DISCUSSION PAPER SERIES

No. 4502

BIGGER AND BETTER: A DYNAMIC REGULATORY MECHANISM FOR OPTIMUM QUALITY

Gianni De Fraja and Alberto Iozzi

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP4502.asp**

ISSN 0265-8003

BIGGER AND BETTER: A DYNAMIC REGULATORY MECHANISM FOR OPTIMUM QUALITY

Gianni De Fraja, University of York and CEPR
Alberto Iozzi, Università di Roma Tor Vergata

Discussion Paper No. 4502
August 2004

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 4502

July 2004

ABSTRACT

Bigger and Better: A Dynamic Regulatory Mechanism for Optimum Quality*

Vogelsang and Finsinger's seminal paper (*Bell Journal of Economics*, 1979) proposes a mechanism for price regulation with some desirable properties, such as convergence to a second best optimum. This mechanism applies to situations where quality is fixed: in practice, quality can be varied by the firm, and regulators have typically imposed constraints on the firm's quality choice. This Paper lays a rigorous theoretical foundation to the inclusion of quality measures in the constraints faced by a regulated firm. We identify a potential pitfall in the approach taken in practice by regulators, and show that, in order to avoid it, the regulated firm should be subject to an additional constraint, which, loosely speaking, requires firms' choices not to be too erratic.

JEL Classification: L430, L510
Keywords: price cap, quality, regulation and RPI-X

Gianni De Fraja
University of York
Department of Economics and Related Studies
Heslington
York
YO10 5DD
Email: gd4@york.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=113071

Alberto Iozzi
Università di Roma 'Tor Vergata'
Dip. SEFEMEQ
Via Columbia 2
I-00133 Rome
ITALY
Email: alberto.iozzi@uniroma2

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=161111

*We would like to thank the seminar audiences in Padua, Turin and Brescia for helpful comments on earlier drafts.

Submitted 22 June 2004

NON TECHNICAL SUMMARY

Regulated firms tend to provide a quality of their services lower than socially optimal. This has been shown theoretically, as well as being observed in practice.

Utility regulators spend therefore a great deal of effort trying to affect the quality supplied by their regulated firms. However, there is little rigorous theoretical analysis of applied mechanisms for quality choices in regulated firms. In the absence of sound theoretical guidance, regulators have typically taken an *ad hoc* approach towards quality. They have done so either by the imposing quality standards, enforced through legal sanctions, or by creating a link between the quality provided by the firm and its allowed revenues. The latter generates a trade-off between prices and qualities, thereby inducing a form of market response by the firm, which can "sell" higher quality to consumers. The link between quality and prices may take the form of the explicit inclusion of a quality correction term in the price cap formula (Italian motorway franchisees and the gas distribution company, and British water companies). In other cases, a quality composite index is not explicitly included in the price cap, but still affects the firm's revenues (Qwest, the New Mexico local telecommunications company and the British energy distribution companies).

The *ad hoc* approach towards quality regulation constitutes a sharp contrast with the relative sophistication of RPI-X price cap regulation which has become increasingly popular since its first use in the UK at the beginning of the '80s. The RPI-X price cap is the practical counterpart of Vogelsang and Finsinger's seminal theoretical contribution published in the Bell Journal of Economics in 1979. They study a dynamic model of price regulation (with fixed qualities), and show that an appropriately designed price cap mechanism aligns the firm's trade-offs to the regulator's, and induces the firm to choose, as time goes by, prices closer and closer to those which would be chosen by an omniscient regulator concerned solely with consumers' welfare. The principles underlying Vogelsang and Finsinger's analysis, and its practical counterpart, RPI-X price cap regulation, is that the firm is the decision maker in the best position to evaluate the trade-offs between costs and demand for its different products, and therefore to take decision regarding relative prices.

In this paper we lay a theoretical foundation of the regulation of quality on the same principle as Vogelsang and Finsinger's. Extending their analysis, we propose a mechanism that exploits the firm's superior knowledge of its cost function, and as no special information requirements on the regulator's part. In our mechanism, the firm is constrained to choose, in each period, qualities and prices satisfying two constraints: the *quality adjusted Vogelsang-Finsinger* constraint, and the *distance* constraint.

The first of these constraints, the *quality adjusted Vogelsang-Finsinger* constraint, creates trade-offs between the allowed change in the firm's prices and the change in the firm's quality measures. The firm may choose an average price higher than under the Vogelsang and Finsinger mechanism (or

RPI-X regulation) if it increases the quality of its output: the regulated firm can "sell" higher quality to consumers; conversely, of course, the firm can lower the quality of its output, but must "bribe" the consumers via lower average prices. Deviations from the average price which the Vogelsang and Finsinger mechanism would impose are permitted through a quality adjustment term added to the X factor: this term is a weighted average of the marginal effects of the quality changes on consumers' welfare and it is not in any way dependent on the firm's cost, typically unknown to the regulator. In analogy to the RPI-X price cap and the Vogelsang and Finsinger mechanism, where the weight of each price in the price average is given by the social valuation in the previous period, the weight of each quality measure in the quality adjustment term added to the X factor is given by the social valuation of that dimension of quality, measured by the social value at last period's prices and qualities. Our mechanism is therefore reminiscent of the practice of some regulators to include a quality adjustment in the price cap and provide a robust theoretical foundation for this regulatory provision.

However, a simple example illustrates that including a quality index in the price cap does not necessarily ensure that the choices made by the firm lead to a socially optimal outcome: the firm's choices may follow a Pareto inefficient cycle. This may occur when the welfare function is not convex. While the assumption of convexity of the consumers' surplus in prices made by Vogelsang and Finsinger is a consequence of the natural assumption that the demand functions be decreasing, and is therefore realistic, when the consumers' utility depends on prices and qualities and the firm can choose both (prices and quantities), convexity cannot be warranted. To address this potential problem, we suggest that the firm is required, in each period, to meet an additional constraint. This requires, in a sense made precise in the paper, the firm's today's choices to be sufficiently similar to yesterday's choices.

Our main result is as follows: when the firm is subject to these two constraints, its choices of price and quality tend to those that an omniscient regulator would make. Hence, our regulatory mechanism has, with respect to prices and quality, the same desirable properties as price cap regulation has with respect to prices only, and applies to a much more complex and realistic setting where the regulated firm can also choose qualities alongside prices. It also has the same practical applicability as price cap regulation.

1 Introduction.

The quality of firms' output has received a great deal of attention, both in the economics literature, where the analysis of quality choices by firms has been a cornerstone of the modern theory of industrial organisation since its early days,[1] and in practice: in many countries, utility regulators spend a great deal of effort trying to affect the quality supplied by their regulated firms.

However, there is little rigorous theoretical analysis of applied mechanisms for quality choices in regulated firms. This is especially surprising, in view of the large extent of the theoretical literature examining regulated firms (Armstrong and Sappington, forthcoming), and in particular of Spence's seminal analysis of the strategic interaction between a regulator and a monopolist who choose price and quality (1975).

In the absence of sound theoretical guidance, regulators have typically taken an *ad hoc* approach towards quality, either imposing quality standards, or explicitly or implicitly linking the allowed price level to quality improvements (see, for instance, Ofwat (2002); more examples are discussed in Section 3). This *ad hoc* approach constitutes a sharp contrast with the relative sophistication of RPI–X price cap regulation which has become increasingly popular since its first use in the UK at the beginning of the '80s. As noted by Bradley and Price (1989), the RPI–X price cap is the practical counterpart of Vogelsang and Finsinger's seminal theoretical contribution (1979, VF in what follows). VF study a dynamic model of price regulation (with fixed qualities), and show that an appropriately designed price cap mechanism aligns the firm's trade-offs to the regulator's, and induces the firm to choose, as time goes by, prices closer and closer to those which would be chosen by an omniscient regulator concerned solely with consumers' welfare.

VF's analysis and its practical counterpart, RPI-X price cap regulation, are founded on the principle that the firm is the decision maker in the best position to evaluate the trade-offs between costs and demand for its

[1] Early contributions are Spence (1975) and Mussa and Rosen (1978) for the monopolist, and Shaked and Sutton (1982) and Gabszewicz and Thisse (1979) for the oligopoly case.

different products, and therefore to take decision regarding relative prices. In this paper we lay a theoretical foundation of the regulation of quality on the same principle. Extending VF's analysis, we propose a mechanism that exploits the firm's superior knowledge of its cost function, and has no special information requirements on the regulator's part. In each period, the regulator imposes two constraints on the firm. One constraint creates trade-offs between the allowed change in the firm's prices and the change in the firm's quality measures. These trade-offs reflect the relative social valuations of prices and qualities in the previous period, and are reminiscent of the practice of some regulators to include a quality adjustment in the price cap. However, our paper also illustrates that, unless the welfare function is convex, simply including a quality index in the price cap does not necessarily ensure that the choices made by the firm lead to a socially optimal outcome: the firm's choices may follow a Pareto inefficient cycle. To address this potential pitfall, we propose that the firm is required, in each period, to meet an additional constraint. This requires, in a sense made precise in Section 4.2, the firm's today's choices to be sufficiently similar to yesterday's choices. As Proposition 3 shows, when the firm is subject to these two constraints, its choices of price and quality tend to those that an omniscient regulator would make.

The paper is organised as follows. The model is in Section 2. The practice of quality regulation is briefly presented in Section 3. The formal analysis is in Section 4: a static regulatory mechanism is studied in Section 4.1 and the more applicable dynamic mechanism in Section 4.2. Section 4.3 illustrates the role played by the additional constraint. Section 4.4 modifies the mechanism to make it applicable to situations where quality is observed *ex-post* only. Section 5 is a brief conclusion.

