HOW TO REGULATE AN UNPROFITABLE UTILITY: A SUBSIDY-CAP FOR THE URBAN TRANSPORT INDUSTRY

Andrea Boitani[a] and Carlo Cambini[b]
[a]Università Cattolica del Sacro Cuore, Milan, Italy.
[b]Dipartimento di Sistemi di Produzione e Economia dell'Azienda (DSPEA), Politecnico di Torino, Italy.
andrea.boitani@unicatt.it, carlo.cambini@polito.it

Abstract

The definition of appropriate regulatory policies for unprofitable public utilities is a difficult task. The incentive for unprofitable monopolistic firms, like a urban transport operator, to reduce its cost and increase revenue is low, particularly when the regulator has a reputation for not allowing the firm to shut-down even in the presence of increasing deficits. A possible incentive mechanism that can be used in practice – as it is still done in Italy and Norway – is to introduce a cap on subsidies. This paper analyses the efficiency and inefficiency properties of this new regulatory instrument, that can be called *subsidy cap*, and suggests a non-Bayesian *subsidy cap* mechanism based on the "menu of contracts" principle suggested by Bayesian incentive regulation theory. It is shown that the proposed mechanism, although non Bayesian, shares some desirable properties with more information demanding optimal Bayesian mechanisms and, in addition, provides incentives to the regulated firm to reduce costs and increase the revenue to cost ratio.

Keywords: Urban transport; Local finance; Incentive regulation
Topic area: H7 Reulation / Liberalization

1. Introduction

The Italian highly subsidised local public transport firms – either publicly or privately owned - are good examples of the soft budget constraint disease (Segal, 1998). Historically all these firms enjoyed monopoly protection by means of non-tendered concessions. The financial performance of these firms has deteriorated for more than thirty years: on average, external subsidies represent over 70% of operating costs[1]. Financial distress is only partly explained by declining patronage (lower shares in the private – public transport split) and fares permanently lower than average costs. An important role is also played by low and stagnant productivity, due to weak incentives for efficiency enhancing effort. Weak incentives, in turn, are not surprisingly related to cost-plus contracts, based on individual negotiations between local governments and the (local) monopoly firms. Incentives are even weaker when the firm is publicly owned and the local government can not credibly commit to let the firm go bankrupt in the presence of high and/or increasing deficits.

As convincingly argued by Crew and Kleindorfer (2002), the mechanism-design theory of regulation "promised efficiency as long as the regulator is prepared to allow information rents" (p. 12). The standard Bayesian regulatory mechanism establishes that less efficient firms must be allowed to choose a low powered incentive scheme; which implies that the imperfectly informed regulator must be willing to accept low efficiency (and high subsidies) in order to minimise the

---

[1] Remarkable differences in subsidies can be traced out between northern and southern regions: northern regions having higher than average revenue to cost ratios.

information rents needed to induce more efficient firms to correctly reveal their true "type" (Laffont, Tirole, 1993). "Theorists, however never understood the impossibility of this in practice […] Thus, the promise of X-efficiency was hedged with conditions, which […] make the theory of little significance for real world regulation" (*ibidem*).

This explains why high-powered price cap regulation is so widespread. But price cap regulation is also of little use when the regulated utility is unprofitable because the regulator himself or the local government chooses (for whatever reason) prices below average costs. In order to strengthen the incentives of unprofitable utilities, it seems reasonable to introduce a *subsidy cap*, that is a mechanism aimed at reducing the real value of subsidies over time.

Such a mechanism is indeed one of the provisions of the 1997 reform of Italian local public transport (D.Lgs. 422/97). According to the reform act, regions and local government remain free to set fares to fulfil "social" goals. However, public subsidies should not exceed 35% of operating costs and should decline over time in force of a subsidy cap explicitly aimed at increasing the *X*-efficiency of the industry. Despite the law, only few Italian local governments appear to have reverted to the subsidy cap.

The same mechanism was introduced in Norway in a menu where also yardstick competition and competitive tendering were listed as alternatives to replace individual negotiations. As documented by Dalen and Gomez-Lobo (2003) subsidy cap contracts rapidly outnumbered cost-plus and yardstick competition. The recent introduction of the subsidy cap mechanism "does not allow to make strong conclusions regarding their incentive properties" (p. 383). However, the estimates contained in Table 3 of Dalen and Gomez-Lobo (2003, p. 381) tell us that whilst firms regulated with a negotiated type contract had, throughout the sample period, costs almost 18% above the estimated cost frontier, those regulated with the yardstick were only 7,6% below the frontier, and those regulated with the subsidy cap had on average a cost inefficiency of 13%.

The present paper is addressed at exploring theoretically the efficiency and inefficiency properties of the *subsidy cap* and at suggesting two ways to overcome the cost-manipulability problem the mechanism proves to have. The second solution advanced, although non-Bayesian and easily implementable, is based on the "menu of contracts" principle suggested by the Bayesian incentive regulation theory (Laffont, Tirole, 1993) and tends to resolve the main problem in any cap-mechanism, that is how to fix the *X* factor.

