

# Building Reputation for Contract Renewal: Implications for Performance Dynamics and Contract Duration<sup>\*</sup>

Elisabetta Iossa<sup>†</sup> and Patrick Rey<sup>‡</sup>

First version: 1 July 2009; this version: 28 May 2013

## Abstract

We study how career concerns affect the dynamics of incentives in a multi-period contract, when the agent's productivity is a stochastic function of his past productivity and investment. We show that incentives are stronger and performance is higher when the contract approaches its expiry date. Contrary to common wisdom, long-term contracts may strengthen reputational effects whereas short-term contracting may be optimal when investment has persistent, long-term effects.

*Keywords:* Career concerns, contract duration, contract renewal, reputation and dynamic incentives.

---

<sup>\*</sup>For helpful discussions and comments we thank Luca Anderlini, Luisa Affuso, Per Agrell, Stefano Gagliarducci, Philippe Gagnepain, Bruno Jullien, Andy Newman, David Salant, Stephane Saussier, Giancarlo Spagnolo, Andrzej Skrzypacz, Yossi Spiegel, Tommaso Valletti, Jeffrey Zwiebel, seminars and conference participants at University of Bicocca, EIEF, University of Luxemburg, the International Conference on Infrastructure Economics and Development (Toulouse 2010), the XI CEPR Conference in Applied Industrial Organization (Toulouse, 2010), the International Conference on "Contracts, Procurement and Public-Private Agreements" (Paris, 2010), the 6th Csef-Igier Symposium on Economics and Institutions (Capri 2011). We also thank the editor, George-Marios Angeletos, as well as three referees, for very helpful comments.

<sup>†</sup>University of Rome Tor Vergata, CEPR, EIEF and IEFE-Bocconi. This author acknowledges financial support from the Italian Ministry of Education.

<sup>‡</sup>Toulouse School of Economics (GREMAQ and IDEI).

# 1 Introduction

In January 2006, the university canteen for staff members of an Italian Economics Department was serving overcooked pasta and oily tomato sauce. The canteen was run by a private contractor on a 5-year contract, expiring in December 2006. From November 2006, higher standards were observed; the pasta being cooked *al dente* and some delicious sauces being served. There were rumors that a new chef had been hired. In the following December the university decided to renew the contract with the catering company, but the higher standards did not last: by the end of January 2007, the quality of the food had gone back to its initial level.

Similar features have been observed in other sectors such as sports or public procurement. In professional basketball, using data on individual performance (points scored, rebounds, assists, etc.) covering the 1980s and 1990s, Stiroh (2007) found that performance improved significantly in the year before signing a multi-year contract, and declined after the contract was signed.<sup>1</sup> The same pattern was noted in professional baseball.<sup>2</sup> In the UK railway industry, using a panel of the 25 franchisees providing passenger services in 1997-2000, Affuso and Newbery (2000) found that contractors' voluntary investment in rolling stocks increased – presumably leading to better service – towards the end of the contracts. And in the French water industry, Chong, Huet and Saussier (2006) found that contracts near expiry date had lower prices compared to other contracts, all things being equal.<sup>3</sup>

In this paper, we first provide a rationale for such performance patterns;<sup>4</sup> it is based on information decay, due e.g. to random productivity shocks. For example, a catering company using natural ingredients may benefit from a change in market attitude towards organic food; an athlete may have an accident that affects his ability to play; a private contractor may experience improved productivity thanks to technological change; and so on. Because of such information

---

<sup>1</sup>He also found that players were strongly rewarded for improvements in performance in their contract year: a one-point increase in a player's scoring average, for example, was associated with an annual salary increase of over \$300,000. Sen and Rice (2008) document a similar pattern, using performance and contract data from a panel of NBA players from 1991 to 2006.

<sup>2</sup>Using a sample of free agent contracts signed in the US between 2000 and 2004, baseball analysts observed that, with the exception of superstar players who had already established themselves as perennial caliber players, hitting players' performance was exceeding expectations in the last year of their contract and declining drastically in the following three years. See [http://baseballanalysts.com/archives/2006/03/longterm\\_free\\_a.php](http://baseballanalysts.com/archives/2006/03/longterm_free_a.php) and also <http://expertvoices.nsd.org/cornell-info204/2010/03/10/winners-curse-in-mlb-free-agency/>

<sup>3</sup>While the study relies on a cross-section analysis of 1102 French local public authorities in 2001, their finding suggests that operators raise the performance of their public-service provision, and thus the payoff of the public authority, by reducing their prices as the expiry date of the contract approaches.

<sup>4</sup>The labour literature has provided ample empirical evidence and rationale as to why performance may be higher (or absenteeism may be lower) before tenure or promotion, but lower afterwards (see e.g. Lazear, (2004), Ichino and Riphahn (2005)). Less attention, however, has been devoted to performance patterns within the pre-promotion or pre-tenure period.

decay, recent performance provides more information than past performance about the agent’s (current) productivity. This in turn induces the agent to improve his performance when a new contract is to be signed, e.g. by investing (hiring a new chef for a catering company, adopting a better training technique for an athlete, investing in a new technology for a company, and so on), or by exerting effort (spending more time in the kitchen for a chef, undertaking more intense training for an athlete, hiring temporary staff for a company, and so on).

To capture this insight in a simple way, we consider in Section 2 a baseline model in which an agent provides a good or service for two periods, followed by competition for the agent’s services in the next period. In each period, the agent’s performance is observable but nonverifiable and depends on his unobservable innate productivity (the agent’s “type”). Over time, this productivity may change exogenously (through random shock), or endogenously through nonverifiable investment.

In this setting, the agent cannot be explicitly incentivized to invest, as performance and investment are nonverifiable, but career concerns create an implicit incentive, as by providing a good performance the agent enhances his future market value. Indeed, in the last period of the contract the agent invests if needed when information is sufficiently persistent (so that current performance provides a good enough signal of future productivity and performance), the future matters (that is, the agent’s discount factor is high enough) and the investment brings great value for its cost. By contrast, the agent never invests in the first period of the contract. This is due to information decay (as earlier past performance is less informative about future productivity and performance), discounting (as the agent would rather incur the cost of investment later than sooner), as well as *crowding out*: the agent has less incentives to invest if investment occurs anyway in the following period. As a result, expected performance improves towards the end of the contract.

We show that these insights apply as well when the agent interacts with a single buyer, as it is often the case in public procurement.<sup>5</sup> We consider as well the case where the agent can also exert effort to improve current performance. In the second period, the agent then favours effort over investment whenever the former is cheaper; this, in turn, creates incentives to invest in the first period, although expected performance still weakly improves over time. Thus, for products or services with large potential for performance improvement, such as catering services, public services, consultancy, or sports, both long-term investment and short-term effort can take place

---

<sup>5</sup>This links our paper to the repeat-purchase mechanism for experience goods pioneered by Klein and Leffler (1981). The incentives generated by contract renewal were initially investigated by Taylor and Wiggings (1997).

– the former at the beginning of a contract and the latter towards the end.

To derive the implications for optimal contract duration, we consider an infinite repetition of the model. Increasing the duration of the contracts has then two opposite effects on performance. On the one hand, it reduces the frequency of investment, because as before the agent only invests in the last period of a contract. On the other hand, it increases the intensity of the investment in these last periods: there is less crowding out, as future investment is further away. The optimal contract duration depends on the relative importance of these two effects: for example, longer contracts are desirable when it is important to foster the intensity of investment in the last contracting periods, which is the case when investment is costly, information decay is high, or the weight given to the future (as measured by the discount factor) is low. Thus, contrary to common wisdom, long-term contracting is optimal when there is a need for greater implicit incentives, whereas short-term contracting is optimal when noncontractible investment has persistent, long term effects.

Our paper incorporates information decay in a multi-period contract with career concerns. In Holmström (1982), the agent exerts effort to increase performance, in an attempt to influence market beliefs about his type. However, whilst this literature also explores information decay due to switching types,<sup>6</sup> the focus has been on stationary environments.<sup>7</sup> We focus instead on performance patterns over the life of the contracts.<sup>8</sup> In addition, by allowing for endogenous type-switching, we are able to study the choice between such (long-term) productivity investment and (short-term) performance improvement effort, as well as their interaction over time. Board and Meyer-ter-Vehn (2010) also introduce endogenous switching types in a career concerns settings, but their focus is on the value of reputation under different market learning hypotheses, rather than on performance patterns.

Our paper also contributes to the literature on the effects of contract duration, or the frequency of evaluations, on incentives.<sup>9</sup> On the one hand, longer contracts alleviate moral hazard problems by facilitating consumption smoothing (Lambert (1983)) and ease hold-up

---

<sup>6</sup>See for example Mailath and Samuelson (2001) and Phelan (2006). As discussed in depth by Bar-Isaac and Tadelis (2008), for reputational concerns to be sustained over time, the market must never fully learn the type of the agent. This may hold for example if types exogenously change over time or if there is finite memory.

<sup>7</sup>Hidden action models of relational contracting have also focused on stationary environments rather than on performance dynamics; see MacLeod (2007) for a survey.

<sup>8</sup>Information decay and career concerns have been used to explain performance dynamics in other contexts – e.g., earnings smoothing in Fudenberg and Tirole (1995); pre-election policies in Cukiermann and Meltzer (1986) and Martinez (2009); and political budget cycle in Rogoff (1990).

<sup>9</sup>A link also exists with the literature in repeated games with imperfect monitoring that looks at the effect of frequency of interaction (or the effect of shortening the period length) on incentives; see e.g. Abreu, Milgrom, and Pearce (1991).

and ratchet effects in the presence of specific investment (Laffont and Tirole (1993)). On the other hand, shorter contracts increase the flexibility to use new information as it comes along (Ellman (2006)) and reduce the gain from defection from implicit agreements (Shapiro (1983), Strausz (2009) and Calzolari and Spagnolo (2009)), making them optimal when the agent’s hidden actions do not generate persistent effects.<sup>10</sup> This literature has mostly focused on moral hazard issues. By allowing also for heterogeneous productivity types and information decay, we obtain quite different implications.

The structure of the paper is as follows. Section 2 presents the baseline model where there is competition for the service of an agent. The agent can invest to enhance his productivity; we mainly focus on constant returns to scale, but also discuss the case of decreasing returns to scale, as well as the case of a principal-agent relationship. Section 3 studies optimal contract duration in an infinitely repeated version of the baseline model with investment only. Section 4 extends the baseline model to the case where the agent can choose between effort and investment. Section 5 concludes and provides some testable predictions. Some proofs are relegated in the Appendix, and the more technical developments as well as some extensions are provided in an online Appendix.

## 2 Baseline model

We first focus here on the dynamics of incentives provided by career concerns, when the agent’s productivity can evolve both exogenously (random shocks) and endogenously (investment). We consider a baseline model in which an agent provides a good or service for two periods ( $t = 1, 2$ ), followed by competition for the agent’s services in the next period ( $t = 3$ ). As in Holmström (1982):

- the agent’s “performance” (e.g., output, quality, and so forth) is observable but nonverifiable; it depends on the agent’s innate productivity (his “type”), which for technological reasons may exogenously change over time;
- at the end of the two periods, the agent’s payoff depends on market beliefs about his productivity.

---

<sup>10</sup>Evidence on the determinants of contract duration shows that contracts are longer when relationship-specific investment is important (Joskow (1987)) and shorter in periods of higher uncertainty (Masten and Crocker (1985) and Saussier (1999)), which is consistent with the benefit of shorter contracts in the presence of information decay. The findings of Bandiera (2007) are also consistent with the idea that contract length is chosen to provide incentives.

We depart from Holmström’s model, however, by allowing the agent to make a nonverifiable (and possibly unobservable) investment to enhance his productivity.

## 2.1 Framework

*Productivity and performance.* In each period  $t$ , the agent’s performance  $q_t$  can take two values: high ( $H$ ) or low ( $L$ ). The realized performance depends on: (i) the agent’s productivity type,  $\theta_t$ , which can also take two values, high ( $H$ ) or low ( $L$ ), and (ii) investment, which costs  $c$  to the agent; we denote by  $i_t \in [0, 1]$ , the probability that the agent actually invests in period  $t$ . We assume that the initial productivity of the agent,  $\theta_1$ , is randomly drawn and equally likely to be  $H$  or  $L$ ; absent investment, the productivity  $\theta_t$  evolves over time according to a Markov process, with transition probability  $\rho \geq 1/2$ . We further assume that current type and investment affect current performance and future types in a similar way:

- if  $\theta_t = H$  or  $i_t = 1$ , then

$$q_t = H \text{ and } \theta_{t+1} = \begin{cases} H & \text{with probability } \rho \\ L & \text{with probability } 1 - \rho \end{cases} ;$$

- if instead  $\theta_t = L$  and  $i_t = 0$ , then

$$q_t = L \text{ and } \theta_{t+1} = \begin{cases} L & \text{with probability } \rho \\ H & \text{with probability } 1 - \rho \end{cases} .$$

An interpretation of this process is that the agent can invest to improve his productivity, which enhances both current performance and the distribution of his productivity in the next period. It follows that current performance constitutes a “sufficient statistic” for the agent’s future productivity and performance. The parameter  $\rho$  reflects information decay: a lower value of  $\rho$  denotes faster changing technologies and therefore a lower probability that, absent any investment, the agent’s productivity in period  $t + 1$  remains the same as in period  $t$ .

*Payoff.* As productivity and performance are nonverifiable, they do not affect the agent’s payoff; this payoff thus depends only on the agent’s market value in period 3, which increases with his expected performance in that period,  $q_3^e \equiv E[q_3 \mid q_1, q_2]$ . For the sake of exposition, we will simply assume that the agent appropriates  $q_3^e$  – later on, the agent’s market value will also depend on exogenous variables such as the parties’ discount factor and endogenous variables such as the expected investment in future periods. Denoting by  $\delta$  the discount factor of the agent, his objective is thus:

$$\Pi = \sum_{t=1}^{t=3} \delta^{t-1} \pi_t = -ci_1 - \delta ci_2 + \delta^2 q_3^e.$$

*Timing.* The timing of the game is as follows. Periods  $t = 1$  and  $t = 2$ : (i)  $\theta_t$  is realized and observed by the agent; (ii) the agent chooses whether to invest; (iii)  $q_t$  is realized and publicly observed.<sup>11</sup> Then in period  $t = 3$ , the agent obtains  $q_3^e$ .

## 2.2 Performance dynamics

As performance and investment are nonverifiable, the agent cannot be explicitly incentivized to provide high performance. However, career concerns create an implicit incentive, as past performance affects the market value of the agent in period 3. In our setting, only the performance in the last period of the contract actually matters. Let  $\Delta \equiv H - L$ ; we have:

**Proposition 1** *The payoff of the agent in period 3 increases with his performance in the previous period,  $q_2$ :  $\pi_3 = q_3^e(q_2)$ , given by*

$$q_3^e(q_2) = E[q_3 | q_2] = \begin{cases} L + \rho\Delta & \text{if } q_2 = H, \\ L + (1 - \rho)\Delta & \text{if } q_2 = L. \end{cases} \quad (1)$$

**Proof.** This follows directly from the fact that: (i)  $q_2$  provides a sufficient statistic for  $\theta_3$ , the agent's productivity at  $t = 3$ ; and (ii) period 3 being the last period, the agent never invests in that period. ■

In period 2 a low-type agent ( $\theta_2 = L$ ) thus chooses  $i_2$  so as to maximize:

$$\Pi_2^L = -ci_2 + \delta \{L + [i_2\rho + (1 - i_2)(1 - \rho)]\Delta\} = \delta [L + (1 - \rho)\Delta] + (b_2 - c)i_2,$$

where

$$b_2 \equiv (2\rho - 1)\delta\Delta. \quad (2)$$

The optimal investment is thus  $i_2^* = 1$  if  $c < b_2$ , and  $i_2^* = 0$  if  $c > b_2$ .

The expected payoff of a high-type agent ( $\theta_2 = H$ ) is instead equal to  $\Pi_2^H = \delta(L + \rho\Delta)$ . Therefore, in period 1 a low-type agent ( $\theta_1 = L$ ) will choose  $i_1$  so as to maximize:

$$\begin{aligned} \Pi_1^L &= -ci_1 + \delta \{i_1 [\rho\Pi_2^H + (1 - \rho)\Pi_2^L] + (1 - i_1) [(1 - \rho)\Pi_2^H + \rho\Pi_2^L]\} \\ &= (1 - \rho)\Pi_2^H + \rho\Pi_2^L + (b_1 - c)i_1, \end{aligned}$$

where the marginal benefit is now

$$b_1 \equiv (2\rho - 1)\delta(\Pi_2^H - \Pi_2^L) = (2\rho - 1)\delta[b_2 - (b_2 - c)i_2^*]. \quad (3)$$

---

<sup>11</sup>The analysis does not depend here on whether the principal observes the productivity  $\theta_t$  and/or the investment decision  $i_t$ : observing the performance  $q_t$  eliminates any relevant information asymmetry. In later sections, in which the agent may temporarily increase performance, observing only that performance maintains some ambiguity.

Thus, in each period  $t = 1, 2$ , the agent chooses his investment  $i_t$  so as to maximize  $(b_t - c) i_t$ ; the marginal benefit  $b_t$  is however lower in period 1 (that is,  $b_1 < b_2$ ), for two reasons:

- information decay ( $\rho < 1$ ) and discounting ( $\delta < 1$ ) reduce the direct return from investment, as reflected by the additional factor  $(2\rho - 1)\delta$  in (3);
- in addition, there is crowding out of investments across periods: the investment expected in period 2,  $i_2^*$ , reduces the incentives to invest in period 1 (as  $b_2 \geq c$  when  $i_2^* > 0$ ).

This leads to:

**Proposition 2** *The agent never invests in period 1, and he invests in period 2 if and only if he has a low type and the cost is sufficiently low:*

$$i_t^* = 1 \text{ if and only if } t = 2, \theta_2 = L, \text{ and } c \leq c^*,$$

where  $c^* \equiv (2\rho - 1)\delta\Delta > 0$ . The incentive to invest in period 2 therefore increases with the degree of information persistence  $\rho$ , the discount factor  $\delta$ , and the performance differential  $\Delta$  relative to the cost  $c$ .

