 literature on *implicit contracts* also explores the optimal behavior of firms in the case of asymmetric information on the marginal profitability of employment (see Hart (1983), Azariadis (1983), or Grossman and Hart (1983)). There as well, inefficiently low employment in bad states

---

<sup>6</sup>See Malcomson (2012) for an overview of the literature on relational contracts.



of the world serves as a commitment device not to under-report the state of the world. This, however, is the consequence of an optimal risk-sharing arrangement between a risk-averse firm and its risk-averse workers.

The rest of the paper is set up as follows: Section 2 introduces the model; Section 3 reviews some benchmarks, in particular the case of public information; Section 4 presents the main results; Section 5 explores extensions; Section 6 concludes. Proofs not given in the text can be found in the Appendix.

## 2 The Model

### 2.1 Environment and Information

There is one principal (“she”) and one agent (“he”), who are both risk neutral and who interact repeatedly in periods  $t = 1, 2, \dots$ . At the beginning of every period  $t$ , the principal makes an employment offer to the agent, consisting of a contractible 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 a cost of  $cn_t$ , where  $c > 0$  is his marginal cost of effort.<sup>7</sup> The agent’s effort is observed by the principal. The effort level  $n_t$  translates into output  $g(n_t)$ , where  $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is  $C^2$ , satisfies  $g' > 0 > g''$  and  $\lim_{n \downarrow 0} g'(n) = \infty$ ,  $\lim_{n \rightarrow \infty} g'(n) = 0$ . 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.

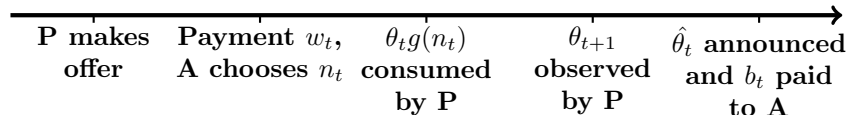
The principal additionally knows her *type* in period  $t + 1$ ,  $\theta_{t+1} \in \{\theta^l, \theta^h\}$ , before deciding on the bonus payment  $b_t$  in period  $t$ ; the agent never learns the realizations of the principal’s types. The values satisfy  $\theta^h > \theta^l > 0$  and are commonly known. Period- $t$  output  $g(n_t)$  translates into revenues  $\theta_t g(n_t)$  for the principal. We write  $\theta^t = \{\theta_\tau\}_{\tau=1}^t$  for the sequence of realizations of the principal’s types up to, and including, period  $t$ . While  $\theta_1 = \theta^h$ , the principal’s types  $\{\theta_t\}_{t=2}^\infty$  are i.i.d. across periods (except in Section 5.2); for all  $t = 2, 3, \dots$ , the probability that  $\theta_t = \theta^h$  is  $q \in (0, 1)$ . The probability  $q$ , as

---

<sup>7</sup>As discussed in the Introduction, the agent’s effort choice  $n_t$  can be interpreted as the aggregate labor input by the firm’s workforce as well. Thus, changes in  $n_t$  can be interpreted as adjustments of either the size of the workforce or individual workers’ effort levels.

well as the principal's type in the first period, are common knowledge, as are all other parameters and the structure of the game. When paying the bonus, the principal can also send a non-verifiable cheap-talk message  $\hat{\theta}_t \in \{\theta^l, \theta^h\}$  to the agent.<sup>8</sup>

The following figure summarizes the timing within each period:



## 2.2 Histories and Strategies

The public events (i.e. those that can be observed by both the principal and the agent) in period  $t$  are given by  $h_t = (w_t, d_t, n_t, b_t, \hat{\theta}_t)$ . 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  $\{\sigma_t^A\}_{t=1}^\infty$ , where, for each  $t \in \mathbb{N}$ ,  $\sigma_t^A = (d_t, n_t)$ , and  $d_t : \mathcal{H}^{t-1} \times [-\bar{w}, \bar{w}] \rightarrow \{0, 1\}$ ,  $(h^{t-1}, w_t) \mapsto d_t(h^{t-1}, w_t)$  and  $n_t : \mathcal{H}^{t-1} \times [-\bar{w}, \bar{w}] \times \{0, 1\} \rightarrow \mathbb{R}_+$ ,  $(h^{t-1}, w_t, d_t) \mapsto n_t(h^{t-1}, w_t, d_t)$ .

The *principal events* in period  $t$  are given by  $\mathfrak{h}_t = (w_t, d_t, n_t, \theta_{t+1}, b_t, \hat{\theta}_t)$ ; 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 her wage offers as a function of the previous principal history, as well as her bonus payment and report as a function of the previous history, current-period wages and effort, as well as her type in the next period. Formally, it is a sequence of mappings

<sup>8</sup>Given our focus on pure strategies (see below for details), the restriction to binary messages is without loss.

$\{\sigma_t^P\}_{t=1}^\infty$ , where, for each  $t \in \mathbb{N}$ ,  $\sigma_t^P = (w_t, b_t, \hat{\theta}_t)$ , and  $w_t : \mathfrak{H}^{t-1} \rightarrow [-\bar{w}, \bar{w}]$ ,  $\mathfrak{h}^{t-1} \mapsto w_t(\mathfrak{h}^{t-1})$ ,  $b_t : \mathfrak{H}^{t-1} \times [-\bar{w}, \bar{w}] \times \{0, 1\} \times \mathbb{R}_+ \times \{\theta^l, \theta^h\} \rightarrow [0, \bar{b}]$ ,  $(\mathfrak{h}^{t-1}, w_t, d_t, n_t, \theta_{t+1}) \mapsto b_t(\mathfrak{h}^{t-1}, w_t, d_t, n_t, \theta_{t+1})$ , with the restriction that  $d_t = 0 \Rightarrow b_t(\mathfrak{h}^{t-1}, w_t, d_t, n_t, \theta_{t+1}) = 0$ , and  $\hat{\theta}_t : \mathfrak{H}^{t-1} \times [-\bar{w}, \bar{w}] \times \{0, 1\} \times \mathbb{R}_+ \times \{\theta^l, \theta^h\} \rightarrow \{\theta^l, \theta^h\}$ ,  $(\mathfrak{h}^{t-1}, w_t, d_t, n_t, \theta_{t+1}) \mapsto \hat{\theta}_t(\mathfrak{h}^{t-1}, w_t, d_t, n_t, \theta_{t+1})$ . A pure public strategy by the principal is a pure strategy which does not condition on her *past* private information, which is no longer payoff-relevant. Formally, a strategy  $\{\sigma_t^P\}_{t=1}^\infty$  is said to be a *public strategy* if, for each period  $t \in \mathbb{N}$ , it can be written  $\sigma_t^P = (\tilde{w}_t, \tilde{b}_t, \tilde{\theta}_t)$ , where  $\tilde{w}_t : \mathcal{H}^{t-1} \times \{\theta^l, \theta^h\} \rightarrow [-\bar{w}, \bar{w}]$ ,  $(h^{t-1}, \theta_t) \mapsto \tilde{w}_t(h^{t-1}, \theta_t)$ ,  $\tilde{b}_t : \mathcal{H}^{t-1} \times [-\bar{w}, \bar{w}] \times \{0, 1\} \times \mathbb{R}_+ \times \{\theta^l, \theta^h\} \rightarrow [0, \bar{b}]$  and  $\tilde{\theta}_t : \mathcal{H}^{t-1} \times [-\bar{w}, \bar{w}] \times \{0, 1\} \times \mathbb{R}_+ \times \{\theta^l, \theta^h\} \rightarrow \{\theta^l, \theta^h\}$ .

### 2.3 Payoffs and Equilibrium

If  $d_t = 1$ , the principal's period payoff in period  $t$  is given by  $\theta_t y_t - w_t - b_t$ ; the agent's is given by  $w_t - n_t c + b_t$ . If  $d_t = 0$ , principal and agent get their outside option payoffs in period  $t$ , which are set to zero. Both players discount future payoffs with the discount factor  $\delta \in (0, 1)$ .

Our solution concept is perfect Bayesian equilibrium in (pure) public strategies as defined above, to which in the following we simply refer as equilibrium. 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  $g(n_t)$  is assumed to be non-verifiable.

Our objective is to find an equilibrium that maximizes the principal's ex ante expected profit  $\Pi_1$ . As expected surplus can be transferred freely through  $w_1$ , the fixed wage in the first period, any equilibrium maximizing  $\Pi_1$  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 an equilibrium leading to a joint surplus of  $s \geq 0$ . Then, there exists an equilibrium giving the principal an expected payoff of  $\pi$  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. We shall focus on truth-telling equilibria; i.e., on the equilibrium path, reported types will coincide with the history of past type realizations,  $\theta^t = \{\theta_\tau\}_{\tau=1}^t$ . By our choice of equilibrium concept, this is without loss in our main model of Sections 3-4. 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  $n_t(\mathfrak{h}^{t-1}, w(\theta^t), 1)$ , the agent's effort choice on the equilibrium path in period  $t$  given history  $\theta^t$ . In addition, we shall use superscripts  $h$  or  $l$  to indicate the type in period  $t+1$ , given history  $\theta^t$ , 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  $\theta^t$ , 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, and  $U(\theta^t) = U^i(\theta^{t-1})$  for the agent's expected on-path utility, at the beginning of period  $t$ , given the history of type realizations  $\theta^t$  with  $\theta_t = \theta^i$  ( $i \in \{h, l\}$ ).

Thus, we can write

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

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

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

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

### 3 Some Benchmarks

In this section, we analyze a few natural benchmarks against which to measure our equilibrium.

First, 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  $n_l^{FB} \equiv n^{FB}(\theta^l)$ .

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  $\theta^t$ , 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. In particular, as truth-telling gives her first-best profits, the principal has no incentive to lie.

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 can always guarantee itself this static SPE payoff, which constitutes its minmax-payoff. As we are interested in the best possible equilibrium for the principal, it is without loss for us to focus on equilibria in which any observable deviation triggers this harshest possible punishment.<sup>9</sup>

### 3.1 Public Types

In this section, we 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

---

<sup>9</sup>See Abreu (1988) on the optimality of such simple penal codes.

constraint must hold, for all histories  $\theta^t$ :

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

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 Compatibility Constraint is given by

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

It must also be optimal, after any history  $\theta^t$ , 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

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

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

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

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

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.

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  $\theta^t$ . In Section A in the Appendix (Lemma 5), we show that there exists an optimal equilibrium in which the agent does not get a rent and the (IC) constraint will bind after any history. Moreover, we show 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.

Note that, as is also the case e.g. in Levin (2003) or MacLeod and Malcolmson (1989), *enforceable* effort in any given period does not depend on the current type but only on expected future profits. Indeed, current revenues are already sunk when 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.

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

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

For the proof, see Section A in the Appendix.

If  $\delta$  is high enough, the first best is achievable. 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  $\tau$  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. 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.

## 4 Private Types

Now, let us 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.

In truth-telling equilibrium, the principal needs sufficient incentives to reveal her type in every period. Specifically, after any history  $\theta^t$ , it must be optimal for her 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  $t + 1$ .<sup>10</sup>

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 (see Hendon, Jacobsen, and Sloth (1996)). 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

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

---

<sup>10</sup>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  $t + 1$ .



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

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

Therefore, the principal is willing to tell the truth in equilibrium following history  $\theta^t$  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:

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

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

Thus, the principal's objective is to maximize

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

where  $\theta^1 = \theta_1 = \theta^h$ , subject to (DEh), (DEl), (TTh), (Ttl), (IR) and (IC) at each history  $\theta^t$ .

As we show in Section B.1 in the Appendix, this optimization problem can be substantially simplified. First, the (DEh) constraint can be omitted because it is always more tempting for the principal to underreport tomorrow's type than to shut down. Furthermore, the agent never gets a rent, and the (IC) constraint always holds as an equality. Moreover,  $b^h(\theta^t) \geq b^l(\theta^t)$ , which implies that the principal 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. We further show in Section B.1 in the Appendix that

these constraints can equivalently be combined into one, and that consequently  $n^h(\theta^t)$  will be independent of  $\theta^t$ , while  $n^l(\theta^t)$  will only depend on the number  $i \in \{0, 1, 2, \dots\}$  of consecutive low shocks after the last high period. Therefore, we write  $n_i^l$  to describe low-type effort levels. Thus, the optimization problem boils down to choosing  $(n^h, n_i^l)_{i \in \mathbb{N}}$  so as to maximize

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

subject to

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

and

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

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

As mentioned above, the (EC) constraints are obtained by combining (DEl) and (TTh) constraints for the respective effort levels. The left-hand side of an (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 against discounted expected future profits. 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  $\delta q g(n^l) (\theta^h - \theta^l) \geq 0$ , 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  $\theta^h$  (which happens with probability  $q$ ), because she always has the option of falsely claiming that the type is  $\theta^l$ . If she does so, she will get  $\theta^h g(n^l)$  in the next period, rather than just  $\theta^l g(n^l)$ , 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.

(EC) constraints also imply that optimal efforts are 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. The reason is that 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. Thus, the optimal effort level in a given low period depends on the optimal effort level in the previous period.

In conclusion, 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. Then, 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 aspect adds another dimension to the credibility problem typical for relational contracts, in the sense that 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.

Our first result shows that if the discount factor is close enough to 1, the first best can be achieved.

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

For the proof, see Section B.2 in the Appendix.

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. On the other hand, there is the need to incentivize the principal to tell the truth because a higher enforceable bonus when the 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  $\delta$ , 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  $\delta$  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 presents the first main result of this paper. It characterizes an optimal outcome, given that the discount factor is too low to implement  $n_h^{FB}$  but high enough to implement  $n_i^{FB}$ . Then, the principal needs to be dissuaded from claiming that next period's type is low when it is in fact high. Therefore, 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.

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

For the proof, see Section B.2 in the Appendix.

Note that, for the first-best solution, the (ECh) and (ECl) constraints are identical but for the first term, which is  $n_h^{FB}$  and  $n_i^{FB}$ , respectively. Thus, as  $\delta$  decreases, (ECh) starts binding before the (ECl) 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  $n_i^{FB}$  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.