2 The model.

2.1 Demand and cost conditions.

Time is divided in periods, denoted by $t = 1, 2, \ldots$. We consider a regulated firm which produces n different products. Let x_i^t denote the quantity of good i sold during period t, and p_i^t its price during the period. We assume for simplicity that prices are constant within each period; if prices vary, a shorter period or an appropriate price average can be taken.

The firm's output is characterised by the m-dimensional vector $\mathbf{q}$ of quality indicators q_j, $j = 1, \ldots, m$, $\mathbf{q} = (q_1, \ldots, q_m)$.[2] Without loss of generality, we denote by $[\underline{q}_j, \overline{q}_j] \subseteq \mathbb{R}$ the quality range along dimension j. Let $Q = \prod_{j=1}^{m} [\underline{q}_j, \overline{q}_j] \subseteq \mathbb{R}^m$ be the space of possible quality vectors.

The firm's output is demanded by L consumers with quasi-linear indirect utility given by $v^\ell(\mathbf{p}, \mathbf{q}) + y^\ell$, $\ell = 1, \ldots, L$, where y^ℓ is consumer ℓ's income and $\mathbf{p}$ the price vector $(p_1, \ldots, p_n)$. Clearly, for all $\ell = 1, \ldots, L$, we have $\frac{\partial v^\ell(.)}{\partial p_i} \leqslant 0$, for all $i = 1, \ldots, N$, and $\frac{\partial v^\ell(.)}{\partial q_j} > 0$ for all $j = 1, \ldots, m$: income and quality are goods. The additive formulation implies that there are no income effects: demand for good i by individual ℓ is simply given by $x_i^\ell(\mathbf{p}, \mathbf{q})$, for $i = 1, \ldots, N$ and $\ell = 1, \ldots, L$. The aggregate demand is given by the sum of individual demands: for each $i = 1, \ldots, N$, $x_i(\mathbf{p}, \mathbf{q}) = \sum_\ell x_i^\ell(\mathbf{p}, \mathbf{q})$. To compact notation, we write $\mathbf{x}(\mathbf{p}, \mathbf{q})$ to denote the vector $(x_1(\mathbf{p}, \mathbf{q}), \ldots, x_n(\mathbf{p}, \mathbf{q}))$. The aggregate demand functions are assumed to have the standard properties: for all $i = 1, \ldots, n$, for all $\mathbf{q} \in Q$, $x_i(\cdot)$ is continuous and twice differentiable, with $\frac{\partial x_i}{\partial p_i} < 0$. Also, as in VF (Assumption 3b, p 159), for any given quality vector, expenditure on each good goes to zero when its price increases without bounds: $\lim_{p_i \to \infty} x_i(\mathbf{p}, \mathbf{q}) p_i = 0$, for $i = 1, \ldots, N$. We will be using the inverse demand function for fixed quality $\mathbf{q}$, which we denote by $\mathbf{x}^{-1}(\mathbf{x}, \mathbf{q})$: $\mathbf{x}^{-1}(\mathbf{x}_0, \mathbf{q})$ is the price vector which ensures that demand is $\mathbf{x}_0$ when quality is $\mathbf{q}$.[3]

We describe the firm's technology by the cost function $C(\mathbf{x}, \mathbf{q})$. This satisfies, plausibly, $\frac{\partial C}{\partial x_i} > 0$ for $i = 1, \ldots, n$, and $\frac{\partial C}{\partial q_j} > 0$ for $j = 1, \ldots, m$: like all good things, quality comes at a cost. Moreover, for every $j = 1, \ldots, m$,

[2] In general, there is no necessary correspondence between products and quality dimensions: q_j could be a quality measure for product i, or a general measure of the firm's output. Among the measures of quality used in practice there is the quality of drinking water, which can be attributed to a specific product (Ofwat, 2002, p. 28), but also the promptness with which complaints are addressed (Ofgem, 2001, p. 9), which cannot.

[3] Mathematically, $\mathbf{x}^{-1}$ is the projection of the inverse demand correspondence on the cartesian product $\mathbb{R}_+^n \times \{\mathbf{q}\}$: for every $\mathbf{x}_0 \in \mathbb{R}_+^n$, $\mathbf{x}^{-1}(\mathbf{x}_0, \mathbf{q}) \in \mathbb{R}_+^n$ is such that: $\mathbf{x}\left(\mathbf{x}^{-1}(\mathbf{x}_0, \mathbf{q}), \mathbf{q}\right) = \mathbf{x}_0$.

$\lim_{q_j \to \overline{q}_j} C(\mathbf{x}, \mathbf{q}) = +\infty$: perfection – in the sense of maximal quality - is beyond reach. The firm's technology also exhibits decreasing (quantity) ray average cost: for every $\mathbf{q} \in Q$, $r \geqslant 1$ implies $C(r\mathbf{x}, \mathbf{q}) \leqslant rC(\mathbf{x}, \mathbf{q})$. This captures formally the idea that the firm has economies of scale in production and justifies the presence of a regulator. The firm's profits is

$$\pi(\mathbf{p}, \mathbf{q}) = \mathbf{x}(\mathbf{p}, \mathbf{q}) \cdot \mathbf{p} - C(\mathbf{x}(\mathbf{p}, \mathbf{q}), \mathbf{q}), \tag{1}$$

where the dot "$\cdot$" denotes the inner product: $\mathbf{x} \cdot \mathbf{p} = \sum_i x_i p_i$.

We follow VF and most of the literature in our assumption that the regulator is a benevolent utilitarian: his objective function $v(\mathbf{p}, \mathbf{q})$ is the unweighted sum of individuals' utility: $v(\mathbf{p}, \mathbf{q}) = \sum_\ell v^\ell(\mathbf{p}, \mathbf{q})$ (we normalise away the sum of total income $\sum_\ell y^\ell$). The quasi-linear nature of the individuals' preferences implies that Roy's identity can be written as

$$\frac{\partial v(\mathbf{p}, \mathbf{q})}{\partial p_i} = -x_i(\mathbf{p}, \mathbf{q}) \text{ for every } i = 1, \ldots, n, \text{ for all } \mathbf{q} \in Q. \tag{2}$$

This implies (see Bös, 1981, p. 5-11) that we can write the regulator's objective function as the area below the demand curve and above the price, summed over the various goods:

$$\begin{aligned} v(\mathbf{p}, \mathbf{q}) &= \sum_{\ell=1}^{L} \left[\int_{p_1}^{\infty} \ldots \int_{p_i}^{\infty} \ldots \int_{p_n}^{\infty} \sum_{i=1}^{n} x_i^\ell(\mathbf{z}, \mathbf{q}) \mathrm{d}z_n \ldots \mathrm{d}z_i \ldots \mathrm{d}z_1 \right] \\ &= \int_{p_1}^{\infty} \ldots \int_{p_i}^{\infty} \ldots \int_{p_n}^{\infty} \sum_{i=1}^{n} x_i(\mathbf{z}, \mathbf{q}) \mathrm{d}z_n \ldots \mathrm{d}z_i \ldots \mathrm{d}z_1. \end{aligned} \tag{3}$$

We take the regulator's objective to be the consumers' surplus mainly for definiteness: the analysis of the paper would not be altered qualitatively if the regulator had an objective function different from (3), for example, one satisfying only some minimal requirements, such as $\frac{\partial v}{\partial p_i} < 0$ and $\frac{\partial v}{\partial q_i} > 0$. The regulator's objective function could differ from (3) because of the regulator's difficulty to determine the consumers' evaluation of quality, or of her distributional concern: the welfare of low income consumers may weigh more heavily in the regulator's payoff. A more radical departure from (3) would be the assumption that the regulator is captured by the regulated firm (Stigler, 1971), or that environmental concerns also enter the regulator's objective function (Oates and Portney, 2003).

The regulator gives the firm, in each period of time, freedom to choose its prices and qualities subject to a set of constraints; if the firm is unwilling to satisfy these constraints, it may leave the market. In line with most of the literature on regulation, we assume that the firm behaves myopically: in each period of time t, the firm maximises current profits.[4]

2.2 Quality provision with traditional price cap regulation.

In the absence of constraints on its quality, the firm chooses in each period the $(n+m)$–dimensional vector $(\mathbf{p}, \mathbf{q})$ of prices and qualities which maximises its current profit, (1), subject to a regulatory constraint which, in general terms, can be written $F(\mathbf{p}) \leq \overline{I}$: $F(\mathbf{p})$ is a real valued function of the firm's prices and $\overline{I}$ is a ceiling chosen by the regulator. The first proposition establishes formally that this form of price cap regulation is ineffective in providing incentives for the provision of quality.

