THE NEW ECONOMICS OF REGULATION TEN YEARS AFTER

BY JEAN-JACQUES LAFFONT

The new economics of regulation is an application of the principal-agent methodology to the contractual relationship between regulators and regulated firms. After a critique of the traditional paradigms of regulation from the point of view of information economics a canonical model of regulation under asymmetric information is developed. A survey of the main results obtained in the new economics of regulation is then provided, in particular concerning the implementation of optimal contracts by a menu of linear contracts, the dichotomy between pricing and cost reimbursement rules, the auctioning of incentive contracts, the dynamics of contracting under limited commitment, and the hierarchical problems in regulation. Empirical implications are then discussed and avenues of further research are described in the conclusion.

KEYWORDS: Regulation, incentives, asymmetric information, contract theory.

INDUSTRIAL ORGANIZATION IS THE STUDY OF ECONOMIC ACTIVITY at the level of a firm or an industry when the paradigm of perfect competition appears inadequate. By giving up ambitions of general equilibrium analysis this field has provided us with a large number of insights into the strategies of economic actors within firms and in environments of imperfect competition. Jean Tirole's book (1988) synthesizes in a superb way the progress of the last ten years that game theory made possible.

Regulation is the public economics face of industrial organization. It explores the various ways in which governments interfere with industrial activities for the good or for the bad. From the start this field has had a less naïve view of the nature of public intervention than traditional public economics as examplified by Atkinson-Stiglitz's manual (1980).

Regulation is becoming a major area of economics because in a world which has given up the debates between socialism and capitalism, it is going to be the major battleground of the opposition between more or less governmental interference in economies, very much like at the end of the eighteenth century when Turgot and Quesnay were debating with Galiani and Diderot the trade regulation of corn. This is quite apparent in Europe where the political debates of the seventies have been replaced by policy debates such as: which part of the telecommunication industry should remain regulated, can generation of electricity be privatized, how should environmental problems be dealt with, how should banks be regulated?

In this lecture I would like to summarize some of the main achievements of ten years of what I call the new economics of regulation.

The starting point of the analysis is a natural monopoly situation where, under complete information, technological efficiency requires a single operating
firm. The need to avoid monopolistic behavior calls for public intervention. Instead of focusing on particular regulatory institutions, the new economics of regulation, in the tradition of the mechanism design literature, aims at characterizing optimal regulation.

The framework is a principal-agent set up in which the principal is the State or the regulatory institution and the agent is the regulated firm. The principal maximizes social welfare under incentive constraints which result from the informational advantage of the agent and its strategic behavior. The regulation problem is essentially a control problem under incomplete information and the results obtained in this literature have a much broader interest than regulation itself, as we will argue in the conclusion.

In Section 1 the paradigms of the traditional economics of regulation are briefly criticized. Section 2 introduces a canonical model which enables us to synthesize most of the literature. Section 3 presents some major results of this new field. Empirical implications are discussed in Section 4. The limits of the current approach and directions for further research conclude the paper.

1. THE TRADITIONAL REGULATORY PARADIGMS OF NATURAL MONOPOLIES

Until the beginning of the eighties the economics of regulation was taught within two quite different paradigms that are reviewed below.

1.1. Cost of Service Regulation

The regulation of utilities has been implemented in most countries by constraining the rate of return on capital. The premise was that it was necessary to attract capital to utilities while avoiding excessive exercise of monopoly power. A few main features characterize this type of regulation.

- A fair rate of return on investment above the market rate is guaranteed as long as investments are prudent.
- Prices are determined to equate average costs with this imputed charge for capital.
- Prices remain fixed during the regulatory lag until some party, the regulatory commission, the firm or the consumers, initiates a new regulatory review leading to new prices.
- The regulatory review is a process of checks and balances in which the conflicts between the firm which wishes high prices and the consumers who wish low prices are arbitrated by the regulatory commission.

This process led to the following criticisms. The allocation of fixed costs among products became an accounting nightmare for a doubtful pricing method. By equating prices to average costs, no incentives for cost minimization (through a choice of effort levels, or a policy of innovation) existed, except in a very

2 See, for example, Berg and Tschirhart (1988), Breyer (1982), Kahn (1971), Schmalensee (1979), Spulber (1989).
limited way through the regulatory lag. No incentive for the provision of quality existed either, except that the reimbursement of all costs made the firm indifferent about incurring monetary expenses for quality.

The reward being related to capital led to the well known overcapitalization or Averch-Johnson effect. This phenomenon has been overemphasized as in most cases investment is directly regulated.

Finally, the process has been criticized as being extremely slow and costly.

From the theoretical point of view the main remark called for is the total lack of a normative framework. The enormous literature written on cost of service regulation is a list of defects of this procedure with no clue whatsoever on whether those defects are the outcomes of optimal trade-offs. For example the choice of a rate of return to capital higher than the market rate reflects a need to give the firm some sort of rent without any clear justification. Because of this lack of a normative framework it is impossible to appraise this type of regulation or to suggest improvements.

This much criticized procedure has nonetheless some basic advantages. First, it is imbedded in a positive view of government which recognizes the need for an organization of checks and balances far away from the naïve view of a benevolent State. Second it defines a feasible procedure which avoids any risk of bankruptcy for utilities. Third, and relatedly, by promising a fair rate of return on capital it offers a type of long run commitment which is crucial for the long run investments needed in utilities such as electric or telecommunications utilities.

1.2. The Ramsey-Boiteux Regulation

In sharp contrast with COS regulation, the Ramsey-Boiteux regulation is the outcome of a well defined optimization process. The regulator maximizes social welfare by choosing tariffs under a budget constraint. The focus is clearly here on pricing. For example optimal departures from marginal cost pricing must be, in the simple case of separable utility functions without income effects, simply inversely proportional to price elasticities.

The coefficient of proportionality is related to the Lagrange multiplier of the budget constraint which itself depends on both demand functions and cost function. The informational burden on the regulator is enormous and explains why this type of regulation has never been used even in the nationalized Electricité de France ruled by M. Boiteux where the head of the firm was a benevolent social maximizer. The simpler Allais doctrine of uniformly increasing prices above marginal costs was chosen.

The second remark is that the role of the firm’s budget constraint has never been theoretically established. Why should the regulators be constrained to not using transfers and to distorting prices because of fixed costs? The third remark is that this approach, very much in the tradition of a benevolent welfare State,

---

3 See Boiteux (1956), Baumol, Panzar, and Willig (1982), Spulber (1989).
completely sets aside incentive problems, which today play a central role in policy debates.

1.3. The Neglected Role of Asymmetric Information

The conclusions I want to derive from this brief review are as follows: the traditional cost of service regulation lacks a theoretical vision which would make sense of its practice. The Ramsey-Boiteux formulation raises implicitly a crucial difficulty. Regulators, whatever their objectives, are fundamentally constrained by their lack of information on the firms they are regulating. It is only by making explicit these constraints that some intellectual progress can be made in this field.

The first step of the new theory of regulation has been to formulate regulation as an agency problem.\(^4\) By stressing the role of informational asymmetries and making use of incentive theory developed in the seventies, it has considerably extended the relevance of the Ramsey-Boiteux paradigm while at the same time making progress in providing a normative appraisal of cost of service regulation.

Of course this particular innovation of theoretical economics does not cover all the issues raised by regulation. In my conclusion I will discuss the weak points of the theory constructed so far.

In the ten years from the 1982 Baron-Myerson piece to my book with Jean Tirole (1993), *A Theory of Incentives in Procurement and Regulation*, a lot has been achieved enabling us to make some sense of such phenomena as the price-cap revolution in the regulation of utilities or the incentive reform in procurement.

It is this story that I would like to tell in this lecture for three reasons.

First, it is a good example of theoretical innovations driven by the effort to understand very concrete phenomena.

Second, it has provided a whole set of very general results of contract theory that should be used in other fields.

Third, despite the lack of comprehensive coverage of all the issues of regulation, it has provided a theoretical framework for empirical studies which now need to be done.

