

CIRPÉE

Centre interuniversitaire sur le risque, les politiques économiques et l'emploi

Cahier de recherche/Working Paper **06-10**

Dynamic Regulation of Public Good Quality

Stéphane Auray
Thomas Mariotti
Fabien Moizeau

Mars/March 2006

Auray: Université de Lille 3, GREMARS (EA 2459) and CIRPÉE, Maison de la Recherche, Domaine Universitaire du Pont de Bois, BP 149, 59653 Villeneuve d'Ascq cedex, France

stephane.auray@univ-lille3.fr

Mariotti: Université de Toulouse 1, GREMAQ (UMR CNRS 5604) and IDEI, 21 Allée de Brienne, 31000 Toulouse, France; Centre for Economic Policy Research, 90-98 Goswell Road, London EC1V 7RR

mariotti@cict.fr

Moizeau: Université de Toulouse 1, GREMAQ (UMR CNRS 5604), 21 Allée de Brienne, 31000 Toulouse, France

fabien.moizeau@univ-tlse1.fr

We would like to thank Bruno Biais, Arianna Degan, Bruno Julien, Jean-Charles Rochet and Jean Tirole for insightful comments.

Abstract:

We investigate the design of incentives for public good quality provision in a dynamic regulation setting in which maintenance efforts and quality shocks have durable effects. When the regulator contracts with a sequence of agents, asymmetries of information can lead to over-provision of quality under optimal regulation, reflecting a dynamic rent extraction motive. When the regulator hires a single agent to manage public good quality, over-provision of quality can also occur as a result of quality pooling, which typically occurs if quality depreciates slowly and the discount factor is large. We further show that for small levels of asymmetric information, the regulator prefers to hire a single agent rather than to contract with a sequence of agents, provided all parties can commit to a long-term contract. When no such commitment is feasible, the fact that quality physically links periods together leads to a ratchet effect even when private information is recurring, and shorter franchises are beneficial from a social point of view.

Keywords: Quality, Regulation, Asymmetric Information

JEL Classification: D82, L15, L51

1. INTRODUCTION

The provision of quality by public utilities or regulated firms more generally is one of the main objectives of regulatory activity. Besides securing basic services such as electricity supply or voice and data transmission, regulators attempt to incite regulated firms to provide adequate levels of service enhancements that affect final customers' welfare. For instance, the speed and clarity of transmissions are key dimensions of quality in the telecommunications sector, as is the reliability of supply in the energy sector.

The literature on the regulation of quality has mostly focused on static frameworks, with special emphasis on the issue of quality verifiability (Laffont and Tirole (1991), Lewis and Sappington (1991, 1992)). However, in many cases, the management of quality can be properly understood only in a dynamic context. For instance, the maintenance of a road or electricity network requires sustained efforts, while the network itself is subject to exogenous events which affect the quality of the service it provides to final customers. Similarly, the quality of water supply is affected by exogenous polluting activities, and requires continuing cleaning up efforts. In all these examples, public good quality evolves over time as a result of maintenance efforts and exogenous random shocks, and these efforts and shocks have long-lasting effects. As a result of this, quality is a durable characteristic which can be viewed as a capital stock: the current quality of service depends on its past levels. This paper analyzes the problem of optimal regulation of public good quality in such a dynamic environment. In this framework, we endeavor to shed light on the following questions: What is the best incentive scheme when the regulated firm's decisions affect the level of quality available both in the present and in the future? Does private information systematically lead to under-provision of quality, as in a static environment?

Our analysis focuses on the case where the quality of the public good is verifiable. That is, quality can be described *ex ante* in a contract and certified *ex post* in court, and, as a result of this, the regulator can directly impose a quality target on the regulated firm, or more generally reward or punish the firm directly as a function of quality improvements. This is most relevant for industries such as electricity, where the number and intensity of outages can be ascertained in an almost costless way, or water supply, in which the chemical composition of water provides an accurate measure of its quality for final customers.¹ While quality itself is verifiable, we assume that the factors governing its evolution over time cannot be verified separately, and are private information of the regulated firm. Thus, in line with Lewis and Sappington (1991), the regulator cannot determine the portion of overall quality that can be respectively attributed to the regulated firm's maintenance efforts and to the exogenous quality shocks. Quality shocks are assumed to be independently distributed across periods. Hence, the only link between periods is physical rather than informational. For simplicity, we consider a binomial model in which only high and low quality shocks can occur.

¹As pointed out by De Fraja and Iozzi (2004), there have been in practice two types of regulatory responses to the problem of securing an adequate quality service (see also Armstrong, Cowan and Vickers (1994)). First, the imposition of quality standards, enforced through legal sanctions. Second, the imposition of a link between the firm's allowed revenues and prices and the quality of the service it provides. For instance, in the UK water industry, price cap adjustments are based on comparative performance indicators (OFWAT (2002)). Similarly, UK energy distribution companies receive financial compensations according to various quality indicators (OFGEM (2001)). Note that these two types of mechanisms require at least some dimensions of quality to be verifiable by a court. De Fraja and Iozzi (2004) propose an extension of the Vogelsang and Finsinger (1979) dynamic model of price regulation that allows for such price-quality tradeoffs. They do not address the issue of quality durability, which is the objective of the present paper.

Two regulatory frameworks are considered in turn. We first analyze the dynamics of quality in a stationary setup in which the regulator delegates the management of quality to a sequence of firms or agents, one for each period. This allows us to disentangle the question of quality dynamics from that of the provision of dynamic incentives, which is addressed in the second part of the paper. A key assumption of the model is that each agent must receive a non-negative utility in each state of nature. That is, no contract can be enforced whereby a truthful agent could potentially incur losses. Within each period, the timing is as follows. After signing a contract with the regulator, which depends on the current level of quality, the agent in charge privately learns his type, that is, the current realization of the quality shock, and then privately chooses his maintenance effort. Transfers are then effected according to the achieved level of quality.

Because the agents must receive a non-negative utility in each state of nature, the fact that the quality shocks are privately observed by the agents leads to a sequence of non-degenerate moral hazard problems. The main difference with a static framework is that the social value of quality reflects not only the current social benefit of quality, but also its impact on the continuation game played by the regulator and the future agents. Accordingly, maintenance efforts and transfers will vary over time with the quality of the public good. Using standard recursive techniques (Stokey and Lucas (1989)), we characterize the social value of quality under both symmetric and asymmetric information. A key result of our analysis is that the marginal social value of quality is strictly higher under asymmetric information than under symmetric information, reflecting a dynamic rent extraction effect. Indeed, the informational rent that the regulator must leave to the current agent in case of a high quality shock is decreasing in the current level of quality: when quality is high, the agency problem becomes less severe. This implies that, relative to a static environment, the regulator has an additional incentive to enhance quality, namely to reduce future informational rents. In particular, she will take advantage of a high quality shock to demand a higher effort from the current agent. As a result of this, there may be over-provision of quality relative to the symmetric information environment, typically following a sequence of high quality shocks. Sharper predictions are derived using a linear-quadratic specification of the model. It is shown that, while private information leads to a lower average growth rate of quality, it also increases the variance of quality. In the long run, the range of possible public good qualities is larger under asymmetric information than under symmetric information, and thus both over-provision and under-provision of quality can persist asymptotically.

We next turn to the case where the regulator delegates the management of quality to a single agent, which raises the issue of dynamic incentives. To deliver analytical results, we consider a two-period model. In each period, the agent privately learns the realization of the quality shock, and then chooses his effort. In line with the basic model, we assume that the agent must receive a non-negative utility at each date and in each state. We mainly focus on the full commitment case, in which both the regulator and the agent can commit to a long-term contract, subject to incentive compatibility and participation constraints. Our main findings are as follows. First, the optimal long-term contract exhibits memory: the level of distortions in the second period depends on the type of the agent in the first period. With a sequence of agents, by contrast, distortions in the second period would depend on the past only through the level of quality inherited from the first period. Second, the regulator may find it optimal to induce a constant quality of service in the first period, irrespective of the realized quality shock. The reason for this pooling outcome is that increasing the

quality produced following a low quality shock in the first period allows to reduce the second period informational rent. This directly benefits the regulator, and also makes misreporting by the agent in case of a high quality shock less attractive from the first period perspective, thus decreasing the cost of dynamic incentives. Screening is then achieved using only current transfers and continuation rents: following a high quality shock, the agent receives a low transfer in the first period in exchange for a high continuation rent, while the reverse is true following a low quality shock. In these circumstances, there is over-provision of first period quality following both a high and a low quality shock. Using the linear-quadratic specification of the model, we show that pooling over first period quality typically occurs when the future is important relative to the present, that is, when quality depreciates slowly and the discount factor is high. Our final result compares the model with a single agent and full commitment to the model with a sequence of agents, specialized to the two-period case. A natural question is whether it is better for the regulator to hire a single agent, or to contract with a sequence of agents. The benefit of hiring a single agent is that the regulator can directly condition his continuation rent on his first period performance, as in the standard repeated moral hazard problem (Rogerson (1985)). The cost, by contrast, is that the agent correctly anticipates the impact of his first period actions on his future utility, thus making the first period incentive compatibility constraint more stringent than with a sequence of agents. Our main result is that, for small levels of asymmetric information, the first effect dominates, so that it is optimal for the regulator to hire a single agent.

This paper is in line with works that extend the analysis of incentives in regulation to a dynamic framework (Baron and Besanko (1984), Laffont and Tirole (1988, 1990), Lewis and Sappington (1997)). In these papers, the source of the dynamics is that the regulated firm's costs are correlated across periods, so that the regulator progressively learns about the efficiency of the firm. Instead of this, the intertemporal link stressed in this paper is purely physical. This is similar to Lewis and Yildirim (2002), who study the optimal regulation of a firm who learns to use cost-reducing innovations over time.² They show that a light-handed regulation may encourage innovation, by allowing the firm to earn greater informational rents while providing greater service. Moreover, innovation may occur even in the absence of long-term agreements, provided private information is renewed in each period. Their model differs from ours in some important respects. First, private information in our model takes the form of quality shocks that directly affect future consumer surplus and thus have permanent effects, while it is embedded in their model in a transitory cost. Next, while Lewis and Yildirim (2002) only allow for sequences of spot contracts, we analyze the opposite case in which the regulator can commit to a long-term contract, subject to the constraint that the agent must receive a non-negative utility at each date and in each state.³

The paper is organized as follows. Section 2 presents the basic model. In Section 3, we analyze the dynamics of quality under symmetric and asymmetric information when the regulator contracts with a sequence of short-lived agents. In Section 4, we consider the case of a single long-lived agent. Section 5 concludes.

²See also Gaudet, Lasserre and Van Long (1997) and Gärtner (2004) for related models. Gaudet, Lasserre and Van Long (1995) also consider a dynamic contracting model with a physical state variable. In a context of nonrenewable natural resource exploitation, they analyze how the dynamic relationship between the owner and the manager of a mine is affected by the level of reserves.

³We briefly consider the case of short-term contracts in Subsection 4.4. The main insight is that the absence of commitment generates a ratchet effect: an agent experiencing a high quality shock today anticipates that, as a result of this, quality tomorrow will be higher, which in turn leads to lower continuation rents.

2. THE BASIC MODEL

Our model departs in two ways from standard regulation models such as those studied by Baron and Myerson (1982) or Laffont and Tirole (1986). First, we consider a multi-period environment. Second, we assume that there exists a physical state variable that dynamically links periods to each other. This state variable is interpreted as the quality of a public good or service, such as a road or electricity network.

The basic model focuses on the dynamics of quality, leaving aside the complex issues related to the dynamic provision of incentives which are examined in Section 4. For this purpose, we consider a benevolent regulator who delegates the management of the quality of the public good to a sequence of agents, one for each period. The per-period consumer surplus generated by a public good of quality $q \geq 0$ is denoted $S(q)$. The function S is bounded, continuously differentiable, strictly increasing and strictly concave over \mathbb{R}_{++} .

The quality of the public good evolves over time as a function of the agents' effort to maintain it and of exogenous quality shocks. Specifically, if the quality of the public good at date $t = 0, 1, \dots$ is q_t , then its quality at date $t + 1$ is:

$$q_{t+1} = \delta q_t + e_t + \theta_t, \quad (1)$$

where $\delta \in (0, 1)$ is a depreciation factor, e_t is the maintenance effort exerted by the date t agent, and θ_t is the date t quality shock. The parameter δ measures the extent to which maintenance efforts and quality shocks have durable effects. The quality shocks $\{\theta_t\}_{t=0}^{\infty}$ are independently and identically distributed across periods, with support $\{\bar{\theta}, \underline{\theta}\}$ such that $\bar{\theta} > \underline{\theta} \geq 0$, and we let $\Delta\theta = \bar{\theta} - \underline{\theta}$. For any date $t = 0, 1, \dots$, we denote by $\nu \in (0, 1)$ the probability that $\theta_t = \bar{\theta}$, and we let $E_\theta = E[\theta_t]$ and $\text{Var}_\theta = \text{Var}[\theta_t]$.

An agent exerting a maintenance effort e incurs a disutility $\psi(e)$ in monetary units. The function ψ is continuously differentiable, strictly increasing and strictly convex over \mathbb{R}_+ , and satisfies $\psi(0) = \psi'(0) = 0$. It is analytically convenient to extend the function ψ to the whole real line by setting $\psi = 0$ over \mathbb{R}_- . To guarantee that the regulator's objective function is concave, we also assume that ψ has a convex derivative over \mathbb{R}_+ .

Agents are compensated for their efforts by monetary transfers. Given effort level e_t and monetary transfer u_t , the date t agent's overall utility is then:

$$U_t = u_t - \psi(e_t). \quad (2)$$

Each agent's outside opportunity is normalized to zero. In addition, we assume that, to accept working for the regulator, each agent must receive a non-negative utility in each state of nature. An interpretation is that agents have infinite risk aversion below zero wealth (see for instance Laffont and Martimort (2005, §3.5)).

As in Laffont and Tirole (1986), distortionary taxation inflicts a disutility $\$(1 + \lambda)$ on consumers in order to levy $\$1$ for the state, where $\lambda > 0$ is the shadow cost of public funds. Hence, the date t net consumer surplus is $S(q_t) - (1 + \lambda)u_t$, so that by (2), the corresponding utilitarian social welfare is:

$$S(q_t) - (1 + \lambda)u_t + u_t - \psi(e_t) = S(q_t) - (1 + \lambda)\psi(e_t) - \lambda U_t. \quad (3)$$

The regulator is far-sighted and discounts future payoffs with a discount factor $\beta \in (0, 1)$. By (3), the expected discounted social welfare at date zero is thus equal to:

$$E \left[\sum_{t=0}^{\infty} \beta^t [S(q_t) - (1 + \lambda)\psi(e_t) - \lambda U_t] \right]. \quad (4)$$

We assume throughout that quality is verifiable, so that the regulator can reward or punish the agents directly as a function of quality improvements. By (1), this means that, at each date, the sum of the agent's effort and of the quality shock is verifiable. Under symmetric information, efforts and shocks are themselves verifiable. Under asymmetric information, however, neither efforts nor shocks are verifiable, as in Lewis and Sappington (1991), while the shocks are privately observed by the agents.⁴

3. REGULATING QUALITY WITH A SEQUENCE OF AGENTS

3.1. Regulation under Symmetric Information

As a benchmark, we consider the symmetric information situation in which not only quality improvements, but also the extent to which these can be attributed to agents' efforts or to quality shocks are verifiable variables. As a result of this, regulatory contracts can be directly made contingent on agents' efforts. The existence of a shadow cost of public funds implies that agents receive no rent at the optimum, that is $u_t = \psi(e_t)$ for all $t = 0, 1, \dots$. By (1)–(2) and (4), the regulator's problem is then to find a sequence of quality levels $\{q_t\}_{t=1}^{\infty}$, where each q_t is contingent on the history of shocks $(\theta_0, \dots, \theta_{t-1})$ up to date t , that solves:

$$V_1(q_0) = \sup \left\{ \mathbb{E} \left[\sum_{t=0}^{\infty} \beta^t [S(q_t) - (1 + \lambda)\psi(q_{t+1} - \delta q_t - \theta_t)] \right] \right\},$$

given any initial quality level q_0 . This is a standard dynamic programming problem that can be tackled by usual techniques. The symmetric information social value function $V_1 : \mathbb{R}_+ \rightarrow \mathbb{R}$ is the unique bounded solution to the Bellman equation:

$$V_1(q) = \max_{(\bar{e}, \bar{q}, \underline{e}, \underline{q})} \{ S(q) - \nu(1 + \lambda)\psi(\bar{e}) - (1 - \nu)(1 + \lambda)\psi(\underline{e}) + \nu\beta V_1(\bar{q}) + (1 - \nu)\beta V_1(\underline{q}) \}, \quad (5)$$

subject to the state transition constraints:

$$\bar{q} = \delta q + \bar{e} + \bar{\theta}, \quad (6)$$

$$\underline{q} = \delta q + \underline{e} + \underline{\theta}, \quad (7)$$

and the feasibility constraints:

$$\bar{q} \geq 0, \quad (8)$$

$$\underline{q} \geq 0. \quad (9)$$

Standard considerations (see for instance Stokey and Lucas (1989)) yield our first result.

Lemma 1. *The symmetric information social value function V_1 is bounded, continuously differentiable, strictly increasing and strictly concave over \mathbb{R}_{++} .*

An implication of $\psi'(0) = 0$ is that it is always optimal for the regulator to induce strictly positive effort on the part of the current agent, no matter the current level of quality or the

⁴This modelling assumption is naturally reminiscent of Laffont and Tirole's (1986) regulation model.

current quality shock. That is, the feasibility constraints (8)–(9) are never binding. Denote by $\bar{e}_1(q)$, $\bar{q}_1(q)$, $\underline{e}_1(q)$, $\underline{q}_1(q)$ the optimal choices in (5), which are uniquely determined. Then the first-order conditions for (5) read as:

$$\beta V_1'(\bar{q}_1(q)) = (1 + \lambda)\psi'(\bar{q}_1(q) - \delta q - \bar{\theta}), \quad (10)$$

$$\beta V_1'(\underline{q}_1(q)) = (1 + \lambda)\psi'(\underline{q}_1(q) - \delta q - \underline{\theta}). \quad (11)$$

Since V_1 is strictly concave and ψ strictly convex over \mathbb{R}_+ , it follows from (10)–(11) that $\bar{q}_1(q) > \underline{q}_1(q)$: for a given level of public good quality, a high quality shock today leads to a higher quality tomorrow than a low quality shock. Along with (6)–(7), (10)–(11) further imply that $\bar{e}_1(q) < \underline{e}_1(q)$: an agent facing a high quality shock exerts less effort than one facing a low quality shock, and therefore receives a lower transfer. Note that, as a result of this, $-(1 + \lambda)\psi(\bar{e}_1(q)) + \beta V_1(\bar{q}_1(q)) > -(1 + \lambda)\psi(\underline{e}_1(q)) + \beta V_1(\underline{q}_1(q))$, so that the regulator benefits from a high quality shock.

