Regulating a multi-attribute/multi-type Monopolist*

Dezsö Szalay
University of Bonn

May 26, 2009

Abstract

I study the regulation of a firm producing a good with two attributes, e.g. quantity and quality. The firm has private information about its cost of production and the maximum quality it is able to produce. If quality and quantity are complementary and the reverse hazard rate of the conditional distribution of cost is increasing in the quality parameter, then asymmetric information about the quality capacity is irrelevant, the optimal allocation is the same as if the quality capacity were known to the regulator. If quality and quantity are substitutable and the reverse hazard rate is decreasing in the quality parameter, then shirking on quality provision becomes relevant. To prevent such shirking, high quality firms price below marginal cost and all of them - including the highest cost producer among them - receive a rent; low quality producers all price above marginal costs.

JEL: D82, L21, Asymmetric Information, Multi-dimensional Screening, Regulation.

1 Introduction

When duplication of fixed costs is wasteful, a service is efficiently provided by a natural monopoly. To keep the service provider from abusing its monopoly power, the pricing of the firm is regulated. If the regulator had access to the firm's information, regulation would be a trivial matter. As is well known, the firm should follow a marginal cost pricing rule and should be subsidized for the losses it makes on a lump-sum basis. However, the problem is precisely that the firm has better information than the regulator has about payoff relevant circumstances. Baron and Myerson [1982] first analyzed this problem when the firm's costs are unknown to the regulator but known to the firm. Marginal cost pricing is no longer optimal as this rule gives firms with relatively low marginal costs incentives to exaggerate their costs in order to get larger subsidies. To make such exaggeration unattractive, prices are distorted upwards for all but the most efficient firms.

In many industries regulation is more complex than deciding how much to produce at what price. In rail transport not only the number of trains per hour is important but also their timeliness; cleanness of water is important in addition to its quantity and price; likewise, the reliability of

*This paper supersedes an earlier version, which was joint with Charles Blarckorby and entitled "Regulating a Monopolist with unknown costs and unknown quality capacity". The earlier paper contains the analysis of the case where asymmetric information about quality information is irrelevant; the analysis of the case with a binding quality-shirking constraint was done after Chuck opted out of the project. I thank him for his contribution to this project. Special thanks also to Benny Moldovanu for pointing me to hazard rate orders, which I had derived endogenously at the unnecessarily high cost of assuming affiliation in the earlier version. I also thank Paul Beaudry, Claudio Mezzetti, and seminar participants at HEC Lausanne and the University of Bonn for helpful conversations.
Correspondence can be sent to Dezso Szalay, Chair of Economic Theory I, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany, or to szalay@uni-bonn.de

electricity supply or the safety of air traffic control are vital issues. Thus it seems both natural and important to address the regulation of firms in a multidimensional world, where service provision has at least two dimensions. Moreover, it is just as natural to suppose that firms have better information about their production possibility frontier with respect to the production of both attributes. This is the problem I address in this paper.

In my extension of the Baron and Myerson [1982] model a monopolist provides a two-attribute service. To fix ideas I interpret the first attribute as a quantity and the second one as a quality, but the model is not at all confined to this interpretation. The monopolist knows the marginal costs of production the maximum quality he is able to produce. The regulator has neither the marginal cost nor the capacity constraint information. The regulator can observe the quantity, the quality, and the price of the service the consumers buy. Hence, the regulator's problem is to choose these three instruments, and in addition a lump-sum subsidy to the firm to cover its losses. Although the regulator can observe the quality the firm produces, he cannot observe whether the firm produces at its quality capacity or below; that is, whether or not the firm shirks on quality. Moreover, the firm's true cost of production are not observable to the regulator. Among other things, my model allows me to address the following list of questions:

How do incentives to overstate costs interact with incentives to shirk on quality? How does the presence of a second source of asymmetric information affect the pricing rules? Is private information about the quality capacity a source of additional rents for the firms? How does the presence of the quality dimension affect incentives to exclude some inefficient firms? The answers to these questions depend crucially on how the two attributes interact in the social surplus function and on how the two informational variables interact in their joint distribution.

If quality and quantity are net complements and the distribution of marginal cost conditional on a high quality capacity has a higher reversed hazard rate than the distribution of marginal cost conditional on a low quality capacity, then the presence of quality information comes at no cost at all. Everything is as if firms' capacity bounds were known; pricing rules are exactly the ones for known quality bounds; all firms except the least cost producer price above marginal cost, and the least efficient firm receives no rent. Moreover, firms voluntarily select into "their" pricing scheme. The intuition for this self selection result is that firms have no incentives to shirk on quality simply because that would reduce their rents. Finally, as everything is exactly as if the quality capacity were common knowledge, the presence of quality information cannot affect the shutdown decision either.

If quality and quantity are net substitutes the reverse hazard rate is decreasing in the quality capacity, then nothing is as if firms' capacity bounds were known. At the optimum some high quality producers price below marginal costs. It is even possible that all but the least cost efficient high quality producers are induced to set prices below marginal cost; all low quality producers are induced to set prices above marginal cost. Moreover, it is possible that producers of high quality *all* receive a rent. That is, there is *a rent at the bottom* of high quality producers. In short, all of the well known features of one-dimensional asymmetric information models no longer hold in two dimensions. Obviously, the intuition for these results goes back to the shirking on quality constraint that becomes binding in this case. The informational rents producers obtain depend on the quantities they sell; hence to prevent high quality producers from shirking on quality, their production quantities are raised relative to the quantities produced by low quality producers. To induce such choices by the consumers, prices of high quality producers are lowered and prices of low quality producers are raised. Although shutdown is ruled out by a boundary condition in my model, this reasoning suggests that the addition of the second dimension would affect the shutdown decision absent such a boundary condition: some high-cost/low-quality producers might be excluded even though they would be allowed to produce when their quality capacity were known to the regulator.

These results confirm findings by Lewis and Sappington [1988] and Armstrong [1999], although I obtain mine using very different verifiability assumptions and techniques. Lewis and Sappington [1988] assume that the firm knows the intercept of a linear demand function and the value of its marginal cost parameter, while the regulator does not have any of this information. If the regulator were able to observe quantity and price in their model, demand information would not add anything to the one-dimensional case; the regulator could simply infer the demand intercept from quantity and price information. Therefore, they investigate the case where the regulator cannot observe the quantity consumers purchase, but he can verify whether or not customers have been served. This problem is amenable to techniques developed in Laffont, Maskin and Rochet [1987]. In particular, the regulator has only one instrument - the marginal price - to screen firms, but information has two dimensions. Hence, firms whose demand and cost parameter add up to the same value behave the same way; that is, these firms are bunched together. As I do, they show that optimal pricing is strikingly different in this world; some firms are induced to set prices below marginal costs. However, due to the bunching feature of their model, the reasons for this pricing below marginal cost cannot be traced back to individual dimensions of the problem. This is different in the present problem which can be analyzed as a family of one-dimensional problems constrained by each other.

Armstrong [1999] reinvestigates the Lewis and Sappington [1988] problem, showing that the optimal pricing policy in the Lewis and Sappington [1988] model induces exclusion of certain high cost-low demand firms. Technically, exclusion is optimal because the density of the sum of two random variables goes to zero at the bounds of its support. Armstrong [1996] shows that exclusion is robust in these kind of settings under more general assumptions. Armstrong [1999] confirms in a two-by-two-type model that optimal pricing can indeed be below marginal costs for some types in this context. An important insight from the present investigation is that these findings remain intact even under very different verifiability conditions.

Some of the techniques I am using here were introduced in Beaudry, Blackorby, and Szalay [2009] in the context of a taxation model where workers could choose jobs requiring different abilities in addition to the number of hours worked on the job. This paper extends the previous work in many important dimensions. The characterization of incentive compatible allocations in section 4.1.of the present paper is given for arbitrary allocations while Beaudry, Blackorby, and Szalay [2009] exploit that capacity constraints are binding in that model at the optimum. The most important addition to the previous work is that I am able to solve the case of a binding shirking constraint which was not discussed in Beaudry, Blackorby, and Szalay [2009].

This study is related to a growing literature on multi-dimensional screening. Closest in terms of focus - the Lewis and Sappington [1988] and Armstrong [1999] papers notwithstanding - is Matthews and Moore [1987]. The main difference to Matthews and Moore is that in their paper information is one-dimensional but there are two instruments to choose; in my paper both information and the set of instruments is two-dimensional. Other studies in multi-dimensional screening include McAfee and MacMillan [1988], and Jehiel, Moldovanu, and Stacchetti [1999]. Wilson [1993] offers a general solution (the demand profile approach) to the multi-dimensional pricing problem. The most general results to date have been obtained by Armstrong [1996] and Rochet and Choné

[1998].[1]

The paper is organized as follows. In Section two I lay out the model, explain the regulator's problem and explain its solution for the case where the regulator has perfect information. In Section three, I study the case where the regulator knows the firms' quality capacity, but does not have access to the firm's cost information. I show that my problem is nicely amenable to monotone comparative statics methods, and use these methods to investigate how the pricing policy depends on the firm's quality capacity. In Section four, I address the full problem when the regulator knows neither the cost parameter nor the firm's quality capacity. In particular, this section contains equivalent representations of the incentive and participation constraints. In Section five, I provide closed form solutions for a number of interesting cases with binding capacity constraints. In Section six, I discuss the case of nonbinding capacity constraints. Depending on the support of the quality information and on the joint-distribution of quality capacity and marginal costs, I am able to fully characterize the optimal pricing policy. The final section concludes. All proofs, with the exception of one, have been relegated to the appendix.

2 The model and the main assumptions

Let $V(x,q)$ denote consumers' valuations for a good with attributes x and q. If x is the quantity of the good and q its quality, then $V(x,q)$ is the area below the inverse demand function $P(x,q)$, that is $V(x,q)$ is defined as the gross consumer surplus of a consumer who buys x units of a good of quality q at a constant marginal price

$$V(x,q) \equiv \int_0^x P(z,q)\,dz.$$

If the two attributes refer to quantities of goods that could just as well be traded independently of each other, then I take

$$V(x,q) \equiv \int_0^x P_1(z)\,dz + \int_0^q P_2(z)\,dz$$

where $P_1(x)$ and $P_2(q)$ denote the inverse demand functions for each good separately. I assume that $\lim_{x\to 0} P_1(x) = \lim_{q\to 0} P_2(q) = \infty$ and that for any $q>0$, $\lim_{x\to 0} P(x,q) = \infty$, as would be the case, e.g., for a constant elasticity of demand function. $V(x,q)$ is increasing in x and q. I mostly stick to the quantity/quality interpretation, but this is just to ease language; all the results apply to both settings.

The good is produced by a monopoly firm subject to price regulation. The firm's cost of producing the good in quantity x and quality q is

$$\mathcal{C}(x,q,\theta,\eta) = \begin{cases} C(x,q,\theta) & \text{for } q \le \eta \\ \infty & \text{for } q > \eta. \end{cases}$$

where θ and η are parameters that are known only to the firm but not to the regulator and

$$C(x,q,\theta) = K + c(x,\theta) + k(x,q).$$

[1]For excellent surveys of multidimensional problems, see Armstrong and Rochet [1999] and Rochet and Stole [2003].

I assume that V and C are differentiable to the order needed with respect to all their arguments.[2] K is a fixed cost and the functions c and k satisfy $c(0,\theta) = k(0,q) = k(x,0) = 0$. θ shifts the cost function on the relevant domain upwards, $c_\theta(x,\theta) > 0$ for all θ and $x > 0$. Costs are strictly increasing in x and weakly increasing in q. The additive separability of the cost function rules out any interaction between θ and q, which greatly simplifies the analysis. I impose the standard Spence-Mirrlees condition on the cost function, that is $c_{x\theta}(x,\theta) > 0$ for all (x,θ). Moreover, to ensure that stationary points of the model are maxima I assume throughout that $c_{x\theta\theta}(x,\theta) \geq 0$ and $c_{xx\theta}(x,\theta) \geq 0$ for all (x,θ).

η defines an upper bound on qualities that the firm is capable of producing. x and q are verifiable; hence contracts can be written on these variables. The parameters θ and η are known only to the firm; the regulator knows only the joint distribution of these variables. θ and η are distributed on a product set $\mathbf{\Theta} \times \mathbf{H}$ with probability density function $f(\theta,\eta) > 0$ for all θ,η. The set $\mathbf{\Theta}$ is taken as the interval $\left[\underline{\theta},\overline{\theta}\right]$ throughout the paper, where $\underline{\theta} > 0$. The density of the marginal distribution of θ is denoted $f(\theta)$. The set $\mathbf{H}$ can either be discrete or continuous. In the latter case I take it as the interval $\left[\underline{\eta},\overline{\eta}\right]$ where $\underline{\eta} > 0$. Let $G(\eta)$ denote the cdf of η. Given the full support assumption, for each η that has $dG(\eta) > 0$, the conditional distribution of θ given η has full support. The density and cdf of this distribution are denoted $f(\theta\,|\eta)$ and $F(\theta\,|\eta)$, respectively. Let $\mathcal{E}$ denote the expectation operator and let $f(\theta) \equiv \mathcal{E}_H\left[f(\theta\,|\eta)\right]$ and $F(\theta)$ denote the density and the cdf of the marginal distribution, respectively. I assume throughout the paper that the (inverse) reversed hazard rates of these distributions satisfy monotonicity conditions. In particular, I assume that $\frac{\partial}{\partial\theta}\frac{F(\theta|\eta)}{f(\theta|\eta)} \geq 0$ and $\frac{\partial}{\partial\theta}\frac{F(\theta)}{f(\theta)} \geq 0$.[3] Moreover, $\frac{F(\theta|\eta)}{f(\theta|\eta)}$ can be ordered with respect to η.

The firm is subject to price and quality regulation. However, it is equivalent and notationally much more convenient to analyze the model directly in terms of quantity and quality regulation. If the firm produces a quantity x and a quality q then it receives a subsidy t and its profit is for $q \leq \eta$

$$t + P(x,q)\,x - C(x,q,\theta).$$

For $q > \eta$, the profit becomes minus infinity.

Define the sum of consumer and producer surplus as

$$S(x,q,\theta) \equiv V(x,q) - C(x,q,\theta) \text{ for } q \leq \eta.$$

Note that $S(x,q,\theta)$ satisfies the boundary condition, $\lim_{x\to 0} S_x(x,q,\theta) = \infty$ for all θ and all $q > 0$. In addition, I assume that either i) the surplus function is concave in x and strictly increasing in q for all x below some upper bound $\bar{x}$; or ii) that the surplus function is strictly concave in x and q jointly, that is, for all x,θ, and $q \leq \eta$, $S_{xx}(x,q,\theta), S_{qq}(x,q,\theta) < 0$, and $S_{xx}(x,q,\theta)\,S_{qq}(x,q,\theta) - \left(S_{xq}(x,q,\theta)\right)^2 > 0$. The former case gives rise to corner solutions for the optimal quality provision; in the latter case interior solutions arise.

I maintain these assumptions throughout the paper without listing them in each result I state. Perhaps with the exception of additive separability of the cost function, all of them are standard in the literature to ensure the regularity of the maximization problems I solve.[4] Additive separability of the cost function is of course quite restrictive, although the assumption is frequently and for

[2] The differentiability of surplus and costs with respect to x and θ is crucial for the approach; the differentiability with respect to q is not and could be dispensed with at the cost of additional notational clutter.

[3] Interestingly, the former condition is not sufficient for the latter.

[4] See, e.g., Laffont and Tirole [1993].

good reasons imposed in the literature.

2.1 The Regulator's Problem

I think of the regulator's problem in term's of a direct revelation mechanism, which is a triple of functions $\{q(\theta,\eta), x(\theta,\eta), t(\theta,\eta)\}$ for all $(\theta,\eta) \in \boldsymbol{\Theta} \times \mathbf{H}$ that satisfy incentive compatibility constraints.[5] The regulator maximizes a weighted sum of net consumer surplus and producer surplus under incentive constraints. If a firm announces costs $\hat{\theta}$ and quality capacity $\hat{\eta}$, where $q\left(\hat{\theta},\hat{\eta}\right) > \eta$, then its profits are minus infinity; if it announces a type such that $q\left(\hat{\theta},\hat{\eta}\right) \leq \eta$, then its profits are given by

$$\Pi\left(\hat{\theta},\theta,\hat{\eta}\right) \equiv t\left(\hat{\theta},\hat{\eta}\right) + P\left(x\left(\hat{\theta},\hat{\eta}\right), q\left(\hat{\theta},\hat{\eta}\right)\right) x\left(\hat{\theta},\hat{\eta}\right) - C\left(x\left(\hat{\theta},\hat{\eta}\right), q\left(\hat{\theta},\hat{\eta}\right), \theta\right).$$

Under a truthful mechanism, the weighted joint surplus for a given pair (θ,η) is equal to

$$W(\theta,\eta) \equiv V\left(x\left(\hat{\theta},\hat{\eta}\right), q\left(\hat{\theta},\hat{\eta}\right)\right) - P\left(x\left(\hat{\theta},\hat{\eta}\right), q\left(\hat{\theta},\hat{\eta}\right)\right) x\left(\hat{\theta},\hat{\eta}\right) - t(\theta,\eta) + \alpha\Pi(\theta,\theta,\eta)$$

where $\alpha \in (0,1)$. Since α is kept constant throughout the paper, I suppress the dependence of the welfare function on α in what follows. I let Θ and H denote the random variables with typical realizations θ and η, respectively, and let $\mathcal{E}_{\Theta H}$ denote the expectation operator taken over the random variables Θ and H. The regulator solves problem

$$\mathrm{P} \equiv \max_{x(\cdot,\cdot),q(\cdot,\cdot),t(\cdot,\cdot)} \mathcal{E}_{\Theta H} W(\theta,\eta) \tag{1}$$

s.t. for all θ,η and all $\hat{\theta},\hat{\eta}$ such that $q\left(\hat{\theta},\hat{\eta}\right) \leq \eta$

$$\Pi(\theta,\theta,\eta) \geq \Pi\left(\hat{\theta},\theta,\hat{\eta}\right), \tag{2}$$

and for all θ,η

$$\Pi(\theta,\theta,\eta) \geq 0, \tag{3}$$

and

$$q(\theta,\eta) \leq \eta. \tag{4}$$

(2) is the incentive compatibility condition, requiring that a firm of type θ,η must have no incentive to mimic any other type which itself produces a quality the firm is able to produce too. Which types fall into this category depends on the allocation of qualities, $q(\theta,\eta)$. Of course, the firm must also have no incentive to mimic types such that $q\left(\hat{\theta},\hat{\eta}\right) > \eta$. However, the profit arising from such a choice is equal to minus infinity, which is obviously never tempting. In fact, the constraint (3) requires that each firm in equilibrium obtain a non-negative profit, which is surely better than imitating a type $\left(\hat{\theta},\hat{\eta}\right)$ such that $q\left(\hat{\theta},\hat{\eta}\right) > \eta$.[6] Finally, condition (4) requires that the allocation must be technically feasible.