To develop an intuition for this result, it is helpful to split up (ECh) again and to take a look at its individual components, the (TTh) and (DEl)

constraints for  $n^h$ :

$$-b^h(n^h) + \delta\Pi^h \geq -b^l(n^h) + \delta\Pi_0^l + \delta g(n_0^l) (\theta^h - \theta^l) \quad (\text{TTh})$$

$$-b^l(n^h) + \delta\Pi_0^l \geq 0 \quad (\text{DEl})$$

Consider an arbitrary effort level  $n^h < n_h^{FB}$  together with bonuses  $b^h(n^h)$  and  $b^l(n^h)$  such that (TTh) and (DEl) hold as equalities, i.e., (ECh) binds. (Indeed, if only one of them was binding, for example (TTh), a first response would be to reduce  $b^h(n^h)$  by  $\varepsilon > 0$  and to increase  $b^l(n^h)$  by  $\frac{q}{(1-q)}\varepsilon$ , which would allow for a further increase in  $n^h$ .) A natural, surplus-neutral, way of relaxing (TTh) would be for the principal to pay a rent  $R$  to the agent following an announcement of a low state at the end of the period, an arrangement equivalent to making such a payment at the beginning of the next period. This reduces the right-hand side of (TTh) by  $R$ , thereby relaxing (TTh) and allowing the principal to increase  $b^h(n^h)$  by  $R$  as well. However, the principal also needs an incentive to pay  $R$ . Hence, the (DEl) constraint becomes  $-b^l(n^h) - R + \delta\Pi_0^l \geq 0$  and is tightened. As (DEl) was binding before,  $b^l(n^h)$  must be reduced by  $R$  in order to keep (DEl) satisfied. But this once again increases the right-hand side of (TTh) by  $R$ , making it necessary to reduce  $b^h(n^h)$  by the same amount (and thus to its original level) – and, at the end, nothing has been gained because (EC) constraints are not relaxed.

Thus, (ECh), the combination of (TTh) and (DEl) constraints, can only be relaxed by downsizing costs if those hit a lying off-path principal harder than a principal who truthfully claims next period's type to be low. Mere transfers cannot achieve this goal as we have just seen. However, 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 and therefore relaxes (ECh). To see that, assume that in the situation considered in the previous paragraph, effort after an announcement of a low state is reduced by a small  $\varepsilon > 0$  in the following period. This reduces  $\Pi_0^l$  – and consequently  $b^l(n^h)$  – by  $\varepsilon(\theta^l g' - c)$ . However, the right-hand side of (TTh) is decreased by  $\varepsilon\delta g'(\theta^h - \theta^l)$ , which allows for an increase in  $b^h(n^h)$  by the same amount. As  $\theta^l g'(n_i^{FB}) - c = 0$ , the resulting surplus destruction, as well as the necessary reduction in  $b^l(n^h)$ , are only of second order at  $n_i^{FB}$ . The possible increase of  $b^h(n^h)$ , though, is of first order, and (ECh) is eventually relaxed. Therefore, it is optimal to use a reduction of  $n_0^l$  in order to

implement a larger  $n^h$ . Thus, the game exhibits memory, and the equilibrium is not sequentially optimal, in that  $n_i^{FB} (> n_0^l)$  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 is sequentially optimal. In our iid model, this distortion in effort levels only lasts a single period, and  $n_i^l = n_i^{FB}$  for  $i \geq 1$ . This is due to two reasons. First, reducing  $n_i^l$  for  $i \geq 1$  would not allow for a further increase in  $n^h$  because the resulting distortions in later periods would hit on-path and off-path principals alike.<sup>11</sup> 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.

Given  $\delta$  is below  $\underline{\delta}$ , ECl constraints also bind for  $i \geq 1$ . This considerably complicates our maximization problem because all (ECl) constraints potentially interact: A higher  $n_{i+1}^l$  tightens (ECl), whereas a higher  $n_i^l$  might require a reduction of  $n_{i+1}^l$  and consequently relax (ECl+1). Therefore, we have to consider infinitely many constraints. In the following, we derive a number of properties of effort levels  $n_i^l$  if  $\delta < \underline{\delta}$ . Due to the complexity of the problem, we restrict ourselves to the case  $q\theta^h \geq \theta^l$ :

**Proposition 5** *Assume  $q\theta^h \geq \theta^l$ . There exists a left-neighborhood of  $\underline{\delta}$  such that optimal effort levels  $n_i^l < n_i^{FB}$ ,  $i \geq 1$ , are characterized by one of the following cases:*

- $n_j^l = n_1^l$  for all odd  $j$  and  $n_\iota^l = n_2^l$  for all even  $\iota$ , with  $n_1^l > n_2^l$ ;
- $n_1^l > n_3^l > n_5^l > \dots$  and  $n_2^l < n_4^l < n_6^l < \dots$ , with  $\sup_{j \in \mathbb{N}} n_{2j}^l \leq \inf_{j \in \mathbb{N}} n_{2j-1}^l$ ;
- $q\theta^h = \theta^l \Leftrightarrow n_i^l = n_{i+1}^l$  for all  $i \geq 1$ .

We prove this proposition by Lemmata 10 - 23 in Section B.2 the Appendix. It shows that, unless  $q\theta^h = \theta^l$ , effort levels oscillate, with either a constant or a decreasing amplitude, starting at their highest level  $n_1^l$ . If  $q\theta^h = \theta^l$ , by contrast, effort levels  $n_i^l$  ( $i \geq 1$ ) remain constant, as for intermediate discount factors. We still observe overshooting in this region, as  $n_0^l$  is constrained only by ECh, while ECl0 is slack. Furthermore,  $n^h < n_h^{FB}$  and  $n_0^l < n_1^l$ .

---

<sup>11</sup>As we shall see in Section 5.2, distortions last longer when types are (fully or partially) persistent.

## 5 Extensions and Robustness

In this section, we explore several extensions of our main model. First, we analyze different assumptions regarding the timing of events within a given period. Second, we show that overshooting also occurs with persistent shocks.

### 5.1 Timing

Here, we vary the timing of the revelation of next period's type. We present a summary of our results; a more detailed description can be found in Section C in the Appendix. First, we show that a later revelation – the type of period  $t$  is revealed at the beginning of period  $t$  – increases the principal's profits compared to our main case. Then, private information is not costly and the outcome equivalent to the case of public information. Second, we assume that the principal observes the type of period  $t+1$  already at the beginning of period  $t$ , before the agent exerts period- $t$  effort. In this case, private information is costly, but the nature of the costs and the principal's response substantially differs. Thus, our overshooting result in the previous section relies on there being no possibility of monetary transfers in between the time of the agent's effort choice and the revelation of private information to the principal.

**Late Revelation** Assume that the type of period  $t$  is revealed at the beginning of period  $t$  (this is equivalent to having  $\theta_{t+1}$  revealed in period  $t$ , but *after*  $b_t$  has been paid). In contrast to before, the bonus  $b_t$  is not a function of next period's type anymore and hence is certain (on the equilibrium path) at the time of effort choice. However, in period  $t$ , the agent's future compensation might still depend on  $\theta_{t+1}$  through the fixed wage  $w_{t+1}$ . But it turns out that it remains (weakly) optimal to use only certain period- $t$  bonus payments to reward period- $t$  effort. This implies that effort and compensation levels in a profit-maximizing equilibrium for the case of public information also satisfy the truth-telling constraints under private information.

**Proposition 6** *Assume  $\theta_t$  is privately revealed at the beginning of period  $t$ . Then, private information is not costly and the outcome equivalent to the case of public information.*

For the proof, see Section C.1 in the Appendix.

To understand the intuition behind this result, note that in our benchmark case, it would also be feasible to make the agent's compensation independent of the realization of next period's type. But such a payment structure would leave some slackness in the dynamic enforcement constraints, which could be utilized in order to increase implementable effort. At some point, however, truth-telling constraints start to bind, leading to the structure of the profit-maximizing equilibrium that we have derived in Section 4. Here, by contrast, the agent's compensation can be independent of next period's type while fully exhausting dynamic enforcement constraints. Thus, implementable effort cannot be further increased. Therefore, it is optimal to make the agent's compensation (conditional on effort) independent of the realization of next period's type, and the principal's private information does not matter.

**Early Revelation** Now, we describe the properties of a profit-maximizing equilibrium for the case that the type of period  $t+1$  is already revealed at the beginning of period  $t$ , before effort  $n_t$  is chosen by the agent. If  $\theta_{t+1}$  is *publicly* revealed at the beginning of period  $t$ ,  $n_t$  will be a function not only of today's, but generally also of tomorrow's, type. This is because enforceable effort in a given period is a function of expected future profits. A high type tomorrow is associated with higher expected future profits and thus a higher enforceable effort level today. Now, we use superscripts to indicate equilibrium values as functions of this and next period's types. For example,  $n^{hh}$  is equilibrium effort in case today's and tomorrow's type are high,  $n^{hl}$  is equilibrium effort if today's type is high and tomorrow's type is low, and so on. In the following, we focus on *quasi-stationary* equilibria, where, after all histories, actions depend only on today's and tomorrow's types.

**Lemma 1** *Assume  $\theta_{t+1}$  is publicly revealed at the beginning of period  $t$ . Then, there are levels of the discount factor,  $\bar{\delta}$ ,  $\tilde{\delta}$  and  $\underline{\delta}$ , with  $0 < \underline{\delta} < \tilde{\delta} < \bar{\delta} < 1$ , such that*

- $n^{hh} = n^{hl} = n_h^{FB} > n^{ll} = n^{lh} = n_l^{FB}$  for  $\delta \geq \bar{\delta}$ ;
- $n^{hl} < n^{hh} = n_h^{FB}$  and  $n^{ll} = n^{lh} = n_l^{FB}$  for  $\tilde{\delta} \leq \delta < \bar{\delta}$ .
- For  $\underline{\delta} \leq \delta < \tilde{\delta}$ , there are levels of the discount factor  $\delta^h$  and  $\delta^l$ , such that



- $n^{hl} < n^{hh} < n_h^{FB}$  for  $\delta < \delta^h$ ;
- $n^{ll} < n^{lh} = n_l^{FB}$  for  $\delta < \delta^l$ ;
- $n^{hl} < n^{hh} < n_h^{FB}$  and  $n^{ll} < n^{lh} < n_l^{FB}$  for  $\delta < \underline{\delta}$ ; in this case,  $n^{ll} = n^{hl} < n^{lh} = n^{hh}$ .

For the proof, see Section C.2 in the Appendix.

The early revelation of information is costly compared to a later revelation – because no “cross-subsidization” of high future profits to low future profits is feasible anymore. If information is revealed later (like in the previous section), the resulting uncertainty allows us to use potential high future profits to motivate effort also in case future profits are actually low. Here, by contrast, a binding (DEhl) constraint cannot be relaxed by a potential slackness of (DEhh), as would be the case if information was revealed later.

If next period’s type is only privately revealed to the principal at the beginning of the present period, the relevant trade-off in truth-telling equilibrium is different from the main part of this paper. There, the principal is tempted to underreport her type because this results in a lower bonus payment to the agent in the present period, at the cost of a distorted production in the next period. The *current* effort level is unaffected by a lie of the principal, as the corresponding output has already been realized. If next period’s type is revealed at the beginning of the present period, however, under-reporting tomorrow’s type already results in a lower output today. Therefore, a lie is associated with *present and future* production inefficiencies. The resulting costs make the principal’s temptation to under-report her type vanish, and overshooting as a consequence of downsizing is not needed to induce truth-telling. Truth-telling constraints can still severely constrain profits, though, due to an issue that was absent before: Because having a high type in the next period potentially allows for higher effort, and consequently higher profits, today, the principal might be tempted falsely to claim that tomorrow’s type is high – and then to renege on the promised payment. It turns out that this constraint in fact prevents the principal from achieving higher profits if tomorrow’s type is high. Profits will only be a function of today’s type, and will always be constrained by  $\delta \bar{\Pi}^l$ , the discounted expected profits given tomorrow’s type was low – no matter if it is actually high or low. Proposition 7 characterizes an equilibrium maximizing  $\bar{\Pi}^h$ , the expected profits if the current type is high and tomorrow’s type is high with probability  $q$ .

**Proposition 7** *Assume  $\theta_{t+1}$  is privately revealed at the beginning of period  $t$ . Then, among the class of quasi-stationary equilibria in which the agent does not get a rent and faces binding (IC) constraints following any history,  $\bar{\Pi}^h$  is maximized when  $n^{hh} = n^{hl} \equiv n^h$  and  $n^{lh} = n^{ll} \equiv n^l$ . Moreover, there exist discount factors  $\bar{\delta}$  and  $\underline{\delta}$ , with  $0 < \underline{\delta} < \bar{\delta} < 1$ , such that*

- $n^h = n_h^{FB}$  and  $n^l = n_l^{FB}$  for  $\delta \geq \bar{\delta}$ ;
- $n^l = n_l^{FB} < n^h < n_h^{FB}$  for  $\underline{\delta} < \delta < \bar{\delta}$ ;
- $n^l = n^h \leq n_l^{FB}$  for  $\delta \leq \underline{\delta}$

*in this equilibrium.*

For the proof, see Section C.2 in the Appendix.

Importantly, effort is only a function of today's type. High future profits cannot be used to implement higher effort today. If this were the case, the principal would have an incentive to misreport her type and then shut down. Moreover, the temptation now lies in over-reporting one's type because this would be associated with higher productivity.

Note that for discount factors such that the first-best effort cannot be implemented, it only matters that the principal's *profits* not be larger if tomorrow's type is high. Instead of equalizing effort levels, we could also have  $n^{hh} > n^{hl}$ , together with a payment to the agent before his effort choice, in the form of a positive fixed wage  $w^{hh}$ .<sup>12</sup> Thus, in order to make use of higher future profits and induce the agent to work harder, the principal immediately has to pay him for the extra effort. If the principal is able to extract the expected value of these payments at the beginning of the game, this equilibrium generates higher expected profits than that of Proposition 7.

## 5.2 Persistent Types

So far, we have assumed that the principal's types are iid across periods. In this section, we show that implicit downsizing costs may also obtain if shocks are persistent – within our initial setup where  $\theta_{t+1}$  is revealed to the principal

<sup>12</sup>Then, TThl2, which has been defined in the proof of Proposition 7, is the tighter constraint, and becomes  $\theta^h g(n^{hl}) - n^{hl}c + \delta \bar{\Pi}^l \geq -w^{hh} + \theta^h g(n^{hh})$ . This is satisfied for  $\theta^h g(n^{hh}) - w^{hh} = \theta^h g(n^{hl})$  (given dynamic enforcement constraints hold).

before the period- $t$  bonus is paid, but after effort has been exerted.<sup>13</sup> First, we explore permanent shocks. We assume that the principal starts out with a high type, and 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. Later, we argue that the case with shocks that are persistent but not permanent yields similar results.

With permanent shocks, equilibrium high-type effort remains constant, whereas low-type effort depends on the distance in time to the (now permanent) switch from high to low. All payoffs and constraints are characterized equivalently to the case with iid shocks, with the caveat that a reversion to the high state is impossible once a low type has been reported.