Proposition 1 *If the price cap of a regulated firm is binding, then quality is underprovided, in the sense that there exists a Pareto improving quality increase.*

Proof. With fixed prices, since we have assumed $\lim_{q_j \to \overline{q}_j} C(\mathbf{x}, \mathbf{q}) = +\infty$, $\mathbf{q}$ is in the interior of Q, and therefore the firm's choice of q_j satisfies the first order condition: $\sum_{i=1}^{n} \frac{\partial x_i(\cdot)}{\partial q_j} \left(p_i - \frac{\partial C(\cdot)}{\partial x_i} \right) - \frac{\partial C(\cdot)}{\partial q_j} = 0$. A marginal change in quality j determines a change in gross consumers' surplus given by: $\frac{\partial v(\cdot)}{\partial q_j} > 0$, because of the assumption that quality is a good. Therefore, consumers would be willing to compensate the firm for a marginal increase in quality j, $\mathrm{d}q_j > 0$. This establishes the statement.■

Proposition 1, a natural generalisation of Spence's seminal analysis of quality setting by a regulated monopolist (Spence, 1975), says that, for any given quality choice by the firm, consumers' surplus would increase for a small increase in quality even if they had to compensate the firm fully for

[4] Myopic behaviour implies that firms maximise the present value of future profits with a 0 discount rate. Vogelsang (1989) shows that the qualitative nature of VF's conclusions is not changed by a strictly positive discount rate. On the other hand, if the firm is foresighted and behaves strategically while the regulator is myopic, then, as shown by Sappington (1980), the VF mechanism unravels.

the change in profit that the change in quality would cause.[5] Intuitively, this is because a marginal increase in a quality measure causes a second order change in the firm's profit, and a first order increase in consumers' surplus. An *unregulated* monopolist may of course supply too high a quality level (Spence 1975). In graphical terms, for the case of one product and one quality level, consider Figure 1 below. A firm subject to price cap regulation chooses a point where the tangent to the isoprofit is vertical: at any such point the isowelfare is increasing, implying that both quality and price can be increased to make the firm and the regulator better-off. But, of course, if the isowelfare curves are very "steep", the point of tangency between the zero-profit locus and the isowelfare may be at a quality level lower than q_m, implying that the monopolist is choosing excessive quality (and excessive price) relative to the social optimum.

3 Quality regulation

Proposition 1 comes of course as no surprise to regulators and practitioners, who are well aware that a price capped firm may reduce quality to cut its costs. A textbook example is the noticeable reduction in the quality of British Telecom's services immediately after privatisation, when it was subject to price cap regulation, without specific provisions with regard to quality (see, for instance, Rovizzi and Thompson, 1992).[6] There are two possible regulatory responses to the firm's drive to lower quality. Firstly,

[5] Similar results on the suboptimality of price-regulated monopolist's choice on quality are derived by Sheshinsky (1976), who also concludes that no regulation may be better than only price regulation.

[6] More recently, a number of papers (lucidly surveyed by Sappington, 2003) have argued that the quality of the service provided by regulated telecommunications firms does not suffer with incentive regulation, relative to rate of return regulation. By the same token, the UK regulator has recently reported a high level of consumers' satisfaction with the level of service received (Oftel, 2001). This is probably due to the fact that regulated telecommunications firms operate today in rather competitive environments. Theoretical arguments do suggest that competition may indeed lessen the problem of underprovision of quality by regulated firms (Beil *et al.*, 1995, and Ma and Burgess, 1993).

the imposition of quality standards,[7] enforced through legal sanctions, from fines up to the withdrawal of the licence; this happened to the train operator serving the South-East of England in 2003. Secondly, the creation of a link between the quality provided by the firm and its allowed revenues or prices. This generates a trade-off between prices and qualities, and induces a form of market response by the firm, which can "sell" higher quality to consumers (Waddams Price *et al.*, 2002; Banerjee, 2003). This link between quality and prices may take the form of the explicit inclusion of a quality correction term in the price cap formula: for example, in Italy, higher quality relaxes the price cap for toll motorway franchisees (Iozzi, 2002) and the gas distribution company (AEEG, 2000). In the UK water industry incentives for quality provision are based on comparative performance indicators: the best (worst) performing companies are rewarded (penalised) with a positive (negative) adjustment in the price cap (Ofwat, 1999 and 2002). There are other examples, where quality is not explicitly included in the price cap, but still affects the firm's revenues: the New Mexico Public Regulation Commission approved in 2001 a five-year plan in which the local telecommunications company, Qwest, may increase prices if some target quality measures are met (NMPRC, 2001). Similarly, in the UK, the energy distribution companies receive financial incentives (as a percentage of their annual revenues) according to various quality of service indicators (Ofgem, 2001).

There is, however, no presumption that either of these methods, the imposition of quality standards and the link between the quality levels and the allowed prices or revenues, will lead to efficient price and quality choices. In Section 4, we therefore propose a natural modification of the price cap mechanism which aligns the firm's incentives in choosing prices and qualities to the regulator's thus leading to the efficient choice of price and quality. In

[7] In the UK, there are two different types of standards: Guaranteed Standards, for which direct compensation must be paid to the customer, if the company does not meet them; and Overall Standards, which are similar in nature to the statutory obligations included in the licence. These standards cover very specific issues in the service provision: minimum percentage of letters from customers to be replied to within a given number of working days, minimum percentage of system faults to be corrected within a given period of time, offer of timed appointment with the customer, and so on.

Section 4.1 we consider the case where the price cap is of a static nature, and subsequently the more applicable dynamic price cap.

4 The quality-adjusted price cap

4.1 Static price cap

We begin with a static price cap constraint taking the following form:

$$\mathbf{w} \cdot \mathbf{p} - \mathbf{u} \cdot \mathbf{q} \leqslant \overline{I}, \tag{4}$$

where $\mathbf{w}$ and $\mathbf{u}$ are the weights attributed to the prices and to the quality indicators, and have therefore the same dimensionality as $\mathbf{p}$ and $\mathbf{q}$ respectively: $\mathbf{w} = (w_1, \ldots, w_n)$ and $\mathbf{u} = (u_1, \ldots, u_m)$. (4) requires the firm to choose prices and qualities such that the difference between their weighted averages is lower than the exogenously set level $\overline{I}$.

To interpret (4) note first of all that if the weights u_k are all 0, then we are in the standard case where quality is not regulated. If instead a positive weight is attributed to quality dimension j, then $u_j > 0$, and, by increasing quality j, the firm can alter its price constraint and obtain an increase in the allowed average price $\mathbf{w} \cdot \mathbf{p}$. This can be seen by rewriting (4) as

$$\sum_{i=1}^{n} w_i p_i \leqslant \overline{I} + \sum_{j=1}^{m} u_j q_j. \tag{5}$$

(5) highlights the relationship with RPI–X regulation: the firm is allowed a value of X in the RPI–X formula where the value of X is a decreasing function of the quality level of the firms' output. The next proposition establishes the second best optimality of a regulatory constraint of the type (4) or (5).

Proposition 2 *Let $(\mathbf{p}^*, \mathbf{q}^*)$ be the price-quality vector which maximises consumers' welfare subject to the break even constraint. Let the firm be subject to price cap regulation according to formula (4), where*

$$w_i = x_i(\mathbf{p}^*, \mathbf{q}^*) \qquad i = 1, \ldots, n, \tag{6}$$

$$u_j = \frac{\partial v(\mathbf{p}^*, \mathbf{q}^*)}{\partial q_j} \qquad j = 1, \ldots, m, \tag{7}$$

$$\overline{I} = C(\mathbf{w}, \mathbf{q}^*) - \mathbf{u} \cdot \mathbf{q}^*. \tag{8}$$

Then the firm chooses the price-quality vector $(\mathbf{p}^*, \mathbf{q}^*)$.

Proof. Consider the second best price-quality vector $(\mathbf{p}^*, \mathbf{q}^*)$. This is the solution to the following problem:

$$\max_{\mathbf{p},\mathbf{q}} v(\mathbf{p}, \mathbf{q}) \quad \text{s.t.: } \pi(\mathbf{p}, \mathbf{q}) \geqslant 0. \tag{9}$$

The FOC to problem (9) are:

$$\frac{\partial \mathcal{L}}{\partial p_i} = \frac{\partial v(\mathbf{p}^*, \mathbf{q}^*)}{\partial p_i} - \mu \frac{\partial \pi(\mathbf{p}^*, \mathbf{q}^*)}{\partial p_i} = 0 \qquad i = 1, \ldots, n; \tag{10}$$

$$\frac{\partial \mathcal{L}}{\partial q_j} = \frac{\partial v(\mathbf{p}^*, \mathbf{q}^*)}{\partial q_j} - \mu \frac{\partial \pi(\mathbf{p}^*, \mathbf{q}^*)}{\partial q_j} = 0 \qquad j = 1, \ldots, m. \tag{11}$$

where μ is the Lagrangean multiplier associated with the constraint in (9). Since $\frac{\partial v(\mathbf{p}^*, \mathbf{q}^*)}{\partial p_i} < 0$ whenever $\frac{\partial \pi(\mathbf{p}^*, \mathbf{q}^*)}{\partial p_i} > 0$, μ is strictly positive, and so the non-negativity constraint holds as an equality: $\pi(\mathbf{p}, \mathbf{q}) = 0$.