[5] Again, it is of course equivalent to define a mechanism as the triple of functions $\{q(\theta,\eta), p(\theta,\eta), t(\theta,\eta)\}$. Given $p(\theta,\eta), q(\theta,\eta)$, consumers make their consumption decisions resulting in demand $X(p(\theta,\eta), q(\theta,\eta))$.

[6] It is important to stress that we allow each type to imitate every other type, so there is no limit on what a firm can communicate. This technical difference is important for the proof of the revelation principle in this context. For a formal proof that the revelation principle applies under our verifiability assumptions, see Proposition 6 in Beaudry, Blackorby, and Szalay [2009].

Before I study the solution to this problem, I describe the first-best allocation when types are observable.

2.2 The first-best

Since $\alpha < 1$, the regulator allocates all surplus to the consumer in the first-best allocation; the participation constraint is binding for each type, $\Pi(\theta, \theta, \eta) = 0$, so

$$t(\theta, \eta) = C(x(\theta, \eta), q(\theta, \eta), \theta) - P(x(\theta, \eta), q(\theta, \eta)) x(\theta, \eta).$$

Substituting for $t(\theta, \eta)$ into the regulator's objective function, I obtain

$$\max_{p(\cdot,\cdot), q(\cdot,\cdot)} \mathcal{E}_{\Theta H} (V(x(\theta, \eta), q(\theta, \eta)) - C(x(\theta, \eta), q(\theta, \eta), \theta))$$

Given the boundary condition and that the surplus function is concave in x, the optimal quantity schedule $x^{fb}(\theta, \eta)$ satisfies for each θ, η the first-order condition

$$V_x\left(x^{fb}(\theta, \eta), q^{fb}(\theta, \eta)\right) - C_x\left(x^{fb}(\theta, \eta), q^{fb}(\theta, \eta), \theta\right) = 0.$$

The optimal quality allocation $q^{fb}(\theta, \eta)$ satisfies either $q^{fb}(\theta, \eta) < \eta$ and

$$V_q\left(x^{fb}(\theta, \eta), q^{fb}(\theta, \eta)\right) - C_q\left(x^{fb}(\theta, \eta), q^{fb}(\theta, \eta), \theta\right) = 0,$$

or $q^{fb}(\theta, \eta) = \eta$ and

$$V_q\left(x^{fb}(\theta, \eta), q^{fb}(\theta, \eta)\right) - C_q\left(x^{fb}(\theta, \eta), q^{fb}(\theta, \eta), \theta\right) \geq 0.$$

Since $V_x(x, q) = P(x, q)$, the first of these conditions states that marginal prices should equal marginal cost. The quality allocation is optimal if the marginal benefit of quality is equal to the marginal cost of providing quality, unless of course the quality bound is binding.

I now address the principal's problem when θ and η are not observable to him. It is useful to begin the investigation with the case where the firm can only misrepresent its cost parameter θ, but cannot lie about the parameter η. In this case the regulator faces a family of firms, indexed by their quality capacity η. For each given η, the regulator faces a problem that is now familiar from the analysis of Baron and Myerson [1982].

3 The Case of Observable Quality Bounds

If η is known, the regulator can condition his instruments on η. Although the solution to the regulation problem in this case is now well known from the literature, I reproduce the results for convenience of the reader. A minor novelty of this section, if any, is to show that the model is nicely amenable to monotone comparative statics methods.

I let $x(\theta; \eta)$, $q(\theta; \eta)$, and $t(\theta; \eta)$ for all θ denote the quantity, quality, and transfer schedule conditional on η and define

$$\Pi\left(\hat{\theta}, \theta; \eta\right) \equiv t\left(\hat{\theta}; \eta\right) + P\left(x\left(\hat{\theta}; \eta\right), q\left(\hat{\theta}; \eta\right)\right) x\left(\hat{\theta}; \eta\right) - C\left(x\left(\hat{\theta}; \eta\right), q\left(\hat{\theta}; \eta\right), \theta\right)$$

The principal solves, for each given η, the following problem

$$\mathrm{P}^1 \equiv \max_{p(\cdot;\eta),t(\cdot;\eta),q(\cdot;\eta)} \int_{\underline{\theta}}^{\overline{\theta}} W(\theta;\eta) f(\theta|\eta) d\theta$$

$$\text{s.t., for all } \theta, \eta, \ (2), \ (3), \text{ and } (4) \text{ with } \hat{\eta} \equiv \eta.$$

This is a standard problem and is normally solved by reformulating the incentive and participation constraints. I state a more tractable version of these constraints in the following lemma. I call a triple of quantity schedule $x(\theta;\eta)$, quality schedule $q(\theta;\eta)$, and transfer schedule $t(\theta;\eta)$ implementable if they satisfy constraints (2), (3) with $\hat{\eta} \equiv \eta$. Moreover, I let $\pi(\theta;\eta) \equiv \max_{\hat{\theta}} \Pi\left(\hat{\theta}, \theta; \eta\right)$.

Lemma 1 *The price, quality, and payment schedules, $p(\theta;\eta)$, $q(\theta;\eta)$, and $t(\theta;\eta)$, are implementable if and only if*

$$\begin{aligned} t(\theta;\eta) &= C(x(\theta;\eta), q(\theta;\eta), \theta) \\ &\quad + \int_{\theta}^{\overline{\theta}} c_y(x(y;\eta), y) dy - x(\theta;\eta) P(x(\theta;\eta), q(\theta;\eta)) \end{aligned} \tag{5}$$

and $x(\theta;\eta)$ is non-increasing in θ.

Lemma 1 is standard. Using envelope arguments, the equilibrium profit of the firm is equal to $\pi(\theta;\eta) = \int_{\theta}^{\overline{\theta}} c_y(x(y;\eta), y) dy$, which is just another way of writing condition (5). x has to satisfy a monotonicity condition, which amounts— roughly speaking—to a second order condition, which is typical in incentive problems. Perhaps, it is interesting to note that incentive compatibility places no restrictions on the function $q(\theta;\eta)$.

Define the virtual surplus

$$B(x,q,\theta,\eta) \equiv V(x,q) - C(x,q,\theta) - (1-\alpha) c_\theta(x,\theta) \frac{F(\theta|\eta)}{f(\theta|\eta)}. \tag{6}$$

Substituting the expression for equilibrium profits into the objective function, and integrating by parts, I obtain an equivalent representation of the regulator's problem:

$$\mathrm{P}^2 \equiv \max_{x(\cdot;\eta),q(\cdot;\eta)} \int_{\underline{\theta}}^{\overline{\theta}} B(x(\theta;\eta), q(\theta;\eta), \theta, \eta) f(\theta|\eta) d\theta$$

$$s.t. \ x(\theta;\eta) \text{ non-increasing in } \theta, \text{ and}$$

$$q(\theta;\eta) \le \eta \text{ for all } \theta.$$

Given my maintained assumptions, the monotonicity constraint is automatically satisfied at the solution to this problem. Then, in the absence of this monotonicity constraint, the instruments can effectively be chosen to maximize $B(x(\theta;\eta), q(\theta;\eta), \theta, \eta)$ pointwise. This pointwise problem has a lot of structure. First, the constraint set for each θ is $x(\theta;\eta) \ge 0$ and $q(\theta;\eta) \in [0,\eta]$, a lattice. Second, under natural conditions, the objective function is supermodular. In particular, $B(x,q,\theta,\eta)$ is supermodular (submodular) in (x,q) for each (θ,η) if and only if $S(x,q,\theta)$ is supermodular (submodular) in (x,q) for each θ. Since $B(x,q,\theta,\eta)$ is differentiable to the desired degree, this is equivalent to $B_{xq}(x,q,\theta,\eta) \geq 0$ in the supermodular case and $B_{xq}(x,q,\theta,\eta) \leq 0$ in the submodular case. Moreover, if $B(x,q,\theta,\eta)$ is submodular in (x,q) for each (θ,η), then, with $\hat{q} \equiv q^{-1}$, $B\left(x, \frac{1}{\hat{q}}, \theta, \eta\right)$ is supermodular in $(x,\hat{q})$ for each (θ,η), because $B_{x\hat{q}}\left(x, \frac{1}{\hat{q}}, \theta, \eta\right) < 0$ if and only if $B_{xq}(x,q,\theta,\eta) > 0$. Considering changes in θ, I note that given the maintained assumptions $B(x,q,\theta,\eta)$ has non-increasing differences in (x,q) and θ. Finally, notice that $B(x,q,\theta)$ is strictly concave in (x,q) for all θ.

As usual I will characterize the solution to the problem assuming such a solution exists. Note that given $B(x,q,\theta,\eta)$ is either strictly concave in (x,q) or strictly concave in x alone and increasing in q, the solution is unique and let it be denoted $x^*(\theta;\eta)$ and $q^*(\theta;\eta)$, respectively. I can now characterize this solution:

Proposition 1 *i) The optimal quantity schedule $x^*(\theta;\eta)$ satisfies the first-order condition*

$$V_x(x^*(\theta;\eta), q^*(\theta;\eta)) - C_x(x^*(\theta;\eta), q^*(\theta;\eta), \theta) - (1-\alpha)\, c_{x\theta}(x^*(\theta;\eta), \theta) \frac{F(\theta|\eta)}{f(\theta|\eta)} = 0, \quad (7)$$

and the optimal quality schedule satisfies either

$$q^*(\theta;\eta) < \eta \text{ and } V_q(x^*(\theta;\eta), q^*(\theta;\eta)) - C_q(x^*(\theta;\eta), q^*(\theta;\eta), \theta) = 0, \text{ or} \quad (8)$$

$$q^*(\theta;\eta) = \eta \text{ and } V_q(x^*(\theta;\eta), \eta) - C_q(x^*(\theta;\eta), \eta, \theta) \geq 0; \quad (9)$$

ii) If $S_{xq}(x,q,\theta) \geq 0$ for all (x,q,θ), then $x^*(\theta;\eta)$ and $q^*(\theta;\eta)$ are both non-increasing in θ and the quality capacity constraint is binding for low values of θ if any.
iii) If $S_{xq}(x,q,\theta) \leq 0$ for all (x,q,θ), then $x^*(\theta;\eta)$ is non-increasing and $q^*(\theta;\eta)$ is non-decreasing in θ and the quality capacity constraint is binding for high values of θ if any.

The solution has the remarkable feature that for a *given* quantity schedule, the quality schedule satisfies the same first-order condition as in the first-best. Thus, the only distortions relative to the first-best case arise from the familiar trade-off that causes the quantity schedule to be distorted away from first-best. Using $V_x(x,q) = P(x,q)$, I can rewrite the first-order condition (7) as

$$\begin{aligned} & (P(x^*(\theta;\eta), q^*(\theta;\eta)) - C_x(x^*(\theta;\eta), q^*(\theta;\eta), \theta))\, f(\theta|\eta) \\ = \; & (1-\alpha)\, c_{x\theta}(x^*(\theta;\eta), \theta)\, F(\theta|\eta). \end{aligned}$$

It is useful to review the trade-off behind the optimal pricing in some detail to understand the results that follow. On the left side of the equation I have the principal's desire to implement an efficient solution, which requires that price equals marginal cost. The weight given to this motive is $f(\theta|\eta)$, the likelihood of type $\theta|\eta$. On the right side appears the principal's desire to limit the firm's rents. An increase in $x(\theta;\eta)$ increases the rents that have to be given to all types that are more efficient at producing than type $\theta|\eta$. Since there is a mass $F(\theta|\eta)$ of firms of these types, the weight attached to this motive is $F(\theta|\eta)$. The trade-off is optimally resolved by having all firms except the most efficient one price above marginal costs; the most efficient one prices exactly at marginal costs. Moreover, given the standard assumptions that the type distribution has full support and the surplus function satisfies the boundary condition $\lim_{x\to 0} S_x(x,q,\theta) = \infty$, all types produce a strictly positive quantity at the optimum.

3.1 Comparative Statics

The discussion in the preceding subsection has shown that the regulator's problem amounts to maximizing a supermodular function on a lattice. It is also easy to show that $B(x, q, \theta, \eta)$ has non-decreasing (non-increasing) differences in (x, q) and η for given θ if $\frac{F(\theta|\eta)}{f(\theta|\eta)}$ is non-increasing in η (non-decreasing in η). Thus, the maximand of the regulator's problem depends on η in a monotonic way. I can now state how the solution of the regulation problem depends on η :

Proposition 2 *i) If $\frac{F(\theta|\eta)}{f(\theta|\eta)}$ is non-increasing in η and $S_{xq}(x, q, \theta) \geq 0$ for all (x, q, θ), then $x^*(\theta; \eta)$ and $q^*(\theta; \eta)$ are both non-decreasing in η;*
ii) If $\frac{F(\theta|\eta)}{f(\theta|\eta)}$ is non-decreasing in η and $S_{xq}(x, q, \theta) \leq 0$ for all (x, q, θ), then $x^*(\theta; \eta)$ is non-increasing in η and $q^*(\theta; \eta)$ is non-decreasing in η.

The intuition for these results is as follows. The lower is $\frac{F(\theta|\eta)}{f(\theta|\eta)}$ the smaller is the weight attached to the rent extraction motive relative to the efficiency motive in the regulator's objective function. Hence, the induced quantity that consumers purchase is higher at the optimum. When in addition x and q are complements, then this implies that the quality should also be weakly higher the higher is η. The intuition for the second case is reversed, as far as the quantity of consumption is concerned. But since x and q are substitutes in this case, again the quality is non-decreasing in η.

Proposition 2 covers two out of four possible ways to combine the order of the reverse hazard rates with payoff complementarities. This is because applying monotone comparative statics methods requires that the implied change in quality as a reponse to an increase in η is non-negative. However, for the case where the quality capacity is non-binding for all θ, all four cases can be covered. Since this result is somewhat peripheral to the analysis that follows, the interested reader is referred to Blackorby and Szalay (2008). In contrast Proposition 2 is key to understanding the analysis of the multi-dimensional regulation problem, to which I now turn.

4 The Two-dimensional Problem

I now address the problem when both θ and η are unobservable to the regulator. I start by making some basic observations about the implications of incentive compatibility in my model.

4.1 Incentive Compatibility

One basic obstacle multi-dimensional screening problems face is that the number of deviations to consider is simply too large to deal with. This is a striking difference to the present problem where the number of constraints is simply the "sum" of constraints in each dimension alone. To derive this fundamental result, I first state a technical preliminary needed in the proof of the result.

Lemma 2 *Let $\eta' < \eta''$ and $\theta' < \theta'' < \theta'''$ and suppose $q(\theta, \eta'') \leq \eta'$ for $\theta \in [\theta', \theta''']$. Then, for all $(\theta, \eta) \in \{\eta', \eta''\} \times [\theta', \theta''']$ and all $(\hat{\theta}, \hat{\eta}) \in \{\eta', \eta''\} \times [\theta', \theta''']$:*
i) $\left(x\left(\hat{\theta}, \hat{\eta}\right) - x(\theta, \eta)\right)\left(\hat{\theta} - \theta\right) \leq 0$; and
ii) if $x(\theta, \eta')$ is continuous in θ at θ'', then $x(\theta'', \eta'') = x(\theta'', \eta')$; moreover
iii) $\Pi(\theta, \theta, \eta') = \Pi(\theta, \theta, \eta'')$

Part i) of the Lemma is a straightforward generalization of the monotonicity properties of incentive compatible solutions in one dimension. The difference is that η is allowed to vary as well. As is well known, monotonic functions are continuous almost everywhere. Hence, part ii) applies almost everywhere, stating that bunching of quantity schedules will arise if two types (θ, η') and (θ, η'') both produce qualities that are feasible for both types. Finally, part iii) says that if types (θ, η') and (θ, η'') can mimic each other, then their profits must be the same. Using these properties, I can now show that a mechanism is incentive compatible if and only if no type has any incentive to misrepresent his type in one dimension at a time.

Lemma 3 *The incentive constraint* (2) *is satisfied if and only if the constraints*

$$\Pi(\theta, \theta, \eta) \geq \Pi\left(\hat{\theta}, \theta, \eta\right) \text{ for all } \hat{\theta} \tag{10}$$

and

$$\Pi(\theta, \theta, \eta) \geq \Pi(\theta, \theta, \hat{\eta}) \text{ for all } \hat{\eta} \text{ such that } q(\theta, \hat{\eta}) \leq \eta \tag{11}$$

are satisfied.

The intuition for the result is quite straightforward. Since costs are additively separable in a part related to quality and quantity, and in a part related to quantity and the firm's cost parameter θ, the firm's incentive to misrepresent its cost parameter θ depends solely on the quantity allocation. The intuition why only the one-dimensional constraints are relevant is best seen in the case where a type (θ, η) considers mimicking a type $\left(\hat{\theta}, \hat{\eta}\right)$ with $\hat{\eta} < \eta$. The profit type (θ, η) obtains this way is $\Pi\left(\hat{\theta}, \theta, \hat{\eta}\right)$, just the same profit type $(\theta, \hat{\eta})$ obtains when it mimics type $\left(\hat{\theta}, \hat{\eta}\right)$. But by (10) applied to type $(\theta, \hat{\eta})$, I know that it would be better to state the true cost parameter θ, so $\Pi(\theta, \theta, \hat{\eta}) \geq \Pi\left(\hat{\theta}, \theta, \hat{\eta}\right)$. But then, by constraint (11), being truthful about the quality capacity would result in an even higher profit. Using Lemma 4 above, I can generalize this insight also to the case where the type mimicked, $\left(\hat{\theta}, \hat{\eta}\right)$, is such that $\hat{\eta} > \eta$. If the deviation is feasible in the first place, then the quantity allocation and the profit must be exactly the same as if type $\left(\hat{\theta}, \eta\right)$ was mimicked. But then, by (10), this deviation gives rise to a profit that is weakly lower than the profit obtained by being truthful.