Using again the strict concavity of V_1 and the strict convexity of ψ over \mathbb{R}_+ , it is easy to check from (10)–(11) that the mappings \bar{q}_1 and \underline{q}_1 are strictly increasing, while the mappings \bar{e}_1 and \underline{e}_1 are strictly decreasing. Since $\psi'(0) = 0$ and $\lim_{q \rightarrow \infty} \bar{q}_1(q) = \lim_{q \rightarrow \infty} \underline{q}_1(q) = \infty$ by (6)–(7), and since V_1 is bounded, a further implication of (10)–(11) is that $\lim_{q \rightarrow \infty} \bar{e}_1(q) = \lim_{q \rightarrow \infty} \underline{e}_1(q) = 0$. These properties reflect the fact that the agents' maintenance efforts become less important from the regulator's viewpoint as quality improves. As a result of this, $\lim_{q \rightarrow \infty} V_1(q) = \lim_{q \rightarrow \infty} \frac{S(q)}{1-\beta}$.

3.2. Regulation under Asymmetric Information

We now turn to the case in which neither agents' efforts nor quality shocks are verifiable. An asymmetry of information then arises because, once in charge, agents become privately informed of the current quality shock. Since they must receive a non-negative utility in each state, eliciting this information from them is socially costly.⁵ The task of the regulator is to design appropriate incentive schemes to overcome this hidden information problem and the resulting moral hazard problem.

An incentive contract between the regulator and the date t agent specifies a transfer-quality pair for each realization of the date t quality shock, which will be henceforth referred to as the date t agent's type. For a given level of public good quality q , a contract is thus a 4-tuple $(\bar{u}, \bar{q}, \underline{u}, \underline{q})$. Incentive compatibility requires that:

$$\bar{u} - \psi(\bar{q} - \delta q - \bar{\theta}) \geq \underline{u} - \psi(\underline{q} - \delta q - \bar{\theta}), \quad (12)$$

$$\underline{u} - \psi(\underline{q} - \delta q - \underline{\theta}) \geq \bar{u} - \psi(\bar{q} - \delta q - \underline{\theta}). \quad (13)$$

Since the agent must receive a non-negative utility in each state, an incentive feasible contract must also satisfy the following participation constraints:

$$\bar{u} - \psi(\bar{q} - \delta q - \bar{\theta}) \geq 0, \quad (14)$$

$$\underline{u} - \psi(\underline{q} - \delta q - \underline{\theta}) \geq 0. \quad (15)$$

⁵In the absence of such an ex-post participation constraint, the regulator could achieve the symmetric information outcome through appropriate ex-ante contracting (D'Aspremont and Gérard-Varet (1979)).

It is easy to check that the optimal contract under symmetric information does not satisfy the incentive constraint (12) of the high type agent. As usual with this type of models, the incentive constraint (12) of the high type agent and the participation constraint (15) of the low type agent together imply the participation constraint (14) of the high type agent. To maximize the expected discounted social welfare, we shall momentarily neglect the incentive constraint (13) of the low type agent, and later check that the solution thus obtained satisfies this constraint.

Denote by $\bar{U}(q) = \bar{u} - \psi(\bar{q} - \delta q - \bar{\theta})$ and $\underline{U}(q) = \underline{u} - \psi(\underline{q} - \delta q - \underline{\theta})$ the rents left to the agent under the contract $(\bar{u}, \bar{q}, \underline{u}, \underline{q})$, given current quality q . The participation constraint (15) of the low type agent can be rewritten as:

$$\underline{U}(q) \geq 0, \quad (16)$$

while the incentive constraint (12) of the high type agent can be rewritten as:

$$\bar{U}(q) \geq \underline{U}(q) + \Phi(q, \underline{q}), \quad (17)$$

where the function Φ is defined by:

$$\Phi(q, \underline{q}) = \psi(\underline{q} - \delta q - \underline{\theta}) - \psi(\bar{q} - \delta q - \bar{\theta}). \quad (18)$$

Intuitively, $\Phi(q, \underline{q})$ is the informational rent that must be left to the high type agent when the low type agent improves quality from q to \underline{q} . Since ψ is convex, $\Phi(q, \cdot)$ is increasing for any value of q , while $\Phi(\cdot, \underline{q})$ is decreasing for any value of \underline{q} .⁶ Moreover, since ψ' is convex, $\Phi(q, \cdot)$ is convex for any value of q , and $\Phi(\cdot, \underline{q})$ is convex for any value of \underline{q} . These properties will ensure that the regulator's objective function is concave, and that the social value function is strictly increasing and strictly concave.

Since leaving rents to the agent is socially costly, it will be optimal to let (16)–(17) be binding. That is, one will have:

$$U_t = \mathbb{E}_t[\Phi(q_t, q_{t+1}) | \theta_t = \underline{\theta}] 1_{\{\theta_t = \bar{\theta}\}} \quad (19)$$

for all $t = 0, 1, \dots$. By (1), (4) and (19), the regulator's problem is then to find a sequence of quality levels $\{q_t\}_{t=1}^{\infty}$, where each q_t is contingent on the history of shocks $(\theta_0, \dots, \theta_{t-1})$ up to date t , that solves:

$$V_2(q_0) = \sup \left\{ \mathbb{E} \left[\sum_{t=0}^{\infty} \beta^t \left[S(q_t) - (1 + \lambda)\psi(q_{t+1} - \delta q_t - \theta_t) - \frac{\lambda\nu}{1 - \nu} \Phi(q_t, q_{t+1}) 1_{\{\theta_t = \underline{\theta}\}} \right] \right] \right\},$$

given any initial quality level q_0 . In analogy with (5), the asymmetric information social value function $V_2 : \mathbb{R}_+ \rightarrow \mathbb{R}$ is the unique bounded solution to the Bellman equation:

$$\begin{aligned} V_2(q) = \max_{(\bar{e}, \bar{q}, \underline{e}, \underline{q})} \{ & S(q) - \nu(1 + \lambda)\psi(\bar{e}) - (1 - \nu)(1 + \lambda)\psi(\underline{e}) - \lambda\nu\Phi(q, \underline{q}) \\ & + \nu\beta V_2(\bar{q}) + (1 - \nu)\beta V_2(\underline{q}) \}, \end{aligned} \quad (20)$$

subject to constraints (6)–(9). The following result parallels Lemma 1.

⁶However, for $\underline{q} < \delta q + \underline{\theta}$, these functions are constant and equal to zero.

Lemma 2. *The asymmetric information social value function V_2 is bounded, continuously differentiable, strictly increasing and strictly concave over \mathbb{R}_{++} .*

As in the symmetric information benchmark, the condition $\psi'(0) = 0$ implies that the feasibility constraints (8)–(9) are never binding: it is always optimal for the regulator to induce strictly positive effort on the part of the current agent, no matter his type or the current level of quality. Denote by $\bar{e}_2(q)$, $\bar{q}_2(q)$, $\underline{e}_2(q)$, $\underline{q}_2(q)$ the optimal choices in (20), which are uniquely determined. Then the first-order conditions for (20) read as:

$$\beta V_2'(\bar{q}_2(q)) = (1 + \lambda)\psi'(\bar{q}_2(q) - \delta q - \bar{\theta}), \quad (21)$$

$$\beta V_2'(\underline{q}_2(q)) = (1 + \lambda)\psi'(\underline{q}_2(q) - \delta q - \underline{\theta}) + \frac{\lambda\nu}{1 - \nu} \Phi_2(q, \underline{q}_2(q)). \quad (22)$$

Since V_2 is strictly concave and ψ strictly convex over \mathbb{R}_+ , and since $\Phi(q, \cdot)$ is convex, it follows from (21)–(22) that $\bar{q}_2(q) > \underline{q}_2(q)$. This in turn implies that the neglected constraint (13) is satisfied by our candidate solution. Indeed, because (16)–(17) are binding, (13) is equivalent to $\Phi(q, \bar{q}_2(q)) \geq \Phi(q, \underline{q}_2(q))$, which holds as $\Phi(q, \cdot)$ is increasing. Because of the second term on the right-hand side of (22), which corresponds to the distortion due to asymmetric information, the comparison between the effort levels $\bar{e}_2(q)$ and $\underline{e}_2(q)$ is ambiguous, unlike in the symmetric information benchmark.

Using again the strict concavity of V_2 and the strict convexity of ψ over \mathbb{R}_+ , together with the convexity of ψ' and the definition (18) of Φ , it is easy to check from (21)–(22) that the mappings \bar{q}_2 and \underline{q}_2 are strictly increasing, while the mappings \bar{e}_2 and \underline{e}_2 are strictly decreasing. The strict convexity of ψ over \mathbb{R}_+ , along with the monotonicity of \underline{e}_2 , also implies that the informational rent of a high type agent,

$$\Phi(q, \underline{q}_2(q)) = \psi(\underline{e}_2(q)) - \psi(\underline{e}_2(q) - \Delta\theta), \quad (23)$$

is a strictly decreasing function of public good quality q . It follows from (21)–(22) that, as in the symmetric information benchmark, $\lim_{q \rightarrow \infty} \bar{e}_2(q) = \lim_{q \rightarrow \infty} \underline{e}_2(q) = 0$. By (23), this implies that the informational rent of the high type agent vanishes as quality gets large, $\lim_{q \rightarrow \infty} \Phi(q, \underline{q}_2(q)) = 0$. The intuition for this result is that when quality improves, it becomes less important for the regulator to incite the low type agent to exert effort. In particular, $\lim_{q \rightarrow \infty} V_2(q) = \lim_{q \rightarrow \infty} \frac{S(q)}{1 - \beta}$ as in the symmetric information benchmark.

Remark. Using the Envelope Theorem for (5) and (20), it is easy to check that both V_1' and V_2' are strictly greater than S' . It follows that, in both regulatory contexts, the regulator induces a higher quality service than in the corresponding one-period version of the model.⁷ This reflects that, because quality is durable in our model, the benefits of producing a higher quality today arise directly through an increase of future consumer surpluses, and, indirectly, through a reduced cost of supplying quality in the future.

3.3. Comparing the Two Regulatory Environments

A key insight of our analysis is that the social value of public good quality depends on the regulatory environment. We argue in this section that, as a result of this, asymmetric

⁷This result is in line with Lewis and Yildirim (2002, Proposition 3(i)).

information typically leads to distortions in both types of agents' maintenance efforts relative to their symmetric information levels. To establish this point, we first present some general analytical results, and then examine a linear-quadratic specification of the model.

3.3.1. General Results

To contrast the outcomes of the regulation game under symmetric and under asymmetric information, it is helpful to compare the social value functions V_1 and V_2 . Formally, the only difference between the Bellman equations (5) and (20) that implicitly define V_1 and V_2 lies in the informational rent $\lambda\nu\Phi(q, \underline{q})$ that appears on the right-hand side of (20). As observed above, this rent is a decreasing function of the current public good quality level q . This suggests that under asymmetric information, an additional incentive to increase public good quality is to reduce future informational rents. Accordingly, the marginal social value of quality is strictly higher under asymmetric information than under symmetric information, as shown by the following result.

Proposition 1. *For any $q > 0$, $V_2'(q) > V_1'(q)$.*

Along with (10) and (21), Proposition 1 implies that $\bar{q}_2(q) > \bar{q}_1(q)$ for any $q > 0$. That is, for a given level of public good quality $q > 0$, and conditional on a high quality shock occurring, the regulator induces more effort from the agent under asymmetric information than under symmetric information, $\bar{e}_2(q) > \bar{e}_1(q)$. The intuition is straightforward: given a high quality shock, the marginal cost of effort is the same in both regulatory environments, while by Proposition 1, the marginal benefit of effort is higher under asymmetric information. In line with the dynamic rent extraction effect outlined above, the regulator thus takes advantage of facing a high type agent today to build better quality for tomorrow. It should be noted that this over-provision of quality contrasts with the prediction of a one-period model, in which high type agents would optimally exert the same level of effort under symmetric information as under asymmetric information. Because of the additional distortion term on the right-hand side of (22), it is not possible in general to rank the quality levels $\underline{q}_1(q)$ and $\underline{q}_2(q)$ conditional on a low shock to quality. The intuition is that asymmetric information raises both the marginal benefit and the marginal cost of exerting effort given a low quality shock. A robust prediction of our model is however that if sufficiently many high quality shocks occur, asymmetric information will lead to over-provision of public good quality relative to the symmetric information benchmark.

The differences between the two regulatory environments do not disappear in the long run, as can be shown by studying the asymptotic distribution of quality. Specifically, let P_1 and P_2 be the probability transition functions over quality levels respectively induced by (10)–(11) and (21)–(22). That is, for each $i = 1, 2$, and for any $q \geq 0$, $P_i(q, \cdot)$ is the law of a random variable that takes the value $\bar{q}_i(q)$ with probability ν , and the value $\underline{q}_i(q)$ with probability $1 - \nu$. Then the following holds.

Proposition 2. *P_1 and P_2 have unique invariant probability measures μ_1 and μ_2 with compact supports $\text{supp } \mu_1$ and $\text{supp } \mu_2$ such that $\max \text{supp } \mu_2 > \max \text{supp } \mu_1$.*

For any initial quality level q_0 , the distribution of quality will converge weakly to μ_1 or μ_2 depending on the regulatory environment. Since the upper bound of the support of the

asymptotic distribution of quality is strictly larger under asymmetric information than under symmetric information, one will in the long run observe high quality levels in the former case that cannot be achieved in the latter case. Therefore over-provision of public good quality can persist in the long run under asymmetric information.

3.3.2. The Linear-Quadratic Case

To obtain sharper predictions, we consider a special case of our model, in which the underlying surplus and cost functions are quadratic,

$$S(q) = aq - \frac{b}{2}q^2, \quad (24)$$

$$\psi(e) = \frac{c}{2} \max\{e, 0\}^2, \quad (25)$$

for some strictly positive parameters a , b and c . It should be noted that the surplus function S defined by (24) makes sense only as long $q < \frac{a}{b}$. Along standard lines (Stokey and Lucas (1989, §4.4)), we first solve the Bellman equations (5) and (20) without taking into account this restriction. We then check under which conditions and over which ranges the resulting analytical solutions are economically meaningful. The appropriate parameter restrictions amount to make $\underline{\theta}$ and $\Delta\theta$ close enough to zero, see the Appendix for a precise statement.

The linear-quadratic specification (24)–(25) ensures that the social value functions V_1 and V_2 are themselves quadratic over the relevant ranges, that is, for each $i = 1, 2$,

$$V_i(q) = A_iq - \frac{B_i}{2}q^2 + C_i, \quad (26)$$

for some strictly positive parameters A_i , B_i and C_i , while the optimal policy functions \bar{q}_1 , \underline{q}_1 , \bar{q}_2 and \underline{q}_2 are linear over the relevant ranges,

$$\bar{q}_1(q) = \frac{\beta A_1 + c(1 + \lambda)(\bar{\theta} + \delta q)}{\beta B_1 + c(1 + \lambda)}, \quad (27)$$

$$\underline{q}_1(q) = \frac{\beta A_1 + c(1 + \lambda)(\underline{\theta} + \delta q)}{\beta B_1 + c(1 + \lambda)}, \quad (28)$$

$$\bar{q}_2(q) = \frac{\beta A_2 + c(1 + \lambda)(\bar{\theta} + \delta q)}{\beta B_2 + c(1 + \lambda)}, \quad (29)$$

$$\underline{q}_2(q) = \frac{\beta A_2 - \frac{\lambda\nu}{1-\nu} c\Delta\theta + c(1 + \lambda)(\underline{\theta} + \delta q)}{\beta B_2 + c(1 + \lambda)}. \quad (30)$$

To determine the optimal policy functions, what matters are the values of A_1 , B_1 , A_2 and B_2 . These are easily obtained by inserting the policy functions (27)–(28) and (29)–(30) into the Bellman equations (5) and (20) and then using (26) to identify terms. An immediate result is that the coefficients B_1 and B_2 coincide. This reflects the property that, in the linear-quadratic specification, and over the relevant range, the informational rent $\Phi(q, \underline{q})$ is a linear function of (q, \underline{q}) ,

$$\Phi(q, \underline{q}) = c\Delta\theta \left(\underline{q} - \delta q - \frac{\bar{\theta} + \underline{\theta}}{2} \right). \quad (31)$$

It should be noted that, since the marginal social value of quality is higher under asymmetric information than under symmetric information, $A_2 > A_1$.

An immediate implication of (27)–(30) is that asymmetric information increases the wedge between qualities following a high and a low quality shock, $\bar{q}_2(q) - \underline{q}_2(q) > \bar{q}_1(q) - \underline{q}_1(q)$. Our next result strengthens this insight.

Proposition 3. *Over the relevant range, the spread of public good qualities is larger under asymmetric information than under symmetric information,*

$$\bar{q}_2(q) > \bar{q}_1(q) > \underline{q}_1(q) > \underline{q}_2(q). \quad (32)$$

The intuition for this result can be grasped by comparing the objective functions of the regulator in (5) and (20). Under symmetric information, the marginal value of date $t + 1$ quality is the same following a high or a low quality shock, namely $\beta(A_1 - Bq_{t+1})$. By contrast, under asymmetric information, the need to concede an informational rent to high type agents introduces a wedge between the marginal value of quality following a high quality shock, $\beta(A_2 - Bq_{t+1})$, and the marginal value of quality following a low quality shock, net of the incentive cost, $\beta(A_2 - Bq_{t+1}) - \frac{\lambda\nu}{1-\nu} c\Delta\theta$. It turns out that:

$$\beta A_1 < \beta A_2 < \beta A_1 + \frac{\lambda\nu}{1-\nu} c\Delta\theta,$$

so that the higher marginal social benefit of quality due to asymmetric information does not fully offset the incentive cost given a low quality shock. As shown by (32), this simultaneously leads to over-provision of quality following a high quality shock, $\bar{q}_2(q) > \bar{q}_1(q)$, and to under-provision of quality following a low quality shock, $\underline{q}_1(q) > \underline{q}_2(q)$. While the first effect reflects a dynamic rent extraction motive, as discussed in Subsection 3.3.1, the second effect reflects a static rent extraction motive: the regulator induces less maintenance effort from a low type agent in order to reduce the rent she would have to leave today to a high type agent. Proposition 3 shows that this static rent extraction effect dominates the dynamic effect. As can be checked from (29)–(30), this may cause a high type agent to exert more effort than a low type agent given the same current level of quality, thus reversing the prediction of the symmetric information benchmark.

In light of these results, it is natural to investigate which of the static and the dynamic rent extraction motives dominates from an ex-ante perspective. To this end, we examine the time-series properties of public good quality under the two regulatory regimes. Denote by $\{q_{1,t}\}_{t=0}^{\infty}$ and $\{q_{2,t}\}_{t=0}^{\infty}$ the stochastic processes of quality under symmetric and asymmetric information respectively. Provided initial quality is low, these processes will remain in the appropriate ranges over which our analytical solutions to (5) and (20) are economically meaningful. One has the following result.