In Section D.1 in the Appendix, we show that, in order for effort levels  $n^h$  and  $(n_i^l)_{i \in \mathbb{N}}$  to be enforceable, the following (ECh) and (DEli) constraints are necessary and sufficient

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

$$-n_i^l c + \delta \Pi_{i+1}^l \geq 0, \quad (\text{DEli})$$

for all  $i \geq 0$ .

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  $\delta q (\theta^h - \theta^l) \sum_{i=0}^{\infty} (\delta q)^i g(n_i^l)$ , while it was  $\delta q (\theta^h - \theta^l) g(n_0^l)$  before.

In Lemma 26 in Section D.1 of the Appendix, we show that, as before, constraints are tightened for lower values of  $\delta$ . As  $\delta \rightarrow 1$ , 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  $\frac{q}{1-q} (\theta^h - \theta^l) g(n_i^{FB}) < \infty$ . Thus, if  $\delta$  is sufficiently large, the first-best effort levels  $n_h^{FB}$  and  $n_l^{FB}$  can be implemented. As  $\delta$  leaves this range, it is of interest whether (ECh) or (DEli)

---

<sup>13</sup>A more detailed analysis can be found in Section D in the Appendix.

constraints start binding first. Let  $\delta^l := \frac{n_i^{FB} c}{\theta^l g(n_i^{FB})}$  denote the discount factor at which (DE*l*)-constraints start binding for first-best effort levels, and  $\delta^h$  the corresponding discount factor for the (ECh)-constraint. It can be shown that (ECh) binds first if  $q < \frac{\theta^l g(n_i^{FB})(n_h^{FB} - n_i^{FB})}{\theta^h n_i^{FB} (g(n_h^{FB}) - g(n_i^{FB}))}$ ; 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 8** *Assume  $q < \frac{\theta^l g(n_i^{FB})(n_h^{FB} - n_i^{FB})}{\theta^h n_i^{FB} (g(n_h^{FB}) - g(n_i^{FB}))}$  and  $\delta \in [\delta^l, \delta^h)$ . Then,  $n^h < n_h^{FB}$ . Furthermore, for all  $i \in \mathbb{N}$ ,  $n_i^l < n_{i+1}^l < n_i^{FB}$ , with  $\lim_{i \rightarrow \infty} n_i^l = n_i^{FB}$ .*

For the proof, see Section D.1 in the Appendix.

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 is always 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. As with iid shocks, our optimal self-enforcing contract is thus not sequentially optimal since  $n_i^{FB}$  would satisfy all (DE*l*) constraints.

In Section D.2 in the Appendix, we also analyze a setting where states evolve according to a Markov chain with positive autocorrelation. If today's

type is high, the probability of tomorrow's type being high is  $q^h$ . If today's type is low, this probability is  $q^l < q^h$ . For some parameters, we still observe overshooting in this case if the type moves from high to low, with recovery being gradual and never complete, as in the case of permanent shocks. Concerning the intuition for this result, recall that with iid shocks, the principal was tempted to under-report her type in order to save on her bonus payments today. This led to overshooting in a single period in the optimal equilibrium. This overshooting induced costs of lying that were larger than the principal's costs on the equilibrium path. Now, we face a similar tradeoff, with an additional component, however: After a deviation (i.e., reporting a low type although it is actually high), the probability of observing another high period, implying a larger bonus in the next period, is higher than if the type had indeed been low. Therefore, the costs of a deviation, in relation to those arising when the type is indeed low, increase in next period's high-type bonus. Therefore, with positive auto-correlation, the high-type bonus in period  $t + 1$  should be as high as feasible to deter lying in period  $t$ , an effect that does not arise with iid shocks. Since high-type bonuses are restricted by the respective (TTh) constraints, these will bind in all subsequent periods. Now, binding (TTh) constraints are optimally associated with distortions of low-type effort levels, for the same reasons as above. Therefore, distortions continue in subsequent low periods, albeit at a decreasing intensity.

## 6 Conclusion

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. As voluntary transfers would hurt on-path and off-path principals alike, these costs manifest themselves in a reduction of performance below feasible levels, because this imposes steeper costs on off-path principals. Our model thus provides a novel perspective on the phenomenon of implicit downsizing costs, which we view as complementary to the psychological aspects being advanced by the strategic management literature.

# Appendix

## A Proofs for Section 3

Before proving Proposition 2, we establish the following lemma, which details some characteristics of an optimal solution.

**Lemma 2** *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  $\theta^t$ .

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

**Proof:** We shall first show that there exists an optimal equilibrium such that  $U(\theta^t) = 0$  for all histories  $\theta^t$ . 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  $\theta^t$  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  $\theta^t$  and all predecessor histories unaffected. Furthermore, the principal's profits at history  $\theta^t$  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  $\theta^t$ .

Now, suppose that there exists a history  $\theta^\tau$  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  $\varepsilon$ , the (DE) constraint for history  $\theta^\tau$  is relaxed, and all constraints for all other histories  $\theta^t$  are unaffected by this change. This adjustment potentially increases profits if (DE) for history  $\theta^\tau$  binds, and leaves profits unaffected if (DE) for history  $\theta^\tau$  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  $\theta^t$ .

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

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

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  $\theta^{\bar{\tau}}$  and  $\theta^{\bar{h}}$ , with  $n^h(\theta^{\bar{\tau}}) > n^h(\theta^{\bar{h}})$ . If the profits being produced in the continuation play following  $(\theta^{\bar{\tau}}, \theta^{\bar{h}})$  are higher, it is possible to implement  $n^h(\theta^{\bar{\tau}})$  with the continuation play following  $(\theta^{\bar{\tau}}, \theta^{\bar{h}})$ . In this case, the principal can therefore increase her profits following history  $(\theta^{\bar{\tau}}, \theta^{\bar{h}})$  by increasing the current period's effort level to  $n^h(\theta^{\bar{\tau}})$ , while leaving the continuation play unchanged. Now, suppose that it is not possible to implement  $n^h(\theta^{\bar{\tau}})$  with the continuation play following  $(\theta^{\bar{\tau}}, \theta^{\bar{h}})$ . This implies that the profits created by the continuation play following  $(\theta^{\bar{\tau}}, \theta^{\bar{h}})$  are lower than the continuation play following  $(\theta^{\bar{h}}, \theta^{\bar{h}})$ . Furthermore, because  $n^h(\theta^{\bar{\tau}})$  is enforceable, it is possible to replace the continuation play following  $(\theta^{\bar{\tau}}, \theta^{\bar{h}})$  with the continuation play following  $(\theta^{\bar{h}}, \theta^{\bar{h}})$ , thereby relaxing the (DE) constraint in  $\bar{\tau}$ . It thus becomes possible to increase  $n^h(\theta^{\bar{\tau}})$  to  $n^h(\theta^{\bar{h}})$ . 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.  $\blacksquare$

## 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

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

where  $\lambda_{DE_i}$  denotes the Lagrange multiplier associated with the (DE)-constraint, given the current type is  $\theta^i \in \{\theta^l, \theta^h\}$ .

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

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

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

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  $(n^h, n^l, \lambda_{DE_h}, \lambda_{DE_l})$  is (jointly) continuous in  $(n^h, n^l, \lambda_{DE_h}, \lambda_{DE_l}, \delta)$ , the solutions  $(n^h, n^l, \lambda_{DE_h}, \lambda_{DE_l})$  can be written as continuous functions of  $\delta$ . Thus, profits  $\Pi^h$  and  $\Pi^l$  are continuous in  $\delta$ .

The left-hand sides of the  $(DE_i)$  constraints are increasing in  $\delta$ ,<sup>14</sup> hence maximum enforceable effort increases in  $\delta$  as well.

For  $\delta \rightarrow 1$ ,  $(DE_i)$  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  $\delta$ , for every pair of effort levels  $(n^h, n^l)$  between zero and the respective first-best effort levels  $n_h^{FB}$  and  $n_l^{FB}$ , 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  $\delta$ . 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  $\delta$  is strictly positive and, since  $\delta \leq \bar{\delta} < 1$ , bounded. Therefore  $\hat{n}^h$  is a continuous function of  $\delta$ , and thus, by continuity of  $(DE_l)$  in  $(n^h, \delta)$ , 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

<sup>14</sup>This can be shown formally by an argument analogous to the one underlying the proof of Lemma 9.



$\delta \leq \underline{\delta}$ , both (DE) constraints bind, and hence  $n^h = n^l$ . ■

## B Auxiliary Results and Proofs for Section 4

### B.1 Preliminaries for the iid Model

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 3** *For any history  $\theta^t$ , the (DEh) constraint can be omitted.*

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

The following lemma will be useful to derive some characteristics of an optimal equilibrium.

**Lemma 4** *For any history  $\theta^t$  with  $n^h(\theta^t) \neq n^l(\theta^t)$ , (TTh) and (TTl) are not both binding.*

**Proof.** Assume there is a history  $\theta^\tau$  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  $g(n^l(\theta^\tau)) (\theta^h - \theta^l) = g(n^h(\theta^\tau)) (\theta^h - \theta^l)$ . Since  $\theta^h - \theta^l > 0$  and  $g$  is strictly increasing, this contradicts the claim that both constraints bind for  $n^h(\theta^\tau) \neq n^l(\theta^\tau)$ . ■

Now, we can substantially simplify the problem by establishing some structural properties of an optimal equilibrium.

**Lemma 5** *There exists an optimal equilibrium with the properties that, for every history  $\theta^t$ ,*

- $U(\theta^t) = 0$ ,
- $\Pi^h(\theta^t) \geq \tilde{\Pi}^l(\theta^t)$ ,
- $b^h(\theta^t) \geq b^l(\theta^t)$ ,
- *the (TTl) constraint can be omitted,*

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

**Proof.** We start with proving the first two parts. Suppose to the contrary that there exists a history  $\theta^t$  of length  $t \geq 1$  and an equilibrium such that, following history  $\theta^t$ , 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  $\theta^t$  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) and (IR) constraints as well as all of the principal's constraints at history  $\theta^t$  and all predecessor histories unaffected. Furthermore, the principal's profits at history  $\theta^t$  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  $(\theta^t, \theta^h)$  by play after  $(\theta^t, \theta^l)$ . This leads to on-path profits of  $\hat{\Pi}^h(\theta^t) = \tilde{\Pi}^l(\theta^t)$ . Set  $b_{new}^h(\theta^t) = b_{new}^l(\theta^t) = n(\theta^t)c$ , while increasing  $w(\theta^t)$  by  $\delta q \left( \hat{\Pi}^h(\theta^t) - \Pi^h(\theta^t) \right) + q \left( b_{old}^h(\theta^t) - b_{new}^h(\theta^t) \right) + (1 - q) \left( b_{old}^l(\theta^t) - b_{new}^l(\theta^t) \right)$ . (By Step 1. and the fact that (IC) at history  $\theta^t$  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  $\theta^t$  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  $\theta^t$  continues to hold. As the proof of Lemma 6 shows, the fact that (DEl) and (TTh) previously held at history  $\theta^t$ , together with Step 1, implies

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

As  $\Pi^h(\theta^t) < \tilde{\Pi}^l(\theta^t)$ , 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  $\theta^t$  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  $(\theta^t, \theta^l)$  by play after  $(\theta^t, \theta^h)$ . This leads to on-path profits of  $\hat{\Pi}^l(\theta^t) = \tilde{\Pi}^h(\theta^t)$ . Set  $b_{new}^h(\theta^t) = b_{new}^l(\theta^t) = n(\theta^t)c$ , while increasing  $w(\theta^t)$  by  $\delta(1-q) \left( \hat{\Pi}^l(\theta^t) - \Pi^l(\theta^t) \right) + q \left( b_{old}^h(\theta^t) - b_{new}^h(\theta^t) \right) + (1-q) \left( b_{old}^l(\theta^t) - b_{new}^l(\theta^t) \right)$ . (TTh), (TTl) and (IC) at history  $\theta^t$  now hold with equality. Previous constraints remain unchanged, with the exception of previous (IC) and (IR) constraints, which are relaxed. It remains to be shown that (DEL)-constraint at history  $\theta^t$  continues to hold. As the proof of Lemma 3 shows, the fact that (DEL) and (TTh) previously held at history  $\theta^t$ , 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

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

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  $t-1, t-2, \dots$ . 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  $\theta^t$ , assume to the contrary that there exists a history  $\theta^t$  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  $\theta^t$ . (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 and the (IC) constraint will bind. If  $n^h(\theta^t) \leq n^l(\theta^t)$ , it immediately follows from the fact that  $b^h(\theta^t) \geq b^l(\theta^t)$  and  $\Pi^l(\theta^t) \geq \tilde{\Pi}^h(\theta^t)$  that (TTl) can be omitted. So suppose that  $n^h(\theta^t) > n^l(\theta^t)$ , and suppose that the (TTl) constraint binds. By Lemma 4, 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  $q\varepsilon$ . This leaves

all previous constraints and profits unaffected yet relaxes the current (IC) and (TTI) constraints (while tightening the current (TTh) constraint and leaving the current (DEL) constraint unaffected). Now suppose that the (IC) constraint is slack. If  $b^l(\theta^t) > 0$ , we can decrease  $b^h(\theta^t) > 0$  and  $b^l(\theta^t) > 0$  by some  $\varepsilon > 0$ , while increasing  $w(\theta^t)$  by  $\varepsilon$ . This leaves all previous profits as well as all previous and current constraints unaffected, with the exception of the current (DEL)-constraint, which is relaxed. If now  $b^l(\theta^t) = 0$  and the (IC) and (TTI) constraints are slack, we can decrease  $b^h(\theta^t)$  by some  $\varepsilon > 0$ , while increasing  $w(\theta^t)$  by  $\frac{\varepsilon}{q}$ . This leaves all previous constraints and profits unaffected, yet relaxes the current (TTh) constraint (while tightening the current (TTI) and (IC) constraints and leaving the current (DEL) constraint unaffected). If  $b^l(\theta^t) = 0$  and the (TTI) constraint binds, we can replace play after  $(\theta^t, \theta^l)$  by play after  $(\theta^t, \theta^h)$  while setting  $b_{new}^h(\theta^t) = b_{new}^l(\theta^t) = n(\theta^t)c$  and increasing  $w(\theta^t)$  by  $(1 - q)(\Pi_{new}^l(\theta^t) - \Pi_{old}^l(\theta^t)) + qb_{old}^h(\theta^t) - n(\theta^t)c$ . As  $\Pi_{new}^l(\theta^t) = \tilde{\Pi}^h(\theta^t) \geq \tilde{\Pi}^h(\theta^t) - \frac{b_{old}^h(\theta^t)}{\delta} = \Pi_{old}^l(\theta^t)$ , and  $b_{old}^h(\theta^t) \geq n(\theta^t)c$  by the (IC) constraint, the increase in  $w(\theta^t)$  is positive. Therefore, previous (IC) and (IR) constraints are relaxed while all other previous constraints remain unaffected by our change. Furthermore, the current (TTh), (TTI) and (IC) constraints all hold with equality by construction. It remains to show that the current (DEL) constraint continues to hold, i.e. that  $-n(\theta^t)c + \delta\Pi_{new}^l(\theta^t) = -n(\theta^t)c + \delta\tilde{\Pi}^h(\theta^t) \geq 0$ . Yet, the binding (TTI) implies that  $\delta\Pi_{old}^l(\theta^t) = -b_{old}^h(\theta^t) + \delta\tilde{\Pi}^h(\theta^t) \geq 0$ , which implies that the current (DEL) constraint will hold after our change, as  $b_{old}^h(\theta^t) \geq \frac{n(\theta^t)c}{q} \geq n(\theta^t)c$  by the (IC) constraint.