The crucial property of the problem that allows me to reduce the dimensionality of the problem in this way is that the profit $\Pi\left(\hat{\theta}, \theta, \hat{\eta}\right)$ depends only on messages and the true cost parameter, but not on the true quality capacity parameter.

Building on Lemmas four and five, I can bring the incentive constraints into a more tractable form. Let $\Lambda\left(\hat{\theta}, \hat{\eta}, \eta\right) \equiv \left\{\hat{\theta}, \hat{\eta} : q\left(\hat{\theta}, \hat{\eta}\right) \leq \eta\right\}$ and let $\pi(\theta, \eta) \equiv \max_{\hat{\theta}, \hat{\eta} \in \Lambda\left(\hat{\theta}, \hat{\eta}, \eta\right)} \Pi\left(\hat{\theta}, \theta, \hat{\eta}\right)$.

Lemma 4 *The incentive and participation constraints are satisfied if and only if i)*

$$t(\theta, \eta) = C(x(\theta, \eta), q(\theta, \eta), \theta) + \pi(\theta, \eta) - x(\theta, \eta) P(x(\theta, \eta), q(\theta, \eta)), \tag{12}$$

where

$$\pi(\theta, \eta) = \pi\left(\overline{\theta}, \eta\right) + \int_{\theta}^{\overline{\theta}} c_y(x(y, \eta), y)\, dy; \tag{13}$$

ii) $x(\theta, \eta)$ is non-increasing in θ for all η, and $\pi\left(\overline{\theta}, \eta\right) \geq 0$ for all η;
iii) $\pi(\theta, \eta)$ is non-decreasing in η for all θ, and
iv) if for $\eta' < \eta''$ $q(\theta, \eta'') \leq \eta'$ then $\pi(\theta, \eta') = \pi(\theta, \eta'')$ and $x(\theta, \eta') = x(\theta, \eta'')$.

I can compute the rent of a firm of type (θ, η) by summing the rent of the most inefficient cost type for a given quality capacity, $\pi\left(\overline{\theta}, \eta\right)$, and the marginal changes of the firm's rent with respect to changes in its cost parameter θ. Notice that (13) allows for the case where $\pi\left(\overline{\theta}, \eta\right) > 0$, so some high cost types may receive rents. Apart from this, I can essentially use the same procedure as in Lemma 1. ii) is the usual monotonicity requirement to guarentee incentive compatibility in the θ dimension. iii) is an additional monotonocity requirement that takes care of incentive compatibility in the η-dimension; the firm's rent must be non-decreasing in η because mimicking a firm with a lower η is - by technical feasibility of the quality allocation - always possible. Condition iv) summarizes the essence of Lemma 2; if two types with the same cost parameter θ but different quality capacities can mimic each other, then they must receive the same rent and they must produce the same quantities. This final condition in the Lemma is somewhat difficult to treat analytically. The problem is that the regulator's ability to implement distinct quantity schedules for firms of the same cost but with different quality capacities depends on the allocation of quality for these firms. Thus, to solve the problem, I need to have a good guess about the quality allocation. Before I turn to this question, it is worth discussing some important differences between this problem and other problems of multi-dimensional screening.

A crucial obstacle the multi-dimensional screening problem faces is that there is no natural order of types (see Rochet and Choné (2003). In their problem it would not make sense to compute the rent of a firm simply by integrating up from the most inefficient type for each given η. A deviation of a type (θ, η) to some statement $\left(\hat{\theta}, \hat{\eta}\right)$ must be ruled out irrespective of the particular path connecting (θ, η) and $\left(\hat{\theta}, \hat{\eta}\right)$. In the present problem, only orthogonal incentive constraints play a role, and only those in the θ-dimension are necessarily binding at the optimum. Using these binding incentive constraints, I obtain conditions (12) and (13). Incentive compatibility in the η dimension is then a relatively simple monotonicity condition in the η dimension.

4.2 The optimal allocation of quality

Recall from the analysis of the first-best allocation that the first-best optimal quality allocation, $q^{fb}(\theta, \eta)$, satisfies either a first-order condition stating that the marginal benefit of quality should equal the marginal cost, or the capacity constraint is binding. I can provide a lower bound for the second-best allocation of quality:

Proposition 3 *i) An optimal second-best allocation for any given (not necessarily optimal) quantity schedule $x(\theta, \eta)$ entails either $q^*(\theta, \eta) = \eta$ or*

$$q^*(\theta, \eta) < \eta \text{ and } V_q(x(\theta, \eta), q^*(\theta, \eta)) - C_q(x(\theta, \eta), q^*(\theta, \eta), \theta) \leq 0.$$

ii) If $S(x, q, \theta)$ is strictly increasing in q for all $x \leq \bar{x}$ and the optimal quantity schedule given a quality allocation $q(\theta, \eta) = \eta$ satisfies $x^*(\theta, \eta) \leq \bar{x}$ for all θ, η, then the optimal quality allocation satisfies $q^*(\theta, \eta) = \eta$ for all θ, η.

The intuition for this result is as follows. Given additive separability of the cost function, the regulator can just compensate the firm for its additional cost if it is asked to provide higher quality. The regulator does not have to pay more than the increase in pure economic costs, because there is no interaction between q and θ in the firm's cost function; hence the firm's informational rent with respect to its information θ is not affected by changes in the quality allocation. In addition to that, one needs to be sure that changing the quality allocation does not open new deviation possibilities for the firm. Clearly, that is true if I raise the quality allocation pointwise, starting from a given quantity allocation and an allocation of quality where the marginal benefit of quality is still higher than the marginal cost. The reason is that such a change makes deviations for the firm more difficult, so if the initial allocation is incentive compatible, the new one is as well. Additive separability of the cost function is obviously crucial for this result. If the cost function were not additively separable in θ and q, then changing the quality allocation would change the agent's incentive to mimic other types. The second part of Proposition 3 is a direct consequence of the first part.

I am now ready to characterize the complete solution to the regulation problem in a number of cases. Interesting cases that can be studied are the extreme cases where the quality capacity is either binding for all types or nonbinding for all types. I now study these cases in detail.

5 The case of binding capacity constraints

In this section I assume that social surplus is increasing in q for all $x \leq \bar{x}$ and $\bar{x}$ is sufficiently large. Proposition 3 suggests a simple solution procedure for this case: I can conjecture that the quality capacity is binding for all types and solve for the optimal quantity schedules conditional on that quality allocation rule. Clearly, for $\bar{x}$ sufficiently high, the optimal quantity of production $x^* (\theta, \eta)$ is indeed below $\bar{x}$ for all θ, η, and hence this procedure picks up the optimum.

A binding capacity constraint simplifies the problem dramatically. By definition, the feasibility constraint is always met. Moreover, types will only be tempted to mimic downwards in the η dimension, but not upwards. Hence, $\pi (\theta, \eta)$ must be non-decreasing in η to rule out downward deviations. With $q (\theta, \eta) = \eta$ for all θ, η, problem (1) subject to (2), (3), and (4), is equivalent to the the following problem

$$\mathrm{P}^3 \equiv \max_{x(\cdot;\cdot),\pi\left(\overline{\theta},\cdot\right)} \mathcal{E}_{\mathcal{H}} \int_{\underline{\theta}}^{\overline{\theta}} \left(B\left(x\left(\theta;\eta\right), q\left(\theta;\eta\right), \theta, \eta\right) - (1-\alpha)\,\pi\left(\overline{\theta},\eta\right)\right) f\left(\theta\left|\eta\right.\right) d\theta$$

$$\begin{aligned} s.t.\ & x\left(\theta,\eta\right) \text{ non-increasing in } \theta \text{ for all } \eta \text{ and} \\ & \pi\left(\theta,\eta\right) \text{ non-decreasing in } \eta \text{ for all } \theta, \end{aligned}$$

where $\pi (\theta, \eta)$ is determined by $x (\theta, \eta)$ and $\pi \left(\overline{\theta}, \eta\right)$ through (13). The solution to the problem depends crucially on two things: first whether quality and quantity are net complements or net substitutes in the social surplus function, and second on the order structure of reversed hazard rate, that is whether $\frac{f(\theta|\eta)}{F(\theta|\eta)}$ is increasing or decreasing in η. One can show that $\frac{f(\theta|\eta)}{F(\theta|\eta)}$ is increasing (decreasing) in η if and only if the right truncated distribution of $\theta\left|\eta'\right.$ dominates the right truncated distribution of $\theta\left|\eta'\right.$ for any $\eta' > \eta'$ and any common point of truncation in the sense of First Order Stochastic Dominance (see Shaked and Shantikumar (2007)). Since this obviously implies a positive (negative) correlation between θ and η, I will losely speak of positive or negative correlation in what follows although the conditions I use are that either $\frac{f(\theta|\eta)}{F(\theta|\eta)}$ is increasing or decreasing in η.

5.1 Net complements and positive correlation

I now show that if quality and quantity are net complements and the reversed hazard rate is increasing in η the agent has no incentive to mimic another type who produces a lower quality level. Recall that $x^*(\theta;\eta)$, defined by condition (7), is the optimal quantity schedule for the case where η is known. I have the following result:

Proposition 4 *If* $\frac{f(\theta|\eta)}{F(\theta|\eta)}$ *is nondecreasing in* η *for all* θ *and* $S_{xq}(x,q,\theta) \geq 0$ *for all* (x,q,θ), *then unobservability of* η *does not affect the solution to the principal's problem. Formally, the solution is given by the quality schedule* $q(\theta,\eta) = \eta$ *for all* (θ,η), *by the quantity schedule* (7), *and the transfer schedule* (5).

The formal proof of this statement is omitted, since the argument is obvious. Indeed, suppose one solves problem P^3 under the assumptions in Proposition 4, neglecting both the monotonicity constraint on $x(\theta,\eta)$ and the one on $\pi(\theta,\eta)$ and setting $\pi(\overline{\theta},\eta) = 0$ for all η. Clearly, the quantity schedule that solves this "reduced" problem is $x^*(\theta,\eta) = x^*(\theta;\eta)$. Under the maintained assumptions, $x^*(\theta;\eta)$ is non-increasing in θ. Moreover, by Proposition 2, $x^*(\theta;\eta)$ is non-decreasing in η for each θ so that $\pi(\theta,\eta)$ is automatically non-decreasing in η. Hence, the procedure picks up the solution to the full problem including its constraints.

The intuition is quite simple: there is no conflict of interest with respect to η. It is optimal from the regulator's perspective to let higher η types produce more, because x and q are complements in the social surplus function, and because the higher is η the higher is $\frac{f(\theta|\eta)}{F(\theta|\eta)}$, so the greater is the weight given to the principal's efficiency motive as opposed to the motive to limit the firm's rents. Hence, higher η types receive larger rents so they have no incentive to report a lower value of η.

5.2 Net substitutes and negative correlation

The case where $\frac{f(\theta|\eta)}{F(\theta|\eta)}$ is nonincreasing in η for all θ and $S_{xq}(x,q,\theta) \leq 0$ for all (x,q,θ) is considerably harder to tackle analytically. To avoid a heavy use of variational methods, I assume for the analysis of this section, that $\mathbf{H}$ is discrete. For reasons of space I restrict attention to the special case where there are only two $\eta-$types, i.e. $\eta \in \{\underline{\eta},\overline{\eta}\}$. Even this arguably simplest case becomes quite involved. A generalization to the case of an arbitrary discrete set $\mathbf{H}$ is doable but not presented here to keep the analysis reasonably short. The generalization to the case where $\mathbf{H}$ is a continuum requires a different approach that I pursue in companion work.

5.2.1 Statement of the problem

Let $\beta \equiv \Pr(\eta = \overline{\eta})$. When there are only two η-types, I can ease notation letting $\overline{x}(\theta) \equiv x(\theta,\overline{\eta})$ and $\underline{x}(\theta) \equiv x(\theta,\underline{\eta})$; likewise I use notation $\overline{q}(\theta)$ and $\underline{q}(\theta)$ for the allocation of quality and $\overline{t}(\theta)$ and $\underline{t}(\theta)$ for transfers. $\pi(\theta,\eta)$ is non-decreasing in η for all θ if and only if

$$\rho(\theta,\pi) \equiv \pi + \int_{\theta}^{\overline{\theta}} c_y(\overline{x}(y),y)\,dy - \int_{\theta}^{\overline{\theta}} c_y(\underline{x}(y),y)\,dy \geq 0, \tag{14}$$

where $\pi \equiv \pi(\overline{\theta},\overline{\eta})$ and $\pi(\overline{\theta},\underline{\eta}) = 0$ has been used. (14), requires for all θ that $\rho(\theta,\pi)$, the excess rent of type $(\theta,\overline{\eta})$ over type $(\theta,\underline{\eta})$, be non-negative. $\pi(\overline{\theta},\underline{\eta}) = 0$ is optimal because the regulator´s objective function is decreasing in the rents left to firms and reducing $\pi(\overline{\theta},\underline{\eta})$ relaxes constraint (14) at the same time. On the other hand, it is not a priori clear that $\pi = 0$ is also optimal. When setting π, the regulator faces a trade-off between separating firms with different quality capacities and extracting rents from high quality capacity producers. This trade-off is analyzed in detail below.

Recall the definition of the virtual surplus given in (6). For binding capacity constraints, problem P^3 specializes further to

$$\text{P}^4 \equiv \max_{\overline{x}(\cdot),\underline{x}(\cdot),\pi} \beta \int_{\underline{\theta}}^{\overline{\theta}} B\left(\overline{x}(\theta),\overline{\eta},\theta,\overline{\eta}\right) f\left(\theta\left|\overline{\eta}\right.\right) d\theta - \beta(1-\alpha)\pi + (1-\beta)\int_{\underline{\theta}}^{\overline{\theta}} B\left(\underline{x}(\theta),\underline{\eta},\theta,\underline{\eta}\right) f\left(\theta\left|\underline{\eta}\right.\right) d\theta$$

$$s.t.\ \overline{x}(\theta),\underline{x}(\theta) \text{ non-increasing in } \theta, \text{ and } (14).$$

Problem P^4 has the following structure. If the monotonicity constraints on $\overline{x}(\theta)$ and $\underline{x}(\theta)$ are nonbinding, the problem can be viewed as a control problem with two control variables, $c_\theta(\overline{x}(\theta),\theta)$ and $c_\theta(\underline{x}(\theta),\theta)$, and two state variables, $-\int_\theta^{\overline{\theta}} c_y(\overline{x}(y),y)\,dy$ and $-\int_\theta^{\overline{\theta}} c_y(\underline{x}(y),y)\,dy$. Moreover, the state variables enter the problem through an inequality constraint. This is a relatively complex problem, but solution techniques are available in the literature (see, e.g., Kamien and Schwartz (1981) or Seyerstad and Sydsaeter (1999)). If the monotonicity constraints are binding for some θ, the problem involves second derivatives. This case becomes extremely difficult to analyze. Therefore my approach is to impose assumptions that guarantee that the monotonicity constraints are slack at the solution to problem P^4.

5.2.2 The necessity of bunching

I begin with a few general observations about the problem and its solution. First, I observe that constraint (14) is binding for some θ at the optimum. To see this, suppose (14) were non-binding for all θ. Then the solution to the problem would involve the quantity schedules $x^*(\theta;\overline{\eta})$ and $x^*\left(\theta;\underline{\eta}\right)$ defined by (7) in Proposition 1 for the case where $q^*(\theta;\eta)=\eta$ for $\eta\in\left\{\underline{\eta},\overline{\eta}\right\}$. But this requires that the regulator leaves at least a rent $\overline{\pi}$ to all firms producing high quality, where

$$\overline{\pi} \equiv \int_{\underline{\theta}}^{\overline{\theta}} c_y\left(x^*(y;\overline{\eta}),y\right)dy - \int_{\underline{\theta}}^{\overline{\theta}} c_y\left(x^*\left(y;\underline{\eta}\right),y\right)dy > 0.$$

$\overline{\pi}$ is strictly positive because by Proposition 2 $x^*(\theta;\overline{\eta}) < x^*\left(\theta;\underline{\eta}\right)$ for all θ if either quality and quantity are strict substitutes and/or $\frac{f(\theta|\eta)}{F(\theta|\eta)}$ is decreasing in η for all θ. However, setting $\pi \geq \overline{\pi}$ cannot be optimal. Around $\pi=\overline{\pi}$, the marginal cost of increasing π is equal to $-\beta(1-\alpha)$; a fraction of firms β has quality capacity $\overline{\eta}$ and rents left to firms enter the regulators payoff function with a weight of $-(1-\alpha)$. On the other hand, the benefit of increasing π around $\pi=\overline{\pi}$ is zero, as the regulator is already unconstrained by condition (14) for $\pi=\overline{\pi}$. Hence, at the optimum I must have $0\leq\pi^*<\overline{\pi}$.[7]

Second, I observe that if constraint (14) is binding over an interval, then the schedules $\overline{x}(\theta)$ and $\underline{x}(\theta)$ involve bunching, that is $\overline{x}(\theta)=\underline{x}(\theta)$. This can be seen easily by differentiating (14) with respect to θ over such an interval.

Given that the solution to the problem necessarily involves bunching, it is crucial to determine the precise location of the bunching regions. Moreover, one has to distinguish bunching according to the dimension in which it occurs. I refer to $\eta-$bunching when the schedules $\overline{x}(\theta)$ and $\underline{x}(\theta)$ are bunched together. I speak of $\theta-$bunching when at least one of the schedules $\overline{x}(\theta)$ or $\underline{x}(\theta)$ has flat parts.

[7] This heuristic argument is made more formally in the proof of Proposition 5 below.

5.2.3 No (traditional) $\theta-$bunching

I have argued above that bunching in the θ-dimension adds complexity but not insights. Therefore, I introduce assumptions that rule such bunching out. More specifically, I impose:

Assumption i: a) $c(x,\theta) = \theta x$, b) $\frac{\partial}{\partial\theta}\frac{1-F(\theta|\overline{\eta})}{f(\theta|\overline{\eta})} \leq 0$ and c) $\frac{\partial}{\partial\theta}\frac{\frac{\beta}{1-\beta}+F(\theta|\underline{\eta})}{f(\theta|\underline{\eta})} \geq 0$.