Proposition 4. *Public good quality grows at a lower expected rate under asymmetric information than under symmetric information,*

$$\mathbb{E}[q_{1,t+1} | q_{1,t} = q] > \mathbb{E}[q_{2,t+1} | q_{2,t} = q], \quad (33)$$

while the variance of public good quality growth is higher under asymmetric information than under symmetric information,

$$\text{Var}[q_{2,t+1} | q_{2,t} = q] > \text{Var}[q_{1,t+1} | q_{1,t} = q]. \quad (34)$$

From an ex-ante perspective, the static rent extraction motive therefore dominates the dynamic rent extraction motive, and asymmetric information leads to on average lower and more volatile quality growth. These properties carry over to the long run. For each $i = 1, 2$, let $E_{\mu_i} = \int q d\mu_i(q)$ and $\text{Var}_{\mu_i} = \int (q - E_{\mu_i})^2 d\mu_i(q)$ be the expectation and variance of quality under the invariant measure μ_i . Weak convergence of the distribution of quality to μ_i ensures that $\lim_{t \rightarrow \infty} E[q_{i,t}] = E_{\mu_i}$ and $\lim_{t \rightarrow \infty} \text{Var}[q_{i,t}] = \text{Var}_{\mu_i}$, no matter the initial quality level q_0 . The following result provides the long run analogues of (33) and (34).

Proposition 5. *In the long run, public good quality is on average lower under asymmetric information than under symmetric information,*

$$E_{\mu_1} > E_{\mu_2}, \quad (35)$$

while its variance is higher under asymmetric information than under symmetric information,

$$\text{Var}_{\mu_2} > \text{Var}_{\mu_1}. \quad (36)$$

4. REGULATING QUALITY WITH A SINGLE AGENT

In this section, we suppose that there is a single agent the regulator can contract with over time. In contrast with the regulatory environments studied in Section 3, the regulator can now provide dynamic incentives to exert effort. To simplify the analysis, we assume that the contractual relationship lasts only two periods: the agent exerts effort at dates 0 and 1 to provide public good quality at dates 1 and 2. The model remains otherwise the same, except that we no longer require the discount factor β to be smaller than one.⁸ Our setup therefore differs in two ways from the dynamic regulation models previously studied in the literature (Baron and Besanko (1984), Freixas, Guesnerie and Tirole (1985), Laffont and Tirole (1988, 1990), Lewis and Sappington (1997)). First, the only link between periods is physical and not informational, as quality shocks are independently distributed across periods. Second, as in the basic model, only contracts that leave the agent with a non-negative utility at each date and in each state can be enforced.⁹

4.1. Regulation under Symmetric Information

The symmetric information benchmark does not differ significantly from that derived in Subsection 3.1, except for the finite horizon. In each period, the agent receives no rent at the optimum, that is $u_t = \psi(e_t)$ for each $t = 0, 1$. For each realization (θ_0, θ_1) of the agent's types at dates 0 and 1, we denote by $q_{1,1}(\theta_0)$, $q_{1,2}(\theta_0, \theta_1)$ the optimal contingent public good quality levels at dates 1 and 2, and by $e_{1,0}(\theta_0)$ and $e_{1,1}(\theta_0, \theta_1)$ the corresponding optimal contingent effort levels at dates 0 and 1. Then, using the Envelope Theorem, the first-order conditions for the regulator's problem can be compactly written as follows:

$$\beta S'(q_{1,2}(\theta_0, \theta_1)) = (1 + \lambda)\psi'(e_{1,1}(\theta_0, \theta_1)), \quad (37)$$

$$\beta S'(q_{1,1}(\theta_0)) + \beta\delta(1 + \lambda)E[\psi'(e_{1,1}(\theta_0, \theta_1)) | \theta_0] = (1 + \lambda)\psi'(e_{1,0}(\theta_0)) \quad (38)$$

⁸A large discount factor may reflect the fact that the accounting period for future production stages exceeds that for the current production stage, or that later projects are operated on a larger scale than early projects.

⁹Laffont and Martimort (2002, §8.1.3) study how this constraint affects the outcome of a repeated agency model in which the link between periods is purely informational.

where $q_{1,1}(\theta_0) = \delta q_0 + e_{1,0}(\theta_0) + \theta_0$ and $q_{1,2}(\theta_0, \theta_1) = \delta q_{1,1}(\theta_0) + e_{1,1}(\theta_0, \theta_1) + \theta_1$ for each $(\theta_0, \theta_1) \in \{\bar{\theta}, \underline{\theta}\}^2$. The date 1 first-order condition (37) is self-explanatory. The second term on the left-hand side of the date 0 first-order condition (38) reflects the dynamic benefit of date 0 maintenance efforts, which is to reduce the cost of enhancing quality at date 1.

Remark. As in the stationary setting of Subsection 3.1, the optimal symmetric information quality provision depends at each date only on the current level of quality, as well as on the current quality shock. We denote accordingly by $\bar{Q}_{1,1}$ and $\underline{Q}_{1,1}$ the optimal policy functions describing how date 0 quality is mapped into date 1 quality depending of the realization of the date 0 quality shock. Similarly, $\bar{Q}_{1,2}$ and $\underline{Q}_{1,2}$ denote the optimal policy functions describing how date 1 quality is mapped into date 2 quality depending of the realization of the date 1 quality shock.

4.2. Regulation under Asymmetric Information: Full Commitment

We now turn to the case where neither the agent's efforts nor the quality shocks are verifiable, but long-term regulatory contracts can be perfectly enforced, subject to the constraint that the agent receives a non-negative utility at each date and in each state. Because of full commitment, the Revelation Principle applies. For each pair of reports $(\hat{\theta}_0, \hat{\theta}_1) \in \{\bar{\theta}, \underline{\theta}\}^2$ by the agent, we denote by $u_{2,0}(\hat{\theta}_0)$, $u_{2,1}(\hat{\theta}_0, \hat{\theta}_1)$ the optimal contingent transfer levels at dates 0 and 1, and by $q_{2,1}(\hat{\theta}_0)$, $q_{2,2}(\hat{\theta}_0, \hat{\theta}_1)$ the optimal contingent quality levels at dates 1 and 2.¹⁰

4.2.1. Preliminaries

We solve for the optimal contract in two steps. Consider first the date 1 contract. Since the regulator and the agent are risk-neutral and discount future utility at the same rate, there is no loss of generality in assuming that, whenever $\theta_1 = \underline{\theta}$, the agent receives no rent at date 1 no matter his first period report $\hat{\theta}_0$. Therefore,

$$u_{2,1}(\hat{\theta}_0, \underline{\theta}) = \psi(q_{2,2}(\hat{\theta}_0, \underline{\theta}) - \delta q_{2,1}(\hat{\theta}_0) - \underline{\theta})$$

for each $\hat{\theta}_0 \in \{\bar{\theta}, \underline{\theta}\}$. Standard considerations then imply that, given a date 0 report $\hat{\theta}_0$, the rent that the agent receives at date 1 whenever $\theta_1 = \bar{\theta}$ is given by $\Phi(q_{2,1}(\hat{\theta}_0), q_{2,2}(\hat{\theta}_0, \underline{\theta}))$. In turn, incentive compatibility at date 0 requires that:

$$u_{2,0}(\bar{\theta}) - \psi(q_{2,1}(\bar{\theta}) - \delta q_0 - \bar{\theta}) + \nu\beta\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) \geq u_{2,0}(\underline{\theta}) - \psi(q_{2,1}(\underline{\theta}) - \delta q_0 - \bar{\theta}) + \nu\beta\Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta})), \quad (39)$$

$$u_{2,0}(\underline{\theta}) - \psi(q_{2,1}(\underline{\theta}) - \delta q_0 - \underline{\theta}) + \nu\beta\Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta})) \geq u_{2,0}(\bar{\theta}) - \psi(q_{2,1}(\bar{\theta}) - \delta q_0 - \underline{\theta}) + \nu\beta\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})). \quad (40)$$

¹⁰We hereby restrict the analysis to deterministic mechanisms. Since the objective function of the regulator is not necessarily concave, she may increase social welfare by using stochastic mechanisms (see Laffont and Martimort (2002, §2.13)). Such mechanisms are however difficult to implement in practice, which is why we chose to rule them out. This problem does not arise in the linear-quadratic specification, because the function Φ is then linear over the relevant range, see (31).

Since the agent must receive a non-negative utility at each date and in each state, the optimal contract must also satisfy the following date 0 participation constraints:

$$u_{2,0}(\bar{\theta}) - \psi(q_{2,1}(\bar{\theta}) - \delta q_0 - \bar{\theta}) \geq 0, \quad (41)$$

$$u_{2,0}(\underline{\theta}) - \psi(q_{2,1}(\underline{\theta}) - \delta q_0 - \underline{\theta}) \geq 0. \quad (42)$$

It will be convenient to define the equilibrium levels of effort as $e_{2,0}(\theta_0) = q_{2,1}(\theta_0) - \delta q_0 - \theta_0$ and $e_{2,1}(\theta_0, \theta_1) = q_{2,2}(\theta_0, \theta_1) - \delta q_{2,1}(\theta_0) - \theta_1$ for each $(\theta_0, \theta_1) \in \{\bar{\theta}, \underline{\theta}\}^2$.

It is not a priori clear which of the constraints (39)–(40) and (41)–(42) are binding at the optimum. This is because we require the optimal contract to satisfy not only intertemporal participation constraints, but also the more restrictive constraints (41)–(42) on date 0 utilities. As a result of this, one cannot for instance argue that the incentive constraint of the high type agent and the participation constraint of the low type agent together imply the participation constraint of the high type agent, as would be the case in a standard model. Still, the following result holds.

Proposition 6. *In the optimal long-term contract, $q_{2,1}(\bar{\theta}) \geq q_{2,1}(\underline{\theta})$, and constraints (39) and (42) are binding. Therefore, the date 0 rent of a high type agent is given by:*

$$U_{2,0}(\bar{\theta}) = \Phi(q_0, q_{2,1}(\underline{\theta})) + \nu\beta[\Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta})) - \Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta}))]. \quad (43)$$

Equation (43) has a natural interpretation: the first term on the right-hand side is a standard static informational rent, while the second term represents the expected discounted gain in terms of date 1 rents from pretending to be a low type agent rather than a high type agent at date 0.

The intuition for Proposition 6 is as follows. Suppose it were optimal to leave a low type agent a strictly positive rent at date 0. One can verify that, in analogy with (43), this rent must be given by:

$$-\Phi(q_0, q_{2,1}(\bar{\theta})) + \nu\beta[\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) - \Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta}))]. \quad (44)$$

The incentive for a low type agent to misrepresent his type at date 0 only comes from the bracketed term in (44), which represents the date 1 rent differential. Since the function Φ is decreasing in its first argument and $q_{2,1}(\bar{\theta}) \geq q_{2,1}(\underline{\theta})$, this term can be strictly positive only if $q_{2,2}(\bar{\theta}, \underline{\theta})$ is significantly larger than $q_{2,2}(\underline{\theta}, \underline{\theta})$. Given (44), the total expected discounted rents at date 0 amount to:

$$\begin{aligned} & (1 - \nu)\{-\Phi(q_0, q_{2,1}(\bar{\theta})) + \nu\beta[\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) - \Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta}))]\} \\ & \quad + \nu^2\beta\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) + \nu(1 - \nu)\beta\Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta})) \\ & = -(1 - \nu)\Phi(q_0, q_{2,1}(\bar{\theta})) + \nu\beta\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})). \end{aligned} \quad (45)$$

Since (45) is independent of $q_{2,2}(\underline{\theta}, \underline{\theta})$ and (42) is assumed slack, the optimal determination of $q_{2,2}(\underline{\theta}, \underline{\theta})$ involves no rent-efficiency trade-off. By contrast, since (45) is strictly increasing in $q_{2,2}(\bar{\theta}, \underline{\theta})$, it is optimal to distort $q_{2,2}(\bar{\theta}, \underline{\theta})$ downward. This in turn amplifies the reduction

of date 1 rent due to the fact that $q_{2,1}(\bar{\theta}) \geq q_{2,1}(\underline{\theta})$. As is easy to check, this implies that the bracketed term in (44) is strictly negative, reflecting that a low type agent has no incentive to misrepresent his type at date 0. As a result of this, he cannot earn a strictly positive rent at date 0, contrary to the assumption.

Proposition 6 allows us to simplify the regulator's problem as follows. Given (43), the total expected discounted rents at date 0 amount to:

$$\begin{aligned} & \nu \left\{ \Phi(q_0, q_{2,1}(\underline{\theta})) + \nu\beta [\Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta})) - \Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta}))] \right\} \\ & \quad + \nu^2\beta\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) + \nu(1-\nu)\beta\Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta})) \\ & = \nu\Phi(q_0, q_{2,1}(\underline{\theta})) + \nu\beta\Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta})). \end{aligned} \quad (46)$$

Using the fact that (39) and (42) are binding at the optimum, it is easy to see that (40) is equivalent to the following monotonicity constraint:

$$q_{2,1}(\bar{\theta}) \geq q_{2,1}(\underline{\theta}), \quad (47)$$

while, given (43), (41) can be rewritten as:

$$\Phi(q_0, q_{2,1}(\underline{\theta})) + \nu\beta [\Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta})) - \Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta}))] \geq 0. \quad (48)$$

It thus follows from (46) that, for a given value of q_0 , the pair of functions $(q_{2,1}, q_{2,2})$ solves:

$$\begin{aligned} & \max_{(\tilde{q}_1, \tilde{q}_2)} \{ \mathbb{E} [- (1 + \lambda)\psi(\tilde{q}_1(\theta_0) - \delta q_0 - \theta_0) + \beta S(\tilde{q}_1(\theta_0)) \\ & \quad - (1 + \lambda)\beta\psi(\tilde{q}_2(\theta_0, \theta_1) - \delta\tilde{q}_1(\theta_0) - \theta_1) + \beta^2 S(\tilde{q}_2(\theta_0, \theta_1))] \\ & \quad - \lambda\nu\Phi(q_0, q_{2,1}(\underline{\theta})) - \lambda\nu\beta\Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta})) \}, \end{aligned} \quad (49)$$

subject to the constraint that (47)–(48) hold at the optimum. While constraints (39) and (42) are necessarily binding in the optimal long-term contract, one cannot decide a priori on the status of constraints (47)–(48). Indeed, pooling over date 1 quality may occur, so that (40) or equivalently (47) may be binding. Furthermore, it is unclear that a high type agent receives a strictly positive rent at date 0, and thus that (41) or equivalently (48) are slack. We say that we are in the regular case when the parameter values are such that this latter property holds.

4.2.2. Date 2 Quality Levels

In this subsection, we investigate date 2 quality levels. Let us first assume for simplicity that the regular case obtains, so that constraint (48) can be neglected. It then follows that there is no term of the form $\Phi(\tilde{q}_1(\bar{\theta}), \tilde{q}_2(\bar{\theta}, \underline{\theta}))$ in the program (47)–(49) that determines $q_{2,2}$. This reflects that, while increasing the date 1 rent $\Phi(\tilde{q}_1(\bar{\theta}), \tilde{q}_2(\bar{\theta}, \underline{\theta}))$ by $\varepsilon > 0$ has a direct expected discounted cost $\lambda\nu^2\beta\varepsilon$ for the regulator, it also allows her to reduce the expected cost $\lambda\nu U_{2,0}(\bar{\theta})$ of the date 0 rent by exactly the same amount, see (43). Thus the cost of providing incentives at date 1 given that the agent had a high type at date 0 is already taken

into account in the rent left by the regulator to this agent at date 0, and is therefore perfectly internalized by her. By contrast, increasing the date 1 rent $\Phi(\tilde{q}_1(\underline{\theta}), \tilde{q}_2(\underline{\theta}, \underline{\theta}))$ by $\varepsilon > 0$ has a direct expected discounted cost $\lambda\nu(1-\nu)\beta\varepsilon$ at date 1, to which must be added an indirect expected cost $\lambda\nu^2\beta\varepsilon$ at date 0 due to the increase of the date 0 rent $U_{2,0}(\bar{\theta})$, see again (43). This reflects that raising $\Phi(\tilde{q}_1(\underline{\theta}), \tilde{q}_2(\underline{\theta}, \underline{\theta}))$ makes a deviation more attractive for a high type agent at date 0. It is then easy to characterize the optimal quality levels at date 2. Using the notation of Subsection 4.1, one has:

$$q_{2,2}(\bar{\theta}, \bar{\theta}) = \bar{Q}_{1,2}(q_{2,1}(\bar{\theta})), \quad (50)$$

$$q_{2,2}(\underline{\theta}, \bar{\theta}) = \bar{Q}_{1,2}(q_{2,1}(\underline{\theta})), \quad (51)$$

$$q_{2,2}(\bar{\theta}, \underline{\theta}) = \underline{Q}_{1,2}(q_{2,1}(\bar{\theta})), \quad (52)$$

so that there are no distortions of date 2 quality given the date 1 quality levels if $\theta_0 = \bar{\theta}$ or $\theta_1 = \bar{\theta}$, while the date 2 quality if $(\theta_0, \theta_1) = (\underline{\theta}, \underline{\theta})$ is distorted downward according to:

$$\beta S'(q_{2,2}(\underline{\theta}, \underline{\theta})) = (1 + \lambda)\psi'(q_{2,2}(\underline{\theta}, \underline{\theta}) - \delta q_{2,1}(\underline{\theta}) - \underline{\theta}) + \frac{\lambda\nu}{(1-\nu)^2} \Phi_2(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta})). \quad (53)$$

Note from (47)–(49) that (50)–(51) hold even if we are not in the regular case. Concerning $q_{2,2}(\bar{\theta}, \underline{\theta})$ and $q_{2,2}(\underline{\theta}, \underline{\theta})$, the following general result holds.

Proposition 7. *In the optimal long-term contract, the level of distortions at date 1 is strictly higher following $\theta_0 = \underline{\theta}$ than following $\theta_0 = \bar{\theta}$. In particular, $q_{2,2}(\bar{\theta}, \underline{\theta}) > q_{2,2}(\underline{\theta}, \underline{\theta})$.*

In analogy with the standard repeated moral hazard problem (Rogerson (1985)), the optimal long-term contract therefore exhibits memory, in the sense that the level of distortions at date 1 depends on the type of the agent at date 0. This is because we impose that the agent must receive a non-negative utility at each date and in each state. Thus, although the date 1 contract is signed at date 0 under symmetric information about the date 1 quality shock, the regulator cannot fully extract the agent's expected date 1 rent at date 0, unlike for instance in the Baron and Besanko (1984) regulation model with independent types.

4.2.3. Date 1 Quality Levels and Pooling

Having determined the optimal date 2 quality levels, one can move backward to date 1. Without specifying further the model, however, it is not easy to characterize the date 1 properties of the optimal contract, and in particular to assess under which circumstances (47) and (48) are binding. In this section, we focus on the monotonicity constraint (47), and we determine under which circumstances pooling over date 1 quality may occur.