The paper is organised as follows. Section 2 explores the technical efficiency properties of the subsidy cap mechanism. It is shown that the subsidy cap has the same efficiency properties of the price cap mechanism; in other words, contrary to the rate-of-return mechanism, no distortions on the use of capital factor is induced by the subsidy cap. However, this conclusion depends on how the initial transfer, to which the *X* factor is applied, is determined. If the initial transfer is fixed as equal to the loss in the initial period (i.e. the difference between operating costs and market revenues in the first period), the firm has an incentive to manipulate its initial cost in order to increase the initial transfer (section 3). This results confirms what Segal (1998) suggests about the actual behaviour of an unprofitable monopolistic firm. The opportunistic behaviour of regulated firms leads to low effort in cost-reducing activities. The possibility of cost manipulation induced by the mechanism suggests to develop higher powered incentive regulatory schemes in order to avoid such distortions. The problem tends to coincide with the fixing of the *X* factor.

Following Crew and Kleindorfer (1996), in section 4 a simple, non-Bayesian subsidy cap mechanism is developed, based on the "menu of contracts" principle suggested by Bayesian incentive regulation theory. Since one of the main problems with the local public transport industry is to reduce the huge deficits of firms, a regulatory mechanism is suggested that links the level of the *X* factor to the level market revenue to operating cost ratio. The proposed mechanism, although non Bayesian, shares some desirable properties with more information demanding optimal mechanisms, as the asymmetric information regulatory model developed by Laffont and Tirole (1986, 1993) or Baron and Myerson (1982). Finally, in section 5 some conclusive remarks and suggestions for further research are advanced.

2. Subsidy cap's technical efficiency properties

The first step of our analysis is the assessment of the X-efficiency properties of a regulation mechanism based on the *subsidy cap*. Following Leibenstein (1966) and Bös (1991, 1994), let $q(l, k)$ be the quasiconcave and differentiable production function of the transport firm, where q is output[2], l is labour and k is capital. Let w and r be the unit price of labour and capital, respectively. Suppose that $q^0 = q(l^0,k^0)$ is the quantity of transport services that the local authority wants the firm to produce and assume that this quantity is set in order to let the firm breaks even; a firm subject to a *subsidy cap* has to solve the following minimization problem:

$$\begin{aligned} &\min_{l,k} wl + rk \\ &s.t. \\ &q^0 \leq q(l,k) \\ &wl + rk - \overline{p}q(l,k) \leq t^{cap} \end{aligned} \qquad (1)$$

where t^{cap} is the *subsidy cap* and p is the fare per unit of service, as fixed by the local government. The latter assumption reflects the fact that in Italy, as well as in other European countries, fares for local transport services are not under the control of the firm and are set below the average cost of service for either distributional or environmental considerations. From a regulatory point of view, this implies that throughout the paper the only regulatory instrument considered is the transfer and not the price. From now onwards we shall also assume that in any and each period $p = p_{-1}(1+ RPI)$. This assumption is necessary to ensure that whatever the cap on subsidies, the revenue to cost ratio (R/C) does not decrease because of costs' increases brought about by an inflation rate (RPI) higher than the X factor[3].

Let $S = \{(l,k): q(l,k) \geq q^0, wl + rk \leq t^{cap} + \overline{p}q(l,k)\}$ and $S_1 = \{(l,k): q(l,k) \geq q^0\}$

Consider the problem:

$$\mathbf{P}_1: \ \min_{l,k} wl + rk \in S_1 \qquad (2)$$

It is straightforward to verify that (l^0, k^0) is the optimal solution for $\mathbf{P}_1$ if and only if it verifies the following Kuhn-Tucker conditions. In particular, the Lagrangian of $\mathbf{P}_1$ is:

---

[2] The model can be easy extended to the multiple-output case, differentiating for example between urban and inter-city services. Since the purpose of the paper is to analyse the incentive properties of the a subsidy cap mechanism that covers all the losses of the regulated firm and not to define optimal prices for each services, to simplify, but without loss of generality, we assume throughout the paper that the regulated firm offers only one public transportation service.

[3] Let $\tau = \frac{p \cdot q}{C}$. In logs one has $\ln\tau = \ln p + \ln q - \ln$. The growth rate is then: $\tilde{\tau} = \tilde{p} + \tilde{q} - \tilde{c}$. If pricesare constant and the quantity is fixed at the level requested by the local authority, $\tilde{\tau} = -\tilde{c}$. If $\tilde{c} \approx RPI - X$, then $\tilde{\tau} < 0$ if $RPI > X$. If, instead, $p = RPI$ and q is constant when relative prices are (approximately) constant, then $\tilde{\tau} = X$ and the revenue to cost ratio increases of X.

$$L = wl + rk + \varphi\left[q^0 - q(l,k)\right] \quad (3)$$

It results that $\varphi > 0$ and so $q^0 = q(l, k)$, i.e. the mechanism satisfies the technical efficiency property. From the F.O.C.s with respect to *l* and *k* we get:

$$\frac{\partial q/\partial l}{\partial q/\partial k} = \frac{w}{r} \quad , \quad q(l,k) = q^0. \quad (4)$$

Since $S \supseteq S_1$ and (l^0, k^0) satisfies the second constraint in (1) for assumption, that is $wl^0 + rk^0 \leq t^{cap} + pq(l^0, k^0) = t^{cap} + pq^0$, i.e. the local government sets the quantity in order to let the firm break even, then (l^0, k^0) is an optimal solution the problem in (1).