Assumption i places structure on the virtual surplus functions, once π is set to its optimal level. These modified virtual surplus functions take the form

$$\overline{B}(x,\overline{\eta},\theta) \equiv V(x,\overline{\eta}) - C(x,\overline{\eta},\theta) + (1-\alpha)\, c_\theta(x,\theta)\,\frac{1-F(\theta\,|\overline{\eta})}{f(\theta\,|\overline{\eta})} \tag{15}$$

and

$$\underline{B}\left(x,\underline{\eta},\theta\right) \equiv V\left(x,\underline{\eta}\right) - C\left(x,\underline{\eta},\theta\right) - (1-\alpha)\, c_\theta(x,\theta)\,\frac{\frac{\beta}{1-\beta}+F\left(\theta\,|\underline{\eta}\right)}{f\left(\theta\,|\underline{\eta}\right)}. \tag{16}$$

Part a) of the assumption ensures that both expressions are strictly concave in x; parts b) and c) guarantee that both functions have a negative cross derivative with respect to x and θ. It is worth noting that property b) is imposed in much of the screening literature; many well known distributions possess this property; c) is somewhat more restrictive than the maintained (standard) assumption of a non-increasing reversed hazard rate. As usual, the precise role of these conditions becomes apparent once the solution to the problem is presented. Clearly, assumption i rules out certain cases; however, it allows me to sidestep technical difficulties that do not add economic insights.

5.2.4 The location of $\eta-$bunching regions

I now turn to locating the regions where the schedules $\overline{x}(\theta)$ and $\underline{x}(\theta)$ are bunched together.

Lemma 5 *Consider the "reduced" problem P^4 absent monotonicity constraints on $\overline{x}(\theta)$ and $\underline{x}(\theta)$. At the solution to this problem, constraint* (14) *is binding at $\theta = \underline{\theta}$.*

The proof is a simple argument by contradiction. Let $\overline{x}^*(\theta)$ and $\underline{x}^*(\theta)$ denote the optimal quantity schedules solving the reduced problem P^4. If constraint (14) were slack at $\theta = \underline{\theta}$, I could use the transversality conditions of the problem to conclude that $\overline{x}^*(\theta) = x^*(\theta;\overline{\eta})$ and $\underline{x}^*(\theta) = x^*\left(\theta;\underline{\eta}\right)$ for all $\theta \leq \theta'$, where θ' is the smallest θ where (14) is binding. By definition of the point θ', the excess rent of type $\left(\theta',\overline{\eta}\right)$ is zero, that is $\rho\left(\theta',\pi\right) = 0$. By Proposition 2 $x^*(\theta;\overline{\eta}) < x^*\left(\theta;\underline{\eta}\right)$, and hence

$$\int_{\underline{\theta}}^{\theta'} \overline{x}^*(y)\,dy - \int_{\underline{\theta}}^{\theta'} \underline{x}^*(y)\,dy < 0.$$

But this shows that

$$\rho\left(\underline{\theta},\pi\right) = \rho\left(\theta',\pi\right) + \int_{\underline{\theta}}^{\theta'} \overline{x}^*(y)\,dy - \int_{\underline{\theta}}^{\theta'} \underline{x}^*(y)\,dy < 0,$$

for any $\theta' > \underline{\theta}$. Hence, I must have $\theta' = \underline{\theta}$, that is, constraint (14) is binding at $\underline{\theta}$. Finally, suppose that (14) is binding over an interval. Then, differentiation (14) with respect to θ over such an interval, it follows directly that $\overline{x}(\theta) = \underline{x}(\theta)$. The complete statement of the problem and the details of the argument can be found in the formal proof in the appendix.

While Lemma 5 establishes that constraint (14) must be binding at $\theta = \underline{\theta}$, the Lemma does of course not offer a complete characterization of the sets of points where (14) is binding or not, respectively. Moreover, the second part of the lemma shows that there is a close connection between points where (14) is binding and points of bunching between the schedules $\overline{x}(\theta)$ and $\underline{x}(\theta)$. I am able to completely characterize the solution for two cases; first, when (14) binds only at $\theta = \underline{\theta}$ and second when (14) binds over a single interval. More precisely, the first case arises when imposing (14) as an equality for $\theta = \underline{\theta}$ is sufficient for (14). I term this case the "isoperimetric" case, because in this case (14) becomes an integral constraint and the regulator's problem becomes an isoperimetric problem. In the second case, that I term "interval-bunching", (14) binds over an interval at the low end of the support. I now analyze these cases in turn.

5.2.5 The isoperimetric case

By Lemma 5 (14) binds at $\theta = \underline{\theta}$. In fact, it is possible that (14) binds only at $\theta = \underline{\theta}$.

Lemma 6 *For schedules* $\overline{x}(\theta)$ *and* $\underline{x}(\theta)$ *that satisfy*

$$\underline{x}(\theta) = \overline{x}(\theta) \Longrightarrow \frac{d\overline{x}(\theta)}{d\theta} \geq \frac{d\underline{x}(\theta)}{d\theta} \tag{17}$$

condition (14) *is satisfied if*

$$\pi + \int_{\underline{\theta}}^{\overline{\theta}} \overline{x}(y)\, dy - \int_{\underline{\theta}}^{\overline{\theta}} \underline{x}(y)\, dy = 0. \tag{18}$$

Proof. For schedules $\overline{x}(\theta)$ and $\underline{x}(\theta)$ that satisfy (17), $\overline{x}(\theta) - \underline{x}(\theta)$ is non-positive at $\theta = \underline{\theta}$ and changes sign at most once. To see this, note first that the solution schedules cannot satisfy $\overline{x}(\theta) \geq \underline{x}(\theta)$ for all θ, because that would imply that (14) is slack. Condition (17) implies that $\underline{x}(\theta)$ crosses $\overline{x}(\theta)$ from above everytime the schedules cross. Hence, $\overline{x}(\theta) - \underline{x}(\theta)$ changes sign at most once. Moreover, I must have $\overline{x}(\underline{\theta}) \leq \underline{x}(\underline{\theta})$. If I had $\overline{x}(\underline{\theta}) > \underline{x}(\underline{\theta})$, $\underline{x}(\theta)$ would "start" below $\overline{x}(\theta)$ and hence could not cross $\overline{x}(\theta)$ from above. Differentiate now (14) with respect to θ to get $\rho_\theta(\theta, \pi) = \underline{x}(\theta) - \overline{x}(\theta)$. It follows that $\rho(\theta, \pi)$, the excess rent of type $(\theta, \overline{\eta})$ over type $(\theta, \underline{\eta})$ is either non-decreasing in θ for all θ (if $\overline{x}(\theta) \leq \underline{x}(\theta)$ for all θ) or first non-decreasing and then non-increasing in θ. Since $\rho(\underline{\theta}, \pi) = 0$ and $\rho(\overline{\theta}, \pi) = \pi \geq 0$, constraint (18) is sufficient for constraint (14) in both cases. ■

Condition (17) is a single crossing condition on the *endogenous* schedules $\overline{x}(\theta)$ and $\underline{x}(\theta)$. To generate the single-crossing property in the sense of the lemma from first principles, I impose:

Assumption ii: $S_{xxq}(x, q, \theta) \leq 0$ for all (x, q, θ) and

$$\max\left\{\frac{\partial}{\partial\theta}\frac{F(\theta\,|\overline{\eta})}{f(\theta\,|\overline{\eta})}, -\frac{\partial}{\partial\theta}\frac{1-F(\theta\,|\overline{\eta})}{f(\theta\,|\overline{\eta})}\right\} \leq \min\left\{\frac{\partial}{\partial\theta}\frac{F(\theta\,|\underline{\eta})}{f(\theta\,|\underline{\eta})}, \frac{\partial}{\partial\theta}\frac{\frac{\beta}{1-\beta} + F(\theta\,|\underline{\eta})}{f(\theta\,|\underline{\eta})}\right\}.$$ [8]

[8] Notice that assumptions i and ii are both satisfied if the distribution of θ is independent of η and satisfies $f_\theta(\theta) \geq 0$.

Given the lemma and assumptiona i and ii, I can reduce my problem further to the following one:

$$\mathrm{P}^5 \equiv \max_{\overline{x}(\cdot),\underline{x}(\cdot),\pi} \beta \int_{\underline{\theta}}^{\overline{\theta}} B\left(\overline{x}\left(\theta\right),\overline{\eta},\theta,\overline{\eta}\right) f\left(\theta\left|\overline{\eta}\right.\right) d\theta - \beta\left(1-\alpha\right)\pi + \left(1-\beta\right) \int_{\underline{\theta}}^{\overline{\theta}} B\left(\underline{x}\left(\theta\right),\underline{\eta},\theta,\underline{\eta}\right) f\left(\theta\left|\underline{\eta}\right.\right) d\theta$$

$$s.t.\ (18)\,.$$

Let k denote the Lagrange multiplier attached to the constraint (18). I can now characterize the solution to my problem.

Proposition 5 *If $S\left(x,q,\theta\right)$ is increasing in q on the relevant domain, and under Assumptions i and ii, the solution to problem P^4 is given by the quality schedules $\overline{q}^*\left(\theta\right)=\overline{\eta}$ and $\underline{q}^*\left(\theta\right)=\underline{\eta}$, the quantity schedules $\overline{x}^*\left(\theta\right)$ and $\underline{x}^*\left(\theta\right)$ defined by the conditions*

$$\left(V_x\left(\overline{x}^*\left(\theta\right),\overline{\eta}\right) - C_x\left(\overline{x}^*\left(\theta\right),\overline{\eta},\theta\right) - \left(1-\alpha\right)\frac{F\left(\theta\left|\overline{\eta}\right.\right)}{f\left(\theta\left|\overline{\eta}\right.\right)}\right)\beta f\left(\theta\left|\overline{\eta}\right.\right) + k^* = 0 \tag{19}$$

and

$$\left(V_x\left(\underline{x}^*\left(\theta\right),\underline{\eta}\right) - C_x\left(\underline{x}^*\left(\theta\right),\underline{\eta},\theta\right) - \left(1-\alpha\right)\frac{F\left(\theta\left|\underline{\eta}\right.\right)}{f\left(\theta\left|\underline{\eta}\right.\right)}\right)\left(1-\beta\right) f\left(\theta\left|\underline{\eta}\right.\right) - k^* = 0, \tag{20}$$

and the associated transfer schedules $\overline{t}^*\left(\theta\right)$ and $\underline{t}^*\left(\theta\right)$ defined by (12). Moreover, $\pi^* < \overline{\pi}$ and either

$$k^* \leq \beta\left(1-\alpha\right) \ \textit{and}\ \pi^* = 0,\ \textit{or} \tag{21}$$

$$k^* = \beta\left(1-\alpha\right) \ \textit{and}\ \pi^* > 0. \tag{22}$$

The solution is a remarkably simple pair of first-order conditions. Up to the optimal choice of π^* Proposition 5 provides a complete characterization of the optimum. The solution schedules (19) and (20) differ markedly from their counterparts for the case where η is known. In particular, $\overline{x}^*\left(\theta\right)$ is distorted upwards relative to $x^*\left(\theta;\overline{\eta}\right)$, the schedule for when η is known to be high, and $\underline{x}^*\left(\theta\right)$ is distorted downwards relative to $x^*\left(\theta;\underline{\eta}\right)$. The reason is of course that the schedules $x^*\left(\theta;\overline{\eta}\right)$ and $x^*\left(\theta;\underline{\eta}\right)$ violate constraint (14). A particular implication of constraint (14) being binding is that (19) and (20) violate the "no distortion at the top" condition within types of the same quality capacity. That is, $\overline{x}^*\left(\underline{\theta}\right)$ is higher than the efficient quantity for type $\left(\underline{\theta},\overline{\eta}\right)$ and $\underline{x}^*\left(\underline{\theta}\right)$ is lower than the efficient quantity for type $\left(\underline{\theta},\underline{\eta}\right)$.

As shown in the proof of Proposition 5, the problem is concave in the choice variables. Hence, $\pi^* > 0$ if and only if $k\left(0\right) \equiv \left.k\left(\pi\right)\right|_{\pi=0} > \beta\left(1-\alpha\right)$. It is straightforward to compute $k\left(0\right)$ from conditions (19), (20), and (18). I now show by means of an example that $k\left(0\right)$ can indeed be larger than $\beta\left(1-\alpha\right)$ and thus that the optimal π can be positive.

Example 1 *Let $V\left(x,q\right) = l\left(q\right) + \frac{x^{1-a}}{1-a}$ for some increasing function $l\left(q\right)$, $c\left(x,\theta\right) = \theta x$, and $k\left(x,q\right) = K + xq$, where K is a fixed cost. Moreover, let θ be uniformly distributed both for $\eta=\overline{\eta}$ and $\eta=\underline{\eta}$. Then,* (19) *and* (20) *can be solved explicitly. I obtain*

$$\overline{x}^*\left(\theta\right) = \left(\theta + \left(1-\alpha\right)\frac{\left(\theta-\underline{\theta}\right)}{\overline{\theta}-\underline{\theta}} + \overline{\eta} - \frac{k}{\beta}\left(\overline{\theta}-\underline{\theta}\right)\right)^{-\frac{1}{a}}$$

and

$$\underline{x}^*(\theta) = \left(\theta + (1-\alpha)\frac{(\theta - \underline{\theta})}{\overline{\theta} - \underline{\theta}} + \underline{\eta} + \frac{k}{1-\beta}\left(\overline{\theta} - \underline{\theta}\right)\right)^{-\frac{1}{a}}$$

From (18) *with* $\pi = 0$, *I conclude that* $k(0)$ *satisfies the condition* $\overline{\eta} - \frac{k(0)}{\beta}\left(\overline{\theta} - \underline{\theta}\right) = \underline{\eta} + \frac{k(0)}{1-\beta}\left(\overline{\theta} - \underline{\theta}\right)$, *for otherwise I would have either* $\overline{x}^*(\theta) > \underline{x}^*(\theta)$ *for all* θ *or* $\overline{x}^*(\theta) < \underline{x}^*(\theta)$ *for all* θ, *and both possibilities are inconsistent with condition* (18). *Hence, I have* $k(0) = \frac{(1-\beta)\beta}{\overline{\theta}-\underline{\theta}}\left(\overline{\eta} - \underline{\eta}\right)$, *and therefore* $\pi^* > 0$ *for* $\frac{1-\beta}{\overline{\theta}-\underline{\theta}}\left(\overline{\eta} - \underline{\eta}\right) > (1-\alpha)$.

The simple example nicely illustrates the trade-off the regulator faces when setting π. For $\pi = 0$, the same production schedule has to be implemented irrespective of the quality type η. This is the worse the larger is the difference between the levels of quality offered by high quality and low quality producers, respectively. By leaving an additional positive rent to all high quality producers, that is by setting $\pi > 0$, the regulator obtains the possibility to tailor the quantity schedules to the qualities offered, that is to separate different quality producers.

For the case where the regulator leaves a positive rent to all high quality producers, the solution has quite unconventional features. Substituting $k^* = \beta(1-\alpha)$ (from (22)) into (19) and (20), I obtain the optimal quantity schedules for the case where $\pi^* > 0$:

$$V_x\left(\overline{x}^*(\theta), \overline{\eta}\right) = C_x\left(\overline{x}^*(\theta), \overline{\eta}, \theta\right) - (1-\alpha)\frac{1 - F(\theta\,|\overline{\eta})}{f(\theta\,|\overline{\eta})} \tag{23}$$

and

$$V_x\left(\underline{x}^*(\theta), \underline{\eta}\right) = C_x\left(\underline{x}^*(\theta), \underline{\eta}, \theta\right) + (1-\alpha)\frac{\frac{\beta}{1-\beta} + F\left(\theta\,|\underline{\eta}\right)}{f\left(\theta\,|\underline{\eta}\right)}. \tag{24}$$

Since $V_x(x, q) = P(x, q)$, these conditions relate marginal prices to marginal costs and statistics of the type distribution. The optimal marginal prices of the high quality producer (the high η-type) are below marginal costs for all but the the least efficient type $\left(\overline{\theta}, \overline{\eta}\right)$. This is to implement a quantity allocation that is distorted upwards relative to the first-best quantity for all but the least efficient type $\left(\overline{\theta}, \overline{\eta}\right)$. The least efficient type produces the efficient quantity. For producers of low quality, the pricing scheme and the implemented allocation has almost but not quite the traditional features: the low quality producing types all price above marginal costs, even the most efficient of these producers. Obviously, these distortions arise due to the constraint that prevents the high quality producer from shirking on quality, that is from mimicking the low quality producer with the same cost parameter.

I now turn to analyzing a class of cases that violate assumption ii, but still have enough structure to allow for a complete characterization of the optimum.

5.2.6 The case of interval-bunching

If the reverse of condition (17) holds, then bunching in the η dimension occurs only for θ at the low end of the support. I state this result more formally in the following Lemma:

Lemma 7 *If the schedules* $\overline{x}(\theta)$ *and* $\underline{x}(\theta)$ *defined by* (19), *and* (20) *satisfy for any* $k \geq 0$

$$\overline{x}(\theta) = \underline{x}(\theta) \Longrightarrow \frac{d\underline{x}(\theta)}{d\theta} > \frac{d\overline{x}(\theta)}{d\theta}, \tag{25}$$

then at the solution to problem P^4, *constraint* (14) *is binding on a set* $\left[\underline{\theta}, \theta'\right]$ *for some* $\theta' \geq \underline{\theta}$.

The proof of the Lemma is straightforward. On any subinterval of the type space between any two points where constraint (14) binds, the incremental changes of the extra rent type a high η type receives relative to his low η counterpart, must sum to zero. But then, I face on this subinterval a problem that is essentially identical to problem P^4. Hence, if the solution schedules (19) and (20) satisfy a single crossing condition, then on any subinterval where constraint (14) is slack, the schedules $\overline{x}(\theta)$ and $\underline{x}(\theta)$ cross at most once. They do have to cross at least once so as to guarantee the incremental rents of high versus low η types add up to zero. Suppose now there were two points θ_1 and θ_2 such that (14) binds for all $\theta \leq \theta_1$ and for all $\theta \geq \theta_2$, but not for any θ in between θ_1 and θ_2. As I have just argued, this would imply that $\overline{x}(\theta_1)$ is above $\underline{x}(\theta_1)$ so that $\rho_\theta(\theta_1, \pi) = \underline{x}(\theta_1) - \overline{x}(\theta_1) < 0$. However, since (14) binds at θ_1, I have $\rho(\theta_1, \pi) = 0$, so $\rho(\theta, \pi) < 0$ for θ close to but larger than θ_1.