The reason for why this may arise at the optimum is that increasing date 1 quality following $\theta_0 = \underline{\theta}$ does not only affect consumer surplus, but also the informational rents that must be left to the agent. As shown by (46), these rents consist of two terms. The first term is proportional to $\Phi(q_0, q_{2,1}(\underline{\theta}))$, which is increasing in $q_{2,1}(\underline{\theta})$. This reflects a standard static rent extraction motive. By contrast, the second term is proportional to $\Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta}))$, which is decreasing in $q_{2,1}(\underline{\theta})$. As in the multiple agent model of Section 3, this reflects a

dynamic rent extraction motive. The logic is slightly different, however. In the multiple agent model, the regulator typically wanted to increase the quality produced by a *high* type agent at date 0 only in order to decrease the rents that he had to concede at date 1, regardless of date 0 incentives. By contrast, in the single agent model, the regulator may want to increase the quality produced by a *low* type agent at date 0 in order to reduce the date 1 rents following a report $\hat{\theta}_0 = \underline{\theta}$, and thus make a deviation less attractive for a high type agent at date 0. The fact that, as discussed in Subsection 4.2.2, date 1 rents are treated asymmetrically by the regulator, depending on the type of the agent at date 0, explains why this dynamic rent extraction effect now affects the quality produced by a low type agent at date 0, and not directly that produced by a high type agent.

Overall, pooling over date 1 quality takes place if the dynamic rent extraction motive prevails, a situation that can intuitively arise only if the discount factor β and the depreciation factor δ are high enough. This is confirmed by our next result.

Proposition 8. *If q_0 is close enough to zero, pooling over date 1 quality can occur in the optimal long-term contract only if $\beta\delta > 1$.*

In order to show that pooling over date 1 quality can indeed occur in equilibrium in the circumstances delineated in Proposition 7, we turn to the linear-quadratic specification (24)–(25). As in the basic model, special care must be taken in specifying parameters for the model such that the solution is economically meaningful. In particular, it turns out that for date 1 quality levels to remain smaller than $\frac{a}{b}$, the discount factor β cannot be too large for a given value of δ . Specifically, it can be shown that if q_0 , $\underline{\theta}$ and $\Delta\theta$ are close enough to zero, it must be the case that $\beta < \bar{\beta}(\delta)$, where $\bar{\beta}(\delta)$ is the positive solution to:

$$\frac{\beta^2\delta(1-\delta)b}{\beta b + c(1+\lambda)} = 1. \quad (54)$$

Note that $\bar{\beta}(\delta)$ goes to infinity when δ goes to one, and thus one can choose $\beta\delta$ to be greater than one as required by Proposition 8. One then has the following result.¹¹

Proposition 9. *For any δ close enough to one, there exists a $\underline{\beta}(\delta) \in (0, \bar{\beta}(\delta))$ such that, for any $\beta \in [\underline{\beta}(\delta), \bar{\beta}(\delta))$, and for any q_0 , $\underline{\theta}$ and $\Delta\theta$ close enough to zero, pooling over date 1 quality occurs in the optimal long-term contract and (48) is slack.*

Pooling over date 1 quality distorts the allocation of resources relative to the symmetric information benchmark. Consider the case of a high quality shock at date 0, and suppose that we are in the regular case, as in Proposition 9. We know from (50) and (52) that, in this case, there will be no distortions of date 2 quality levels given the date 1 quality level $q_{2,1}(\bar{\theta})$. In turn, $q_{2,1}(\bar{\theta})$ satisfies the first-order condition:

$$\beta S'(q_{2,1}(\bar{\theta})) + \beta\delta(1+\lambda)\mathbb{E}[\psi'(e_{2,1}(\bar{\theta}, \theta_1))] = (1+\lambda)\psi'(e_{2,0}(\bar{\theta})) - \frac{\xi_1}{\nu}, \quad (55)$$

where ξ_1 is the multiplier associated to (47). Using the notation of Subsection 4.1, it follows from (38) and (55) that $q_{2,1}(\bar{\theta}) \geq q_{1,1}(\bar{\theta})$, with a strict inequality if $\xi_1 > 0$. Thus pooling

¹¹This result hinges on the assumption that δ is close to one. One can show that, in the linear-quadratic specification of the model, pooling actually requires such a high degree of quality durability. Specifically, if δ is close enough to zero, then no pooling over date 1 qualities can occur in the optimal contract, even if β is equal to its maximum admissible level $\bar{\beta}(\delta)$. We refer to the Appendix for details.

over date 1 quality leads to over-provision of date 1 quality when $\theta_0 = \bar{\theta}$, and a fortiori when $\theta_0 = \underline{\theta}$, relative to the symmetric information benchmark.

Finally, it should be emphasized that pooling over date 1 quality does not prevent the regulator from screening the agent's type at date 0. Rather, screening is achieved through date 0 transfers $u_{2,0}(\hat{\theta}_0)$ and date 1 continuation rents $\Phi(q_{2,1}(\hat{\theta}_0), q_{2,2}(\hat{\theta}_0, \underline{\theta}))$. Specifically, let $q_{2,1}^p = q_{2,1}(\bar{\theta}) = q_{2,1}(\underline{\theta})$. Then, one has:

$$\begin{aligned} u_{2,0}(\bar{\theta}) &= \psi(q_{2,1}^p - \delta q_0 - \bar{\theta}) + \Phi(q_0, q_{2,1}^p) + \nu\beta \left[\Phi(q_{2,1}^p, q_{2,2}(\underline{\theta}, \underline{\theta})) - \Phi(q_{2,1}^p, q_{2,2}(\bar{\theta}, \underline{\theta})) \right] \\ &= u_{2,0}(\underline{\theta}) + \nu\beta \left[\Phi(q_{2,1}^p, q_{2,2}(\underline{\theta}, \underline{\theta})) - \Phi(q_{2,1}^p, q_{2,2}(\bar{\theta}, \underline{\theta})) \right] \\ &< u_{2,0}(\underline{\theta}), \end{aligned}$$

where we have used the fact that (39) and (42) are binding, and that $q_{2,2}(\bar{\theta}, \underline{\theta}) > q_{2,2}(\underline{\theta}, \underline{\theta})$ by Proposition 7. At date 0, a high type agent therefore receives a lower transfer than a low type agent, $u_{2,0}(\bar{\theta}) < u_{2,0}(\underline{\theta})$, in exchange for a higher date 1 continuation rent, $\Phi(q_{2,1}^p, q_{2,2}(\bar{\theta}, \underline{\theta})) > \Phi(q_{2,1}^p, q_{2,2}(\underline{\theta}, \underline{\theta}))$. Note that $q_{2,2}(\bar{\theta}, \underline{\theta}) > q_{2,2}(\underline{\theta}, \underline{\theta})$ although $q_{2,1}(\bar{\theta}) = q_{2,1}(\underline{\theta})$. Thus, unlike in the multiple agent model, date 1 quality does not constitute a sufficient statistic for the continuation of the contractual relationship: while a high type and a low type agent having delivered the same level of date 1 quality have the same preferences at date 1, they are treated differently in the continuation contract. The reason of this apparent paradox is that the regulator does not treat the date 1 rents of a former high type agent in the same way as those of a former low type agent. This endogenously leads to a sorting condition that makes it optimal for her to screen the agent at date 0, in spite of pooling over date 1 quality.

4.3. Single versus Multiple Agents

It is instructive to compare regulation with a single agent and with a sequence of agents. The key difference between the two setups is that, when the regulator contracts with a single agent, she is no longer constrained to use sequentially rational date 2 Markovian quality policies contingent on date 1 quality and the agent's date 1 type. To illustrate this point, suppose for simplicity that the regular case obtains. Then, as shown by (52), the regulator commits to an efficient quality level when $(\theta_0, \theta_1) = (\bar{\theta}, \underline{\theta})$, whereas she would find it optimal to distort downward quality were she facing multiple agents in the same circumstances.¹² Conversely, as shown by (53), she commits to a higher level of distortions when $(\theta_0, \theta_1) = (\underline{\theta}, \underline{\theta})$, reflecting the cost of providing dynamic incentives.

Although general comparison results are hard to obtain, one can unambiguously rank date 1 quality levels in the two regulatory environments. Specifically, let $\bar{q}_{2,1}(q_0)$ and $\underline{q}_{2,1}(q_0)$ be the quality levels at date 1 following respectively a high and low quality shock at date 0 when there are multiple agents. Then the following holds.

Proposition 10. *If the regular case obtains and no pooling over date 1 qualities occurs in optimal long-term contract with a single agent, the spread of public good qualities at date 1 is wider with a sequence of agents than with a single agent,*

$$\bar{q}_{2,1}(q_0) > q_{2,1}(\bar{\theta}) > q_{2,1}(\underline{\theta}) > \underline{q}_{2,1}(q_0). \quad (56)$$

¹²The optimal distortion level is given by equation (84) in the Appendix, see footnote 14.

This result reflects the difference in date 1 distortions that occur in the two regulatory environments. With a sequence of agents, as in the basic model of Section 3, there is over-provision of date 1 quality following a high quality shock at date 0. This is because the date 1 informational rent that has to be left when there is a further high quality shock at date 1 is not internalized by the regulator, as opposed to when there is a single agent. Increasing date 1 quality allows the regulator to reduce this rent, which leads to the first half of (56). By contrast, the date 1 distortions following a low quality shock at date 0 are larger with a single agent than with a sequence of agents, reflecting the cost of making a deviation for a high type agent less attractive at date 0. In turn, increasing date 1 quality allows the regulator to alleviate this cost, which leads to the second half of (56).

A natural question is whether it is better from the regulator's viewpoint to hire a single agent or to contract with a sequence of agents.¹³ In our model, each option comes with costs and benefits. On the one hand, by hiring a single agent, the regulator can directly condition his date 1 compensation on his date 0 performance, while this occurs only indirectly through date 1 quality when there is a sequence of agents. This in turn allows the regulator to internalize some of the costs of providing incentives at date 1, as explained in Subsection 4.2.2. On the other hand, the fact that the same agent is rewarded over two periods raises the cost of providing incentives at date 0, because the agent anticipates the impact of his date 0 actions on his date 1 utility. Which of these effects prevails is a priori unclear. However, for small but strictly positive levels of asymmetric information, it is better to hire a single agent, as the following result indicates.

Proposition 11. *Fix all parameters of the model except $\bar{\theta}$. When $\bar{\theta}$ is close enough to $\underline{\theta}$, the ex-ante social welfare is strictly higher in the optimal long-term contract with a single agent than in the optimal contract with a sequence of agents.*

The intuition for this result is that, when $\bar{\theta}$ is close to $\underline{\theta}$, one can always replicate the optimal contract with a sequence of agents in the environment with full commitment and a single agent. To check that, one needs only to verify that the analogues of constraints (39)–(42) are satisfied when one substitutes in the solution for the optimal contract with a sequence of agents. The only constraint that turns out to matter is (39). Letting $\underline{q}_{2,2}(q_1)$ be the optimal quality level at date 2 given a date 1 quality level q_1 and a low quality shock at date 1, this can be rewritten as:

$$\Phi(q_0, \underline{q}_{2,1}(q_0)) \geq \nu\beta \left[\Phi(\underline{q}_{2,1}(q_0), \underline{q}_{2,2}(\underline{q}_{2,1}(q_0))) - \Phi(\bar{q}_{2,1}(q_0), \underline{q}_{2,2}(\bar{q}_{2,1}(q_0))) \right]. \quad (57)$$

When $\bar{\theta}$ is close to $\underline{\theta}$ and thus $\Delta\theta$ is close to zero, the left-hand side of (57) is of order $\Delta\theta$, while the right-hand side of (57) is of order $\Delta\theta[\bar{q}_{2,1}(q_0) - \underline{q}_{2,1}(q_0)]$. When $\bar{\theta}$ is close to $\underline{\theta}$, this term is negligible relative to $\Delta\theta$ as both optimal date 1 policy functions converge pointwise

¹³In our context, the answer to this question is not a priori obvious because quality physically links periods together, and agents must receive a non-negative utility at each date and in each state. By contrast, in Baron and Besanko (1984), long-term contracts can be signed that do not need to satisfy this condition, and there is no physical intertemporal link. When types are independent across periods, sequential contracting with multiple agents is then suboptimal because it does not allow the regulator to extract the expected date 1 rent at date 0. This result still holds if, in addition, one modifies their model by imposing our restrictions on utilities: in the absence of a physical intertemporal link, the optimal contract with a sequence of agents is then simply the repetition of the optimal static contract, which can obviously be duplicated in a long-term contract with a single agent.

to a function $q_{2,1}$ corresponding to the limit case where $\bar{\theta} = \underline{\theta}$. It follows that (57) is satisfied for $\bar{\theta}$ close enough to $\underline{\theta}$, which implies the result.

4.4. No Commitment and the Ratchet Effect

Let us briefly consider the case where there is a single agent, but neither the regulator nor the agent can commit to a long-term contract. At date 1, the regulator offers the same contract than with a sequence of agents. At date 0, and for an initial level of quality q_0 , a contract is simply a 4-tuple $(\bar{u}_0, \bar{q}_1, \underline{u}_0, \underline{q}_1)$. Incentive compatibility at date 0 requires that:

$$\begin{aligned} \bar{u}_0 - \psi(\bar{q}_1 - \delta q_0 - \bar{\theta}) + \nu\beta\Phi(\bar{q}_1, \underline{q}_{2,2}(\bar{q}_1)) \\ \geq \underline{u}_0 - \psi(\underline{q}_1 - \delta q_0 - \bar{\theta}) + \nu\beta\Phi(\underline{q}_1, \underline{q}_{2,2}(\underline{q}_1)), \end{aligned} \quad (58)$$

$$\begin{aligned} \underline{u}_0 - \psi(\underline{q}_1 - \delta q_0 - \underline{\theta}) + \nu\beta\Phi(\underline{q}_1, \underline{q}_{2,2}(\underline{q}_1)) \\ \geq \bar{u}_0 - \psi(\bar{q}_1 - \delta q_0 - \underline{\theta}) + \nu\beta\Phi(\bar{q}_1, \underline{q}_{2,2}(\bar{q}_1)). \end{aligned} \quad (59)$$

Because ψ is convex, an immediate implication of (58)–(59) is that $\bar{q}_1 \geq \underline{q}_1$. In this context, if a low type agent misrepresents his type at date 0, he will still contract with the regulator at date 1 and earn a positive rent on average: a “take-the-money-and-run” strategy is not profitable, reflecting the fact that types are independent across periods. This represents a key difference between this setup and a standard dynamic adverse selection model with perfectly correlated types (Laffont and Tirole (1988)).

Since the agent must receive a non-negative utility in each state, an incentive feasible contract must also satisfy the following date 0 participation constraints:

$$\bar{u}_0 - \psi(\bar{q}_1 - \delta q_0 - \bar{\theta}) \geq 0, \quad (60)$$

$$\underline{u}_0 - \psi(\underline{q}_1 - \delta q_0 - \underline{\theta}) \geq 0. \quad (61)$$

Given that, in analogy with the basic model, the date 1 rent is a decreasing function of date 1 quality, it is easy to check that (58) and (61) imply (60) provided that $\bar{q}_1 \geq \underline{q}_1$. Adding this as a constraint to the regulator’s problem, one can then solve for the optimal date 0 contract in a completely standard way. The incentive compatibility constraint (58) of the high type agent is binding, as well as the participation constraint (61) of the low type agent. Straightforward manipulations imply that the incentive compatibility constraint (59) of the low type agent can be rewritten as $\Phi(q_0, \bar{q}_1) \geq \Phi(q_0, \underline{q}_1)$, which is satisfied when $\bar{q}_1 \geq \underline{q}_1$.

Although the complications that usually arise in models of dynamic adverse selection are absent in this model, the lack of commitment still generates a ratchet effect. Indeed, since the regulator cannot commit to a compensation scheme at date 1, a high type agent anticipates at date 0 that, by revealing his type, he will reduce his date 1 rent, since:

$$\Phi(\bar{q}_1, \underline{q}_{2,2}(\bar{q}_1)) \leq \Phi(\underline{q}_1, \underline{q}_{2,2}(\underline{q}_1)).$$

That is, the agent anticipates that being efficient today will increase the level of quality tomorrow and therefore jeopardize his continuation rent. The corresponding date 0 rent of a

high type agent is therefore given by:

$$\Phi(q_0, \underline{q}_1) + \nu\beta \left[\Phi(\underline{q}_1, \underline{q}_{2,2}(\underline{q}_1)) - \Phi(\bar{q}_1, \underline{q}_{2,2}(\bar{q}_1)) \right].$$

To reduce this rent, it is typically optimal for the regulator to induce a pooling outcome at date 0 when β is large.

It is straightforward to compare social welfare with a single agent and no commitment and with a sequence of agents. Indeed, the optimal date 1 policies are the same in both problems, while the incentive compatibility constraint at date 0 is more stringent with a single agent, due to the ratchet effect. Therefore, when no commitment is feasible, it is better from a social viewpoint to contract with a sequence of agents rather than to hire a single agent. This contrasts with the findings of Lewis and Yildirim (2002, Proposition 4), who show in a learning-by-doing model that when supply costs decrease with past production, the regulator will prefer dealing with a single supplier rather than relying on less durable franchises. The reason is that, in their model, the regulated firm's informational rents are higher when supply costs falls and the regulator demands more service, while these rents are lower in our model when public good quality increases.

5. CONCLUDING REMARKS

In this paper, we explored the design of incentives for public good quality provision in a dynamic regulation framework in which maintenance efforts and quality shocks have long-lasting effects. We considered in turn two regulatory frameworks, one with a sequence of short-term agents, and one with a single long-term agent. When the regulator contracts with a sequence of agents, asymmetries of information can result in over-provision of quality under optimal regulation, reflecting a dynamic rent extraction motive. This contrasts with the standard prediction of a static model of quality provision. In the linear-quadratic example, asymmetries of information lead to higher (lower) maintenance efforts in the case of a high (low) quality shock, which translates into more volatile quality growth than under symmetric information. When the regulator hires a single agent to manage public good quality, over-provision of quality can also occur, but through a different mechanism: if quality depreciates slowly and the discount factor is large, optimal regulation can lead to quality pooling, which results in over-provision of quality irrespective of the realization of the quality shock. For small levels of asymmetric information, the regulator prefers to hire a single agent rather than to contract with a sequence of agents, provided they can commit to a long-term contract. By contrast, when no such commitment is feasible, the durability of quality leads to a ratchet effect even though private information is recurring, and shorter franchises are beneficial from a social point of view.

This paper abstracts from several important features of quality regulation, which should be investigated in future work. We have considered a project of fixed size, implicitly focusing on quality as the sole dimension of differentiation. It would be interesting to extend the analysis to the case where consumers care both about the quantity and the quality of output, in order to explore the trade-off between the level and the quality of service, as well as the interaction between the regulated firm's prices and changes in quality. Another meaningful extension of the analysis would be to relax the assumption that quality is verifiable by allowing only imperfect signals of quality to be ascertained in court.