**Proof.** From the above analysis, in period 2 the agent invests if  $c < c^* = b_2$ , and does not invest if  $c > c^*$  (in the limit case where  $c = c^*$ , the agent is indifferent between investing or not); the comparative statics follow directly. We now show that the agent never invests in period 1. When  $c < c^*$ , the agent invests in period 2 whenever  $\theta_2 = L$ ; thus, investing in period 1 would have no impact on  $q_2$  (and thus on  $q_3^e$ ), and although this would reduce the probability of having to invest in period 2, the expected benefit from cost saving is only  $b_1 = (2\rho - 1)\delta c < c$ . Instead, when  $c > c^*$ , the agent prefers not to invest in period 2; but then, it does not pay to invest in period 1 either, as the benefit is even lower (due to information decay) and it comes later – that is, we then have  $b_1 = (2\rho - 1)^2 \delta^2 \Delta = (2\rho - 1)\delta c^* < c$ . ■

The impact of information decay, discounting and crowding out is here extreme and actually nullifies investment incentives in period 1: performance in that period has no impact on market beliefs, as period 2 performance provides a sufficient statistic for the agent's productivity in period 3.<sup>12</sup>

Note that information decay affects not only the temporal pattern of the implicit incentive, but also its strength. In case of full decay ( $\rho = 1/2$ ), past performance tells nothing about

---

<sup>12</sup>In this baseline model, information decay ( $\rho < 1$ ) is not necessary to generate the investment pattern: discounting and crowding out suffice here. As we will see, this is no longer the case when the interaction is repeated over time.



future productivity; it thus has no impact on market beliefs, and the agent then never invests. As  $\rho$  increases, past performance becomes more informative about the agent’s future productivity, which in turn gives the agent greater incentives. Investment incentives remain however suboptimal: although the agent fully internalizes here the impact of his investment on his performance in period 3, he ignores the short-term impact on the performance in periods 1 and 2 – this is particularly clear in period 1, where the agent never invests, but applies as well to period 2.

An implication of Proposition 2 is that, on average, the agent’s performance improves as the market evaluation approaches: the expected performance in period 1 is equally likely to be  $H$  or  $L$ , as the agent never invests in that period; in the second period, the performance is instead equal to  $H$  with probability 1 when  $c < c^*$ . Therefore:

**Corollary 1** *The expected performance (weakly) increases as the market evaluation approaches:  $q_1^e \leq q_2^e$ , with a strict inequality when  $c < c^*$ .*

**Proof.** See Appendix A. ■

## 2.3 Discussion

Before concluding this section, we briefly comment on the assumptions of this baseline model. That performance and productivity levels are all binary simplifies the exposition but does not play a critical role. The assumption that the agent knows his type is also not necessary to generate the investment and performance patterns that we have described. If  $\theta_t$  were not observed by the agent either, we would still observe investment only in period 2, as performance in that period would still provide a sufficient statistic for predicting performance in period 3.<sup>13</sup> Likewise, whilst the “sufficient statistic” feature greatly simplifies the analysis, we would expect similar increasing investment and performance patterns with alternative processes determining current productivity as a function of performance history, such as moving averages or other higher-order Markov process, as long as recent performance provides more reliable information about future productivity.

Our baseline model keeps in line with standard career concern models, in that competition for the agent’s services in period 3 enables the agent to appropriate the expected value of his

---

<sup>13</sup>We provide a formal analysis of such a variant of our baseline model in Online Appendix A. Interestingly, in this alternative baseline model, the incentive to invest increases over time even in the absence of information decay or of discounting, as there is an additional “option value” in waiting and learning about the agent’s type before deciding whether to invest.

investment for that period. This setup is relevant for situations (e.g., the market for managers) in which several “buyers” (the firms) compete for the same agent (the manager). However, our analysis carries over to contexts (e.g., public procurement for refuse collection, water distribution, prison services, and so forth) where the agent (a private contractor) provides specialized services that have mainly one buyer (the public authority), as long as the agent obtains some benefit from enhancing the principal’s beliefs about his expected performance. We illustrate this in Appendix B, using a simple model based on information rents: the agent invests to make the principal more eager to contract; as a result, investment and performance exhibit the same patterns as in our baseline model.

In a previous version of this paper (Iossa and Rey (2010)), we also considered investment aiming at reducing cost rather than improving performance. Investment incentives then decrease, rather than increase, as the renewal date approaches, and this for two reasons. First, the agent benefits more directly from cost reductions than from performance improvements. Second, good performance may signal that the agent faces low cost, giving an incentive to the principal to lower the price.<sup>14</sup>

Our baseline model also assumes for simplicity that investment is only valuable in case of low productivity; more generally, investment could be socially desirable for both types of agents, and both types could have an incentive to invest. For example, consider a variant where the probability of a high performance is additively separable in past productivity and current investment (e.g., it is of the form  $\psi\theta_t + \varphi i_t$ , with  $\psi, \varphi > 0$ ). Enhancing their market value then gives *both* types of agents the same investment incentive;<sup>15</sup> however, current performance still provides a sufficient statistic for future productivity, and thus these investment incentives remain concentrated in period 2.

In the same vein, the binary investment decision introduces “constant returns to scale”, which exacerbates the investment dynamics. Yet, although decreasing returns to scale create a motive for investment smoothing, the pattern remains “backloaded”.<sup>16</sup> To see this, we briefly consider the case where investment costs  $k(i)$  are convex ( $k'' \geq 0$ ) and satisfy the boundary conditions  $\{k(0) = 0, k'(0) < (2\rho - 1)\delta\Delta\}$ , as well as the particular case of *isoelastic costs*,

---

<sup>14</sup>This is also related to the ratchet effect identified in regulatory settings – see e.g. Pint (1992) and Laffont and Tirole (1993); Lewis (1986) stresses that the threat of cancellation may also induce the contractor to make nonverifiable cost-reducing investment in the early periods of a contract.

<sup>15</sup>Introducing cost or benefit complementarities between productivity and investment could even generate greater incentives for a high productivity agent.

<sup>16</sup>The same holds when investment returns are stochastic, in which case the agent may invest in period 1, as period 2 investment no longer guarantees high performance in the future.

where  $k(i) = ci^\varepsilon$  for some  $\varepsilon > 1$  (in the baseline model,  $\varepsilon = 1$ ):

**Proposition 3** *In case of decreasing returns to scale, a low-type agent may invest in both periods, but still invests more in period 2 than in period 1:  $0 \leq i_1^* < i_2^* \leq 1$ ; hence the expected performance still increases as the market evaluation approaches:  $q_1^e < q_2^e$ . Furthermore, for isoelastic costs the ratio  $i_2^*/i_1^*$  increases with the value of investment (i.e., when  $\Delta$  increases), information decay (i.e., when  $\rho$  decreases) and discounting (i.e., when  $\delta$  decreases).*

**Proof.** See Appendix C. ■

As in the baseline model, in each period  $t$  a low-type agent chooses  $i_t$  so as to maximize  $b_t i_t - k(i_t)$ , where the investment benefits  $b_1$  and  $b_2$  are respectively given by (3) and (2). When  $k(\cdot)$  is convex, these optimization problems have unique solutions, which are both positive when  $k(0) = k'(0) = 0$ : due to decreasing returns, it is worth investing also in period 1, in order to smooth investment cost over the lifetime of the contract. These investment benefits are moreover lower in the first period ( $b_1 < b_2$ ), due to information decay, discounting and crowding out; hence the agent still invests less in period 1 than in period 2 ( $i_1^* < i_2^*$ ), and all the more so as  $\rho$  or  $\delta$  decreases, or  $\Delta$  increases.

### 3 Optimal contract duration

In this section we extend the analysis to study the impact of contract duration on incentives and performance dynamics. For this purpose we plug the baseline model of Section 2 into an infinitely repeated framework; that is, in each period  $t = 1, 2, \dots$ : if the agent is already productive ( $\theta_t = H$ ) or invests ( $i_t = 1$ ), then current performance is high ( $q_t = H$ ) and, in the next period, the productivity is high ( $\theta_{t+1} = H$ ) with probability  $\rho$ ; otherwise, the current performance is low ( $q_t = L$ ) and, in the next period, the productivity is high only with probability  $1 - \rho$ .

Suppose that the agent renegotiates the terms of his contract every  $n$  periods. At the beginning of each contracting period  $\tau = 1, n+1, 2n+1, \dots$ , competition for the agent's services then allows him to appropriate the full expected value generated by the contract, of the form:

$$v(\mathbf{q}_{\tau-1}) = E \left[ \sum_{t=1}^{t=n} \delta^{t-1} q_{\tau+t-1} \mid \mathbf{q}_{\tau-1} \right],$$

where  $\mathbf{q}_{\tau-1} = \{q_1, \dots, q_{\tau-1}\}$  denotes the agent's performance history at the beginning of period  $\tau$ .

We will focus on Markov perfect equilibria, in which the agent's investment strategy, and thus the market expectations, are stationary. Consider a contract running from  $t = 1$  to  $t = n$ , following a performance  $q \in \{H, L\}$  in the last period of the previous contract. As in the baseline model, the most recent performance provides a sufficient statistic for the entire performance history, and thus the agent never invests before the last period of a contract:  $i_1 = \dots = i_{n-1} = 0$ ; in the last period, the agent may however invest in case of low productivity (i.e.,  $i_n > 0$ ) if it has a sufficient impact on the next market evaluation. The following Lemma describes the corresponding productivity and performance pattern:

**Lemma 1** *The agent never invests before the last period of a contract. The probability  $p_t^q$  of a high productivity in period  $t$  ( $\theta_t = H$ ), as a function of the performance  $q \in \{H, L\}$  observed in the last period of the previous contract, is then given by:*

$$p_t^H \equiv \frac{1 + (2\rho - 1)^t}{2}, p_t^L \equiv 1 - p_t^H.$$

The probability of a high performance in period  $t \in \{1, \dots, n-1\}$  is  $\mu_t^q = p_t^q$ ; in the last period ( $t = n$ ), this probability becomes  $\mu_n^q = p_n^q + (1 - p_n^q) i_n$ .

**Proof.** See Appendix D. ■

The value  $v^q$  generated by the contract is therefore given by:

$$v^q = \sum_{t=1}^n \delta^{t-1} (L + \mu_t^q \Delta) \quad (4)$$

Hence, the agent's overall expected payoff, evaluated at the beginning of the contract, can be expressed as, for  $q \in \{H, L\}$ :

$$\Pi^q = v^q - \delta^{n-1} (1 - p_n^q) c i_n + \delta^n [\mu_n^q \Pi^H + (1 - \mu_n^q) \Pi^L]. \quad (5)$$

This equation determines recursive conditions that determine  $\Pi^H$  and  $\Pi^L$ , which we can use to analyze the agent's incentives to invest in the last period of the contract. When facing a low productivity in period  $n$ , the agent will then make his investment decision so as to maximize  $(b_n - c) i_n$ , where

$$b_n \equiv \delta (\Pi^H - \Pi^L).$$

Using (5), in equilibrium the expected benefit  $b_n$  can be written as:

$$b_n = \delta \frac{v^H - v^L + \delta^{n-1} (p_n^H - p_n^L) c i_n}{1 - \delta^n (\mu_n^H - \mu_n^L)}. \quad (6)$$

where  $i_n$  corresponds here to the equilibrium investment level in the last periods of future contracts. Good performance in period  $t = n$  brings three types of benefit to the agent:

- It increases the expected performance of the next contract by

$$\delta (v^H - v^L) = \sum_{t=1}^n \delta^t (\mu_t^H - \mu_t^L) \Delta = \sum_{t=1}^n (2\rho - 1)^t \delta^t \Delta - i_n (2\rho - 1)^n \delta^n \Delta,$$

As before, this benefit increases with the performance differential  $\Delta$  but suffers from information decay (it decreases as  $\rho$  decreases), discounting (it decreases as  $\delta$  decreases) and crowding out (reflected in the last term).

- It saves on the cost of future investment,  $ci_n$ , as it reduces the probability of having to invest in the last period of the next contract by:

$$(p_n^H - p_n^L) \delta^n = (2\rho - 1)^n \delta^n.$$

This benefit also decreases with information decay and discounting.

- It raises the probability of enjoying  $\Pi^H$  rather than  $\Pi^L$  in the following market evaluation by:

$$\delta^n (\mu_n^H - \mu_n^L) = (2\rho - 1)^n \delta^n (1 - i_n).$$

- This benefit, too, decreases with information decay and discounting, and suffers as well from crowding out.

Building on this analysis, there exists a unique equilibrium, characterized as follows :

**Proposition 4** *Under  $n$ -period contracting, there is a unique Markov perfect equilibrium; in this equilibrium, the agent never invests during the first  $n - 1$  periods of a contract, and in the  $n^{\text{th}}$  period:*

- If  $c \geq \bar{c} \equiv \frac{(2\rho-1)\delta\Delta}{1-(2\rho-1)\delta}$ , the agent never invests ( $i_n^* = 0$ );
- If  $c \leq \underline{c}_n \equiv \frac{1-(2\rho-1)^{n-1}\delta^{n-1}}{1-(2\rho-1)^n\delta^n} \bar{c}$ , the agent invests with probability 1 in case of low productivity ( $i_n^* = 1$ );
- If  $\underline{c}_n < c < \bar{c}$ , the agent invests with probability  $i_n^* = \hat{i}_n \in (0, 1)$  in case of low productivity, where

$$\hat{i}_n = \frac{1 - (2\rho - 1)^n \delta^n}{(2\rho - 1)^n \delta^n} \left( \frac{(2\rho - 1) \delta}{1 - (2\rho - 1) \delta} - \frac{c}{\Delta} \right). \quad (7)$$

This equilibrium level of investment  $i_n^*$  increases with the contract duration  $n$ , the degree of information persistence  $\rho$ , the discount factor  $\delta$  and the relative performance benefit  $\Delta/c$ .

**Proof.** Due to crowding out, the expected benefit  $b_n = b_n(i_n)$  decreases as  $i_n$  increases. Therefore, in equilibrium  $i_n^* = 0$  if  $c > \bar{c} = b_n(0)$ ,  $i_n^* = 1$  if  $c < \underline{c}_n = b_n(1)$ , and if  $c \in (b_n(1), b_n(0))$  then  $i_n^* = \hat{i}_n$ , such that  $b_n(\hat{i}_n) = c$ , which leads to (7). From (7), the comparative statics of  $\hat{i}_n$  with respect to  $n$  and  $\Delta/c$  are straightforward. For those with respect to  $\rho$  and  $\delta$ , it suffices to note that  $b_n$  can be expressed as:

$$b_n(i_n) = \frac{\sum_{t=1}^{n-1} (2\rho - 1)^t \delta^t \Delta + (2\rho - 1)^n \delta^n [(1 - i_n) \Delta + ci_n]}{1 - (2\rho - 1)^n \delta^n (1 - i_n)},$$

which clearly increases with  $\rho$  and  $\delta$ , but decreases with  $i_n$ ; to see this, note that the sign of  $\frac{\partial b_n}{\partial i_n}$  is the same as (using  $\sigma \equiv (2\rho - 1)\delta$ ):

$$\begin{aligned} & - [1 - \sigma^n (1 - i_n)] \sigma^n (\Delta - c) - \sigma^n \left\{ \sum_{t=1}^{n-1} \sigma^t \Delta + \sigma^n [\Delta - (\Delta - c) i_n] \right\} \\ & = -\sigma^n \frac{1 - \sigma^n}{1 - \sigma} [\Delta - (1 - \sigma) c] \end{aligned}$$

which is negative for  $c \leq \bar{c}$ . ■

In equilibrium, the agent invests when the performance differential  $\Delta$  is large enough, compared to the investment cost  $c$ , as well as when there is not too much information decay (that is,  $\rho$  is high enough) or discounting ( $\delta$  high enough). To interpret the bound  $\bar{c}$ , note that for low investment levels: (i) there is little crowding out; and (ii) the benefit of saving on future investment costs is also low. Therefore, the agent invests with positive probability whenever the cost of doing so is lower than its “full” impact on future performance:

$$c < \delta \sum_{k=1}^{\infty} \delta^{(k-1)n} (v^H - v^L) \Big|_{i_n^*=0} = \sum_{t=1}^{\infty} \delta^t (2\rho - 1)^t \Delta = \bar{c}.$$

Obviously, the threshold  $\bar{c}$  increases with  $\delta$ ,  $\rho$  and  $\Delta$ , and thus it is natural to expect investment to increase as well with these parameters. Increasing the investment levels however generates crowding out, which tends to reduce the impact of current investment on future performance. Crowding out is maximal for one-period contracting, where the agent never invests with probability 1 ( $\underline{c}_n = 0$ ). For longer contracts, however crowding out never completely eliminates the benefit of investment: even if the agent invests with probability 1 at the end of the next contract, investing today still has a positive effect on expected performance during the first  $n - 1$  periods of the next contract (but only that one); in addition, investing today now also allows the agent to save on future investment costs. As a result, even if the agent always invests in

case of low productivity at the end of the next contracts, it is still worth investing if  $\delta$ ,  $\rho$  and  $\Delta$  are large enough, namely, if:

$$c < \delta (v^H - v^L) \Big|_{i_n^*=1} + (2\rho - 1)^n \delta^n c$$

$$\Leftrightarrow c < \frac{\delta (v^H - v^L) \Big|_{i_n^*=1}}{1 - (2\rho - 1)^n \delta^n} = \frac{\sum_{t=1}^{n-1} (2\rho - 1)^t \delta^t \Delta}{1 - (2\rho - 1)^n \delta^n} = \underline{c}_n,$$

where  $\underline{c}_n > 0$  for  $n > 1$ . Finally, extending the length of the contracts limits crowding out – as future investment comes later and less often; as a result,  $i_n^*$  increases with the contract duration  $n$ .

We now study the implications for optimal contract duration. As competition for the agent's services allows him to appropriate all the surplus, in equilibrium the expected welfare coincides with the agent's *ex ante* expected payoff:

$$W_n = \frac{\Pi_n^H + \Pi_n^L}{2},$$

where  $\Pi_n^q$  denotes the equilibrium expected payoff at the beginning of a contract under  $n$ -period contracting, given the performance  $q \in \{H, L\}$  in the last period of the previous contract. Using (5), we have:

$$W_n = v_n - \delta^{n-1} c \frac{i_n^*}{2} + \delta^n \left[ W_n + \frac{i_n^*}{2} (\Pi_n^H - \Pi_n^L) \right]$$

$$= \frac{v_n}{1 - \delta^n} + \frac{\delta^{n-1} i_n^*}{1 - \delta^n} \frac{1}{2} [\delta (\Pi_n^H - \Pi_n^L) - c], \quad (8)$$

where  $v_n = \frac{v_n^H + v_n^L}{2}$  denotes the equilibrium expected performance generated by the first contract under  $n$ -period contracting.

When investment is too costly ( $c \geq \bar{c}$ ), the agent never invests:  $i_n^* = 0$  for any  $n \geq 1$ . The expected welfare is then equal to

$$W_n = \frac{1}{1 - \delta} \frac{L + H}{2},$$

whatever the duration of the contracts. In what follows, we thus focus on the case where  $c < \bar{c}$ .