In telling this story I will overemphasize my own results for expositional convenience.\(^5\)

2. A CANONICAL MODEL

Consider a natural monopoly which is a multiproduct firm with cost function

\[
C = C(\beta, e, q_1, \ldots, q_n) + \varepsilon
\]

where the cost \(C\) is observable, \(q_k\) is the production level of good \(k\) \((k = 1, \ldots, n)\), \(\beta\) is a cost parameter known only to the firm (adverse selection

---

\(^4\) Loeb and Magat (1979) deserves credit for this first formulation.

\(^5\) Valuable complementary surveys include Baron (1989), Besanko and Sappington (1987).
parameter), $e$ is an unobservable action of the firm (a moral hazard variable: an
effort for example), and $\tilde{e}$ is a random variable which can be interpreted as a
shock on cost or as an observational error of the regulator.

By choosing effort level $e$ the firm incurs a nonmonetary disutility $\psi(e)$ with
$\psi' > 0$, $\psi'' > 0$, and $\psi'' > 0$.\(^6\)

The utility derived by consumers from these commodities is denoted

$$S(\theta, s, q_1, \ldots, q_n)$$

where $\theta$ is a parameter which, depending on context, can be private information
of the firm or of consumers, and where $s$ can be an effort level of the firm or of
consumers. In the first case the disutility of the firm is $\psi(e, s)$ and $s$ is to be
interpreted as a nonverifiable quality variable.

We make the accounting convention\(^7\) that cost is reimbursed to the firm and
that revenues from sales, $R(q)$, accrue directly to the regulator.

If $\hat{t}$ is the monetary transfer to the firm from the regulator, the firm’s
expected utility is then

$$EU = \hat{t} + R(q) - C(\beta, e, q) - \psi(e, s) = t - \psi(e, s)$$

where $t$ is the net transfer. Later we will discuss the case with no direct transfer
from the regulator to the firm.

The consumers’ utility is defined as

$$S(\theta, s, q) - R(q) - (1 + \lambda)\hat{t}$$

where $\lambda$ is the social cost of public funds (because of distortionary taxation).

Expected social welfare for a utilitarian regulator\(^8\) is then\(^9\)

$$W = S(\theta, s, q) - R(q) - (1 + \lambda)\hat{t} + EU$$

$$= S(\theta, s, q) + \lambda R(q) - (1 + \lambda)(C(\beta, e, q) + \psi(e, s)) - \lambda EU.$$

**Remarks**

1. $\beta$ and $e$ can be multidimensional without changing the basic intuitions.
   Multidimensional $\beta$ raises essentially technical difficulties.\(^10\) Multidimensional
   $e$ raises the more interesting questions of allocation of effort across various
tasks.\(^11\)

2. The goods produced can be marketed goods, public goods or public bads,
   and also verifiable dimensions of quality.

\(^6\) $\psi'' > 0$ is a concavity assumption; $\psi''' > 0$ makes stochastic mechanisms irrelevant.
\(^7\) It is only an accounting convention because these quantities are here assumed observable.
\(^8\) The assumption of a utilitarian regulator is by no means essential. It is enough for the analysis
   that the transfers from consumers to the natural monopoly be socially costly. A specificity of
   regulation with respect to the principal-agent problem is that the costly nature of these transfers is
   not as obvious as usual.
\(^9\) This formulation of social welfare shows that giving up a rent to the firm is socially costly. This
   is essential; otherwise the first best can be achieved by giving to the firm a transfer equal to
   the social surplus.
Special Cases

(a) **Procurement of Public Good** (e.g., military equipment, pollution abatement program): Consider a public project which has utility $S$ and cost

$$C = C(\beta, e) + \tilde{e}.$$ 

(b) **Simple Natural Monopoly**:

$$C = (\beta - e)q + \alpha + \tilde{e}$$

where $\alpha$ is a known fixed cost.

Net utility for consumers is

$$S(q) - p(q)q.$$ 

(c) **The Baron-Myerson Model**:

$$C = \beta q - e + \alpha, \quad \psi(e) = e.$$ 

$C$ becomes de facto nonobservable.

(d) **Multiproduct Regulation**:

$$C = C(\beta, e, q_1, \ldots, q_n) + \tilde{e}$$

with consumers' utility

$$S(q) - R(q).$$

The revenue function $R(q)$ encompasses linear pricing, two part tariffs, and nonlinear pricing.

(e) **Service Quality**:

$$C = (\beta - e)q + \alpha + \tilde{e}$$

with consumers' utility

$$S(\theta, s, q)$$

and firms's disutility of effort

$$\psi = \psi(e + s).$$

By effort $s$ the firm can improve quality for consumers. Efforts to minimize cost or to improve quality are substitutes.

(f) **Polluting Natural Monopoly**:

$$C = C(\beta, e, q_1, q_2) + \tilde{e}$$

with consumers' utility

$$S(q_1) - p(q_1)q_1 - V(q_2).$$

$q_2$ is now interpreted as a pollution level.

Other Interpretations

1. **Sharecropping**: Interpret $C$ as a production level, $q_1, \ldots, q_n$ as input levels with prices $p_1, \ldots, p_n$. The tenant's welfare is

$$t - \psi(e).$$
The landlord's welfare is

\[ S(C) - \sum_{k=1}^{n} p_k q_k - t. \]

2. Managerial Compensation: Interpret \( C \) as profit, \( q_1, \ldots, q_n \) as production and input levels. The manager's utility is

\[ t - \psi(e). \]

The stockholders' utility is

\[ C - t. \]

3. Private Procurement between Two Firms: \( S \) is now the utility of firm 1 which obtains quantities \( q \) from firm 2.

4. Internal Contract within an Organization: The two parties can be two divisions within a firm. This leads to a theory of transfer pricing.

3. MAIN RESULTS

3.1. Implementation by a Menu of Linear Contracts

Let us consider the simplest case of a fixed size project which is a public good valued \( S \) by consumers and with cost

\[ (1) \quad C = \beta - \bar{e} + \bar{e}, \]

and let us ignore momentarily the random variable \( \bar{e} \). It is common knowledge at the time of the contract that \( \beta \) has a distribution function \( F(\beta) \) on \( [\beta, \bar{\beta}] \) with a density function \( f \) strictly bounded away from zero. We also assume that \( d(F(\beta)/f(\beta))/d\beta > 0 \) to avoid bunching.

Assuming that cost is paid directly by the regulator, the firm's utility level is

\[ (2) \quad U = t - \psi(e) \]

or

\[ (3) \quad U = t - \psi(\beta - C). \]

Under complete information of the regulator with respect to \( \beta \) and \( e \), maximization of social welfare under the firm's participation constraint \( U \geq 0 \) leads immediately to

\[ \psi'(e) = 1 \quad \text{or} \quad e = e^*, \]

\[ t = \psi(e^*) \quad \text{or} \quad U(\beta) = 0 \quad \text{for any} \ \beta. \]

Effort is optimal and the firm's individual rationality constraint is binding for any \( \beta \).

12 For these examples the assumption of risk neutrality may seem inappropriate. See Laffont and Rochet (1994) for an extension of the analysis to risk averse agents.

13 Bunching means that firms with different \( \beta \) characteristics are treated equally. It complicates the analysis without changing the basic economic intuitions (see Guesnerie and Laffont (1984)).
Under incomplete information, we face an adverse selection problem where $\beta$ is the parameter of private information, $C$ is the action variable, and $t$ the compensatory payment. From the revelation principle, any method of regulating the firm is equivalent to a mechanism \( \{t(\beta), C(\beta)\}_{\beta \in [\underline{\beta}, \overline{\beta}]} \) which induces truthful revelation of the private information parameter.

Such a mechanism induces truthful behavior iff

\[ \beta \in \operatorname{Arg\max}_{\beta} \left\{ t(\beta) - \psi(\beta - C(\beta)) \right\}. \tag{4} \]

Before characterizing revelation mechanisms let us pause to stress the key underlying assumptions: costless communication is assumed, information can be described in a Bayesian fashion, the regulator has the ability to commit to a mechanism. The weakening of those assumptions will be considered later.

The first order and second order conditions\(^{14} \) associated with program (4) are

\[ i + \psi'(\beta - C)\dot{C} = 0, \tag{5} \]
\[ \dot{C} \geq 0. \tag{6} \]

Let $U(\beta) = t(\beta) - \psi(\beta - C(\beta))$ be the firm’s utility level when it is truthful.