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  $\theta^t$ . ■

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

**Lemma 6** *Maximum profits in the problem in which (TTh) and (DEL) are replaced by the following constraint (EC) equal maximum equilibrium profits:*

$$-n(\theta^t)c + \delta(q\Pi^h(\theta^t) + (1 - q)\Pi^l(\theta^t)) \geq \delta qg(n^l(\theta^t))(\theta^h - \theta^l). \quad (EC)$$

*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  $b^h(\theta^t) = \frac{1}{q}[n(\theta^t)c - \delta(1 - q)\Pi^l(\theta^t)] > \delta\Pi^l(\theta^t) = b^l(\theta^t)$  otherwise.*

**Proof.** By Lemma 5, we can without loss focus on equilibria in which

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

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

To prove that (EC) implies (TTh) and (DEl) given (B.1), assume that we are at an optimum satisfying the properties of Lemma 5 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 (B.1) holds, and that (DEl) and (TTh) are both satisfied. Toward this purpose, we set  $b^l(\theta^t) = \min \{ \delta \Pi^l(\theta^t), n(\theta^t)c \} \geq 0$ . 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  $\delta \Pi^h(\theta^t) \geq \delta g(n^l(\theta^t)) (\theta^h - \theta^l) + \delta \Pi^l(\theta^t)$ , which is implied by the second part of Lemma 5. Now suppose that  $n(\theta^t)c > \delta \Pi^l(\theta^t)$ . In this case, we set  $b^h(\theta^t) = \frac{1}{q} [n(\theta^t)c - \delta(1 - q)\Pi^l(\theta^t)] > 0$ . 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). ■

While effort dynamics in the case of public types are completely stationary (see Lemma 2), 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  $\theta^t := (\theta^h, \theta_2, \theta_3, \dots, \theta_t)$ , the function

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

which indicates the number of consecutive low periods immediately preceding period  $t$  along a given history  $\theta^t$ .

**Lemma 7** *There exists an optimal equilibrium with the property that, for every two histories  $\theta^t$  and  $\tilde{\theta}^t$ ,  $n^h(\theta^t) = n^h(\tilde{\theta}^t)$ . Furthermore, for every history  $\theta^t$ ,  $n^l(\theta^t) = n_{i(\theta^t)}^l$ , for some sequence of effort levels  $(n_j^l)_{j \in \mathbb{N}}$ .*

**Proof.** Consider an optimum satisfying the properties of Lemmas 5 and 6. Suppose that there exists a history  $\theta^t$  such that  $\Pi^h(\theta^t) < \max_{\hat{\theta}^\tau} \Pi^h(\hat{\theta}^\tau)$ . Replace the continuation play following  $(\theta^t, \theta^h)$  by the continuation play following  $(\tilde{\theta}, \theta^h)$ , where  $\tilde{\theta} \in \arg \max_{\hat{\theta}^\tau} \Pi^h(\hat{\theta}^\tau)$ . 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  $\Pi^h(\theta^t) = \bar{\Pi}^h$  for all  $\theta^t$  (if two different continuation plays lead to  $\arg\max_{\hat{\theta}^r} \Pi^h(\hat{\theta}^r)$ , we select one to be played after all histories  $(\theta^t, \theta^h)$ ). Therefore, there exists an optimum in which for any history  $\theta^t$ ,  $n^h(\theta^t) = \bar{n}^h$  and  $n^l(\theta^t) = n_{i(\theta^t)}^l$ . ■

In the following, we shall write  $n^h := n(\theta^t)$  for all  $\theta^t$  such that  $\theta_t = \theta^h$ ; we shall write  $n_i^l = n_{i(\theta^t)}^l = n(\theta^{t+1})$  for all  $\theta^{t+1} = (\theta^t, \theta^l)$ . By the same token, we shall write  $\Pi^h$  and  $\Pi_i^l$  for the corresponding optimal profits. These results allow us to restate our problem as in Section 4 in the main text.

The following two lemmata 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  $\delta$ .

**Lemma 8** *There exists an optimal equilibrium with the property that  $n_i^l \leq n_i^{FB}$  and  $n^h \leq n_h^{FB}$ .*

**Proof.** Consider an optimum satisfying the properties of Lemmas 5, 6 and 7. Suppose there exists a history  $\theta^t$  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. ■

**Lemma 9** *For every history  $\theta^t$ , maximal profits  $\Pi(\theta^t)$  are weakly increasing in  $\delta$ . Furthermore, a higher  $\delta$  relaxes (EC) constraints.*

**Proof.** Consider a given discount factor  $\hat{\delta}$  and the associated sequence of optimal actions  $(n^h(\hat{\delta}), n_i^l(\hat{\delta}))_{i \in \mathbb{N}}$ . We first show that a higher  $\delta$  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  $-n^h c + \delta q [\Pi^h - g(n_0^l) (\theta^h - \theta^l)] + \delta(1-q)\Pi_0^l \geq 0$ . In Lemma 5 we showed that, at our optimum,  $\Pi^h(\theta^t) \geq \Pi^l(\theta^t) + g(n^l(\theta^t)) (\theta^h - \theta^l)$  for all histories  $\theta^t$ . 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 8, these actions lead to (weakly) higher profits. The argument for the induction step is analogous. ■

## B.2 Proofs of Propositions 3–5

### Proof of Proposition 3

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

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

The left-hand side can be bounded from below by

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

Since  $\theta^h g(n_h^{FB}) - n_h^{FB}c > 0$  by assumption and because  $g(n_i^{FB})$  is finite, this expression diverges to infinity as  $\delta \rightarrow 1$ . Since, by Lemma 9, (EC) constraints are relaxed by larger values of  $\delta$ , the claim follows.  $\blacksquare$

### Proof of Proposition 4

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

$$-n_h^{FB}c + \bar{\delta} (q\Pi^{h,FB} + (1-q)\Pi^{l,FB}) = \bar{\delta} qg(n_i^{FB}) (\theta^h - \theta^l).$$

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

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

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

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

By our assumptions on  $g$ , the objective function and the constraint are twice continuously differentiable in the choice variables  $(n^h, n_i^l)_{i \in \mathbb{N}}$ . 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.<sup>15</sup>

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

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

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

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

Furthermore, optimality requires  $\frac{\partial \mathcal{L}}{\partial n_i^l} = 0$ , 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  $n_i^{FB}$  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

<sup>15</sup>In this case, one can show that a global optimum exists and that it entails  $n^h \in (0, n_h^{FB})$  by substituting the binding (ECh) constraint into the objective. Indeed, considering  $n_0^l$  as 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  $n_0^l \leq n_1^{FB}$  by Lemma 8, decreasing at  $n^h = n_h^{FB}$ . 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.



function of  $\delta$ . Moreover, define  $\tilde{n}^l(\delta)$  by  $g'(\tilde{n}^l(\delta)) = c(1-q)\frac{1-\delta(1-q)}{1-\delta} \left[ \frac{1-\delta(1-q(1-q))}{1-\delta}\theta^l - q\theta^h \right]^{-1}$  and  $\underline{n}^l(\delta)$  by

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

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  $\delta$ . Clearly, the solution  $(\hat{n}^h, \hat{n}_0^l)(\delta)$  to the problem in which only (ECh) is imposed entails  $(\hat{n}^h, \hat{n}_0^l)(\delta) \in \mathcal{I}$ , where  $\mathcal{I} := [\underline{n}^h(\delta), n_h^{FB}] \times [\underline{n}^l(\delta), n_l^{FB}]$ .<sup>16</sup> 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  $\delta$  is bounded. Hence, it is feasible to have a policy  $(\hat{n}^h, \hat{n}_0^l)$  that is continuous in  $\delta$ , implying that the optimal profits  $\hat{\Pi}^h$  in this problem are a continuous function of  $\delta$ . As  $(n^h, n_0^l)$  impacts the (ECli) constraints only via the profits  $\Pi^h$ , and since these constraints are continuous in  $\Pi^h$ , all (ECli) constraints hold for the solutions of this reduced problem in a neighborhood of  $\bar{\delta}$ .<sup>17</sup> By the argument underlying the proof of Lemma 9, the (ECh) constraint becomes tighter as the discount factor  $\delta$  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  $n_i^l = n_i^{FB}$ , 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  $n^h > n_i^{FB}$ . Suppose to the contrary that  $n^h \leq n_i^{FB}$ . Yet this solution is dominated by  $\hat{n}^h = \hat{n}_0^l = n_i^l = n_i^{FB}$ , which leads to higher profits and is feasible since all (ECli)-constraints (for  $i \geq 1$ ) hold for  $n_i^l = n_i^{FB}$  even for the initial  $n^h$  and  $n_0^l$ . ■

## Proof of Proposition 5

By definition of  $\underline{\delta}$ , some ECli ( $i \geq 1$ ) will bind in some left-neighborhood of  $\underline{\delta}$ , while EC10 remains slack. In this neighborhood, the profit-maximizing  $n_i^l$  ( $i \geq 1$ ) are obtained by maximizing  $\Pi_1^l$ .

<sup>16</sup>One shows that  $\underline{n}^l < n_i^{FB}$  ( $\underline{n}^h < n_h^{FB}$ ) by showing that the partial derivative of (ECh) with respect to  $n_0^l$  ( $n^h$ ) is always strictly negative at  $n_0^l = n_i^{FB}$  ( $n^h = n_h^{FB}$ ).

<sup>17</sup>As the only exception, there is a direct impact of  $\hat{n}_0^l$  in (EC10). Yet, as  $\hat{n}_0^l \leq n_i^{FB}$ , (EC10) is slacker than the other (ECli) constraints, and thus continues to hold as well.

Thus, we maximize

$$\Pi_1^l = \sum_{\tau=1}^{\infty} (\delta(1-q))^{\tau-1} (\theta^\tau g(n_\tau^l) - n_\tau^l c) + \delta q \Pi^h \frac{1}{1 - \delta(1-q)}$$

subject to

$$-n_i^l c + \delta q \Pi^h (1 + \delta(1-q)) + \delta ((\theta^i - q\theta^h) g(n_{i+1}^l) - (1-q)n_{i+1}^l c) + \delta^2 (1-q)^2 \Pi_{i+2}^l \geq 0$$

for all  $i \geq 1$ . We proceed in several steps.

**Lemma 10** For any  $i \geq 1$ ,  $\Pi_1^l \geq \Pi_i^l$ .

**Proof.** Suppose to the contrary that  $\Pi_j^l > \Pi_1^l$ , for some  $j > 1$ . For all  $i \geq 1$ , replace  $n_i$  by  $n_{j+i-1}$ . (This operation is feasible because all (ECli) were satisfied by assumption.) Thus, our previous  $\Pi_1^l$  cannot solve our maximization problem. ■

**Lemma 11**  $n_1^l \geq n_i^l$  for all  $i \geq 1$ .

**Proof.** Suppose to the contrary that there is a  $j > 1$  with  $n_j^l > n_1^l$ . Replace  $n_j^l$  with  $n_1^l$  and the continuation play following  $n_j^l$  with the continuation play following  $n_1^l$ . This is clearly feasible and (weakly) profitable (as  $\Pi_j^l \leq \Pi_1^l$  by Lemma 10). ■

**Lemma 12** For all odd  $i \geq 1$ ,  $\Pi_i^l \geq \Pi_{i+2}^l$  and  $n_i^l \geq n_{i+2}^l$ .

**Proof.** We proceed by induction over  $i$ . That  $\Pi_3^l \leq \Pi_1^l$  follows from Lemma 10. For the induction step, suppose that  $\Pi_j^l \geq \Pi_{j+2}^l$ , for some odd integer  $j$ . We have to show that  $\Pi_{j+2}^l \geq \Pi_{j+4}^l$ . Suppose to the contrary that  $\Pi_{j+2}^l < \Pi_{j+4}^l$ . Now, for all  $i \geq j+2$ , replace  $n_i^l$  by  $n_{i+2}^l$ . This is feasible if  $n_{j+1}^l \leq n_{j+3}^l$ . Therefore, our operation increases  $\Pi_{j+2}^l$  and hence  $\Pi_1^l$ . If  $n_{j+1}^l > n_{j+3}^l$ , by contrast, we distinguish two cases: (1.) If  $\Pi_{j+1}^l \leq \Pi_{j+3}^l$ , we can replace  $n_i^l$  by  $n_{i+2}^l$  for all  $i \geq j+1$ . This replacement is feasible and weakly increases  $\Pi_1^l$ . (2.) If, however,  $\Pi_{j+1}^l > \Pi_{j+3}^l$ , we replace  $n_{i+2}^l$  by  $n_i^l$  for all  $i \geq j+1$ . This is feasible if  $n_j^l \geq n_{j+2}^l$ . If, however,  $n_j^l < n_{j+2}^l$ , we can replace  $n_{i+2}^l$  by  $n_i^l$  for all  $i \geq j$ . Because, by the induction hypothesis,  $\Pi_j^l \geq \Pi_{j+2}^l$ , this increases  $\Pi_1^l$ .