When instead contracts are sufficiently long that  $i_n^* = 1$ , the expected welfare decreases as the contract length  $n$  further increases, which then only delays the next periods of investment. As the agent never invests in the early periods of a contract, and moreover fails to internalize the impact of investment on current performance even in the last period, there is always underinvestment. It is thus never optimal to choose a duration  $n$  larger than what is needed to

induce  $i_n^* = 1$ , as this would only make the investment less frequent, without any off-setting effect on the investment intensity in the last periods of the contracts.

For shorter durations where  $i_n^* = \hat{i}_n < 1$ , the second term in (8) is nil,<sup>17</sup> in that case, the equilibrium expected welfare coincides with the *ex ante* expected performance generated by the first contract, which from (4) is given by:

$$W_n = \frac{v_n}{1 - \delta^n} = \frac{\sum_{t=1}^n \delta^{t-1} \frac{L+H}{2} + \delta^{n-1} \hat{i}_n \frac{\Delta}{2}}{1 - \delta^n} = \frac{1}{1 - \delta} \frac{L+H}{2} + \frac{\delta^{n-1}}{1 - \delta^n} \hat{i}_n \frac{\Delta}{2}. \quad (9)$$

Increasing the contract duration  $n$  has then two opposite effects on expected welfare:

- it *increases investment intensity* in the last period of a contract:  $\hat{i}_n$  increases with  $n$ ;
- but it *reduces the frequency of investment*:  $\delta^{n-1}/(1 - \delta^n)$  decreases as  $n$  increases.

The following proposition shows that the intensity effect dominates here: it is always optimal to increase  $n$  so as to increase  $\hat{i}_n$  despite the reduced frequency of its occurrence. It is therefore optimal to make contracts just long enough to induce  $i_n^* = 1$ . As this is more difficult to achieve when investment is costly or not too valuable, or when there is substantial information decay and discounting, the optimal contract duration increases with the relative cost  $c/\Delta$  and decreases as  $\rho$  or  $\delta$  increases.

We show in the Appendix that a similar pattern arises for *expected performance*, leading to:

**Proposition 5** *When  $c \geq \bar{c}$ , the agent never invests; the duration of the contracts thus has no impact on performance or welfare. When instead  $c < \bar{c}$ , expected welfare and performance are both maximized by making the contracts just long enough to yield  $i_n^* = 1$ . As an increase in the discount factor ( $\delta$ ) and/or persistence ( $\rho$ ) encourages investment, whereas an increase in the relative cost  $\frac{c}{\Delta}$  discourages it, the optimal duration  $\hat{n}((2\rho - 1)\delta, \Delta/c)$  decreases as the discount factor  $\delta$ , the relative benefit  $\Delta/c$ , and the degree  $\rho$  of information persistence increase.*

**Proof.** See Appendix E. ■

The above “corner solution” ( $i_n^* = 1$ ) is partly driven by the assumption of constant returns to scale. As we have seen, with decreasing returns to scale (i.e., convex investments costs), the agent may choose to limit his investment, and has moreover an incentive to smooth it over the duration of the contract; yet, he still has more incentives to invest in later periods than

<sup>17</sup>This is obvious if  $i_n^* = 0$ ; if  $i_n^* = \hat{i}_n \in (0, 1)$ , the agent is indifferent between investing or not in the last period of a contract, and thus  $\delta(\Pi_n^H - \Pi_n^L) = c$ .



in previous ones, and longer contracts still foster investment incentives in the last periods of a contract.<sup>18</sup> In addition, even with convex costs, we would expect more weight being put on *frequency* when investment levels are likely to be substantial even with short-term contracts, that is, when the investment is valuable ( $\Delta$  high), marginal costs are low, and there is little information decay or discounting ( $\rho$  and  $\delta$  close to 1); in such a case, we would thus expect short-term contracts to be optimal. When instead the relative costs are high, or there is a lot of information decay or discounting, then longer contracts will become desirable, in order to foster the agent's incentive to invest in the last periods of these contracts.<sup>19</sup>

Note that longer contracts may generate greater implicit incentives is opposite to what current wisdom suggests. The key here is that, as the agent never invests in the first periods of a contract, increasing the length of the contracts attenuates the crowding out effect of future investment, which in turn fosters the incentives to invest in the last period of a contract.

That short-term contracting is optimal when information persistence is high is also counterintuitive. When  $\rho$  is high, investments have long-term effects, and one may have expected long-term contracts to be optimal. This would for example be the case if the agent benefitted from the investment during the execution of the contract (e.g., investment reduces operating costs); he would then invest at the beginning of the contracts, and even more so if the contracts get longer. Here, however, the agent invests only in the last periods of the contracts, to enhance his market value for the next contracts; as information persistence makes it easier to induce the agent to invest in these last periods, it becomes more important to increase their frequency, through shorter contracts.

## 4 Extension: Signal jamming

This section extends our baseline model by allowing also for temporary performance improvements. That is, in each period  $t = 1, 2$ , a low-type agent ( $\theta_t = L$ ) can choose between investing as before in lasting improvements, or exerting effort to generate short-term effects:<sup>20</sup> exerting effort costs  $\gamma < c$  but only improves current performance ( $q_t = H$ ), and does not affect the dis-

---

<sup>18</sup>For instance, with two-period contracting the agent has more incentives to invest in the second period of a contract than in the first one, and has also more incentives to invest in that period than under one-period contracting. See Online Appendix F.

<sup>19</sup>See Online Appendix F for illustrations.

<sup>20</sup>For example, the catering company running the university canteen may increase the quality of the food served in a given period by asking its chef to spend more time in the kitchen in that period (effort). Alternatively, the company could hire a new chef (investment) bringing long-term quality improvements. Effort may also only affect *perceived* performance – for example, athletes may use more Performance Enhancing Drugs (PEDs); we thank Andy Skrzypacz for pointing this out to us.

tribution of the agent's type in the next period (that is,  $\theta_{t+1}$  coincides with  $\theta_t$  with probability  $\rho$ , regardless of the agent's effort).<sup>21</sup>

Obviously, the agent will never invest or exert effort in period 3 (as this is the last period), and in the previous periods he will never do so either when having a high productivity ( $\theta_t = H$ ). Consider now the behavior of a low productivity agent in period 2 ( $\theta_2 = L$ ). The agent will never invest in that period, as exerting effort would have the same impact on the market belief and would be less costly to produce. He may however exert effort in order to enhance market beliefs:

- If  $q_2 = L$ , the market infers that the agent has a low productivity in period 2, and thus expects in period 3 a performance

$$E[q_3 \mid q_1, q_2 = L] = L + (1 - \rho) \Delta.$$

- By contrast,  $q_2 = H$  entertains ambiguity, since a high performance can be the result of productivity or effort. Let  $\nu_{q_1}$  denote the market belief, as a function of the performance  $q_1 \in \{H, L\}$  observed in period 1, that the agent has a high productivity in period 2 when observing  $q_2 = H$ . The expected performance in period 3, is then given by:

$$\begin{aligned} E[q_3 \mid q_1, q_2 = H] &= \nu_{q_1} (\rho H + (1 - \rho) L) + (1 - \nu_{q_1}) ((1 - \rho) H + \rho L) \\ &= L + (1 - \rho) \Delta + (2\rho - 1) \nu_{q_1} \Delta. \end{aligned}$$

The expected payoff of a low-type agent, as a function of his past performance  $q_1$  and of his current effort  $e_2$ , is therefore:

$$\Pi_2^L(e_2; q_1) = \delta [L + (1 - \rho) \Delta] + ((2\rho - 1) \delta \nu_{q_1} \Delta - \gamma) e_2.$$

As a result, in period 2, effort will be undertaken if:

$$\gamma < (2\rho - 1) \delta \nu_{q_1} \Delta. \quad (10)$$

Intuitively, the market will be more optimistic if it observes a high performance in period 1 ( $\nu_H \geq \nu_L$ ); it follows that the agent has more incentive to exert effort in that case: denoting by  $e_{q_1}$  the equilibrium probability of effort in period 2, as a function of the performance  $q_1 \in \{H, L\}$  observed in period 1, we expect  $e_H \geq e_L$ .

---

<sup>21</sup>If  $\gamma \geq c$ , the agent would always favor investment over effort, and the analysis would be the same as in the baseline model.

When instead the agent has a high productivity in period 2, his expected payoff is  $\Pi_2^H(q_1) = \delta E[q_3 | q_1, q_2 = H]$ . Consider now the behavior of a low-type agent in period 1 ( $\theta_1 = L$ ). If the agent does neither exert effort nor invest, then  $q_1 = L$  and, in period 2:

- with probability  $1 - \rho + \rho e_L$ , the performance is  $q_2 = H$ , leading the market to expect a high productivity with probability  $\nu_L$ ;
- with probability  $\rho(1 - e_L)$ , the performance is  $q_2 = L$ , thus revealing the low productivity of the agent.

Thus, the agent's expected payoff from neither investing nor exerting effort in period 1 is:

$$\begin{aligned} \Pi_1^L \Big|_{i_1=e_1=0} &= \delta [(1 - \rho) \Pi_2^H(L) + \rho \Pi_2^L(e_L; L)] \\ &= \delta^2 [L + (1 - \rho) \Delta] + (1 - \rho + \rho e_L) \nu_L (2\rho - 1) \delta^2 \Delta - \rho e_L \delta \gamma. \end{aligned} \quad (11)$$

If instead the agent invests in period 1, he enhances both his current performance  $q_1$  from  $L$  to  $H$  and his future productivity; there is then no point exerting effort, and his expected payoff becomes:

$$\begin{aligned} \Pi_1^L \Big|_{\substack{i_1=1 \\ e_1=0}} &= \delta [\rho \Pi_2^H(H) + (1 - \rho) \Pi_2^L(e_H; H)] - c \\ &= \delta^2 [L + (1 - \rho) \Delta] + [\rho + (1 - \rho) e_H] \nu_H (2\rho - 1) \delta^2 \Delta - (1 - \rho) e_H \delta \gamma - c. \end{aligned} \quad (12)$$

Finally, the agent may choose to exert effort rather than to invest, in which case he only enhances his current performance  $q_1$  and his expected payoff becomes:

$$\begin{aligned} \Pi_1^L \Big|_{\substack{i_1=0 \\ e_1=1}} &= \delta [(1 - \rho) \Pi_2^H(H) + \rho \Pi_2^L(e_H; H)] - \gamma \\ &= \delta^2 [L + (1 - \rho) \Delta] + [1 - \rho + \rho e_H] (2\rho - 1) \nu_H \delta^2 \Delta - \rho e_H \delta \gamma - \gamma. \end{aligned} \quad (13)$$

Comparing these options yields:

**Proposition 6** *Suppose that the agent can invest at cost  $c$  and/or exert effort at cost  $\gamma < c$ ; then:*

- *In period 2, the agent never invests but exerts effort with positive probability in case of low productivity ( $\theta_2 = L$ ) when this is not too costly (namely, if  $\gamma \leq c^* = (2\rho - 1) \delta \Delta$ ; in period 1, the agent favours investment whenever its relative cost is not excessive (namely, if  $c/\gamma < 1 + (2\rho - 1) \delta$ ), and effort otherwise.*

- *The incentive to enhance performance increases in each period with the discount factor  $\delta$  and the performance differential  $\Delta$ ; in addition, it increases:*
  - *over time:  $e_1^* > 0$  or  $i_1^* > 0$  imply  $e_H^* = 1$  and  $e_L^* > 0$ , whereas we can have  $e_H^* > 0$ ,  $e_L^* > 0$  and  $e_1^* = i_1^* = 0$ ;*
  - *and with the agent’s reputation:  $e_H^* \geq e_L^*$ , with a strict inequality whenever  $0 < e_L^* < 1$ .*
- *In the absence of information decay ( $\rho = 1$ ), performance remains constant over time:  $q_1 = q_2$ ; with information decay ( $\rho < 1$ ), performance increases over time:  $q_1^e \leq q_2^e$ , with a strict inequality whenever costs are in a moderate range.<sup>22</sup>*

**Proof.** See Online Appendix B. ■

Long-term investment and short-term effort are substitutes in each period, and in the second period the agent favors effort over investment, as the former is a cheaper way to deliver high performance. However, in contrast with our baseline model, here the effort exerted in period 2 does not “crowd out” the incentives in the previous period: as  $q_2$  becomes less informative about the agent’s type, delivering good performance becomes valuable also in period 1.<sup>23</sup> In the first period, the agent favors investment over effort when the cost difference is small enough, as investment reduces the likelihood of having to exert effort in the next period. In the second period, the agent has more incentives to hide bad news when he is supposed to be good:  $e_H \geq e_L$ .<sup>24</sup>

Contrary to the baseline model where the agent can only invest, expected performance increases as the market evaluation approaches only when there is information decay; discounting alone does not suffice.<sup>25</sup> Indeed, if  $\rho = 1$ , then when observing a bad performance in period

---

<sup>22</sup>Namely, when (i)  $\gamma \leq (2\rho - 1)\delta\Delta$ , so that the agent exerts effort in period 2, but (ii)  $c$  and  $\gamma$  are not too low, so that  $e_1^*, i_1^* < 1$ .

<sup>23</sup>Our formulation allows for “asymmetric ambiguity”: a bad performance  $L \ll$  reveals  $\gg$  a low productivity, whereas a good performance  $H$  entertains some ambiguity, as it may be the result of the agent’s effort. More generally, incentives to exert effort remain stronger in the second period whenever, due to information decay, productivity estimates are ranked as “ $LL \leq HL < LH \leq HH$ ”.

<sup>24</sup>The literature on life-cycle effects in reputational models shows that incentives may either decrease or increase with existing reputation; see Bar-Isaac and Tadelis (2008) for a survey. Board and Meyer-ter-Vehn (2010) highlight the role of the learning process: when bad news reveal the agent’s type, as in our framework, agents with better reputation have higher incentives to invest. The opposite result obtains when instead there is perfect learning upon good news.

<sup>25</sup>In Grossman and Shapiro (1985), discounting alone explains why firms may devote more resources to a project as the project nears completion. In their model, investment yields no return until the project is completed, thus the marginal discounted benefit of investing increases as the project nears completion. In Rice and Sen (2008), where both the agent and the principal progressively learn the agent’s type, discounting suffices to generate increasing effort over time; however, the implications on the performance pattern are ambiguous.

1 ( $q_1 = L$ , thus revealing a bad productivity in that period), the market will expect for sure a low productivity in the following periods, as investing in period 2 is not credible; this, in turn, discourages the agent from exerting any effort in period 2:  $e_L = 0$ , and thus performance remains low if it is low in the first period.

It is unfortunately difficult to plug this extended version of our baseline model into a repeated interaction setting, as now the market needs to keep track of the entire history of the agent’s performance – the performance observed in the last period of a contract is no longer a sufficient statistic of the agent’s past performance when the agent may exert effort in that period. In particular, if the interaction is infinitely repeated (even if only with positive probability), then the environment is no longer stationary, as the agent’s relevant past history keeps growing over time.

We can however draw on our analysis to make some general predictions for the case of infinitely repeated interaction. If investment is not substantially more costly than effort, then the equilibrium remains the same as in Section 3: in the last period of a contract the agent will favor investment over effort, so as to save on future costs; but then, the performance in that last period is a sufficient statistic for the agent’s past performance, which eliminates any effort or investment incentives in the previous periods of the contract. When instead effort is significantly cheaper than investment, then the agent will rather exert effort in the last period of a contract, and may then choose to exert effort as well in the previous periods, so as to entertain the ambiguity, or to invest in the early periods, so as to save on effort costs in later periods. Thus, within the course of a contract, investment can precede effort but it cannot be the other way round. The above analysis also suggests that incentives to “behave” will be greater in the last periods of a contract than in the first ones, and that this pattern will be more pronounced with longer contracts.

## 5 Conclusion

Underperformance in agency relationships may be a serious problem when performance is non-verifiable and therefore cannot be contracted upon. However, when performance is (at least partially) the result of the agent’s productivity, the agent may wish to build a reputation by working harder or by investing to enhance his type; career concerns may then ease the moral hazard problem.

We build on these insights to study the performance dynamics that career concerns generate in multi-period contracts with volatile environments, and derive the implications for contract

duration. Environment volatility generates information decay. Together with discounting and crowding out, this reduces the agent's incentives to invest in earlier periods of the contract. As a result, expected performance improves towards the end of the contract.

Further, and opposite to what the current wisdom suggests, longer contracts generate greater implicit incentives, as they reduce crowding out effects. As a result, longer contracts can be optimal when the agent would otherwise be reluctant to invest, e.g., when the environment is highly volatile, making the benefits from investment rather short-lived. Conversely, shorter contracts can be optimal when investment has long-term effects on performance.

Keeping in line with standard career concern models, we mainly focused here on the case where several principals compete for the same agent; however, it is shown in Appendix B that the analysis applies as well to situations in which the agent interacts with a single principal, as in the case of public procurement.<sup>26</sup> These insights highlight the importance of granting some discretion to public authorities in the selection of their contractors. By making past performance relevant for future contract opportunities, this discretion fosters the contractor's incentives to deliver good performance. This can be particularly important for public services, such as educational services, clinical services and nursing homes, which involve many noncontractible dimensions.<sup>27</sup>

Our model provides a number of testable predictions. A first prediction is that performance (and/or investment or effort) improves significantly as contract renewal date approaches. A second prediction is that performance falls after the contract is renewed, as the agent then faces less reward for improving performance. Other things being equal, these patterns are exacerbated with longer term contracts, as the incentive to enhance performance before the contract is signed is further increased. Third, when investment exhibits decreasing returns to scale, the performance pattern is exacerbated in environments with greater information decay, as in sectors with high technological progress, unstable environments, or in sports. Fourth, in procurement contexts we should expect these patterns to be more pronounced when the procurement authority is allowed to exert her discretion, and thus to rely on nonverifiable information about the contractor's past performance, as is more often the case for private than

---

<sup>26</sup>In a previous version of the paper, we also considered the infinite repetition of the principal-agent relationship presented in that Appendix; see Iossa and Rey (2010).

<sup>27</sup>The importance of granting discretion to public procurer to allow reputational forces to operate was initially stressed by Kelman (1990). When he was the head of public procurement during the Clinton administration, Kelman adopted more flexible purchasing practices common in the private sector, among which giving more weight to suppliers' past performance. Since the Federal Acquisitions Streamlining Act in 1994, US Federal Departments and Agencies are expected to record past contractors' performance evaluations and share them through common platforms for use in future contractor selection. See Spagnolo (2012) for further details.

for public procurement (e.g. this is the case for example in European public procurement). Fifth, for industries in which both short-term performance-enhancing efforts and long-term productivity investments are relevant options, effort should take place towards the end of the contractual relationship whereas investment should rather take place earlier in the relationship. Finally, for contracts where performance is difficult to verify and explicit incentives are hard to enforce, longer contracts should be observed when the agent's stakes are small, investment is costly, and the environment is volatile.

Throughout the paper we have restricted our attention to a single agent's investment incentives. It would be interesting to extend the analysis to competing agents with heterogeneous types and study how the duration of the contract may be affected by the competition among agents.