Finally, the techniques and insights developed in this paper might be also applied to

the study of other dynamic agency relationships in which the cost of providing incentives is affected by an endogenous state variable as in our model. For instance, the owner of a durable good such as a flat or a building might be concerned with the maintenance efforts exerted by the successive renters, while the latter may differ with respect to the kind of use they will make of the property—they may be for instance careful or careless. The state of the property would then vary over time as a function of the successive renters' types and maintenance efforts, while the value of the property would be endogenously determined by the whole sequence of rental contracts offered by the owner. As in our model, a succession of careful renters may increase the state of the property over and above the first-best level. In this context, it might be interesting to endogenize the sequence of renters' types by studying the incentives of different types of renters to select properties of different states, and the impact of this self-selection mechanism on the value of property over time. These important questions are left for future research.

APPENDIX

Proof of Proposition 1. Denote by T_1 and T_2 the Bellman operators associated to (5) and (20) respectively. The following lemma holds.

Lemma A.1. *Let f_1 and f_2 be two real-valued functions over \mathbb{R}_+ that are both bounded, strictly increasing, strictly concave and continuously differentiable over \mathbb{R}_{++} . Then, if $f_2' \geq f_1'$ over \mathbb{R}_{++} , $(T_2 f_2)' > (T_1 f_1)'$ over \mathbb{R}_{++} .*

Proof. Standard considerations (see for instance Stokey and Lucas (1989, §4)) imply that $T_1 f_1$ and $T_2 f_2$ are bounded, strictly increasing, strictly concave and continuously differentiable over \mathbb{R}_{++} . The condition $\psi'(0) = 0$ ensures that the optimal effort levels remain strictly positive. Fix $q > 0$. The first-order conditions corresponding to the program that defines $(T_1 f_1)(q)$ are:

$$\beta f_1'(\overline{Q}_1(q)) = (1 + \lambda)\psi'(\overline{Q}_1(q) - \delta q - \overline{\theta}), \quad (62)$$

$$\beta f_1'(\underline{Q}_1(q)) = (1 + \lambda)\psi'(\underline{Q}_1(q) - \delta q - \underline{\theta}). \quad (63)$$

By the Envelope Theorem,

$$(T_1 f_1)'(q) = S'(q) + \nu\delta(1 + \lambda)\psi'(\overline{Q}_1(q) - \delta q - \overline{\theta}) + (1 - \nu)\delta(1 + \lambda)\psi'(\underline{Q}_1(q) - \delta q - \underline{\theta}).$$

Using (62)–(63), this can be rewritten as:

$$(T_1 f_1)'(q) = S'(q) + \nu\beta\delta f_1'(\overline{Q}_1(q)) + (1 - \nu)\beta\delta f_1'(\underline{Q}_1(q)). \quad (64)$$

Similarly, the first-order conditions corresponding to the program that defines $(T_2 f_2)(q)$ are:

$$\beta f_2'(\overline{Q}_2(q)) = (1 + \lambda)\psi'(\overline{Q}_2(q) - \delta q - \overline{\theta}), \quad (65)$$

$$\beta f_2'(\underline{Q}_2(q)) = (1 + \lambda)\psi'(\underline{Q}_2(q) - \delta q - \underline{\theta}) + \frac{\lambda\nu}{1 - \nu} \Phi_2(q, \underline{Q}_2(q)). \quad (66)$$

By the Envelope Theorem,

$$(T_2 f_2)'(q) = S'(q) + \nu\delta(1 + \lambda)\psi'(\overline{Q}_2(q) - \delta q - \overline{\theta}) + (1 - \nu)\delta(1 + \lambda)\psi'(\underline{Q}_2(q) - \delta q - \underline{\theta}) - \nu\lambda\Phi_1(q, \underline{Q}_2(q)).$$

Using (65)–(66) together with $\Phi_1 = -\delta\Phi_2$, this can be rewritten as:

$$(T_2 f_2)'(q) = S'(q) + \nu\beta\delta f_2'(\overline{Q}_2(q)) + (1 - \nu)\beta\delta f_2'(\underline{Q}_2(q)). \quad (67)$$

Using the fact that f_1 and f_2 are strictly concave, that ψ is strictly convex over \mathbb{R}_+ and that $f_2' \geq f_1'$, one can check from (62) and (65) that $f_2'(\overline{Q}_2(q)) \geq f_1'(\overline{Q}_1(q))$. Since $\Phi_2(q, \underline{Q}_2(q)) > 0$, it follows in a similar way from (63) and (66) that $f_2'(\underline{Q}_2(q)) > f_1'(\underline{Q}_1(q))$. Therefore, by (64) and (67), one obtains that $(T_2 f_2)'(q) > (T_1 f_1)'(q)$, which implies the result since q is arbitrarily chosen. \square

We are now ready to complete the proof of Proposition 1. The fact that T_1 and T_2 are contractions with unique fixed points V_1 and V_2 over the space of bounded continuous functions over \mathbb{R}_+ ensures that $V_1 = \lim_{n \rightarrow \infty} T_1^n S$ and $V_2 = \lim_{n \rightarrow \infty} T_2^n S$ pointwise over \mathbb{R}_+ . Since V_1 and V_2 are finite, concave and differentiable over \mathbb{R}_{++} , this implies that $V_1' = \lim_{n \rightarrow \infty} (T_1^n S)'$ and $V_2' = \lim_{n \rightarrow \infty} (T_2^n S)'$ pointwise over \mathbb{R}_{++} (Rockafellar (1970, Theorem 25.7)). By Lemma A.1, $(T_2^n S)' > (T_1^n S)'$ over \mathbb{R}_{++} for each

$n \in \mathbb{N} \setminus \{0\}$. Taking limits, it follows that $V_2' \geq V_1'$ over \mathbb{R}_{++} . Applying Lemma A.1 again and using the fact that $T_1V_1 = V_1$ and $T_2V_2 = V_2$, one obtains that $V_2' > V_1'$ over \mathbb{R}_{++} , as claimed. \blacksquare

Proof of Proposition 2. For each $i = 1, 2$, the effort mappings \bar{e}_i and \underline{e}_i are strictly decreasing, or, equivalently, the mappings $q \mapsto \bar{q}_i(q) - \delta q$ and $q \mapsto \underline{q}_i(q) - \delta q$ are strictly decreasing. Since $\delta \in (0, 1)$ and both $\bar{q}_i(0)$ and $\underline{q}_i(0)$ are strictly positive, and since the mappings \bar{q}_i and \underline{q}_i are strictly increasing with $\bar{q}_i > \underline{q}_i$, this implies that there are exactly two points $q_i^+ > q_i^- > 0$ such that $\bar{q}_i(q_i^+) = q_i^+$ and $\underline{q}_i(q_i^-) = q_i^-$. Note that $q_2^+ > q_1^+$ since $\bar{q}_2 > \bar{q}_1$ over \mathbb{R}_{++} . It is easy to verify that both $[0, q_i^-)$ and $(q_i^+, \infty]$ are transient sets for the transition function P_i , so that one can restrict the analysis to $[q_i^-, q_i^+]$. Because the mappings \bar{q}_i and \underline{q}_i are strictly increasing and continuous, P_i is monotone and satisfies the Feller property. Moreover, since $\lim_{n \rightarrow \infty} \bar{q}_i^n(q_i^-) = q_i^+$ and $\lim_{n \rightarrow \infty} \underline{q}_i^n(q_i^+) = q_i^-$, it follows that for any $q_i \in (q_i^-, q_i^+)$, there exists an integer $n \geq 1$ such that $P_i^n(q_i^-, [q_i, q_i^+]) \geq \nu^n$ and $P_i^n(q_i^+, [q_i^-, q_i]) \geq (1 - \nu)^n$. Thus P_i satisfies the mixing Assumption 12.1 in Stokey and Lucas (1989) over $[q_i^-, q_i^+]$. As a result, there exists a unique probability measure μ_i over this interval that is invariant under P_i in the sense that $\mu_i(A) = \int P_i(q, A) \mu_i(dq)$ for any Borel subset A (Stokey and Lucas (1989, Theorem 12.12)). We now prove that $\max \text{supp } \mu_i = q_i^+$, which concludes the proof as $q_2^+ > q_1^+$. Suppose instead that $q_i^+ > \max \text{supp } \mu_i$. Then, since $\bar{q}_i(q) > q$ for any $q < q_i^+$, $\bar{q}_i(q) > \max \text{supp } \mu_i$ for any q close enough to $\max \text{supp } \mu_i$. By definition of $\text{supp } \mu_i$, it follows that $\mu_i(\bar{q}_i^{-1}(\max \text{supp } \mu_i, q_i^+)) > 0$. However, as μ_i is invariant under P_i , one then obtains that:

$$\mu_i(\max \text{supp } \mu_i, q_i^+) = \int P_i(q, (\max \text{supp } \mu_i, q_i^+)) \mu_i(dq) \geq \nu \mu_i(\bar{q}_i^{-1}(\max \text{supp } \mu_i, q_i^+)) > 0,$$

a contradiction. Therefore $\max \text{supp } \mu_i = q_i^+$, as claimed, and similarly $\min \text{supp } \mu_i = q_i^-$. Finally, note that Theorem 12.12 in Stokey and Lucas (1989) also ensures that for any initial quality level, the distribution of quality converges weakly to μ_i . This concludes the proof. \blacksquare

Proof of Proposition 3. We first solve (5) and (20) without restrictions on quality levels, and verify that (32) holds. Using (27)–(28) and (29)–(30) together with (26), and identifying terms in q^2 in (5) and (20), we first obtain that $B_1 = B_2 = B$, where B is the positive solution to:

$$B = b + \frac{\beta \delta^2 c(1 + \lambda)B}{\beta B + c(1 + \lambda)}. \quad (68)$$

Next, identifying terms in q in (5) and (20), we obtain that:

$$A_1 = \frac{a[\beta B + c(1 + \lambda)] - \beta \delta c B(1 + \lambda)E_\theta}{\beta B + c(1 + \lambda)(1 - \beta \delta)} \quad (69)$$

and:

$$A_2 = \frac{a[\beta B + c(1 + \lambda)] - \beta \delta c B(1 + \lambda)E_\theta + \nu \beta \delta c B \lambda \Delta \theta}{\beta B + c(1 + \lambda)(1 - \beta \delta)}. \quad (70)$$

Hence $A_2 > A_1$, and since $B_1 = B_2 = B$, it follows from (27)–(29) that $\bar{q}_2(q) > \bar{q}_1(q) > \underline{q}_1(q)$, as expected. Next, from (28) and (30), $\underline{q}_1(q) > \underline{q}_2(q)$ if and only if $\beta A_1 > \beta A_2 - \frac{\lambda \nu}{1 - \nu} c \Delta \theta$. Using (69)–(70), this condition can be rewritten as:

$$\frac{\beta B + c(1 + \lambda)(1 - \beta \delta)}{1 - \nu} > \beta^2 \delta B,$$

which clearly holds as $(\beta, \delta, \nu) \in (0, 1)^3$ and $B > 0$. Hence $\underline{q}_1(q) > \underline{q}_2(q)$, and (32) follows. We now provide parameter and range restrictions under which the social value functions V_1 and V_2 thus

obtained are economically meaningful. Define q_1^+ and q_2^+ as in the proof of Proposition 2. By construction, the interval $[0, q_2^+]$ is invariant under both the transition functions P_1 and P_2 . We shall therefore assume that q_0 initially belongs to $[0, q_2^+]$ and restrict our attention to this interval. A necessary condition for our solutions to (5) and (20) to be economically meaningful over $[0, q_2^+]$ is that $q_2^+ < \frac{a}{b}$. Using (29) to solve for q_2^+ , and taking advantage of (68) and (70), one can check after tedious but straightforward algebraic manipulations that a sufficient condition for this to occur is:

$$\frac{\bar{\theta} - \beta\delta E_\theta}{(1-\delta)(1-\beta\delta)} + \frac{\nu\beta\delta\lambda\Delta\theta}{(1+\lambda)(1-\delta)(1-\beta\delta)} < \frac{a}{b}. \quad (71)$$

We shall hereafter assume that (71) holds. This condition immediately implies that:

$$(1-\delta)\frac{a}{b} > \bar{\theta}. \quad (72)$$

Using (72) together with (68)–(70), one can in turn check that:

$$\frac{A_2}{B} > \frac{A_1}{B} > \frac{a}{b}. \quad (73)$$

The inequalities (73) capture the intuitive fact that, as long as the surplus function S is increasing, so are the social value functions V_1 and V_2 . In particular, the latter are increasing over $[0, q_2^+]$. We now investigate under which conditions efforts remain strictly positive over this interval. It is easy to check from (27) and (29) that for each $i = 1, 2$ and $q \in [0, q_2^+]$,

$$\bar{e}_i(q) = \frac{\beta B}{\beta B + c(1+\lambda)} \left(\frac{A_i}{B} - \delta q - \bar{\theta} \right) > \frac{\beta B}{\beta B + c(1+\lambda)} \left[(1-\delta)\frac{a}{b} - \bar{\theta} \right] > 0,$$

where the first inequality follows from (73) and the fact that $q_2^+ < \frac{a}{b}$, and the second from (72). Hence efforts conditional on a high quality shock are strictly positive, and since $\underline{e}_1(q) > \bar{e}_1(q)$, so are efforts conditional on a low quality shock under symmetric information. To conclude, one must check that $\underline{e}_2(q) \geq \Delta\theta$ for all $q \in [0, q_2^+]$, so that we are justified in using the quadratic specification of ψ when writing down the informational rent of the high type agent, see (31). From (30), one has:

$$\underline{e}_2(q) = \frac{\beta B}{\beta B + c(1+\lambda)} \left(\frac{A_2}{B} - \delta q - \underline{\theta} \right) - \frac{\lambda\nu c\Delta\theta}{(1-\nu)[\beta B + c(1+\lambda)]}.$$

Proceeding as above, we obtain that for each $q \in [0, q_2^+]$,

$$\frac{\beta B}{\beta B + c(1+\lambda)} \left(\frac{A_2}{B} - \delta q - \underline{\theta} \right) > \frac{\beta B}{\beta B + c(1+\lambda)} \left[(1-\delta)\frac{a}{b} - \underline{\theta} \right].$$

Hence a sufficient condition for $\underline{e}_2(q) \geq \Delta\theta$ to hold for all $q \in [0, q_2^+]$ is that:

$$\left[(1-\delta)\frac{a}{b} - \underline{\theta} \right] \beta B > \left[\frac{\lambda\nu c}{1-\nu} + \beta B + c(1+\lambda) \right] \Delta\theta. \quad (74)$$

Note from (68) that B is independent of $\bar{\theta}$, $\underline{\theta}$ and $\Delta\theta$. Hence, given (72), (74) typically holds if $\Delta\theta$ is close enough to zero. It is straightforward to find parameter values such that (71) and (74) simultaneously hold. \blacksquare

Proof of Proposition 4. Fix some $t = 0, 1, \dots$. From (27)–(30), $E[q_{1,t+1} | q_{1,t} = q] > E[q_{2,t+1} | q_{2,t} = q]$ if and only if $\beta A_1 > \beta A_2 - \lambda\nu c\Delta\theta$. Using (69)–(70), this condition can be rewritten as:

$$\beta B + c(1+\lambda)(1-\beta\delta) > \beta^2\delta B,$$

which clearly holds as $(\beta, \delta, \nu) \in (0, 1)^3$ and $B > 0$. Therefore (33) follows. As for (34), it is an immediate consequence of (32). For further reference, note that:

$$\text{Var} [q_{1,t+1} | q_{1,t} = q] = \left[\frac{c(1+\lambda)}{\beta B + c(1+\lambda)} \right]^2 \text{Var}_\theta, \quad (75)$$

$$\text{Var} [q_{2,t+1} | q_{2,t} = q] = \left[\frac{c(1+\lambda)}{\beta B + c(1+\lambda)} \right]^2 \text{Var}_\theta + \frac{\lambda \nu^2 c^2 \Delta \theta^2 [\nu \lambda + 2(1-\nu)(1+\lambda)]}{(1-\nu)[\beta B + c(1+\lambda)]^2}, \quad (76)$$

as is easy to check from (27)–(30). ■

Proof of Proposition 5. From (27)–(28) and (29)–(30), it follows that for each $t = 0, 1, \dots$,

$$\begin{aligned} \mathbb{E} [q_{1,t+1}] &= \frac{\beta A_1 + c(1+\lambda)(\mathbb{E}_\theta + \delta \mathbb{E} [q_{1,t}])}{\beta B + c(1+\lambda)}, \\ \mathbb{E} [q_{2,t+1}] &= \frac{\beta A_2 - \lambda \nu c \Delta \theta + c(1+\lambda)(\mathbb{E}_\theta + \delta \mathbb{E} [q_{2,t}])}{\beta B + c(1+\lambda)}. \end{aligned}$$

