

Planning under Incomplete Information and the Ratchet Effect

XAVIER FREIXAS

Université des Sciences Sociales, Toulouse

ROGER GUESNERIE

CERAS and EHESS, Paris

and

JEAN TIROLE

Massachusetts Institute of Technology

Central planning of production is usually performed under asymmetric information which leads to use of an incentive scheme. As the planner revises the scheme over time to take into account information provided by the firm's performance, this induces firms to underproduce to avoid more demanding schedules in the future—the ratchet effect. This paper explores this phenomenon under the realistic assumption that the planner cannot commit himself to a revision procedure. We show that the ratchet effect exists, in the sense that the planner may choose a scheme which is suboptimal from a static viewpoint in order to induce revelation, with the marginal price of output exceeding its optimal static value.

1. INTRODUCTION

Central Planning of production is usually performed under asymmetric information. The firm in general has more information about its productive possibilities than the Central Planner (CP). This justifies the use of incentive schemes in which the CP does not directly fix activity levels. Casual observation of such existing schemes suggests that the time dimension is crucial. Typically the CP revises the incentive scheme over time to take into account the information provided by the firm's performance. Managers of centrally planned economies as well as economists have long recognized that this revision induces firms to underproduce in order to avoid demanding schemes in the future. This is the ratchet effect.

Although the study of incentive schemes in centrally planned economies has become popular in the late seventies, few researches have focused on the ratchet effect. Various authors have investigated how schemes may induce the firm to reveal its information before producing (Fan (1975), Bonin (1976), Gidin (1971)) or the effect of a given incentive scheme on the firm's behaviour (see Weitzman (1976) on the new Soviet Incentive Scheme). Also, Weitzman (1980) and Holmstrom (1982) gave examples of the ratchet effect assuming that the CP commits himself in advance to an intertemporal incentive scheme chosen in a suboptimal class. Although the exploration of the phenomenon under this latter assumption is useful, it must be stressed that most often, in practice, the CP does not commit himself to a revision procedure; moreover non-commitment turns out to be crucial to the ratchet effect.¹

This paper assumes non-commitment, which we believe to be more realistic. The inadequacy of conceptual tools made it difficult in the past to consider such an assumption. Recently basic research on dynamic games with incomplete information resulted in new ideas which look particularly relevant for a theoretical analysis of the problem.

The dynamic theory of the relationship CP-firm that we develop relies on these modern game-theoretic developments. Although our model is highly stylized, it is designed to incorporate the central positive and normative aspects of the problem, as they have been more or less explicitly described in previous theoretical literature.

The content of the paper can be summarized as follows: Section 2 introduces the model. The CP is assumed to have incomplete information about the productivity of the firm (we shall suppose that this productivity can take only two values). We explain why the CP does not decentralize the “first best” optimum by having the firm face the social value of its output. And as a starting point, we study the influence of incomplete information on the optimal incentive scheme in a static framework (Section 3). We show that the slope of the reward scheme is lower than the shadow price of the output and we explain why. Section 4 introduces dynamics and the possibility for the CP to revise the incentive scheme over time. To simplify we consider a two-period model where the essence of the phenomenon can be captured while leaving the argument reasonably simple. To be consistent with casual observation, we assume that the CP does not commit himself to an intertemporal incentive scheme at the start (reasons why commitment may not be possible are discussed in Section 6). The sequential choice of reward schemes by the CP and outputs by the firm is most naturally modelled as a game between the two partners. We show that this game has a unique welfare outcome. In the optimal incentive scheme, the firm may hide its information or reveal it. Even in the latter case the ratchet effect exists in the sense that the CP may choose a scheme which is suboptimal from a static point of view in order to induce revelation. We also show that we should expect the marginal price of output in the reward scheme to exceed its optimal static marginal price. This is what intuition would suggest: the CP should be more “generous” in a dynamic context in order to obtain information. These results are derived in Section 5.

This work is not only relevant for Central Planning as most types of principal/agent relationships under asymmetric information (delegation within a firm, taxation, etc.) is not one-shot. As the principal then learns over time, some form of ratchet effect necessarily obtains. Section 6 mentions these applications and stresses analogies. But much work remains to be done to formalize them.

2. THE MODEL

2.1. *The set up*

(a) Technology: A firm produces one good whose “social value” per unit p is exogenously given. y denotes the firm’s output. The cost to the firm² of producing y units of the good is a function $\psi(y, \theta)$ of y and of a productivity parameter θ ; θ stands for the ability of workers and managers, true capacity of the firm, etc., and can only be observed by the firm. The CP has some probability distribution $f(\theta)$ about it, which is common knowledge. In this paper, the probability distribution will be confined to two values $\underline{\theta}$ and $\bar{\theta}$ ($\underline{\theta} < \bar{\theta}$) with respective probabilities ν and $(1 - \nu)$.³ The cost function will then be written $\underline{\psi}(y)$ or $\bar{\psi}(y)$.

The production cost can be thought of as embodying the managers and workers’ effort and cannot be observed by the CP. We shall make the following assumptions on

ψ : the cost increases with and is convex in output: $\psi_y > 0$, $\psi_{yy} > 0$; and it decreases with productivity: $\psi_\theta < 0$. Furthermore the marginal cost of production decreases with productivity: $\psi_{y\theta} < 0$.

(b) Incentive scheme: The CP rewards the firm with an incentive scheme $R(y)$, and derives a utility py from output y .

Standard first best analysis would suggest to reward the firm according to $R(y) = py$; this conclusion will be incorrect here where it is assumed that there is a social cost of $\$ \lambda$ to transferring $\$1$ to the firm. A possible interpretation relates λ to the social cost of raising funds. This is a standard assumption, precise justifications of which have been given when, for example, funds are raised through distortionary taxation. Justifications of a similar kind could be given in the context of Central Planning.

We formalize the firm's objective function π^F as the difference between its revenue and its cost:

$$\pi^F(y, \theta) = R(y) - \psi(y, \theta). \quad (1)$$

It is convenient to split the CP's objective function into two parts. First, the CP is concerned with the "rest of the world" welfare, i.e. with the difference between the social value of output and the social cost of transferring $R(y)$ to the firm. The first part will be referred to as the CP's "social concern" utility function and denoted π^{CP} :

$$\pi^{CP}(y) = py - (1 + \lambda)R(y). \quad (2)$$

The CP also takes into account the firm's welfare π^F .

Thus aggregate welfare W can be written:

$$W(y, \theta) = \pi^F + \pi^{CP} = py - \psi(y, \theta) - \lambda R(y). \quad (3)$$

At this stage two remarks are in order. First, another interpretation of the formula might be given where λ would account for the distortion associated with the firm's distribution of its revenue rather than a cost of raising funds. The social value of the good and the internal cost of production would be weighted equally while excess income transfer R would be penalized. This interpretation may be particularly relevant in the context of Central Planning which, particularly in its Soviet-type version, stems from a distrust of allocational and distributional effects of market mechanisms. The surplus R if distributed may generate uncontrolled inequalities or if it is invested lead to relative capital accumulation considered as inappropriate. We have not formalized these ideas. Second, note that if we define

$$s \stackrel{\text{def}}{=} \frac{p}{1 + \lambda},$$

an alternative formalization of social welfare obtains

$$W = (1 + \lambda)(sy - R) + (R - \psi). \quad (3')$$

This is a more traditional form of objectives in a social planning context (see Baron-Myerson (1982)) where total welfare appears as a weighted sum with coefficients $(1, 1/(1 + \lambda))$ of a pseudo planner's welfare $(sy - R)$ and of the firm's welfare $(R - \psi)$. Although this form (3') is equivalent to (3) and consistent with the former interpretations,⁴ it may also suggest a different interpretation putting the emphasis on $1/(1 + \lambda)$ as the firm's social weight. In the following we will stick to the first interpretation (λ as a cost of

raising funds) but the reader is invited to see how comments should be modified in view of the other two interpretations.