To justify the new single-crossing condition (25) from first principles, I impose:

Assumption iii: $S_{xxq}(x, q, \theta) > 0$ for all (x, q, θ) and (to be verified!)

$$\min\left\{\frac{\partial}{\partial\theta}\frac{F(\theta\,|\overline{\eta})}{f(\theta\,|\overline{\eta})}, -\frac{\partial}{\partial\theta}\frac{1-F(\theta\,|\overline{\eta})}{f(\theta\,|\overline{\eta})}\right\} \geq \max\left\{\frac{\partial}{\partial\theta}\frac{F(\theta\,|\underline{\eta})}{f(\theta\,|\underline{\eta})}, \frac{\partial}{\partial\theta}\frac{\frac{\beta}{1-\beta}+F(\theta\,|\underline{\eta})}{f(\theta\,|\underline{\eta})}\right\}.$$

Notice that assumption iii is met if the conditional distributions of θ satisfy $f_\theta(\theta\,|\underline{\eta}) \geq f_\theta(\theta\,|\overline{\eta}) \geq 0$ and $\frac{\partial}{\partial\theta}\frac{F(\theta|\overline{\eta})}{f(\theta|\overline{\eta})} \geq \frac{\partial}{\partial\theta}\frac{F(\theta|\underline{\eta})}{f(\theta|\underline{\eta})}$.

The solution to the regulator´s problem takes the following form:

Proposition 6 *If $S(x, q, \theta)$ is increasing in q on the relevant domain, and under Assumptions i and iii, the solution to problem P^4 is given by the quality schedules $\overline{q}^*(\theta) = \overline{\eta}$ and $\underline{q}^*(\theta) = \underline{\eta}$ in conjunction with either*
i) $\pi^* = 0$, quantity schedules $\underline{x}^*(\theta) = \overline{x}^*(\theta) \equiv x^*(\theta)$ where $x^*(\theta)$ satisfies

$$\mathcal{E}_{H|\theta}\left[V_x(x^*(\theta), \eta) - C_x(x^*(\theta), \eta, \theta) \mid \Theta = \theta\right] - (1-\alpha)\frac{F(\theta)}{f(\theta)} = 0, \tag{26}$$

and the associated transfer schedules $\overline{t}^*(\theta)$ and $\underline{t}^*(\theta)$ defined by (12); or
ii) $\pi^* > 0$, quantity schedules $\underline{x}^*(\theta)$ and $\overline{x}^*(\theta)$ that satisfy (26) for $\theta < \theta^*$ and (23) and (24) for $\theta \geq \theta^*$, and the associated transfer schedules given by (12). θ^* is the unique intersection of schedules (23) and (24), if such an intersection exists; otherwise $\theta^* = \underline{\theta}$.

There is bunching of different η types who have the same marginal cost parameter θ at the low end of the θ support; at the high end there is separation of such types. Moroever, at the point where the regime changes from bunching to separation of different η types, the solution schedules switch continuously from one regime to the other. The idea to show this is the following. The value of the regulator's payoff function at an optimum should be invariant to small changes in the switch-point θ'. This requires that, conditional on $\theta = \theta'$, the expected value of the objective at θ' - where there is bunching - should be the same as the expected value of the objective just after the switch point, that is at $\theta = \theta' + \varepsilon$ for ε positive but arbitrarily small. This value matching condition essentially boils down to requiring continuity of the solution schedules.

I can now look into the regulator's interest, if any, to leave a positive rent to type $(\overline{\theta}, \overline{\eta})$. While this would never happen in the case of known $\eta-$types, the advantage in the current context is that a higher $\pi = \pi(\overline{\theta}, \overline{\eta})$ shifts the switch point θ' to the left. In other words, there is a new trade-off between rent extraction and efficiency. The solution schedules are constrained efficient given the informational asymmetry about θ. Raising π allows the regulator to get closer to the optimal solution schedules that are not constrained by the informational asymmetry about η. Recall from condition (7) in Proposition 1 that $x^*\left(\overline{\theta};\underline{\eta}\right)$ is the optimal quantity allocation for type $\overline{\theta}$ if the quality capacity is known to be $\underline{\eta}$. Moreover, let $x^{fb}\left(\overline{\theta},\overline{\eta}\right)$ denote the first-best efficient quantity allocation for type $\left(\overline{\theta},\overline{\eta}\right)$. I have the following result.

Proposition 7 *At the solution to problem P^4, $\pi^* > 0$ for β sufficiently small if and only if* $x^*\left(\overline{\theta};\underline{\eta}\right) > x^{fb}\left(\overline{\theta},\overline{\eta}\right).$

Using standard envelope theorems I observe that the effect of a small increase in π on the regulator's payoff is equal to $-(1-\alpha)\beta + k(\pi)$. The marginal cost of setting $\pi > 0$ is that an additional rent of π has to be left to all firms with a high quality capacity. There is a measure β of such firms and the cost enters the regulator's objective with a weight of $(1-\alpha)$. On the other hand, there is a benefit to raising π which is measured by $k(\pi)$, the value of the multiplier attached to constraint (14) over the separation region $\left[\theta',\overline{\theta}\right]$. Clearly, $k(\pi)$ is the higher the more the presence of firms with quality capacity $\overline{\eta}$ impinges the regulator from pursuing an optimal regulation policy for firms with a lower quality capacity. In particular, if the regulator would, when he knew η but not θ, have the firm of type $\left(\overline{\theta};\underline{\eta}\right)$ produce more than the first-best efficient amount for a firm of type $\left(\overline{\theta},\overline{\eta}\right)$, then the value of the multiplier is larger than the shadow cost of raising π. Since the regulator's objective is concave in π, this argument shows that the unique optimal π is positive under the conditions given in the proposition.

6 The case of non-binding capacity constraints

When $\underline{\eta}$ is large enough, the parameter η loses its relevance altogether: the solution schedules for quality, quantity, and transfer all become independent of η. To see this formally, simply observe that when $q(\theta,\eta) \leq \underline{\eta}$, I can apply Lemma 2 to all types (θ,η). Hence, it follows that for each θ, $x(\theta,\eta)$ and $\pi(\theta,\eta)$ must be independent of η. Consider now the quality schedule. It is easy to see that the schedule must satisfy the condition $V_q(x(\theta,\eta), q^*(\theta,\eta)) = C_q(x(\theta,\eta), q^*(\theta,\eta), \theta)$ for all θ,η. While Proposition 3 only provided a lower bound on the quality allocation, this bound must be tight if $q^*(\theta,\eta) < \eta$ for all θ,η. The reason is as follows. In Proposition 3 I can only derive a lower bound on the quality allocation, because I must make sure not to introduce additional possibilities for deviations for any type. If I increase the quality schedule pointwise, I will never increase these possibilities, but will at most reduce them or leave them the same. But among allocations where every type can mimic any other type, this caveat is irrelevant: Hence an allocation cannot be optimal if a marginal change in the quality allocation raises surplus.

For a nonbinding capacity constraint, the capacity information becomes irrelevant altogether. More formally, I have the following result:

Proposition 8 *Consider the case where $S(x,q,\theta)$ is strictly concave in x and q and $\underline{\eta}$ is sufficiently high so that the solution satisfies $q^*(\theta;\eta) < \eta$ for all θ,η. Then, the optimal quantity schedule satisfies*

$$\left(V_x(x^*(\theta), q(\theta)) - C_x(x^*(\theta), q(\theta), \theta) - (1-\alpha)\, c_{x\theta}(x^*(\theta), \theta)\frac{F(\theta)}{f(\theta)}\right) = 0, \tag{27}$$

and the optimal quality schedule satisfies

$$V_q\left(x\left(\theta\right),q^{*}\left(\theta\right)\right)-C_q\left(x\left(\theta\right),q^{*}\left(\theta\right),\theta\right)=0. \tag{28}$$

The intuition for the result is straightforward. If the capacity constraint on q is never binding for any type, the regulator does not need to have this information, because the optimal allocation cannot depend on it. All types can mimic others both by exaggerating their quality capacity and by understating it. But then, to keep firms from misrepresenting their quality capacity, the firms' profits must be independent of η for each θ. Hence, the total cost of serving customers does not depend on η. But if all firms with the same θ receive the same profits, then it is also optimal to offer them the same allocation, that is to make them produce identical quality, quantity bundles.

7 Concluding Remarks

I have solved a regulation problem featuring two dimensional asymmetric information in some detail. A firm knows its marginal cost of production and the maximum level of quality it is able to provide, the regulator does not have either of this information. Depending on whether quality and quantity are substitutable or complementary in the social surplus function, and on whether quality and cost information are positively or negatively correlated (in the sense of reversed hazard rates of the distribution of marginal costs conditional on capacity information being increasing or decreasing in capacity information), pricing rules are either completely standard or completely different from their counterparts in one-dimensional models. In particular, in the complements/increasing reversed hazard rates case everything is exactly as it is in the one-dimensional case. However, in the substitutes/decreasing reversed hazard rates case, nothing is the same as in the one-dimensional case; some firms price below marginal costs and even the least efficient producers of a given quality level may obtain rents at the optimum.

The result, that under some conditions asymmetric information about one dimension of the problem is irrelevant, has also been found in other contexts. Malakhov and Vohra [2005a,2005b] and Iyengar and Kumar [2006] have studied auction problems where bidders' valuations and capacities for consumption are unknown. They show that the solution to the problem when only valuations are private information remains incentive compatible when the second dimension of private information is added. Beaudry, Blackorby, and Szalay [2009] have first obtained a result in the spirit of Proposition 6 of the present paper. The main difference between the Malakhov/Vohra and Iyengar/Kumar approaches to the present one is as follows. In their problem the principal has two choice variables, and one of them - the quantity allocated to an agent - interacts non-trivially with the agent's types. In contrast, the present results apply to problems where the principal has three choices to make and two of them - the quantity and the quality allocated to the agent - interact non-trivially with the agent's types. Taken together, these results demonstrate the usefulness of the model structures to obtain insights into the problem of multi-dimensional screening. In ongoing work, I apply this approach to a problem proposed by Che and Gale [2000] where the seller faces possibly budget-constrained buyers.

8 Appendix

Proof of Lemma 1. I begin showing that $x(\theta;\eta)$ must be non-increasing. Incentive compatibility requires that type $(\theta;\eta)$ has no incentive to mimic type $\left(\hat{\theta};\eta\right)$

$$\begin{aligned}&t(\theta;\eta)+x(\theta;\eta)P(x(\theta;\eta),q(\theta;\eta))-C(x(\theta;\eta),q(\theta;\eta),\theta)\\&\geq t\left(\hat{\theta};\eta\right)+x\left(\hat{\theta};\eta\right)P\left(x\left(\hat{\theta};\eta\right),q\left(\hat{\theta};\eta\right)\right)-C\left(x\left(\hat{\theta};\eta\right),q\left(\hat{\theta};\eta\right),\theta\right)\end{aligned}$$

and that type $\left(\hat{\theta};\eta\right)$ has no incentive to mimic type $(\theta;\eta)$

$$\begin{aligned}&t\left(\hat{\theta};\eta\right)+x\left(\hat{\theta};\eta\right)P\left(x\left(\hat{\theta};\eta\right),q\left(\hat{\theta};\eta\right)\right)-C\left(x\left(\hat{\theta};\eta\right),q\left(\hat{\theta};\eta\right),\hat{\theta}\right)\\&\geq t(\theta;\eta)+x(\theta;\eta)P(x(\theta;\eta),q(\theta;\eta))-C\left(x(\theta;\eta),q(\theta;\eta),\hat{\theta}\right).\end{aligned}$$

Summing these inequalities gives

$$C\left(x\left(\hat{\theta};\eta\right),q\left(\hat{\theta};\eta\right),\theta\right)-C\left(x\left(\hat{\theta};\eta\right),q\left(\hat{\theta};\eta\right),\hat{\theta}\right)\geq C(x(\theta;\eta),q(\theta;\eta),\theta)-C\left(x(\theta;\eta),q(\theta;\eta),\hat{\theta}\right).$$

Writing as integrals, I have

$$\int_{\hat{\theta}}^{\theta}C_z\left(x\left(\hat{\theta};\eta\right),q\left(\hat{\theta};\eta\right),z\right)dz\geq\int_{\hat{\theta}}^{\theta}C_z(x(\theta;\eta),q(\theta;\eta),z)\,dz.$$

Since $C_{q\theta}=0$, this is equivalent to

$$\int_{\hat{\theta}}^{\theta}c_z\left(x\left(\hat{\theta};\eta\right),z\right)dz\geq\int_{\hat{\theta}}^{\theta}c_z(x(\theta;\eta),z)\,dz.$$

Given that $c_{x\theta}\geq 0$ this requires for $\theta>\hat{\theta}$ that $x\left(\hat{\theta};\eta\right)\geq x(\theta;\eta)$.

Since $x(\theta;\eta)$ is monotonic in θ it is differentiable almost everywhere in θ. Then, by the envelope theorem, almost everywhere

$$\pi_\theta(\theta;\eta)=-c_\theta(x(\theta;\eta),\theta).$$

Since $\pi_\theta(\theta;\eta)<0$, the participation constraint must be binding for type $\overline{\theta}$, so $\pi\left(\overline{\theta};\eta\right)=0$. Thus, I can write

$$\pi(\theta;\eta)=\int_{\theta}^{\overline{\theta}}c_y(x(y;\eta),y)\,dy.$$

Since $\pi(\theta;\eta)=t(\theta;\eta)+x(\theta;\eta)P(x(\theta;\eta),q(\theta;\eta))-C(x(\theta;\eta),q(\theta;\eta),\theta)$, I can write

$$t(\theta;\eta)=C(x(\theta;\eta),q(\theta;\eta),\theta)+\int_{\theta}^{\overline{\theta}}c_y(x(y;\eta),y)\,dy-x(\theta;\eta)P(x(\theta;\eta),q(\theta;\eta)),$$

which is the expression (5) given in the Lemma.

Consider now the sufficiency part. Almost everywhere $\Pi_{\hat{\theta}}\left(\hat{\theta},\theta;\eta\right)\Big|_{\hat{\theta}=\theta}=0$. Differentiating totally with respect to $\hat{\theta}$ and θ, and equating $d\hat{\theta} = d\theta$, I find

$$\Pi_{\hat{\theta}\hat{\theta}}\left(\hat{\theta},\theta;\eta\right)\Big|_{\hat{\theta}=\theta} = -\,\Pi_{\hat{\theta}\theta}\left(\hat{\theta},\theta;\eta\right)\Big|_{\hat{\theta}=\theta}.$$

Since

$$-\,\Pi_{\hat{\theta}\theta}\left(\hat{\theta},\theta;\eta\right)\Big|_{\hat{\theta}=\theta} = c_{\theta x}\left(x\left(\theta;\eta\right),\theta\right)\frac{dx}{d\theta} \leq 0,$$

I find that the single crossing condition in conjunction with $\frac{dx}{d\theta} \leq 0$ implies that the local second order condition is satisfied. Finally, consider non-local deviations to $\hat{\theta}$. I can write

$$\Pi\left(\hat{\theta},\theta;\eta\right) = \Pi\left(\hat{\theta},\hat{\theta};\eta\right) + C\left(x\left(\hat{\theta};\eta\right),q\left(\hat{\theta};\eta\right),\hat{\theta}\right) - C\left(x\left(\hat{\theta};\eta\right),q\left(\hat{\theta};\eta\right),\theta\right).$$

Hence, incentive compatibility requires that

$$\Pi\left(\theta,\theta;\eta\right) \geq \Pi\left(\hat{\theta},\hat{\theta};\eta\right) + C\left(x\left(\hat{\theta};\eta\right),q\left(\hat{\theta};\eta\right),\hat{\theta}\right) - C\left(x\left(\hat{\theta};\eta\right),q\left(\hat{\theta};\eta\right),\theta\right).$$

Subtracting $\Pi\left(\hat{\theta},\hat{\theta};\eta\right)$ from both sides, substituting for $\pi\left(\theta;\eta\right)$ and $\pi\left(\hat{\theta};\eta\right)$, respectively, and using

$$\int_{\hat{\theta}}^{\theta} \pi_y\left(y;\eta\right)dy = -\int_{\hat{\theta}}^{\theta} c_y\left(x\left(y;\eta\right),y\right)dy,$$

I can write this as

$$-\int_{\hat{\theta}}^{\theta} c_y\left(x\left(y;\eta\right),y\right)dy \geq C\left(x\left(\hat{\theta};\eta\right),q\left(\hat{\theta};\eta\right),\hat{\theta}\right) - C\left(x\left(\hat{\theta};\eta\right),q\left(\hat{\theta};\eta\right),\theta\right).$$

Using additive separability of the cost function, I can substitute $-\int_{\hat{\theta}}^{\theta} c_y\left(x\left(\hat{\theta};\eta\right),y\right)dy$ for the difference on the right hand side. Rearranging the resulting inequality, I have

$$\int_{\hat{\theta}}^{\theta} c_y\left(x\left(\hat{\theta};\eta\right),y\right)dy \geq \int_{\hat{\theta}}^{\theta} c_y\left(x\left(y;\eta\right),y\right)dy.$$

For $\hat{\theta} > \theta$, $x\left(\hat{\theta};\eta\right) \geq x\left(y;\eta\right)$ for all $y \leq \hat{\theta}$, and hence the inequality is satisfied. For $\hat{\theta} < \theta$, $x\left(\hat{\theta};\eta\right) \leq x\left(y;\eta\right)$ for all $y \geq \hat{\theta}$. Switching the lower and upper bound of integration yields the desired insight. ■

Proof of Proposition 1. I first prove as a preliminary result, that for a given η, $B\left(x,q,\theta,\eta\right)$ has non-decreasing differences (x,q) and θ. I observe that for given η, $B\left(x,q,\theta,\eta\right)$ has non-increasing differences in (x,q) and θ if and only if $c\left(x,\theta\right)+\left(1-\alpha\right)c_\theta\left(x,\theta\right)\frac{F(\theta|\eta)}{f(\theta|\eta)}$ has non-decreasing differences in x and θ for given η. To see this, recall that, for given η, $B\left(x,q,\theta,\eta\right)$ is said to have non-decreasing differences in (x,q) and θ if $B\left(x,q,\theta\right) - B\left(x,q,\theta'\right)$ is non-decreasing in (x,q) for every $\theta \geq \theta'$.