To these first- and second-order conditions of incentive compatibility must be added an individual rationality constraint

\[ U(\beta) \geq 0 \quad \forall \beta \tag{7} \]

where $0$ is the normalized reservation utility of the firm.\(^{15} \)

Social welfare for an utilitarian regulator is

\[ S = (1 + \lambda)(t + C) + U \tag{8} \]
\[ = S - (1 + \lambda)(\beta - e + \psi(e)) - \lambda U. \tag{9} \]

Optimal regulation is the outcome of the maximization of expected social welfare under incentive and individual rationality constraints, i.e.,

\[ \text{Max} \int_{\beta}^{\overline{\beta}} \left\{ S - (1 + \lambda)(\beta - e + \psi(e)) - \lambda U(\beta) \right\} dF(\beta), \tag{10} \]
\[ \dot{U}(\beta) = -\psi'(e(\beta)) \quad \forall \beta, \tag{11} \]
\[ U(\beta) \geq 0 \quad \forall \beta. \tag{12} \]

In the above program (11) is a rewriting of the first order incentive constraint (5) using $U$ instead of $t$ as a variable and the second order incentive constraint has been neglected.

\(^{14} \) These conditions are here necessary and sufficient because the Spence-Mirrlees condition (here equivalent to $\psi'' > 0$) is satisfied. Indeed the local second order condition is $\psi''$. $C \geq 0$ and local incentive compatibility implies global incentive compatibility (see Guesnerie and Laffont (1984)).

\(^{15} \) It is often legitimate to assume that the reservation utility of a firm depends on its $\beta$ characteristics. This complicates substantially the analysis because it is not always easy to know where the I.R. constraint is binding (see, for example, Lewis and Sappington’s (1989) paper on countervailing incentives).
(11) tells us that truth-telling imposes giving up to the firm a rent which decreases with $\beta$ at the rate $\psi'(e(\beta))$ if $e(\beta)$ is the effort level that the regulator wishes to implement.

Optimal regulation is described by the first order conditions of the above program, namely:

\begin{equation}
\psi'(e^*(\beta)) = 1 - \frac{\lambda F(\beta)}{1 + \lambda f(\beta)} \psi''(e^*(\beta)),
\end{equation}

\begin{equation}
U^*(\beta) = \int_\beta^\bar{\beta} \psi'(e^*(\tilde{\beta})) d\tilde{\beta}.
\end{equation}

The most efficient firm, firm $\beta$, exerts an effort level which is the full information optimal level (since $\psi'(e^*(\beta)) = 1$ as $F(\beta) = 0$) and obtains a rent $U(\beta)$. The least efficient firm, firm $\bar{\beta}$, has no rent and a distorted effort level.

The intuition for this result is as follows. From (14) we see that the higher the effort levels, the higher the rents of asymmetric information. As rents are costly to the regulator because of the social cost of public funds (see (9)), effort levels are depressed to decrease those rents. Inefficiencies of the choice of effort levels enable the regulator to mitigate somewhat the losses due to the rents of asymmetric information.

This fundamental trade-off between rents and allocative inefficiencies is in my view one of the basic insights of economics. It is here achieved by a revelation mechanism. A more concrete implementation is obtained by a nonlinear transfer as follows.

From (13), $e^*(\beta)$ is decreasing and therefore $C^*(\beta) = \beta - e^*(\beta)$ increasing strictly and invertible. Let $\beta(C)$ be the inverse function. By definition

\[ t(\beta) = U^*(\beta) + \psi(e^*(\beta)) = t(\beta(C)) = T(C). \]

Optimal regulation can be achieved by offering the transfer function $T(C)$ and by letting the firm select itself by choosing a cost level. It is easy to check that $T(C)$ is here convex (see Figure 1).

Firm $\beta$ picks the contract $A$, firm $\hat{\beta}$ picks $B$, firm $\bar{\beta}$ picks $D$, etc.

As $T(\cdot)$ is convex it can be replaced by the family of its tangents without affecting the choices of the agents. A parametric representation of the family of tangents is

\[ t(C, C^a) = a(C^a) - b(C^a)(C - C^a). \]

By announcing his cost characteristic $\beta$ the firm "announces" a cost $C^a = C^*(\beta)$. It receives a lump sum transfer $a(C^a)$ and if the realized cost $C$ differs from the announced cost, the firm and the regulator share the overrun in fractions $(b, 1 - b)$.

More generally, when the principal does not maximize expected social welfare but another criterion, the trade-off is between rents given up to the agents and distortions from the principal's preferred allocation.

As seen in Figure 1, a high share of cost reimbursed is associated with a low lump sum transfer.
Without the shock $\bar{\epsilon}$, the realized cost is always the announced cost and no adjustment needs to be made.

Suppose now that we reintroduce the error term $\bar{\epsilon}$ which has a probability distribution independent of $\beta$ and $\epsilon$. Because implementation was achieved via a menu of linear contracts and the firm is risk neutral, the error term does not affect incentives and the same effort levels are implemented with the same transfer functions.

Thanks to this linear implementation, observational errors in the "action variable" $C$, on which is based the contract, have no cost. We obtain the following proposition.

**Proposition 1**—Implementation by a Menu of Linear Contracts (Laffont-Tirole (1986)): Under the assumptions of this section, the optimal procurement contract is implemented by a menu of linear cost sharing rules in which firms selfselect themselves. Accordingly randomness in the cost evaluation, whatever its origin, entails no welfare loss as long as the firm is risk neutral.

This straightforward but important result was given a further interpretation in the case of a team of firms or agents by McAfee and McMillan (1991) and Picard and Rey (1990).

Suppose that we have several independent firms with costs $C^i = \beta^i - e^i$ which all realize a project. If the $\beta^i$ are stochastically independent, and costs observable, optimal regulation is identical for each firm. Suppose however that we observe only total costs $C = \Sigma_i C^i$ because accounting manipulations between firms are not controllable.

---

18 See Melumad and Reichelstein (1989) for cases where noisy observation of actions is costly.

19 See Laffont and Tirole (1990a) for more general assumptions. What is needed is the convexity of the optimal nonlinear transfer $T(C)$. Approximate convexity leads to small departures from optimality by using the menu of linear contracts derived from the convexification of $T(C)$.
If the implementation by menus of linear contracts is possible for each firm, $C^i$ can be replaced by $C^i + \sum_{j \neq i} C^j$ and the same allocation is implementable. Indeed $\sum_{j \neq i} C^j$ plays the same role as the random variable $\tilde{\varepsilon}$ above. Since it is only a function of the $\beta^j$, $j \neq i$, at the Nash equilibrium of the game, it is treated by firm $i$ as an exogenous random variable.

So, as long as risk aversion is negligible and there is no limited liability constraint, the lack of observability of individual performances in a team does not affect incentives. Clearly, to be practically meaningful such a result requires a small number of team members. With a large number any slight risk aversion would change drastically the result since each member of the team would have a very large uncertainty over the others' aggregate costs.

A similar interpretation can be used to say that under similar circumstances nonpoint source or diffuse pollution does not raise incentive difficulties beyond those of point pollution.\textsuperscript{20}

Procurement has for a long time been carried out with fixed price contracts (corresponding to $b(C^a) = 1$) or cost plus contracts (corresponding to $b(C^a) = 0$). Fixed price contracts excel in giving incentives for effort but lead to high rents. At the other extreme cost plus contracts give no incentive for effort but may avoid any rent.

What theory tells us is that the best procurement method is to offer a menu of sharing rules among which firms selfselect themselves. Incentive contracts with some sharing of overruns are familiar in the defense industry. It is only recently on the basis of those theoretical results that a more systematic use of menus of contracts is made, for example in the electricity industry.\textsuperscript{21}

### 3.2. The Dichotomy between Pricing and Cost Reimbursement Rules

Consider a regulated firm with aggregate cost function

\begin{equation}
C = C(\beta, e, q) \quad \text{with} \quad q = (q_1, \ldots, q_n).
\end{equation}

The $n$ commodities are sold by the firm and let $p_k(q)$, $k = 1, \ldots, n$, be the inverse demand functions. Assuming that costs are paid to and revenues from sales are collected by the regulator, the consumers' welfare is

\begin{equation}
S(q) - \sum_k p_k(q) q_k - (1 + \lambda) \left( t + C - \sum_k p_k(q) q_k \right)
\end{equation}

\begin{equation}
= S(q) + \lambda \sum_k p_k(q) q_k - (1 + \lambda) (t + C).
\end{equation}

\textsuperscript{20} See Laffont (1992).