Taking limits as t goes to infinity, one obtains that:

$$\begin{aligned} \mathbb{E}_{\mu_1} &= \frac{\beta A_1 + c(1+\lambda)\mathbb{E}_\theta}{\beta B + c(1+\lambda)(1-\delta)}, \\ \mathbb{E}_{\mu_2} &= \frac{\beta A_2 - \lambda \nu c \Delta \theta + c(1+\lambda)\mathbb{E}_\theta}{\beta B + c(1+\lambda)(1-\delta)}, \end{aligned}$$

from which (35) follows as $\beta A_1 > \beta A_2 - \lambda \nu c \Delta \theta$. Next, for each $i = 1, 2$ and $t = 0, 1, \dots$,

$$\begin{aligned} \mathbb{E} [\text{Var} [q_{i,t+1} | q_{i,t}]] &= \mathbb{E} [q_{i,t+1}^2] - \mathbb{E} [\mathbb{E} [q_{i,t+1} | q_{i,t}]^2] \\ &= \mathbb{E} [q_{i,t+1}^2] - \mathbb{E} \left[\left[\mathbb{E} [q_{i,t+1}] + \frac{c\delta(1+\lambda)(q_{i,t} - \mathbb{E} [q_{i,t}])}{\beta B + c(1+\lambda)} \right]^2 \right] \\ &= \text{Var} [q_{i,t+1}] - \left[\frac{c\delta(1+\lambda)}{\beta B + c(1+\lambda)} \right]^2 \text{Var} [q_{i,t}], \end{aligned}$$

where the second equality follows from (27)–(28) and (29)–(30). Using (75)–(76), this yields:

$$\begin{aligned} \text{Var} [q_{1,t+1}] &= \left[\frac{c(1+\lambda)}{\beta B + c(1+\lambda)} \right]^2 (\text{Var}_\theta + \delta^2 \text{Var} [q_{1,t}]), \\ \text{Var} [q_{2,t+1}] &= \left[\frac{c(1+\lambda)}{\beta B + c(1+\lambda)} \right]^2 (\text{Var}_\theta + \delta^2 \text{Var} [q_{2,t}]) + \frac{\lambda \nu^2 c^2 \Delta \theta^2 [\nu \lambda + 2(1-\nu)(1+\lambda)]}{(1-\nu)[\beta B + c(1+\lambda)]^2} \end{aligned}$$

for each $t = 0, 1, \dots$. Taking limits as t goes to infinity, one obtains that:

$$\begin{aligned} \text{Var}_{\mu_1} &= \frac{c^2(1+\lambda)^2 \text{Var}_\theta}{[\beta B + c(1+\lambda)]^2 - c^2 \delta^2 (1+\lambda)^2}, \\ \text{Var}_{\mu_2} &= \frac{c^2(1+\lambda)^2 \text{Var}_\theta}{[\beta B + c(1+\lambda)]^2 - c^2 \delta^2 (1+\lambda)^2} + \frac{\lambda \nu^2 c^2 \Delta \theta^2 [\nu \lambda + 2(1-\nu)(1+\lambda)]}{(1-\nu)\{[\beta B + c(1+\lambda)]^2 - c^2 \delta^2 (1+\lambda)^2\}}, \end{aligned}$$

from which (36) follows. Hence the result. ■

Proof of Proposition 6. Using date 1 and 2 qualities and date 0 rents as control variables, the regulator's problem can be formulated as follows:

$$\begin{aligned} \max_{(\tilde{q}_1, \tilde{q}_2, \tilde{U}_0)} \{ & \mathbb{E}[-(1 + \lambda)\psi(\tilde{q}_1(\theta_0) - \delta q_0 - \theta_0) + \beta S(\tilde{q}_1(\theta_0)) \\ & - (1 + \lambda)\beta\psi(\tilde{q}_2(\theta_0, \theta_1) - \delta\tilde{q}_1(\theta_0) - \theta_1) + \beta^2 S(\tilde{q}_2(\theta_0, \theta_1))] \\ & - \lambda\nu\tilde{U}_0(\bar{\theta}) - \lambda(1 - \nu)\tilde{U}_0(\underline{\theta}) - \lambda\nu^2\beta\Phi(\tilde{q}_1(\bar{\theta}), \tilde{q}_2(\bar{\theta}, \underline{\theta})) - \lambda\nu(1 - \nu)\beta\Phi(\tilde{q}_1(\underline{\theta}), \tilde{q}_2(\underline{\theta}, \underline{\theta}))\}, \end{aligned} \quad (77)$$

subject to the constraints:

$$\tilde{U}_0(\bar{\theta}) + \nu\beta\Phi(\tilde{q}_1(\bar{\theta}), \tilde{q}_2(\bar{\theta}, \underline{\theta})) \geq \tilde{U}_0(\underline{\theta}) + \Phi(q_0, \tilde{q}_1(\underline{\theta})) + \nu\beta\Phi(\tilde{q}_1(\underline{\theta}), \tilde{q}_2(\underline{\theta}, \underline{\theta})), \quad (78)$$

$$\tilde{U}_0(\underline{\theta}) + \nu\beta\Phi(\tilde{q}_1(\underline{\theta}), \tilde{q}_2(\underline{\theta}, \underline{\theta})) \geq \tilde{U}_0(\bar{\theta}) - \Phi(q_0, \tilde{q}_1(\bar{\theta})) + \nu\beta\Phi(\tilde{q}_1(\bar{\theta}), \tilde{q}_2(\bar{\theta}, \underline{\theta})), \quad (79)$$

$$\tilde{U}_0(\bar{\theta}) \geq 0, \quad (80)$$

$$\tilde{U}_0(\underline{\theta}) \geq 0. \quad (81)$$

Let $(q_{2,1}, q_{2,2}, U_{2,0})$ be the solution to this problem. One first has the following lemma.

Lemma A.2. $q_{2,1}(\bar{\theta}) \geq q_{2,1}(\underline{\theta})$.

Proof. An immediate implication of (78)–(79) is that $\Phi(q_0, q_{2,1}(\bar{\theta})) \geq \Phi(q_0, q_{2,1}(\underline{\theta}))$. Since $\Phi(q, \underline{q})$ is strictly increasing in \underline{q} for $\underline{q} > \delta q + \underline{\theta}$, we need only to show that $q_{2,1}(\bar{\theta}) > \delta q_0 + \underline{\theta}$. Suppose the contrary holds, and consider the following modification of $(q_{2,1}, q_{2,2}, U_{2,0})$. First, keep $\tilde{q}_1(\underline{\theta})$, \tilde{q}_2 , and $\tilde{U}_0(\underline{\theta})$ the same as in the original contract, so that (81) is preserved. Second, let $\tilde{q}_1(\bar{\theta}) = \delta q_0 + \underline{\theta} + \varepsilon$ for some $\varepsilon > 0$, and let $\tilde{U}_0(\bar{\theta}) = U_{2,0}(\bar{\theta}) + \nu\beta\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) - \nu\beta\Phi(\tilde{q}_1(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta}))$ so as to maintain the same level of intertemporal rents for a good type agent at date 0. As a result of this, (78) is preserved. Next, since $\tilde{q}_1(\bar{\theta}) > q_{2,1}(\bar{\theta})$ and Φ is increasing with respect to its second argument, (79) is also preserved. Finally, since $\tilde{q}_1(\bar{\theta}) > q_{2,1}(\bar{\theta})$ and Φ is decreasing with respect to its first argument, one obtains that $\tilde{U}_0(\bar{\theta}) \geq U_{2,0}(\bar{\theta})$ so that (80) is preserved. The new contract is therefore incentive feasible, and yields the same expected intertemporal rent to the agent as the original contract. The condition $\psi'(0) = 0$ then implies that, for ε close enough to zero, this contract yields a strictly higher expected social welfare at date 0 than the original contract, a contradiction. \square

It turns out that which constraints are binding at the optimum depends on the sign of the following quantities:

$$\Delta_r = \Phi(q_0, q_{2,1}(\underline{\theta})) + \nu\beta[\Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta})) - \Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta}))], \quad (82)$$

$$\Delta_i = -\Phi(q_0, q_{2,1}(\bar{\theta})) + \nu\beta[\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) - \Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta}))]. \quad (83)$$

Δ_r is the date 0 rent that must be left to a high type agent whenever (39) and (42) are the binding constraints. For (41) to be satisfied, one must then have $\Delta_r \geq 0$. Similarly, Δ_i is the date 0 rent that must be left to a low type agent whenever (40) and (41) are the binding constraints. For (42) to be satisfied, one must then have $\Delta_i \geq 0$. The following lemma holds.

Lemma A.3. Δ_r and Δ_i cannot be both non-negative unless they are both equal to zero. Moreover, at least one of the following assertions is true:

- (i) $\Delta_r \geq 0 \geq \Delta_i$, in which case (39) and (42) are binding;
- (ii) $\Delta_i \geq 0 \geq \Delta_r$, in which case (40) and (41) are binding.

Proof. We first establish that Δ_r and Δ_i cannot be both non-negative unless they are both equal to zero. Suppose the contrary holds. Then, by (82)–(83),

$$\Phi(q_0, q_{2,1}(\underline{\theta})) \geq \nu\beta[\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) - \Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta}))] \geq \Phi(q_0, q_{2,1}(\bar{\theta})),$$

with at least one strict inequality. This however implies that $\Phi(q_0, q_{2,1}(\underline{\theta})) > \Phi(q_0, q_{2,1}(\bar{\theta}))$, which contradicts (39)–(40). This implies the claim.

Suppose next that $\Delta_r \geq 0$ at the optimum contract. Observing that one can rewrite (39) as $u_{2,0}(\bar{\theta}) - \psi(e_{2,0}(\bar{\theta})) \geq u_{2,0}(\underline{\theta}) - \psi(e_{2,0}(\underline{\theta})) + \Delta_r$ using (82), it necessarily follows that (42) is binding, since otherwise reducing $u_{2,0}(\bar{\theta})$ and $u_{2,0}(\underline{\theta})$ by some small amount $\varepsilon > 0$ would preserve all constraints and increase social welfare. Similarly, (39) must be binding since otherwise reducing $u_{2,0}(\bar{\theta})$ by some small amount $\varepsilon > 0$ would preserve all constraints and increase social welfare. Similar arguments imply that if $\Delta_i \geq 0$ at the optimum contract, then (40) and (41) must be binding.

To conclude the proof, we must only check that Δ_r and Δ_i cannot be both strictly negative. Suppose the contrary holds. Then, observing that one can rewrite (39) as $u_{2,0}(\bar{\theta}) - \psi(e_{2,0}(\bar{\theta})) \geq u_{2,0}(\underline{\theta}) - \psi(e_{2,0}(\underline{\theta})) + \Delta_r$ using (82), and (40) as $u_{2,0}(\underline{\theta}) - \psi(e_{2,0}(\underline{\theta})) \geq u_{2,0}(\bar{\theta}) - \psi(e_{2,0}(\bar{\theta})) + \Delta_i$ using (83), it is clear that (39) and (42) cannot be both binding, as this would violate (41) since $\Delta_r < 0$, and that (40) and (41) cannot be both binding, as this would violate (42) since $\Delta_i < 0$. Now, suppose that (41) is slack. Then (42) must be binding, since otherwise reducing $u_{2,0}(\bar{\theta})$ and $u_{2,0}(\underline{\theta})$ by some small amount $\varepsilon > 0$ would preserve all constraints and increase social welfare. Thus, as (39) and (42) cannot be both binding, (39) must be slack. But (39) and (41) cannot both be slack, since otherwise reducing $u_{2,0}(\bar{\theta})$ by some small amount $\varepsilon > 0$ would preserve all constraints and increase social welfare. Hence, (41) must be binding, contrary to the assumption. A similar reasoning implies that (42) must be binding as well. Since (39) and (42) cannot be both binding, and similarly for (40) and (41), this implies that (39)–(40) are slack. Therefore, for a given value of q_0 , the function $q_{2,2}$ solves:

$$\begin{aligned} \max_{\tilde{q}_2} \{ & \mathbb{E}[-(1 + \lambda)\psi(\tilde{q}_2(\theta_0, \theta_1) - \delta q_{2,1}(\theta_0) - \theta_1) + \beta S(\tilde{q}_2(\theta_0, \theta_1))] \\ & - \lambda\nu^2\Phi(q_{2,1}(\bar{\theta}), \tilde{q}_2(\bar{\theta}, \underline{\theta})) - \lambda\nu(1 - \nu)\Phi(q_{2,1}(\underline{\theta}), \tilde{q}_2(\underline{\theta}, \underline{\theta})) \}. \end{aligned}$$

It follows that $q_{2,2}(\bar{\theta}, \underline{\theta}) = \underline{\mathcal{Q}}_{2,2}(q_{2,1}(\bar{\theta}))$ and $q_{2,2}(\underline{\theta}, \underline{\theta}) = \underline{\mathcal{Q}}_{2,2}(q_{2,1}(\underline{\theta}))$, where for each $q_1 \geq 0$, $\underline{\mathcal{Q}}_{2,2}(q_1)$ is implicitly defined by:¹⁴

$$\beta S'(\underline{\mathcal{Q}}_{2,2}(q_1)) = (1 + \lambda)\psi'(\underline{\mathcal{Q}}_{2,2}(q_1) - \delta q_1 - \underline{\theta}) + \frac{\lambda\nu}{1 - \nu} \Phi_2(q_1, \underline{\mathcal{Q}}_{2,2}(q_1)). \quad (84)$$

The condition $\psi'(0) = 0$ ensures that $\underline{\mathcal{E}}_{2,1}(q_1) = \underline{\mathcal{Q}}_{2,2}(q_1) - \delta q_1 - \underline{\theta}$ remains strictly positive for each $q_1 \geq 0$. Using now standard arguments, it is easy to check from (84) that the function $\underline{\mathcal{E}}_{2,1}$ is strictly decreasing, which implies in turn that the informational rent:

$$\Phi(q_1, \underline{\mathcal{Q}}_{2,2}(q_1)) = \psi(\underline{\mathcal{E}}_{2,1}(q_1)) - \psi(\underline{\mathcal{E}}_{2,1}(q_1) - \Delta\theta)$$

¹⁴Note that (84) is formally analogous to the first-order condition (22) for $\underline{q}_2(q)$ derived in the stationary model with multiple agents. Indeed, it is easy to see that (84) gives the sequentially rational date 2 optimal quality policy given that $\theta_1 = \underline{\theta}$.

is a strictly decreasing function of q_1 . Since $\Delta_r < 0$ and $q_{2,1}(\bar{\theta}) \geq q_{2,1}(\underline{\theta})$ by Lemma A.2, one has:

$$\begin{aligned}
0 &\leq \Phi(q_0, q_{2,1}(\underline{\theta})) \\
&< \nu\beta [\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) - \Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta}))] \\
&= \nu\beta [\Phi(q_{2,1}(\bar{\theta}), \underline{Q}_{2,2}(q_{2,1}(\bar{\theta}))) - \Phi(q_{2,1}(\underline{\theta}), \underline{Q}_{2,2}(q_{2,1}(\underline{\theta})))] \\
&\leq 0,
\end{aligned}$$

a contradiction. Hence the result. \square

To determine which of cases (i) or (ii) of Lemma A.3 holds, we first need the following result.

Lemma A.4. $\Delta_i \leq 0$.

Proof. Suppose by way of contradiction that $\Delta_i > 0$. Then, by Lemma A.3, (40) and (41) are binding. Moreover, a low type agent is left with a strictly positive rent Δ_i at date 0, so that (42) is slack. Therefore, for a given value of q_0 , the function $q_{2,2}$ solves:

$$\max_{\tilde{q}_2} \{E[-(1+\lambda)\psi(\tilde{q}_2(\theta_0, \theta_1) - q_{2,1}(\theta_0) - \theta_1) + \beta S(\tilde{q}_2(\theta_0, \theta_1))] - \lambda\nu\Phi(q_{2,1}(\bar{\theta}), \tilde{q}_2(\bar{\theta}, \underline{\theta}))\}.$$

It follows that $q_{2,2}(\bar{\theta}, \underline{\theta}) = \underline{Q}_{3,2}(q_{2,1}(\bar{\theta}))$ and $q_{2,2}(\underline{\theta}, \underline{\theta}) = \underline{Q}_{1,2}(q_{2,1}(\underline{\theta}))$, where for each $q_1 \geq 0$, $\underline{Q}_{3,2}(q_1)$ and $\underline{Q}_{1,2}(q_1)$ are implicitly defined by:

$$\beta S'(\underline{Q}_{3,2}(q_1)) = (1+\lambda)\psi'(\underline{Q}_{3,2}(q_1) - \delta q_1 - \underline{\theta}) + \frac{\lambda}{1-\nu} \Phi_2(q_1, \underline{Q}_{3,2}(q_1)). \quad (85)$$

$$\beta S'(\underline{Q}_{1,2}(q_1)) = (1+\lambda)\psi'(\underline{Q}_{1,2}(q_1) - \delta q_1 - \underline{\theta}), \quad (86)$$

The condition $\psi'(0) = 0$ ensures that $\underline{\mathcal{E}}_{1,1}(q_1) = \underline{Q}_{1,2}(q_1) - \delta q_1 - \underline{\theta}$ and $\underline{\mathcal{E}}_{3,1}(q_1) = \underline{Q}_{3,2}(q_1) - \delta q_1 - \underline{\theta}$ remain strictly positive for each $q_1 \geq 0$. Since one clearly has $\underline{Q}_{3,2}(q_1) < \underline{Q}_{1,2}(q_1)$ by (86)–(85), this implies that $\Phi(q_1, \underline{Q}_{3,2}(q_1)) < \Phi(q_1, \underline{Q}_{1,2}(q_1))$ for each $q_1 \geq 0$. Using now standard arguments, it is easy to check from (86) that the function $\underline{\mathcal{E}}_{1,1}$ is strictly decreasing, which implies in turn that the informational rent:

$$\Phi(q_1, \underline{Q}_{1,2}(q_1)) = \psi(\underline{\mathcal{E}}_{1,1}(q_1)) - \psi(\underline{\mathcal{E}}_{1,1}(q_1) - \Delta\theta)$$

is a strictly decreasing function of q_1 . Since $\Delta_i > 0$ and $q_{2,1}(\underline{\theta}) \geq q_{2,1}(\bar{\theta})$ by Lemma A.2, one has:

$$\begin{aligned}
0 &\leq \Phi(q_0, q_{2,1}(\bar{\theta})) \\
&< \nu\beta [\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) - \Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta}))] \\
&= \nu\beta [\Phi(q_{2,1}(\bar{\theta}), \underline{Q}_{3,2}(q_{2,1}(\bar{\theta}))) - \Phi(q_{2,1}(\underline{\theta}), \underline{Q}_{1,2}(q_{2,1}(\underline{\theta})))] \\
&< \nu\beta [\Phi(q_{2,1}(\bar{\theta}), \underline{Q}_{1,2}(q_{2,1}(\bar{\theta}))) - \Phi(q_{2,1}(\underline{\theta}), \underline{Q}_{1,2}(q_{2,1}(\underline{\theta})))] \\
&\leq 0,
\end{aligned}$$

a contradiction. Hence the result. \square

Finally, the following lemma rules out the case $\Delta_i = 0 > \Delta_r$.

Lemma A.5. $\Delta_r \geq 0$.