Given additive separability of the cost function, $C(x,q,\theta) = K + c(x,\theta) + k(x,q)$, I have

$$\begin{aligned} & B(x,q,\theta) - B\left(x,q,\theta'\right) \\ = \quad & c\left(x,\theta'\right) + (1-\alpha)\, c_\theta\left(x,\theta'\right) \frac{F\left(\theta' \,|\eta\right)}{f\left(\theta' \,|\eta\right)} - \left(c(x,\theta) + (1-\alpha)\, c_\theta(x,\theta) \frac{F(\theta\,|\eta)}{f(\theta\,|\eta)} \right), \end{aligned}$$

which proves indeed the result. Writing this difference in terms of integrals, I obtain

$$\int_{\theta'}^{\theta} \left(c_\tau(x,\tau) + (1-\alpha)\, c_{\tau\tau}(x,\tau) \frac{F(\tau\,|\eta)}{f(\tau\,|\eta)} + (1-\alpha)\, c_\tau(x,\tau) \frac{\partial}{\partial\tau} \frac{F(\tau\,|\eta)}{f(\tau\,|\eta)} \right) d\tau.$$

Differentiating the integrand with respect to x, I find that

$$c_{\tau x}(x,\tau) + (1-\alpha)\, c_{\tau\tau x}(x,\tau) \frac{F(\tau\,|\eta)}{f(\tau\,|\eta)} + (1-\alpha)\, c_{\tau x}(x,\tau) \frac{\partial}{\partial\tau} \frac{F(\tau\,|\eta)}{f(\tau\,|\eta)} \geq 0$$

implies that $c(x,\theta) + (1-\alpha)\, c_\theta(x,\theta) \frac{F(\theta|\eta)}{f(\theta|\eta)}$ has non-decreasing differences in x and θ.

i) are the first-order conditions resulting from pointwise maximization of the integrand in problem $\hat{\mathrm{P}}^1$; given the preliminary result, $x^*(\theta;\eta)$ is monotonic in θ and therefore $x^*(\theta;\eta)$ characterizes the optimum. ii) follows from Topkis (1978); see also Theorem 2.8.1 in Topkis (1998). iii) follows also from Theorem 2.8.1 in Topkis (1998) when I maximize over x and $\hat{q}$. ■

Proof of Proposition 2. As a preliminary, I first show that $B(x,q,\theta,\eta)$ has non-decreasing (non-increasing) differences in (x,q) and η for given θ if $\frac{F(\theta|\eta)}{f(\theta|\eta)}$ is non-increasing in η (non-decreasing in η). I can write

$$B(x,q,\theta,\eta) - B(x,q,\theta,\eta') = (1-\alpha)\, c_\theta(x,\theta) \left(\frac{F(\theta\,|\eta')}{f(\theta\,|\eta')} - \frac{F(\theta\,|\eta)}{f(\theta\,|\eta)} \right),$$

Differentiating with respect to x, I have

$$\frac{\partial}{\partial x}\left(B(x,q,\theta,\eta) - B(x,q,\theta,\eta')\right) = (1-\alpha)\, c_{x\theta}(x,\theta) \left(\frac{F(\theta\,|\eta')}{f(\theta\,|\eta')} - \frac{F(\theta\,|\eta)}{f(\theta\,|\eta)} \right),$$

which proves the statement.

The proof of the proposition is now a direct application of Theorem 2.3 part ii) and remark 10 in Vives [1999]. i) is direct; to see ii) recall that $B(x,\hat{q},\theta,\eta)$ is supermodular in this case, which implies the result follows again from Vives theorem. ■

Proof of Lemma 2. i) From the fact that $q(\theta,\eta'') \leq \eta'$ for $\theta \in \left[\theta', \theta'''\right]$ I know that any type (θ,η') for $\theta \in \left[\theta', \theta'''\right]$ can mimic any type (θ,η'') for $\theta \in \left[\theta', \theta'''\right]$. Hence, incentive compatibility requires that

$$\begin{aligned} & t(\theta,\eta) + x(\theta,\eta)\, P\left(x(\theta,\eta), q(\theta,\eta)\right) - C\left(x(\theta,\eta), q(\theta,\eta), \theta\right) \\ & \geq t\left(\hat{\theta},\hat{\eta}\right) + x\left(\hat{\theta},\hat{\eta}\right) P\left(x\left(\hat{\theta},\hat{\eta}\right), q\left(\hat{\theta},\hat{\eta}\right)\right) - C\left(x\left(\hat{\theta},\hat{\eta}\right), q\left(\hat{\theta},\hat{\eta}\right), \theta\right) \end{aligned}$$

and that type $\left(\hat{\theta},\hat{\eta}\right)$ has no incentive to mimic type (θ,η)

$$\begin{aligned} & t\left(\hat{\theta},\hat{\eta}\right)+x\left(\hat{\theta},\hat{\eta}\right)P\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right)\right)-C\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right),\hat{\theta}\right) \\ & \geq t(\theta,\eta)+x(\theta,\eta)P(x(\theta,\eta),q(\theta,\eta))-C\left(x(\theta,\eta),q(\theta,\eta),\hat{\theta}\right). \end{aligned}$$

Summing these inequalities gives

$$C\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right),\theta\right)-C\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right),\hat{\theta}\right)\geq C(x(\theta,\eta),q(\theta,\eta),\theta)-C\left(x(\theta,\eta),q(\theta,\eta),\hat{\theta}\right).$$

Writing as integrals, I have (using $C_{q\theta}=0$)

$$\int_{\hat{\theta}}^{\theta} c_z\left(x\left(\hat{\theta},\hat{\eta}\right),z\right)dz\geq\int_{\hat{\theta}}^{\theta} c_z\left(x(\theta,\eta),z\right)dz.$$

Given that $c_{x\theta}\geq 0$ this requires for $\theta>\hat{\theta}$ that $x\left(\hat{\theta},\hat{\eta}\right)\geq x(\theta,\eta)$.

ii) Suppose $x(\theta,\eta')$ is continuous in θ at θ''. From the analysis of the one-dimensional problem, I know that $x(\theta',\eta')\geq x(\theta'',\eta')\geq x(\theta''',\eta')$. Taking limits, I have, by continuity, $\lim_{\theta'\nearrow\theta''}x(\theta',\eta')=\lim_{\theta'''\searrow\theta''}x(\theta''',\eta')=x(\theta'',\eta')$. I have just shown in part i) that $x(\theta',\eta')\geq x(\theta'',\eta'')\geq x(\theta''',\eta')$. But, this can be true for all $\theta'\leq\theta''\leq\theta'''$ only if $x(\theta'',\eta')=x(\theta'',\eta'')$.

iii) For any $\theta\in[\theta',\theta''']$, $q(\theta,\eta'')\leq\eta'$, and hence incentive compatibility for type (θ,η') requires that he has no incentive to mimic type (θ,η''), so

$$\Pi(\theta,\theta,\eta')\geq\Pi(\theta,\theta,\eta'').$$

For type (θ,η''), incentive compatibility requires that

$$\Pi(\theta,\theta,\eta'')\geq\Pi(\theta,\theta,\eta').$$

Both inequalities can hold only if

$$\Pi(\theta,\theta,\eta'')=\Pi(\theta,\theta,\eta').$$

■

Proof of Lemma 3. It is easy to see that the two one dimensional constraints are necessary for incentive compatibility. I now show they are also sufficient for incentive compatibility. Suppose a type (θ,η) mimics a type $\left(\hat{\theta},\hat{\eta}\right)$ with $q\left(\hat{\theta},\hat{\eta}\right)\leq\eta$, so that the deviation is feasible for type (θ,η). I will show that (10), (11), imply that such a deviation is suboptimal. There are two types of deviations to consider. i) $\hat{\eta}\leq\eta$; ii) $\hat{\eta}>\eta$.

The proof for case i) is very simple. By mimicking type $\left(\hat{\theta},\hat{\eta}\right)$, type (θ,η) he obtains profit $\Pi\left(\hat{\theta},\theta,\hat{\eta}\right)\equiv t\left(\hat{\theta},\hat{\eta}\right)+x\left(\hat{\theta},\hat{\eta}\right)P\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right)\right)-C\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right),\theta\right)$. But by incentive compatibility for type $(\theta,\hat{\eta})$ in the θ dimension, type (θ,η) could obtain more by mimicking type $(\theta,\hat{\eta})$, since

$$\Pi(\theta,\theta,\hat{\eta})\geq\Pi\left(\hat{\theta},\theta,\hat{\eta}\right).$$

But then, by incentive compatibility for type (θ, η) in the η dimension, I have

$$\Pi(\theta, \theta, \eta) \geq \Pi(\theta, \theta, \hat{\eta}).$$

Hence, the deviation gives a weakly lower profit than announcing the true type.

ii) From Lemma 2 $x\left(\hat{\theta}, \hat{\eta}\right) = x\left(\hat{\theta}, \eta\right)$ and $\Pi\left(\hat{\theta}, \hat{\theta}, \eta\right) = \Pi\left(\hat{\theta}, \hat{\theta}, \hat{\eta}\right)$. From (10) for type (θ, η), I have

$$\Pi(\theta, \theta, \eta) \geq \Pi\left(\hat{\theta}, \theta, \eta\right)$$

Now, using the additive separability of the cost function,

$$\begin{aligned}
\Pi\left(\hat{\theta}, \theta, \eta\right) &= \Pi\left(\hat{\theta}, \hat{\theta}, \eta\right) + C\left(x\left(\hat{\theta}, \eta\right), q\left(\hat{\theta}, \eta\right), \hat{\theta}\right) - C\left(x\left(\hat{\theta}, \eta\right), q\left(\hat{\theta}, \eta\right), \theta\right) \\
&= \Pi\left(\hat{\theta}, \hat{\theta}, \eta\right) + c\left(x\left(\hat{\theta}, \eta\right), \hat{\theta}\right) - c\left(x\left(\hat{\theta}, \eta\right), \theta\right).
\end{aligned}$$

Similarly, I can write

$$\Pi\left(\hat{\theta}, \theta, \hat{\eta}\right) = \Pi\left(\hat{\theta}, \hat{\theta}, \hat{\eta}\right) + c\left(x\left(\hat{\theta}, \hat{\eta}\right), \hat{\theta}\right) - c\left(x\left(\hat{\theta}, \hat{\eta}\right), \theta\right)$$

Since $\Pi\left(\hat{\theta}, \hat{\theta}, \hat{\eta}\right) = \Pi\left(\hat{\theta}, \hat{\theta}, \eta\right)$ and $x\left(\hat{\theta}, \hat{\eta}\right) = x\left(\hat{\theta}, \eta\right)$ it follows that

$$\Pi\left(\hat{\theta}, \theta, \hat{\eta}\right) = \Pi\left(\hat{\theta}, \theta, \eta\right).$$

Hence, I have shown that $\Pi(\theta, \theta, \eta) \geq \Pi\left(\hat{\theta}, \theta, \eta\right)$ implies also $\Pi(\theta, \theta, \eta) \geq \Pi\left(\hat{\theta}, \theta, \hat{\eta}\right)$. ■

Proof of Lemma 4. By Lemma 3, the one dimensional constraints are necessary and sufficient for the constraint ruling out two-dimensional deviations. Applying the same procedure as in Lemma 1 to (10), I get conditions (12) and (13). By Lemma 1 (12) and (13) in conjunction with monotonicity of x are necessary and sufficient for (10). ii) follows directly from (11), because mimicking a firm with a lower quality capacity is always feasible. iii) is simply restating conditions ii) and iii) from Lemma 2. ■

Proof of Proposition 3. Take any incentive compatible allocation given by the triple of schedules $x\left(\hat{\theta}, \hat{\eta}\right), q\left(\hat{\theta}, \hat{\eta}\right)$, and $t\left(\hat{\theta}, \hat{\eta}\right)$ for all $\hat{\theta}, \hat{\eta}$. Suppose, contrary to the Proposition, that for some $\left(\hat{\theta}, \hat{\eta}\right)$, I have $q^*\left(\hat{\theta}, \hat{\eta}\right) < \eta$ and

$$V_q\left(x\left(\hat{\theta}, \hat{\eta}\right), q^*\left(\hat{\theta}, \hat{\eta}\right)\right) - C_q\left(x\left(\hat{\theta}, \hat{\eta}\right), q^*\left(\hat{\theta}, \hat{\eta}\right), \theta\right) > 0.$$

Let Φ denote the set of $\left(\hat{\theta}, \hat{\eta}\right)$ for which this inequality holds. Then, I can change the allocation to the new triple of schedules $\tilde{x}\left(\hat{\theta}, \hat{\eta}\right), \tilde{q}\left(\hat{\theta}, \hat{\eta}\right)$, and $\tilde{t}\left(\hat{\theta}, \hat{\eta}\right)$ for all $\hat{\theta}, \hat{\eta}$ as follows. I set $\tilde{q}\left(\hat{\theta}, \hat{\eta}\right)$ to either $\tilde{q}\left(\hat{\theta}, \hat{\eta}\right) = \hat{\eta}$ or, if feasible to $\tilde{q}\left(\hat{\theta}, \hat{\eta}\right)$ satisfying

$$V_q\left(\tilde{x}\left(\hat{\theta}, \hat{\eta}\right), \tilde{q}\left(\hat{\theta}, \hat{\eta}\right)\right) - C_q\left(\tilde{x}\left(\hat{\theta}, \hat{\eta}\right), \tilde{q}\left(\hat{\theta}, \hat{\eta}\right), \theta\right) = 0$$

Moreover, I set $\tilde{x}\left(\hat{\theta}, \hat{\eta}\right) = x\left(\hat{\theta}, \hat{\eta}\right)$ for all $\hat{\theta}, \hat{\eta}$. For all $\left(\hat{\theta}, \hat{\eta}\right) \in \Phi$ I adjust the transfers from the initial transfers $t\left(\hat{\theta},\hat{\eta}\right)$, to the new transfers

$$\begin{aligned}\tilde{t}\left(\hat{\theta},\hat{\eta}\right) &= t\left(\hat{\theta},\hat{\eta}\right)+x\left(\hat{\theta},\hat{\eta}\right)P\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right)\right)-\tilde{x}\left(\hat{\theta},\hat{\eta}\right)P\left(\tilde{x}\left(\hat{\theta},\hat{\eta}\right),\tilde{q}\left(\hat{\theta},\hat{\eta}\right)\right)\\ &\quad +C\left(\tilde{x}\left(\hat{\theta},\hat{\eta}\right),\tilde{q}\left(\hat{\theta},\hat{\eta}\right),\hat{\theta}\right)-C\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right),\hat{\theta}\right).\end{aligned}$$

I first show that the new allocation with the new transfers is incentive compatible. Then, I show that the surplus to the principal has increased under the new allocation.

By assumption, the functions p,q and t are incentive compatible, that is for all θ,η and all $\hat{\theta},\hat{\eta}$ such that $q\left(\hat{\theta},\hat{\eta}\right)\leq\eta$

$$\begin{aligned}& t(\theta,\eta)+x(\theta,\eta)P(x(\theta,\eta),q(\theta,\eta))-C(x(\theta,\eta),q(\theta,\eta),\theta)\\ &\geq t\left(\hat{\theta},\hat{\eta}\right)+x\left(\hat{\theta},\hat{\eta}\right)P\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right)\right)-C\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right),\theta\right).\end{aligned} \tag{29}$$

Incentive compatibility of the new functions $\tilde{p},\tilde{q}$, and $\tilde{t}$ is given if and only if for all θ,η and all $\hat{\theta}$ and all $\tilde{q}\left(\hat{\theta},\hat{\eta}\right)\leq\eta$

$$\begin{aligned}& \tilde{t}(\theta,\eta)+\tilde{x}(\theta,\eta)P(\tilde{x}(\theta,\eta),\tilde{q}(\theta,\eta))-C(\tilde{x}(\theta,\eta),\tilde{q}(\theta,\eta),\theta)\\ &\geq \tilde{t}\left(\hat{\theta},\hat{\eta}\right)+\tilde{x}\left(\hat{\theta},\hat{\eta}\right)P\left(\tilde{x}\left(\hat{\theta},\hat{\eta}\right),\tilde{q}\left(\hat{\theta},\hat{\eta}\right)\right)-C\left(\tilde{x}\left(\hat{\theta},\hat{\eta}\right),\tilde{q}\left(\hat{\theta},\hat{\eta}\right),\theta\right).\end{aligned} \tag{30}$$

Substituting for $\tilde{t}\left(\hat{\theta},\hat{\eta}\right)$, and simplifying, this amounts to

$$\begin{aligned}& t(\theta,\eta)+x(\theta,\eta)P(x(\theta,\eta),q(\theta,\eta))-C(x(\theta,\eta),q(\theta,\eta),\theta)\\ &+C(\tilde{x}(\theta,\eta),\tilde{q}(\theta,\eta),\theta)-C(\tilde{x}(\theta,\eta),\tilde{q}(\theta,\eta),\theta)\\ &\geq t\left(\hat{\theta},\hat{\eta}\right)+x\left(\hat{\theta},\hat{\eta}\right)P\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right)\right)-C\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right),\hat{\theta}\right)\\ &+C\left(\tilde{x}(\theta,\eta),\tilde{q}(\theta,\eta),\hat{\theta}\right)-C\left(\tilde{x}(\theta,\eta),\tilde{q}\left(\hat{\theta},\hat{\eta}\right),\theta\right).\end{aligned}$$

Adding and subtracting $C\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right),\theta\right)$ on the right-hand side, and noting (29), the new incentive constraint is satisfied if

$$\begin{aligned}0 &\geq C\left(\tilde{x}\left(\hat{\theta},\hat{\eta}\right),\tilde{q}\left(\hat{\theta},\hat{\eta}\right),\hat{\theta}\right)-C\left(\tilde{x}\left(\hat{\theta},\hat{\eta}\right),\tilde{q}\left(\hat{\theta},\hat{\eta}\right),\theta\right)\\ &\quad +C\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right),\theta\right)-C\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right),\hat{\theta}\right).\end{aligned}$$

Using $C(x,q,\theta)=K+c(x,\theta)+k(x,q)$, and $X\left(p\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right)\right)=X\left(\tilde{p}\left(\hat{\theta},\hat{\eta}\right),\tilde{q}\left(\hat{\theta},\hat{\eta}\right)\right)\equiv x\left(\hat{\theta},\hat{\eta}\right)$, I can write this as

$$\begin{aligned}0 &\geq c\left(x\left(\hat{\theta},\hat{\eta}\right),\hat{\theta}\right)-c\left(x\left(\hat{\theta},\hat{\eta}\right),\theta\right)+c\left(x\left(\hat{\theta},\hat{\eta}\right),\theta\right)-c\left(x\left(\hat{\theta},\hat{\eta}\right),\hat{\theta}\right)\\ &\quad +k\left(x\left(\hat{\theta},\hat{\eta}\right),\tilde{q}\left(\hat{\theta},\hat{\eta}\right)\right)-k\left(x\left(\hat{\theta},\hat{\eta}\right),\tilde{q}\left(\hat{\theta},\hat{\eta}\right)\right)+k\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right)\right)-k\left(x\left(\hat{\theta},\hat{\eta}\right),q\left(\hat{\theta},\hat{\eta}\right)\right)\\ &= 0.\end{aligned}$$

Hence (30) is satisfied.