\textsuperscript{21} See, for example, the incentive formula that the Federal Energy Regulatory Commission proposed for the Palisades Generating Company (see Lorenzo Brown's testimony on behalf of the FERC in the Palisades Generating Company (1990) case).
Consequently social welfare is

\[ W = S(q) + \lambda \sum_k p_k(q)q_k - (1 + \lambda)(C + \psi(e)) - \lambda U. \]

Let \( E(\beta, C, q) \) represent the solution for \( e \) of (15). It is the effort level needed to realize cost \( C \) while producing \( q \) when the efficiency parameter is \( \beta \).

A crucial quantity is \( E_\beta = -C_\beta/C_e \) which is the rate at which a manipulation of \( \beta \) enables the firm to decrease its effort level. The incentive compatibility conditions show that \( E_\beta \) plays a major role in defining the informational rent. Indeed we have

\[ U = t - \psi(e) = t - \psi(E(\beta, C, q)). \]

From the envelope theorem

\[ \dot{U} = -\psi'(e)E_\beta. \]

The higher \( E_\beta \), the higher the rate at which rent must be given up to more efficient firms.

If we assume linear pricing and optimize expected social welfare under incentive constraints we obtain a generalization of the Ramsey-Boiteux equations:\textsuperscript{22}

\[ L_k = \frac{p_k - C_k}{p_k} = \frac{\lambda}{1 + \lambda} \frac{1}{\hat{\eta}_k} + \left[ \frac{\lambda F(\beta)\psi'(e)}{(1 + \lambda)f(\beta)p_k} \right] \frac{d}{dq_k}(E_\beta) \]

\[ (k = 1, \ldots, n) \]

as well as the optimal effort levels:

\[ \psi'(e) = -C_e - \frac{\lambda}{1 + \lambda} \frac{F(\beta)}{f(\beta)} \frac{d}{de}[\psi'(e)E_\beta]. \]

Equation (20) defines the optimal pricing rule. The Lerner index, \( L_k \), equates the Ramsey-Boiteux index

\[ R_k = \frac{\lambda}{1 + \lambda} \frac{1}{\hat{\eta}_k} \]

(\( \hat{\eta}_k \) is the superelasticity of good \( k \) and \( 1 + \lambda \) is the social price of public funds which replaces the usual multiplier of the budget constraint), plus the incentive correction

\[ I_k = \left[ \frac{\lambda F(\beta)\psi'(e)}{(1 + \lambda)f(\beta)p_k} \right] \frac{d}{dq_k}(E_\beta). \]

\textsuperscript{22} For \( n = 2 \) the superelasticity

\[ \hat{\eta}_1 = \eta_1 \left[ \frac{\eta_1\eta_2 - \eta_1\eta_2}{\eta_1\eta_2 + \eta_1\eta_1} \right] < \eta_1 \]

where \( \eta_i \) is good \( i \)'s price elasticity and \( \eta_{ij} \) the cross price elasticity of goods \( i \) and \( j \).
Equation (21) defines implicitly the cost reimbursement rule since by maximizing
\[ t(C) - \psi(e) \]
the firm obtains
\[ \psi'(e) = t'(C)C_e. \] (23)

If \( t'(C) = -1 \), the total transfer received from the regulator is independent of \( C \) (remember \( t \) is the net transfer). It is analogous to a "fixed price" contract in procurement. (Beware: do not confuse this "fixed price" with the prices of commodities. It is better to say a "fixed reimbursement" contract.)

If \( t'(C) = 0 \), the global transfer received from the regulator is \( C \) up to a constant. It is analogous to a "cost plus" contract in procurement.

What (21) tells us is that
\[ t'(C) = -1 + \frac{\lambda}{1 + \lambda} \frac{F(\beta)}{f(\beta)} \frac{d}{de}(\psi'(e)E_{\beta}) \] (24)

Costs are reimbursed with a nonlinear transfer rule \( t(C) \) and firm \( \beta \) selects a share of global reimbursement which is
\[ \frac{\lambda}{1 + \lambda} \frac{F(\beta)}{f(\beta)} \frac{d}{de}(\psi'(e)E_{\beta}) \] (25)

The most efficient type \( \beta \) selects a fixed reimbursement contract \( (F(\beta) = 0) \) and less efficient types choose to share their costs with the regulator.

Under appropriate restrictions on the cost function and the distribution \( F(\cdot) \), this nonlinear rule can also be implemented by a menu of linear contracts.

Regulation has two main parts characterized above, pricing and cost reimbursement rules. A dichotomy obtains when the incentive corrections of the pricing equation vanish. We then obtain a clear benchmark case. Optimal pricing should be Ramsey-Boiteux pricing for the cost function induced by the incentives associated with the optimal cost reimbursement rule. The cost reimbursement rule deals with incentives for cost minimization and contrary to what happens in the Baron-Myerson model, prices are not used to extract the rent of asymmetric information.

**Proposition 2**—The Incentive-Pricing Dichotomy (Laffont-Tirole (1990a)):

The incentive-pricing dichotomy holds if and only if there exists a function \( \xi(\cdot) \) such that
\[ C = C(\xi(\beta, e), q). \]
The intuition behind this proposition is simple. The incentive correction vanishes iff
\[ \frac{d}{dq_k}(E_\beta) = 0 \]
for any \( k \). But as \( E_\beta = -C_\beta/C_e \) it holds iff the "marginal rate of substitution" between \( e \) and \( \beta \) in \( C \) is unaffected by quantities, hence the result by the Leontief theorem on aggregation.

If changing quantities (and therefore prices) does not affect \( E_\beta \), i.e. the rent of asymmetric information, there is no point in distorting prices for incentive reasons.

**Budget Balance**: In many cases, direct transfers to the firm are not allowed.\(^{23}\) All the income of the firm is entirely provided by sales revenue,\(^ {24}\) hence the "budget constraint:"
\[ U \leq \sum_k p_k(q)q_k - C - \psi(e). \]

Maximization of expected social welfare under incentive and individual rationality constraints becomes
\[
\text{Max} \int_\beta \left\{ S(q(\beta)) - \sum_k p_k(q(\beta))q_k(\beta) + U(\beta) \right\} dF(\beta),
\]
\[
\dot{U}(\beta) = -\psi'(E(\beta, C(\beta), q(\beta))) E_\beta \quad \forall \beta,
\]
\[
\sum_k p_k(q(\beta))q_k(\beta) - C(\beta) - \psi(E(\beta, C(\beta), q(\beta)))
\]
\[ - U(\beta) \geq 0 \quad \forall \beta. \]

If we denote \((1 + \lambda(\beta))f(\beta)\) the multiplier of the individual rationality constraint we obtain the same equations as above except that
\[
\frac{\lambda F(\beta)}{(1 + \lambda)f(\beta)} \quad \text{becomes} \quad \int_\beta^{\beta} \lambda(\bar{\beta}) f(\bar{\beta}) d\bar{\beta}
\]
\[ \frac{(1 + \lambda(\beta))f(\beta)}{(1 + \lambda(\beta))f(\beta)}. \]

In particular, under the price-incentive dichotomy, relative prices follow the same rules as before but the absolute level of prices is set in order to satisfy the budget balance rationality constraint. By giving up the instrument of transfer and therefore imposing budget balance, even when the cost function would allow the price-incentive dichotomy, the pricing rules are strongly affected by the need to provide incentives for participation to the firm.

\(^{23}\) See Laffont and Tirole (1993, Ch. 15) for a tentative explanation of why one might want to exclude transfers.