Proof. Suppose by way of contradiction that $\Delta_r < 0$. Then, by Lemmas A.3 and A.4, $\Delta_i = 0$, so that (40)–(42) are binding. Observing that one can rewrite (39) as $u_{2,0}(\bar{\theta}) - \psi(e_{2,0}(\bar{\theta})) \geq u_{2,0}(\underline{\theta}) - \psi(e_{2,0}(\underline{\theta})) + \Delta_r$ using (82), it necessarily follows that (39) is slack. Now turn to the formulation (77)–(81) of the regulator’s problem. We know from the proof of Lemma A.2 that $q_{2,1}(\bar{\theta}) > \delta q_0 + \underline{\theta}$. It is easy to check that this implies that (78)–(81) satisfy the Kuhn-Tucker constraint qualification conditions at the optimum. Denote by η_1, η_2, η_3 and η_4 the corresponding multipliers. As (78) is slack at the optimum, $\eta_1 = 0$. It follows that $q_{2,2}(\bar{\theta}, \underline{\theta}) = \underline{\mathcal{Q}}_{4,2}(q_{2,1}(\bar{\theta}))$ and $q_{2,2}(\underline{\theta}, \underline{\theta}) = \underline{\mathcal{Q}}_{5,2}(q_{2,1}(\underline{\theta}))$, where for each $q_1 \geq 0$, $\underline{\mathcal{Q}}_{4,2}(q_1)$ and $\underline{\mathcal{Q}}_{5,2}(q_1)$ are implicitly defined by:

$$\beta S'(\underline{\mathcal{Q}}_{4,2}(q_1)) = (1 + \lambda)\psi'(\underline{\mathcal{Q}}_{4,2}(q_1) - \delta q_1 - \underline{\theta}) + \frac{\lambda\nu + \eta_2}{1 - \nu} \Phi_2(q_1, \underline{\mathcal{Q}}_{4,2}(q_1)), \quad (87)$$

$$\beta S'(\underline{\mathcal{Q}}_{5,2}(q_1)) = (1 + \lambda)\psi'(\underline{\mathcal{Q}}_{5,2}(q_1) - \delta q_1 - \underline{\theta}) + \frac{[\lambda(1 - \nu) - \eta_2]\nu}{(1 - \nu)^2} \Phi_2(q_1, \underline{\mathcal{Q}}_{5,2}(q_1)). \quad (88)$$

Optimizing with respect to $\tilde{U}_0(\underline{\theta})$ yields $\eta_2 + \eta_4 = \lambda(1 - \nu)$, so that $\eta_2 \leq \lambda(1 - \nu)$. The condition $\psi'(0) = 0$ then ensures that $\underline{\mathcal{E}}_{5,1}(q_1) = \underline{\mathcal{Q}}_{5,2}(q_1) - \delta q_1 - \underline{\theta}$ remains strictly positive for each $q_1 \geq 0$. Since one clearly has $\underline{\mathcal{Q}}_{4,2}(q_1) \leq \underline{\mathcal{Q}}_{5,2}(q_1)$ by (87)–(88), $\Phi(q_1, \underline{\mathcal{Q}}_{4,2}(q_1)) \leq \Phi(q_1, \underline{\mathcal{Q}}_{5,2}(q_1))$ for each $q_1 \geq 0$. Using now standard arguments, it is easy to check from (88) that the function $\underline{\mathcal{E}}_{5,1}$ is strictly decreasing, which implies in turn that the informational rent:

$$\Phi(q_1, \underline{\mathcal{Q}}_{5,2}(q_1)) = \psi(\underline{\mathcal{E}}_{5,1}(q_1)) - \psi(\underline{\mathcal{E}}_{5,1}(q_1) - \Delta\theta)$$

is a strictly decreasing function of q_1 . Since $\Delta_r < 0$ and $q_{2,1}(\bar{\theta}) \geq q_{2,1}(\underline{\theta})$ by Lemma A.2, one has:

$$\begin{aligned} 0 &\leq \Phi(q_0, q_{2,1}(\underline{\theta})) \\ &< \nu\beta [\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) - \Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta}))] \\ &= \nu\beta [\Phi(q_{2,1}(\bar{\theta}), \underline{\mathcal{Q}}_{4,2}(q_{2,1}(\bar{\theta}))) - \Phi(q_{2,1}(\underline{\theta}), \underline{\mathcal{Q}}_{5,2}(q_{2,1}(\underline{\theta})))] \\ &\leq \nu\beta [\Phi(q_{2,1}(\bar{\theta}), \underline{\mathcal{Q}}_{5,2}(q_{2,1}(\bar{\theta}))) - \Phi(q_{2,1}(\underline{\theta}), \underline{\mathcal{Q}}_{5,2}(q_{2,1}(\underline{\theta})))] \\ &\leq 0, \end{aligned}$$

a contradiction. Hence the result. \square

To conclude the proof, simply observe that Lemmas A.4 and A.5 imply that case (i) of Lemma A.3 obtains. Thus (39) and (42) are binding, and the result follows. \blacksquare

Proof of Proposition 7. We first prove that $q_{2,2}(\bar{\theta}, \underline{\theta}) > q_{2,2}(\underline{\theta}, \underline{\theta})$. In the regular case, this follows from (52)–(53) together with $q_{2,1}(\bar{\theta}) \geq q_{2,1}(\underline{\theta})$. Consider next the irregular case. Two subcases must

be distinguished. Suppose first that (79) or equivalently (47) is binding. From the proof of Lemma A.2, we know that $q_{2,1}(\bar{\theta}) > \delta q_0 + \underline{\theta}$, and hence $\Phi(q_0, q_{2,1}(\underline{\theta})) = \Phi(q_0, q_{2,1}(\bar{\theta})) > 0$. Since $U_{2,0}(\bar{\theta}) = 0$, it follows that $\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) > \Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta}))$, which implies that $q_{2,2}(\bar{\theta}, \underline{\theta}) > q_{2,2}(\underline{\theta}, \underline{\theta})$ as claimed. Suppose next that (79) or equivalently (47) is slack. Since $U_{2,0}(\bar{\theta}) = 0$, $\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) \geq \Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta}))$. As (47) is slack, we need only to prove that this implies that $q_{2,2}(\bar{\theta}, \underline{\theta}) - q_{2,2}(\underline{\theta}, \underline{\theta}) \geq \delta[q_{2,1}(\bar{\theta}) - q_{2,1}(\underline{\theta})]$. Again, this will be the case if $q_{2,2}(\bar{\theta}, \underline{\theta}) > \delta q_{2,1}(\bar{\theta}) + \underline{\theta}$. Suppose the contrary holds, and consider the following modification of $(q_{2,1}, q_{2,2}, U_{2,0})$. First, keep \tilde{q}_1 , $\tilde{q}_2(\bar{\theta}, \bar{\theta})$, $\tilde{q}_2(\underline{\theta}, \bar{\theta})$, $\tilde{q}_2(\underline{\theta}, \underline{\theta})$ and \tilde{U}_0 the same as in the original contract, so that (80)–(81) are preserved. Second, let $\tilde{q}_2(\bar{\theta}, \underline{\theta}) = \delta q_{2,1}(\bar{\theta}) + \underline{\theta} + \varepsilon$ for some $\varepsilon > 0$. Since $\tilde{q}_2(\bar{\theta}, \underline{\theta}) > q_{2,2}(\bar{\theta}, \underline{\theta})$ and Φ is increasing in its second argument, (78) is preserved. Moreover, since (79) is slack, it is preserved in the new contract if ε is close enough to zero, which ensures that the new contract is incentive feasible. The condition $\psi'(0) = 0$ then implies that, for ε close enough to zero, this contract yields a strictly higher expected social welfare at date 0 than the original contract, a contradiction. Thus $q_{2,2}(\bar{\theta}, \underline{\theta}) > \delta q_{2,1}(\bar{\theta}) + \underline{\theta}$, which implies that $q_{2,2}(\bar{\theta}, \underline{\theta}) > q_{2,2}(\underline{\theta}, \underline{\theta})$ as claimed. Note that $(\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})), \Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta}))) \neq (0, 0)$, so that (47)–(48) satisfy the Kuhn-Tucker constraint qualification conditions at the optimum.

We now show that the level of distortions at date 1 is strictly higher following $\theta_0 = \underline{\theta}$ than following $\theta_0 = \bar{\theta}$. Let ξ_2 be the multiplier associated to (48). Then, from (47)–(49), the first-order conditions for $q_{2,2}(\bar{\theta}, \underline{\theta})$ and $q_{2,2}(\underline{\theta}, \underline{\theta})$ are given by:

$$\beta S'(q_{2,2}(\bar{\theta}, \underline{\theta})) = (1 + \lambda)\psi'(q_{2,2}(\bar{\theta}, \underline{\theta}) - \delta q_{2,1}(\bar{\theta}) - \underline{\theta}) + \frac{\xi_2}{1 - \nu} \Phi_2(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})), \quad (89)$$

$$\beta S'(q_{2,2}(\underline{\theta}, \underline{\theta})) = (1 + \lambda)\psi'(q_{2,2}(\underline{\theta}, \underline{\theta}) - \delta q_{2,1}(\underline{\theta}) - \underline{\theta}) + \frac{(\lambda - \xi_2)\nu}{(1 - \nu)^2} \Phi_2(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta})). \quad (90)$$

We need to prove that $(\lambda - \xi_2)\nu > \xi_2(1 - \nu)$ or equivalently $\lambda\nu > \xi_2$. If $\xi_2 = 0$ this is immediate. Suppose next that $\xi_2 > 0$, so that (48) is binding. Then, proceeding as above, one obtains that $q_{2,2}(\bar{\theta}, \underline{\theta}) - \delta q_{2,1}(\bar{\theta}) \geq q_{2,2}(\underline{\theta}, \underline{\theta}) - \delta q_{2,1}(\underline{\theta})$. Since ψ' is convex, it therefore follows from (18) that $\Phi_2(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta})) \geq \Phi_2(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta}))$. Suppose now that $\xi_2 \geq \lambda\nu$. Then the right-hand side of (89) is greater than or equal to the right-hand side of (90). This is however impossible since $q_{2,2}(\bar{\theta}, \underline{\theta}) > q_{2,2}(\underline{\theta}, \underline{\theta})$ and S is strictly concave. Thus $\lambda\nu > \xi_2$, and the result follows. ■

Proof of Proposition 8. Suppose that pooling over date 1 qualities occurs in equilibrium, and let ξ_1 and ξ_2 be the multipliers associated to (47)–(48). Using the Envelope Theorem, the first-order conditions for $q_{2,1}(\bar{\theta}) = q_{2,1}(\underline{\theta}) = q_{2,1}^p$ can be written as follows:

$$\begin{aligned} \beta S'(q_{2,1}^p) &= (1 + \lambda)\psi'(q_{2,1}^p - \delta q_0 - \bar{\theta}) - \beta\delta(1 + \lambda)\mathbb{E}[\psi'(q_{2,2}(\bar{\theta}, \theta_1)) - \delta q_{2,1}^p - \theta_1] \\ &\quad - \frac{\xi_1}{\nu} + \xi_2\beta\Phi_1(q_{2,1}^p, q_{2,2}(\bar{\theta}, \underline{\theta})), \end{aligned} \quad (91)$$

$$\begin{aligned} \beta S'(q_{2,1}^p) &= (1 + \lambda)\psi'(q_{2,1}^p - \delta q_0 - \underline{\theta}) - \beta\delta(1 + \lambda)\mathbb{E}[\psi'(q_{2,2}(\underline{\theta}, \theta_1)) - \delta q_{2,1}^p - \theta_1] \\ &\quad + \frac{\lambda\nu}{1 - \nu} \Phi_2(q_0, q_{2,1}^p) + \frac{\lambda\nu\beta}{1 - \nu} \Phi_1(q_{2,1}^p, q_{2,2}(\underline{\theta}, \underline{\theta})), \\ &\quad + \frac{\xi_1}{1 - \nu} - \frac{\xi_2}{1 - \nu} \Phi_2(q_0, q_{2,1}^p) - \frac{\xi_2\nu\beta}{1 - \nu} \Phi_1(q_{2,1}^p, q_{2,2}(\underline{\theta}, \underline{\theta})), \end{aligned} \quad (92)$$

The following lemma holds.

Lemma A.6. *If q_0 is close enough to 0, $q_{2,1}^p - \delta q_0 > q_{2,2}(\underline{\theta}, \underline{\theta}) - \delta q_{2,1}^p$.*

Proof. Using the definition of Φ , and in particular the fact that $\Phi_1 = -\delta\Phi_2$, one can rewrite (91) as:

$$\begin{aligned} \beta S'(q_{2,1}^p) &= (1 + \lambda)\psi'(q_{2,1}^p - \delta q_0 - \underline{\theta}) - (1 + \lambda)\Phi_2(q_0, q_{1,2}^p) - \beta\delta(1 + \lambda)\mathbb{E}[\psi'(q_{2,2}(\bar{\theta}, \theta_1)) - \delta q_{2,1}^p - \theta_1] \\ &\quad - \frac{\xi_1}{\nu} - \xi_2\beta\delta\Phi_2(q_{2,1}^p, q_{2,2}(\bar{\theta}, \underline{\theta})). \end{aligned} \quad (93)$$

Since S is concave and ψ is convex, and since $\lambda\nu > \xi_2$ as shown in the proof of Proposition 7, it follows from (90) and (93) that if $q_{2,2}(\underline{\theta}, \underline{\theta}) - \delta q_{2,1}^p \geq q_{2,1}^p - \delta q_0$, then $q_{2,1}^p \geq q_{2,2}(\underline{\theta}, \underline{\theta})$ and therefore $q_0 \geq q_{2,1}^p$. But if q_0 is close to zero, this is inconsistent with a positive effort being exerted at date 0 by a high type agent, in contradiction to (91). Hence $q_{2,1}^p - \delta q_0 > q_{2,2}(\underline{\theta}, \underline{\theta}) - \delta q_{2,1}^p$, as claimed. \square

We are now ready to complete the proof. Using (50)–(51) together with the definition of Φ , one can rearrange (91)–(92) to obtain:

$$\begin{aligned} \xi_1 \left(\frac{1}{\nu} + \frac{1}{1 - \nu} \right) &= -(1 - \nu)\beta\delta(1 + \lambda) [\psi'(q_{2,2}(\bar{\theta}, \underline{\theta}) - \delta q_{2,1}^p - \underline{\theta}) - \psi'(q_{2,2}(\underline{\theta}, \underline{\theta}) - \delta q_{2,1}^p - \underline{\theta})] \\ &\quad - \left(1 + \lambda + \frac{\lambda\nu}{1 - \nu} \right) \Phi_2(q_0, q_{2,1}^p) + \frac{\lambda\nu\beta\delta}{1 - \nu} \Phi_2(q_{2,1}^p, q_{2,2}(\underline{\theta}, \underline{\theta})) \\ &\quad - \xi_2\beta\delta\Phi_2(q_{2,1}^p, q_{2,2}(\bar{\theta}, \underline{\theta})) + \frac{\xi_2}{1 - \nu} \Phi_2(q_0, q_{2,1}^p) - \frac{\xi_2\nu\beta\delta}{1 - \nu} \Phi_2(q_{2,1}^p, q_{2,2}(\underline{\theta}, \underline{\theta})). \end{aligned} \quad (94)$$

Since $q_{2,2}(\bar{\theta}, \underline{\theta}) > q_{2,2}(\underline{\theta}, \underline{\theta})$ by Proposition 7, and since the functions ψ and $\Phi(q_{2,1}^p, \cdot)$ are convex, it therefore follows that:

$$\frac{\xi_1}{\nu} < (\lambda\nu - \xi_2) [\beta\delta\Phi_2(q_{2,1}^p, q_{2,2}(\underline{\theta}, \underline{\theta})) - \Phi_2(q_0, q_{2,1}^p)]. \quad (95)$$

By Lemma A.6, $q_{2,1}^p - \delta q_0 > q_{2,2}(\underline{\theta}, \underline{\theta}) - \delta q_{2,1}^p$ for q_0 close enough to 0. Hence, since a high type agent exerts a positive effort at date 0, $\Phi_2(q_0, q_{2,1}^p) > \Phi_2(q_{2,1}^p, q_{2,2}(\underline{\theta}, \underline{\theta}))$. Moreover, $\lambda\nu > \xi_2$ as shown in the proof of Proposition 7. Hence (95) implies that if $\beta\delta \leq 1$, then $\xi_1 < 0$, a contradiction. Thus pooling over date 1 qualities can occur only if $\beta\delta > 1$, as claimed. \blacksquare

Proof of Proposition 9. To construct the optimal contract, we first conjecture that (48) is slack at the optimum, which we will verify ex-post. As a result of this, the multiplier ξ_2 of (48) is equal to zero. From (50)–(53), we obtain:

$$q_{2,2}(\bar{\theta}, \bar{\theta}) = \frac{\beta a + c(1 + \lambda)[\bar{\theta} + \delta q_{2,1}(\bar{\theta})]}{\beta b + c(1 + \lambda)}, \quad (96)$$

$$q_{2,2}(\underline{\theta}, \bar{\theta}) = \frac{\beta a + c(1 + \lambda)[\bar{\theta} + \delta q_{2,1}(\underline{\theta})]}{\beta b + c(1 + \lambda)}, \quad (97)$$

$$q_{2,2}(\bar{\theta}, \underline{\theta}) = \frac{\beta a + c(1 + \lambda)[\underline{\theta} + \delta q_{2,1}(\bar{\theta})]}{\beta b + c(1 + \lambda)}, \quad (98)$$

$$q_{2,2}(\underline{\theta}, \underline{\theta}) = \frac{\beta a - \frac{\lambda\nu}{(1-\nu)^2} c\Delta\theta + c(1 + \lambda)[\underline{\theta} + \delta q_{2,1}(\underline{\theta})]}{\beta b + c(1 + \lambda)}. \quad (99)$$

For (96)–(99) to hold, one must ensure that whatever $q_{2,1}(\bar{\theta})$ and $q_{2,1}(\underline{\theta})$ might be, provided they remain smaller than $\frac{a}{\beta}$, the date 1 efforts stay positive, and that $e_{2,1}(\bar{\theta}, \underline{\theta}) \geq \Delta\theta$ and $e_{2,1}(\underline{\theta}, \underline{\theta}) \geq \Delta\theta$, so

that we are justified in using the quadratic specification of ψ when writing down the informational rents $\Phi(q_{2,1}(\bar{\theta}), q_{2,2}(\bar{\theta}, \underline{\theta}))$ and $\Phi(q_{2,1}(\underline{\theta}), q_{2,2}(\underline{\theta}, \underline{\theta}))$, see (31). It is easy to check that a sufficient condition for this to be true is that (72) holds and that:

$$\left[(1 - \delta) \frac{a}{b} - \underline{\theta} \right] \beta b > \left[\frac{\lambda \nu c}{(1 - \nu)^2} + \beta b + c(1 + \lambda) \right] \Delta \theta, \quad (100)$$

which, given (72), typically holds if $\Delta \theta$ is close enough to zero. To show that pooling can occur, let us evaluate the multiplier ξ_1 of (47) using (94). This yields:

$$\xi_1 = -c \Delta \theta \nu (1 - \nu) \varphi_\delta(\beta), \quad (101)$$

where:

$$\varphi_\delta(\beta) = 1 + \lambda + \frac{\lambda \nu}{1 - \nu} \left[\frac{\beta \delta c (1 + \lambda)}{\beta b + c(1 + \lambda)} + 1 - \beta \delta \right]. \quad (102)$$

For each $\delta > 0$, the function φ_δ is strictly decreasing, with $\varphi_\delta(0) > 0$ and $\lim_{\beta \rightarrow \infty} \varphi_\delta(\beta) = -\infty$. Let $\underline{\beta}(\delta) = \varphi_\delta^{-1}(0)$. Clearly $\underline{\beta}(\delta) > 1$ and $\lim_{\delta \rightarrow 1} \underline{\beta}(\delta) < \infty$. Thus $\xi_1 > 0$ if $\beta > \underline{\beta}(\delta)$. For the corresponding allocation to be economically meaningful, $q_{2,1}(\bar{\theta}) = q_{2,1}(\underline{\theta}) = q_{2,1}^p$ must remain smaller than $\frac{a}{b}$. For q_0 , $\underline{\theta}$ and $\Delta \theta$ close enough to zero, it is easy to check this will be the case if $\beta < \bar{\beta}(\delta)$, where $\bar{\beta}(\delta)$ is the positive solution to (54). From (54), one has:

$$\varphi_\delta(\bar{\beta}(\delta)) = 1 + \lambda + \frac{\lambda \nu}{1 - \nu} \left[\frac{c(1 + \lambda)}{\bar{\beta}(\delta)(1 - \delta)b} + 1 - \bar{\beta}(\delta)\delta \right], \quad (103)$$

as well as $\lim_{\delta \rightarrow 1} \bar{\beta}(\delta) = \infty$ and $\lim_{\delta \rightarrow 1} \bar{\beta}(\delta)(1 - \delta) = 1$. Thus (103) implies that $\lim_{\delta \rightarrow 1} \varphi_\delta(\bar{\beta}(\delta)) = -\infty$. Moreover, $\bar{\beta}(\delta) > \underline{\beta}(\delta)$ whenever δ is close enough to one. Thus, pooling can occur if $\beta \in [\underline{\beta}(\delta), \bar{\beta}(\delta)]$.¹⁵ To check that this can indeed be the case, one needs only to find parameter restrictions such that $U_{2,0}(\bar{\theta}) > 0$, so that (48) is slack as postulated. From (31), (43), and (98)–(99), one has:

$$U_{2,0}(\bar{\theta}) = c \Delta \theta \left(q_{2,1}^p - \delta q_0 - \frac{\bar{\theta} + \underline{\theta}}{2} \right) - (c \Delta \theta)^2 \frac{\lambda \nu^2 \beta}{(1 - \nu)^2 [\beta b + c(1 + \lambda)]},$$

which is strictly positive for q_0 , $\underline{\theta}$ and $\Delta \theta$ close enough to zero. Since φ_δ is independent from q_0 , $\underline{\theta}$ and $\Delta \theta$, the result follows. \blacksquare