Use Roy's identity, (2), and (6) in (10), and substitute (7) into (11), to write (10) and (11) as:

$$\frac{\partial \pi(\mathbf{p}^*, \mathbf{q}^*)}{\partial p_i} = \frac{w_i}{\mu} \qquad i = 1, \ldots, n,$$

$$\frac{\partial \pi(\mathbf{p}^*, \mathbf{q}^*)}{\partial q_j} = \frac{u_j}{\mu} \qquad j = 1, \ldots, m.$$

Now consider the problem of a profit maximising firm subject to a regulatory constraint given by (4):

$$\max_{\mathbf{p},\mathbf{q}} \pi(\mathbf{p}, \mathbf{q}) \quad \text{s.t.: } \mathbf{w} \cdot \mathbf{p} - \mathbf{u} \cdot \mathbf{q} \leq \overline{I}. \tag{12}$$

Let $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$ be the solutions to this problem. They satisfy:

$$\frac{\partial \mathcal{L}}{\partial p_i} = \frac{\partial \pi\left(\hat{\mathbf{p}}, \hat{\mathbf{q}}\right)}{\partial \hat{p}_i} + \nu w_i = 0 \qquad i = 1, \ldots, n; \tag{13}$$

$$\frac{\partial \mathcal{L}}{\partial q_j} = \frac{\partial \pi\left(\hat{\mathbf{p}}, \hat{\mathbf{q}}\right)}{\partial \hat{q}_j} - \nu u_j = 0 \qquad i = 1, \ldots, m. \tag{14}$$

Again, since profit increases in at least one price or decreases in at least one quality, ν, the Lagrangean multiplier, is strictly positive. Therefore, at the solution of problems (9) and (12), the gradients of the profit function are proportional. This, combined with the fact that the firm obtains zero profit in both problems, implies $(\mathbf{p}^*, \mathbf{q}^*) = (\hat{\mathbf{p}}, \hat{\mathbf{q}})$. ■

Proposition 2 is a natural extension of the standard argument that pricing decisions can be delegated to the firm, provided that an appropriate

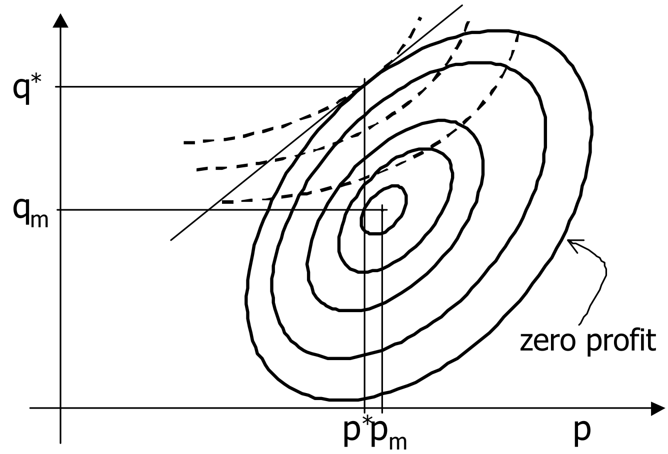

Figure 1: The profit maximising and the welfare maximising price-quality vectors.

constraint is imposed (Armstrong *et al.*, 1994). The vector of n prices is replaced by the $(n+m)$-dimensional price-quality vector. The firm must respect the constraint that the weighted average of its choices is below a certain level. Both for prices and for qualities, the weights are the changes in consumers' surplus at the optimum. Social optimality follows from the fact that, at the second best welfare optimum, the isowelfare and the isoprofit are tangent to each other, since they are both tangent to the hyperplane with slope $(\mathbf{w}, -\mathbf{u})$.

Figure 1 illustrates the situation in the price-quality cartesian plane, for the case where there is only one good and one quality measure. The solid curves are the isoprofit lines. The dashed curves are the isowelfare lines, upward sloping because welfare is increasing in quality and decreasing in price; we have drawn them as convex functions to reflect the natural assumption that consumers' willingness to pay for increases in quality should be higher when quality is low than when quality is already high.[8] The vector $(\mathbf{p}_m, \mathbf{q}_m)$ is the profit maximising price-quality pair and $(\mathbf{p}^*, \mathbf{q}^*)$ is the second best optimum: at this point the zero-profit isoprofit line is tangent

[8] For further discussion, see Sheshinski (1976, p 130-1). Recall that we have ruled out income effects. With income effects, the argument would be strengthened: when price is already high, consumers prefers consuming the other goods relatively more than they prefer increases in quality.

to the isowelfare map and both are tangent to the hyperplane with slope $(\mathbf{w}, -\mathbf{u})$, drawn as the line aa. The argument would not change if the isowelfare loci were instead concave, in which case the welfare function would be quasi-convex (see Section 4.3 for the role of the convexity of the welfare function).

4.2 Dynamic price and quality regulation

Implementation of the mechanism given in Proposition 2 requires detailed global knowledge of the cost and demand conditions. This informational requirement is, in practice, beyond the regulator's reach: it is the firm's private information, and the firm has no incentive to reveal it. The 'new' regulation literature models this information advantage by positing that the regulator maximises her expected payoff, given her beliefs about the firm's private information (Baron and Myerson, 1982; Laffont, 1994). This approach applies of course to the case where the firm's output includes quality measures (Laffont and Tirole, 1991), but the algebra gets rapidly out of hand when the dimensionality of the price-quality vector increases, limiting the practical applicability to multiproduct firms. Moreover, accountability of the regulators and certainty of the economic environment in which regulated firms operates may advice against using regulatory schemes where the outcome depends heavily on the regulator's subjective beliefs.

If quality is fixed, these limitations are addressed successfully in VF's seminal paper. They assume a limited information requirement on the part of the regulator, who neither has knowledge of the cost function, nor needs to form beliefs on the distribution of the unknown parameter, and has only local knowledge of the demand function. The VF mechanism is a dynamic version of (4), with $\mathbf{u} = 0$. In each period, the weights in (4) are proportional to the quantities produced in the previous period, and the parameter $\overline{I}$ is the previous period's total cost. That is, in period t, the firm can charge prices $\mathbf{p}^t$ satisfying (qualities are fixed at $\mathbf{q}^{t-1}$):

$$\mathbf{x}^{t-1} \cdot \mathbf{p}^t \leq C(\mathbf{x}^{t-1}, \mathbf{q}^{t-1}). \tag{15}$$

VF show that the prices chosen by a firm subject to this type of regulatory mechanism converge to the second best Ramsey prices: for each product, the mark-up over the marginal cost is proportional to the inverse of the elasticity of the demand for that product (the Ramsey rule), and the firm makes no profit (second best optimality). Constraint (15) can be re-written as an RPI-X formula: divide each side by period $t-1$ revenues, $R^{t-1} = \mathbf{x}^{t-1} \cdot \mathbf{p}^{t-1}$, and write period $t-1$ profit as $\pi^{t-1} = R^{t-1} - C(\mathbf{x}^{t-1}, \mathbf{q}^{t-1})$:

$$\frac{\mathbf{x}^{t-1} \cdot \mathbf{p}^{t}}{\mathbf{x}^{t-1} \cdot \mathbf{p}^{t-1}} \leq 1 - \frac{\pi^{t-1}}{R^{t-1}}. \tag{16}$$

If there is no inflation, RPI is 1, and the X in the RPI formula is given by the rate of profit over sales for the previous period. Geometrically, constraint (16) is constructed in a two-step procedure. In the first step, we take the hyperplane (in the price space) tangent to the isowelfare surface through the point representing the price vector chosen by the firm in the previous period. This has gradient $\mathbf{x}^{t-1}$. In the second step, we shift this hyperplane parallelly in such a way that it goes through a price vector which determines a demand vector with two features: (i) it is a proportional increase in the quantities sold in the previous period, hence it is given by $\mathbf{x}^{-1}\left(r\mathbf{x}^{t-1}, \mathbf{q}^{t-1}\right)$ for some $r > 1$, and (ii) if the quantities sold in the previous period had been sold at these prices, the revenues generated would have exactly covered the total cost incurred in the previous period costs: (15) holds an equality. VF explain that at this value of r the increase in total cost caused by the increase in quantity from $\mathbf{x}^{t-1}$ to $\mathbf{x}(\mathbf{p}^{t}, \mathbf{q}^{t-1})$ is more than compensated by the increase in revenues, because, by assumption, the average cost (along the ray) is decreasing[9]. This parallel shift, which occurs only when the firm made strictly positive profits in the previous period, is introduced by VF to tighten the constraint, and it ensures convergence to a situation where the firm cannot make positive profits. Without it, the process would converge to Ramsey prices, but the firm's profit would be strictly positive in the limit (Brennan 1989).

When quality can vary, a price cap is shown in Proposition 1 to give insuf-

[9] That this r must exist is shown by VF (p 163), by noting that at $r = 1$ the constraint is clearly violated, and that, when r is increased without bounds, total revenue goes to 0, and, therefore, the constraint is satisfied. Hence, by the continuity of $\mathbf{x}^{-1}$, there exists an r such that the constraint is satisfied as an equality. If there is more than one such r, then take the smallest.

ficient incentive for the provision of quality. We illustrated in Section 3 that regulators have tried to extend the VF mechanism by creating a formal or informal link between the allowed prices and the quality supplied by the firm. This link has been typically *ad hoc*. We show here that this *ad hoc* approach may form part of a regulatory mechanism with similar desirable properties as VF's, provided that the quality correction of the price cap formula reflects the social valuation of quality changes and that is supplemented by an additional constraint. We label this the "distance" constraint, and to derive it we begin by defining the following function: for a given vector of non-negative weights (not all 0) $(\rho_1, \ldots, \rho_n, \omega_1, \ldots, \omega_m) \in \mathbb{R}_+^{n+m} \setminus \{0\}$, and for a given vector of strictly positive parameters $(\xi_1, \ldots, \xi_n, \zeta_1, \ldots, \zeta_m) \in \mathbb{R}_{++}^{n+m}$, given two price-quality vectors $(\mathbf{p}^A, \mathbf{q}^A)$ and $(\mathbf{p}^B, \mathbf{q}^B)$, let their *distance*[10] be given by:

$$d\left(\left(\mathbf{p}^A, \mathbf{q}^A\right), \left(\mathbf{p}^B, \mathbf{q}^B\right)\right) = \sum_{i=1}^{n} \rho_i \left|p_i^A - p_i^B\right|^{\xi_i} + \sum_{j=1}^{m} \omega_j \left|q_j^A - q_j^B\right|^{\zeta_j}. \quad (17)$$

The first of the two constraints forming the mechanism we propose is the following:

The distance constraint *In each period, the firm must choose a price-quality vector satisfying:*

$$d\left(\left(\mathbf{p}^t, \mathbf{q}^t\right), \left(\mathbf{p}^{t-1}, \mathbf{q}^{t-1}\right)\right) \leqslant d\left(\left(\mathbf{p}^{t-1}, \mathbf{q}^{t-1}\right), \left(\mathbf{p}^{t-2}, \mathbf{q}^{t-2}\right)\right) \phi(\pi^{t-1}), \quad (18)$$

where $\phi(\pi)$ is a function satisfying $\phi(0) = 1$ and $\phi(\pi) \in (0, 1)$ for every $\pi > 0$.