Finally, the system is overall incentive compatible. This follows, because the schedule $\tilde{q}(\theta,\eta)$

is weakly higher than the schedule $q(\theta,\eta)$, which implies that the set of messages $\left(\hat{\theta},\hat{\eta}\right)$ such that $\tilde{q}\left(\hat{\theta},\hat{\eta}\right)\leq\eta$ is a subset of the set of messages $\left(\hat{\theta},\hat{\eta}\right)$ such that $q\left(\hat{\theta},\hat{\eta}\right)\leq\eta$. Hence, for each type at most as many deviations are feasible in the new system as in the old system.

Consider now the government's surplus. By construction, the firm's equilibrium payoffs are unchanged, as

$$\begin{aligned}&t(\theta,\eta)+x(\theta,\eta)P(x(\theta,\eta),q(\theta,\eta))-\tilde{x}(\theta,\eta)P(\tilde{x}(\theta,\eta),\tilde{q}(\theta,\eta))\\&+C(\tilde{x}(\theta,\eta),\tilde{q}(\theta,\eta),\theta)-C(x(\theta,\eta),q(\theta,\eta),\theta)\\&+\tilde{x}(\theta,\eta)P(\tilde{x}(\theta,\eta),\tilde{q}(\theta,\eta))-C(\tilde{x}(\theta,\eta),\tilde{q}(\theta,\eta),\theta)\\&=t(\theta,\eta)+x(\theta,\eta)P(x(\theta,\eta),q(\theta,\eta))-C(x(\theta,\eta),q(\theta,\eta),\theta).\end{aligned}$$

In other words, the firm is just compensated for the increase in his cost of production, and the entire additional surplus goes to the principal. Hence, holding the quantity schedule $x(\theta,\eta)$ constant, any allocation that satisfies

$$V_q(x(\theta,\eta),q(\theta,\eta))-C_q(x(\theta,\eta),q(\theta,\eta),\theta)>0$$

and $q(\theta,\eta)<\eta$ can be improved by raising $q(\theta,\eta)$. ■

Proof of Lemma 5. Let $\beta=\Pr[\eta=\overline{\eta}]$. The "reduced" problem where I neglect the monotonicity constraints on $\overline{x}(\theta)$ and $\underline{x}(\theta)$ can be written as follows:

$$\Gamma(\pi)=\max_{\overline{u}(\theta),\underline{u}(\theta)}\left[\begin{array}{c}\beta\int_{\underline{\theta}}^{\overline{\theta}}\left(B\left(c_\theta^{-1}(\overline{u}(\theta),\theta),\overline{\eta},\theta,\overline{\eta}\right)\right)f(\theta\,|\overline{\eta})\,d\theta-\beta(1-\alpha)\pi\\+(1-\beta)\int_{\underline{\theta}}^{\overline{\theta}}B\left(c_\theta^{-1}(\underline{u}(\theta),\theta),\underline{\eta},\theta,\underline{\eta}\right)f\left(\theta\,|\underline{\eta}\right)d\theta\end{array}\right]$$

$$s.t.\int_{\theta}^{\overline{\theta}}\underline{u}(y)\,dy-\int_{\theta}^{\overline{\theta}}\overline{u}(y)\,dy\leq\pi$$

where the function B is defined in (6) and where I have defined the variables $\overline{u}(\theta)\equiv c_\theta(\overline{x}(\theta),\theta)$ and $\underline{u}(\theta)\equiv c_\theta(\underline{x}(\theta),\theta)$ and substituted the identities $\overline{x}(\theta)\equiv c_\theta^{-1}(\overline{u}(\theta),\theta)$ and $\underline{x}(\theta)\equiv c_\theta^{-1}(\underline{u}(\theta),\theta)$. Letting $\underline{z}\equiv-\int_\theta^{\overline{\theta}}c_\theta(\underline{x}(y),y)\,dy$ and $\overline{z}\equiv-\int_\theta^{\overline{\theta}}c_\theta(\overline{x}(y),y)\,dy$ I can note further that $\underline{u}=\underline{z}_\theta$ and $\overline{u}=\overline{z}_\theta$.

I can view this as a control problem with Hamiltonian of the following form:

$$\begin{aligned}H\;=\;&\left(B\left(c_\theta^{-1}(\overline{u}(\theta),\theta),\overline{\eta},\theta,\overline{\eta}\right)\right)\beta f(\theta\,|\overline{\eta})+\left(B\left(c_\theta^{-1}(\underline{u}(\theta),\theta),\underline{\eta},\theta,\underline{\eta}\right)\right)(1-\beta)f\left(\theta\,|\underline{\eta}\right)\\&+\overline{\kappa}\overline{u}+\underline{\kappa}\underline{u}+\mu(\pi-(\overline{z}-\underline{z}))\end{aligned}$$

Differentiating with respect to state variables, I get the conditions of optimality

$$\begin{aligned}\frac{\partial H}{\partial\overline{z}}&=-\mu=-\overline{\kappa}_\theta\\\frac{\partial H}{\partial\underline{z}}&=\mu=-\underline{\kappa}_\theta;\end{aligned}$$

differentiating with respect to the controls I get

$$\frac{\partial H}{\partial \overline{u}} = \left(\frac{V_{\overline{x}}\left(c_\theta^{-1}\left(\overline{u},\theta\right),\overline{\eta}\right) - C_{\overline{x}}\left(c_\theta^{-1}\left(\overline{u},\theta\right),\overline{\eta},\theta\right)}{c_{\theta x}\left(c_\theta^{-1}\left(\overline{u},\theta\right),\theta\right)} - (1-\alpha)\frac{F\left(\theta\left|\overline{\eta}\right.\right)}{f\left(\theta\left|\overline{\eta}\right.\right)} \right)\beta f\left(\theta\left|\overline{\eta}\right.\right) + \overline{\kappa} = 0 \tag{31}$$

$$\frac{\partial H}{\partial \underline{u}} = \left(\frac{V_{\underline{x}}\left(c_\theta^{-1}\left(\underline{u},\theta\right),\underline{\eta}\right) - C_{\underline{x}}\left(c_\theta^{-1}\left(\underline{u},\theta\right),\underline{\eta},\theta\right)}{c_{\theta x}\left(c_\theta^{-1}\left(\underline{u},\theta\right),\theta\right)} - (1-\alpha)\frac{F\left(\theta\left|\underline{\eta}\right.\right)}{f\left(\theta\left|\underline{\eta}\right.\right)} \right)(1-\beta) f\left(\theta\left|\underline{\eta}\right.\right) + \underline{\kappa} = 0$$

The Kuhn-Tucker conditions are

$$\pi - \left(\overline{z} - \underline{z}\right) \leq 0,\ \ \mu \geq 0,\ \text{ and } \mu\left(\pi - \left(\overline{z} - \underline{z}\right)\right) = 0.$$

For the transversality conditions, I have to distinguish two cases. If $\mu\left(\underline{\theta}\right) = 0$, then $\overline{z}\left(\underline{\theta}\right)$ and $\underline{z}\left(\underline{\theta}\right)$ are both free and the transversality conditions are

$$\overline{\kappa}\left(\underline{\theta}\right) = \underline{\kappa}\left(\underline{\theta}\right) = 0.$$

If $\mu\left(\underline{\theta}\right) > 0$, then $\overline{z}\left(\underline{\theta}\right)$ is fully determined once $\underline{z}\left(\underline{\theta}\right)$ is given and vice versa. Hence, I do not impose any transversality condition in this case.

Suppose that $\mu\left(\underline{\theta}\right) = 0$ and that $\mu\left(\theta\right) = 0$ on a set of positive measure $\left[\underline{\theta}, \theta'\right]$. From conditions (31) it is clear that $\overline{\kappa}$ and $\underline{\kappa}$ are continuously differentiable in θ whenever $\overline{u}$ and $\underline{u}$ are continuously differentiable in θ. Using the conditions of optimality for the state variables, $\overline{\kappa}_\theta = \mu$ and $\underline{\kappa}_\theta = -\mu$, and the transversality conditions - which must hold if $\mu\left(\underline{\theta}\right) = 0$ - I have for $\theta \leq \theta'$

$$\overline{\kappa}\left(\theta\right) = \overline{\kappa}\left(\underline{\theta}\right) + \int_{\underline{\theta}}^{\theta} \overline{\kappa}_\tau d\tau = \int_{\underline{\theta}}^{\theta} \mu\left(\tau\right) d\tau = 0$$

and

$$\underline{\kappa}\left(\theta\right) = \underline{\kappa}\left(\underline{\theta}\right) - \int_{\underline{\theta}}^{\theta} \mu\left(\tau\right) d\tau = 0.$$

Hence, for $\theta \in \left[\underline{\theta}, \theta'\right]$, I have

$$\frac{V_x\left(c_\theta^{-1}\left(\overline{u},\theta\right),\overline{\eta}\right) - C_x\left(c_\theta^{-1}\left(\overline{u},\theta\right),\overline{\eta},\theta\right)}{c_{\theta x}\left(c_\theta^{-1}\left(\overline{u},\theta\right),\theta\right)} - (1-\alpha)\frac{F\left(\theta\left|\overline{\eta}\right.\right)}{f\left(\theta\left|\overline{\eta}\right.\right)} = 0$$

and

$$\frac{V_x\left(c_\theta^{-1}\left(\underline{u},\theta\right),\underline{\eta}\right) - C_x\left(c_\theta^{-1}\left(\underline{u},\theta\right),\underline{\eta},\theta\right)}{c_{\theta x}\left(c_\theta^{-1}\left(\underline{u},\theta\right),\theta\right)} - (1-\alpha)\frac{F\left(\theta\left|\underline{\eta}\right.\right)}{f\left(\theta\left|\underline{\eta}\right.\right)} = 0.$$

Of course these conditions are equivalent to those obtained in Proposition 1. Using the comparative statics results from Proposition 4, I know that for $\theta \leq \theta'$

$$\underline{x}\left(\theta\right) > \overline{x}\left(\theta\right).$$

I have shown in the text that $\mu\left(\theta\right) = 0$ for all $\theta \in \left[\underline{\theta}, \theta'\right]$, where $\theta' > \underline{\theta}$, leads to a contradiction. ■

Proof of Proposition 5. The proof is split into two parts. In the first part, (after spelling out the Lagrangian of the problem), I prove that the first-order conditions are sufficient for an optimum. In the second part, I show that the solution is not only a solution to the reduced problem but satisfies also the neglected constraints.

The Lagrangian of the problem is

$$L = \max_{\overline{x}(\theta),\underline{x}(\theta),\pi} \begin{bmatrix} \beta \int_{\underline{\theta}}^{\overline{\theta}} B\left(\overline{x}\left(\theta\right), \overline{\eta}, \theta, \overline{\eta}\right) f\left(\theta \left|\overline{\eta}\right.\right) d\theta - \beta\left(1-\alpha\right)\pi \\ + \left(1-\beta\right)\left[\int_{\underline{\theta}}^{\overline{\theta}} B\left(\underline{x}\left(\theta\right), \underline{\eta}, \theta, \underline{\eta}\right) f\left(\theta \left|\overline{\eta}\right.\right) d\theta\right] \end{bmatrix} \tag{32}$$

$$+k\left(\int_{\underline{\theta}}^{\overline{\theta}} \left(\overline{x}\left(\theta\right) - \underline{x}\left(\theta\right)\right) d\theta + \pi\right)$$

i) The objective is strictly concave in $\overline{x}$ and $\underline{x}$ and the constraint is linear in these variables. By a standard theorem, the first-order conditions in $\overline{x}$ and $\underline{x}$, respectively, are also sufficient for an optimum. Consider now the derivatives with respect to π. I have

$$\frac{\partial L}{\partial \pi} = -\beta\left(1-\alpha\right) + k$$

and

$$\frac{\partial^2 L}{\partial \pi^2} = \frac{dk}{d\pi}.$$

I can compute $\frac{dk}{d\pi}$ from a total differentiation of the constraint. I obtain

$$\frac{dk}{d\pi} = \frac{1}{\int_{\underline{\theta}}^{\overline{\theta}} \left(\underline{x}_k\left(\theta;k\right) - \overline{x}_k\left(\theta;k\right)\right) d\theta}.$$

From a total differentiation of (19) and (20) , I get

$$\frac{d\overline{x}}{dk} = -\frac{1}{B_{xx}\left(\overline{x}\left(\theta\right), \overline{\eta}, \theta, \overline{\eta}\right) \beta f\left(\theta \left|\overline{\eta}\right.\right)} > 0 \tag{33}$$

and

$$\frac{d\underline{x}}{dk} = \frac{1}{B_{xx}\left(\underline{x}\left(\theta\right), \underline{\eta}, \theta, \underline{\eta}\right)\left(1-\beta\right) f\left(\theta \left|\underline{\eta}\right.\right)} < 0. \tag{34}$$

This shows that $\frac{dk}{d\pi} < 0$,and hence the problem is concave in the choice variable π. It follows that the condition of optimality is that either $\pi^* = 0$ and

$$k\left(\pi^*\right) \leq \beta\left(1-\alpha\right)$$

or $\pi^* > 0$ and

$$k\left(\pi^*\right) = \beta\left(1-\alpha\right).$$

Moreover, as $k\left(\overline{\pi}\right) = 0$, I note that $\pi^* < \overline{\pi}$.

ii) Consider first incentive compatibility in the θ dimension. The schedules (19) and (20) are continuous; hence they are differentiable everywhere and if they satisfy $\frac{d\overline{x}(\theta)}{d\theta} \leq 0$ and $\frac{d\underline{x}(\theta)}{d\theta} \leq 0$, they are monotonic. Consider first the schedule $\overline{x}\left(\theta\right)$. From a total differentiation of (19) , I obtain

$$\frac{d\overline{x}}{d\theta} = \frac{\left(1 + \left(1-\alpha\right) \frac{\partial}{\partial\theta} \frac{F\left(\theta|\overline{\eta}\right)}{f\left(\theta|\overline{\eta}\right)}\right) + \frac{k^*}{\beta} \frac{f_\theta\left(\theta|\overline{\eta}\right)}{f\left(\theta|\overline{\eta}\right)^2}}{V_{xx}\left(\overline{x}\left(\theta\right), \overline{\eta}\right) - C_{xx}\left(\overline{x}\left(\theta\right), \overline{\eta}, \theta\right)}. \tag{35}$$

Clearly, for $f_\theta\left(\theta \left|\overline{\eta}\right.\right) \geq 0$, $\frac{d\overline{x}(\theta)}{d\theta} \leq 0$. So, suppose $f_\theta\left(\theta \left|\overline{\eta}\right.\right) < 0$. In that case the numerator is bounded below by

$$1+(1-\alpha)\frac{\partial}{\partial\theta}\frac{F(\theta\,|\overline{\eta})}{f(\theta\,|\overline{\eta})}+(1-\alpha)\frac{f_\theta(\theta\,|\overline{\eta})}{f(\theta\,|\overline{\eta})^2},$$

where I have replaced k^* with its upper bound, $(1-\alpha)\beta$. The result follows from noting that $\frac{\partial}{\partial\theta}\frac{1-F(\theta|\overline{\eta})}{f(\theta|\overline{\eta})}\leq 0$ is equivalent to the requirement that the sum of the second and third term are non-negative.

A total differentiation of (20) gives

$$\frac{d\underline{x}}{d\theta}=\frac{\left(1+(1-\alpha)\frac{\partial}{\partial\theta}\frac{F(\theta|\underline{\eta})}{f(\theta|\underline{\eta})}\right)-\frac{k^*}{1-\beta}\frac{f_\theta(\theta|\underline{\eta})}{f(\theta|\underline{\eta})^2}}{V_{xx}(\underline{x}(\theta),\overline{\eta})-C_{xx}(\underline{x}(\theta),\underline{\eta},\theta)}. \tag{36}$$

Clearly, if $f_\theta(\theta\,|\underline{\eta})\leq 0$ I have $\frac{d\underline{x}}{d\theta}<0$, so suppose $f_\theta(\theta\,|\underline{\eta})>0$. Then again, the numerator above is bounded below by

$$1+(1-\alpha)\frac{\partial}{\partial\theta}\frac{F(\theta\,|\underline{\eta})}{f(\theta\,|\underline{\eta})}-(1-\alpha)\frac{\beta}{1-\beta}\frac{f_\theta(\theta\,|\underline{\eta})}{f(\theta\,|\underline{\eta})^2};$$

Now condition $\frac{\partial}{\partial\theta}\frac{\frac{\beta}{1-\beta}+F(\theta|\underline{\eta})}{f(\theta|\underline{\eta})}\geq 0$ is equivalent to the requirement that the sum of the second and third term are non-negative.