Analogously to the *price cap*, the *subsidy cap* mechanism is compatible with the *X*-efficiency properties of Leibeinstein (1966). Otherwise, if the transport firm is free to decide the quantity to offer, i.e. local government does not intervenes to set the quantity, the results partially change. In particular, since the objective function in (1) is linear in (*l*, *k*), it does not have stationary points and this implies that the optimal solution is binding at least one of the constraints. In particularly, it results that the optimal solution (l^*, k^*) is necessarily binding to the second constraint and so it satisfies the following Kuhn-Tucker conditions:[4]

$(1+\lambda)\, w = p\, \partial q/\partial l$; $(1+\lambda)\, r = p\, \partial q/\partial k$; $wl^* + rk^* = t^{cap} + p\, q(l^*, k^*)$.

where λ is the lagrangian multiplier of the second constraints in (1). Let us note that the optimal solution satisfies, once again, the first condition in (4) with $q(l^*, k^*) > q^0$. This means that the firm increases the quantity with respect to the level stated by local government but it does not distort the use of production factor. In conclusion, the subsidy cap mechanism induces the regulated firms to efficiently use its production factor and in addition it could give the incentive to increase the patronage. In practice, however, the real impact of a subsidy cap strictly depends on the way it is implemented. In particularly, the above results strictly depend a) on the extent to which the firm is able to manipulate the initial transfer it receives from the regulator; and b), on the determination (and quantification) of the *X* factor in order to give appropriate incentives to *X*-efficiency. In the following section we try to respond to the question (a), while we attempt at answering question (b) in section 4.

3. Incentives in cost reduction

3.1 The level of effort under cost-plus regulation

Let $C(e, q)$ be the total cost function of the transport firm which depends on the vector of intertemporal effort levels in cost reducing activities ($e = e(e0, e1, \ldots, en)$) – where the subscripts denotes the period in which the effort is exerted - and on the level of output (*q*); these efforts level are independent over time, i.e. $\partial e_{t+1}/\partial e_t = 0, \forall t = 0,\ldots,n$. Without loss of generality, we assume that the firm faces constant returns to scale, so that the total cost function can be rewritten as $C(e, q) = c(e)q$, where $c(e)$ is the unit cost, with $e_i \in [e_i^-, e_i^+], i = 0,..n$.

[4] In fact, if $q(l^*, k^*) = q^0$, then, from the second constraint, we have $wl^* + rk^* \leq t^{cap} + p\, q^0 < wl^0 + rk^0$, since total cost in (l^*, k^*) is the minimun; but this is absurd since min $wl + rk$, $(l, k) \in S_1 \leq$ min $wl + rk$, $(l, k) \in S$. Thus, it results that $q(l^*, k^*) > q^0$

Suppose that the regulatory lag is exogenously set (i = 0,1,...J), the price is indexed to the *RPI* (as seen above), and the effort level exerted in period i, e_i, can influence total cost in the same in period i and in the following periods.

With no subsidy, the firm's objective is given by utility:

$$U = \sum_{i=0}^{J} \frac{1}{(1+\delta)^i}\left[\bar{p}q_i(\bar{p}) - c_i(e_0, e_1, \ldots e_i)q_i(\bar{p}) - \psi(e_i)\right] \quad (5)$$

where p is the indexed service fare, $\psi(.)$ is the disutility of effort, with $\psi'(.) > 0$ and $\psi''(.) > 0$, $\psi(e_i)>0$ and δ is the discount rate. Assume $\partial c_0/\partial e_0 < 0$; $\partial c_0/\partial e_i = 0$, $\forall e_i \neq e_0$; $\partial c_i/\partial e_0 < 0$, $\forall i=1,\ldots J$. This latter assumption implies that a low effort in the first period increase the unit cost in the following periods.

Absent a regulatory mechanism and transfer, the *first best* effort level in period 0, e_0^*, is:

$$-\sum_{i=0}^{J} \frac{1}{(1+\delta)^i}\frac{\partial c_i}{\partial e_0} q_i(\bar{p}) = \psi'(e_0) \quad (6)$$

i.e. e_0^* is the effort level that makes the discounted flow of cost savings equal to the marginal disutility of effort. The same first best solution on effort applies in every period *
(i.e. $e_i = e_i$, = $\forall i$,..,1 n).

Consider now the opposite case in which the regulator and the regulated firm enter an *individual negotiation* - in the words of Dalen and Gomez-Lobo (2003) - which in turn implies the application of a cost plus contract: each year the regulator sets a transfer equal to the loss incurred by the firm, i.e. $T_i = c_i q_i(\bar{p}) - \bar{p} q_i(\bar{p}), \forall i = 0,.., n$. In this case, the optimisation problem of the regulated firm becomes:

$$\max_{e_i} U = \sum_{i=0}^{J} \frac{1}{(1+\delta)^i}\left[-\psi(e_i)\right] \Rightarrow \min_{e_i} \varphi = \sum_{i=0}^{J} \frac{\psi(e_i)}{(1+\delta)^i} \quad (5bis)$$

The application of a cost plus contract implies that utility maximisation of a regulated firm is *de facto* tantamount to the minimisation of the (discounted flow of) disutility of effort, hence the minimisation of the effort level: from (5bis) it is straightforward to show that the optimal effort level reaches always the minimum level, i.e. $e^*_i = e_i$, i = ,..0 n .