Coming back to (3) it is useful to represent indifference curves for π^F and π^{CP} in the revenue-output space. This is done in Figures 1 and 2. We shall assume that the CP is risk neutral. We also assume that the firm can guarantee itself some minimum utility, for example as suggested above by producing at a minimum acceptable level. Without loss of generality we shall assume that the minimum utility level is associated with a zero production and is equal to 0. Thus the firm maximizes $\pi^F(y, \theta)$ subject to the individual rationality constraint $\pi^F(y, \theta) \geq 0$. If $\max_y \pi^F(y, \theta) < 0$ the firm's output is zero.

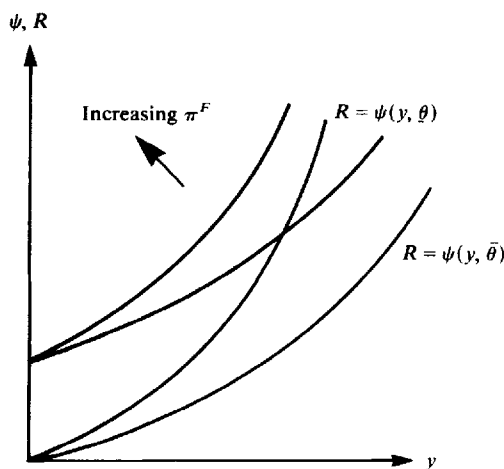


FIGURE 1
 $\theta < \bar{\theta}$

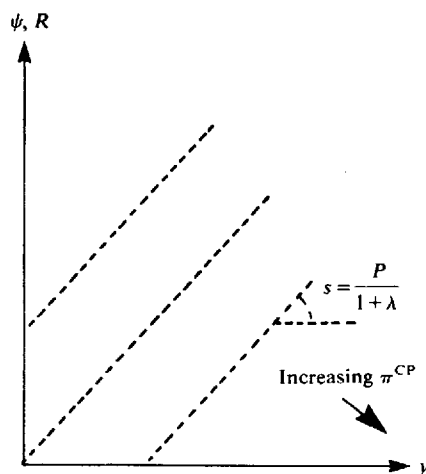


FIGURE 2

Note that $s = p/(1 + \lambda)$ is the slope of the CP's social concern indifference curves. For reasons that will be given shortly, s will be called the shadow price of output.

This paper is concerned with the class of linear schemes, i.e. schemes of the type: $R(y) = a + by$. A linear scheme can also be written as: $R(y) = b(y - y_0)$ where $y_0 = -a/b$ can be interpreted as a quota, and b as a bonus per unit of output above the quota. Linear schemes are mainly used for expositional purposes. They allow focusing on a single marginal incentive (the slope of the reward function). Also they are commonly used.

2.2. Static incentive scheme under complete information

In order to motivate the following, we first consider the simple case of complete information about θ . Let us derive the optimal linear scheme (which in this case is also the optimal non-linear scheme). Let $y(b, \theta)$ denote the output that maximizes $\pi^F(y, \theta)$. For any given bonus b , it is desirable from the viewpoint of aggregate welfare to choose the lump-sum transfer a so as to make the firm's individual rationality constraint binding: lowering it by $\delta a > 0$, other things being equal, does not affect the firm's production decision and increases aggregate welfare W by $\delta W = \lambda \delta a$. Hence $a = \psi(y(b, \theta), \theta) - by(b, \theta)$. What bonus will the CP choose? b must be such that the (marginal) social price of output net of transfer costs equals the marginal value of output for the firm (b). The former is nothing but the difference between the social value of one unit of output p and the shadow cost of transferring $\$b$, λb . Thus

$$p - \lambda b = b \quad (4)$$

or

$$b = \frac{p}{1 + \lambda} = s. \quad (5)$$

The formal proof of this latter point is as simple as the heuristics: given that we have shown that the individual rationality constraint of the firm has to be binding at the optimum, the CP's problem reduces to the maximization of $\{sy - R\}$ over the indifference curve of the firm $R - \psi(y, \theta) = 0$, i.e. to the unconstrained maximization of $sy - \psi(y, \theta)$. The optimum is uniquely determined by $\psi_y = s$ and is decentralized through a linear bonus function with $b = s$.

It follows:

Proposition 1. *The optimal bonus under complete information is equal to the shadow price of output s .*

3. STATIC INCENTIVE SCHEMES UNDER INCOMPLETE INFORMATION

We analyse the optimal linear scheme when the relationship between the firm and the CP is short-run and there are two types of firms. We assume that the CP can choose only in the class of linear schemes and that the firm has two potential productivities θ and $\bar{\theta}$ with probability ν and $(1 - \nu)$. We shall also assume that the proportion of low productivity firms is not "too low" relative to the productivity differential, so that the CP never wants to induce such a firm not to operate (otherwise the problem is trivial). Let

$y(b)$ and $\bar{y}(b)$ the firm's utility-maximizing outputs of both types of firms when the bonus is b . Note that $\bar{y}(b) > y(b)$. Given cost functions $\underline{\psi}$ and $\bar{\psi}$ for the two types, the CP's optimization problem can be written:

$$W = \nu[p\underline{y}(b) - \underline{\psi}(\underline{y}(b)) - \lambda(a + b\underline{y}(b))] + (1 - \nu)[p\bar{y}(b) - \bar{\psi}(\bar{y}(b)) - \lambda(a + b\bar{y}(b))]. \quad (6)$$

Clearly, the lump-sum transfer must make the low-productivity firm individual rationality constraint binding (the argument of Section 2.2 is transposed here):

$$a = \underline{\psi}(\underline{y}(b)) - b\underline{y}(b). \quad (7)$$

Substituting (7) into (6) and differentiating (6) gives:

$$\frac{dW}{db} = (1 + \lambda) \left[(s - b) \left[\nu \frac{dy}{db} + (1 - \nu) \frac{d\bar{y}}{db} \right] - \frac{\lambda(1 - \nu)}{1 + \lambda} [\bar{y}(b) - \underline{y}(b)] \right]. \quad (8a)$$

The first-order condition obtains when:

$$s - b = \frac{\lambda(1 - \nu)}{1 + \lambda} \left[\frac{\bar{y}(b) - \underline{y}(b)}{\nu(dy/db)(b) + (1 - \nu)(d\bar{y}/db)(b)} \right]. \quad (8b)$$

Since output grows with the productivity and the bonus, we conclude from (8a, b):

Proposition 2. *Under incomplete information about the productivity of the firm,*

- (i) *W is a decreasing function of b for $b \geq s$. And therefore:*
- (ii) *the optimal bonus is strictly smaller than the shadow price of output.*

We can give a more intuitive proof of part (ii) of the proposition as follows:

Assume that for the optimal scheme $b > s$. Let $\bar{B} = (\bar{R}, \bar{y})$, $B = (R, y)$ be the choices of firms $\bar{\theta}$ and θ . Now assume that the CP offers the linear scheme with slope s that leaves the low productivity firms with a zero surplus. If the firm is of type θ , it is indifferent to this change of scheme; since the CP reaches a better "social concern" indifference curve, ex post aggregate welfare, in case the firm has a low productivity, increases. Assume now that the firm has type $\bar{\theta}$. The firm reacts to the new scheme of slope s by choosing $B' = (R', y')$. This change can be decomposed into two steps.