# Appendix

## A Proof of Corollary 1

The agent's productivity is initially high or low with equal probability. As he never invests in period 1, his expected performance is:

$$q_1^e = \frac{L + H}{2} = L + \frac{\Delta}{2}.$$

By contrast, in period 2, the performance is low only if the agent's type is low *and* he does not invest; as a result, the expected performance is

$$q_2^e = L + (1 + i_2^*) \frac{\Delta}{2}.$$

The conclusion ( $q_2^e \geq q_1^e$ ) follows from  $i_2^* \geq 0$  (with a strict inequality when  $c > c^*$ ).

## B Principal-agent relationship

We sketch here a variant in which a principal  $P$  can either delegate the provision of a service to an agent  $A$  or keep it in-house. If in period  $t$   $A$  provides the service at price  $p_t$ , then  $P$  obtains a net payoff  $q_t - p_t$  whereas the agent obtains  $p_t - \phi_t - ci_t$ , where  $\phi_t \in \{\underline{\phi}, \bar{\phi}\}$  denotes  $A$ 's operating cost, with  $\Pr(\underline{\phi}) = \alpha$ . Keeping the provision in-house yields instead  $V$  for  $P$ , where  $V$  is uniformly distributed over  $[\underline{V}, \bar{V}]$ , and 0 for  $A$ . The realization of  $\phi_t$  (resp.,  $V$ ) is privately observed by  $A$  (resp.,  $P$ ) before contracting takes place.

At the beginning of period 3, upon observing the value  $V$ ,  $P$  makes a take-it-or-leave-it price offer to  $A$ . If  $P$  is optimistic about  $A$ 's future performance, she will offer a high price  $p_t = \bar{\phi}$ , to ensure that  $A$  accepts the contract;  $A$  then enjoys a rent  $\bar{\phi} - \underline{\phi}$  in case of low cost. If instead  $P$  is pessimistic about  $A$ 's future performance, she will either offer a low price  $\underline{p}_t = \underline{\phi}$ , or keep the provision in-house; either way, the agent obtains no rent. This gives  $A$  an incentive to enhance  $P$ 's beliefs, which yields the same investment and performance patterns as in the baseline model. For instance, in the baseline model in which the contract negotiation occurs at  $t = 3$ , we have:

**Proposition 7** *The expected rent of the agent in period 3 increases with his previous performance:*

$$E[\pi_3 \mid q_1, q_2] = \lambda(q_3^e(q_2) - \hat{q}),$$



where  $q_3^e(q_2)$  is given as before by (1), and  $\lambda \equiv \alpha(\bar{\phi} - \underline{\phi}) / (\bar{V} - \underline{V})$  measures the sensitivity of the agent's rent to his expected performance, compared with a target  $\hat{q} \equiv \underline{V} + 2\bar{\phi} - \underline{\phi}$ . In equilibrium, the agent never invests in period 1, whereas in period 2 he invests in case of low productivity if  $c < c_\lambda^* \equiv \lambda(2\rho - 1)\delta\Delta$ ; investment and performance thus exhibit the same pattern as in the baseline model.

**Proof.** In period 3,  $P$  will favour offering a high price  $\bar{\phi}$  over a low price  $\underline{\phi}$  (and thus a fortiori over in-house provision) when:

$$q_3^e - \bar{\phi} > \max\{\alpha(q_3^e - \underline{\phi}) + (1 - \alpha)V, V\} \iff q_3^e > V + \underline{\phi} + \frac{\bar{\phi} - \underline{\phi}}{1 - \alpha}.$$

$A$ 's expected rent is thus:

$$\begin{aligned} E[\pi_3 | q_1, q_2] &= \Pr\left(V \leq q_3^e - \underline{\phi} - \frac{\bar{\phi} - \underline{\phi}}{1 - \alpha}\right) \alpha(\bar{\phi} - \underline{\phi}) \\ &= \frac{\alpha(\bar{\phi} - \underline{\phi})}{\bar{V} - \underline{V}} \left(q_3^e - \underline{\phi} - \frac{\bar{\phi} - \underline{\phi}}{1 - \alpha} - \underline{V}\right), \end{aligned}$$

which is of the form  $\lambda(q_3^e - \hat{q})$ , where the ratio  $\lambda \equiv \alpha(\bar{\phi} - \underline{\phi}) / (\bar{V} - \underline{V})$  and the performance threshold  $\hat{q} \equiv \underline{V} + \underline{\phi} + \frac{\bar{\phi} - \underline{\phi}}{1 - \alpha}$  are independent of  $A$ 's behavior. As before,  $q_2$  provides a sufficient statistic for the agent performance in period 3:  $q_3^e = q_3^e(q_2)$ , as given by (1). The analysis is then formally the same as in the baseline model, scaling the expected benefits from investment,  $b_1$  and  $b_2$ , by a factor  $\lambda$ : that is, in each period  $t$  a low-productivity agent ( $\theta_t = L$ ) now seeks to maximize  $(\lambda b_t - c)i_t$ . The other conclusions follow. ■

## C Proof of Proposition 3

In period 2, the agent chooses  $i_2$  so as to maximize  $b_2 i_2 - k(i_2)$ , where as before  $b_2 = (2\rho - 1)\delta\Delta$ ; hence  $i_2^* > 0$  (as  $k(0) = 0$  and  $k'(0) < b_2$ ) and increases with  $b_2$  (strictly so as long as it remains lower than 1). In period 1, the agent maximizes  $b_1 i_1 - k(i_1)$ , where  $b_1$  :

$$b_1 = (2\rho - 1)\delta [b_2 - \{b_2 i_2^* - k(i_2^*)\}].$$

As  $(2\rho - 1)\delta < 1$  and  $b_2 i_2^* - k(i_2^*) = \max_{i_2} \{b_2 i_2 - k(i_2)\} > 0$ ; hence,  $b_1 < b_2$ . In addition,  $i_1^* < 1$ :  $i_1^* = 1$  would imply  $i_2^* = 1$  as well, in which case  $b_1 = (2\rho - 1)\delta k(1) < k'(1)$ , where the inequality stems from  $(2\rho - 1)\delta < 1$  and  $k(1) > k'(1)$  by convexity; it follows that  $i_1^* < i_2^*$ .

Suppose now that  $k(i) = ki^\varepsilon$ , where  $k > 0$  and  $\varepsilon > 1$ . In that case,

$$i_t^* = \left( \frac{b_t}{\varepsilon k} \right)^{\frac{1}{\varepsilon-1}},$$

and thus the ratio  $i_1^*/i_2^*$  evolves as  $b_1/b_2$ , which is equal to (with  $\sigma \equiv (2\rho - 1)\delta$ ):

$$\begin{aligned} \frac{b_1}{b_2} &= \frac{\sigma [\sigma \Delta - \{\sigma \Delta i_2^* - k(i_2^*)\}]}{\sigma \Delta} \\ &= \sigma - \frac{\max_{i_2} \{\sigma \Delta i_2 - k(i_2)\}}{\Delta}. \end{aligned}$$

Hence,

$$\frac{\partial}{\partial \Delta} \left( \frac{b_1}{b_2} \right) = -\frac{k(i_2^*)}{\Delta^2} < 0,$$

so that  $i_1^*/i_2^*$  decreases with  $\Delta$ , and (using the envelope theorem, i.e.,  $k'(i_2^*) = b_2 = \sigma \Delta$ ):

$$\frac{\partial}{\partial \sigma} \left( \frac{b_1}{b_2} \right) = 1 - i_2^* > 0,$$

so that  $i_1^*/i_2^*$  increases with both  $\rho$  and  $\delta$ .

## D Proof of Lemma 1

By construction,  $p_1^H = \rho$  and, for  $t > 1$ :

$$p_t^H = \rho p_{t-1}^H + (1 - \rho)(1 - p_{t-1}^H) \implies 2p_t^H - 1 = (2\rho - 1)(2p_{t-1}^H - 1).$$

Therefore, by iteration:

$$2p_t^H - 1 = (2\rho - 1)^{t-1}(2p_1^H - 1) = (2\rho - 1)^t \implies p_t^H = \frac{1}{2} + \frac{(2\rho - 1)^t}{2}.$$

Similarly,  $p_1^L = 1 - \rho$  and, for  $t > 1$ :

$$\begin{aligned} p_t^L &= \rho p_{t-1}^L + (1 - \rho)(1 - p_{t-1}^L) \implies 2p_t^L - 1 = (2\rho - 1)^{t-1}(2p_1^L - 1) \\ &\implies p_t^L = \frac{1}{2} - \frac{(2\rho - 1)^t}{2} = 1 - p_t^H. \end{aligned}$$

## E Proof of Proposition 5

When the contracts are sufficiently lengthy to ensure that  $i_n^* = 1$ , from (8) the expected welfare is equal to:

$$\begin{aligned} W_n &= \frac{v_n}{1 - \delta^n} + \frac{\delta^{n-1}}{1 - \delta^n} \frac{1}{2} [\delta (\Pi_n^H - \Pi_n^L) - c] \\ &= \frac{\frac{1-\delta^n}{1-\delta} \frac{L+H}{2} + \delta^{n-1} \frac{\Delta}{2}}{1 - \delta^n} + \\ &\quad + \frac{\delta^{n-1}}{2(1 - \delta^n)} \left[ \frac{1 - (2\rho - 1)^n \delta^n}{1 - (2\rho - 1) \delta} (2\rho - 1) \delta \Delta - (2\rho - 1)^n \delta^n (\Delta - c) - c \right] \\ &= \frac{1}{1 - \delta} \frac{L + H}{2} + \chi_n \frac{1}{2\delta} \left( \frac{\Delta}{1 - (2\rho - 1) \delta} - c \right), \end{aligned}$$

where

$$\chi_n \equiv \frac{\delta^n (1 - (2\rho - 1)^n \delta^n)}{1 - \delta^n}.$$

In Online Appendix C we show that  $\chi_n$  decreases when  $n$  increases; hence it is never optimal to choose a duration  $n$  larger than what is needed to induce  $i_n^* = 1$ .

When instead  $i_n^* = \hat{i}_n < 1$  using (7) and (9) the equilibrium welfare is of the form:

$$W_n = \frac{1}{1 - \delta} \frac{L + H}{2} + \xi_n \left( \frac{(2\rho - 1) \delta}{1 - (2\rho - 1) \delta} - \frac{c}{\Delta} \right) \frac{\Delta}{2\delta},$$

where

$$\xi_n \equiv \frac{\delta^n}{1 - \delta^n} \frac{1 - (2\rho - 1)^n \delta^n}{(2\rho - 1)^n \delta^n} = \frac{1 - (2\rho - 1)^n \delta^n}{(2\rho - 1)^n (1 - \delta^n)}$$

increases with  $n$  – see Online Appendix D. Therefore, the optimal duration consists of choosing  $n$  so that  $\hat{i}_n = 1$ . The comparative statics follows from Proposition 4, and in particular from  $i_n^*$  increasing with  $n$ ,  $\Delta/c$ ,  $\delta$  and  $\rho$ .

A similar analysis can be made for the evaluation of the expected performance generated by the contract. As shown in Online Appendix E, as long as the equilibrium investment is  $i_n^* = \hat{i}_n < 1$ , the expected total discounted value of performance is of the form:

$$\begin{aligned} V_n &= \frac{1}{1 - \delta} \frac{L + H}{2} + \frac{\delta^n}{1 - \delta^n} \hat{i}_n \frac{1}{1 - (1 - (2\rho - 1) \delta) \frac{c}{\Delta}} \frac{\Delta}{2\delta} \\ &= \frac{1}{1 - \delta} \frac{L + H}{2} + \xi_n \left( \frac{(2\rho - 1) \delta \Delta}{1 - (2\rho - 1) \delta} - c \right) \frac{1}{\Delta - (1 - (2\rho - 1) \delta) c} \frac{\Delta}{2\delta}, \end{aligned}$$

whereas for  $i_n^* = 1$ , it is equal to:

$$V_n = \frac{1}{1 - \delta} \frac{L + H}{2} + \chi_n \frac{1}{1 - (2\rho - 1) \delta} \frac{\Delta}{2\delta}.$$

As  $\xi_n$  increases whereas  $\chi_n$  decreases when  $n$  increases, it follows that, to maximize this expected total performance, it is optimal to choose  $n$  so that  $\hat{i}_n = 1$ .

# References

## References

- [1] Abreu, Dilip, Paul, Milgrom and David Pearce (1991). “Information and Timing in Repeated Partnerships.” *Econometrica*, 59, 1713–1733.
- [2] Affuso, Luisa and David Newbery (2000). “Investment, Reprourement and Franchise Contract Length in The British Railway industry.” *CEPR Discussion Papers* 2619.
- [3] Bandiera, Oriana (2007). “Contract Duration and Investment Incentives: Evidence from Land Tenancy Agreements.” *Journal of the European Economic Association*, 5, 953-986.
- [4] Bar-Isaac, Heski, and Steven Tadelis (2008). “Seller Reputation.” *Foundations and Trends in Microeconomics*, 4, 273-351.
- [5] Board, Simon and Meyer-Ter-Vehn Moritz (2010). “Reputation for Quality,” *mimeo UCLA*.
- [6] Calzolari, Giacomo and Giancarlo Spagnolo (2009). “Relational Contracts and Competitive Screening.” CEPR Discussion Paper n° 7434.
- [7] Chong, Eshien, Freddy, Huet and Stephane Saussier (2006). “Auctions, Ex Post Competition and Prices: The Efficiency of Public-Private Partnerships.” *Annals of Public and Cooperative Economics*, 77, 524-561.
- [8] Cukierman, Alex and Allan H. Meltzer (1986). “A positive Theory of Discretionary Policy. The Cost of Democratic Government and the Benefit of a Constitution,” *Economic Inquiry*, 24, 367-388.
- [9] Ellman, Matthew (2006). “The Optimal Length of Contracts with Application to Outsourcing.” Economics Working Paper n° 965, Universitat Pompeu Fabra.
- [10] Fudenberg, Drew and Jean Tirole (1995). “A theory of Dividends Smoothing Based on Incumbency Rents.” *Journal of Political Economy*, 103, 75-93.
- [11] Grossman, David and Carl Shapiro (1986). “Optimal Dynamic R&D Programs.” *The Rand Journal of Economics*, 17, 581-593.

- [12] Holmström, Bengt (1982). "Managerial Incentive Problems - A Dynamic Perspective" in *Economics and Management in Honor of Lars Wahlbeck*, Helsinki: Swedish School of Economics, reprinted in *Review of Economic Studies*, 66, 169–182.
- [13] Ichino, Andrea and Regina T. Riphahn (2005). "The Effect of Employment Protection on Worker Effort - A Comparison of Absenteeism During and After Probation." *Journal of the European Economic Association*, 3, 120-143.
- [14] Iossa, Elisabetta and Patrick Rey (2010). "Building Reputation for Contract Renewal: Implications for Performance Dynamic and Contract duration." CEIS Research Paper n° 155, Tor Vergata University.
- [15] Joskow, Paul (1987). "Contract Duration and Relationship-Specific Investments: Empirical Evidence from Coal Markets." *American Economic Review*, 77, 168-85.
- [16] Kelman, Steven (1990). *Procurement and Public Management: The Fear of Discretion and the Quality of Government Performance*. AEI Press.
- [17] Klein, Benjamin and Keith B. Leffler (1981). "The Role of Market Forces in Assuring Contractual Performance." *Journal of Political Economy*, 89, 615-641.
- [18] Laffont, Jean Jacques and Jean Tirole (1993). *A Theory of Incentives in Procurement and Regulation*. MIT Press.
- [19] Lambert, Richard (1983). "Long-term Contracts and Moral Hazard." *Journal of Economics*, 14, 441-452.
- [20] Lazear, Edward (2004). "The Peter's Principle: A Theory of Decline." *Journal of Political Economy*, Papers in Honor of Sherwin Rosen: A Supplement to Volume 112, S141-S163.
- [21] Lewis, Tracy (1986). "Reputation and Contractual Performance in Long-term Projects." *Rand Journal of Economics*, 17, 141-157.
- [22] Macleod, Bentley (2007). "Reputations, Relationships, and Contract Enforcement." *Journal of Economic Literature*, 45, 595-628.
- [23] Mailath, George and Larry Samuelson (2001). "Who Wants a Good Reputation?." *Review of Economics Studies*, 68, 415-441.

- [24] Martinez, Leonardo (2009). “A theory of Political Cycles.” *Journal of Economic Theory*, 144, 1166-1186.
- [25] Masten, Scott E. and Keith J., Crocker (1985). “Efficient Adaptation in Long Term Contracts: Take-or-pay Provisions for Natural Gas.” *The American Economic Review*, 75, 1083-1093.
- [26] Phelan, Christopher (2006). “Public Trust and Government Betrayal?.” *Journal of Economic Theory*, 130, 27-43.
- [27] Pint, Ellen (1992). “Price-Cap versus Rate-of-Return Regulation in a Stochastic-Cost Model.” *RAND Journal of Economics*, 23, 564-578.
- [28] Rice, Arup and J. Bradford Sen (2008). “Moral hazard in long-term guaranteed contracts: theory and evidence from the NBA.” mimeo, Boston University.
- [29] Rogoff, Kenneth (1990), “Equilibrium Political Budget Cycle.” *American Economic Review*, 80, 21-36.
- [30] Saussier, Stephane (1999), “Transaction Cost Economics and Contract Duration: An Empirical Analysis of EDF Coal Contracts.” *Louvain Economic Review*, 65, 3-21.
- [31] Shapiro, Carl (1983). “Premiums for High Quality Products as Returns on Reputation.” *Quarterly Journal of Economics*, 98, 659-80.
- [32] Spagnolo, Giancarlo (2012). “Reputation, Competition and Entry in Procurement.” *International Journal of Industrial Organization*, 30(3), 291-296.
- [33] Stiroh, Kevin J. (2007). “Playing for Keeps: Pay and Performance in the NBA.” *Economic Inquiry*, 45, 145–161.
- [34] Strausz, Roland (2009). “Planned Obsolescence and the Provision of Unobservable Quality.” *The Economic Journal*, 119, 1405-1421.
- [35] Taylor, Curtis, and Steven Wiggings (1997). “Competition or Compensation: Supplier Incentives Under the American and Japanese Subcontracting Systems.” *The American Economic Review*, 87, 598-618.

# Online Appendix

## Not for publication

### A Alternative baseline model without adverse selection

We consider here a variant of our baseline model in which the type  $\theta_t$  is not observable by the agent either; in each period, the agent thus bases his investment decision only on his performance history.