In the following paragraphs of this section we shall show how the objective function of the regulated firm changes according to the introduction of a cap on subsidies and consequently we shall show how the effort level varies in different scenarios.

3.2 Introducing a subsidy cap

How does firm's behaviour change when a cap on subsidies is introduced? In order to simplify the analysis, suppose that the retail price index and the X factor are constant over time: RPI_1 = RPI_2 = ...RPI_i = RPI ; X_1 = X_i = ...X . The subsidy cap mechanism can then be written as:

$$T_n = T_{n-1}(1 + RPI - X) \quad (7)$$

where Tn is the transfer received in period n.

In addition, let $T0$ be the "initial transfer", that is the subsidy received in the first year of application of the subsidy cap. Suppose that this transfer is fixed as equal to the loss in the initial period (i.e. the difference between the cost of service and the market revenue in period 0), i.e. $T_0 = c_0(e_0)q_0 - qp_0(p)$. This the simplest subsidy cap one can conceive: just a starting point to explore the mechanism. Under subsidy cap regulation, firm's utility is given by:

$$U = \sum_{i=0}^{J} \frac{1}{(1+\delta)^i}\left[\bar{p}q_i(\bar{p}) - c_i(e_0, e_1, \ldots e_i)q_i(\bar{p}) - \psi(e_i) + T_0(1+RPI-X)^i\right] \qquad (8)$$

Then, the optimal effort level exerted in period 0 is given by:

$$\frac{\partial U}{\partial e_0} = -\psi'(e_0) + \sum_{i=0}^{J} \frac{1}{(1+\delta)^i}\frac{\partial c_0}{\partial e_0} q_0(\bar{p})[1+RPI-X]^i - \sum_{i=0}^{J} \frac{1}{(1+\delta)^i}\frac{\partial c_i}{\partial e_0} q_i(\bar{p}) = 0 \qquad (9)$$

Rearranging, we get:

$$\underbrace{\left(\frac{\partial c_0}{\partial e_0}\right) q_0 \left(\sum_{i=0}^{J} \frac{1}{(1+\delta)^i}[1+RPI-X]^i\right)}_{\Omega} - \underbrace{\sum_{i=0}^{J} \frac{1}{(1+\delta)^i}\left(\frac{\partial c_i}{\partial e_0}\right) q_i}_{\Xi} = \psi'(e_0) \qquad (10)$$

Since $\partial c_i / \partial e_0 < 0$, $i = 0 \ldots J$, Ω denotes the negative effect that the effort exerted in period 0 has on future transfers, due to the reduction of the initial transfer $T0$ to which the cap mechanism is applied. Ξ represents the discounted flow of cost savings deriving from an increase in effort in period 0[5].

We can summarise the results above in the following:

Proposition 1: *In presence of a "pure" subsidy cap where the initial transfer is equal to the difference between cost of service and market revenue in period 0 – i.e. the first year of application of the regulatory contract between the firm and the regulator – the effort exerted by the firm is always lower than the first best effort. More precisely:*

- *if* $|\Omega| \geq |\Xi|$, *effort level in period 0 is always at its minimum,* $e0 = e0-$;
- *if* $|\Omega| < |\Xi|$, *the optimal effort level is higher than the minimum but lower than the first best,* $e^- \leq \tilde{e}_0 < e_0^*$

As a consequence, the structure of the initial transfer can generate negative incentives in cost reduction activities[6]. The results confirm what suggested by Segal (1998): a monopolistic firm that receives public subsidies from a benevolent regulator - i.e. a regulator that could be induced ex post to bail out an unprofitable firm – could make inefficient *ex ante* decisions, aimed at increasing losses in order to get increasing subsidies over time.

Note that if $\partial c_i / \partial e_0 = 0$, $\forall\ i \neq 0$, a corner solution is reached, that is $\tilde{e_0} = \bar{e}$. In fact, condition (9) becomes:

$$\frac{\partial U}{\partial e_0} = -\psi'(e_0) + \sum_{i=0}^{T} \frac{1}{(1+\delta)^i}\frac{\partial c_0}{\partial e_0} q_0[1+RPI-X]^i < 0 \qquad (9bis)$$

[5] The results obtained would only differ for a multiplicative constant would the firm minimise the disutility of effort under a budget constraint embedding the subsidy cap.
[6] Similarly, Sappington (1980) show how the *price cap* mechanism is also characterised by this kind of difficulties.

On the other hand, in any other period $i \neq 0$, the effort level is given by:

$$-\frac{\partial c_i}{\partial e_i} q_i(\overline{p}) = \psi'(e_i) \quad (11)$$

i.e. it is always the *first best* effort level; from period 1 the cap mechanism triggers an efficiency enhancing behaviour.

3.3 The use of past information to define a subsidy cap

Following Vogelsang and Finsinger (1979), a regulator can define a subsidy using past information about the regulated firm. In other words, a possible solution to the manipulation problem examined in the previous paragraph could be found by applying the *cap* not to the loss incurred in period 0, but to the loss incurred in period –1. For example, the regulator could propose a contract such that in period 0 the firm will be given a subsidy such that the gap between market revenue and operating costs is filled; in period 1 the subsidy cap mechanism will start to be applied, making the initial transfer equal to the loss in period –1.