Constraint (18) is a rule requiring that if profit is positive, *the distance between the price-quality vector chosen in one period and the price-quality*

[10] This is consistent both with the everyday use of the word "distance", and with the mathematical notion of distance on a cartesian space $\mathbb{R}^k$ as a function d from $\mathbb{R}^{2k}$ into $\mathbb{R}_+$, satisfying reflexivity (for every $z \in \mathbb{R}^k$, $d(z, z) = 0$), symmetry, (for every $z, z' \in \mathbb{R}^k$, $d(z, z') = d(z', z)$), and the Cauchy-Schwartz inequality (for every $z, z', z'' \in \mathbb{R}^k$, $d(z, z') + d(z', z'') \geqslant d(z, z'')$)(Kelley 1955). The weights ρ_i and ω_j and the parameters ξ_i and ζ_j are unrestricted. They could reflect the different units in which the firm's choices are measured, or some assessment of the relative importance of the different quality aspects, though this is not necessary.

vector chosen in the previous period must fall with time. More succinctly, the choices of the firm must not be too erratic. The role of the distance constraint (18) is to ensure convergence of the price-quality choices of the firm: as the next section shows, without it, the mechanism may be locked in a Pareto inefficient cycle. Note also that the rich parameter set in the definition of distance (17) allows the regulator to design this constraint in terms of any subset of prices and/or qualities, setting the parameters of all the other dimensions equal to 0.

The second constraint is the analogue to the VF constraint.

The quality-adjusted VF constraint *In each period, the firm must choose a price-quality vector satisfying:*

$$\frac{\mathbf{x}^{t-1}\cdot\mathbf{p}^{t}}{\mathbf{x}^{t-1}\cdot\mathbf{p}^{t-1}} \leqslant 1-\left(\frac{\pi^{t-1}}{R^{t-1}}\gamma^{t}-\frac{\mathbf{u}^{t-1}\cdot\left(\mathbf{q}^{t}-\mathbf{q}^{t-1}\right)}{R^{t-1}}\right), \tag{19}$$

where $\mathbf{u}^{t-1}=\left(u_1^{t-1},\ldots,u_m^{t-1}\right)$ *and*

$$u_j^{t-1}=\frac{\partial v_i\left(\mathbf{p}^{t-1},\mathbf{q}^{t-1}\right)}{\partial q_j^{t}} \qquad j=1,\ldots,m, \qquad t=1,\ldots. \tag{20}$$

The *quality-adjusted VF constraint* is obtained in much the same way as the original VF constraint: take the hyperplane tangent to the isowelfare surface through the point representing the price-quality vector chosen in the previous period, and tighten this constraint by means of a parallel shift, in order to reduce the choice set of the firm. To see this rewrite (19) as

$$\mathbf{x}^{t-1}\cdot\mathbf{p}^{t}-\mathbf{u}^{t-1}\cdot\mathbf{q}^{t}\leqslant C(\mathbf{x}^{t-1},\mathbf{q}^{t-1})+\left(1-\gamma^{t}\right)\pi^{t-1}+\mathbf{u}^{t-1}\cdot\mathbf{q}^{t-1};$$

the LHS describes the hyperplane with slope $\left(\mathbf{x}^{t-1},\mathbf{u}^{t-1}\right)$, which is the gradient of the isowelfare through point $\left(\mathbf{p}^{t-1},\mathbf{q}^{t-1}\right)$. The RHS determines how the hyperplane should be shifted. The last term is a necessary normalisation to compensate for the inclusion of quality measures on the LHS. The substantive difference with the VF mechanism is that the rate of profit which determines the allowed price level is multiplied in period t by a factor γ^t. This is necessary to ensure that, in each period, there exist allowed price-quality vectors which give the firm non-negative profit, or else the firm would leave the market. The exact way in which this parallel shift occurs in our mechanism is described in the following four steps. Throughout these four steps, qualities are fixed at the previous period levels $\mathbf{q}^{t-1}$.

Step 1: Choose r_V^t. r_V^t is the smallest value of $r \geq 1$ that solves the following equation:

$$\mathbf{x}^{t-1} \cdot \mathbf{x}^{-1}\left(r\mathbf{x}^{t-1}, \mathbf{q}^{t-1}\right) = \mathbf{x}^{t-1} \cdot \mathbf{p}^{t-1} - \pi^{t-1}.$$

Step 2: Choose r_0^t. r_0^t is the value of r that solves the following equation:

$$d((\mathbf{x}^{-1}(r\mathbf{x}^{t-1}, \mathbf{q}^{t-1}), \mathbf{q}^{t-1}), (\mathbf{p}^{t-1}, \mathbf{q}^{t-1})) = d\left(\left(\mathbf{p}^{t-1}, \mathbf{q}^{t-1}\right), \left(\mathbf{p}^{t-2}, \mathbf{q}^{t-2}\right)\right) \phi(\pi^{t-1}).$$

Step 3: Choose γ_0^t. Let γ_0^t be the value of γ that solves:

$$\mathbf{x}^{t-1} \cdot \mathbf{x}^{-1}\left(r_0^t\mathbf{x}^{t-1}, \mathbf{q}^{t-1}\right) = \mathbf{x}^{t-1} \cdot \mathbf{p}^{t-1} - \gamma\pi^{t-1}.$$

Step 4: Choose γ^t. $\gamma^t = \min\left\{1, \gamma_0^t\right\}$.

Figure 2 shows the determination of the firm's constraints in the case $n = 1$ and $m = 1$. Let A be the price-quality vector chosen in period $t-1$; point B is the price-quality vector where the *quality-adjusted VF constraint* (19) with $\mathbf{q}^t = \mathbf{q}^{t-1}$ and $\gamma^t = 1$ holds as an equality. At point B, the firm at least breaks even; this follows from the assumption of decreasing ray returns, as discussed above in the case of fixed qualities (see the paragraph following (16)). If the mechanism to regulate quality were the exact analogue to the VF mechanism, this would be it: γ^t would be 1, and the firm would face only a constraint given by the parallel shift of the hyperplane tangent to the isowelfare from point A to point B, and could choose any point in the half-space to the left of the shifted hyperplane (the grey area on panel a). However, as mentioned above, when quality varies, the distance constraint (18) (drawn as the ellipsis centred in A in panel a) is also needed. With this additional constraint, the feasible region is only the darker region in panel a.

It may be the case that, as drawn in panel a, point B belongs to the feasible region, and so there are points with non-negative profits in the

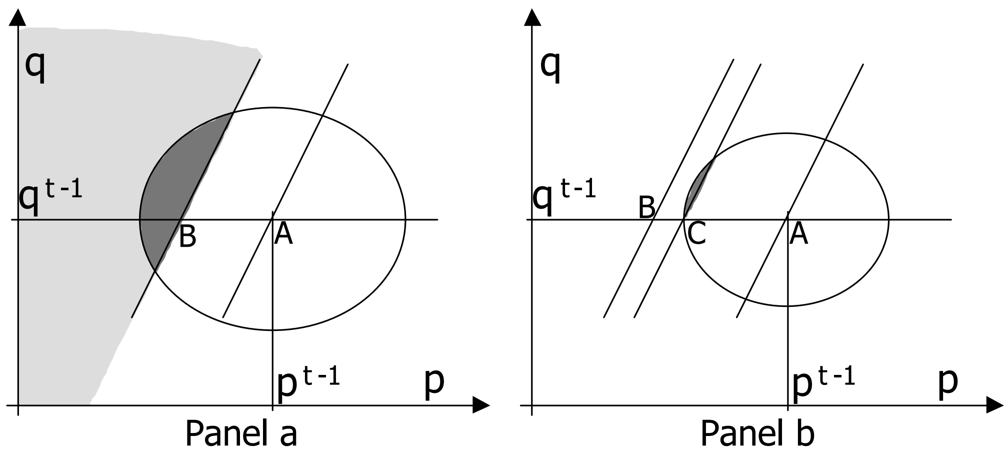

Figure 2: The determination of the factor γ.

feasible area. The distance constraint may however be too demanding, as depicted in panel b in Figure 2, where the ellipsis is smaller. In this case, if the hyperplane with slope $\left(\mathbf{x}^{t-1}, \mathbf{u}^{t-1}\right)$ through point A were shifted all the way to point B, that is, if γ^t were 1, the admissible region would be empty. We therefore take a lower γ^t to stop the shift at point C: this is point where, keeping qualities constant, prices are changed in such a way to reduce proportionally all quantities so as to satisfy the distance constraint (18) as an equality. This ensures that the firm's feasible choice set is non-empty: it is the dark area in panel b. Moreover, at point C the firm's profit is higher than at point B (this follows again from ray decreasing costs) and so the firm can choose price-quality vectors giving non-negative profits. Step 2 determines point C algebraically: r_0^t is the factor by which the output vector sold in $t-1$ must be scaled up at time t to satisfy constraint (18) as an equality, when the qualities are the same as in period $t-1$. Steps 3 and 4 simply ensure that if the point B is to right of C, then γ^t is 1, and vice versa, if point C is to the right of B, then the shift is to the intersection of the distance constraint and the ray where quantities are reduced proportionally.