\(^{24}\) We assume here linear pricing. A similar analysis can be carried out with nonlinear pricing (see Laffont and Tirole (1993, Ch. 2)).
Note that the above regulation gives up transfers but not the right to choose costs.
If this policy instrument is given up as well, the firm becomes the residual claimant for its cost as in a fixed price contract. Effort level will follow the efficient rule $\psi'(e) = -C_e$. The main consequence will be that a high rent

$$U(\beta) = \int_{\tilde{\beta}} \bar{\psi}(e^{**}(\tilde{\beta})) E_\beta d\beta$$

will have to be given up to firms through high prices as a high profile $e^{**}()$ of effort will be implemented.

This is basically the hidden side of the famous price cap proposals which for a while were presented as a free lunch guaranteeing efficiency of utilities. More efficiency can only be induced at the cost of higher rents.25

Any attempt to redistribute partially ex post those rents, as in the last schemes of the Federal Communications Commission, is bound to decrease the incentives for effort supposedly provided by a price cap.

The second main consequence of giving up the instrument of cost observability is that pricing rules, being then the only instruments to extract the rent of asymmetric information, are strongly affected by incentives.

Note finally that an advantage of a price cap is that the choice of prices can be decentralized to firms if the right weights $(w_k)$ and the right maximal level $(\pi)$ are chosen in the formula:26

$$\sum_k w_k p_k \leq \pi.$$  

### 3.3. Auctions of Incentive Contracts

Auctioning the right to be the natural monopoly has been early proposed (Demsetz (1968), Posner (1972)) as a way to have both efficiency and limited rents.

Fixed costs are presumed here to be so high that only one firm should be used.

Suppose that the technology of firm $i$ is described by

$$(26) \quad C^i = C(\beta^i, e^i, q^i) \quad (i = 1, \ldots, I)$$

where it is common knowledge that the $\beta^i$'s are i.i.d. random variables drawn from the same distribution $F(\cdot)$. Moreover $\beta^i$ is private information of firm $i$ and let $\beta = (\beta^1, \ldots, \beta^I)$.

---

25 It has sometimes been argued that the price cap was higher than the monopoly price (AT&T case).
26 See also Vogelsang and Finsinger (1979).
The optimal revelation mechanism is obtained under the assumption of Bayesian-Nash behavior. The incentive constraints are translated into constraints on expected rents \( U^i(\beta^i) \) which are

\[
U^i(\beta^i) = -E_{\beta^i} \left\{ x^i(\beta) \psi'(e^i(\beta)) E_{\beta^i}(\beta^i, C(\beta^i, e^i(\beta), q^i(\beta)), q^i(\beta)) \right\}
\]

where \( x^i(\beta) \) is the probability that firm \( i \) wins the auction, i.e. is selected to produce.

Expected welfare is

\[
E^i \left\{ \sum x^i(\beta) \left( S(q^i(\beta)) + \lambda p(q^i(\beta)) \right) q^i(\beta) \right\}
\]

\[-(1 + \lambda) \left( C(\beta^i, e^i(\beta), q^i(\beta)) \right) + \psi(e^i(\beta)) \right\} - \lambda \sum i U^i(\beta^i) \right\}.
\]

Solving (27) with the individual rationality constraint for the rent \( U^i(\beta^i) \) and inserting in (28) we obtain the maximization program

\[
\max_{(x^i(\cdot), q^i(\cdot), e^i(\cdot))} E_{\beta^i} \left\{ \sum x^i(\beta) \left( S(q^i(\beta)) + \lambda p(q^i(\beta)) \right) q^i(\beta) \right\}
\]

\[-(1 + \lambda) \left( C(\beta^i, e^i(\beta), q^i(\beta)) \right) + \psi(e^i(\beta)) \right\}
\]

\[-\lambda \frac{F(\beta^i)}{f(\beta^i)} \psi'(e^i(\beta)) \right\} \right\}
\]

subject to

\[1 - \sum x^i(\beta) \geq 0 \quad \text{for any} \ \beta,\]

\[x^i(\beta) \geq 0 \quad \text{for any} \ \beta, \text{any} \ i.\]

Maximization with respect to \((q^i, e^i)\) shows that, under appropriate concavity assumptions, \((S(q) + \lambda p(q)q - (1 + \lambda)(C(\beta, e, q) + \psi(e)) - \lambda(F/f)\psi'(e)E_{\beta}(\beta, C(\beta, e, q), q))\) concave in \((q, e)\), the optimal production and effort levels are determined for firm \( i \) as in the case of a single firm.

Maximizing with respect to \((x^i(\cdot))\) shows that the firm with the lowest \( \beta^i \) is selected if

\[-(1 + \lambda)C_{\beta^i} - \lambda \frac{d}{d\beta^i} \left( \frac{F(\beta^i)}{f(\beta^i)} \right) \psi'(\beta^i) - \lambda \frac{F}{f} \psi'(E_{\beta^i} + E_{\beta^iC}(\beta^i)C_{\beta^i})\]

is negative (which is true in particular for \( \lambda \) small enough).

An interesting separability obtains. The contract obtained by the winning firm implements the same quantity and effort levels as if the winning firm was alone.
facing the regulator. The gain of the auction is first to select the firm with the highest intrinsic efficiency and second to pay lower rents.\(^{27}\)

This second point can be particularly well illustrated by noticing that the optimal regulation can be implemented also in dominant strategy by a Vickrey-like auction.\(^{28}\) The firm is asked to announce his willingness to pay to be selected and be allowed to choose a contract in the menu of contracts usually offered to a single firm, knowing that it will only pay the second highest bid.

The firm's dominant strategy bid is the level of the rent it would obtain in a monopoly situation, i.e. from (27)

\[
\int_{\beta'}^\beta \psi' \Psi \, d\beta
\]

but it will only pay

\[
\int_{\beta'}^\beta \psi' \Psi \, d\beta
\]

where \(\beta^i\) is the second lowest cost parameter.

The net gain of the winning firm is then

\[
\int_{\beta'}^\beta \psi' \Psi \, d\beta.
\]

The auction enables the regulator to truncate the size of the asymmetry of information from \([\beta, \bar{\beta}]\) to \([\underline{\beta}, \beta^i]\) where \(\beta^i\) is the second lowest cost parameter.

**Proposition 3**—The Separability Property (Laffont-Tirole (1987)): *Auctions enable the regulator to choose the most efficient firm and to pay a lower informational rent than if it was facing only the winning firm. A separability property obtains: the winning firm chooses the same effort level and quantities as if it were alone.*

As the number of firms grows, \(\beta^i\) goes to \(\beta\) with probability one as well as \(\beta^j\). Asymptotically, the auction solves both the moral hazard and the adverse selection problems. By selecting a firm close to the most efficient type \(\beta\) it implements the contract with no efficiency distortion and as \(\beta^j\) is also close to \(\beta\) it pays a negligible rent to elicit information.

Several questions deserve further analysis. When the contract is awarded for a long period, it is desirable to be able to benefit from the possibility of technological progress brought about by future potential entrants. The simple

---

\(^{27}\) See Riordan and Sappington (1987) and McAfee and McMillan (1986) for similar results.

\(^{28}\) See Mookherjee and Reichelstein (1992) for a more general characterization of regulatory environments for which the optimal Bayesian mechanism is implementable by dominant strategy mechanisms.
solution suggested by the Chicago school of reauctioning the contract every year, say, faces the difficulty of inducing under-investment in specific nonobservable assets as stressed by Williamson (1976). Optimal dynamic regulation involves an optimal specification of break-out rules which favor in general the incumbent to encourage those investments (see Laffont and Tirole (1988a)).

Auctions presume noncooperative behavior between firms. This assumption is somewhat naïve once it is realized that an auction is a deus ex machina aimed at extracting a maximal surplus from firms. A natural reaction of those firms is to protect themselves by collusion. Further analysis requires here a major breakthrough of contract theory. The study of collusion under incomplete information is a major open problem.

Next, we have taken here an extreme case where selecting a single firm was clearly optimal. More generally the market may be split between different firms for various reasons (including differentiated products, decreasing returns to scale) and optimal regulation entails an appropriate sharing of the markets (see the literature on split-award auctions as well as Auriol and Laffont (1992), and Riordan and McGuire (1991)).

Finally, introducing correlation between the firms' characteristics opens the still largely neglected domain of yardstick competition (Shleifer (1985), Auriol and Laffont (1992)).