With such a contract, the opportunistic behaviour is reduced and the effort in cost reducing activities correspondingly increases. This can be easily proved as follows. Given that $T_0 = c(e_0)\overline{q}_0 - \overline{p}_0\overline{q}_0$ and $T_1 = T_{-1}(1 + RPI - X)$, where $T_{-1} = c_{-1}(\overline{e}_{-1})\overline{q}_{-1} - \overline{p}_{-1}\overline{q}_{-1}$, the firm's utility function becomes:

$$U = -\psi(e_0) + \sum_{i=1}^{J} \frac{1}{(1+\delta)^i} \left[\overline{p}_i q_i - c_i(e_0, e_1 ... e_i) q_i - \psi(e_i) + (c_{-1}(\overline{e}_{-1}) q_{-1} - \overline{p}_{-1} q_{-1})(1 + RPI - X)^i \right] \quad (12)$$

Thus the optimal effort level in period 0 is given by:

$$\frac{\partial c_0}{\partial e_0} q_0 - \sum_{i=0}^{J} \frac{1}{(1+\delta)^i} \left(\frac{\partial c_i}{\partial e_0} \right) q_i = \psi'(e_0) \quad (13)$$

By comparison of condition (13) with condition (10) it turns out that *e*0 is higher with the second type of contract. Notice that the use of past information on costs and prices makes the suggested mechanism resemble the Vogelsang and Finsinger (1979)'s mechanism which is generally acknowledged to be easily implementable.

Another way to use past information to regulate the firm is to define a dynamic mechanism according to which the subsidy varies over time and is continuously revised according to the past performance of the firm. Let γ be the subsidy per unit of service (in terms of vehicles*km). Suppose that the regulator uses – in a myopic way like in Vogelsang and Finsinger (1979) - his past information regarding the level of service's quantity (q^{t-1}), traffic revenues (R^{t-1}) and costs (C^{t-1}). The subsidy cap can then be implemented in this way:

$$\gamma^t \leq \frac{C^{t-1} - R^{t-1}}{q^{t-1}} = \gamma^{t-1} + RPI - X$$

which means that the subsidy per unit of traffic must decrease over time, when $X \geq RPI$, and this is possible if and only if the regulated firm a) reduces its costs, b) can increase the service's fare beyond *RPI* and/or c) increases the service's quality and so the amount of vehicles *km "sold".

The regulatory scheme suggested in this section could create problems of commitment between the regulator and the firm, it does not completely eliminate the *rachet effect* and in addition it does not allow to deal with the main difficulty of cap-based regulation, i.e. how to define the *X* factor. The following section is devoted to this issue.

4. An alternative subsidy cap mechanism

The new theory of regulation has deeply focused on the problem of asymmetric information between the regulator and the regulated firm. The main problem the regulator has to face is how to induce regulated firms to reveal its information in order to increase resources' allocation in the market and reduce the information rent of the firm (Baron and Myerson, 1982; Laffont and Tirole, 1986 and 1993).

The main conclusion is – as Sappington and Weisman (1996) put it – "one contract does not fit all". Welfare could be increased were the regulator able to define a menu of contracts to the regulated firm, a menu from which the firm ought to select an incentive regulatory contract. In other words, leaving discretion to the firm and providing appropriate pecuniary incentive could lead the firm to reveal its private information.

The approach followed in this literature is defined as "Bayesian" (Laffont, 1994), since it is based on the theory of games under incomplete information. As already mentioned in section 1 of the present paper, despite the theoretical relevance of this theory, its prescriptions are far from being implementable. "Indeed, the entire mechanism design literature […] is based in one way or another on assumptions like common knowledge that endow the regulator with information that cannot have without a contested discovery process. […] Common knowledge is the Achilles heel of mechanism design theory" (Crew and Kleindorfer, 2002, p. 11).

In this section we propose a possible and implementable non-Bayesian regulatory mechanism rooted in the "menu of contract" principle typical of the Bayesian approach. Following Crew and Kleindorfer (1996), we suggest to deal with the problem of determining the *X* factor in the *subsidy cap* mechanism by defining a menu of contracts according to which the firm can self-select the *X* factor more appropriate to its characteristics. In the words of Acton and Vogelsang (1989, p. 9), we use the *X* factor as a "means to introduce in a non-Bayesian approach a Bayesian parameter".

As one of the main problem in the local public transport industry is reducing the huge deficits of transport firms, we suggest that in the menu of *subsidy cap* contracts the *X* factor is related to the traffic revenues to operating costs ratio of the firm (R/C). Let $\hat{\tau}$ be such a ratio, the regulator can define a menu of contract $\{T(\hat{\tau}), X(\hat{\tau}), \hat{\tau}\}$ such that the regulator sets the initial transfer $T(\hat{\tau})$ and the *X* factor according to the R/C ratio that the regulated transport firm decide to announced, i.e. declare as its – true or not - level of coverage.

Since transport fares are set by local governments and are not under control of the regulated firms, and so revenues can be increased only by increasing service quality, it is reasonable to assume that the higher is the initial transfer $\hat{T}$ asked by the firm, the higher should be the *X* factor. Given a certain level of revenue, a high transfer implies that the level of total cost is very high and so, increasing the *X* factor is justified in order to favour cost reductions. On the other hand, if the firm asks for a low initial transfer, this reveals that the firm is sufficiently efficient (compared to the previous case) and thus the regulator must give a reward to the firm lowering the *X* to zero or even below zero, i.e. giving a sort of grant to the efficient transport operator.