Suppose that  $n_j^l < n_{j+2}^l$  for some odd integer  $j$ . Replace all  $n_{i+2}^l$  by  $n_i^l$  for all  $i \geq j$ . This is clearly feasible and (weakly) profitable (as  $\Pi_j^l \geq \Pi_{j+2}^l$ ). ■

**Lemma 13** For all even  $i \geq 2$ ,  $\Pi_i^l \leq \Pi_{i+2}^l$  and  $n_i^l \leq n_{i+2}^l$ .

**Proof.** Suppose to the contrary that  $\Pi_{j+4}^l < \Pi_{j+2}^l$  for some even integer  $j$ . Then, we can replace all  $n_{i+2}^l$  by  $n_i^l$  for all  $i \geq j+2$ . This is feasible as  $n_{j+1}^l \geq n_{j+3}^l$  by Lemma 12. Suppose that  $n_j^l > n_{j+2}^l$  for some even integer  $j$ . Replace all  $n_i^l$  by  $n_{i+2}^l$  for all  $i \geq j$ . This is clearly feasible and (weakly) profitable (as  $\Pi_{j+2}^l \geq \Pi_j^l$ ). ■

**Lemma 14**  $n_i^l \neq n_{i+2}^l \Rightarrow n_j^l \neq n_{j+2}^l \forall j \leq i$ .

**Proof.** Suppose to the contrary that  $n_i^l \neq n_{i+2}^l$  but  $n_j^l = n_{j+2}^l$  for some integer  $j < i$ . Consider the biggest such integer  $j$ , i.e.,  $n_{j+1}^l \neq n_{j+3}^l$ . First, assume that  $j$  is even, i.e.,  $j+1$  is odd and, by Lemma 12,  $n_{j+1}^l > n_{j+3}^l$ . Replace all  $n_{i+2}^l$  by  $n_i^l$  for all  $i \geq j+1$ . This is feasible as  $n_j^l = n_{j+2}^l$  and (weakly) profitable (as  $\Pi_{j+1}^l \geq \Pi_{j+3}^l$ ). Second, assume that  $j$  is odd, i.e.,  $j+1$  is even and, by Lemma 13,  $n_{j+1}^l < n_{j+3}^l$ . Replace all  $n_i^l$  by  $n_{i+2}^l$  for all  $i \geq j+1$ . This is feasible as  $n_j^l = n_{j+2}^l$  and (weakly) profitable (as  $\Pi_{j+1}^l \leq \Pi_{j+3}^l$ ). ■

**Lemma 15**  $n_i^l = n_{i+2}^l \Rightarrow n_j^l = n_{j+2}^l \forall j \geq i$ .

**Proof.** Suppose to the contrary that  $n_i^l = n_{i+2}^l$  but  $n_j^l \neq n_{j+2}^l$  for some integer  $j > i$ . Consider the smallest such integer  $j$ , i.e.,  $n_{j-1}^l = n_{j+1}^l$ . First, assume that  $j-1$  is even, i.e.,  $j$  is odd and, by Lemma 12,  $n_j^l > n_{j+2}^l$ . Replace all  $n_{i+2}^l$  by  $n_i^l$  for all  $i \geq j$ . This is feasible as  $n_{j-1}^l = n_{j+1}^l$  and (weakly) profitable (as  $\Pi_j^l \geq \Pi_{j+2}^l$ ). Second, assume that  $j-1$  is odd, i.e.,  $j$  is even and, by Lemma 13,  $n_j^l < n_{j+2}^l$ . Replace all  $n_i^l$  by  $n_{i+2}^l$  for all  $i \geq j$ . This is feasible as  $n_{j-1}^l = n_{j+1}^l$  and (weakly) profitable (as  $\Pi_{j-1}^l \leq \Pi_{j+1}^l$ ). ■

**Lemma 16**  $n_1^l = n_2^l \Rightarrow n_i^l = n_1^l \forall i \geq 1$ .

**Proof.** By Lemma 13,  $n_1^l = n_2^l \Rightarrow n_j^l = n_1^l$  for all even  $j$ . Hence, by Lemma 15,  $n_i^l = n_{i+2}^l$  for all odd  $i \geq 3$ . Suppose to the contrary that  $n_1^l > n_3^l$ . Replace  $n_3^l$  with  $n_1^l$  and the continuation play following  $n_3^l$  with the continuation play following  $n_1^l$ . This is feasible and (weakly) profitable (as  $\Pi_j^l \leq \Pi_1^l$  by Lemma 10). ■

**Lemma 17** Assume there is one  $i$  for which the (Ecli) constraint is slack. Then, the (Ecli+1) constraint binds.

**Proof.** To the contrary, assume that the (Ecli+1) constraint is slack. Increase  $n_{i+1}^l$  by a small  $\varepsilon > 0$ . This is feasible and increases  $\Pi_1^l$ . ■

**Lemma 18** *Assume there is one odd  $i > 1$  for which the (Ecli) constraint is slack. Then,  $n_j^l = n_i^{FB} \forall j \geq 1$ .*

**Proof.** Suppose (Ecli) is slack for  $i$  odd, with  $i > 1$ . Then, there must exist an optimum with  $n_{i+j} = n_j \forall j \geq 1$ . This implies that  $\Pi_{i+1}^l = \Pi_1^l$ ,  $\Pi_{i+2}^l = \Pi_2^l$ , ...,  $\Pi_{2i-1}^l = \Pi_{i-1}^l$ . By Lemma 13,  $\Pi_2^l \leq \Pi_4^l \leq \dots \leq \Pi_{i+1}^l \leq \Pi_{i+3}^l \leq \dots$ . Since  $\Pi_{i+1}^l = \Pi_1^l \leq \Pi_{i+3}^l = \Pi_3^l \leq \Pi_{i+5}^l = \Pi_5^l \leq \dots \leq \Pi_{2i}^l = \Pi_i^l \leq \dots \leq \Pi_1^l$ ,  $\Pi_j^l = \Pi_1^l$  for all even  $j$ .

By Lemma 12, we have  $\Pi_1^l \geq \Pi_3^l \geq \dots \geq \Pi_i^l = \Pi_{2i}^l = \Pi_1^l$ , and hence  $\Pi_j^l = \Pi_1^l$  for all odd  $j$ . Thus,  $n_j^l = n_1^l$  for all  $j \geq 1$ . Therefore, the Lagrange parameters satisfy  $\lambda_j = \lambda_{j+1} = 0$  for all  $j$ , and  $n_1^l = n_i^{FB}$ . ■

Lemma 18 implies that, in our left-neighborhood of  $\delta$ , all odd-numbered constraints will bind, i.e. the Lagrange parameters satisfy  $\lambda_j > 0$  for all odd integers  $j$ .

**Lemma 19** *Assume there is one even  $i$  for which the (Ecli) constraint is slack. Then, the (Eclj) constraints are slack for any even  $j$ . Moreover,  $n_j^l = n_{j+2}^l = \dots = n_1^l$  for all odd  $j$ , and  $n_\nu^l = n_{\nu+2}^l = \dots = n_2^l$  for any even  $\nu$ .*

**Proof.** Suppose (Ecli) is slack for  $i$  even. Then, there must exist an optimum with  $n_{i+j} = n_j \forall j \geq 1$ . This implies that  $\Pi_{i+1}^l = \Pi_1^l$ ,  $\Pi_{i+2}^l = \Pi_2^l$ , ...,  $\Pi_{2i-1}^l = \Pi_{i-1}^l$ . By Lemma 13,  $\Pi_2^l \leq \Pi_4^l \leq \dots \leq \Pi_i^l \leq \Pi_{i+2}^l = \Pi_2^l$ , and hence  $\Pi_\nu = \Pi_2^l$  for all even  $\nu$ . It follows that (Ecl $\nu$ ) is slack for all even  $\nu$ . Thus,  $n_\nu^l = n_2^l$  for all even  $\nu$ .

By Lemma 12, we have  $\Pi_1^l \geq \Pi_3^l \geq \dots \geq \Pi_{i+1}^l = \Pi_1^l$ , and hence  $\Pi_j^l = \Pi_1^l$  for all odd  $j$ . Thus,  $n_j^l = n_1^l$  for all odd  $j$ . Therefore, the Lagrange parameters  $\lambda_\nu = \lambda_{\nu+2} = 0$  for all even  $\nu$ . ■

The previous lemmata imply that there are two possibilities for an optimum. Either, all even (Ecli) constraints are slack, in which case  $n_j^l = n_1^l$  for all odd  $j$  and  $n_\nu^l = n_2^l$  for all even  $\nu$ . Otherwise, all (Ecli) constraints will bind. In the following, we characterize effort levels  $n_i^l$  ( $i \geq 1$ ) for the latter possibility.

**Lemma 20** *Assume all (Ecli) constraints bind. Then, either  $n_1^l = n_3^l = n_5^l = \dots$  and  $n_2^l = n_4^l = n_6^l = \dots$ , or  $n_1^l > n_3^l > n_5^l > \dots$  and  $n_2^l < n_4^l < n_6^l < \dots$*

**Proof.** To the contrary, assume that  $n_{j+2}^l > n_j^l$  for  $j$  even, but that  $n_{j+3}^l = n_{j+1}^l$ . By Lemma 15, this implies that  $n_{j+2}^l = n_{j+4}^l = \dots$  and  $n_{j+3}^l = n_{j+5}^l = \dots$ , and in particular also that  $\Pi_{j+2}^l = \Pi_{j+4}^l$ . But then, (ECl $j$ ) can not bind, a contradiction. The same logic can be applied to show that  $n_{j+2}^l < n_j^l$  for  $j$  odd, but that  $n_{j+3}^l = n_{j+1}^l$ , is not feasible. ■

**Lemma 21**  $n_i^l > n_j^l \Rightarrow \Pi_i^l \geq \Pi_j^l$ .

**Proof.** Suppose to the contrary that there exist integers  $i$  and  $j$  such that  $n_i^l > n_j^l$  yet  $\Pi_i^l < \Pi_j^l$ . Then,

$$\begin{aligned} \Pi_j^l - \Pi_i^l &= (\theta^l g(n_j^l) - cn_j^l) + \delta q \Pi^h + \delta(1-q) \Pi_{j+1}^l - [(\theta^l g(n_i^l) - cn_i^l) + \delta q \Pi^h + \delta(1-q) \Pi_{i+1}^l] \\ &= [(\theta^l g(n_j^l) - cn_j^l) - (\theta^l g(n_i^l) - cn_i^l)] + \delta(1-q) (\Pi_{j+1}^l - \Pi_{i+1}^l) \geq 0. \end{aligned}$$

Because  $n_i^l > n_j^l$ ,  $[(\theta^l g(n_j^l) - cn_j^l) - (\theta^l g(n_i^l) - cn_i^l)] < 0$ . Therefore,  $\Pi_{j+1}^l - \Pi_{i+1}^l > 0$ . Hence, replacing the history  $n_i^l$  by  $n_j^l$  and the continuation play after  $n_i^l$  by the continuation play after  $n_j^l$  is feasible, and also strictly profitable. ■

**Lemma 22** Suppose all (ECl $i$ ) constraints bind. Then,  $\sup_{j \in \mathbb{N}} n_{2j}^l \leq \inf_{j \in \mathbb{N}} n_{2j-1}^l$ .

**Proof.** Suppose to the contrary that  $\sup_{j \in \mathbb{N}} n_{2j}^l > \inf_{j \in \mathbb{N}} n_{2j-1}^l$ . Then, by Lemmata 12 and 13, this implies  $\limsup_{j \in \mathbb{N}} n_{2j}^l > \liminf_{j \in \mathbb{N}} n_{2j-1}^l$ . Therefore, there exists an integer  $i$  such that  $n_{2i}^l > n_{2i-1}^l \geq n_{2i+1}^l$ . By Lemma 21, this implies that  $\Pi_{2i}^l > \Pi_{2i+1}^l$ . Yet, as all constraints (ECl $i$ ), and in particular (ECl $2i-2$ ), are binding,  $n_{2i}^l > n_{2i-1}^l$  and  $\Pi_{2i}^l > \Pi_{2i+1}^l$  implies that  $n_{2i-2}^l > n_{2i-1}^l$ , which, by Lemma 21, implies  $\Pi_{2i-2}^l > \Pi_{2i-1}^l$ . As furthermore  $\Pi_{2i}^l \geq \Pi_{2i-2}^l$  by Lemma 13 and all constraints, in particular (ECl $i-2$ ) and (ECl $i-3$ ), are binding, we can conclude that  $n_{2i-2}^l > n_{2i-3}^l$  and thus, by Lemma 21,  $\Pi_{2i-2}^l > \Pi_{2i-3}^l$ . Iterating this argument finally yields  $n_2^l > n_1^l$ , a contradiction to Lemma 11. ■

**Lemma 23**  $n_1^l = n_2^l \Leftrightarrow q\theta^h = \theta^l$ .

**Proof.** Recall that for  $\delta < \underline{\delta}$ , the values  $n_i^l$ ,  $i \geq 1$ , can be obtained by maximizing

$$\Pi_1^l = \sum_{\tau=1}^{\infty} (\delta(1-q))^{\tau-1} (\theta^l g(n_\tau^l) - n_\tau^l c) + \delta q \Pi^h \frac{1}{1-\delta(1-q)}, \text{ s.t. (ECl}i\text{) constraints for } i \geq 1, \text{ and treating } \Pi^h \text{ as a constant.}$$

The Lagrange function of



$$\begin{aligned}
& \frac{\delta q}{1-\delta(1-q)} \Pi^h \\
&= n_1^l c - \delta \left( (\theta^l - q\theta^h) g(n_2^l) - (1-q)n_2^l c \right) \\
& - \frac{(\delta(1-q))^2}{1-\delta(1-q)^2} \left[ (\theta^l g(n_1^l) - n_1^l c) + \delta(1-q) (\theta^l g(n_2^l) - n_2^l c) \right] \text{ into (EC12) gives}
\end{aligned}$$

$$(g(n_1^l) - g(n_2^l)) \delta (\theta^l - q\theta^h - q\delta(1-q) (\theta^h - \theta^l)) + c (n_1^l - n_2^l) \geq 0.$$

For  $q\theta^h = \theta^l$  and  $n_1^l = n_2^l$ , the left hand side equals zero, hence (EC12) is satisfied. ■

This concludes the proof of Proposition 5. ■

## C Timing – Details

### C.1 $\theta_t$ Revealed at Beginning of Period $t$

First, we analyze the case of public information. There, we consider a quasi-stationary equilibrium in the sense that bonus and effort are only a function of today's type. The wage might be a function of today's and yesterday's type, if it is used to provide incentives for yesterday's effort. We use left and right superscripts to describe wages (and profits) as functions of  $\theta_{t-1}$  (left) and  $\theta_t$  (right). For example, if the type in both periods is high, profit is  ${}^h\Pi^h$  and wages are  ${}^h w^h$ . Profits can thus be written as

$$\begin{aligned}
{}^h\Pi^h &= \theta^h g(n^h) - b^h - {}^h w^h + \delta \bar{\Pi}^h \\
{}^l\Pi^h &= \theta^h g(n^h) - b^h - {}^l w^h + \delta \bar{\Pi}^h \\
{}^h\Pi^l &= \theta^l g(n^l) - b^l - {}^h w^l + \delta \bar{\Pi}^l \\
{}^l\Pi^l &= \theta^l g(n^l) - b^l - {}^l w^l + \delta \bar{\Pi}^l,
\end{aligned}$$

with  $\bar{\Pi}^h = q {}^h\Pi^h + (1-q) {}^h\Pi^l$  and  $\bar{\Pi}^l = q {}^l\Pi^h + (1-q) {}^l\Pi^l$ . The agent's utilities are described accordingly.