As before, enhancing performance in period 2 from  $q_2 = L$  to  $q_2 = H$  increases the agent's expected payoff by  $\Delta\Pi_3 = (2\rho - 1)\delta\Delta$ . Consider now period 2, following a given investment decision  $\hat{i}_1 \in \{0, 1\}$  taken in period 1 (here  $\hat{i}_1$  denotes the decision eventually taken by the agent in period ; it is equal to 1 with probability  $i_1$  and to 0 with probability  $1 - i_1$ ), a given performance  $q_1$  observed in that period (which is thus such that  $q_1 = H$  if  $\theta_1 = H$  or  $\hat{i}_1 = 1$ , and  $q_1 = L$  otherwise).

If  $q_1 = L$ , then the agent and the market infer  $\theta_1 = L$ , and thus anticipate  $\theta_2 = L$  with probability  $\rho$ . The agent thus chooses  $i_2$  so as to maximize:

$$\begin{aligned}\Pi_2^L &= -ci_2 + \delta \{(1 - \rho + \rho i_2)(L + \rho\Delta) + \rho(1 - i_2)(L + (1 - \rho)\Delta)\} \\ &= \delta(L + \rho\Delta) - ci_2 - \rho(1 - i_2)(2\rho - 1)\delta\Delta \\ &= \delta(L + \rho\Delta) - (2\rho - 1)\rho\delta\Delta + (b_2^L - c)i_2,\end{aligned}\tag{14}$$

where:

$$b_2^L \equiv (2\rho - 1)\rho\delta\Delta.\tag{15}$$

The optimal investment is thus  $i_2^L = 1$  if  $c < b_2^L$ , and  $i_2^L = 0$  if  $c > b_2^L$ .

If instead  $q_1 = H$ , then the agent and the market anticipate  $\theta_2 = H$  with probability  $\rho$ . The agent thus chooses  $i_2$  so as to maximize:

$$\begin{aligned}\Pi_2^H &= -ci_2 + \delta \{(\rho + (1 - \rho)i_2)(L + \rho\Delta) + (1 - \rho)(1 - i_2)(L + (1 - \rho)\Delta)\} \\ &= \delta(L + \rho\Delta) - ci_2 - (1 - \rho)(1 - i_2)(2\rho - 1)\delta\Delta \\ &= \delta(L + \rho\Delta) - (2\rho - 1)(1 - \rho)\delta\Delta + (b_2^H - c)i_2,\end{aligned}\tag{16}$$

where:

$$b_2^H \equiv (2\rho - 1)(1 - \rho)\delta\Delta.\tag{17}$$

The optimal investment is thus  $i_2^H = 1$  if  $c < b_2^H$ , and  $i_2^H = 0$  if  $c > b_2$ . Note that  $b_2^H < b_2^L$  whenever there is information decay ( $\rho < 1$ ):

$$b_2^L - b_2^H = (2\rho - 1)^2 \delta \Delta > 0.$$

Therefore, substituting for the equilibrium values of  $i_2$  in (14) and (16):

- If  $c < b_2^H$ , the agent always invests in period 2:  $i_2^H = i_2^L = 1$ ; his payoff is then:

$$\Pi_2^L = \Pi_2^H = \delta(L + \rho\Delta) - c.$$

- If instead  $b_2^H < c < b_2^L$ , the agent invests in period 2 only when  $q_1 = L$ :  $i_2^H = 0 < i_2^L = 1$ ; his payoff is then:

$$\begin{aligned} \Pi_2^L &= \delta(L + \rho\Delta) - c, \\ \Pi_2^H &= \delta(L + (1 - 2\rho(1 - \rho))\Delta). \end{aligned}$$

- Finally, if  $c > b_2^L$ , the agent never invests in period 2:  $i_2^H = i_2^L = 0$ ; his payoff is then:

$$\begin{aligned} \Pi_2^L &= \delta(L + 2\rho(1 - \rho)\Delta), \\ \Pi_2^H &= \delta(L + (1 - 2\rho(1 - \rho))\Delta). \end{aligned}$$

Consider now period 1. The agent chooses  $i_1$  so as to maximize:

$$\Pi_1 = -ci_1 + \delta \left\{ \frac{1 + i_1}{2} \Pi_2^H + \frac{1 - i_1}{2} \Pi_2^L \right\} = \delta \frac{\Pi_2^H + \Pi_2^L}{2} + (b_1 - c) i_1.$$

where

$$b_1 \equiv \delta \frac{\Pi_2^H - \Pi_2^L}{2}. \quad (18)$$

The expected payoff differential  $\Pi_2^H - \Pi_2^L$  is by construction non-negative<sup>28</sup> and, using (14) and (16), can be expressed as:

$$\begin{aligned} \Pi_2^H - \Pi_2^L &= (2\rho - 1)^2 \delta \Delta + (b_2^H - c) i_2^H - (b_2^L - c) i_2^L \\ &= (1 - i_2^L) (2\rho - 1)^2 \delta \Delta + (c - b_2^H) (i_2^L - i_2^H). \end{aligned}$$

Therefore:

---

<sup>28</sup>We have:

$$\begin{aligned} \Pi_2^L &= \max_{i_2} \{ \delta(L + \rho\Delta) - ci_2 - \rho(1 - i_2)(2\rho - 1)\delta\Delta \} \\ &\leq \max_{i_2} \{ \delta(L + \rho\Delta) - ci_2 - (1 - \rho)(1 - i_2)(2\rho - 1)\delta\Delta \} = \Pi_2^H. \end{aligned}$$



- If  $c < b_2^H$ , the agent always invests in period 2:  $i_2^H = i_2^L = 1$ , and thus  $\Pi_2^L = \Pi_2^H$ ; hence, the agent does not invest in period 1:  $i_1^* = 0$ .

- If instead  $b_2^H < c < b_2^L$ , the agent invests in period 2 only when  $q_1 = L$ :  $i_2^H = 0 < i_2^L = 1$ , and thus:

$$b_1 = \delta \frac{c - b_2^H}{2} < c,$$

implying again that the agent does not invest in period 1:  $i_1^* = 0$ .

- Finally, if  $c > b_2^L$ , the agent never invests in period 2:  $i_2^H = i_2^L = 0$ , and thus:

$$b_1 = \frac{(2\rho - 1)^2 \delta^2 \Delta}{2} < b_2^L = (2\rho - 1) \rho \delta \Delta, \quad (19)$$

implying again that the agent does not invest in period 1 (as here  $c > b_2^L > b_1$ ).

This leads to:

**Proposition 8** (i) *The agent never invests in period 1, but invests in period 2 if the cost is sufficiently low:*

$$i_t^* = 1 \text{ if and only if } t = 2 \text{ and } c \leq b_2^{q_1},$$

where  $q_1$  denotes the performance observed in period 1 and

$$b_2^L \equiv (2\rho - 1) \rho \delta \Delta > b_2^H \equiv (2\rho - 1) (1 - \rho) \delta \Delta > 0.$$

(ii) *The incentive to invest in period 2 therefore increases with the discount factor  $\delta$  and the performance differential  $\Delta$ , relative to the cost  $c$ ; it however decreases with the performance observed in the first period.*

(iii) *Following a bad performance, the incentive to invest in period 2 increases with the degree of information persistence  $\rho$ ; following a good performance, however, it increases with information persistence only when this is initially low (namely, in the range  $\rho < 3/4$ ).*

The insights of our baseline model thus carry over. Part (i), for instance, follows directly from the fact that the performance in period 2 is a sufficient statistic for predicting the performance in period 3. Interestingly, however, the incentive to invest increases over time even in the absence of information decay or of discounting: there is an additional “option value” in waiting and observing  $q_1$  before deciding whether to invest.<sup>29</sup> Likewise, in part (ii) the comparative

---

<sup>29</sup>This is reflected in the term  $\rho - \frac{1}{2} < \rho$  in (19), rewritten as:

$$b_1 = \delta \left( \rho - \frac{1}{2} \right) (2\rho - 1) \delta \Delta < b_2^L = (2\rho - 1) \rho \delta \Delta.$$

statics on  $\delta$ ,  $\Delta$  and  $c$  build on the same intuition as before; in addition, incentives decrease with existing reputation as the payoff from investing (net of  $c$ ) is the same regardless of past performance, whilst the payoff from not investing is lower when past performance is low. Part (iii) follows from information persistence raising the payoff from investing, and lowering (resp., raising) the payoff from not investing when past performance is low (resp., high).

An implication of Proposition 8 is that, on average, the agent's performance improves as the market evaluation approaches: the expected performance in period 1 is equally likely to be  $H$  or  $L$ , as the agent never invests in that period; in the second period, the performance is more likely to be good when the agent invests with positive probability, which is the case when  $c < b_2^L$ . Therefore:

**Corollary 2** *The expected performance (weakly) increases as the market evaluation approaches:  $q_1^e \leq q_2^e$ , with a strict inequality when  $c < b_2^L$ .*

**Proof.** The agent's productivity is initially high or low with equal probability. As he never invests in period 1, his expected performance is:

$$q_1^e = \frac{L + H}{2} = L + \frac{\Delta}{2}.$$

By contrast, in period 2, the performance is low only if the agent's type is low *and* he does not invest; as a result, the expected performance is

$$\begin{aligned} q_2^e &= \frac{[\rho + (1 - \rho) i_2^H] H + (1 - \rho) (1 - i_2^H) L}{2} + \frac{[1 - \rho + \rho i_2^L] H + \rho (1 - i_2^L) L}{2} \\ &= L + (1 + (1 - \rho) i_2^H + \rho i_2^L) \frac{\Delta}{2}. \end{aligned}$$

The conclusion ( $q_2^e \geq q_1^e$ ) follows from  $i_2^H, i_2^L \geq 0$  (with a strict inequality for  $i_2^L$  when  $c > b_2^L$ ).

■

## B Proof of Proposition 6

As already noted: (i) the agent never invests in period 2, as exerting effort is less costly and produces the same benefit on his expected payoff; and (ii) in period 1, the agent may either invest or exert effort, but never does both. What remains to be determined is: (i) the probability  $e_q$  of effort in period 2, as a function of the performance  $q \in \{L, H\}$  observed in period 1; and (ii) the probability of effort  $e_1$ , or of investment  $i_1$ , in the first period.

## B.1 No information decay ( $\rho = 1$ )

We first consider the case where  $\rho = 1$ . When  $\theta_1 = H$ , the agent then remains productive forever and thus never needs to invest or exert effort. If instead  $\theta_1 = L$ , and the agent neither invests nor exerts efforts in period 1 then, observing  $q = L$  in period 1 reveals that the agent has a low productivity; as investing is not credible in period 2, the market will then anticipate that the agent will still have a low productivity in period 3, and thus the agent never exerts effort in period 2. The agent can instead enhance his performance and the market beliefs, either by investing in period 1, or by exerting effort in both periods (exerting effort in period 1 only would be a dominated strategy, as a low performance in period 2 would reveal his low productivity).

In a candidate equilibrium in which the agent invests in period 1, the market then believes that the agent is highly productive and the gain from investment is  $\delta^2\Delta - c$ . As the agent could deviate either by exerting effort in both periods or do nothing, it follows that investing constitutes an equilibrium strategy if

$$\delta^2\Delta - c \geq 0, \delta^2\Delta - (1 + \delta)\gamma.$$

In a candidate equilibrium in which the agent exerts effort with probability  $e_1$  in period 1, the market then remains uncertain about the agent's productivity and the gain from investment is  $\delta^2\Delta/(1 + e_1) - (1 + \delta)\gamma$ . As the agent could deviate either by investing or by doing nothing, it follows that exerting effort constitutes an equilibrium strategy if

$$\frac{\delta^2\Delta}{1 + e_1} - (1 + \delta)\gamma \geq 0, \frac{\delta^2\Delta}{1 + e_1} - c.$$

Finally, in a candidate equilibrium in which the agent neither invests nor exerts effort, by deviating the agent would convince the market that he has a high productivity; therefore, doing nothing constitutes an equilibrium strategy if

$$0 \geq \delta^2\Delta - c, \delta^2\Delta - (1 + \delta)\gamma.$$

Comparing these conditions, we obtain that:

- if  $c < (1 + \delta)\gamma$ , then the agent never exerts effort, and invests in period 1 when he has a low productivity if  $c < \delta^2\Delta$  (in the limit case where  $c = \delta^2\Delta$ , then the agent invests with any probability  $i_1 \in [0, 1]$ );
- if  $c > (1 + \delta)\gamma$ , then the agent never invests, and exerts effort in both periods when he has a low productivity if  $\gamma < \frac{\delta^2\Delta}{2(1+\delta)}$  (in the limit case where  $\gamma = \frac{\delta^2\Delta}{2(1+\delta)}$ , then the agent

exerts effort in period 1 with any probability  $e_1 \in [0, 1]$ , in which case he exerts effort also in period 2);

- if  $c = (1 + \delta) \gamma$ , then:
  - if  $c < \delta^2 \Delta$ , there exists an equilibrium in which the agent invests in period 1 (or with any probability  $i_1^* \in [0, 1]$  in the limit case where  $c = \delta^2 \Delta$ );
  - if  $\gamma < \frac{\delta^2 \Delta}{1 + \delta}$ , there exists an equilibrium in which the agent exerts effort in period 1 with some probability  $e_1^* > 0$ , in which case he exerts effort also in period 2 (if  $\gamma < \frac{\delta^2 \Delta}{2(1 + \delta)}$ , then  $e_1^* = 1$ );
- in all cases, the performance remains constant over time:  $q_1 = q_2 = H$  if the agent invests in period 1 or exerts effort in both periods, and  $q_1 = q_2 = L$  otherwise.

## B.2 Information decay ( $\rho < 1$ )

We now turn to the case of information decay:  $\rho \in (1/2, 1)$ . In period 2 a low-productivity agent exerts effort if condition (10) holds, which in equilibrium amounts to:

$$\gamma \leq G_q(e_q) \equiv (2\rho - 1) \delta \nu_q(e_q) \Delta, \quad (20)$$

where  $\nu_q(e_q)$ , the equilibrium market belief following  $q \in \{L, H\}$  in period 1 and  $H$  in period 2, is given by:<sup>30</sup>

$$\nu_H(e_H) \equiv \frac{\rho + (1 - \rho) e_1}{\rho + (1 - \rho) e_1 + (1 - \rho + \rho e_1) e_H}, \quad (21)$$

$$\nu_L(e_L) \equiv \frac{1 - \rho}{1 - \rho + \rho e_L}. \quad (22)$$

As  $\nu_q$  decreases when  $e_q$  increases, the equilibrium effort of a low-productivity agent in period 2, following a performance  $q \in \{L, H\}$  in period 1, is given by:

$$e_q^* = \begin{cases} 0 & \text{if } \gamma \geq G_q(e_q = 0), \\ 1 & \text{if } \gamma \leq G_q(e_q = 1), \\ \hat{e}_q \text{ s.t. } G_q(\hat{e}_q) = \gamma & \text{otherwise.} \end{cases} \quad (23)$$

Furthermore,  $\nu_H(0) = \nu_L(0)$  and, for  $e > 0$ ,  $\nu_H(e) > \nu_L(e)$ :

$$\frac{\nu_H(e)}{\nu_L(e)} = \frac{1 + \frac{\rho}{1 - \rho} e}{1 + \frac{1 - \rho + \rho e_1}{\rho + (1 - \rho) e_1} e} > 1, \quad (24)$$

---

<sup>30</sup> $\nu_H$  and thus  $G_H$  depend on both  $e_H$  and  $e_1$ ; when there is no risk of confusion, we drop the argument  $e_1$  to simplify the exposition.

where the inequality stems from  $\frac{1-\rho+\rho e_1}{\rho+(1-\rho)e_1} < \frac{\rho}{1-\rho}$  (which amounts to  $(1-\rho)^2 < \rho^2$ ). Therefore,  $G_H(0) = G_L(0) = (2\rho - 1)\delta\Delta$  and  $G_H(e) > G_L(e)$  for  $e > 0$ . It follows that: (i)  $e_H^*$  and  $e_L^*$  are both positive when  $\gamma < (2\rho - 1)\delta\Delta$ , and they are both equal to 0 if  $\gamma \geq (2\rho - 1)\delta\Delta$ ; (ii)  $e_H^* \geq e_L^*$ , with a strict inequality whenever  $0 < e_L^* < 1$ .

Next, we note that the agent neither invests nor exerts effort in period 1 when, in case of a high performance in period 1, it has little incentive to exert effort in period 2:

**Lemma 2** *Suppose  $e_H^* < 1$ . Then  $e_1^* = i_1^* = 0$ .*

**Proof.** To see this, note first that  $e_H^* < 1$  implies  $e_L^* (\leq e_H^*) < 1$ ; therefore, in period 2, and for any given performance  $q \in \{L, H\}$  observed in period 1, the equilibrium market belief  $\nu_q^*$  must be such that agent either does not exert effort (whenever  $e_q^* = 0$ ), or is indifferent between exerting effort or not (if  $e_q^* = \hat{e}_q$ ); it follows that in period 1, the expected payoffs from investing, exerting effort, and doing nothing are respectively given by:

$$\begin{aligned}\Pi_1^L(i_1 = 1, e_1 = 0) &= \delta^2 [L + (1 - \rho)\Delta] + \rho\nu_H^* (2\rho - 1)\delta^2\Delta - c, \\ \Pi_1^L(i_1 = 0, e_1 = 1) &= \delta^2 [L + (1 - \rho)\Delta] + (1 - \rho)\nu_H^* (2\rho - 1)\delta^2\Delta - \gamma, \\ \Pi_1^L(i_1 = e_1 = 0) &= \delta^2 [L + (1 - \rho)\Delta] + (1 - \rho)\nu_L^* (2\rho - 1)\delta^2\Delta.\end{aligned}$$

Therefore, the benefit from investing or exerting effort cannot exceed

$$\rho\nu_H^* (2\rho - 1)\delta^2\Delta < \nu_H^* (2\rho - 1)\delta\Delta,$$

where the right-hand side does not exceed  $\gamma$  (and thus  $c$ ) when  $e_H^* < 1$ . ■

As  $e_H^* < 1$  implies  $e_1^* = 0$ , this requires:

$$\gamma > G_H(e_H = 1)|_{e_1^*=0} = \rho(2\rho - 1)\delta\Delta. \quad (25)$$

Conversely, under (25), we must have  $e_1^* = 0$ : to see this, note that

$$G_H(e_H = 1) = (2\rho - 1)\delta\nu_H(e_H = 1)\Delta,$$

where:

$$\nu_H(e_H = 1) = \frac{\rho + (1 - \rho)e_1}{1 + e_1}$$

decreases as  $e_1$  increases; but then  $e_1^* > 0$  would imply  $e_H^* = 1$  and thus require  $\gamma < G_H(e_H = 1) \leq G_H(e_H = 1)|_{e_1^*=0} = \rho(2\rho - 1)\delta\Delta$ , a contradiction. It follows that, under (25), there exists a unique equilibrium, in which the agent (i) never invests, (ii) never exerts effort in period 1, and (iii) in period 2:

- if  $\gamma \geq (2\rho - 1) \delta \Delta$ , the agent never exerts effort;
- if  $(2\rho - 1) \delta \Delta > \gamma > \rho(2\rho - 1) \delta \Delta$  then, given the performance  $q \in \{L, H\}$  observed in period 1, the agent exerts effort with probability  $\hat{e}_q$  in case of low productivity.