Following Crew e Kleindorfer (1996), a simple formulation of such a mechanism could be written as:

$$X(\tau) = X^{\max} - \frac{(X^{\max} - X^{\min})(\tau - \tau^{\min})}{\tau^{\max} - \tau^{\min}} \quad (14)$$

where X^{max} and X^{min} are the maximum and the minimum values of the X factor – fixed by the local regulator - respectively, and τ^{min} is the lower R/C ratio accepted by the local government.[7]

In general, we could think that, despite the regulator were benevolent, he could oblige the firm not to fall below a certain R/C ratio. As already noted in section 1, one of the provisions of the recent Italian reform of local public transports is that each firm must max reach a R/C not lower than 35%. Hence τ^{min} could be set equal to 35%; τ is the maximum R/C ratio, ideally equal to 100%, but possibly lower than this level.[8] Then, the mechanism in (14) implies that:

- If the announced R/C ratio is the lowest, i.e. $\tau = \tau^{min}$ – that is, the firm needs a huge amount of money to cover its loss - then the regulator sets $X = X^{max}$, in order to induce the regulated firm to increase its effort and to reduce costs and/or to try to gain more customers increasing the service's quality;
- If the announced R/C ratio is the highest, $\tau = \tau^{max}$, the level of the X factor is the lowest, $X = X^{min}$, near to zero or even negative; thus, the contract mimics a sort of fixed-price contract, which induces high effort whilst yielding information rents to the regulated firm.

With regard to the Italian local transport industry, we can assume $\tau \in$ [35%, 100%], X^{min} = 0% and X^{max} = 5%. The suggested mechanism leads to the menu of contracts illustrated in Table 1.

Table 1 – A possible menu of contracts

τ(announced)	***X***
τmin = 35%	*X*max = 5%
40%	4.6%
50%	3.85%
65%	2.7%
80%	1.55%
90%	0.8%
τmax = 100%	*X*min = 0%

It is better to stress the fact that prices and quantities of transport services are given and set by local government as it is usual almost in all Europe. In particular, they are declared ex ante before the bidding process or the designing of the contract.

The proposed mechanism is, of course, not optimal from a Bayesian point of view, since it is not derived from a welfare maximization process. In addition, this mechanism does not lead – as the optimal models do (Baron and Myerson, 1982; Laffont and Tirole, 1986, 1993) - to the

---

[7] The relation between the X factor and τcould also differ from the linearity assumption considered for simplicity in (14); the incentive properties of this mechanism depends only on the inverse relation between the two variables.

[8] How to exactly set τ^{max} level is a difficult operative task that local regulators have to face. The use of past information about the level of revenue to cost ratio of the firm or of other similar (in terms of traffic area) transport firms could give useful benchmarking to this purpose. Otherwise, these threshold levels could be defined using a sort of *yardistick* measure of the *R/C* ratio. In this case, the incentive effect of the proposed mechanism could be even reinforced.

truthful revelation of firms' private information. However, the well known difficulties in implementing the optimal mechanisms, may make our proposal attractive for practical purposes.

Another useful mechanism could be a sort of yardstick subsidy cap where the single parameters can be estimated using data from all the regional areas of a country (Dalen and Gomez-Lobo, 2003). Unfortunately, this instrument can be applied with success only when the economic conditions in different areas can be considered quite homogeneous in terms of level of demand, quality offered, geographical condition in the transport area, fares' level and so on, otherwise - as stated by Shleifer (1985) - the introduction of *yardstick competition* could be detrimental and lead to inefficient allocations.

It is important to stress that, despite its simplicity, our mechanism has some interesting properties, which are summarised in the following propositions.

Proposition 2: *If we assume that* $X^{max} > RPI$ *and* $X^{min} < RPI$*, then the mechanism in (14) induces the regulated firm to increase its R/C ratio above a certain minimum level* τ^{min}.

Proof. See Appendix.

Proposition 2 implies that once a minimum τ is fixed, a reasonable mechanism of the type shown in (14) may be implemented which provides incentives for the regulated firm to increase its R/C ratio, with a positive welfare effect due to the reduction in the local fiscal burden and, more generally in the use of public funds.

The increase in the R/C ratio derives essentially from a higher level of effort in cost reduction activities. In fact we can prove:

Proposition 3: *The menu in (14) leads to a higher effort level and so it is more efficient than the "pure" subsidy cap mechanism shown in section 3.*

Proof. τ is defined as the ratio between market revenue, *R*, which is constant as fares are set by the local government, and operating expenses. Given the cost function shown in section 3, we have: $\tau = R/c(e_i)q_i$, $i = 0,1\ldots$ J. Then, $\tau'(e) > 0$ and since, from (14), $X'(\tau) < 0$ it implies that $X'(e) < 0$. Suppose, for simplicity, that $\partial c_i / \partial e = 0$, $\forall\, i \neq 0$. Then, firm's utility is:

$$U = \sum_{i=0}^{J} \frac{1}{(1+\delta)^i} \left[\bar{p} q_i(\bar{p}) - c_i(e_i) q_i(\bar{p}) - \psi(e_i) + T_0 (1 + RPI - X(e_i))^i \right]$$