3.4. Dynamics and Commitment

An important issue in the current price cap debate concerns the length of the regulator's commitment. Within the traditional rate of return regulation a major question is the extent of the regulator's commitment to allow a fair rate of return. In industries with long run investments these questions are essential. In particular in the electricity sector the ability of the industry to invest in the most efficient technologies is probably the most important question: For example, it is a serious concern to know if the current investments to produce electricity from gas turbines in the USA is the optimal investment choice or if it is the outcome of a bad regulatory system which has lost any credibility to guarantee reasonable rates of returns for long run investments such as those required by nuclear energy.

A good understanding of these questions obviously requires a good grasp of the role of commitment in dynamic contracts. Economists of regulation have provided results on this topic which are of general interest.

3.4.1. Dynamics with Commitment

Consider the model of Section 3.1 repeated twice. The cost of the public good in period $\tau$ is

$$C_{\tau} = \beta - e_{\tau} \quad \text{with} \quad \tau = 1, 2.$$
If $\delta$ is the common discount factor, the firm's utility level is
\[ U = t_1 - \psi(e_1) + \delta(t_2 - \psi(e_2)) \]
and consumers' welfare is
\[ S(1 + \delta) - (1 + \lambda)(t_1 + C_1 + \delta(t_2 + C_2)). \]
Social welfare is accordingly
\[ W = S(1 + \delta) - (1 + \lambda)[\beta - e_1 + \psi(e_1) + \delta(\beta - e_2 + \psi(e_2)) - \lambda U]. \]
Optimal regulation under complete information entails
\[ \psi'(e_1) = \psi'(e_2) = 1 \quad \text{or} \quad e_1 = e_2 = e^*, \]
\[ t_1 = t_2 = \psi(e^*). \]

Let us now characterize the optimal regulation under incomplete information about $\beta, e_1, e_2$ under various commitment assumptions.

If the regulator can commit to a two-period mechanism, we obtain the following proposition:

**Proposition 4**—The Optimal Long Run Contract (Baron-Besanko (1984b)):29

The optimal contract with commitment is the repetition of the optimal static contract.

It is optimal for the regulator to commit not to exploit the information acquired at the end of period 1. The best way to arbitrate between the efficiency costs of effort distortions and the informational rents is to have "small" distortions each period.

Note that in period 2 the effort level of the inefficient firm $e_2$ is below the full information level $e^*$. It is then common knowledge that such a contract is Pareto inefficient.

### 3.4.2. Short Run Commitment

Suppose that, because of political constraints, the regulator cannot commit to period 2. The firm then knows that if it reveals its type in period 1, it will be put at its individual rationality level tomorrow.

The regulator offers a first period contract $t_1(C_1)$ in period one. If the contract is accepted, then, on the basis of the observed cost, the regulator offers a second period contract $t_2(C_1, C_2)$. We are interested in the Perfect Bayesian Equilibria of the game induced by these offers.

The first feature of equilibria is that firms never reveal themselves completely in the case of a continuum $[\beta, \bar{\beta}]$ of types. We obtain the following proposition.

29 See also K. Roberts (1983).
Proposition 5—No Separation Result (Laffont-Tirole (1988b)): For any first period scheme $t_1(\cdot)$, there exists no fully separating continuation equilibrium.

The intuition for this result, strikingly different from the case of commitment, is as follows: Suppose on the contrary that firm $\beta$ reveals itself in period 1. In period 2 it has no rent and therefore chooses its optimal first period contract. By slightly underestimating its efficiency the firm does not lose anything to the first order (since it was choosing its best contract by the envelope theorem) and gains a second period rent proportional to its underestimation.

The family of equilibria is actually quite complex and is not limited to the partition equilibria one could have expected from Crawford and Sobel (1982), for example. The reason comes from the existence of the take the money and run strategy which can be explained simply in the two state case. As the regulator increases payments for a low cost in period 1 to induce the efficient type to select a low cost, the inefficient type is tempted to select this contract. If it does, in period 2 the regulator believes that he is facing an efficient type and is very demanding; but, then, the inefficient type rejects the second period contract. Technically, the difficulty is that incentive constraints can be binding in both directions and not only upwards as in the static model.

The equilibria exhibit the ratchet effect, i.e., information unfolds slowly, but the complexity of equilibria defies further characterization. Despite the complexity of equilibria one can obtain the following intuitive result in the case of two types $\{\beta, \bar{\beta}\}$. In an equilibrium with a lot of pooling in the first period, the inefficient type's effort may be set above the commitment level. The reason for this result is that the efficient types often pool with the inefficient types. This would mean a large departure from efficiency of effort for the efficient type if the effort required from the inefficient type was already low. To limit this inefficiency the effort of the inefficient type is set very high. In the case of two types $\{\beta, \bar{\beta}\}$, the following intuitive results can be proved. For a discount factor $\delta$ small enough, the first period contract becomes separating. As $\delta$ becomes large, the first period contract converges to pooling without ever reaching it.

3.4.3. Renegotiation-Proof Commitment

As already mentioned above optimal contracts with commitment are not renegotiation-proof (see Dewatripont (1986)). It is possible for both parties to gain by renegotiating the initial contract when the firm turns out to be inefficient.

When, for institutional reasons or more explicit fears of collusion, third parties cannot be invoked to justify the ability to commit to inefficiencies, it is reasonable to assume that it is possible to commit except that it is not possible to commit not to renegotiate. The next result characterizes in the 2-type case the sequence of contracts $t_1(C_1), t_2(C_1, C_2)$ such that $t_2(\cdot)$ is not renegotiated.
when, after observing $C_1$, the regulator can make a take or leave it renegotiation offer.

**Proposition—The Optimal Renegotiation-Proof Menu of Contracts** (Laffont-Tirole (1990b)): In the case of two firm types the regulator offers the firm a choice between two contracts. The first one is a long-term contract that yields the efficient cost in both periods. It is picked only by the efficient type and is a long-term fixed-price contract. The second contract is such that both types produce at the same cost level in the first period (i.e., the efficient type randomizes between the two contracts) and is conditionally optimal given the posterior beliefs in period two. This second contract is therefore like a short-term contract with no further commitment.

The contract described in Figure 2 is indeed renegotiation-proof: at the end of period 1, in $A$, it is known that the firm is efficient and the efficient cost $\beta - e^*$ is required in period 2. In $B$, the posterior is $\nu_2(x)$, and the contract offered is the best contract the regulator can choose given his beliefs.

A separating contract would clearly be possible. The best one for the principal would consist of a fixed price contract given at the end of the two periods with a payment $(1 + \delta)(\beta - e^* + \psi(e^*))$. Both types would accept this contract and choose efficient levels of effort leading to costs $\beta - e^*$ and $\bar{\beta} - e^*$ for the efficient and inefficient type. This contract is clearly renegotiation-proof. However it is very costly for the regulator who can decrease his payments by inducing some pooling in period 1.

![Figure 2](image-url)
The rent he has to give up in the top branch is equal to the rent the efficient type obtains by deviating to the bottom branch. The best choice \( x \) of randomization for the regulator is obtained as follows: the contract following \( B \) is the best static contract conditionally on the posterior \( \nu_2(x) \); the contract \( C_1(x) \) is the best first period pooling contract knowing that the efficient firm pools with probability \((1-x)\nu\) and the inefficient firm pools with probability \(1-\nu\); \( x \) is then selected to arbitrate optimally between the inefficiencies associated with \( C_2(\nu_2(x)) \) and \( C_1(x) \) and the rents given up to the good type.

The great simplification with respect to the short-run commitment case is the much simpler structure of equilibria. This is due to the elimination by commitment of the take the money and run strategy: consequently, incentive constraints are binding as in a static problem. The comparative statics in \( \delta \) is as in the noncommitment case for the two types: for \( \delta \) small enough the first period contract is separating and, as \( \delta \) goes to infinity, the first period contract converges to pooling without ever reaching it. Finally, in the case of a continuum of types, separating equilibria exist but are never optimal.\(^{30}\) The characterization of optimal contracts in this case is an open problem.