From now on, we focus on the case  $\gamma \leq \rho(2\rho - 1) \delta \Delta$ . In period 2 we must then have

- $e_H^* = 1$ :  $e_H^* < 1$  would imply  $e_1^* = 0$  and thus require  $\gamma > G_H(e_H = 1)|_{e_1^*=0} = \rho(2\rho - 1) \delta \Delta$ , a contradiction.
- $e_L^* > 0$ :  $e_H^* = 1 > 0$  implies  $G_L(0) = G_H(0) > \gamma$ ; we thus have:
  - if  $\gamma \leq G_L(1) = (1 - \rho)(2\rho - 1) \delta \Delta$ ,  $e_L^* = 1$ ;
  - otherwise  $e_L^* = \hat{e}_L > 0$  is such that  $\gamma = G_L(\hat{e}_L) = (2\rho - 1) \delta \nu_L(\hat{e}_L) \Delta$ , or:

$$\gamma = \frac{1 - \rho}{1 - \rho + \rho \hat{e}_L} (2\rho - 1) \delta \Delta \iff \hat{e}_L = \frac{1 - \rho}{\rho} \left( \frac{(2\rho - 1) \delta \Delta}{\gamma} - 1 \right). \quad (26)$$

In period 1, the comparison between the expected payoffs from investing and from exerting effort, given respectively by (12) and (13) shows, using  $e_H^* = 1$ , that the agent favors effort over investment when:

$$c > [1 + (2\rho - 1) \delta] \gamma, \quad (27)$$

and favors instead investment over effort when the inequality is reversed.

Suppose first that (27) holds, in which case the agent never invests ( $i_1^* = 0$ ). Comparing (11) and (13), in period 1 the agent then exerts effort if:

$$\gamma \leq \frac{(1 - \rho + \rho e_H^*) \nu_H^* - (1 - \rho + \rho e_L^*) \nu_L^*}{1 + (e_H^* - e_L^*) \rho \delta} (2\rho - 1) \delta^2 \Delta,$$

which, using  $e_H^* = 1$  and (22), can be written as:

$$\gamma \leq G_E(e_1; e_L^*) \equiv \frac{\nu_H^*(e_1) - (1 - \rho)}{1 + (1 - e_L^*) \rho \delta} (2\rho - 1) \delta^2 \Delta,$$

where:

$$\nu_H^*(e_1) \equiv \nu_H(e_H^* = 1; e_1) = \frac{\rho + (1 - \rho) e_1}{1 + e_1}.$$

As  $\nu_H^*$ , and thus  $G_E$ , decreases as  $e_1$  increases, it follows that  $e_1^* > 0$  when (using  $\nu_H^*(0) = \rho$ ):

$$\gamma < G_E(0; e_L^*) = \frac{(2\rho - 1)^2 \delta^2 \Delta}{1 + (1 - e_L^*) \rho \delta}.$$

Two cases can be distinguished, depending on the value of  $e_L^*$ :

- If  $\gamma \leq G_L(1) = (1 - \rho)(2\rho - 1)\delta\Delta$ , then  $e_L^* = 1$  and :

$$G_E(0; e_L^* = 1) = (2\rho - 1)^2 \delta^2 \Delta,$$

which lies below  $(1 - \rho)(2\rho - 1)\delta\Delta$  if and only if  $(2\rho - 1)\delta < 1 - \rho$ , or

$$\rho < \hat{\rho} \equiv \frac{1 + \delta}{1 + 2\delta}.$$

- If  $(1 - \rho)(2\rho - 1)\delta\Delta < \gamma \leq (2\rho - 1)\delta\Delta$ , then  $e_L^* = \hat{e}_L$  and:

$$\gamma < G_E(0; e_L^* = \hat{e}_L) = \frac{(2\rho - 1)^2 \delta^2 \Delta}{1 + \delta \rho \left(1 - \frac{1 - \rho}{\rho} \left(\frac{(2\rho - 1)\delta\Delta}{\gamma} - 1\right)\right)} \iff \gamma < \frac{\rho(2\rho - 1)\delta^2 \Delta}{1 + \delta},$$

where the right-hand side is lower than  $\rho(2\rho - 1)\delta\Delta$ , and lies above  $(1 - \rho)(2\rho - 1)\delta\Delta$  if and only if  $\rho > \hat{\rho}$ :

$$\frac{\rho(2\rho - 1)\delta^2 \Delta}{1 + \delta} - (1 - \rho)\delta(2\rho - 1)\Delta = (\rho - \hat{\rho})\delta(2\rho - 1)\Delta \frac{1 + 2\delta}{1 + \delta}.$$

Summing-up, we have:  $e_1^* > 0$  if

$$\gamma < \hat{\gamma} \equiv \begin{cases} (2\rho - 1)\delta^2 \Delta & \text{if } \rho < \hat{\rho} \\ \frac{\rho(2\rho - 1)\delta^2 \Delta}{1 + \delta} & \text{if } \rho \geq \hat{\rho} \end{cases},$$

and  $e_1^* = 0$  if  $\gamma \geq \hat{\gamma}$ .

Likewise, we have  $e_1^* = 1$  when

$$\gamma \leq G_E(1; e_L^*) = \frac{(2\rho - 1)^2 \delta^2 \Delta}{1 + (1 - e_L^*)\rho\delta}.$$

Therefore:

- If  $\gamma \leq G_L(1) = (1 - \rho)(2\rho - 1)\delta\Delta$ , then  $e_L^* = 1$  and:

$$G_E(1; e_L^* = 1) = (2\rho - 1)^2 \delta^2 \frac{\Delta}{2},$$

which lies below  $(1 - \rho)(2\rho - 1)\delta\Delta$  if and only if

$$\rho < \tilde{\rho} \equiv \frac{2 + \delta}{2(1 + \delta)},$$

as can be seen from:

$$\begin{aligned} (2\rho - 1)^2 \delta^2 \frac{\Delta}{2} - (1 - \rho)(2\rho - 1)\delta\Delta &= [(2\rho - 1)\delta - 2(1 - \rho)] \frac{(2\rho - 1)\delta\Delta}{2} \\ &= (\rho - \tilde{\rho})(2\rho - 1)\delta(1 + \delta)\Delta, \end{aligned}$$

- If  $(1 - \rho)(2\rho - 1)\delta\Delta < \gamma \leq (2\rho - 1)\delta\Delta$ , then  $e_L^* = \hat{e}_L$  and:

$$\gamma < G_E(1; e_L^* = \hat{e}_L) = \frac{(2\rho - 1)^2 \delta^2}{1 + \delta\rho \left(1 - \frac{1-\rho}{\rho} \left(\frac{(2\rho-1)\delta\Delta}{\gamma} - 1\right)\right)} \frac{\Delta}{2} \iff \gamma < (2\rho - 1) \frac{\delta^2 \Delta}{2(1 + \delta)},$$

where the right-hand side is lower than  $\rho(2\rho - 1)\delta\Delta$ , and lies above  $(1 - \rho)(2\rho - 1)\delta\Delta$  if and only if  $\rho > \tilde{\rho}$ :

$$(2\rho - 1) \frac{\delta^2 \Delta}{2(1 + \delta)} - (1 - \rho)(2\rho - 1)\delta\Delta = (\rho - \tilde{\rho})(2\rho - 1)\delta\Delta.$$

Summing-up, we have:  $e_1^* = 1$  if:

$$\gamma \leq \tilde{\gamma} \equiv \begin{cases} (2\rho - 1) \frac{\delta^2 \Delta}{2} & \text{if } \rho < \tilde{\rho} \\ (2\rho - 1) \frac{\delta^2 \Delta}{1 + \delta} & \text{if } \rho \geq \tilde{\rho} \end{cases},$$

and  $e_1^* < 1$  if  $\gamma > \tilde{\gamma}$ .

Suppose now that:

$$c < [1 + (2\rho - 1)\delta]\gamma, \quad (28)$$

in which case the agent never exerts effort in period 1 (i.e.,  $e_1^* = 0$ ), as he favors investment over effort in that period. We thus have:

$$\nu_H^* = \frac{\rho}{\rho + (1 - \rho)e_H^*}, \nu_L^* = \frac{1 - \rho}{1 - \rho + \rho e_L^*}.$$

Using this and  $e_H^* = 1$ , in period 1 the expected payoffs from investing and from not investing are respectively given by:

$$\Pi_1^L(i_1 = e_1 = 0) = \delta^2 [L + (1 - \rho)\Delta] + (1 - \rho)(2\rho - 1)\delta^2 \Delta - \delta\rho\gamma e_L^*,$$

$$\Pi_1^L(i_1 = 1, e_1 = 0) = -c + \delta^2 [L + (1 - \rho)\Delta] + \rho(2\rho - 1)\delta^2 \Delta - \delta(1 - \rho)\gamma.$$

Therefore, the agent indeed invests in period 1 if:

$$c \leq G_I(e_L^*) \equiv (2\rho - 1)^2 \delta^2 \Delta + [\rho e_L^* - (1 - \rho)]\delta\gamma. \quad (29)$$

Two cases can again be distinguished, depending on the value of  $e_L^*$ :

- If  $\gamma < G_L(1) = (1 - \rho)(2\rho - 1)\delta\Delta$ , then  $e_L^* = 1$  and (29) becomes:

$$c \leq G_I(e_L^* = 1) = (2\rho - 1)^2 \delta^2 \Delta + (2\rho - 1)\delta\gamma.$$

This condition is less demanding than (28) as long as:

$$\gamma < (2\rho - 1)^2 \delta^2 \Delta.$$

Hence, in the range  $\gamma < (1 - \rho)(2\rho - 1)\delta\Delta$ :



- if  $\rho \geq \hat{\rho}$ , which implies  $1 - \rho < (2\rho - 1)\delta$ , (29) is always less demanding than (28);
- if  $\rho < \hat{\rho}$ , (29) is less demanding than (28) as long as  $\gamma < (2\rho - 1)^2 \delta^2 \Delta$ , and more demanding for  $(2\rho - 1)^2 \delta^2 \Delta < \gamma < (1 - \rho)(2\rho - 1)\delta\Delta$ .

- If  $(1 - \rho)(2\rho - 1)\delta\Delta < \gamma \leq (2\rho - 1)\delta\Delta$ , then  $e_L^* = \hat{e}_L$  and (29) becomes:

$$c \leq G_I(e_L^* = \hat{e}_L) = \delta^2(2\rho - 1)\rho\Delta - 2(1 - \rho)\delta\gamma.$$

As the right-hand side decreases when  $\gamma$  increases, this condition is more demanding than (28) for

$$\gamma > \frac{\rho(2\rho - 1)\delta^2\Delta}{1 + \delta},$$

where the right-hand side is lower than  $(2\rho - 1)\delta\Delta$  and, as already noted, lies above (resp., below)  $(1 - \rho)(2\rho - 1)\delta\Delta$  if  $\rho > \hat{\rho}$  (resp.,  $\rho < \hat{\rho}$ ). Hence, in the range  $(1 - \rho)(2\rho - 1)\delta\Delta < \gamma \leq (2\rho - 1)\delta\Delta$ :

- if  $\rho \geq \hat{\rho}$ , (29) is less demanding than (28) as long as  $(1 - \rho)(2\rho - 1)\delta\Delta < \gamma < \frac{\rho(2\rho - 1)\delta^2\Delta}{1 + \delta}$ , and more demanding for  $\frac{\rho(2\rho - 1)\delta^2\Delta}{1 + \delta} < \gamma < (2\rho - 1)^2 \delta^2 \Delta$ ;
- if  $\rho < \hat{\rho}$ , (29) is always more demanding than (28).

Summing-up:

- If  $\gamma \geq (2\rho - 1)\delta\Delta$ , the agent never invests nor exerts effort.
- If  $(2\rho - 1)\delta\Delta > \gamma > \rho(2\rho - 1)\delta\Delta$  then the agent (i) never invests, (ii) never exerts effort in period 1, and (iii) in period 2, given the performance  $q \in \{L, H\}$  observed in period 1, the agent exerts effort with probability  $\hat{e}_q$  in case of low productivity.
- If  $\gamma < \rho(2\rho - 1)\delta\Delta$ , then the agent exerts effort with positive probability in case of low productivity in period 2:  $e_H^* = 1$  and  $e_L^* > 0$ ; in addition, in case of low productivity in period 1:
  - If  $\rho > \hat{\rho}$ , then the agent (see Figure 1a):

- \* invests ( $i_1^* = 1$ ) if

$$\frac{c}{\gamma} < 1 + (2\rho - 1)\delta \text{ and } c < \rho(2\rho - 1)\delta^2\Delta - 2(1 - \rho)\delta\gamma,$$

\* exerts instead effort with positive probability ( $e_1^* > 0$ ) if

$$\frac{c}{\gamma} > 1 + (2\rho - 1)\delta \text{ and } \gamma < \frac{\rho(2\rho - 1)\delta^2\Delta}{1 + \delta},$$

with  $e_1^* = 1$  if  $\gamma \leq \tilde{\gamma}$ , and  $e_1^* < 1$  if  $\gamma > \tilde{\gamma}$ ,

\* neither invest nor exert effort otherwise.

– If  $\rho < \hat{\rho}$ , then the agent (see Figure 1b):

\* invests ( $i_1^* = 1$ ) if  $c/\gamma < 1 + (2\rho - 1)\delta$  and

$$c < \max \{ (2\rho - 1)^2 \delta^2 \Delta + (2\rho - 1)\delta\gamma, \rho(2\rho - 1)\delta^2\Delta - 2(1 - \rho)\delta\gamma \},$$

\* exerts effort with positive probability ( $e_1^* > 0$ ) if  $c/\gamma > 1 + (2\rho - 1)\delta$  and

$$\gamma < (2\rho - 1)^2 \delta^2 \Delta,$$

with  $e_1^* = 1$  if  $\gamma \leq \tilde{\gamma}$ , and  $e_1^* < 1$  if  $\gamma > \tilde{\gamma}$ ,

\* and neither invest nor exert effort otherwise.

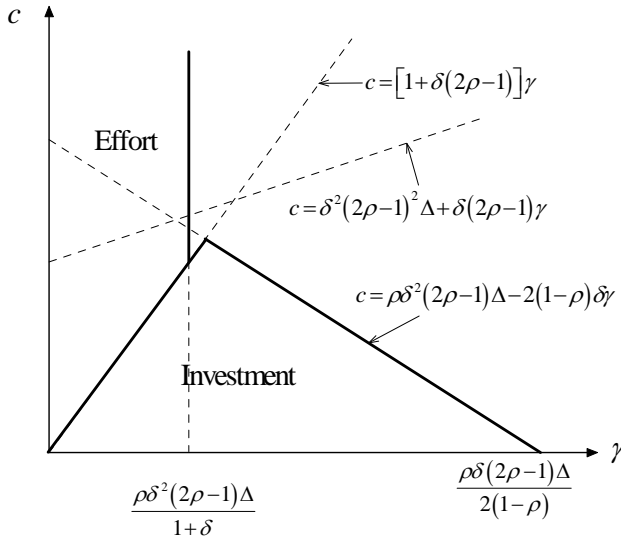


Figure 1a:  $\rho > \hat{\rho}$

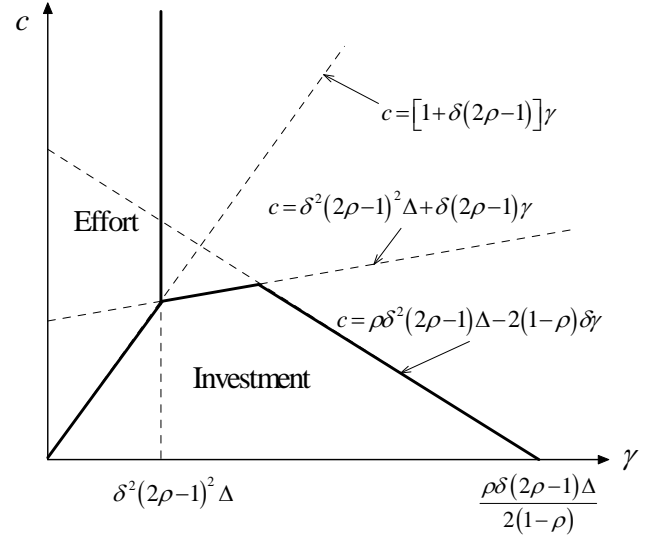


Figure 1b:  $\rho < \hat{\rho}$

The comparative statics follow. In particular, it can be checked that the region in which there is either investment or effort in case of low productivity in period 1 expands as either  $\delta$  or  $\Delta$  increases.

We now compare the expected performance in the two periods. Obviously, the expected performance remains unchanged and equal to  $q_t^e = \frac{L+H}{2} = L + \frac{\Delta}{2}$  in the absence of any investment or effort, which occurs when  $\gamma \geq (2\rho - 1) \delta \Delta$ . Let us now focus on the case where  $\gamma < (2\rho - 1) \delta \Delta$ , in which case  $e_H^*, e_L^* > 0$ .

- If  $e_1^* = i_1^* = 0$ , then

$$\begin{aligned} q_1^e &= L + \frac{\Delta}{2} \\ q_2^e &= L + \frac{1}{2} [\rho + (1 - \rho) e_H^*] \Delta + \frac{1}{2} (1 - \rho + \rho e_L^*) \Delta \\ &= L + [1 + (1 - \rho) e_H^* + \rho e_L^*] \frac{\Delta}{2}. \end{aligned}$$

Therefore:

$$q_2^e - q_1^e = [(1 - \rho) e_H^* + \rho e_L^*] \frac{\Delta}{2} > 0,$$

where the inequality follows from  $e_H^*, e_L^* > 0$ .

- If  $e_1^* > i_1^* = 0$ , then  $e_H^* = 1$  and  $e_L^* > 0$ ; therefore:

$$\begin{aligned} q_1^e &= L + \frac{1 + e_1^*}{2} \Delta, \\ q_2^e &= L + \frac{1}{2} \Delta + \frac{1}{2} [1 - \rho + \rho (e_1^* + (1 - e_1^*) e_L^*)] \Delta, \end{aligned}$$

and thus:

$$q_2^e - q_1^e = (1 - e_1^*) [1 - \rho (1 - e_L^*)] \frac{\Delta}{2}.$$

Therefore,  $q_2^e \geq q_1^e$ , with a strict inequality whenever  $e_1^* < 1$ .