And the optimal effort at time 0 is:

$$\frac{\partial U}{\partial e_0} = -\psi'(e_0) + \sum_{i=0}^{T} \frac{1}{(1+\delta)^i} \frac{\partial c_0}{\partial e_0} q_0 \left[1 + RPI - X(e_i) \right]^i - T_0 X'(e_0) \; . \qquad (15)$$

Then, since the last term in (15) is positive, by comparing condition (15) with the expression in (9bis), it results that the effort level is unambiguously higher. *Q.E.D.*

The above mechanism could appear in contrast with the optimal incentive contracts defined by Laffont and Tirole (1993): according to these authors, an optimal regulatory mechanism implies that an inefficient type firm should face a low powered incentive contract while a highly efficient one should face a high powered contract. Intuitively, since the informational rent depends on the effort level exerted by the regulated firm, the regulator prefers to reduce those costly rents by lowering the effort level requested from the inefficient firm. That is, it is better - from the social point of view - to lower the incentive of the less efficient firm's type and raise the incentive of the more efficient one in order to minimize the public expenditure needed to provide incentives to the regulated firm.

The proposed mechanism goes in the opposite direction: since urban transport industry is characterised in many countries – and especially in Italy – by huge deficits and lack of efficiency, it is extremely important to introduce high powered incentives also for poorly efficient firms in order to limit subsidisation and increase the in the industry's efficiency.

Moreover, the empirical analysis of Dalen and Gomez-Lobo (2003) shows that in Norway "companies regulated under the subsidy cap system started out being very inefficient, which is reasonable considering that countries applied this type of regulation first were the ones that had not switched to a yardstick type contract earlier" (pg. 382). These firms have shown also strong efficiency gains from the introduction of the subsidy cap, meaning that the empirical efficiency properties of the mechanism are relevant. In other words, the empirical evidence shows that applying a high powered incentive contract, like the subsidy cap, also to inefficient firms can give really positive results in term of efficiency and reduction in the use of local public funds; our proposed mechanism reflects this empirical evidence.

5. Conclusions

The definition of appropriate regulatory policies for unprofitable public utilities is a difficult task. As argued by Segal (1998), the incentive for unprofitable monopolistic firms, like a local public transport operator, to reduce its cost and increase revenue is low, particularly when the regulator has a reputation for not allowing the firm to shut-down even in the presence of increasing deficits. In other words, if the regulator is benevolent, the firm tends to under-invest (i.e. reduce its effort) in order to become unprofitable and extract higher public subsidies. This is, quite obviously, the *soft budget constraint* problem the welfare losses of which are typically higher than the dead-weight cost of monopoly.

Then, the main difficulty when governments have to regulate unprofitable industries is how to define a subsidy mechanism that provides incentives to reduces losses and increase efficiency. This is precisely the route followed by the recent reform of local public transport in Italy and Norway, where a *subsidy cap* mechanism is imposed as a regulatory instrument. In the present paper we show that this mechanism shares the same efficiency and manipulability problems of the best known *price cap* mechanism. We are also able to show that by pegging public subsidies to some non manipulable variable (such as past deficits) the regulator may obtain higher effort in cost reducing activities on the part of the firm. As this way out of the manipulability problem suffers, in turns, from non commitment problems we suggest a non-Bayesian menu of contract solution which – despite being non-optimal - not only increases effort in cost reducing activities, but also shares some incentive properties of more complex Bayesian mechanisms.

One of the provision of the Italian reform is that contracts for local transport services are to be tendered out by January the first, 2004. How to combine tendering and the menu of contracts approach to *subsidy cap* could be matter for further research, as well as exploring whether fare regulation together with subsidy regulation might increase the efficiency property of the proposed scheme.

References

Acton, J.P. and Vogelsang, I.,1989. Symposium on Price Cap Regulation: Introduction, Rand Journal of Economics, 20, 369-372.

Baron, D. P., and Myerson, R. B., 1982. Regulating a Monopolist with Unknown Costs, Econometrica, 50, 911-930.

Bös, D., 1991. Privatisation: A Theoretical Treatment, Clarendon Press, Oxford.

Bös, D., 1994. Pricing and Price Regulation, Elsevier Publisher, North Holland, Amsterdam.

Crew, M.A., and Kleindorfer, P.R., 1996. Incentive Regulation in the United Kingdom and the United States: Some Lessons, Journal of Regulatory Economics, 9(3) 211-225.

Crew, M.A., and Kleindorfer, P.R., 2002. Regulatory Economics: Twenty Years of Progress?, Journal of Regulatory Economics, 21(1) 5-22.

Dalen, D.M., and Gómez-Lobo, A., 2003. Yardsticks on the road: regulatory contracts and cost efficiency in the Norwegian bus industry, Transportation, 30, 371-386.

Laffont, J.J., 1994. The New Economics of Regulation Ten Years After, Econometrica, 62 (3) 507-537.

Laffont, J.J., and Tirole, J., 1986. Using Cost Observation to Regulate Firms, Journal of Political Economy, 94 (3) 614-641.

Laffont, J.J. and Tirole, J., 1993. A Theory of Incentives in Procurement and Regulation, The MIT Press, Cambridge (MA).

Leibenstein, H., 1966. Allocative Efficiency vs. ‘X-efficiency’, American Economic Review, 56, 392-415.

Sappington, D., 1980. Strategic Firm Behaviour under a Dynamic Regulatory Adjustment Process, Bell Journal of Economics, 11, 360-372.