A change along the indifference curve through the original point $\bar{B}$. This corresponds to a "substitution effect". For the same reason as above, the $\bar{\theta}$ firm being indifferent to this move and the CP's social concern utility increasing, aggregate welfare increases.

A change from the upper indifference curve through $\bar{B}$ to the lower indifference curve through B' . This corresponds to an "income effect" and given our separability assumption for the firm's objective function it only involves an income transfer from the firm to the society and according to (3) this move is beneficial from the point of view of social welfare.

Hence if the firm is of type $\bar{\theta}$, aggregate welfare increases when moving from a bonus scheme $b > s$ to a bonus scheme with $b = s$. Hence $b > s$ cannot be optimal.⁵ The argument conveys the following intuitive idea: raising the bonus above the shadow price of output creates a productive misallocation and increases the high productivity firm's surplus and both effects are socially undesirable.

It will be very useful for the dynamic study to consider what happens when the proportion of high and low productivity firms changes. Let $\bar{\pi}^F(\nu)$ denote the high

productivity firm's profit when the proportion of low productivity firms is ν and the CP chooses the optimal linear scheme $(a(\nu), b(\nu))$ (so as to keep both types active—remember that the profit of the low productivity firm π^F is equal to 0).

We can show:

Proposition 3. (i) $b(\nu)$ increases with ν . (ii) $\bar{\pi}^F(\nu)$ is continuous and increasing in ν .

Proof. Let us prove that the optimal bonus grows with the proportion of low productivity firms.

Let ν' be greater than ν and suppose that $b(\nu') < b(\nu)$. Decompose the “population” $\{\nu', 1 - \nu'\}$ of low and high productivity firms into two subpopulations.

The first one has all high productivity firms and $\nu'' = (\nu/(1 - \nu))(1 - \nu')$ low productivity firms.

The second has $(\nu' - \nu'')$ low productivity firms (with $\nu' - \nu'' = (1/1 - \nu)(\nu' - \nu) > 0$). For the first population, which has the same relative composition as the economy $(\nu, 1 - \nu)$, $b(\nu)$ is better from the welfare viewpoint than $b(\nu')$.

For the second population of low productivity firms, welfare increases with b , when $b < s$; hence $b(\nu)$ is also better than $b(\nu')$.

These remarks contradict the fact that $b(\nu') < b(\nu)$.

To show that $\bar{\pi}^F$ is increasing in ν , it is enough to remark that from the firm first order conditions $d\bar{\pi}^F/db = \bar{y}(b) - \underline{y}(b) < 0$. ||

Proposition 3 says that the optimal bonus and the profit of high productivity firms increase with the probability of a low productivity firm. This latter fact is indeed the basic reason why the firm would like to pretend it has a productivity lower than its true one.

Of course an implicit assumption in Proposition 3 is that the CP wants both types of firm to operate, which is not optimal if the probability of a low productivity firm is low (in the latter case only the high productivity firm operates and $b = s$).

4. LONG-RUN RELATIONSHIP AND THE RATCHET EFFECT: EXISTENCE OF EQUILIBRIUM

4.1 The dynamic model. We now want to allow for the possibility that the CP learns over time about the firm's constant productivity. Our framework is reminiscent of that employed by Fudenberg and Tirole (1983) in their study of sequential bargaining under incomplete information. There are two types of firms, θ and $\bar{\theta}$ ($\theta < \bar{\theta}$) with prior probabilities ν_1 and $(1 - \nu_1)$. There are two periods, $t = 1, 2$. The social price of output p remains constant over time. We assume that the CP chooses schemes in the linear class. The timing is the following: At the start of the first period the CP chooses a scheme (a_1, b_1) ; the firm reacts to this scheme by producing y_1 at cost $\psi(y_1, \theta)$ and is rewarded $(a_1 + b_1 y_1)$. Then at the start of the second period the CP chooses a new scheme (a_2, b_2) , and the firm chooses a second-period output y_2 . Let δ be the common discount factor.⁶ With straightforward notation the discounted profit and aggregate social welfare are:

$$\pi^F = \pi_1^F + \delta \pi_2^F = \{a_1 + b_1 y_1 - \psi(y_1, \theta)\} + \delta \{a_2 + b_2 y_2 - \psi(y_2, \theta)\} \quad (9)$$

$$W = W_1 + \delta W_2 = (\pi_1^F + \delta \pi_2^F) + (\pi_1^{CP} + \delta \pi_2^{CP})$$

$$= \{p y_1 - \psi(y_1, \theta) - \lambda(a_1 + b_1 y_1)\} + \delta \{p y_2 - \psi(y_2, \theta) - \lambda(a_2 + b_2 y_2)\}. \quad (10)$$

The situation depicted above is a dynamic game between the firm and the CP. The CP strategy is a sequence of schemes (a_1, b_1) and $(a_2, b_2)_{(a_1, b_1, y_1)}$ where the second period scheme is a function of his information at that date, i.e. that the firm has chosen to produce y_1 under the incentive scheme (a_1, b_1) ⁷ (we insist on the fact that the CP does not commit himself in advance to a given revision process). The firm's strategy is a sequence of outputs $y_1(a_1, b_1, \theta)$ and $y_2(a_1, b_1, y_1, a_2, b_2, \theta)$, where the output at a given period is a function of the firm's information at that date (clearly, y_2 will depend only on (a_2, b_2) and θ in equilibrium). Mixed strategies are allowed. To introduce the equilibrium notion, we must also describe the CP's beliefs about the productivity of the firm at the beginning of the second period. Let $\nu_2(a_1, b_1, y_1)$ denote the CP's posterior probability that the firm has a low productivity. The equilibrium notion—Perfect Bayesian, or Sequential Equilibrium in the terminology of Kreps–Wilson (1982) to which we refer for a careful analysis⁸ requires that strategies and beliefs be consistent; i.e. that the strategies be optimal given the beliefs and that the beliefs be derived from strategies using Bayes rule.

A Perfect Bayesian Equilibrium is a set of (possibly mixed) strategies $\{(a_1, b_1), y_1, (a_2, b_2), y_2\}$ and of beliefs $\{\nu_2\}$ satisfying (we shall often omit the arguments of strategies and beliefs for clarity):

- P1. $\forall \theta, y_2$ maximizes π^F .
- P2. (a_2, b_2) maximizes the expectation of W_2 , given the CP's beliefs $\nu_2(a_1, b_1, y_1)$ and the firm's second-period strategy y_2 .
- P3. $\forall \theta, y_1$ maximizes π^F given the second-period strategies.
- P4. (a_1, b_1) maximizes the expectation of W , given the firm's and CP's subsequent strategies.
- B. $\nu_2(a_1, b_1, y_1)$ is Bayes-consistent with the prior probability ν_1 and the firm's first-period strategy $y_1(a_1, b_1, \theta)$ and observed action.

P1 to P4 only describe the principles of Kuhn's algorithm for finding a dynamic perfect Nash equilibrium when the probabilities of reaching the nodes of the tree are exogenously given. B endogenizes these probabilities in conformity with Bayes' rule.