- Finally, if  $i_1^* > e_1^* = 0$ , then again  $e_H^* = 1$  and  $e_L^* > 0$ , and thus

$$\begin{aligned} q_1^e &= L + \frac{1}{2} (1 + i_1^*) \Delta, \\ q_2^e &= L + \frac{1}{2} (1 + i_1^*) \Delta + \frac{1}{2} (1 - i_1^*) (1 - \rho + \rho e_L^*) \Delta, \end{aligned}$$

which yields:

$$q_2^e - q_1^e = (1 - i_1^*) [1 - \rho (1 - e_L^*)] \frac{\Delta}{2}.$$

Therefore,  $q_2^e \geq q_1^e$ , with a strict inequality whenever  $i_1^* < 1$ .

## C The factor $\chi_n$ decreases with $n$

The factor  $\chi_n$  is of the form

$$\chi_n = \delta^n x_n,$$

where

$$x_n \equiv \frac{1 - (2\rho - 1)^n \delta^n}{1 - \delta^n}$$

decreases as  $n$  increases; to show this, it suffices to note that  $\log(x_n)$  indeed decreases as  $n$  increases:<sup>31</sup>

$$\begin{aligned} \frac{d \log(x_n)}{dn} &= \frac{d}{dn} (\log(1 - (2\rho - 1)^n \delta^n) - \log(1 - \delta^n)) \\ &= \frac{\delta^n \log(\delta)}{1 - \delta^n} - \frac{(2\rho - 1)^n \delta^n \log((2\rho - 1)\delta)}{1 - (2\rho - 1)^n \delta^n} \\ &= f(\delta) - f(\delta(2\rho - 1)), \end{aligned}$$

where  $f(z) \equiv \frac{z^n}{1-z^n} \log(z)$  is an increasing function of  $z$  for  $z < 1$ :

$$f'(z) = \frac{z^n (nz^{n-1} \log z + z^{n-1}) - nz^{n-1} \log z}{(1-z^n)^2} = \frac{z^n z^{n-1} - (1-z^n)nz^{n-1} \log z}{(1-z^n)^2} > 0.$$

## D The factor $\xi_n$ increases with $n$

$$\xi_n \equiv \frac{\delta^n}{1 - \delta^n} \frac{1 - (2\rho - 1)^n \delta^n}{(2\rho - 1)^n \delta^n} = \frac{1 - (2\rho - 1)^n \delta^n}{(2\rho - 1)^n (1 - \delta^n)}$$

$$\log(\xi_n) \equiv \log(\delta^n) - \log((2\rho - 1)^n \delta^n) - \log(1 - \delta^n) + \log(1 - (2\rho - 1)^n \delta^n)$$

$$\begin{aligned} \frac{d \log(\xi_n)}{dn} &= \frac{\delta^n \log(\delta)}{\delta^n} - \frac{(2\rho - 1)^n \delta^n \log((2\rho - 1)\delta)}{(2\rho - 1)^n \delta^n} + \frac{\delta^n \log(\delta)}{1 - \delta^n} - \frac{(2\rho - 1)^n \delta^n \log((2\rho - 1)\delta)}{1 - (2\rho - 1)^n \delta^n} \\ &= \log(\delta) - \log((2\rho - 1)\delta) + f(\delta) - f((2\rho - 1)\delta), \end{aligned}$$

where the functions  $\log(z)$  and  $f(z) = \frac{z^n}{1-z^n} \log(z)$  are both increasing in  $z$ .

## E Expected quality

Let  $V_n^H$  (resp.,  $V_n^L$ ) denote the expected value of the quality generated by the contract when the quality was  $H$  or the agent invested in period 0. We have:

$$V_n^H = v_n^H + \delta^n [\mu_n^H V_n^H + (1 - \mu_n^H) V_n^L], \quad (30)$$

$$V_n^L = v_n^L + \delta^n [\mu_n^L V_n^H + (1 - \mu_n^L) V_n^L], \quad (31)$$

<sup>31</sup>Note:

$$\frac{da^n}{dn} = \frac{d(\exp(n \log a))}{dn} = (\exp(n \log a)) (\log a) = a^n (\log a).$$

and thus:

$$\begin{aligned}
V_n &= \frac{V_n^H + V_n^L}{2} \\
&= v_n + \delta^n \left[ V_n + \left( \mu_n - \frac{1}{2} \right) (V_n^H - V_n^L) \right] \\
&= \frac{v_n + \delta^n \left( \mu_n - \frac{1}{2} \right) (V_n^H - V_n^L)}{1 - \delta^n} \\
&= \frac{v_n + \delta^n \frac{i_n}{2} (V_n^H - V_n^L)}{1 - \delta^n}.
\end{aligned} \tag{32}$$

From (30) and (31), we have:

$$\begin{aligned}
V_n^H - V_n^L &= v_n^H - v_n^L + \delta^n (\mu_n^H - \mu_n^L) (V_n^H - V_n^L) \\
&= \frac{v_n^H - v_n^L}{1 - (\mu_n^H - \mu_n^L)} \\
&= \frac{v_n^H - v_n^L}{1 - (2\rho - 1)^n \delta^n (1 - i_n)},
\end{aligned}$$

which, using

$$\begin{aligned}
v_n^H - v_n^L &= \sum_{t=1}^n \delta^{t-1} (p_t^H - p_t^L) \Delta - \delta^{n-1} (p_n^H - p_n^L) i_n \Delta, \\
&= \sum_{t=1}^n (2\rho - 1)^t \delta^{t-1} \Delta - (2\rho - 1)^n \delta^{n-1} i_n \Delta \\
&= \frac{1 - (2\rho - 1)^n \delta^n}{1 - (2\rho - 1) \delta} (2\rho - 1) \Delta - (2\rho - 1)^n \delta^{n-1} i_n \Delta \\
&= \left( \frac{1 - (2\rho - 1)^n \delta^n}{1 - (2\rho - 1) \delta} - (2\rho - 1)^{n-1} \delta^{n-1} i_n \right) (2\rho - 1) \Delta,
\end{aligned}$$

leads to:

$$\begin{aligned}
V_n^H - V_n^L &= \frac{\frac{1 - (2\rho - 1)^n \delta^n}{1 - (2\rho - 1) \delta} - (2\rho - 1)^{n-1} \delta^{n-1} i_n}{1 - (2\rho - 1)^n \delta^n (1 - i_n)} (2\rho - 1) \Delta. \\
&= \frac{v_n + \delta^n \frac{i_n}{2} (V_n^H - V_n^L)}{1 - \delta^n}
\end{aligned} \tag{33}$$

Combining (32) and (33) then yields:

$$V_n = \frac{v_n + \delta^n i_n \frac{\frac{1 - (2\rho - 1)^n \delta^n}{1 - (2\rho - 1) \delta} - (2\rho - 1)^{n-1} \delta^{n-1} i_n}{1 - (2\rho - 1)^n \delta^n (1 - i_n)} (2\rho - 1) \frac{\Delta}{2}}{1 - \delta^n},$$

where

$$v_n = \sum_{t=1}^n \delta^{t-1} \frac{L + H}{2} + \delta^{n-1} \hat{i}_n \frac{\Delta}{2} = \frac{1 - \delta^n}{1 - \delta} \frac{L + H}{2} + \delta^{n-1} \hat{i}_n \frac{\Delta}{2}.$$

This leads to:

$$\begin{aligned}
V_n &= \frac{\frac{1-\delta^n}{1-\delta} \frac{L+H}{2} + \delta^{n-1} i_n \frac{\Delta}{2} + \delta^n i_n \frac{\frac{1-(2\rho-1)^n \delta^n - (2\rho-1)^{n-1} \delta^{n-1} i_n}{1-(2\rho-1)\delta} - (2\rho-1)^{n-1} \delta^{n-1} i_n}{1-(2\rho-1)^n \delta^n (1-i_n)}}{1-\delta^n} (2\rho-1) \frac{\Delta}{2} \\
&= \frac{1}{1-\delta} \frac{L+H}{2} + \frac{\delta^{n-1} i_n \frac{\Delta}{2} + \delta^{n-1} i_n \frac{\frac{1-(2\rho-1)^n \delta^n - (2\rho-1)^{n-1} \delta^{n-1} i_n}{1-(2\rho-1)\delta} - (2\rho-1)^{n-1} \delta^{n-1} i_n}{1-(2\rho-1)^n \delta^n (1-i_n)}}{1-\delta^n} (2\rho-1) \delta \frac{\Delta}{2} \\
&= \frac{1}{1-\delta} \frac{L+H}{2} + \frac{\delta^{n-1} i_n \frac{\Delta}{2}}{1-\delta^n} \left( 1 + \frac{\frac{1-(2\rho-1)^n \delta^n}{1-(2\rho-1)\delta} - (2\rho-1)^{n-1} \delta^{n-1} i_n}{1-(2\rho-1)^n \delta^n (1-i_n)} (2\rho-1) \delta \right).
\end{aligned}$$

For  $i_n = i_n^* = 1$ , this boils down to:

$$\begin{aligned}
V_n^* &= \frac{1}{1-\delta} \frac{L+H}{2} + \frac{\delta^{n-1} \frac{\Delta}{2}}{1-\delta^n} \left( 1 + \left( \frac{1-(2\rho-1)^n \delta^n}{1-(2\rho-1)\delta} - (2\rho-1)^{n-1} \delta^{n-1} \right) (2\rho-1) \delta \right) \\
&= \frac{1}{1-\delta} \frac{L+H}{2} + \frac{\delta^{n-1} \frac{\Delta}{2}}{1-\delta^n} \left( \frac{1-(2\rho-1)^n \delta^n}{1-(2\rho-1)\delta} \right) \\
&= \frac{1}{1-\delta} \frac{L+H}{2} + \chi_n \frac{1}{1-(2\rho-1)\delta} \frac{\Delta}{2\delta}.
\end{aligned}$$

As  $\chi_n$  decreases when  $n$  increases, the expected quality  $V_n^*$  decreases as contracts last longer than needed to induce  $i^* = 1$ .

For  $i_n = i_n^* = \hat{i}_n \leq 1$ , we have

$$V_n = \frac{1}{1-\delta} \frac{L+H}{2} + \frac{\delta^{n-1} \hat{i}_n \frac{\Delta}{2}}{1-\delta^n} \left( 1 + \frac{\frac{1-(2\rho-1)^n \delta^n}{1-(2\rho-1)\delta} - (2\rho-1)^{n-1} \delta^{n-1} \hat{i}_n}{1-(2\rho-1)^n \delta^n (1-\hat{i}_n)} (2\rho-1) \delta \right),$$

where, using (7) and after simplification

$$1 + \frac{\frac{1-(2\rho-1)^n \delta^n}{1-(2\rho-1)\delta} - (2\rho-1)^{n-1} \delta^{n-1} \hat{i}_n}{1-(2\rho-1)^n \delta^n (1-\hat{i}_n)} (2\rho-1) \delta = 1 + \frac{\frac{c}{\Delta}}{\frac{1}{1-(2\rho-1)\delta} - \frac{c}{\Delta}}.$$

Therefore:

$$\begin{aligned}
V_n &= \frac{1}{1-\delta} \frac{L+H}{2} + \frac{\delta^n}{1-\delta^n} \hat{i}_n \frac{1}{1-(1-(2\rho-1)\delta) \frac{c}{\Delta}} \frac{\Delta}{2\delta} \\
&= \frac{1}{1-\delta} \frac{L+H}{2} + \xi_n \left( \frac{(2\rho-1)\delta}{1-(2\rho-1)\delta} - \frac{c}{\Delta} \right) \frac{1}{1-(1-(2\rho-1)\delta) \frac{c}{\Delta}} \frac{\Delta}{2\delta},
\end{aligned}$$

which, as  $\xi_n$ , increases with  $n$ .

## F Contract duration, investment and welfare with convex costs

To check the robustness of our insights on optimal contract duration, we consider here an infinitely repeated setting with convex investment costs: improving productivity with probability

$i$  costs  $c(i)$ , where  $c(0) = 0$ ,  $c'(0) \geq 0$  and  $c''(\cdot) > 0$ ; for some of the results, we will further index the cost function by some parameter  $\lambda$ , where  $\frac{\partial^2 c(i, \lambda)}{\partial i \partial \lambda}$ ,  $\frac{\partial^3 c(i, \lambda)}{\partial i^2 \partial \lambda} \geq 0$ , with at least one strict inequality. In this setting, we compare investment patterns and welfare under one-period contracting and two-period contracting.

## F.1 Investment patterns

We first show that investment exhibits similar patterns as in the case of constant returns to scale.

• **One-period contracting.** Under one period contracting, the agent's expected payoff at the beginning of a contract, conditional on the performance  $q \in \{H, L\}$  in the previous contract, is given by:

$$\begin{aligned}\Pi^H &= v^H - (1 - \rho) c(i) + \delta [(\rho + (1 - \rho)i) \Pi^H + (1 - \rho)(1 - i) \Pi^L], \\ \Pi^L &= v^L - \rho c(i) + \delta [(1 - \rho + \rho i) \Pi^H + \rho(1 - i) \Pi^L],\end{aligned}$$

where:

$$\begin{aligned}v^H &= L + [\rho + (1 - \rho)i] \Delta, \\ v^L &= L + (1 - \rho + \rho i) \Delta.\end{aligned}$$

The equilibrium level of investment,  $i^*$ , solves  $c'(i^*) = b(i^*)$ , where:

$$\begin{aligned}b(i) &\equiv \delta (\Pi^H - \Pi^L) = \delta [v^H - v^L + (2\rho - 1) c(i) + (2\rho - 1)(1 - i) \delta (\Pi^H - \Pi^L)] \\ &= \sigma \frac{(1 - i) \Delta + c(i)}{1 - \sigma(1 - i)},\end{aligned}$$

with  $\sigma \equiv \delta(2\rho - 1) < 1$ . Note that, as in the case of constant returns to scale, crowding out limits the investment level:  $i^* < 1$ , as it is never optimal to invest merely to reduce the investment cost in the next period:

$$b(1) = \sigma c(1) < c(1) < c'(1).$$

The equilibrium condition  $c'(i^*) = b(i^*)$  amounts to  $\varphi(i^*, \Delta, \sigma) = 0$ , where

$$\varphi(i; \Delta, \sigma) \equiv \sigma [(1 - i) \Delta + c(i)] - [1 - \sigma(1 - i)] c'(i), \quad (34)$$

with

$$\begin{aligned}
\frac{\partial \varphi}{\partial i} &= -\sigma \Delta - [1 - \sigma(1 - i)] c''(i) < 0, \\
\frac{\partial \varphi}{\partial \Delta} &= \sigma(1 - i) > 0, \\
\frac{\partial \varphi}{\partial \sigma} &= (1 - i) \Delta + c(i) + (1 - i) c'(i) > 0, \\
\frac{\partial \varphi}{\partial \lambda} &= \sigma \frac{\partial c}{\partial \lambda}(i, \lambda) - [1 - \sigma(1 - i)] \frac{\partial^2 c}{\partial i \partial \lambda}(i, \lambda) \\
&= -(1 - \sigma) \frac{\partial^2 c}{\partial i \partial \lambda}(i, \lambda) + \sigma \frac{\partial}{\partial \lambda} \left[ c(i, \lambda) - i \frac{\partial c}{\partial i}(i, \lambda) \right] \\
&= -(1 - \sigma) \frac{\partial^2 c}{\partial i \partial \lambda}(i, \lambda) + \sigma \frac{\partial}{\partial \lambda} \left\{ \int_0^i \left[ \frac{\partial c}{\partial i}(x, \lambda) - \frac{\partial c}{\partial i}(i, \lambda) \right] dx \right\} \\
&= -(1 - \sigma) \frac{\partial^2 c}{\partial i \partial \lambda}(i, \lambda) - \sigma \frac{\partial}{\partial \lambda} \left[ \int_0^i \int_x^i \frac{\partial^2 c}{\partial i^2}(y, \lambda) dy dx \right] \\
&= -(1 - \sigma) \frac{\partial^2 c}{\partial i \partial \lambda}(i, \lambda) - \sigma \int_0^i \int_x^i \frac{\partial^3 c}{\partial i^2 \partial \lambda}(y, \lambda) dy dx < 0.
\end{aligned}$$

Hence, the equilibrium investment level  $i^*$  increases with the investment benefit  $\Delta$  as well as with  $\sigma$  (that is, with the discount rate  $\delta$  and with the degree of information persistence  $\rho$ ), and decreases with the cost parameter  $\lambda$ . In particular,  $i^* = 0$  when  $c'(0) \geq b(0) = \frac{\sigma \Delta}{1 - \sigma}$ , whereas  $i^*$  tends to 1 as  $\Delta$  goes to  $+\infty$ .<sup>32</sup>

• **Two-period contracting.** Let us now turn to two-period contracting; we will use the subscript  $t = 1, 2$  to refer to the first or second period of a contract. The agent's expected payoff at the beginning of a contract, conditional on the performance  $q \in \{H, L\}$  delivered at the end of the previous contract, is given by:

$$\begin{aligned}
\Pi_1^H &= v_1^H - (1 - \rho) c(i_1) - \delta(1 - \rho)(1 - i_1) c(i_2) + \delta [\mu_2^H \Pi_1^H + (1 - \mu_2^H) \Pi_1^L], \\
\Pi_1^L &= v_1^L - \rho c(i_1) - \delta \rho(1 - i_1) c(i_2) + \delta [\mu_2^L \Pi_1^H + (1 - \mu_2^L) \Pi_1^L],
\end{aligned}$$

---

<sup>32</sup>This is because  $b(i)$  tends to infinity with  $\Delta$  when  $i$  remains bounded away from 1. For instance, when  $c(i) = \lambda i + i^2/2$ , then  $i^* = 0$  as long as  $\frac{\sigma \Delta}{1 - \sigma} \leq \lambda$  and, for  $\frac{\sigma \Delta}{1 - \sigma} > \lambda$ :

$$i^* = \frac{1 - \sigma}{\sigma} \left( 1 + \frac{\sigma \Delta}{1 - \sigma} \right) \left( \sqrt{1 + 2 \frac{1 - \sigma}{\sigma} \frac{\frac{\sigma \Delta}{1 - \sigma} - \lambda}{\left( 1 + \frac{\sigma \Delta}{1 - \sigma} \right)^2}} - 1 \right),$$

which decreases with  $\lambda$  and increases from 0 to 1 as  $\Delta$  goes from  $\lambda(\sigma/(1 - \sigma))^{-1}$  to infinity.



where:

$$\begin{aligned}
v_1^H &= L + [\rho + (1 - \rho) i_1] \Delta + \delta (L + \mu_2^H \Delta), \\
v_1^L &= L + (1 - \rho + \rho i_1) \Delta + \delta (L + \mu_2^L \Delta), \\
\mu_2^H &= [\rho + (1 - \rho) i_1] [\rho + (1 - \rho) i_2] + (1 - \rho) (1 - i_1) (1 - \rho + \rho i_2), \\
\mu_2^L &= (1 - \rho + \rho i_1) [\rho + (1 - \rho) i_2] + [\rho + (1 - \rho) i_1] (1 - \rho + \rho i_2).
\end{aligned}$$