We maximize  $\bar{\Pi}^h$ , subject to the following constraints:<sup>18</sup>

---

<sup>18</sup>Note that maximizing any other of the above profit streams would yield identical outcomes because the equilibrium – as we will see below – is now sequentially efficient.

$$-n^h c + b^h + \delta \bar{U}^h \geq 0 \quad (\text{ICh})$$

$$-n^l c + b^l + \delta \bar{U}^l \geq 0 \quad (\text{ICl})$$

$$-b^h + \delta \bar{\Pi}^h \geq 0 \quad (\text{DEh})$$

$$-b^l + \delta \bar{\Pi}^l \geq 0. \quad (\text{DEl})$$

First, we show that it is weakly optimal only to use the bonus to provide incentives, while setting wages equal to zero: If any fixed wages were strictly positive, a reduction accompanied by a corresponding increase of the respective bonus would leave all constraints unaffected (for example, if  ${}^h w^h > 0$ , reducing  ${}^h w^h$  by a small  $\varepsilon > 0$  and increasing  $b^h$  by  $\delta q \varepsilon$  has no effect on ICh and DEh) and not decrease profits. Furthermore, as in Lemma 2, we can show that it is feasible and optimal to set  $b^h = n^h c$  and  $b^l = n^l c$ . Then, the two remaining constraints are

$$\begin{aligned} -n^h c + \delta \frac{q (\theta^h g(n^h) - n^h c) + (1-q) (\theta^l g(n^l) - n^l c)}{(1-\delta)} &\geq 0 \\ -n^l c + \delta \frac{q (\theta^h g(n^h) - n^h c) + (1-q) (\theta^l g(n^l) - n^l c)}{(1-\delta)} &\geq 0, \end{aligned}$$

which are the same as in our main setting with public information. Therefore, profit-maximizing effort levels are also characterized by Proposition 2, with levels of the discount factor,  $\bar{\delta}$  and  $\underline{\delta}$  ( $0 < \underline{\delta} < \bar{\delta} < 1$ ), such that  $n^h = n_h^{FB}$  and  $n^l = n_l^{FB}$  for  $\delta \geq \bar{\delta}$ ;  $n^l = n_l^{FB} < n^h < n_h^{FB}$  for  $\underline{\delta} < \delta < \bar{\delta}$ ; and  $n^h = n^l \leq n_l^{FB}$  for  $\delta \leq \underline{\delta}$ .



## Proof of Proposition 6

If the type is the principal's private information, additional truth-telling constraints, now imposed at the beginning of a period, must hold:

$$\theta^h g(n^h) - b^h - {}^h w^h + \delta \bar{\Pi}^h \geq \theta^h g(n^l) - b^l - {}^h w^l + \delta \bar{\Pi}^l \quad (\text{TThh})$$

$$\theta^h g(n^h) - b^h - {}^l w^h + \delta \bar{\Pi}^h \geq \theta^h g(n^l) - b^l - {}^l w^l + \delta \bar{\Pi}^l \quad (\text{TTlh})$$

$$\theta^l g(n^l) - b^l - {}^h w^l + \delta \bar{\Pi}^l \geq \theta^l g(n^h) - b^h - {}^h w^h + \delta \bar{\Pi}^h \quad (\text{TThl})$$

$$\theta^l g(n^l) - b^l - {}^l w^l + \delta \bar{\Pi}^l \geq \theta^l g(n^h) - b^h - {}^l w^h + \delta \bar{\Pi}^h \quad (\text{TTll})$$

To show that these constraints can be omitted, we plug the results from the case with public information,  ${}^h w^h = {}^h w^l = {}^l w^h = {}^l w^l = 0$  and  $b^h = n^h c$  and  $b^l = n^l c$ , into the conditions. Then,  $\bar{\Pi}^h = \bar{\Pi}^l = \frac{q(\theta^h g(n^h) - n^h c) + (1-q)(\theta^l g(n^l) - n^l c)}{(1-\delta)}$ , and the constraints become

$$\theta^h g(n^h) - n^h c \geq \theta^h g(n^l) - n^l c \quad (\text{TThh})$$

$$\theta^h g(n^h) - n^h c \geq \theta^h g(n^l) - n^l c \quad (\text{TTlh})$$

$$\theta^l g(n^l) - n^l c \geq \theta^l g(n^h) - n^h c \quad (\text{TThl})$$

$$\theta^l g(n^l) - n^l c \geq \theta^l g(n^h) - n^h c \quad (\text{TTll})$$

which are satisfied for the respective effort levels. ■

## C.2 $\theta_{t+1}$ Revealed at Beginning of Period $t$

### C.2.1 Public Types

We use superscripts to indicate equilibrium values as functions of this and next period's types. For example,  $n^{hh}$  is equilibrium effort in case today's and tomorrow's types are high,  $n^{hl}$  is equilibrium effort if today's type is high and tomorrow's type is low, and so on. By standard arguments, with public information it is without loss to analyze equilibria where, after all histories, actions depend only on today's and tomorrow's types. In the following, we call these equilibria *quasi-stationary*.

Then, on-path profit streams can take one of the four values

$$\begin{aligned}
\Pi^{hh} &= \theta^h g(n^{hh}) - w^{hh} - b^{hh} + \delta \bar{\Pi}^h \\
\Pi^{hl} &= \theta^h g(n^{hl}) - w^{hl} - b^{hl} + \delta \bar{\Pi}^l \\
\Pi^{lh} &= \theta^l g(n^{lh}) - w^{lh} - b^{lh} + \delta \bar{\Pi}^h \\
\Pi^{ll} &= \theta^l g(n^{ll}) - w^{ll} - b^{ll} + \delta \bar{\Pi}^l,
\end{aligned}$$

where  $\bar{\Pi}^h \equiv q\Pi^{hh} + (1-q)\Pi^{hl}$  and  $\bar{\Pi}^l \equiv q\Pi^{lh} + (1-q)\Pi^{ll}$ . The agent's utilities are defined equivalently. Bonus payments are bounded by dynamic enforcement constraints,

$$-b^{hh} + \delta \bar{\Pi}^h \geq 0 \quad (\text{DEhh})$$

$$-b^{hl} + \delta \bar{\Pi}^l \geq 0 \quad (\text{DEhl})$$

$$-b^{lh} + \delta \bar{\Pi}^h \geq 0 \quad (\text{DElh})$$

$$-b^{ll} + \delta \bar{\Pi}^l \geq 0, \quad (\text{DEll})$$

whereas effort levels are bounded by incentive compatibility constraints,

$$-n^{hh}c + b^{hh} + \delta \bar{U}^h \geq 0 \quad (\text{IChh})$$

$$-n^{hl}c + b^{hl} + \delta \bar{U}^l \geq 0 \quad (\text{IChl})$$

$$-n^{lh}c + b^{lh} + \delta \bar{U}^h \geq 0 \quad (\text{IClh})$$

$$-n^{ll}c + b^{ll} + \delta \bar{U}^l \geq 0. \quad (\text{ICll})$$

Now, although the bonus is a function of next period's type, it is certain at the time of the agent's effort choice. This is different from the main part of our paper, where next period's type is revealed immediately before today's bonus is paid, and therefore *uncertain* at the time of the agent's effort choice. Furthermore, for reasons similar to above (Lemma 2), it is feasible and weakly optimal to set  $w^{hh} = w^{hl} = w^{lh} = w^{ll} = 0$  and let (IC) constraints hold as equalities. Therefore,  $b^{hh} = n^{hh}c$ ,  $b^{hl} = n^{hl}c$ ,  $b^{lh} = n^{lh}c$  and  $b^{ll} = n^{ll}c$ .

Again, our objective is to maximize  $\bar{\Pi}^h$ , now subject to

$$-n^{hh}c + \delta\bar{\Pi}^h \geq 0 \quad (\text{DEhh})$$

$$-n^{hl}c + \delta\bar{\Pi}^l \geq 0 \quad (\text{DEhl})$$

$$-n^{lh}c + \delta\bar{\Pi}^h \geq 0 \quad (\text{DElh})$$

$$-n^{ll}c + \delta\bar{\Pi}^l \geq 0. \quad (\text{DEll})$$

## Proof of Lemma 1

It is immediate that  $\bar{\Pi}^h \geq \bar{\Pi}^l$ , i.e. that a high type is associated with higher profits. Therefore,  $\theta_{t+1} = \theta^h$  allows for a credible promise of a higher bonus, and therefore for the implementation of a higher effort level, in period  $t$ . The *desired* effort levels if today's type is high ( $n^{hh}$  and  $n^{hl}$ ) are also larger than if today's type is low. If the discount factor is sufficiently close to 1, none of the constraints bind and first-best levels  $n^{hh} = n^{hl} = n_h^{FB}$  and  $n^{ll} = n^{lh} = n_l^{FB}$  can be implemented. For a lower discount factor, (DEhl) will eventually bind, and  $n^{hl} < n^{hh} = n_h^{FB}$ . For even lower discount factors, (DEhh) and/or (DEll) will at some point bind as well. This yields the result. ■

### C.2.2 Private Types

With private types, we keep the notation from our analysis with public types (proof of Lemma 1). Though this restriction is not without loss of generality here, we continue to focus on the same kind of quasi-stationary equilibria as with public types (that is, actions depend only on today's and tomorrow's types), where fixed wages equal zero and (IC) constraints bind. We will show below that, in contrast to before, the relevant truth-telling constraints can now either be satisfied by a reduction of effort levels, or by an ex-ante payment made to the agent. If these payments can be extracted by the principal at the beginning of the game, such an agreement would indeed maximize the principal's profits.

Now, two types of truth-telling constraints arise. First, the principal might misreport her type and then proceed with play as prescribed by equi-

librium (like in our main case). This yields the constraints

$$\Pi^{hh} \geq \tilde{\Pi}^{hl} \quad (\text{TThh})$$

$$\Pi^{hl} \geq \tilde{\Pi}^{hh} \quad (\text{TThl})$$

$$\Pi^{lh} \geq \tilde{\Pi}^{ll} \quad (\text{TTlh})$$

$$\Pi^{ll} \geq \tilde{\Pi}^{lh}, \quad (\text{TTll})$$

where

$$\begin{aligned} \tilde{\Pi}^{hh} &= \theta^h g(n^{hh}) - n^{hh}c + \delta \left[ q \left( \theta^l g(n^{hh}) - n^{hh}c + \delta \bar{\Pi}^h \right) + (1-q) \left( \theta^l g(n^{hl}) - n^{hl}c + \delta \bar{\Pi}^l \right) \right] \\ &= \Pi^{hh} - \delta (\theta^h - \theta^l) [qg(n^{hh}) + (1-q)g(n^{hl})] \end{aligned}$$

are the principal's profits in case today's type is high and tomorrow's type is low, but where she falsely reports tomorrow's type to be high.

The respective values  $\tilde{\Pi}^{hl}$ ,  $\tilde{\Pi}^{lh}$  and  $\tilde{\Pi}^{ll}$  are obtained in similar fashion. The second kind of truth-telling constraints prevent the principal from misreporting her type and subsequently shutting down.

These constraints are

$$\theta^h g(n^{hh}) - n^{hh}c + \delta \bar{\Pi}^h \geq \theta^h g(n^{hl}) \quad (\text{TThh2})$$

$$\theta^h g(n^{hl}) - n^{hl}c + \delta \bar{\Pi}^l \geq \theta^h g(n^{hh}) \quad (\text{TThl2})$$

$$\theta^l g(n^{lh}) - n^{lh}c + \delta \bar{\Pi}^h \geq \theta^l g(n^{ll}) \quad (\text{TTlh2})$$

$$\theta^l g(n^{ll}) - n^{ll}c + \delta \bar{\Pi}^l \geq \theta^l g(n^{lh}) \quad (\text{TTll2})$$

Note that these kinds of constraints are not needed in our main case. There, next period's type is revealed after today's effort and output have been realized. They are thus sunk when the principal's announces next period's type. Therefore, these constraints coincide with the respective dynamic enforcement constraints.

Finally, (DE) constraints as specified in the proof to Lemma 1, the case with public information, must hold.

## Proof of Proposition 7

First, we show that  $n^{hh} = n^{hl} \equiv n^h$  and  $n^{lh} = n^{ll} \equiv n^l$ . To do so, we omit (TT) constraints and solve the problem only subject to (TT2) and (DE) constraints. Then, we show that the solution to this relaxed problem also satisfies (TT) constraints.

The reduced problem maximizes  $\bar{\Pi}^h$ , subject to

$$-n^{hh}c + \delta\bar{\Pi}^h \geq 0 \quad (\text{DEhh})$$

$$-n^{hl}c + \delta\bar{\Pi}^l \geq 0 \quad (\text{DEhl})$$

$$-n^{lh}c + \delta\bar{\Pi}^h \geq 0 \quad (\text{DElh})$$

$$-n^{ll}c + \delta\bar{\Pi}^l \geq 0. \quad (\text{DEll})$$

$$\theta^h g(n^{hh}) - n^{hh}c + \delta\bar{\Pi}^h \geq \theta^h g(n^{hl}) \quad (\text{TThh2})$$

$$\theta^h g(n^{hl}) - n^{hl}c + \delta\bar{\Pi}^l \geq \theta^h g(n^{hh}) \quad (\text{TThl2})$$

$$\theta^l g(n^{lh}) - n^{lh}c + \delta\bar{\Pi}^h \geq \theta^l g(n^{ll}) \quad (\text{TTlh2})$$

$$\theta^l g(n^{ll}) - n^{ll}c + \delta\bar{\Pi}^l \geq \theta^l g(n^{lh}) \quad (\text{TTll2})$$

Note that effort is never above the respective first-best effort level. Now, assume to the contrary that  $n^{hh} > n^{hl}$ . If (DEhl) binds, plugging  $-n^{hl}c + \delta\bar{\Pi}^l = 0$  into (TThl2) yields  $\theta^h g(n^{hl}) \geq \theta^h g(n^{hh})$  which is violated for  $n^{hh} > n^{hl}$ . If (DEhl) does not bind, increase  $n^{hl}$  by a small  $\varepsilon > 0$ . This operation increases  $\bar{\Pi}^h$ , relaxes (TThl2), and does not violate (TThh2), (DEhl) or any other constraint. Continue until either  $n^{hh} = n^{hl}$  or (DEhl) binds. In the latter case, recall that (TThl2) is violated for a binding (DEhl) constraint and  $n^{hh} > n^{hl}$ .