3.5. Hierarchies and Regulation

The analysis of hierarchies is a natural next step in the study of organizations following the advances of the principal-agent literature. Little has been written yet on the new problems raised by, say, three level hierarchies.\(^{31}\) The study of regulatory capture has contributed to some better understanding of issues in contract theory which remain largely unsolved.

Two approaches to the analysis of public intervention in the industry can be distinguished, the public interest approach and the interest-group or capture approach. The new theory of regulation offers a framework to reconsider the question of capture by stressing the role of the regulatory agency as a collector of information for the State. Then the agency can favor an interest group by concealing information.

Consider a hierarchy composed of a State, an agency and a firm. The State is a utilitarian social maximizer but suffers from an asymmetry of information about the firm's cost function \( C = \beta - e \) with \( \beta \in \{\beta, \bar{\beta}\} \). The agency's role is to partially bridge the informational gap by making an observation \( \sigma \) which can either reveal the true value of \( \beta(\sigma = \beta) \) or reveal nothing (\( \sigma = \emptyset \)). When the agency discovers the value of \( \beta \) it obtains verifiable information and therefore

\(^{30}\) See also the literature on dynamic contracts with moral hazard: Rogerson (1985), Rey and Salanié (1990), Malcomson and Spinnewyn (1988), Fudenberg, Holmstrom, and Milgrom (1990), and the survey paper Chiappori, Macho, Rey, and Salanié (1992). These papers have clarified the role of credit markets and have shown that the most serious issues for the dynamics of contracts arise from adverse selection rather than moral hazard per se (see, however, Fudenberg and Tirole (1990), Ma (1991), Hermalin and Katz (1991)).

\(^{31}\) See however Tirole (1986), Laffont (1988).
has no discretion of claiming it is $\beta$ rather than $\bar{\beta}$ or $\overline{\beta}$ rather than $\beta$. However it can claim it has observed nothing, i.e. make a report $r = \emptyset$ instead of $\bar{\beta}$ or $\beta$.

If the agency is benevolent it will always reveal the truth, but, if not, we must worry about the ability of the firm to capture the agency, i.e. to bribe the agency for hiding information. When the State is informed about $\beta$ it can regulate the firm without inefficiency as explained above. The efficient effort level is imposed by a fixed price contract which leaves no rent. On the contrary when $r = \emptyset$ (so the State remains uninformed about $\beta$), we know that optimal regulation prescribes a rent to the efficient $\beta$ and too low an effort level $\bar{e}$ for the inefficient type $\overline{\beta}$ and no rent. Therefore the efficient type has a stake in convincing the agency to withdraw the information it has obtained. This enables the firm to secure a rent denoted $\Phi(\bar{e})$ which depends on the level of effort required from the inefficient firm.

The State has two ways to fight this type of collusion. One is to provide direct incentives to the agency which are greater than any bribe the firm might be willing to offer. If $(1 + \lambda_f)$ is the cost of side transfers between the agency and the firm and $S$ is a monetary payment made to the agency only when it reveals that $\beta = \bar{\beta}$, the collusion-proof constraint is

$$S \geq \frac{\Phi(\bar{e})}{1 + \lambda_f}.$$

A first result in this set up is that optimal regulation entails no collusion. The general sufficient conditions for such a result are still unknown. There are circumstances where it is better to let some collusion happen.

The second instrument for fighting collusion is to suppress or at least decrease the stakes of collusion to limit the costly incentive payments to the agency. Here it is desirable to decrease $\bar{e}$ in order to decrease $\Phi(\bar{e})$, i.e., the incentive contract provides less powerful incentives for the inefficient firm. The intuition is clear. The more powerful an incentive scheme is, the more rent it provides to efficient types, the greater the stakes of collusion for those. We obtain Proposition 7.

** Proposition 7—Collusion-Proof Regulation (Laffont-Tirole (1991)):** When the regulatory agency can be captured, optimal regulation yields a lower social welfare. The regulatory response consists of creating a costly incentive scheme for the agency and choosing a less powerful incentive scheme which yields less profit and worse cost performance.

This framework can be extended to consider multiple interest groups. Its advantage is to model the stakes of collusion as well as the regulatory response and therefore the observable variables as functions of fundamental economic data. It is therefore amenable to welfare analysis as well as empirical analysis.

Regulatory capture can take many forms and lead, for example, to undesirable restrictions on entry or favoritism in auctions. Clearly this notion of capture
can be extended to any type of intermediary agent and provides an agency approach to many problems of corruption.

A major weakness of this approach is the black-box treatment of side transfers. It is implicitly assumed that the firm and the agency can write contracts despite their illegal nature. Only a dynamic analysis could provide a firmer foundation for this assumption by appealing to reputation arguments.\\footnote{See Tirole (1992) for a survey on collusion in organizations.}

4. EMPirical IMPLICATIONS

The main empirical advantage of the theories I have described is that they provide a structural approach to model and then estimate various second best situations in a field which has relied so far on nontransparent reduced forms. It also enables us to model the inefficiencies left so far unmodelled in error terms such as in the econometrics of production frontiers.

I will take two examples to illustrate why and how the new economics of regulation can provide an impetus to the econometrics of industrial organization and later, hopefully, to economic policy.

4.1. Strategic Equilibria in Auctions

A substantial number of economic contracts are the outcomes of auctions. There is a large variety of auction rules: first price and second price auctions with announced or secret reservation price are the best known, but there are many others involving participation fees, compensation payments to losers, etc. The object to be contracted can be multidimensional, involving many characteristics that must be somehow aggregated, and the payment mechanism can also be complex since it may involve sharing rules agreements in which, for example, the principal and the agent share cost overruns.

There is clearly no well defined procedure to select auction rules in practice. Empirical analysis based on a clear theoretical vision should be able to improve the design of auction rules. It is remarkable that despite the large data sets available even the simplest cases have not been empirically explored using the full implications of economic theory (with a few exceptions; see in particular the work of Paarsch (1991, 1992), Donald and Paarsch (1991)).\\footnote{There is a well known empirical literature associated with the names of Hendricks and Porter which has used some implications of auction theory in reduced form models. See also the experimental work of Chen and Plott (1991).}

Recently, Laffont, Ossard, and Vuong (1991) has proposed to analyze descending auctions in agricultural products with a simulation-based econometric method which tackles successfully the large nonlinearities derived from contract and game theory. Surprisingly it was the first general method and application analyzing auction data within an econometric model based on the full restrictions of Bayesian Nash equilibria of the game defined by the auction.
It is interesting to note that in the case of auctions the heterogeneity needed for estimation is provided by the variable of asymmetric information itself. This will be a common feature of models with asymmetric information which will combine in general the random variables specifying the asymmetries of information with the more traditional errors of econometrics. Such combination raises identification issues which have been addressed in Laffont and Vuong (1993).

It is too early to ascertain the usefulness of the restrictions imposed by game theory for prediction purposes, but, as Laffont, Ossard, and Vuong (1991) shows, the estimated models track fairly well the highly variable winning bids. By extending the theory to auctions or repeated auctions of contracts as explained above, the new economics of regulation has provided a large set of testable implications which can enrich the fairly mechanical econometrics of second sourcing and take over bids. For example, in a repeated auction with commitment where cost is affected by both adverse selection and moral hazard, the incumbent's performance should be positively correlated with the slope of his incentive scheme and with the cancellation fee if second sourcing occurs; it should be negatively correlated with the probability of second sourcing, the intertemporal increase in the slope of the incumbent's incentive scheme, or the entry fee (see Laffont and Tirole (1988a)).

I see no serious difficulty in generalizing the estimation methods described above for dealing with the more complex auctions of regulated activities, such as local CATV markets as described in Zupan (1989).

One additional difficulty will often come from the unspecified objective functions of the buyers, who aggregate in a discretionary way the various dimensions of the contracts. The econometric work will have to estimate also the implicit weights used by the seller in its aggregation procedure. In practice most auctions are asymmetric and dynamic. The asymmetries prevent the explicit derivation of equilibrium strategies. Numerical integration of the differential equations defining strategies is needed (see Laffont, Loisel, and Vuong (1994)). The dynamic character of auctions raises more serious problems. An auction is essentially an adverse selection contractual problem with several agents. We have seen in Section 3.4 how complex equilibrium strategies are even in a one-agent dynamic situation.