4.1 *Resolution.* Note that a *given* $\{a_1, b_1\}$ defines a game between the CP and the firm which is a *subform* of the initial game. Perfectness requires that equilibrium strategies in the larger initial game induce an equilibrium for this game, which we call a continuation equilibrium. More precisely, a continuation equilibrium is a set of strategies and beliefs satisfying P1, P2, P3 and B for an exogenously fixed $\{a_1, b_1\}$.

We will say that a continuation equilibrium is unique if the observed equilibrium actions are unique.

In this section, we focus on existence and uniqueness of a continuation equilibrium. We shall denote by $\underline{y}(b_1)$ and $\bar{y}(b_1)$ the outputs that maximize the *one-period* surplus of the low and high productivity firms when the bonus is b_1 and by $y_1(b_1)$ and $\bar{y}_1(b_1)$ the *first-period* outputs when the firms take into account the effect of their choice on the second-period incentive schemes; under the assumption that the firms operate, these outputs do not depend on a_1 , as we shall see below.

We first notice that from P1 and P2 the second period actions of the firm and the CP are the same as in the one-period game with prior ν_2 . They have been analysed in Section 3. Let us now consider the first-period actions. We assume that both types of firms operate. For the moment, the statement “for any bonus b_1 ” should be qualified by “assuming that a_1 is high enough so that both firms operate”.

We first derive two *necessary* conditions for a continuation equilibrium.

Lemma 1. $\forall b_1: \underline{y}_1(b_1) = \underline{y}(b_1)$.

In other words, the optimal first period action of the low productivity firm, in the dynamic context and in equilibrium, can only be its optimal one-period action.

Proof of Lemma 1. From Section 3, we know that, whatever his beliefs ν_2 the CP will not allow the low productivity firm to make a strictly positive surplus in the second period. Thus this type of firm is only concerned with its first-period surplus π_1^F , and chooses $\underline{y}_1(b_1)$ so as to maximize it. $\parallel$

Lemma 2. $\forall b_1: \bar{y}_1(b_1) \in \{\bar{y}(b_1), \underline{y}(b_1)\}$.

Proof of Lemma 2. If output y_1 belongs to the support of the high productivity firm strategy, from Bayes rule and Lemma 1, the CP's posterior beliefs ν_2 must be equal to zero unless $y_1 = \underline{y}(b_1)$. Assume $y_1 \neq \underline{y}(b_1)$. Since the CP knows that the firm has a high productivity with certainty, the latter does not obtain a positive second-period surplus. Since, whatever the CP's beliefs, the high productivity firm can guarantee itself a non-negative profit in the second period, switching from any y_1 in the support of its strategy and different from both $\bar{y}(b_1)$ and $\underline{y}(b_1)$ to $\bar{y}(b_1)$ would give such a firm a strictly higher profit. This shows that the support can only consist of two values, as asserted in Lemma 2. $\parallel$

From Lemmas 1 and 2, it follows that there can only be three kinds of continuation equilibria:

(1) *Pooling equilibrium*: Both types of firm choose $\underline{y}(b_1)$. The CP's posterior belief is then: $\nu_2(a_1, b_1, \underline{y}(b_1)) = \nu_1$.

(2) *Separating equilibrium*: The low productivity firm produces $\underline{y}(b_1)$ and the high productivity firm $\bar{y}(b_1)$. Then: $\nu_2(a_1, b_1, \underline{y}(b_1)) = 1$ and $\nu_2(a_1, b_1, \bar{y}(b_1)) = 0$.

(3) *Semi-separating equilibrium*: The low productivity firm produces $\underline{y}(b_1)$ and the high productivity firm randomizes between $\underline{y}(b_1)$ and $\bar{y}(b_1)$. Then: $\nu_2(a_1, b_1, \underline{y}(b_1)) \in (\nu_1, 1)$ and $\nu(a_1, b_1, \bar{y}(b_1)) = 0$.

The determination of the equilibrium type relies on the consideration of the (first period) cost of concealing productivity for a high productivity firm when first-period bonus is b_1 . It is denoted $\Delta(b_1)$.

$$\Delta(b_1) \equiv b_1[\bar{y}(b_1) - \underline{y}(b_1)] - [\bar{\psi}(\bar{y}(b_1)) - \bar{\psi}(\underline{y}(b_1))]. \quad (11)$$

Note that $\Delta(0) = 0$ and that $\Delta(b_1) > 0$ for $b_1 > 0$.

Clearly, this short-run cost of concealment has to be compared with the long-run gain. The latter relates with $\delta\bar{\pi}^F(\nu)$ which, in the notation of Section 3, will be the discounted surplus of the second period high productivity firm when the CP's *a posteriori* beliefs are ν .

The interest of the above comparison is precisely confirmed by the next proposition.

Proposition 4. $\forall b_1$, there exists a unique continuation equilibrium. If b_1 is such that: $\Delta(b_1) \leq \delta\bar{\pi}^F(\nu_1)$ the equilibrium is pooling, $\delta\bar{\pi}^F(1) > \Delta(b_1) > \delta\bar{\pi}^F(\nu_1)$ the equilibrium is semi-separating, $\Delta(b_1) \geq \delta\bar{\pi}^F(1)$ the equilibrium is separating.

Proof. Consider successively the three possible types of equilibria:

Pooling equilibrium. In a pooling equilibrium, the long-run loss of playing $\bar{y}(b_1)$ instead of $y(b_1)$ is $\delta\bar{\pi}^F(\nu_1)$ since the latter is the high productivity firm's discounted second-period surplus when the CP does not acquire information in the first period. Thus a necessary condition for the existence of a pooling equilibrium is:

$$\delta\bar{\pi}^F(\nu_1) \leq \Delta(b_1). \quad (12)$$

Conversely, let us show that if (12) holds, there exists a pooling equilibrium. For that, we must complete equilibrium actions and beliefs with out of equilibrium ones. Since the second-period strategies of the firm and the CP are uniquely defined in Section 3 for a given ν_2 , we just have to describe the CP's beliefs for out-of-equilibrium outputs. There are several ways to choose these beliefs so that the firms actually want to pool at $y(b_1)$. The simplest way to do so is to take "optimistic" (for the CP) beliefs: $\forall y_1 \neq y(b_1): \nu_2 = 0$, i.e. the CP believes that the firm which chooses $y_1 \neq y(b_1)$ has a high productivity; the corresponding second-period scheme has slope s and lump-sum transfer a_2 such that the high productivity firm makes a zero surplus ($a_2 = \bar{\psi}(\bar{y}(s)) - s\bar{y}(s)$). Now it is clear that the efficient firm does not want to reveal its productivity, since the short-run gain to doing so does not exceed $\Delta(b_1)$ and the long-run loss equals $\delta\bar{\pi}^F(\nu_1)$.

We conclude that there exists a pooling equilibrium if and only if (12) is satisfied, and that this equilibrium is unique. The firm, whatever its type, produces $y(b_1)$ in the first period; in the second period the CP chooses a bonus $b_2 = b(\nu_1)$ and a lump-sum transfer such that the low productivity firm has a zero surplus. The firm then produces its optimal static output given b_2 .