The second-period equilibrium level of investment,  $i_2^*$ , thus solves  $c'(i_2^*) = b_2(i_1^*, i_2^*)$ , where (using  $v_2^H - v_2^L = [(2\rho - 1) + \delta(\mu_2^H - \mu_2^L)] \Delta$  and  $\mu_2^H - \mu_2^L = (2\rho - 1)^2 (1 - i_1) (1 - i_2)$ ):

$$\begin{aligned}
b_2(i_1, i_2) &\equiv \delta (\Pi_1^H - \Pi_1^L) = \delta \{v_1^H - v_1^L + (2\rho - 1) c(i_1) + \delta (2\rho - 1)^2 (1 - i_1) c(i_2) \\
&\quad + \delta^2 (\mu_2^H - \mu_2^L) (\Pi_1^H - \Pi_1^L)\} \\
&= \delta \frac{v_1^H - v_1^L + (2\rho - 1) c(i_1) + \delta (2\rho - 1)^2 (1 - i_1) c(i_2)}{1 - \delta^2 (\mu_2^H - \mu_2^L)} \\
&= \sigma \frac{(1 - i_1) [1 + \sigma (1 - i_2)] \Delta + c(i_1) + \sigma (1 - i_1) c(i_2)}{1 - \sigma^2 (1 - i_1) (1 - i_2)}.
\end{aligned}$$

In the first period of a contract, when the productivity is low (i.e.,  $\theta_1 = L$ ), the agent's expected payoff in the second period of a contract is given by:

$$\begin{aligned}
&-c(i_1) + i_1 \delta \{ \rho \delta \Pi_1^H + (1 - \rho) [\delta \Pi_1^L - c(i_2) + \delta (\Pi_1^H - \Pi_1^L) i_2] \} \\
&\quad + (1 - i_1) \delta \{ (1 - \rho) \delta \Pi_1^H + \rho [\delta \Pi_1^L - c(i_2) + \delta (\Pi_1^H - \Pi_1^L) i_2] \} \\
&= \delta \{ (1 - \rho) \delta \Pi_1^H + \rho [\delta \Pi_1^L - c(i_2) + \delta (\Pi_1^H - \Pi_1^L) i_2] \} \\
&-c(i_1) + i_1 (2\rho - 1) \delta \{ \delta (\Pi_1^H - \Pi_1^L) (1 - i_2) + c(i_2) \}.
\end{aligned}$$

Therefore, the first-period equilibrium level of investment,  $i_1^*$ , solves  $c'(i_1^*) = b_1(i_1^*, i_2^*)$ , where, using  $\delta (\Pi_1^H - \Pi_1^L) = b_2(i_1, i_2)$ :

$$b_1(i_1, i_2) \equiv \sigma [c(i_2) + b_2(i_1, i_2) (1 - i_2)].$$

Letting  $b_t^* \equiv b_t(i_1^*, i_2^*)$  denote the equilibrium investment benefits, for  $t = 1, 2$ , we thus have:

$$b_1^* = \sigma b_2^* - \sigma [b_2^* i_2^* - c(i_2^*)] = \sigma b_2^* - \sigma \max_{i_2} [b_2^* i_2 - c(i_2)] \leq \sigma b_2^* < b_2^*.$$

Hence, as in the case of constant returns to scale, there is less investment in the first period than in the second period: either  $i_1^* = i_2^* = 0$ , or  $i_1^* < i_2^*$ ; however, as in Proposition 3, the pattern can be less pronounced than with constant returns (namely,  $i_1^*$  can be positive), as it is here desirable to smooth investment costs across periods. Furthermore,  $i_2^* = 0$  (which implies  $i_1^* = 0$ ) if and only if  $c'(0) \geq b_2(0, 0) = \frac{\sigma \Delta}{1 - \sigma}$ .

• **Comparison.** Under two-period contracting, there is thus less investment in the first period than in the second period; this, in turn, limits crowding out and fosters investment incentives in the last contracting periods. Indeed, we now show that, as in the case of constant returns to scale, there is more investment in the last period of second-period contracts than under one-period contracting:

**Proposition 9** *When  $c'(0) \geq \frac{\sigma\Delta}{1-\sigma}$ ,  $i_1^* = i_2^* = i^* = 0$ . When instead  $c'(0) < \frac{\sigma\Delta}{1-\sigma}$ ,  $i_2^* > i^* > 0$ .*

**Proof.** We already noted that both  $i^*$  and  $i_2^*$  are 0 when  $c'(0) \geq \frac{\sigma\Delta}{1-\sigma}$ , in which case  $i_1^* \leq i_2^*$  implies  $i_1^* = 0$  as well. Conversely,  $i_2^*, i^* > 0$  when  $c'(0) < \frac{\sigma\Delta}{1-\sigma}$ . Suppose now that  $0 < i_2^* \leq i^*$ . From the above, this implies  $i_1^* < i_2^* (\leq i^*)$ . We then have:

- $b_2^* > b(i_2^*)$ . By construction:

$$b_2(i_2^*, i_2^*) = b(i_2^*).$$

Therefore, it suffices to note that  $b_2(i_1, i_2^*)$  strictly decreases with  $i_1$  in the range  $i_1 \in [i_1^*, i_2^*]$ :

$$\begin{aligned} \frac{\partial b_2}{\partial i_1}(i_1, i_2^*) &= \sigma \frac{\left\{ \begin{array}{l} [1 - \sigma^2(1 - i_1)(1 - i_2^*)] \{-[1 + \sigma(1 - i_2^*)]\Delta + c'(i_1) - \sigma c(i_2^*)\} \\ -\sigma^2(1 - i_2^*) \{(1 - i_1)[1 + \sigma(1 - i_2^*)]\Delta + c(i_1) + \sigma(1 - i_1)c(i_2^*)\} \end{array} \right\}}{[1 - \sigma^2(1 - i_1)(1 - i_2^*)]^2} \\ &= \sigma \frac{-[1 + \sigma(1 - i_2^*)]\Delta - \sigma c(i_2^*) + A(i_1)}{[1 - \sigma^2(1 - i_1)(1 - i_2^*)]^2} \end{aligned}$$

where

$$\begin{aligned} A(i_1) &\equiv [1 - \sigma^2(1 - i_1)(1 - i_2^*)] c'(i_1) - \sigma^2(1 - i_2^*) c(i_1), \\ A'(i_1) &= [1 - \sigma^2(1 - i_1)(1 - i_2^*)] c''(i_1) > 0. \end{aligned}$$

Therefore, for  $i_1 \leq i_2^* \leq i^*$ , we have  $A(i_1) \leq A(i_2^*)$  and thus:

$$\begin{aligned} \frac{\partial b_2}{\partial i_1}(i_1, i_2^*) &\leq \sigma \frac{-[1 + \sigma(1 - i_2^*)]\Delta - \sigma c(i_2^*) + A(i_2^*)}{[1 - \sigma^2(1 - i_1)(1 - i_2^*)]^2} \\ &= \sigma \frac{-[1 + \sigma(1 - i_2^*)] \{\Delta + \sigma c(i_2^*) - [1 - \sigma(1 - i_2^*)] c'(i_2^*)\}}{[1 - \sigma^2(1 - i_1)(1 - i_2^*)]^2} \\ &= \frac{-\sigma [1 + \sigma(1 - i_2^*)]}{[1 - \sigma^2(1 - i_1)(1 - i_2^*)]^2} B(i_2^*), \end{aligned}$$

where

$$\begin{aligned} B(i_2) &\equiv \Delta + \sigma c(i_2) - [1 - \sigma(1 - i_2)] c'(i_2), \\ B'(i_2) &= -[1 - \sigma(1 - i_2)] c''(i_2) < 0. \end{aligned}$$

Therefore, for  $i_1 \leq i_2^* \leq i^*$ , we have (using  $c'(i^*) = b(i^*)$ ):

$$B(i_2^*) \geq B(i^*) = \Delta + \sigma c(i^*) - [1 - \sigma(1 - i^*)] c'(i^*) = (1 - \sigma)\Delta + \sigma i^* \Delta > 0,$$

and thus:

$$\frac{\partial b_2}{\partial i_1}(i_1, i_2^*) \leq \frac{-\sigma[1 + \sigma(1 - i_2^*)]}{[1 - \sigma^2(1 - i_1)(1 - i_2^*)]^2} B(i^*) < 0.$$

- $b(i_2^*) \geq b(i^*)$ . To see this, it suffices to check that  $b(i)$  decreases with  $i$  in the range  $i \leq i^*$ :

$$\begin{aligned} b'(i) &= \sigma \frac{[1 - \sigma(1 - i)] [-\Delta + c'(i)] - \sigma \{(1 - i)\Delta + c(i)\}}{[1 - \sigma(1 - i)]^2} \\ &= \sigma \frac{-\Delta + [1 - \sigma(1 - i)] c'(i) - \sigma c(i)}{[1 - \sigma(1 - i)]^2} \\ &= \frac{\sigma C(i)}{[1 - \sigma(1 - i)]^2}, \end{aligned}$$

- where

$$C(i) \equiv -\Delta + [1 - \sigma(1 - i)] c'(i) - \sigma c(i),$$

$$C'(i) = [1 - \sigma(1 - i)] c''(i) > 0.$$

Therefore, for  $i \leq i^*$  (using  $c'(i^*) = b(i^*)$ ):

$$\begin{aligned} b'(i) &\leq \frac{\sigma C(i^*)}{[1 - \sigma(1 - i)]^2} \\ &= \frac{\sigma \{-\Delta + [1 - \sigma(1 - i^*)] c'(i^*) - \sigma c(i^*)\}}{[1 - \sigma(1 - i)]^2} \\ &= -\frac{1 - \sigma(1 - i^*)}{[1 - \sigma(1 - i)]^2} \sigma \Delta < 0. \end{aligned}$$

- But this implies:

$$c'(i_2^*) = b_2^* > b(i_2^*) \geq b(i^*) = c'(i^*),$$

contradicting the working assumption  $i_2^* \leq i^*$ .

■

By contrast,  $i_1^*$  can be lower than  $i^*$ , and may well be equal to 0 even when  $i^*, i_2^* > 0$ :

**Proposition 10** When  $\frac{\sigma\Delta}{1-\sigma} > c'(0) \geq \frac{\sigma^2\Delta}{1-\sigma}$ , then  $i_2^* > i^* > i_1^* = 0$ .

**Proof.** From the above,  $i_2^* > i^* > 0$  when  $\frac{\sigma\Delta}{1-\sigma} > c'(0)$ . We then have  $i_1^* = 0$  when:

$$\frac{\sigma\Delta}{1-\sigma} > c'(0) \geq b_1(0, i_2^*) = \sigma [c(i_2^*) + b_2^*(1 - i_2^*)],$$

where  $b_2^*$  and  $i_2^*$  are jointly determined by

$$c'(i_2^*) = b_2^* = \sigma \frac{[1 + \sigma(1 - i_2^*)] \Delta + \sigma c(i_2^*)}{1 - \sigma^2(1 - i_2^*)}.$$

In particular, this is always the case when:

$$\frac{\sigma\Delta}{1-\sigma} > c'(0) \geq \frac{\sigma^2\Delta}{1-\sigma}.$$

To see this, note that

$$b_1(0, i_2^*) = \sigma b_2(0, i_2^*) - \sigma \max_{i_2} [b_2(0, i_2^*) i_2 - c(i_2)] \leq \sigma b_2(0, i_2^*) < \sigma b_2(0, 0) = \frac{\sigma^2\Delta}{1-\sigma},$$

as  $b_2(0, i_2)$  decreases with  $i_2$  in the range  $i_2 \in [0, i_2^*]$ :

$$\begin{aligned} \frac{\partial b_2(0, i_2)}{\partial i_2} &= \sigma \frac{[1 - \sigma^2(1 - i_2)] \sigma [-\Delta + c'(i_2)] - \sigma^2 \{[1 + \sigma(1 - i_2)] \Delta + \sigma c(i_2)\}}{[1 - \sigma^2(1 - i_2)]^2} \\ &= \sigma^2 \frac{-(1 + \sigma) \Delta + [1 - \sigma^2(1 - i_2)] c'(i_2) - \sigma^2 c(i_2)}{[1 - \sigma^2(1 - i_2)]^2} \\ &= \frac{-\sigma^2 D(i_2)}{[1 - \sigma^2(1 - i_2)]^2}, \end{aligned}$$

where:

$$\begin{aligned} D(i_2) &= (1 + \sigma) \Delta + \sigma^2 c(i_2) - [1 - \sigma^2(1 - i_2)] c'(i_2), \\ D'(i_2) &= -c''(i_2) < 0. \end{aligned}$$

Therefore, for  $i_2 \leq i_2^*$  (using  $c'(i_2^*) = b_2(0, i_2^*)$ ):

$$\begin{aligned} \frac{\partial b_2(0, i_2)}{\partial i_2} &\leq \frac{-\sigma^2 D(i_2^*)}{[1 - \sigma^2(1 - i_2)]^2} \\ &= \frac{-\sigma^2 \{(1 + \sigma) \Delta + \sigma^2 c(i_2^*) - [1 - \sigma^2(1 - i_2^*)] c'(i_2^*)\}}{[1 - \sigma^2(1 - i_2)]^2} \\ &= -\sigma^2 \frac{(1 + \sigma) \Delta + \sigma^2 c(i_2^*) - \sigma \{[1 + \sigma(1 - i_2^*)] \Delta + \sigma c(i_2^*)\}}{[1 - \sigma^2(1 - i_2)]^2} \\ &= -\frac{[1 - \sigma^2(1 - i_2^*)]}{[1 - \sigma^2(1 - i_2)]^2} \sigma^2 \Delta < 0. \end{aligned}$$

■

Note finally that, contrary to  $i^*$ ,  $i_2^*$  may well be equal to 1. For instance, if  $\frac{\sigma^2\Delta}{1-\sigma} \leq c'(0)$  (so that  $i_1^* = 0$ ), this is the case when

$$c'(1) \leq b_2(0, 1) = \sigma\Delta + \sigma^2c(1).$$

Hence, if  $\sigma$  is large enough that

$$\sigma [c'(1) - \sigma^2c(1)] \leq (1 - \sigma) c'(0),$$

we have  $i_2^* = 1 > i^* > i_1^* = 0$  whenever  $\Delta$  is such that  $\sigma^2\Delta$  lies in the range  $[\sigma (c'(1) - \sigma^2c(1)), (1 - \sigma) c'(0)]$ .

## F.2 Welfare

We now compare welfare in the two scenarios, and draw the implications for optimal contract duration. As in the case of constant returns to scale, increasing contract duration (from one period to two periods) fosters investment in the last contracting period, but reduces that in the first period. For the sake of exposition, we will focus on the case where  $c'(0) \geq \frac{\sigma^2\Delta}{1-\sigma}$ , which implies  $i_1^* = 0$ . As long as  $c'(0) < \frac{\sigma\Delta}{1-\sigma}$ , however,  $i_2^*$  and  $i^*$  are positive (and in that case  $i_2^* > i^*$ ). The trade-off involved is then exactly the same as with constant returns to scale: compared with one-period contracting, two-period contracting yields less frequent investment (as it only occurs in the second period of a contract), but greater intensity (that is, in the second contracting period there is more investment than with one-period contracts). Intuitively, the balance will be in favor of one-period contracting when it already yields substantial investment intensity, and can be in favor of longer contracts otherwise.

- **One-period contracting.**

Expected welfare is given by (where the superscript 1 refers to one-period contracting):

$$\begin{aligned} W^1 &= \frac{\Pi^H + \Pi^L}{2} \\ &= \frac{v^H + v^L}{2} - \frac{c(i^*)}{2} + \delta \frac{\Pi^H + \Pi^L}{2} + \delta \frac{i^*}{2} (\Pi^H - \Pi^L), \end{aligned}$$

where  $\frac{v^H + v^L}{2} = \frac{L+H}{2} + \frac{i^*}{2}\Delta$  and  $\delta (\Pi^H - \Pi^L) = c'(i^*)$ ; hence:

$$(1 - \delta) W^1 = \frac{L + H}{2} + \frac{w(i^*)}{2}, \tag{35}$$

where

$$w(i) \equiv i\Delta - c(i) + ic'(i)$$

increases with  $i$ :

$$w'(i) = \Delta + ic''(i) > 0.$$

- **Two-period contracting.** As  $i_1^* = 0$ , expected welfare is given by:

$$\begin{aligned} W_2 &= \frac{\Pi_1^H + \Pi_1^L}{2} \\ &= \frac{v_1^H + v_1^L}{2} - \frac{\delta c(i_2^*)}{2} + \delta^2 \frac{\Pi_1^H + \Pi_1^L}{2} + \delta^2 \frac{i_2^*}{2} (\Pi_1^H - \Pi_1^L), \end{aligned}$$

where  $\frac{v_1^H + v_1^L}{2} = (1 + \delta) \frac{L+H}{2} + \delta \frac{i_2^*}{2} \Delta$  and  $\delta (\Pi_1^H - \Pi_1^L) = c'(i_2^*)$ ; hence:

$$(1 - \delta^2) W_2 = (1 + \delta) \frac{L + H}{2} + \frac{\delta}{2} w(i_2^*),$$

or:

$$(1 - \delta) W_2 = \frac{L + H}{2} + \frac{\delta w(i_2^*)}{2(1 + \delta)}. \quad (36)$$

- **Comparison.** From (35) and (36), welfare is higher under two-period contracting if and only if:

$$\frac{\delta}{1 + \delta} w^*(i_2^*) > w^*(i^*).$$

Note that  $\frac{\delta}{1 + \delta}$  is bounded above by  $\frac{1}{2}$ . Therefore, if  $i^*$  is large, because  $\Delta$  is large or  $\rho$  and  $\delta$  are close to 1, then one-period contracting will lead to greater welfare. If instead  $i^*$  is small, then two-period contracting be preferable.<sup>33</sup>

---

<sup>33</sup>For instance, for the quadratic cost function  $c(i) = \lambda i + i^2/2$  considered in footnote 32, two-period contracting generates more welfare than one-period contracting when, e.g.,  $\lambda = 1$ ,  $\delta \simeq 1$ ,  $\sigma = \frac{1}{2}$  (which, for  $\delta = 1$ , amounts to  $\rho = \frac{3}{4}$ ), and  $\Delta \geq \frac{13}{4}$ , in which case  $i_2^* = 1 > i_1^* = 0$  and  $\frac{w(i_2^*)}{2} = \frac{15}{8} > w(i^*) = \frac{7}{4}$ .