As another example of potential application, in countries where favoritism in auctions may trigger appeal procedures, the capture approach can provide explanatory variables for the number of successful appeal procedures (see Marshall, Meurer, and Richard (1989)).

**4.2. Regulation with Asymmetric Information**

In practice regulatory rules are subject to various kinds of constraints beyond the informational constraints studied here: political constraints, transactional constraints, in particular. It is therefore quite unlikely that the theory will be

---

34 This is done in Laffont, Oustry, Simioni, and Vuong (1994).
able to compare the normative performance of various second best regulations such as Schmalensee's (1989) "good regulatory rules" versus the FERC's price cap cum profit sharing rules. Simulations performed on calibrated models can help us in this task. Gasmi, Ivaldi, and Laffont (1991) compares in the rent-consumers' surplus space the achievements of various regulatory rules for a variable weight of the firm's rent in the social welfare function. The main findings are that pure price cap regulation leaves a substantial rent to the monopoly relative to Schmalensee's linear rule, to the profit sharing rule, and to the optimal mechanism. Introducing room for downward price flexibility improves the efficiency of a price cap beyond Schmalensee's linear rules. By correcting in part the distributional distortion of a price-cap, the profit-sharing mechanism often yields levels of welfare comparable to those obtained by the optimal regulation.

Beyond these simulations little has been done empirically.

Wolak (1991) is to my knowledge the first econometric study making use of structural restrictions derived from the new economics of regulation. He analyzes data concerning the privately-owned and regulated water utilities in California. Regulation is based on three variables, capital invested, the access fee that each consumer must pay, and the marginal price of water. There is no direct transfer from the regulator to the firm, but this role is played by the access fee which, for simplicity, is assumed to discourage no consumer. Under complete information regulation entails marginal cost pricing and marginal productivity of capital equal to the interest rate which lead to the first best. Under incomplete information those prices are distorted as in a Baron-Myerson (1982) model because no use is made of ex post observed costs.

The paper then estimates the structural forms obtained from the first-order conditions of optimal regulation restricted to the instruments described above with and without asymmetric information about labor efficiency. The estimation method combines a non-parametric estimation of the distribution characterizing the asymmetric information with a parametric estimation of the other parameters of the cost function. A nonnested test concludes with the superiority of the model with asymmetric information which also leads to slightly decreasing returns to scale. The model without heterogeneity of labor efficiency obtains unrealistically high increasing returns to scale.

I believe that this paper is the first in a long series of applied works which will renew the econometrics of regulation with the help of the new theory of regulation. Wolak's paper proves the very good point that "the complete regulatory process should be modeled in order to guarantee consistent parameter estimates of the underlying production technology." However, a better description of the regulatory process should take into account the ex post observability of costs and the dynamic dimensions of regulation.

---

35 See also the experimental work of Cox and Isaacs (1987).
36 This empirical finding confirms the theoretical predictions made in Feinstein and Wolak (1990). It is particularly disturbing given that the whole notion of natural monopoly itself is based on the alleged existence of increasing returns.
5. CONCLUDING REMARKS

Research directions are multiple. Let us distinguish between research topics which can be pursued within the traditional complete contracting paradigm and those which require more fundamental breakthroughs.

A first line of research is to broaden the scope of the new economics of regulation by applying the approach to questions such as environmental regulation, the economics of research and development, consumers' protection, strategic regulation of international trade, regulation of financial activities, etc. Another extension concerns the regulation of oligopolies with controlled entry such as hospitals and generation of electricity or long distance telecommunications in some countries. When should the regulation of entry be given up and the control of a sector be reduced to favoring competition by the anti-trust policy? Regulation is one dimension of public economics. What should be the intersection of regulation and taxation, in particular optimal income taxation? What should be the contents of taxation by regulation (Posner (1971))?

The incentive schemes of regulation have been to a large extent developed separately from the auditing procedures. A better integration of both instruments would be welcome as it has been partly done in the optimal income tax area. A related issue is the maintained assumption of costless communication. The integration of incentive theory and constraints on communication is progressing slowly. The economics of regulation as any other application of incentive theory requires much more in this direction.

Most of the literature on regulation has been developed with little attention paid to financial constraints. As the experience of Eastern Europe clearly shows, the imperfections of capital markets require drastic changes in procedures such as auctions and more generally in incentive schemes when agents have strict budget constraints. The regulator is then quickly transformed into a banking institution and the deeper question of the reasons for a dual regulation by separate banks and regulatory commission arises.

Similarly a better positive political economy of regulation requires an explanation for the noncomprehensive character of regulations defined at the constitutional level.

The new economics of regulation is sometimes criticized for its unrealistic Bayesian description of asymmetric information. Even though I share partially this view, I am convinced that there is no alternative meaningful theoretical approach. It is a necessary first step which should be completed in various directions: absence of common knowledge, manipulability of beliefs, robustness, etc.

---

37 See Baron and Besanko (1984a), Crémer and Khalil (1992), Laffont and Tirole (1993, Ch. 13) for a start.


39 Difficulties similar to those encountered when risk aversion of regulated firms is considered will arise.

These last directions of research clearly point towards the more realistic world of incomplete contracting.

No clear methodology is available yet. Two approaches can be distinguished. The first one can be referred to as the nonverifiability paradigm.

Within a given model some variables are assumed ex post observable by both contracting parties but nonverifiable by third parties at a reasonable cost and contracts contingent on those variables are excluded. Institutions such as property rights appear then as imperfect substitutes to these missing clauses, very much like judicial rules for breach of contracts were analyzed in the traditional law and economics literature.

This trend of the literature faces two difficulties: one is the accusation of ad hocery. The selection of those nonverifiable variables gives an enormous degree of freedom to theorists and a substantive body of empirical evidence should always be required to justify the selection of the nonverifiable variables. The other difficulty is that, either by the use of incentive contracts, or by reputation arguments in repeated relationships, or by the construction of sophisticated contracts involving third parties, it is in general possible to develop more powerful contractual forms. The limit put on these contractual forms by simply excluding any dependence on nonverifiable variables appears often excessive. Unfortunately, no satisfactory theory of what are the appropriate limits is available today. The definition of such limits will depend in particular on a missing theory of collusion under incomplete information. It is likely that such a theory is bound to be history dependent through the role of beliefs and, therefore, strongly affected by institutional details (Green and Laffont (1992)).

Besides this important but still unclear paradigm other approaches are developed. They are ad hoc in the sense of taking as given some features of the contractual relationships that we often observe. Limits on commitment as analyzed above are an example of this approach. Clearly, an explanation of the optimal length or degree of commitment would be welcome but it is not uninteresting to study the consequences of such limits. Similarly, the emerging literature on multiple principals does not explain why the various principals are not coordinated by a central regulator, but the externalities created by their decentralized choice of contracts are worth understanding (Martimort (1991), Stole (1990), Laffont and Tirole (1993, Ch. 17)). Also, it may be interesting to study the working of a dual regulation by banks providing credit and a regulatory commission prohibited from using transfers, before understanding more deeply the reason for this dual regulation.

Let me conclude on a personal note derived from my recent experience with the world of regulation in Europe. On the one hand the justified requirements

---

41 A reputation argument can also be used in the complete contracting framework to eliminate incentive constraints. Here the incentives to build reputation are stronger since no use at all of the nonverifiable information is made.

42 A better integration of accounting costs might provide the link between the traditional approach and the nonverifiability approach.
for rigor imposed on economic theory have led to a dramatic neglect of political constraints in economics. On the other hand social choice theory has become so specialized that it has been gently pushed out of mainstream journals. The constraints put on the allocation of resources by political systems and in particular by the functioning of our democracies are enormous. In areas such as regulation, no hope for satisfactory explanations of what we observe exists without a better integration of political constraints. Our idealized normative models of public economics are a necessary first step. Without a second step we have little chance of being taken seriously. I think it is particularly damaging because our basic message for carefully taking into account incentives in social organizations is highly valuable and still needs to be taught.

*IDEI, Université des Sciences Sociales, 31042 Toulouse Cedex, France*

*Manuscript received May, 1992; final revision received August, 1993.*

**REFERENCES**