Separating equilibrium. In a separating equilibrium, the high productivity firm produces $\bar{y}(b_1)$ in the first period and makes a zero surplus in the second. If instead it decided to produce $y(b_1)$ and pool with the low productivity firm, it would lose $\Delta(b_1)$ in the first period and gain $\delta\bar{\pi}^F(1)$, as the CP is convinced that he faces firm θ when he observes $y(b_1)$. Thus a necessary condition for the existence of a separating equilibrium is:

$$\delta\bar{\pi}^F(1) \leq \Delta(b_1). \quad (13)$$

Conversely, it is easy to show that if (13) holds, there exists a unique separating equilibrium path that can be supported by out-of-equilibrium beliefs and strategies (take optimistic beliefs for output that differs from $y(b_1)$ and $\bar{y}(b_1)$). Along this equilibrium path, the firm, whatever its type, produces its static optimum in the first period, the CP then has complete information and chooses $b_2 = s$ and a_2 so as to extract the whole surplus from the firm. The firm then produces its static optimum for bonus s .

Semi-separating equilibrium. In a semi-separating equilibrium, the efficient firm is indifferent between revealing its productivity by producing $\bar{y}(b_1)$, and pooling with the low cost firm by producing $y(b_1)$. The short-run loss of pooling is $\Delta(b_1)$. The long-run gain is $\delta\bar{\pi}^F(\nu_2)$ where ν_2 is the CP's second period belief when he observes $y(b_1)$. Note that ν_2 must belong to $(\nu_1, 1)$ since the low-productivity firm produces $y(b_1)$ with probability one. Thus a necessary condition for existence of a semi-separating equilibrium is that there exists ν_2 in $(\nu_1, 1)$ such that:

$$\delta\bar{\pi}^F(\nu_2) = \Delta(b_1). \quad (14)$$

From Proposition 3, $\bar{\pi}^F(\nu_2)$ is continuous and strictly increasing in ν_2 , thus b_1 must satisfy:

$$\delta\bar{\pi}^F(\nu_1) < \Delta(b_1) < \delta\bar{\pi}^F(1). \quad (15)$$

Conversely, assume that (15) holds. Again from Proposition 3, there exists a unique ν_2 in $(\nu_1, 1)$ such that (14) is satisfied. And therefore there exists a unique real number x_1 in $(0, 1)$ such that, if the efficient firm produces $\bar{y}(b_1)$ with probability x_1 and $\underline{y}(b_1)$ with probability $(1-x_1)$, the CP's posterior belief when he observes $\underline{y}(b_1)$, $\nu_2 = \nu_1/(\nu_1 + (1-x_1)(1-\nu_1))$, satisfies (14). Again it is easy to construct supporting beliefs and strategies out-of-equilibrium (take optimistic beliefs for outputs that differ from $\bar{y}(b_1)$ and $\underline{y}(b_1)$).⁹

To conclude, if b_1 satisfies (15), there exists a unique semi-separating equilibrium path. The low productivity firm produces $\underline{y}(b_1)$ and the high productivity firm randomizes between $\underline{y}(b_1)$ and $\bar{y}(b_1)$. When observing the non-fully revealing output $\underline{y}(b_1)$, the CP chooses a bonus $b_2 = b(\nu_2)$ (where ν_2 is defined by (14)) and a_2 so as to extract the low productivity firm's surplus. When observing $\bar{y}(b_1)$, he chooses bonus s and a_2 so as to extract the high productivity firm's surplus.

To finish the proof of Proposition 4, we observe that for a given b_1 , one of the three mutually exclusive conditions (12), (13) and (15) must hold. $\parallel$

The continuation equilibrium depends on the first-period cost of concealing productivity $\Delta(b_1)$, that is represented in Figure 3:

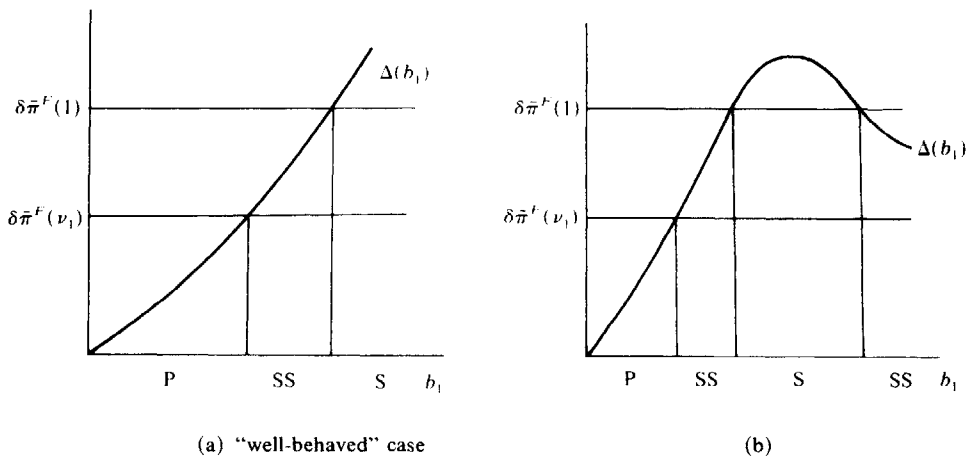


FIGURE 3

Regions for continuation equilibrium (P = pooling, SS = semi-separating, S = separating)

We shall say that $\Delta(b_1)$ is "well-behaved" if it increases with b_1 (Figure 3(a)); this means that when the bonus increases, the high productivity firm must incur higher and higher losses in order to mimic the low productivity firm. Section 5(c) treats in detail the example of a quadratic cost function. It is shown there that $\Delta(b_1)$ is increasing and convex.¹⁰ For a "well-behaved" function, a higher bonus leads to more separation. However $\Delta(b_1)$ need not be increasing in b_1 . Its shape depends on third derivatives of the cost function, on which we have little information. This makes the study of the optimal first-period incentive scheme analytically complicated. Note that for δ small enough the continuation equilibrium separates the firms.

5. LONG-RUN RELATIONSHIP AND THE RATCHET EFFECT: THE OPTIMAL DYNAMIC POLICY

We now study the CP's optimal first-period decision. We will throughout assume that the CP chooses a scheme such that even a low productivity firm operates in the first period. One could also consider "strongly selective" schemes in which only the high productivity firm produces (a few interesting things can be said in a dynamic context about the desirability of such schemes. In particular it may be optimal to leave the inefficient firm out even if the CP wants both types of firm to operate in a static context: the elimination of "lame ducks" may prove useful to induce revelation).

5. *The CP's first-period problem: key facts*

This is a technical section and the proofs of the five lemmas (Lemmas 3 to 7) which are presented can be skipped in a first reading; however their understanding is needed for a comprehensive analysis of the ratchet phenomenon. Results and comments are postponed to the next section.

The existence of a unique continuation equilibrium for any given b_1 makes the CP's first-period maximization problem well-defined. We were not able to prove that in general the solution to this problem is unique; however, a multiplicity of first-period optimal schemes is not too severe a problem, since, once the CP has chosen (a_1, b_1) , the continuation equilibrium is unique (there is no problem of coordination between equilibria). The welfare outcome W is also unique.

One important issue, on which one would expect clarifications from the analysis, concerns the relative magnitude of the optimal first-period bonus in a dynamic context and the static optimal bonus. One is tempted to believe *a priori* that the Planner should be more "generous" in a dynamic context in order to obtain information. In fact, the CP's problem is rather complex and in order to understand it, we have to single out a certain number of key facts which are gathered in the next lemmas.