We can now present the main contribution of the paper. As stated in Proposition 3, the mechanism constituted by constraints (18) and (19) has the same desirable properties in terms of convergence of the VF formula: a myopic firm which, period after period, maximises profit under these constraints chooses price-quality vectors which form a sequence converging to the zero-profit Ramsey price-quality vector.[11]

Proposition 3 *Let* $(\mathbf{p}^t, \mathbf{q}^t)$ *be the price-quality vector chosen by a firm subject in each period* t *to constraints (19) and (18). The sequence* $\left\{(\mathbf{p}^t, \mathbf{q}^t)\right\}_{t=1}^{\infty}$ *converges to the second best optimal price-quality vector* $(\mathbf{p}^*, \mathbf{q}^*)$.

Proof. The proof proceeds along the following steps:

Lemma 1 $\pi^t \geqslant 0$ *for every* $t > 0$.

Proof. We need to show that there exists a feasible price-quality vector which allows the firm to make non-negative profits. Rewrite constraint (19) as

$$\mathbf{x}^{t-1} \cdot \mathbf{p}^t \leqslant \mathbf{x}^{t-1} \cdot \mathbf{p}^{t-1} - \left(\pi^{t-1}\gamma - \mathbf{u}^{t-1} \cdot \left(\mathbf{q}^t - \mathbf{q}^{t-1}\right)\right). \tag{21}$$

Let the firm, in period t, choose qualities $\mathbf{q}^{t-1}$, and prices $\mathbf{x}^{-1}\left(r_V^t \mathbf{x}^{t-1}, \mathbf{q}^{t-1}\right)$, where r_V^t is given in Step 1 above. (21) can be re-written as:

$$\mathbf{x}^{t-1} \cdot \mathbf{x}^{-1}\left(r_V^t \mathbf{x}^{t-1}, \mathbf{q}^{t-1}\right) = \mathbf{x}^{t-1} \cdot \mathbf{p}^{t-1} - \gamma\pi^{t-1},$$

from which it follows that

$$\begin{aligned}
\mathbf{x}^{t-1} \cdot \mathbf{x}^{-1}(r_V^t \mathbf{x}^{t-1}, \mathbf{q}^{t-1}) &= \mathbf{x}^{t-1} \cdot \mathbf{p}^{t-1} - \pi^{t-1} + (1-\gamma^t)\pi^{t-1}, \\
\mathbf{x}^{t-1} \cdot \mathbf{x}^{-1}(r_V^t \mathbf{x}^{t-1}, \mathbf{q}^{t-1}) &= C(\mathbf{x}^{t-1}, \mathbf{q}^{t-1}) + (1-\gamma^t)\pi^{t-1}, \\
r_V \mathbf{x}^{t-1} \cdot \mathbf{x}^{-1}(r_V^t \mathbf{x}^{t-1}, \mathbf{q}^{t-1}) &= r_V^t C(\mathbf{x}^{t-1}, \mathbf{q}^{t-1}) + r_V^t(1-\gamma^t)\pi^{t-1}, \\
r_V \mathbf{x}^{t-1} \cdot \mathbf{x}^{-1}(r_V^t \mathbf{x}^{t-1}, \mathbf{q}^{t-1}) &\geqslant C(r_V^t \mathbf{x}^{t-1}, \mathbf{q}^{t-1}) + r_V^t(1-\gamma^t)\pi^{t-1}.
\end{aligned}$$

The last line follows from the assumption of decreasing (quantity) ray average costs. Rearranging,

$$r_V^t \mathbf{x}^{t-1} \cdot \mathbf{x}^{-1}\left(r_V^t \mathbf{x}^{t-1}, \mathbf{q}^{t-1}\right) - C(r_V^t \mathbf{x}^{t-1}, \mathbf{q}^{t-1}) \geqslant r_V^t \left(1-\gamma^t\right)\pi^{t-1} \geqslant 0.$$

In the above, the LHS is the profit obtained by the firm when it offers the vector $r_V^t \mathbf{x}^{t-1}$. The second inequality follows from the fact that $\gamma^t \in [0,1]$. This proves that the vector $r_V^t \mathbf{x}^{t-1}$ gives non-negative profits. The manner in which r_V^t is determined ensures that this price-quality vector is always in the firm's allowed region. ■

Clearly, the sequence $\{(\mathbf{p}^t, \mathbf{q}^t)\}_{t=1}^{\infty}$ is converging, by construction of the distance constraint (18), and the next lemma shows that it converges to a point where the firm makes no profit.

[11] As with VF, initial conditions are needed. If past data are missing or not suitable, the regulator can leave the firm unregulated in period 1, impose constraint (19) from period 2 and constraint (18) from period 3.

Lemma 2 *Let the sequence* $\{(\mathbf{p}^t, \mathbf{q}^t)\}_{t=1}^{\infty}$ *converge to the vector* $(\overline{\mathbf{p}}, \overline{\mathbf{q}})$. *Then,* $\pi(\overline{\mathbf{p}}, \overline{\mathbf{q}}) = 0$.

Proof. Assume not. Let $\pi(\overline{\mathbf{p}}, \overline{\mathbf{q}}) > 0$. At time $t+1$, because of constraint (19), the price-quality vector $(\overline{\mathbf{p}}, \overline{\mathbf{q}})$ is not allowed; a contradiction. ■

The following lemma is the final step in the argument.

Lemma 3 $(\overline{\mathbf{p}}, \overline{\mathbf{q}})$ *is a second best price-quality optimal vector.*

Proof. If the sequence of price-quality vectors has converged to point $(\overline{\mathbf{p}}, \overline{\mathbf{q}})$, then, in every period, the distance constraint (18) becomes redundant, in the sense that it identifies the single point $(\overline{\mathbf{p}}, \overline{\mathbf{q}})$ and is trivially satisfied. In what follows, therefore, we simply ignore constraint (18). The weights in constraint (19) satisfy $\overline{x}_i = x_i(\overline{\mathbf{p}}, \overline{\mathbf{q}})$ for all $i = 1, .., n$, and $\overline{u}_j = \frac{\partial v(\overline{\mathbf{p}}, \overline{\mathbf{q}})}{\partial q_j}$ for all $j = 1, .., m$. In period t, the firm's problem without constraint (18) is:

$$\max_{\mathbf{p}^t, \mathbf{q}^t} \pi(\mathbf{p}^t, \mathbf{q}^t) \quad \text{s.t.:} \quad \overline{\mathbf{x}} \cdot \mathbf{p}^t - \overline{\mathbf{u}} \cdot \mathbf{q}^t \leqslant C(\overline{\mathbf{x}}, \overline{\mathbf{q}}) - \overline{\mathbf{u}} \cdot \overline{\mathbf{q}}, \tag{22}$$

where $\overline{\mathbf{x}} = (\overline{x}_1, \ldots, \overline{x}_n)$ and $\overline{\mathbf{u}} = (\overline{u}_1, \ldots, \overline{u}_n)$. Denote with $\hat{\mathbf{p}}^t$ and $\hat{\mathbf{q}}^t$ the solutions to problem (22). These satisfy the following FOC:

$$\frac{\partial \pi(\hat{\mathbf{p}}^t, \hat{\mathbf{q}}^t)}{\partial \hat{p}_i^t} = \lambda \overline{x}_i \qquad \text{for all } i = 1, .., n;$$
$$\frac{\partial \pi(\hat{\mathbf{p}}^t, \hat{\mathbf{q}}^t)}{\partial \hat{q}_i^t} = \lambda \overline{u}_j \qquad \text{for all } j = 1, .., m;$$

where λ is the Lagrangean multiplier associated with the constraint in (22). By the assumption that $(\overline{\mathbf{p}}, \overline{\mathbf{q}})$ is a point of convergence, $(\hat{\mathbf{p}}^t, \hat{\mathbf{q}}^t) = (\overline{\mathbf{p}}, \overline{\mathbf{q}})$. Hence, at $(\hat{\mathbf{p}}^t, \hat{\mathbf{q}}^t)$, we must have that, as in Step 3 in VF (p 164), the surfaces $\{S(\mathbf{p}, \mathbf{q}) | S = S(\overline{\mathbf{p}}, \overline{\mathbf{q}})\}$, $\{\pi(\mathbf{p}, \mathbf{q}) | \pi = \pi(\hat{\mathbf{p}}^t, \hat{\mathbf{p}}^t)\}$, and the constraint in (22) are all tangent to each other. This, together with the result in Lemma 2 that $\pi(\overline{\mathbf{p}}, \overline{\mathbf{q}}) = 0$, establishes the Lemma and therefore Proposition 3. ■

Note that the distance constraint (18) shrinks the feasible region in each period, whereas the factor γ^t widens it: the combined effect on the feasible region, and therefore on the speed of convergence, is not determined in general.