Sappington, D., and Weisman, D.L., 1996. Designing Incentive Regulation for the Telecommunications Industry, The MIT Press, Cambridge (MA).

Segal, I. R. 1998. Monopoly and Soft Budget Constraint, RAND Journal of Economics, 29, 596-609.

Shleifer, A., 1985. A Theory of Yardstick Competition, Rand Journal of Economics, 16 (3) 319-327.

Vogelsang, I., and Finsinger, J., 1979. A Regulatory Adjustment Process for Optimal Pricing by Multiproduct Monopoly Firms, Bell Journal of Economics, 10, 157-171.

APPENDIX

In this appendix we show that the incentive non-Bayesian mechanism proposed in (14) leads the firm to increase its efficiency and thus its R/C ratio. Let π be the *accounting* profit of a firm subjected to a *subsidy cap* of the form $T = \hat{T}(\tau)(1 + IP - X(\tau))$, where $\tau \in [\tau_{min}, \tau_{max}]$ is the *R/C* ratio defined as the ration *R/C* between revenues from traffic and total operative cost (at a fixed final price *p*), with $X'(\tau) < 0$. In other words, since a regulator generally uses in the implementation of his regulatory tools information derived from the existing accounting system of the regulated firm, we considerer that the disutility of effort can be neglected for the moment. The ex post incentive impact of the proposed mechanism is shown in Proposition 3. The firm's' profit is:

$$\pi = R - C + \hat{T}(\tau)(1 + IP - X(\tau))$$

That is

$$\pi(\tau) = R\left(1 - \frac{1}{\tau}\right) + \hat{T}(\tau)(1 + IP - X(\tau)) \qquad \text{(A.1)}$$

We assume:

Assumption

1. The menu of contracts in (20) satisfies the following condition:

- $X_{max} > RPI$, if the firm announces $\tau = \tau_{min}$
- $X_{min} < RPI$, if the firm announces $\tau = \tau_{max}$

2. In order to avoid problems of *commitment*, the initial transfer is set in order to cover all the loss in the first period, i.e. $\hat{T} = C - R$.

Given these assumptions, we have the following (see Crew and Kleindorfer, 1996):

Proposition

Suppose the above assumption hold, then there is a unique τ^* *such that* $\tau^* \in \arg\max_{\tau} \pi(\tau)$. *In particular,* τ^* *would be freely selected by a maximasing company from the menu* $[\hat{T}(\tau), X(\tau), \tau]$ *in (14).*

Proof

Since $\pi(\tau)$ is continuous in a closed set, there will be a τ such that it maximizes the firm's profit. Deriving the profit in (A.1) and rearranging, since $\hat{T} = C - R$, we have:

$$\pi'(\tau) = \frac{R}{\tau^2}(IP - X(\tau)) - X'(\tau)\hat{T}(\tau) \qquad \text{(A.2)}$$

Condition (A.2) could be positive or negative. Since the proposed mechanism implies $X'(\tau) < 0$, it results:

- for $\tau = \tau_{min}$, condition (A.2) is always positive, since for assumption 1 $X_{max} > IP$;
- for $\tau = \tau_{max}$, $\hat{T}(\tau_{max})$ is very low, near to zero; then, condition (A.2) is negative since, from assumption 1, $X_{min} < IP$[9].

[9] In general, the proof is valid also when $\tau_{max} < 100\%$; in this case, the *X* factor should be fixed such that $\pi'(\tau_{max}) < 0$. In the following we assume that $\tau_{max} \approx 100\%$.

By continuity there exists at least one value of τ in which condition (A.2) is equal to zero. The S.O.C. is the following:

$$\pi''(\tau) = \frac{2R}{\tau^3}(IP - X(\tau)) + 2X'(\tau)\frac{R}{\tau^2} - X''(\tau)\frac{R}{\tau}(1-\tau)$$

rearranging, we have:

$$\pi''(\tau) = \frac{R}{\tau}\left[\frac{2}{\tau^2}(IP - X(\tau)) + 2X'(\tau)\frac{1}{\tau} - X''(\tau)(1-\tau)\right] \qquad \text{(A.3)}$$

If $\pi''(\tau) < 0$, $\forall\tau \in [\tau]$, then there is a unique solution which is a maximum. Considering our mechanism in (20), condition (A.3) becomes:

$$\pi''(\tau) = \frac{IP - X(\tau)}{\tau} + X'(\tau) \qquad \text{(A.4)}$$

In order to have $\pi'(\tau) < 0$, we must have $-\tau \cdot X'(\tau) > IP - X(\tau)$. Denoting with $\Psi = (X^{\max} - X^{\min}) / (\tau^{\max} - \tau^{\min}) > 0$, condition (20) can be rewritten as $X = -\Psi\tau + \Psi\tau^{\min}$. Since $-\tau \cdot X'(\tau)\Psi = \tau$, condition (A.4) is:

$$X^{\max} - X(\tau) + \Psi\tau^{\min} > IP - X(\tau)$$

that is always satisfied since, from assumption 1, $\Psi\tau^{\min} > 0$. The profit function is strictly concave and so there is a unique interior maximum in the interval $[\tau^{\min}, \tau^{\max}]$. Then, it is possible for the regulator to set the *subsidy cap* mechanism in order to lead the firm to increase its coverage above its lowest level.