A preliminary lemma asserts that more information is indeed desirable to the Centre.

Lemma 3. *The expectation of the optimal second-period aggregate social welfare W_2 , for a given first-period scheme, increases with the proportion of revealing high productivity firms x_1 (or decreases with the a posteriori probability that a pooling firm is of high productivity).*

Proof. The argument goes as follows: For the firms that switch from pooling in the reference situation to revealing, second-period social welfare is increased since the second-period scheme is now the full information scheme ($a_2 = \bar{\psi}(\bar{y}(s)) - s\bar{y}(s)$, $b_2 = s$). For the pooling firms, the CP can always duplicate his previous second-period incomplete information scheme in the reference situation (which is now sub-optimal) and leave social welfare unaffected. ||

We now consider the effect of the first-period choice between pooling, semi-separating and separating on first-period aggregate welfare.

First, let us assume that firms are pooling. First-period aggregate welfare is:

$$W^P(b_1) = py(b_1) - \lambda(\underline{a}(b_1) + b_1 \underline{y}(b_1)) - \nu_1 \underline{\psi}(\underline{y}(b_1)) - (1 - \nu_1) \bar{\psi}(\underline{y}(b_1))$$

where $\underline{a}(b_1)$ is given by $\underline{a}(b_1) + b_1 \underline{y}(b_1) - \underline{\psi}(\underline{y}(b_1)) = 0$.

We prove:

Lemma 4. *For $b_1 \leq s$, $W^P(b_1)$ increases with b_1 .*

Proof of Lemma 4.

$$\frac{dW^P}{db_1} = [p - \lambda b_1 - \nu_1 \psi' - (1 - \nu_1) \bar{\psi}'] \frac{dy}{db_1} - \lambda \left[y(b_1) + \frac{da}{db_1} \right].$$

As $\bar{\psi}'(y(b_1)) < \psi'(y(b_1)) = b_1$, it follows from the computation of da/db_1 that:

$$\frac{dW^P}{db_1} > [\bar{p} - (1 + \lambda)b_1] \frac{dy}{db_1}. \quad \parallel$$

Define

$$\begin{aligned} \tilde{W}_1(b_1, x_1) = & [\nu_1 + (1 - \nu_1)(1 - x_1)] W^P(b_1) \\ & + [x_1(1 - \nu_1)] [p\bar{y}(b_1) - \lambda(\underline{a}(b_1) + b_1\bar{y}(b_1)) - \bar{\psi}'(\bar{y}(b_1))]. \end{aligned}$$

$\tilde{W}_1(b_1, x_1)$ is the aggregate first period welfare when the first-period bonus is b_1 and the proportion of revealing high productivity firms is x_1 .

Lemma 5. *The first-period aggregate welfare $\tilde{W}_1$ increases with the proportion x_1 of revealing high productivity firms, for any $b_1 \leq s$. (Note that Lemmas 4 and 5 hold independently of the value of the parameter ν_1 .)*

Proof of Lemma 5. Assume that for $b_1 \leq s$, one high productivity firm switches from pooling to separating. This move increases the first-period profit of the firm. But it also increases the CP's "social concern" utility for the first period: social concern indifference curves are straight lines of slope s (see Figure 2) and separation moves the firm's bundle to the right along a line of slope b_1 , i.e. on a higher social concern indifference curve when $b_1 < s$. $\parallel$

The following corollary is a straightforward implication of Lemmas 3 and 5.

Corollary 1. *For any $b_1 \leq s$, intertemporal aggregate social welfare increases with the number of revealing high productivity firms.*

We now show that if $b_1 > s$, $\tilde{W}_1(s, 1) > \tilde{W}_1(b_1, x_1)$, $x_1 \neq 1$.

Lemma 6. *A separating equilibrium with bonus s would be superior from the first-period aggregate welfare point of view to a semi-separating (or pooling) equilibrium with $b_1 > s$.*

Proof. Consider the reference situation, a semi-separating or pooling equilibrium with $b_1 > s$ and compare it to a situation of separation occurring with $b'_1 = s$ and $a'_1 = \underline{a}(b_1) + b_1\bar{y}(b_1) - s\bar{y}(b_1)$ (the new budget line passes through $(\bar{y}(b_1), \underline{a}(b_1) + b_1\bar{y}(b_1))$).

In this new situation, consider successively:

- The first-period aggregate welfare restricted to low productivity firms: "social concern" first-period utility remains constant but the individual profit increases when the firm reacts to the new offer.
- The same argument holds for the first-period aggregate welfare restricted to high productivity firms which were pooling in the initial situations.
- For high productivity firms which separate in the initial situation, the total effect can be decomposed (as in the proof of Proposition 1) into two parts: a substitution effect along the initial indifference curve which by definition does not affect the firm's utility and increases the (first-period) social concern utility; an income effect which is negative from the firm's point of view but positive in terms of aggregate welfare.

Finally, the impact on first-period aggregate welfare of switching to the new scheme is unambiguously positive.

Still assuming separation, we can now switch to a scheme with $b_1 = s$, bringing the low productivity firm to its minimum utility, a change which still increases first-period aggregate welfare (such a change has only positive income effects in terms of aggregate welfare). ||

The preceding facts hold independently of any assumption on $\Delta(b_1)$. The next property relies on the fact that $\Delta(b_1)$ increases with b_1 .

Lemma 7. *Assume that $\Delta(b_1)$ is an increasing function of b_1 . Then if $b_1 < b(\nu_1)$ induces a semi-separating or pooling continuation equilibrium, it cannot be a first-period optimal dynamic bonus.*

Proof. Let $\bar{x}_1$ denote the proportion of revealing high productivity firms with the scheme $\bar{b}_1 < b(\nu_1)$, where $b(\nu_1)$ is the optimal static bonus associated with ν_1 . Consider $\bar{\nu} = \nu_1 / [\nu_1 + (\bar{x}_1)(1 - \nu_1)]$ and $b(\bar{\nu})$ the static optimal bonus associated with $\bar{\nu}$.

From Proposition 3, $\bar{\nu} \geq \nu_1$ implies $b(\bar{\nu}) \geq b(\nu_1)$ (but $b(\bar{\nu}) \leq s$). (Note that for $b(\bar{\nu})$ the proportion of high productivity revealing firms $\bar{x}_1$ is strictly greater than $\bar{x}_1$: this follows from the property of $\Delta(b_1)$, Proposition 3, and the consideration of formula (14).)

Now instead of $\bar{b}_1$, take $b(\bar{\nu})$ as first-period bonus. The induced change can be decomposed by a thought experience into two parts:

- A change from $\bar{b}_1$ to $b(\bar{\nu})$ leaving the number of high productivity revealing firms unaffected. For the group of low productivity firms and revealing high productivity firms, this change, by definition of $b(\bar{\nu})$, is favourable from the CP's viewpoint in the first period and nothing is changed at the second period. For the subgroup of non-revealing high productivity firms, the move is also socially favourable in the first period (use Lemma 4 for $\nu_1 = 0$) and indifferent in the second. So, the intertemporal aggregate welfare increases.
- A change at $b_1 = b(\bar{\nu})$ in the proportion of revealing high productivity firms. This change is favourable to the CP in the first period (Lemma 5) and in the second period (Lemma 3).

The conclusion follows. ||

5.2 The optimal first-period bonus: Results

We now have enough understanding of the factors governing the CP's intertemporal welfare to be able to single out conclusions on the optimal first-period dynamic bonus.