It is worth commenting on the information requirements of our proposed mechanism. The quality weight u_j is the marginal social valuation of quality dimension j evaluated at the previous period price-quality vector. Graphically, it is the increase in the area between the demand curve and the price for all goods. We assume that the regulator knows this increase for the various goods. At a theoretical level, this assumption simply implies that the regulator knows its own payoff function: a standard minimal requirement of rationality. In practice, however, accurate acquisition of this information may be difficult. Indeed, regulators do exert considerable effort to learn the preferences of consumers regarding the quality of the regulated services;[12] however, it seems likely that introspection or some other source of the regulator's priori information will also influence the actual values $\mathbf{u}^t$ offered by the regulator to the firm. This is not a problem: the logic of our proposed mechanism (and of course of VF's too) applies also to the case where the regulator maximises a utility function which does not necessarily match the consumers', because of the difficulty to measure the willingness to pay for quality, or because the regulator is captured by the industry interest. Our suggested regulatory mechanism would still ensure convergence to the price-quality vector which maximises the regulator's payoff subject to the break even constraint. The quality weights u_j in this case would, analogously, be equal to the marginal effect of the change in the quality dimension j on the regulator's objective function, evaluated at the previous period price-quality vector.

4.3 The role of the distance constraint

Why do we need the distance constraint (18)? In special crcumstances, it is legitimate to leave it out, as shown by Proposition 4 below. However, with plausible assumption on consumers' preferences, in the absence of the distance constraint (18), the regulatory mechanism may backfire, as we show with a simple geometric example.

Proposition 4 *Let $v(\mathbf{p}, \mathbf{q})$ be convex. Let the regulated firm be subject in each period to the quality adjusted VF constraint only, (19), with $\gamma^t = 1$.*

[12] For instance, Ofgem has recently carried out market research to "identify the areas that consumers are most concerned with and their expectations and priorities for improvement" and to "determine consumers' willingness to pay for improvements in the key outputs identified in the first stage of the consumer research" (Ofgem, 2003, p 44). For a similar study in the water industry see MORI (2002). More generally, UK regulators appoint consumers' councils as recognised participants in the regulatory process.

The price-quality vector converges to the price-quality vector $(\mathbf{p}^*, \mathbf{q}^*)$.

Proof. The proof is an extension of the proof given by VF for their Proposition 1, to the case when quality also adjusts. Consider the set $\Pi_0 = \{(\mathbf{p}, \mathbf{q}) | \pi(\mathbf{p}, \mathbf{q}) \geqslant 0\}$: this is the set of price-quality vectors giving non-negative profit. This set is compact. An argument analogous to the proof of Lemma 1 in Proposition 3 shows that when the feasible region for the firm is given simply by the shift of (19) by r_V^t (that is in the absence of the distance constraint (18) and for $\gamma^t = 1$), there are points giving positive profit, or, in other words, the sequence $\{(\mathbf{p}^t, \mathbf{q}^t)\}_{t=1}^{\infty}$ is in Π_0. We now show that the sequence $\{v(\mathbf{p}^t, \mathbf{q}^t)\}_{t=1}^{\infty}$ is monotonic and therefore that it has a unique accumulation point in Π_0. The convexity of $v(\mathbf{p}, \mathbf{q})$ implies that

$$v(\mathbf{p}^t, \mathbf{q}^t) - v(\mathbf{p}^{t-1}, \mathbf{q}^{t-1}) \geqslant \nabla v(\mathbf{p}^{t-1}, \mathbf{q}^{t-1}) \begin{bmatrix} \mathbf{p}^t - \mathbf{p}^{t-1} \\ \mathbf{q}^t - \mathbf{q}^{t-1} \end{bmatrix}. \tag{23}$$

Where $\nabla v(\mathbf{p}^{t-1}, \mathbf{q}^{t-1})$ is the gradient of the function v at $(\mathbf{p}^{t-1}, \mathbf{q}^{t-1})$. This satisfies $\nabla v(\mathbf{p}^{t-1}, \mathbf{q}^{t-1}) = (-\mathbf{x}^{t-1}, \mathbf{u}^{t-1})$. Multiplying through the RHS, (23) can be rewritten as:

$$v(\mathbf{p}^t, \mathbf{q}^t) - v(\mathbf{p}^{t-1}, \mathbf{q}^{t-1}) \geqslant -\mathbf{x}^{t-1} \cdot (\mathbf{p}^t - \mathbf{p}^{t-1}) + \mathbf{u}^{t-1} \cdot (\mathbf{q}^t - \mathbf{q}^{t-1}). \tag{24}$$

Next note that (19) can be written as $-\mathbf{x}^{t-1} \cdot (\mathbf{p}^t - \mathbf{p}^{t-1}) + \mathbf{u}^{t-1} \cdot (\mathbf{q}^t - \mathbf{q}^{t-1}) \geqslant \pi^{t-1}$, and therefore (24) implies $v(\mathbf{p}^t, \mathbf{q}^t) - v(\mathbf{p}^{t-1}, \mathbf{q}^{t-1}) \geqslant \pi^{t-1} \geqslant 0$. This shows that the sequence $\{S(\mathbf{p}^t, \mathbf{q}^t)\}_{t=1}^{\infty}$ is monotonic and therefore it converges. Moreover, we have

$$0 = \lim_{t \to \infty} \left[v\left(\mathbf{p}^t, \mathbf{q}^t\right) - v\left(\mathbf{p}^{t-1}, \mathbf{q}^{t-1}\right) \right] \geqslant \lim_{t \to \infty} \pi^{t-1},$$

and, so in the limit, the firm makes zero profit. All there remains to show is that the firm's choice converges to a second best optimum. This is identical to Step 3 in VF (p 164): if $(\overline{\mathbf{p}}, \overline{\mathbf{q}})$ is the accumulation point of the sequence $\{(\mathbf{p}^t, \mathbf{q}^t)\}_{t=1}^{\infty}$, then the surfaces $\{v(\mathbf{p}, \mathbf{q}) | S = S(\overline{\mathbf{p}}, \overline{\mathbf{q}})\}$, $\{\pi(\mathbf{p}, \mathbf{q}) | \pi = \pi(\overline{\mathbf{p}}, \overline{\mathbf{q}}) = 0\}$, and the constraint are all tangent to each other. Finally, if the sequence $\{(\mathbf{p}^t, \mathbf{q}^t)\}_{t=1}^{\infty}$ does not converge, then the argument used in Iozzi *et al.* (2002, Proposition 2a, p 113) shows that the limit points of any convergent subsequence of the sequence $\{(\mathbf{p}^t, \mathbf{q}^t)\}_{t=1}^{\infty}$ must satisfy all the first order conditions for a second best optimum. ∎

Thus, if the welfare function is convex, including quality in the price cap in the way proposed in (19) is sufficient for the firm's choices to tend to the second best. The geometric reason is that, exactly as in the VF mechanism (which assumes that the welfare function is convex), the hyperplane tangent

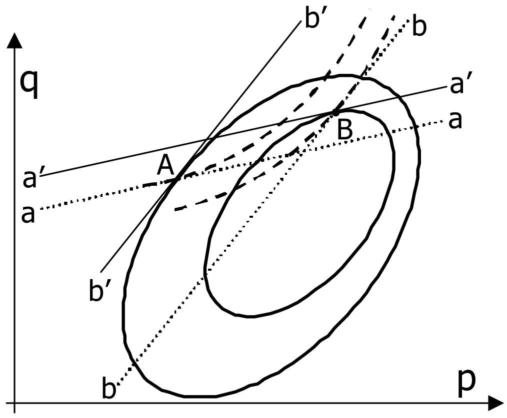

Figure 3: The possibility of regulatory cycles.

to the isowelfare locus at a given point B contains only points which are "welfare better" than point B. Therefore, welfare increases in every period (the proof uses this fact, see (24)).

While the assumption of convexity of the consumers' surplus in prices made by VF is a consequence of the natural assumption that the demand functions be decreasing, and is therefore realistic, when the consumers' utility depends on prices and qualities and the firm can choose both (prices and quantities), convexity cannot be warranted. If the regulator objective function is not convex, the isowelfare loci need not be concave, as argued in the discussion of Figure 1, and the half-space defined by constraint (19) – which is an upward shift of the hyperplane tangential to the isowelfare through the point chosen in the previous period – contains points which are "welfare worse" than the previous period choice.

Without the additional constraint (18), the regulatory process may converge to a cycle: a geometric example suffices to illustrate the point. Consider Figure 3. As before, the isoprofit (isowelfare) loci are the solid (dashed) curves. Suppose the firm starts from point A. Here tangent to the isowelfare is the line aa, and, because profit is strictly positive, the constraint is shifted up, to, say, position $a'a'$. With this constraint, the firm will choose point B. Note that welfare is lower at point B than at point A; as mentioned above, this could not happen if the welfare function were convex. At B, the isowelfare is tangent to line bb; profit at point B is higher than at A, and therefore strictly positive, and the constraint is shifted up, say to position $b'b'$. From here the firm would choose again point A, and the cycle is repeated. Note that, apart from having positive profit, the cycle constituted by points A and B is such that a Pareto improvement is possible: both the firm and the regulator can be made better off in each period if a different point were chosen.