Proof of Proposition 10. For each $t = 1, 2$, let $\bar{q}_{2,t}$ and $\underline{q}_{2,t}$ be the functions mapping date $t - 1$ quality into date t quality following respectively a high and low quality shock at date $t - 1$ in the model with a sequence of agents. Note that $\bar{q}_{2,2} = \bar{\mathcal{Q}}_{1,2}$, while $\underline{q}_{2,2} = \underline{\mathcal{Q}}_{2,2}$ as defined by (84). When there is a single agent, it follows from (53) that $q_{2,2}(\underline{\theta}, \underline{\theta}) = \underline{\mathcal{Q}}_{6,2}(q_{2,1}(\bar{\theta}))$, where for each $q_1 \geq 0$, $\underline{\mathcal{Q}}_{6,2}(q_1)$ is implicitly defined by:

$$\beta S'(\underline{\mathcal{Q}}_{6,2}(q_1)) = (1 + \lambda) \psi'(\underline{\mathcal{Q}}_{6,2}(q_1) - \delta q_1 - \underline{\theta}) + \frac{\lambda \nu}{(1 - \nu)^2} \Phi_2(q_1, \underline{\mathcal{Q}}_{6,2}(q_1)). \quad (104)$$

We first compare $\bar{q}_{2,1}(q_0)$ and $q_{2,1}(\bar{\theta})$. It is helpful to define two functions $\bar{\Omega}$ and $\bar{\Xi}$ as follows:

$$\begin{aligned} \bar{\Omega}(q_1) &= (1 + \lambda) \psi'(q_1 - \delta q_0 - \bar{\theta}) - \nu \beta \delta (1 + \lambda) \psi'(\bar{\mathcal{Q}}_{1,2}(q_1) - \delta q_1 - \bar{\theta}) \\ &\quad - (1 - \nu) \beta \delta (1 + \lambda) \psi'(\underline{\mathcal{Q}}_{2,2}(q_1) - \delta q_1 - \underline{\theta}) - \lambda \nu \beta \delta \Phi_2(q_1, \underline{\mathcal{Q}}_{2,2}(q_1)), \end{aligned}$$

¹⁵Consider by contrast what happens when δ is close to zero. From (54), one has $\lim_{\delta \rightarrow 0} \bar{\beta}(\delta) = \infty$ and $\lim_{\delta \rightarrow 0} \bar{\beta}(\delta)\delta = 1$. It follows from (103) that $\lim_{\delta \rightarrow 0} \varphi_\delta(\bar{\beta}(\delta)) = 1 + \lambda$. Since φ_δ is decreasing, we obtain that, for δ close enough to zero, $\varphi_\delta(\beta) > 0$ for any $\beta \leq \bar{\beta}(\delta)$, and thus pooling over date 1 qualities cannot occur.

and:

$$\begin{aligned}\bar{\Xi}(q_1) &= (1 + \lambda)\psi'(q_1 - \delta q_0 - \bar{\theta}) - \nu\beta\delta(1 + \lambda)\psi'(\bar{\mathcal{Q}}_{1,2}(q_1) - \delta q_1 - \bar{\theta}) \\ &\quad - (1 - \nu)\beta\delta(1 + \lambda)\psi'(\underline{\mathcal{Q}}_{1,2}(q_1) - \delta q_1 - \underline{\theta}).\end{aligned}$$

Thus the difference $\bar{\Omega}(q_1) - \bar{\Xi}(q_1)$ is proportional to:

$$(1 + \lambda)\psi'(\underline{\mathcal{Q}}_{1,2}(q_1) - \delta q_1 - \underline{\theta}) - (1 + \lambda)\psi'(\underline{\mathcal{Q}}_{2,2}(q_1) - \delta q_1 - \underline{\theta}) - \frac{\lambda\nu}{1 - \nu}\Phi_2(q_1, \underline{\mathcal{Q}}_{2,2}(q_1)),$$

which is negative from the first-order conditions (37) and (84) that define $\underline{\mathcal{Q}}_{1,2}(q_1)$ and $\underline{\mathcal{Q}}_{2,2}(q_1)$ for a given value of q_1 . Now, from the Envelope Theorem, one has:

$$\beta S'(\bar{q}_{2,1}(q_0)) = \bar{\Omega}(\bar{q}_{2,1}(q_0)),$$

$$\beta S'(q_{2,1}(\bar{\theta})) = \bar{\Xi}(q_{2,1}(\bar{\theta})).$$

Since S is strictly concave, it follows that $\bar{q}_{2,1}(q_0) > q_{2,1}(\bar{\theta})$, as claimed.

We next compare $\underline{q}_{2,1}(q_0)$ and $q_{2,1}(\underline{\theta})$. It is helpful to define two functions $\underline{\Omega}$ and $\underline{\Xi}$ as follows:

$$\begin{aligned}\underline{\Omega}(q_1) &= (1 + \lambda)\psi'(q_1 - \delta q_0 - \underline{\theta}) - \nu\beta\delta(1 + \lambda)\psi'(\bar{\mathcal{Q}}_{1,2}(q_1) - \delta q_1 - \bar{\theta}) \\ &\quad - (1 - \nu)\beta\delta(1 + \lambda)\psi'(\underline{\mathcal{Q}}_{2,2}(q_1) - \delta q_1 - \underline{\theta}) - \lambda\nu\beta\delta\Phi_2(q_1, \underline{\mathcal{Q}}_{2,2}(q_1)) + \frac{\lambda\nu}{1 - \nu}\Phi_2(q_0, q_1),\end{aligned}$$

and:

$$\begin{aligned}\underline{\Xi}(q_1) &= (1 + \lambda)\psi'(q_1 - \delta q_0 - \underline{\theta}) - \nu\beta\delta(1 + \lambda)\psi'(\bar{\mathcal{Q}}_{1,2}(q_1) - \delta q_1 - \bar{\theta}) \\ &\quad - (1 - \nu)\beta\delta(1 + \lambda)\psi'(\underline{\mathcal{Q}}_{6,2}(q_1) - \delta q_1 - \underline{\theta}) + \frac{\lambda\nu}{1 - \nu}\Phi_2(q_0, q_1) + \frac{\lambda\nu\beta}{1 - \nu}\Phi_1(q_1, \underline{\mathcal{Q}}_{6,2}(q_1)).\end{aligned}$$

Since $\Phi_1 = -\delta\Phi_2$, the difference $\underline{\Omega}(q_1) - \underline{\Xi}(q_1)$ is proportional to:

$$\begin{aligned}(1 + \lambda)\psi'(\underline{\mathcal{Q}}_{6,2}(q_1) - \delta q_1 - \underline{\theta}) + \frac{\lambda\nu}{(1 - \nu)^2}\Phi_2(q_1, \underline{\mathcal{Q}}_{6,2}(q_1)) \\ - (1 + \lambda)\psi'(\underline{\mathcal{Q}}_{2,2}(q_1) - \delta q_1 - \underline{\theta}) - \frac{\lambda\nu}{1 - \nu}\Phi_2(q_1, \underline{\mathcal{Q}}_{2,2}(q_1)),\end{aligned}$$

which is positive from the first-order conditions (84) and (104) that define $\underline{\mathcal{Q}}_{2,2}(q_1)$ and $\underline{\mathcal{Q}}_{6,2}(q_1)$ for a given value of q_1 . Now, from the Envelope Theorem, one has:

$$\beta S'(\underline{q}_{2,1}(q_0)) = \underline{\Omega}(\underline{q}_{2,1}(q_0)),$$

$$\beta S'(q_{2,1}(\underline{\theta})) = \underline{\Xi}(q_{2,1}(\underline{\theta})).$$

Since S is strictly concave, it follows that $q_{2,1}(\underline{\theta}) > \underline{q}_{2,1}(q_0)$, as claimed. ■

Proof of Proposition 11. Since $\lambda\nu > \xi_2$ in (89)–(90), the optimal contract with a sequence of agents does not satisfy the necessary optimality conditions (89)–(90) for the optimal long-term contract with

a single agent. Thus, it suffices to show that it satisfies the analogues of constraints (39)–(42) when $\bar{\theta}$ is close enough to $\underline{\theta}$. The participation constraints are satisfied by construction. One thus only needs to check that the dynamic incentive compatibility constraints hold.

Given a high quality shock, the incentive compatibility constraint can be written as:

$$\Phi(q_0, \underline{q}_{2,1}(q_0)) \geq \nu\beta \left[\Phi(\underline{q}_{2,1}(q_0), \underline{q}_{2,2}(\underline{q}_{2,1}(q_0))) - \Phi(\bar{q}_{2,1}(q_0), \underline{q}_{2,2}(\bar{q}_{2,1}(q_0))) \right]. \quad (105)$$

To show that (105) holds when $\bar{\theta}$ is close enough to $\underline{\theta}$, we proceed as follows. From (18) and the convexity of ψ , one has:

$$\Phi(q_0, \underline{q}_{2,1}(q_0)) \geq \Delta\theta\psi'(\underline{q}_{2,1}(q_0) - \delta q_0 - \bar{\theta}). \quad (106)$$

Consider now the right-hand side of (105). The function $\underline{q}_{2,2}$ is differentiable. It is easy to check that its derivative is uniformly bounded above by δ , reflecting the fact that date 1 efforts are decreasing with respect to date 1 quality, in analogy with the basic model. In particular, $\Phi(q_1, \underline{q}_{2,2}(q_1))$ is a strictly decreasing function of q_1 . Since $\bar{q}_{2,1}(q_0) > \underline{q}_{2,1}(q_0)$, one has:

$$\begin{aligned} & \Phi(\underline{q}_{2,1}(q_0), \underline{q}_{2,2}(\underline{q}_{2,1}(q_0))) - \Phi(\bar{q}_{2,1}(q_0), \underline{q}_{2,2}(\bar{q}_{2,1}(q_0))) \\ & \leq \left[\bar{q}_{2,1}(q_0) - \underline{q}_{2,1}(q_0) \right] \sup_{q_1 \in [\underline{q}_{2,1}(q_0), \bar{q}_{2,1}(q_0)]} \left\{ \left| \frac{d}{dq_1} \Phi(q_1, \underline{q}_{2,2}(q_1)) \right| \right\}. \end{aligned} \quad (107)$$

Moreover, since $\underline{q}'_{2,2} \leq \delta$ and $\Phi_1 = -\delta\Phi_2$, one has, for each q_1 ,

$$\left| \frac{d}{dq_1} \Phi(q_1, \underline{q}_{2,2}(q_1)) \right| \leq 2\delta\Phi_2(q_1, \underline{q}_{2,2}(q_1)) \leq 2\delta\Delta\theta\psi''(\underline{q}_{2,2}(q_1) - \delta q_1 - \underline{\theta}), \quad (108)$$

where the second inequality follows from (18) and the convexity of ψ' . Thus, from (106)–(108),

$$\begin{aligned} & \Phi(\underline{q}_{2,1}(q_0), \underline{q}_{2,2}(\underline{q}_{2,1}(q_0))) - \Phi(\bar{q}_{2,1}(q_0), \underline{q}_{2,2}(\bar{q}_{2,1}(q_0))) \\ & \leq \left[\bar{q}_{2,1}(q_0) - \underline{q}_{2,1}(q_0) \right] \frac{2\delta \sup_{q_1 \in [\underline{q}_{2,1}(q_0), \bar{q}_{2,1}(q_0)]} \left\{ \psi''(\underline{q}_{2,2}(q_1) - \delta q_1 - \underline{\theta}) \right\}}{\psi'(\underline{q}_{2,1}(q_0) - \delta q_0 - \bar{\theta})} \Phi(q_0, \underline{q}_{2,1}(q_0)). \end{aligned} \quad (109)$$

It is easy to check that when $\bar{\theta}$ converges to $\underline{\theta}$ from above, the policy functions $\bar{q}_{2,1}$ and $\underline{q}_{2,1}$ converge pointwise to a function $q_{2,1}$, and the policy function $\underline{q}_{2,2}$ converges pointwise to a function $q_{2,2}$. The policy functions $q_{2,1}$ and $q_{2,2}$ correspond to the limit case where $\Delta\theta = 0$. In particular, the difference $\bar{q}_{2,1}(q_0) - \underline{q}_{2,1}(q_0)$ goes to zero as $\bar{\theta}$ converges to $\underline{\theta}$, while the ratio:

$$\frac{\psi'(\underline{q}_{2,1}(q_0) - \delta q_0 - \bar{\theta})}{2\delta \sup_{q_1 \in [\underline{q}_{2,1}(q_0), \bar{q}_{2,1}(q_0)]} \left\{ \psi''(\underline{q}_{2,2}(q_1) - \delta q_1 - \underline{\theta}) \right\}}$$

stays bounded away from zero. It thus follows from (109) that:

$$\Phi(\underline{q}_{2,1}(q_0), \underline{q}_{2,2}(\underline{q}_{2,1}(q_0))) - \Phi(\bar{q}_{2,1}(q_0), \underline{q}_{2,2}(\bar{q}_{2,1}(q_0))) = o(\Phi(q_0, \underline{q}_{2,1}(q_0)))$$

as $\bar{\theta}$ goes to $\underline{\theta}$, so that the incentive compatibility constraint (105) holds for $\bar{\theta}$ close enough to $\underline{\theta}$.

Given a low quality shock, the incentive compatibility constraint can be written as:

$$\Phi(q_0, \bar{q}_{2,1}(q_0)) - \Phi(q_0, \underline{q}_{2,1}(q_0)) \geq \nu\beta \left[\Phi(\bar{q}_{2,1}(q_0), \underline{q}_{2,2}(\bar{q}_{2,1}(q_0))) - \Phi(\underline{q}_{2,1}(q_0), \underline{q}_{2,2}(\underline{q}_{2,1}(q_0))) \right]. \quad (110)$$

Since $\bar{q}_{2,1}(q_0) > \underline{q}_{2,1}(q_0)$, the left-hand side of (110) is positive, while the right-hand side is negative as $\Phi(q_1, \underline{q}_{2,2}(q_1))$ is a strictly decreasing function of q_1 . Hence (110) is automatically satisfied no matter the parameter values. The result follows. \blacksquare

REFERENCES

- [1] Armstrong, M., S. Cowan, and J. Vickers (1994): *Regulatory Reform: Economic Analysis and British Experience*, Cambridge, MIT Press.
- [2] Baron, D., and D. Besanko (1984): “Regulation and Information in a Continuing Relationship,” *Information Economics and Policy*, 1, 447–470.
- [3] Baron, D., and R. Myerson (1982): “Regulating a Monopolist with Unknown Costs,” *Econometrica*, 50, 911–930.
- [4] D’Aspremont, C., and L.-A. Gérard-Varet (1979): “Incentives and Incomplete Information,” *Journal of Public Economics*, 11, 25–45.
- [5] De Fraja, G., and A. Iozzi (2004): “Bigger and Better: A Dynamic Regulatory Mechanism for Optimum Quality,” CEPR Discussion Paper No. 4502, London.
- [6] Freixas, X., R. Guesnerie, and J. Tirole (1985): “Planning under Incomplete Information and the Ratchet Effect,” *Review of Economic Studies*, 52, 173–191.
- [7] Gärtner, D. (2004): “Endogenizing Private Information: Incentive Contracts under Learning By Doing,” SOI Working Paper 0413, Universität Zürich.
- [8] Gaudet, G., P. Lasserre, and N. Van Long (1995): “Optimal Resource Royalties with Unknown and Temporally Independent Cost Structures,” *International Economic Review*, 36, 715–749.
- [9] Gaudet, G., P. Lasserre, and N. Van Long (1996): “Dynamic Incentive Contracts with Uncorrelated Private Information and History Dependent Outcomes,” *Japanese Economic Review*, 47, 321–334.
- [10] Laffont, J.-J., and D. Martimort (2002): *The Theory of Incentives: The Principal-Agent Model*, Princeton, Princeton University Press.
- [11] Laffont, J.-J., and J. Tirole (1986): “Using Cost Observation to Regulate Firms,” *Journal of Political Economy*, 94, 614–641.
- [12] Laffont, J.-J., and J. Tirole (1988): “The Dynamics of Incentive Contracts,” *Econometrica*, 56, 1153–1175.
- [13] Laffont, J.-J., and J. Tirole (1990): “Adverse Selection and Renegotiation in Procurement,” *Review of Economic Studies*, 75, 597–626.
- [14] Laffont, J.-J., and J. Tirole (1991): “Provision of Quality and Power of Incentive Schemes in Regulated Industries,” in J.J. Gabszewicz and A. Mas-Collel (Eds.), *Equilibrium Theory and Applications: Proceeding of the Sixth International in Economic Theory and Econometrics*, pp. 161–193, Cambridge, Cambridge University Press.
- [15] Lewis, T.R., and D.E.M. Sappington (1991): “Incentives for Monitoring Quality,” *RAND Journal of Economics*, 22, 370–384.
- [16] Lewis, T.R., and D.E.M. Sappington (1992): “Incentives for Conservation and Quality-Improvements in Public Utilities,” *American Economic Review*, 82, 1321–1340.

- [17] Lewis, T.R., and D.E.M. Sappington (1997): “Penalizing Success in Dynamic Incentive Contracts: No Good Deed Goes Unpunished?” *RAND Journal of Economics*, 28, 346–358.
- [18] Lewis, T.R., and H. Yildirim (2002): “Learning by Doing and Dynamic Regulation,” *RAND Journal of Economics*, 33, 22–36.
- [19] Rockafellar, R.T. (1970): *Convex Analysis*, Princeton, Princeton University Press.
- [20] OFGEM (2001): *Information and Incentives Project. Incentive Schemes: Final Proposals*, London, OFGEM.
- [21] OFWAT (2002): *Linking Service Levels to Prices*, London, OFWAT.
- [22] Rogerson, W. (1985): “Repeated Moral Hazard,” *Econometrica*, 53, 69–76.
- [23] Stokey, N.L., and R.E. Lucas, Jr., with E.C. Prescott (1989): *Recursive Methods in Economic Dynamics*, Cambridge, Harvard University Press.
- [24] Vogelsang, I., and J. Finsinger (1979): “A Regulatory Adjustment Process for Optimal Pricing by Multiproduct Monopoly Firms,” *Bell Journal of Economics*, 10, 157–171.