First, the fact that the static optimal bonus is smaller than the social value of the good, has a somewhat weaker dynamic counterpart:

Proposition 5. *If the continuation equilibrium associated with the bonus s is separating, then b_1^D , the optimal dynamic first-period bonus, satisfies $b_1^D \leq s$.*

Proof. Suppose that $b_1^D > s$. Then:

- either b_1^D is separating, a fact ruled out by Proposition 2(i),
- or b_1^D is not separating but from Lemma 6, the scheme with bonus s would be better from the CP's first-period point of view. As it is also better for the second-period aggregate welfare (this is a variant of Lemma 3), a contradiction obtains. ||

Let us now derive a lower bound for b_1^D . The following is not very surprising:

Proposition 6. *If the continuation equilibrium associated with $b_1 = b(\nu_1)$ is separating, then: $b_1^D = b(\nu_1)$.*

Proof. Assume the contrary: $\bar{b}_1 \neq b(\nu_1)$ is optimal.

- $\bar{b}_1$ cannot be separating: given that the informational contents of $\bar{b}_1$ and $b(\nu_1)$ are the same, this would contradict the definition of $b(\nu_1)$.
- $\bar{b}_1$ cannot be a pooling or semi-separating equilibrium with $\bar{b}_1 \leq s$. According to Lemma 5, a separating equilibrium with $b_1 = \bar{b}_1$ would be superior for the first-period aggregate welfare to the equilibrium under consideration, it is also superior for the second-period point of view. However, this hypothetical separating equilibrium associated with $b_1 = \bar{b}_1$ is itself dominated by the scheme $b_1 = b(\nu_1)$, a contradiction.
- $\bar{b}_1$ cannot be a pooling or semi-separating equilibrium with $\bar{b}_1 \geq s$. According to Lemma 6, such an equilibrium would be dominated by an hypothetical separating equilibrium with $b_1 = s$, which in turn is dominated by $b_1 = b(\nu_1)$, a contradiction. $\parallel$

Proposition 6 looks intuitively reasonable. If the second best static bonus induces in addition full revelation in a dynamic context, it is the dynamic optimum. What is surprising is rather that the proof is more intricate than expected. The reason is that pooling in the first period (or semi-separation), a possibility which is not open in the second best static problem, may be an interesting option from the viewpoint of the first-period “social concern”. The fact that it is not so favourable after all—the argument on which the proof relies—is a consequence of Lemmas 5 and 6.

In Proposition 7, below, we will see that another intuitively appealing assertion, i.e. that the dynamic optimal first-period bonus is higher than the second-best optimal static bonus, holds only in favourable cases and for the same reason as above requires a non-straightforward proof. But beforehand we give a corollary to Proposition 6:

Corollary 2. $\exists \delta_0$ such that for $\delta \leq \delta_0$, the first-period dynamic optimum is $b(\nu_1)$.

Proof. Take δ_0 small enough such that $\Delta(b(\nu_1)) > \delta_0 \bar{\pi}^F(1)$. Hence $b(\nu_1)$ is necessarily in the separating zone and Proposition 6 applies.

Note however that $\delta_0 < 1$; for $\delta = 1$, $b(\nu_1)$ cannot be in the separating zone (remember that $\bar{\pi}^F(1) = \Delta(s)$). $\parallel$

Restricting now our attention to the well-behaved case, i.e. the case in which the incentive to conceal one’s information decreases with the bonus, we have:

Proposition 7. *In the well-behaved case (where $\Delta(b_1)$ is increasing in b_1), the optimal first-period bonus exceeds the optimal one-period bonus: $b_1^D \geq b(\nu_1)$.*

Proof. Suppose $b_1^D \leq b(\nu_1)$. According to Lemma 4 and because of $b(\nu_1) < s$, b_1 cannot be pooling without being at the frontier of semi-separation. But because of Lemma 7, b_1 cannot be semi-separating (weakly or strictly). Hence, if b_1^D is not semi-separating it is greater than $b(\nu_1)$.

From Proposition 6, if b_1^D is separating, it equals $b(\nu_1)$. The conclusion follows. $\parallel$

Lastly:

Proposition 8. *The optimal first-period scheme may induce the firms to pool, semi-separate or separate.*

The proof is by way of an example (see quadratic example below).

As this result can also be shown to hold for non-linear schemes, we see that the revelation principle does not hold (whereas it does with commitment). The fact that pooling may be desirable shows that no direct revelation game in which telling the truth is optimal for the firm may be optimal.

5.3 An example

Let us assume that the cost function has the following simple form: $\psi(y, \theta) = ky^2/2\theta$ (where $k > 0$). The reader can easily check that for this cost function:

$$y(b, \theta) = \frac{b\theta}{k} \quad (16)$$

$$\Delta(b) = \frac{b^2}{2k\bar{\theta}} [\bar{\theta} - \underline{\theta}]^2 \quad (17)$$

$$b(\nu) = s \frac{\nu\bar{\theta} + (1-\nu)\underline{\theta}}{\nu\bar{\theta} + (1-\nu)\underline{\theta} \left[1 + \frac{\lambda}{1+\lambda} \frac{\bar{\theta} - \underline{\theta}}{\bar{\theta}} \right]} \quad (18)$$

$$\bar{\pi}^F(\nu) = \frac{b^2(\nu)}{2k} (\bar{\theta} - \underline{\theta}). \quad (19)$$

$\Delta(b)$ is represented in Figure 4. Let $\hat{b}$ and $\hat{\hat{b}}$ be the bonuses that mark the upper limits of the pooling and semi-separating regions. From (17) and (19) and section (b), $\hat{b}$ and $\hat{\hat{b}}$ are defined by:

$$\hat{b} = \mu b(\nu_1) \quad (20)$$

$$\hat{\hat{b}} = \mu s \quad (21)$$

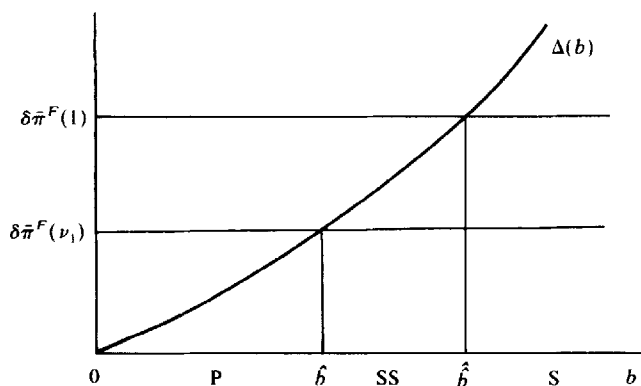


FIGURE 4

where

$$\mu \equiv \sqrt{\frac{\delta}{(\bar{\theta} - \underline{\theta})/\bar{\theta}}}. \quad (22)$$

Note that if $\mu < b(\nu_1)/s$, then $b(\nu_1)$ belongs to the separating region and thus is optimal. This case arises when the future is not too highly valued and the productivity differential is important. On the contrary, if μ is high, $\hat{b}$ and $\hat{b}$ are very high relative to s , and it becomes very costly to separate the two types of firm. Moreover, the second-period social gain of having better information is very small if the firms are very alike. Thus the optimal bonus b_1 leads to pooling, and from Lemma 3, is greater than s . It is also possible to show that, for some values of the parameters, it is socially optimal to induce semi-separation.