4.4 Delay in the observation of quality.

In the VF mechanism, whether the choice of the firm satisfies the regulatory constraint is known at the beginning of each period. This is because the firm can choose exactly the value of the variables which determine the constraint, that is its prices. These variables can therefore be exactly determined for the whole period. This is also true of some quality measures (such as the number of check-in desks at an airport, or whether bills can be paid by credit card), but, by their very nature, some of the quality measures used in practice (such as the punctuality of trains or the number of accidents on motorways) are affected by random elements which are beyond the firm's control, and which cannot be known in advance, but can only be observed *ex-post*.[13] In practice, therefore, regulators use past indicators of quality. The mechanism proposed in Section 4.2 can be easily adapted to this situation by replacing constraints (18) and (19) with:

$$d((\mathbf{p}^t, \mathbf{q}^{t-1}), (\mathbf{p}^{t-1}, \mathbf{q}^{t-2})) \leqslant d((\mathbf{p}^{t-1}, \mathbf{q}^{t-2}), (\mathbf{p}^{t-2}, \mathbf{q}^{t-3}))\phi(\pi^{t-1}), \tag{25}$$

and

$$\frac{\mathbf{x}^{t-1} \cdot \mathbf{p}^t}{\mathbf{x}^{t-1} \cdot \mathbf{p}^{t-1}} \leqslant 1 - \left(\frac{\pi^{t-1}}{R^{t-1}} \gamma^{t-1} - \frac{\mathbf{u}^{t-1} \cdot (\mathbf{q}^{t-1} - \mathbf{q}^{t-2})}{R^{t-1}} \right). \tag{26}$$

The difference with constraints (18) and (19) is simply that all quality measures are delayed by one period: $\mathbf{q}^{t-1}$ replaces $\mathbf{q}^t$, $\mathbf{q}^{t-2}$ replaces $\mathbf{q}^{t-1}$, and $\mathbf{q}^{t-3}$ replaces $\mathbf{q}^{t-2}$.

[13] For an analysis of different regulatory mechanisms for quality in the context of price cap regulation, see Weisman (2000).

Proposition 5 *Let* $(\mathbf{p}^t, \mathbf{q}^t)$ *be the price-quality vector chosen by a firm subject in each period* t *to constraints (25) and (26). The sequence* $\left\{(\mathbf{p}^t, \mathbf{q}^t)\right\}_{t=1}^{\infty}$ *converges to the second best optimal price-quality vector* $(\mathbf{p}^*, \mathbf{q}^*)$.

The proof is essentially identical to the proof of Proposition 3, and is omitted. Convergence under our mechanism is therefore ensured even in the plausible case when quality is observed only *ex-post*, and therefore this need not diminish the applicability of our proposed mechanism in practice.

5 Concluding remarks.

We propose a regulatory mechanism for firms operating in industries where the quality of the output is important. In our mechanism, the firm is constrained to choose, in each period, qualities and prices satisfying two constraints: the *quality adjusted Vogelsang-Finsiger constraint*, given by (19), and the *distance* constraint, (18). Our contribution is an extension of the regulatory mechanism suggested by VF and of its practical counterpart, RPI-X price cap, and it is based on the same principle of delegating economic decisions to the firm: the firm can freely make its choices on the economic variables under its control, subject to a constraint over their average value, with this constraint tightening over time. In the VF mechanism, the tightening of the constraint in each period is automatically determined by the rate of profit over sales in the past period. In the practice of RPI-X price cap regulation, the extent by which the firm must reduce its average price is exogenously given by X and negotiated with the regulator every few years. In some cases, the average price level permitted to the firm, as determined by X, is made dependent also on the quality of the services provided by the firm itself. We provide a robust theoretical foundation to this relationship between the average price and the quality of the services provided by the regulated firm. In the mechanism proposed here, the firm may exceed the value of the index which would be allowed by the VF mechanism, if it increases the quality of its output: the regulated firm can "sell" higher quality to consumers, to the extent allowed by (19); conversely, of course, the firm can lower the quality of its output, but must "bribe" the consumers via lower average prices. Deviations from the average price which the VF mechanism would impose are permitted through a quality adjustment term added to the X factor: this term is a weighted average of the marginal effects of the quality changes on consumers' welfare and it is not in any way dependent on the firm's cost, typically unknown to the regulator. In analogy to the RPI–X price cap and the VF mechanism, where the weight of each price in the price average is given by the social valuation in the previous period, the weight of each quality measure in the quality adjustment term added to the X factor is given by the social valuation of that dimension of quality, measured by the social value at last period's prices and qualities. We also show that, when the regulator's objective function is not quasi-convex, which is well possible when quality affects consumers' utility, this correction of the VF mechanism may not counterproductive: the regulated firm's choices may be locked in a Pareto inefficient cycle. An additional requirement, that the firm's choices not to be "too erratic", addresses this problem.

Our mechanism formalises and gives a rigorous foundation to the practice followed by some regulators to let the factor X in the RPI-X price cap depend on quality improvements. It has the same immediate practical applicability, without some of the disadvantages. One the one hand, the inclusion of the distance constraint addresses the problem caused by hitherto unnoticed possibility of Pareto inefficient cycles. On the other hand, it allows the regulator to offer the firm a formula with which quality changes can be valued, thus avoiding both the need of regular negotiations on the weights of the quality measures, and the uncertainty associated with these.

References

AEEG (2000). *Decision 237/2000: Criteria to determine the tariffs for gas distribution and supply to the customers of the captive market.* Italian Gas and Electricity Authority, Milan.

Armstrong, M., S. Cowan, and J. Vickers (1994). *Regulatory reform: Economic analysis and British experience.* MIT Press, Cambridge, Massachusetts.

Armstrong, M. and D. E. M. Sappington. Recent developments in the theory of regulation. In M. Armstrong and R. Porter (Eds.), *Handbook of Industrial Organization, Vol. III.* North-Holland, New York and Amsterdam.

Armstrong, M. and J. Vickers (2000). Multiproduct price regulation under asymmetric information. *Journal of Industrial Economics 48*, 137–160.

Banerjee, A. (2003). Does incentive regulation 'cause' degradation of retail telephone service quality? *Information Economics and Policy 15*, 243–269.

Baron, D. P. and R. B. Myerson (1982). Regulating a monopolist with unknown cost. *Econometrica 50*, 911–930.

Beil, R. O., D. L. Kaserman, and J. M. Ford (1995). Entry and product quality under price regulation. *Review of Industrial Organization 10*, 361–372.

Bös, D. (1981). *Economic Theory of Public Enterprise.* Springer Verlag, Berlin.

Bradley, I. and C. Price(1989). The economic regulation of private industries by price constraints. *Journal of Industrial Economics*, 99–106.

Brennan, T. J. (1989). Regulating by capping prices. *Journal of Regulatory Economics 1*, 133–47.

Gabszewicz, J. J. and J. F. Thisse (1979). Price competition, quality and income disparities. *Journal of Economic Theory 20*, 340–59.

Iozzi, A. (2002). La riforma della regolazione nel settore autostradale. *Economia Pubblica 32*, 71–93.

Iozzi, A., J. A. Poritz, and E. Valentini (2002). Social preferences and price cap regulation. *Journal of Public Economic Theory 4*, 95–114.

Kelley, J. L. (1955). *General Topology.* D. van Nostrand Company, Amsterdam.

Laffont, J.-J. (1994). The new economics of regulation ten years after. *Econometrica 62*, 507–537.

Laffont, J.-J. and J. Tirole (1991). Provision of quality and power of incentive schemes in regulated industries. In J. J. Gabszewicz and A. Mas-Colell (Eds.), *Equilibrium theory and applications*, pp. 137–56. Cambridge University Press, Cambridge, England.

Laffont, J.-J. and J. Tirole (1993). *A theory of incentives in procurement and regulation*. MIT Press, Cambridge, Massachusetts.

Ma, C. A. and J. F. Burgess, Jr. (1993). Quality competition, welfare, and regulation. *Journal of Economics 58*, 153–173.

MORI (2002). *The 2004 Periodic Review: Research into Customers' Views*.

NMPRC (2002). *Annual report 2002*. Santa Fe (NM): New Mexico Public Regulation Commission.

Oates, W. E. and P. R. Portney (2003). The political economy of environmental policy. In K.-G. Mäler and J. Vincent (Eds.), *Handbook of Environmental Economics*, pp. Ch 8. North-Holland/Elsevier Science, Amsterdam.

OFGEM (1999). *Review of Public Electricity Suppliers 1998 to 2000. Distribution Price Control Review. Final Proposals*. London: OFGEM.

OFGEM (2001). *Information and incentives project. Incentive schemes: Final proposals*. London: OFGEM.

OFGEM (2003). *Electricity Distribution Price Control Review. Initial consultation*. London: OFGEM.

OFTEL (2001). *Competition in the provision of fixed telephony services. . Consultation document issued by the Director General of Telecommunications*. London: OFTEL.

OFWAT (1998). *1997/98 Report on Levels of Service for the Water Industry in England and Wales*. London: OFWAT.

OFWAT (1999). *Future water and sewerage charges 2000-05: Final determinations*. London: OFWAT.

OFWAT (2002). *Linking service levels to prices*. London: OFWAT.

Rovizzi, L. and D. Thompson (1992). The regulation of product quality in the public utilities and the citizen's charter. *Fiscal Studies 13*, 74–95.

Sappington, D. E. M. (2003). The effects of incentive regulation on retail telephone service quality in the united states. *Review of Network Economics 2*.

Shaked, A. and J. Sutton (1982). Relaxing price competition through product differentiation. *Review of Economic Studies 49*, 3–14.

Sheshinski, E. (1976). Price, quality and quantity regulation in monopoly situations. *Economica 43*, 127–137.

Spence, M. (1975). Monopoly, quality and regulation. *Bell Journal of Economics 6*, 417–429.

Stigler, G. J. (1971). The theory of economic regulation. *Bell Journal of Economics 2*, 3–21.

Vogelsang, I. (1989). Price cap regulations of telecommunications services: A long run approach. In M. A. Crew (Ed.), *Deregulation and Diversification of Utilities*. Kluwer, Boston.

Vogelsang, I. and J. Finsinger (1979). A regulatory adjustment process for optimal pricing by multiproduct monopoly firms. *Bell Journal of Economics 10*, 157–71.

Weisman, D. L. (2002). Price regulation and quality. mimeo.