Next, assume  $n^{hh} < n^{hl}$ . If (DEhh) binds, plugging  $-n^{hh}c + \delta\bar{\Pi}^h = 0$  into (TThh2) yields  $\theta^h g(n^{hh}) \geq \theta^h g(n^{hl})$  which is violated for  $n^{hh} < n^{hl}$ . If (DEhh) does not bind, increase  $n^{hh}$  by a small  $\varepsilon > 0$ . This operation increases  $\bar{\Pi}^h$ , relaxes (TThh2), and does not violate (TThl2) and (DEhh) or any other constraint. Continue until either  $n^{hh} = n^{hl}$  or (DEhh) binds. In the latter case, recall that (TThh2) is violated for a binding (DEhh) constraint and  $n^{hh} < n^{hl}$ .

Thus, we have shown that  $n^{hh} = n^{hl} \equiv n^h$  in this reduced problem.

Accordingly, it can be shown that  $n^{lh} = n^{ll} \equiv n^l$ . Taking this into account, the remaining constraints in the reduced problem are

$$-n^h c + \delta \bar{\Pi}^h \geq 0 \quad (\text{DEhh})$$

$$-n^h c + \delta \bar{\Pi}^l \geq 0 \quad (\text{DEhl})$$

$$-n^l c + \delta \bar{\Pi}^h \geq 0 \quad (\text{DElh})$$

$$-n^l c + \delta \bar{\Pi}^l \geq 0. \quad (\text{DEll})$$

Together with  $\bar{\Pi}^h = (\theta^h g(n^h) - n^h c) + \delta q \bar{\Pi}^h + \delta(1-q)\bar{\Pi}^l$  and  $\bar{\Pi}^l = (\theta^l g(n^l) - n^l c) + \delta q \bar{\Pi}^h + \delta(1-q)\bar{\Pi}^l$ , this implies that  $n^h \geq n^l$  and  $\bar{\Pi}^h > \bar{\Pi}^l$ , which allows us to omit (DEhh) and (DElh), and leaves the remaining constraints

$$-n^h c + \delta \bar{\Pi}^l \geq 0 \quad (\text{DEh})$$

$$-n^l c + \delta \bar{\Pi}^l \geq 0. \quad (\text{DEl})$$

Therefore, effort levels to this constrained maximization problem are given by discount factors,  $\bar{\delta}$  and  $\underline{\delta}$ , with  $0 < \underline{\delta} < \bar{\delta} < 1$ , with

- $n^h = n_h^{FB}$  and  $n^l = n_l^{FB}$  for  $\delta \geq \bar{\delta}$
- $n^l = n_l^{FB} < n^h < n_h^{FB}$  for  $\underline{\delta} < \delta < \bar{\delta}$
- $n^l = n^h \leq n_l^{FB}$  for  $\delta \leq \underline{\delta}$

To complete the proof, we have to show that these effort levels do not violate any of the (TT) constraints. These amount to

$$\begin{aligned} & \left( \theta^h g(n^{hh}) - n^{hh} c + \delta \bar{\Pi}^h \right) - \left( \theta^h g(n^{hl}) - n^{hl} c + \delta \bar{\Pi}^l \right) \\ & \quad - \delta (\theta^h - \theta^l) [qg(n^{lh}) + (1-q)g(n^{ll})] \geq 0 \quad (\text{TThh}) \end{aligned}$$

$$\begin{aligned} & - \left[ \left( \theta^h g(n^{hh}) - n^{hh} c + \delta \bar{\Pi}^h \right) - \left( \theta^h g(n^{hl}) - n^{hl} c + \delta \bar{\Pi}^l \right) \right] \\ & \quad + \delta (\theta^h - \theta^l) [qg(n^{hh}) + (1-q)g(n^{hl})] \geq 0 \quad (\text{TThl}) \end{aligned}$$

$$\begin{aligned} & \left( \theta^l g(n^{lh}) - n^{lh} c + \delta \bar{\Pi}^h \right) - \left( \theta^l g(n^{ll}) - n^{ll} c + \delta \bar{\Pi}^l \right) \\ & \quad - \delta (\theta^h - \theta^l) [qg(n^{lh}) + (1-q)g(n^{ll})] \geq 0 \quad (\text{TThh}) \end{aligned}$$

$$\begin{aligned} & - \left[ \left( \theta^l g(n^{lh}) - n^{lh} c + \delta \bar{\Pi}^h \right) - \left( \theta^l g(n^{ll}) - n^{ll} c + \delta \bar{\Pi}^l \right) \right] \\ & \quad + \delta (\theta^h - \theta^l) [qg(n^{hh}) + (1-q)g(n^{hl})] \geq 0. \quad (\text{TTll}) \end{aligned}$$

Plugging  $n^{hh} = n^{hl} = n^h$  and  $n^{lh} = n^{ll} = n^l$  into the (TT) constraints, rearranging and making use of  $\bar{\Pi}^h - \bar{\Pi}^l = (\theta^h g(n^h) - n^h c) - (\theta^l g(n^l) - n^l c)$ , yields an equivalence of (TThh) and (TTlh), as well as of (TThl) and (TTll). Therefore, the remaining (TT) constraints are

$$\delta [(\theta^h g(n^h) - n^h c) - (\theta^h g(n^l) - n^l c)] \geq 0 \quad (\text{TThh})$$

$$\delta [(\theta^l g(n^h) - n^h c) - (\theta^l g(n^l) - n^l c)] \leq 0 \quad (\text{TThl})$$

For  $\delta > \underline{\delta}$ , these conditions hold strictly (the latter because  $n_i^{FB}$  maximizes  $\theta^l g(n^l) - n^l c$ ), for  $\delta \leq \underline{\delta}$  and hence  $n^l = n^h$  they hold as equalities. ■

## D Persistent Types – Details

### D.1 Permanent Shocks

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.

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  $w^h$  ( $w_i^l$ ) and  $n^h$  ( $n_i^l$ ) are the wages and effort levels, respectively, if we are in a high-type period (the  $i+1$ -st subsequent low period).

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

$$-b^h + \delta\Pi^h \geq 0 \quad (\text{DEh})$$

$$-b_i^l + \delta\Pi_i^l \geq 0 \forall i \geq 0. \quad (\text{DEli})$$

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

$$-b^h + \delta\Pi^h \geq -b_0^l + \delta\tilde{\Pi}_0^l \quad (\text{TTh})$$

$$-b_0^l + \delta\Pi_0^l \geq -b^h + \delta\tilde{\Pi}^h. \quad (\text{TTl})$$

After claiming that the state has switched to  $\theta^l$ , the principal has to claim a state of  $\theta^l$  in all subsequent periods. This gives us

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

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

Suppose that the state switches to  $\theta^l$  in period  $t$ . If it is a profitable deviation for the principal falsely to claim that the state has remained  $\theta^h$  in period  $t$ , then, by the One-Deviation principle (Hendon, Jacobsen, and Sloth (1996)), it is a profitable deviation for him to revert to telling the truth by claiming  $\theta_{t+1} = \theta^l$ . Yet, if this single deviation is profitable, then it is also profitable to continue lying in period  $t+1$ . In this case, it will thus also be profitable to continue lying with probability  $q$  while telling the truth with the counter-probability. This gives

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

Finally, the (IC) and (IR) constraints are as in the main model (Section 4).



Lemma 3 and the first three properties of Lemma 5 go through essentially unchanged. Moreover, an equivalent result to Lemma 4 is immediate: (TTh) and (TTI) cannot both bind in the same period if  $\sum_{\tau=0}^{\infty} (\delta q)^\tau g(n_\tau^l) \neq \frac{g(n^h)}{1-\delta q}$ . Furthermore, we can show that there exists an optimal equilibrium with the property  $n(\theta^t) \leq n^{FB}(\theta_t)$ , i.e., a result equivalent to Lemma 8 obtains. Indeed, suppose to the contrary that  $n(\theta^t) > n^{FB}(\theta_t)$  for some history  $\theta^t$ . Then, reduce  $n(\theta^t)$  by some  $\varepsilon > 0$ , while increasing  $w(\theta^t)$  in such a way as to leave the principal's profits unchanged. This relaxes the (IR) and (IC) constraints at history  $\theta^t$  and at all predecessor histories, while all (DE) and (TT) constraints remain unchanged. Thus, reducing  $n(\theta^t)$  cannot do any harm.

We now prove

**Lemma 24** *The effort levels satisfy  $n^h \geq \sup_{i \in \mathbb{N}} n_i^l$ .*

**Proof.** Suppose to the contrary that a policy  $\sigma = (n^h, n_i^l)_{i \in \mathbb{N}}$  such that  $n^h < \sup_{i \in \mathbb{N}} n_i^l =: \bar{n}^l$  is optimal. Then, as  $n(\theta^t) \leq n^{FB}(\theta_t)$ , the policy  $\hat{\sigma} = (\hat{n}^h, \hat{n}_i^l)_{i \in \mathbb{N}}$  given by  $\hat{n}^h = \hat{n}_i^l = \bar{n}^l$ ,  $\hat{w}^h = \hat{w}_i^l = 0$ , and  $\hat{b}^h = \hat{b}_i^l = \bar{n}c$ , 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  $\hat{\Pi}^h$  ( $\Pi^h$ ) and  $\hat{\Pi}^l$  ( $\Pi_i^l$ ) are the profits associated with policy  $\hat{\sigma}$  ( $\sigma$ ), respectively. As policy  $\sigma$  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. Moreover, (TTh) and (TTI) hold with equality. This is a contradiction to policy  $\sigma$  being optimal. ■

Now, we prove that (TTI) can be omitted and (IC) will bind.

**Lemma 25** *There exists an optimal equilibrium with the properties that, for every history  $\theta^t$ ,*

- *the (TTI) constraint can be omitted,*
- *$n(\theta^t)c = qb^h(\theta^t) + (1-q)b^l(\theta^t)$  and  $w(\theta^t) = 0$ .*

**Proof.** Lemma 24 implies that, if both (TTh) and (TTI) bind,  $n^h = n_\tau^l = \bar{n}$  for all  $\tau \in \mathbb{N}$ . In this case,  $b^h = b_0^l \geq \bar{n}c$ .

Now, if there exists a  $\tau \in \mathbb{N}$  such that  $n^h > n_\tau^l$ , Lemma 24 implies that  $\frac{g(n^h)}{1-\delta q} > \sum_{\tau=0}^{\infty} (\delta q)^\tau g(n_\tau^l)$ . Suppose that (TTI) binds. As  $\frac{g(n^h)}{1-\delta q} > \sum_{\tau=0}^{\infty} (\delta q)^\tau g(n_\tau^l)$ , (TTh) is slack. We can therefore increase  $b^h$  by a small  $\varepsilon > 0$  while decreasing  $w^h$  by  $q\varepsilon$ . This leaves all constraints and profits unaffected yet relaxes the

(IC) and (TTI) constraints (while tightening the (TTh) constraint and leaving the (DEli) constraints unaffected). Now suppose that the (IC) constraint is slack. If  $b_0^l > 0$ , we can decrease  $b^h > 0$  and  $b_0^l > 0$  by some  $\varepsilon > 0$ , while increasing  $w^h$  by  $\varepsilon$ . This leaves profits as well as all constraints unaffected, with the exception of the (DEl0)-constraint, which is relaxed. If now  $b_0^l = 0$  and the (IC) and (TTI) constraints are slack, we can decrease  $b^h$  by some  $\varepsilon > 0$ , while increasing  $w^h$  by  $\frac{\varepsilon}{q}$ . This leaves all constraints and profits unaffected, yet relaxes the (TTh) constraint (while tightening the (TTI) and (IC) constraints and leaving the (DEl0) constraint unaffected). If  $b_0^l = 0$  and the (TTI) constraint binds, we can replace  $n_\tau^l$  by  $n^h$  for all  $\tau \in \mathbb{N}$  while setting  $b_{new}^h = b_{\tau,new}^l = n^h c$ . The (TTh), (TTI) and (IC) constraints all hold with equality by construction. It remains to show that the (DEli) constraints continue to hold, i.e. that  $-\bar{n}c + \delta \Pi_{i,new}^l = -\bar{n}c + \delta \tilde{\Pi}^h \geq 0$ . Yet, the binding (TTI) implies that  $\delta \Pi_{0,old}^l = -b_{old}^h(\theta^t) + \delta \tilde{\Pi}^h(\theta^t) \geq 0$ , which implies that the (DEli) constraints will hold after our change, as  $b_{old}^h \geq \frac{n^h c}{q} \geq n^h c$  by the (IC) constraint.

Because  $U^h = w^h - n^h c + q b^h + (1 - q) b_0^l = 0$ , a binding (IC) constraint implies that  $w^h = 0$ . By the same token,  $U_\tau^l = w_\tau^l - n_\tau^l c + b_{\tau+1}^l = 0$ , a binding (IC) constraint implies that  $w_\tau^l = 0$  for all  $\tau \in \mathbb{N}$ .  $\blacksquare$

Lemma 25 implies, inter alia, that  $b^h \geq b_0^l$ ,  $n^h c = q b^h + (1 - q) b_0^l$ ,  $n_i^l c = b_{i+1}^l$  and that (TTh) and (DEli) are the relevant constraints.

The proof of Lemma 6 goes through essentially unchanged. Therefore, we can equivalently replace (TTh) and (DEl0) by the following (ECh) constraint

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

We furthermore need to keep track of

$$-n_i^l c + \delta \Pi_{i+1}^l \geq 0 \quad (\text{DEli})$$

for all  $i \geq 0$ .

We are now ready to show the equivalent of Lemma 9.

**Lemma 26** *Maximal profits  $\Pi^h$  and  $\Pi_i^l$  ( $i \in \mathbb{N}$ ) are weakly increasing in  $\delta$ . Furthermore, a higher  $\delta$  relaxes the (ECh) and (DEli)-constraints.*

**Proof.** 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 (DE*l*) for  $\tilde{\delta}$ . By Lemma 8, these actions lead to weakly higher profits. It thus only remains to show that (ECh) is relaxed as  $\delta$  increases. For this, we compute the derivative  $\mathcal{D}$  of (ECh) with respect to  $\delta$ , 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

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

we have

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

Furthermore, as

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

we have

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

Inserting this gives us

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

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

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

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

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

By Lemma 24, we know that  $n^h \geq \bar{n}^l$ ; by Lemma 8, 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. ■

Finally, we are able to prove Proposition 8.