6. CONCLUSION

(a) The non-commitment assumption: We have analysed the ratchet effect in the context of a game. Another approach to this problem assumes that the CP is able to commit himself to an intertemporal sequence of incentive schemes; in this case he announces the current scheme as well as the revision procedure, and the firm solves its dynamic programming problem given the CP's plans. This is the set-up considered by Weitzman (1980) and Holmstrom (1982). However with commitment and no accrual of information unrestricted schemes, revision is not desirable and there is no ratchet effect (see footnote 1).

Note that the CP always prefers committing himself. Indeed he can always reach our (unique) equilibrium social welfare by duplicating his optimal strategy(ies). The firm's dynamic programming response is then its optimal strategy in the non-commitment game.

On the other hand, casual empiricism suggests that central planners generally do not commit themselves to revising incentive schemes in a given way. We must then ask why they do so. Although this question is out of the scope of this paper, we would like to mention some elements of answer. One possibility, of course, is that the CP is not aware of the benefits of commitment. But there may be deeper reasons. First, commitment may not be credible. The CP is free to design incentive schemes, but also to change them over time at his discretion. Also the planner may be replaced and the new planner may not feel obliged to abide by the scheme designed by his predecessor (remember that the time period in planning is not short). Second, the costs of designing intertemporal plans may be very high, in particular when there is a high number of economic units. Third, it is possible that the CP can over time obtain new information about the firm (other than its output) or about the future social price of output (for example, the CP obtains inside information about effort, observes indicators other than output, etc. . .). This fact by itself does not contradict the general proposition that in the absence of "transaction costs" the CP is better off committing himself if he can. But it greatly increases commitment costs: the CP must then give a detailed account of the future incentive schemes as a function of his information.¹¹

(b) Linear and non-linear schemes: We conducted the analysis under the assumption of a linear incentive scheme. Non linear schemes are considered in Laffont-Tirole [1984] for an alternative model of regulation. The main difference with the linear approach is that low productivity firms may want to pool with high productivity ones, even under the usual sorting assumptions, one must then check that incentives constraints are satisfied "upwards" and "downwards".

(c) Alternative interpretations of the model: The idea that if the relationship between two economic agents is not one-shot, these have incentives to hide their information, is a very general one. For example, the ratchet effect is ubiquitous in all relationships within a firm or institution (a zealous and efficient typist quickly becomes overloaded with work; a complacent employee or manager gets to do the ingrate work, etc.). More generally, it pervades any principal-agent relationship. An example similar to the one presented here is that of an efficient regulated firm which sees the price for its output revised downwards over time. Another example of regulation concerns pollution controls. Yao (1982) considers the dynamics of regulation of automobile emissions control and its effect on the strategy of car manufacturers, also from a non commitment point of view. The analysis of the dynamics of the principal-agent relationship could also be relevant to optimal taxation: the government over time learns the ability of tax-payers and in principle might want to use this information to design individual tax-schemes. Why it in general does not do so is an interesting question, on which this kind of analysis may throw some light.

First version received May 1983; final version accepted August 1984 (Eds.).

We are grateful to two anonymous referees for constructive comments—we bear responsibility for any remaining error.

NOTES

1. Holmstrom's result that revision procedures in the new Soviet incentive scheme may dominate static schemes relies on the piecewise linear aspect of the incentive scheme. Dynamics then allows to enlarge the class of possible schemes. With fully non-linear schemes, commitment and perfect correlation of the firm's characteristic over time, revision does not help and the ratchet effect does not occur; under imperfect correlation and commitment, revision does help, but is done in such a way that the firm does not take future consequences into account, so there is no ratchet effect (see Baron-Besanko (1983) and Roberts (1983)).

2. The formalization of the firm as a unique aggregate agent is standard in the literature. It seems to be an unavoidable first step in the analysis. Note however that facts suggest that the efficiency of incentive schemes do crucially depend upon the way incentives are channelled within the firm through "delegation" systems.

3. An alternative interpretation is that there are a large number of firms and that $f(\theta)$ describes the proportion of firms of parameter θ . Although on many points the following analysis remains valid with such an interpretation, the comments made are put in terms of the one-firm assumption.

4. With our first interpretation of λ , s can be for example considered as the world price of the good; since the shadow cost of \$1 for the government is $\$(1+\lambda)$ because of distortionary taxation, the social value of one unit of good is $p = s(1+\lambda)$; $sy - R$ is also nothing else than government income.

5. In fact a similar argument shows that $b = s$ is not optimal either.

6. The fact that δ is a common discount factor is not crucial for the theory although it simplifies the reasoning. However, with different discount factors, the meaning of W , the sum of discounted profit and "socially concerned" welfare (see (10)), is ambiguous.

7. For notational simplicity, we shall use the same letters for strategies and actions.

8. The sequential equilibrium concept is defined for finite games and is then a refinement of the Perfect Bayesian Equilibrium concept.

9. The mixed strategy equilibrium is mainly an artefact of the extreme stylization of our model and is likely to disappear under the more realistic assumption of a continuum of possible values for θ .

10. Hence for quadratic cost functions, Δ is "very well-behaved". This suggests that the influence of higher order derivatives must be strong to induce a failure of the assumption.

11. Note that these future incentive schemes must be incentive compatible from the point of view of the planner if the firm cannot observe all the CP's new information.

REFERENCES

- BARON, D. and BESANKO, D. (1983), "A Theory of Regulation and Information in a Continuing Relationship" *Information Economics and Policy*, forthcoming.
- BARON, D. and MYERSON, R. (1982), "Regulating a Monopolist with Unknown Costs", *Econometrica*, **52**, 911-930.
- BONIN, J. (1976), "On the Design of Managerial Incentive Structures in a Decentralized Planning Environment", *American Economic Review*, 682-687.

- EKERN, S. (1979), "On the Soviet Incentive Model—Comment", *Bell Journal of Economics*.
- FAN, L. S. (1975), "On the Reward System", *American Economic Review*, 226–229.
- FUDENBERG, D. and TIROLE, J. (1983), "Sequential Bargaining under Incomplete Information", *Review of Economic Studies*, 221–247.
- GIDIN, S. (1970), "A Model of the Soviet Firm", *Economic Planning*, 10, 145–157.
- GUESNERIE, R. and SEADE, J. (1982), "Non-linear Pricing in a Finite Economy", *Journal of Public Economics*, 157–180.
- HOLMSTROM, B. (1982), "Design of Incentive Schemes and the New Soviet Incentive Model", *European Economic Review*, 127–148.
- KREPS, D. and WILSON, R. (1982), "Sequential Equilibria", *Econometrica*, 863–894.
- LAFFONT, J. J. and TIROLE, J. (1984) "Dynamic contracting with cost observability" (Mimeo).
- MIRPLEES, J. (1971), "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies*, 38, 175–208.
- ROBERTS, K. (1983), "Long Term Contracts" (Mimeo).
- SEADE, J. (1977), "On the Shape of Optimal Tax Schedules", *Journal of Public Economics*, 7, 203–236.
- WEITZMAN, M. (1976), "The New Soviet Incentive Model", *Bell Journal of Economics*, 251–257.
- WEITZMANN, M. (1980), "The 'Ratchet Principle' and Performance Incentives", *Bell Journal of Economics*, 302–308.
- YAO, D. (1982), "Strategic Behavior in a Technology Forcing Regulations Setting: Automobile Emissions Control" (Mimeo GSB Stanford).