## Proof of Proposition 8

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

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

yielding first-order conditions

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

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

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

To show that  $n_i^l < n_{i+1}^l$ , rewrite conditions (D.3) for  $n_i^l$  and for  $n_{i+1}^l$  as

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

Dividing the first by the second equality yields the necessary condition

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

which becomes

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

The denominator of this expression must be different from zero:

$(\theta^l g'(n_{i+1}^l) - c) > 0$  because  $n_{i+1}^l < n_i^{FB}$ . 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

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

As  $q < 1$ , this implies  $\frac{(\theta^l g'(n_{i+1}^l) - c)}{(\theta^l g'(n_i^l) - c)} < \frac{g'(n_{i+1}^l)}{g'(n_i^l)}$ . This is equivalent to  $g'(n_i^l) > g'(n_{i+1}^l)$ , which yields  $n_{i+1}^l > n_i^l$  due to the strict concavity of  $g(\cdot)$ .

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

## D.2 Impermanent Shocks

Besides the  $q^h/q^l$  notation introduced in the main text, we shall also write  $q(\theta_t)$  for the probability of next period's type being high given the current-period type  $\theta_t$ . We focus on a subset of the parameter space for which our solution is qualitatively similar to our previous results, with overshooting and gradual recovery.

Here, the truth-telling and dynamic enforcement constraints amount to

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

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

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

$$-b^l(\theta^t) + \delta\Pi^l(\theta^t) \geq 0, \quad (\text{DEl})$$

with

$$\begin{aligned} \Pi^h(\theta^t) &= \theta^h g(n^h(\theta^t)) - w^h(\theta^t) \\ &\quad + q^h (-b^{hh}(\theta^t) + \delta\Pi^{hh}(\theta^t)) + (1 - q^h) (-b^{hl}(\theta^t) + \delta\Pi^{hl}(\theta^t)), \end{aligned}$$

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

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

and

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

Note that our formulations of  $\tilde{\Pi}^h(\theta^t)$  and  $\tilde{\Pi}^l(\theta^t)$  again make use of the One-deviation principle (see Hendon, Jacobsen, and Sloth (1996)).

Now, we maximize  $\Pi(\theta^1)$  subject to the (TTh), (DEl) and (IC) constraints (and omit the (DEh) and (TTl) constraints).

Combining (TTh) and (DEl) yields the following (EC) constraints, which are necessary (but may not be sufficient) for equilibrium:

$$-q^h b^h(\theta^t) - (1 - q^h) b^l(\theta^t) + \delta q^h \left( \Pi^h(\theta^t) - \tilde{\Pi}^l(\theta^t) \right) + \delta \Pi^l(\theta^t) \geq 0, \quad (\text{ECh})$$

$$-q^l b^h(\theta^t) - (1 - q^l) b^l(\theta^t) + \delta q^l \left( \Pi^h(\theta^t) - \tilde{\Pi}^l(\theta^t) \right) + \delta \Pi^l(\theta^t) \geq 0. \quad (\text{ECl})$$

It is straightforward to verify that, if  $\delta$  is large enough, first-best effort levels satisfy these constraints. In contrast to the case of iid shocks, (ECl) might bind for higher discount factors than (ECh) constraints. This is because autocorrelated shocks make not only first-best, but also implementable, effort a function of today's state of the world. We shall, however, focus on the case that (ECh) binds before (ECl) does, as we did for permanent shocks.

While (ECl) constraints can thus be omitted, (TTh) constraints (which constitute one part of (ECl) constraints) will bind for all subsequent histories. Indeed, suppose to the contrary that there exists a subsequent history,  $\hat{\theta}^{t+\tau}$ , such that (TTh) at  $\hat{\theta}^{t+\tau}$  is slack, and (EC) binds. Increase  $b^h(\hat{\theta}^{t+\tau})$  by some  $\varepsilon > 0$  and reduce  $w(\hat{\theta}^{t+\tau})$  by  $q(\hat{\theta}_{t+\tau})\varepsilon$ . This relaxes the (EC) constraint at history  $\theta^t$  and leaves all other (EC) constraints unaffected.

From this observation, it follows that we can plug binding (TTh) constraints into (EC), and rewrite the latter as

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

By the same argument as in the proof of Lemma 7, it follows that  $n^h(\theta^t)$  will be the same for all  $\theta^t$ . By the same token, low-type effort can be written as  $n_i^l$ , where the  $i$  indicates the number of consecutive low periods immediately preceding period  $t$  along a given history  $\theta^t$ .

Furthermore having (IC) constraints hold as equalities and using  $U(\theta^t) = 0$  for all  $\theta^t$ , we solve

$$\max_{n^h, n_i^l} \Pi^h = (1 - \delta(1 - q^l)) \frac{(\theta^h g(n^h) - n^h c) + \delta(1 - q^h) \sum_{i=0}^{\infty} (\delta(1 - q^l))^i (\theta^l g(n_i^l) - n_i^l c)}{(1 - \delta)(1 - \delta(q^h - q^l))}, \quad (\text{D.4})$$

subject to

$$-n^h c + \delta q^h \Pi^h + \delta(1 - q^h) \Pi_0^l \geq \delta q^h (\theta^h - \theta^l) \sum_{i=0}^{\infty} [\delta(q^h - q^l)]^i g(n_i^l) \quad (\text{ECh})$$

**Proposition 9** *The solution to the constrained maximization problem (D.4) has the following features: There exists a  $\delta^h < 1$  such that*

- for  $\delta \geq \delta^h$ ,  $n^h = n_h^{FB}$  and  $n_i^l = n_i^{FB}$  for all  $i$ ;
- for discount factors in some left neighborhood of  $\delta^h$ ,  $n^h < n_h^{FB}$ . Furthermore, for all  $i \in \mathbb{N}$ ,  $n_i^l < n_{i+1}^l < n_i^{FB}$ , with  $\lim_{i \rightarrow \infty} n_i^l = n_i^{FB}$ .

**Proof.** Note that

$$\begin{aligned} \Pi_0^l &= \sum_{i=0}^{\infty} (\delta(1 - q^l))^i (\theta^l g(n_i^l) - n_i^l c) \\ &+ \delta q^l \frac{(\theta^h g(n^h) - n^h c) + \delta(1 - q^h) \sum_{i=0}^{\infty} (\delta(1 - q^l))^i (\theta^l g(n_i^l) - n_i^l c)}{(1 - \delta)(1 - \delta(q^h - q^l))} \end{aligned}$$

and that the (ECh) constraint can be rewritten to

$$\begin{aligned} &-n^h c + \delta \Pi^h \frac{q^h - \delta(q^h - q^l)}{1 - \delta(1 - q^l)} + \delta(1 - q^h) \sum_{i=0}^{\infty} (\delta(1 - q^l))^i (\theta^l g(n_i^l) - n_i^l c) \\ &\geq \delta q^h (\theta^h - \theta^l) \sum_{i=0}^{\infty} [\delta(q^h - q^l)]^i g(n_i^l) \quad (\text{ECh}) \end{aligned}$$

Denoting by  $\lambda$  the Lagrange parameter associated with the (ECh) constraint, the Lagrange function equals



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

yielding first-order conditions

$$\frac{\partial \mathcal{L}}{\partial n^h} = (\theta^h g'(n^h) - c) \frac{(1 - \delta(1 - q^l))}{(1 - \delta)(1 - \delta(q^h - q^l))} \left( 1 + \delta \frac{q^h - \delta(q^h - q^l)}{1 - \delta(1 - q^l)} \lambda \right) - \lambda c = 0 \quad (\text{D.5})$$

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial n_i^l} &= \delta^{i+1} \left\{ (\theta^l g'(n_i^l) - c) \left( \frac{(1 - \delta(1 - q^l))(1 - q^h)}{(1 - \delta)(1 - \delta(q^h - q^l))} (1 - q^l)^i (1 + \lambda) \right) \right. \\
&\quad \left. - \lambda q^h (q^h - q^l)^i (\theta^h - \theta^l) g'(n_i^l) \right\} = 0 \quad (\text{D.6})
\end{aligned}$$

The existence of a  $\delta^h$ , with  $\lambda = 0$  for  $\delta \geq \delta^h$ , follows from the enforceability of first-best effort levels for  $\delta \rightarrow 1$ .

Now, consider a left neighborhood of  $\delta^h$  where  $n_h^{FB}$  and  $n_l^{FB}$  do not satisfy (ECh), and thus  $\lambda > 0$ . Condition (D.5) gives  $n^h < n_h^{FB}$ , whereas (D.6) gives  $n_i^l < n_l^{FB}$ . Condition (D.6) also implies that  $\lim_{i \rightarrow \infty} n_i^l = n_l^{FB}$ : Since  $(q^h - q^l)^i < 1$ ,  $\lim_{i \rightarrow \infty} (q^h - q^l)^i = 0$ , hence  $\lim_{i \rightarrow \infty} (\theta^l g'(n_i^l) - c) = 0$ .

To show that  $n_i^l < n_{i+1}^l$ , rewrite conditions (D.6) for  $n_i^l$  and for  $n_{i+1}^l$  as

$$\begin{aligned}
&(\theta^l g'(n_i^l) - c) \frac{(1 - \delta(1 - q^l))(1 - q^h)(1 - q^l)^i}{(1 - \delta)(1 - \delta(q^h - q^l))} \\
&= -\lambda \left[ (\theta^l g'(n_i^l) - c) \frac{(1 - \delta(1 - q^l))(1 - q^h)}{(1 - \delta)(1 - \delta(q^h - q^l))} (1 - q^l)^i - q^h (q^h - q^l)^i (\theta^h - \theta^l) g'(n_i^l) \right]
\end{aligned}$$

$$\begin{aligned}
& (\theta^l g'(n_{i+1}^l) - c) \frac{(1 - \delta(1 - q^l)) (1 - q^h) (1 - q^l)^{i+1}}{(1 - \delta) (1 - \delta(q^h - q^l))} \\
= & -\lambda \left[ (\theta^l g'(n_{i+1}^l) - c) \frac{(1 - \delta(1 - q^l)) (1 - q^h)}{(1 - \delta) (1 - \delta(q^h - q^l))} (1 - q^l)^{i+1} - q^h (q^h - q^l)^{i+1} (\theta^h - \theta^l) g'(n_{i+1}^l) \right]
\end{aligned}$$

Dividing the first by the second equality and reformulating yields the necessary condition

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

As  $\frac{(q^h - q^l)}{(1 - q^l)} < 1$ , this implies  $\frac{(\theta^l g'(n_{i+1}^l) - c)}{(\theta^l g'(n_i^l) - c)} < \frac{g'(n_{i+1}^l)}{g'(n_i^l)}$ . This is equivalent to  $g'(n_i^l) > g'(n_{i+1}^l)$ , which yields  $n_{i+1}^l > n_i^l$  due to the strict concavity of  $g(\cdot)$ . ■

Proposition 9 suggests that recovery may be gradual and never complete, as in the case of permanent shocks. The solution to the maximization underlying this proposition constitutes an equilibrium for parameter values such that the (EC1) and (TT1) constraints hold at the solution. While we can show that this is the case for an open, non-empty, subset of the parameter space, we leave a complete characterization of this subset outside the scope of this paper.

Concerning the intuition of this result, recall that with persistent shocks, falsely claiming that the type is low forces the principal to stick to announcing the low state forever thereafter. This is not the case with persistent, impermanent, shocks. Indeed, by the One-deviation-principle (Hendon, Jacobsen, and Sloth (1996)), the costs of a deviation today are increasing in the size of tomorrow's high-type bonus  $b^h$  – because the likelihood of having to pay  $b^h$  is larger off the equilibrium path. Therefore, tomorrow's high-type bonus  $b^{lh}$  is set as high as feasible, bounded as it is by the respective truth-telling constraint. This truth-telling constraint is again relaxed by a large high-type bonus the day after tomorrow,  $b^{llh}$ , and so on. In contrast to the iid case, these consecutively binding truth-telling constraints make it optimal to distort later  $n_i^l$  as well. Because of discounting and the decreasing difference between on-path and off-path likelihoods of having to pay high-type bonuses, these distortions decrease with  $i$ , and eventually vanish, as for permanent shocks.

## References

- ABREU, D. (1988): “On the Theory of Infinitely Repeated Games with Discounting,” *Econometrica*, 56(6), 1713–1733.
- AHAMMER, A., D. GRÜBL, M. LAPINSKI, AND R. WINTER-EBMER (2018): “The Hidden Costs of Mass Lay-offs,” Working Paper, JKU Linz.
- AMABILE, T. M., AND R. CONTI (1999): “Changes in the Work Environment for Creativity During Downsizing,” *Academy of Management Journal*, 42(6), 630–640.
- AZARIADIS, C. (1983): “Employment With Asymmetric Information,” *The Quarterly Journal of Economics*, 98, 157–172.
- 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.
- DUBE, A., E. FREEMAN, AND M. REICH (2010): “Employee Replacement Costs,” Working Paper, University of Berkeley.
- 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.
- GROSSMAN, S. J., AND O. D. HART (1983): “Implicit Contracts Under Asymmetric Information,” *The Quarterly Journal of Economics*, 98, 123–156.
- GUTHRIE, J. P., AND D. K. DATTA (2008): “Dumb and Dumber: The Impact of Downsizing on Firm Performance as Moderated by Industry Conditions,” *Organization Science*, 19(1), 108–123.
- HALAC, M. (2012): “Relational Contracts and the Value of Relationships,” *American Economic Review*, 102(2), 750–779.

- HART, O. (1983): “Optimal Labour Contracts under Asymmetric Information: An Introduction,” *The Review of Economic Studies*, 50(1), 3–35.
- HENDON, E., H. J. JACOBSEN, AND B. SLOTH (1996): “The One-Shot-Deviation Principle for Sequential Rationality,” *Games and Economic Behavior*, 12(2), 274–282.
- LEVIN, J. (2002): “Multilateral Contracting and the Employment Relationship,” *The Quarterly Journal of Economics*, 117(3), 1075–1103.
- (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, B., AND J. 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.
- MEUSE, K. P. D., T. J. BERGMANN, P. A. VANDERHEIDEN, AND C. E. RORAFF (2004): “New Evidence Regarding Organizational Downsizing and a Firm’s Financial Performance: A Long-term Analysis,” *Journal of Managerial Issues*, 16(2), 155–177.
- MEUSE, K. P. D., AND G. DAI (2013): “Organizational Downsizing: Its Effect on Financial Performance Over Time,” *Journal of Managerial Issues*, 25(4), 324–344.

MUEHLEMANN, S., AND M. S. LEISER (2018): “Hiring costs and labor market tightness,” *Labour Economics*, 52, 122 – 131.