To demonstrate incentive compatibility in the η dimension, I show that the assumptions imply that the schedules $\overline{x}(\theta)$ and $\underline{x}(\theta)$ satisfy (17). I have $\overline{x}(\theta)=\underline{x}(\theta)\implies\frac{d\overline{x}}{d\theta}\geq\frac{d\underline{x}}{d\theta}$ if for $\overline{x}(\theta)=\underline{x}(\theta)$

$$\frac{1}{(V_{xx}(\overline{x}(\theta),\overline{\eta})-C_{\overline{xx}}(\overline{x}(\theta),\overline{\eta},\theta))}\geq\frac{1}{(V_{xx}(\underline{x}(\theta),\underline{\eta})-C_{xx}(\underline{x}(\theta),\underline{\eta},\theta))},$$

and

$$\left((1-\alpha)\frac{\partial}{\partial\theta}\frac{F(\theta\,|\overline{\eta})}{f(\theta\,|\overline{\eta})}\right)+\frac{k}{\beta}\frac{f_\theta(\theta\,|\overline{\eta})}{f(\theta\,|\overline{\eta})^2}\leq\left((1-\alpha)\frac{\partial}{\partial\theta}\frac{F(\theta\,|\underline{\eta})}{f(\theta\,|\underline{\eta})}\right)-\frac{k}{1-\beta}\frac{f_\theta(\theta\,|\underline{\eta})}{f(\theta\,|\underline{\eta})^2}. \tag{37}$$

The former condition is satisfied if $S_{xxq}(x,q,\theta)\leq 0$. Assumption ii implies the latter condition. To see this, observe that both the left-hand side and the right-hand side of (37) are linear in k. Hence, the maximum of the left-hand side and the minimum of the right-hand side, respectively, are reached either at $k=0$ or at $k=(1-\alpha)\beta$. ■

Proof of Lemma 7. To prove the lemma, I show the bunching region is convex. Together with Lemma 5, this demonstrates that bunching occurs only at the low end of the support, and does so over an interval.

The bunching region is convex if and only if there cannot exist two points θ' and θ'' with $\theta''>\theta'$ such that (14) binds at θ' and θ'' but not in between. Suppose the contrapositive were true, and two such points did exist. Then, I know that

$$\int_{\theta'}^{\theta''}\overline{x}(y)\,dy-\int_{\theta'}^{\theta''}\underline{x}(y)\,dy=0, \tag{38}$$

because (14) is binding at θ' and θ''. But then, I can split the problem into three subproblems, where each subproblem is to choose optimal schedules $\overline{x}(\theta)$ and $\underline{x}(\theta)$ on the (possibly empty) subintervals, $\left[\underline{\theta}, \theta'\right)$, $\left[\theta', \theta''\right]$, and $\left(\theta'', \overline{\theta}\right]$. This is possible, because I solve the "reduced" problem P[4] absent monotonicity constraints on $\overline{x}(\theta)$ and $\underline{x}(\theta)$. On the subinterval $\left[\theta', \theta''\right]$, the problem is identical to problem P[5] with θ' replacing $\underline{\theta}$ and θ'' replacing $\overline{\theta}$, and where $\pi = 0$. Hence, the optimal schedules satisfy conditions (19) and (20).

Suppose (25) holds. Then, the schedules $\overline{x}(\theta)$ and $\underline{x}(\theta)$ can only satisfy condition (38) if they cross at least once. Moreover, schedule $\underline{x}(\theta)$ must cross schedule $\overline{x}(\theta)$ from below. Hence, I must have $\overline{x}\left(\theta'\right) > \underline{x}\left(\theta'\right)$. But then

$$\rho_{\theta}\left(\theta', \pi\right) = \underline{x}\left(\theta'\right) - \overline{x}\left(\theta'\right) < 0$$

contradicting the supposition that (14) is non-binding for $\theta \in \left(\theta', \theta''\right)$. ■

Proof of Proposition 6. The proof is organized as follows. First, I establish continuity of the solution schedules at the switching point between the bunching and the separation region. Then, I demonstrate that the assumptions imply the monotonicity of the solution both in the θ and the η dimension.

i) Consider again the control problem spelled out in the proof of Lemma 5. I first show that for $\theta \leq \theta'$, the optimal schedule satisfies $\overline{x}(\theta) = \underline{x}(\theta) = x^*(\theta)$ and $x^*(\theta)$ solves

$$\mathcal{E}_{H|\theta}\left[V_x\left(x^*(\theta), \eta\right) - C_x\left(x^*(\theta), \eta, \theta\right) \mid \Theta = \theta\right] - (1-\alpha)\frac{F(\theta)}{f(\theta)} = 0. \tag{39}$$

To see this, I can use the conditions for $\overline{\kappa}(\theta)$ and $\underline{\kappa}(\theta)$ and the equations of motion for these costate variables to get

$$\overline{\kappa}(\theta) = \overline{\kappa}(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \overline{\kappa}_{\tau} d\tau = \overline{\kappa}(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \mu(\tau)\, d\tau$$

and

$$\underline{\kappa}(\theta) = \underline{\kappa}(\underline{\theta}) - \int_{\underline{\theta}}^{\theta} \mu(\tau)\, d\tau.$$

Substituting back into (31), using $c_{\theta x} = 1$, I get

$$\left(V_x\left(\overline{x}(\theta), \overline{\eta}\right) - C_x\left(\overline{x}(\theta), \overline{\eta}, \theta\right) - (1-\alpha)\frac{F(\theta\,|\overline{\eta})}{f(\theta\,|\overline{\eta})}\right)\beta f(\theta\,|\overline{\eta}) = -\overline{\kappa}(\underline{\theta}) - \int_{\underline{\theta}}^{\theta} \mu(\tau)\, d\tau$$

and

$$\left(V_x\left(\underline{x}(\theta), \underline{\eta}\right) - C_x\left(\underline{x}(\theta), \underline{\eta}, \theta\right) - (1-\alpha)\frac{F\left(\theta\,|\underline{\eta}\right)}{f\left(\theta\,|\underline{\eta}\right)}\right)(1-\beta) f\left(\theta\,|\underline{\eta}\right) = -\underline{\kappa}(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \mu(\tau)\, d\tau$$

Recall that $\beta f(\theta\,|\overline{\eta}) = f(\theta, \overline{\eta})$ and $(1-\beta) f\left(\theta\,|\underline{\eta}\right) = f\left(\theta, \underline{\eta}\right)$. Letting $g(\eta\,|\theta)$ denote the probability density function for η conditional on θ, I can substitute for $f(\theta, \eta) = g(\eta\,|\theta) f(\theta)$. Moreover, note that $\overline{x} = \underline{x}$ as $\mu > 0$ for $\theta \leq \theta'$. Adding the two conditions of optimality for the control variables, and dividing by $f(\theta)$ I get

$$\mathcal{E}_{H|\theta}\left[V_x\left(x^*\left(\theta\right),\eta\right)-C_x\left(x^*\left(\theta\right),\eta,\theta\right)|\,\Theta=\theta\right]-\left(1-\alpha\right)\frac{F\left(\theta\right)}{f\left(\theta\right)}=\frac{-\underline{\kappa}\left(\underline{\theta}\right)-\overline{\kappa}\left(\underline{\theta}\right)}{f\left(\theta\right)} \tag{40}$$

where I have used the fact that $\beta F\left(\theta\,|\overline{\eta}\right)+\left(1-\beta\right)F\left(\theta\,|\underline{\eta}\right)=F\left(\theta\right).$

To complete the argument I now argue that $-\underline{\kappa}\left(\underline{\theta}\right)-\overline{\kappa}\left(\underline{\theta}\right)=0$. Since $\overline{x}\left(\theta\right)=\underline{x}\left(\theta\right)=x^*\left(\theta\right)$ for $\theta\leq\theta'$, any solution of (40) for given $-\underline{\kappa}\left(\underline{\theta}\right)-\overline{\kappa}\left(\underline{\theta}\right)$ satisfies constraint (14). Moreover, $\underline{\kappa}\left(\underline{\theta}\right)$ and $\overline{\kappa}\left(\underline{\theta}\right)$ have no influence on the value of the objective for $\theta>\theta'$, because the costate variables are allowed to jump at points where the state variable constraint switches from binding to non-binding. Moreover, $\underline{\kappa}\left(\underline{\theta}\right)$ and $\overline{\kappa}\left(\underline{\theta}\right)$ have no influence on the location of the switching point θ' either. Hence, at the optimum $\underline{\kappa}\left(\underline{\theta}\right)$ and $\overline{\kappa}\left(\underline{\theta}\right)$ must be such that, conditional on θ, the expected value of the objective is maximized. Hence, $\overline{\kappa}\left(\underline{\theta}\right)=-\underline{\kappa}\left(\underline{\theta}\right)$, and I obtain the expression in the Proposition.

For $\theta>\theta'$, $\mu\left(\theta\right)=0$, so that $\overline{\kappa}\left(\theta\right)=\overline{k}$ and $\underline{\kappa}\left(\theta\right)=\underline{k}$ for $\theta>\theta'$. A priori it is neither clear how $\overline{k}$ relates to $\underline{k}$, nor is it clear how the values of the costate variables relate to $\overline{\kappa}\left(\theta'\right)$ and $\underline{\kappa}\left(\theta'\right)$. That is there may be jumps in the costate variables at θ'.

I first show that $\overline{k}+\underline{k}=0$. To see this, consider a candidate pair of schedules that give rise to a switch point θ'. Clearly, for the subinterval $\left[\theta',\overline{\theta}\right]$, constraint (14) is binding only at θ'. But then, choosing the optimal schedules $\overline{x}\left(\theta\right)$ and $\underline{x}\left(\theta\right)$ on the subinterval $\left[\theta',\overline{\theta}\right]$ is equivalent to problem P^5 with θ' replacing $\underline{\theta}$. Hence, $k\equiv\overline{k}=-\underline{k}$.

Finally, I show that the solution schedules are continuous at the switch point θ'.

I can write the value of the objective as

$$\Gamma\left(\theta'\right)=W^1\left(\theta'\right)+W^2\left(\theta'\right)-\beta\left(1-\alpha\right)\pi, \tag{41}$$

where

$$W^1\left(\theta'\right)\equiv\mathcal{E}_{\mathcal{H}}\left[\int_{\underline{\theta}}^{\theta'}B\left(x^*\left(\theta\right),\eta,\theta,\eta\right)f\left(\theta\,|\eta\right)d\theta\right]$$

and

$$W^2\left(\theta'\right)\equiv\beta\int_{\theta'}^{\overline{\theta}}B\left(\overline{x}\left(\theta\right),\overline{\eta},\theta,\overline{\eta}\right)f\left(\theta\,|\overline{\eta}\right)d\theta+\left(1-\beta\right)\int_{\theta'}^{\overline{\theta}}B\left(\underline{x}\left(\theta\right),\underline{\eta},\theta,\underline{\eta}\right)f\left(\theta\,|\underline{\eta}\right)d\theta$$
$$+k\left(\pi+\int_{\theta'}^{\overline{\theta}}\overline{x}\left(y\right)dy-\int_{\theta'}^{\overline{\theta}}\underline{x}\left(y\right)dy\right).$$

Clearly θ' must pass the following test: the value of the objective, $\Gamma\left(\theta'\right)$, should not increase through a small change in θ'. Invoking the envelope theorem, the effect of a marginal change in θ' is

$$\Gamma_{\theta'}\left(\theta'\right)=W^1_{\theta'}\left(\theta'\right)+W^2_{\theta'}\left(\theta'\right)$$

where

$$W^1_{\theta'}\left(\theta'\right)\equiv\mathcal{E}_{\mathcal{H}}\left[B\left(x^*\left(\theta'\right),\eta,\theta',\eta\right)f\left(\theta'\,|\eta\right)\right]$$

and $W^2_{\theta'}(\theta')$ is the negative of

$$W^2_{\theta'}(\theta') = -\beta B(\overline{x}(\theta'), \overline{\eta}, \theta', \overline{\eta}) f(\theta' | \overline{\eta}) - (1-\beta) B(\underline{x}(\theta'), \underline{\eta}, \theta', \underline{\eta}) f(\theta' | \underline{\eta}) \\ -k(\overline{x}(\theta') - \underline{x}(\theta')) \tag{42}$$

Clearly, at the optimum I must have $W^1_{\theta'}(\theta') + W^2_{\theta'}(\theta') = 0$, so the values of the objectives evaluated at the bound θ' must match. One solution is clearly reached when $\overline{x}(\theta') = \underline{x}(\theta') = x^*(\theta')$. I now show this solution is unique. To make the dependence of $\overline{x}(\theta')$ and $\underline{x}(\theta')$ on k explicit, I write these schedules as $\overline{x}(\theta'; k) = \underline{x}(\theta'; k)$, respectively. A total differentiation of the integral constraint

$$\pi + \int_{\theta'}^{\overline{\theta}} \overline{x}(y;k)\, dy - \int_{\theta'}^{\overline{\theta}} \underline{x}(y;k)\, dy = 0$$

delivers

$$\frac{dk}{d\theta'} = \frac{\overline{x}(\theta';k) - \underline{x}(\theta';k)}{\int_{\theta'}^{\overline{\theta}} (\overline{x}_k(y;k) - \underline{x}_k(y;k))\, dy}$$

Recall from (33) and (34) that in the proof of Proposition 5 that $\frac{d\overline{x}}{dk} > 0$ and $\frac{d\underline{x}}{dk} < 0$. Hence, the denominator of the expression for $\frac{dk}{d\theta'}$ is positive. Hence, I have $\frac{dk}{d\theta'} < 0$ for $\overline{x}(\theta';k) < \underline{x}(\theta';k)$ and $\frac{dk}{d\theta'} > 0$ for $\overline{x}(\theta';k) > \underline{x}(\theta';k)$. Thus k is minimized when θ' is such that $\overline{x}(\theta';k) = \underline{x}(\theta';k)$. For any other value of θ', I will have $\overline{x}(\theta';k) > \underline{x}(\theta';k)$. However, we know from Proposition 2 that the values

$$\hat{\overline{x}} \equiv \arg\max_x \left\{ \left(V(x,\overline{\eta}) - C(x,\overline{\eta},\theta')\right) f(\theta' | \overline{\eta}) - (1-\alpha) F(\theta' | \overline{\eta}) \right\}$$

and

$$\hat{\underline{x}} \equiv \arg\max_x \left\{ \left(V(x,\underline{\eta}) - C(x,\underline{\eta},\theta')\right) f(\theta' | \underline{\eta}) - (1-\alpha) F(\theta' | \underline{\eta}) \right\}$$

satisfy $\hat{\underline{x}} \geq \hat{\overline{x}}$. Hence, the sum of the terms in the first line in (42) decreases by an increase in k. Moreover, $-k(\overline{x}(\theta') - \underline{x}(\theta'))$ becomes negative. Hence, there can be no other solution.

ii) incentive compatibility of the schedules:

Consider first incentive compatibility in the η dimension:

Recalling the expressions for $\frac{d\overline{x}}{d\theta}$ and $\frac{d\underline{x}}{d\theta}$ from (35) and 36 I have for any θ, $\overline{x}(\theta) = \underline{x}(\theta) \Longrightarrow \frac{d\underline{x}}{d\theta} \geq \frac{d\overline{x}}{d\theta}$ if for $\overline{x}(\theta) = \underline{x}(\theta)$

$$\frac{1}{(V_{xx}(\overline{x}(\theta), \overline{\eta}) - C_{xx}(\overline{x}(\theta), \overline{\eta}, \theta))} \leq \frac{1}{(V_{xx}(\underline{x}(\theta), \underline{\eta}) - C_{xx}(\underline{x}(\theta), \underline{\eta}, \theta))},$$

and

$$\left((1-\alpha)\frac{\partial}{\partial\theta}\frac{F(\theta|\overline{\eta})}{f(\theta|\overline{\eta})}\right) + \frac{k}{\beta}\frac{f_\theta(\theta|\overline{\eta})}{f(\theta|\overline{\eta})^2} \geq \left((1-\alpha)\frac{\partial}{\partial\theta}\frac{F(\theta|\underline{\eta})}{f(\theta|\underline{\eta})}\right) - \frac{k}{1-\beta}\frac{f_\theta(\theta|\underline{\eta})}{f(\theta|\underline{\eta})^2}. \tag{43}$$

These conditions are precisely fulfilled under assumption iii.

Incentive compatibility of the schedules in the θ dimension uses the same arguments as the proof of Proposition 5. ■

Proof of Proposition 7. It is easy to see that the derivative of (41) with respect to π is still given by

$$\Gamma_\pi(\pi) = -\beta(1-\alpha) + k; \tag{44}$$

and the second derivative is still given by $\Gamma_{\pi\pi} = \frac{dk}{d\pi}$. Letting $\overline{x}(\theta;k)$ and $\underline{x}(\theta;k)$ denote the functions defined by (19) and (20), and using the fact that (14) is binding at $\theta'(k)$, I still have

$$\frac{dk}{d\pi} = \frac{1}{\int_{\theta'(k)}^{\overline{\theta}} \left(\underline{x}_k(y;k)\,dy - \overline{x}_k(y;k)\right) dy} < 0,$$

where I have used the fact that effects of k on $\theta'(k)$ exactly cancel out because $\underline{x}\left(\theta'(k);k\right) = \overline{x}\left(\theta'(k);k\right)$. Hence, the optimum features $\pi^* = 0$ if $k(\pi)|_\pi = 0 \leq \beta(1-\alpha)$ and $\pi^* > 0$ if $k(\pi)|_\pi = 0 > \beta(1-\alpha)$.

Due to the single-crossing condition I must have $\theta' = \overline{\theta}$ for $\pi = 0$. This follows directly from substituting $\theta'' = \overline{\theta}$ into condition (38) in the proof of Lemma 7. For $\pi = 0$, constraint (14) is binding at $\overline{\theta}$; by convexity of the bunching region, the constraint is binding for all θ.

In the remainder of this proof, I evaluate the derivative of the regulator's objective function around $\pi = 0$.

For $\theta = \overline{\theta}$, I can write (26) in explicit form as

$$\left\{ \begin{array}{c} V_x\left(x^*\left(\overline{\theta}\right),\overline{\eta}\right) - C_x\left(x^*\left(\overline{\theta}\right),\overline{\eta},\overline{\theta}\right)\beta f\left(\overline{\theta}\,|\overline{\eta}\right) \\ +V_x\left(x^*\left(\overline{\theta}\right),\underline{\eta}\right) - C_x\left(x^*\left(\overline{\theta}\right),\underline{\eta},\overline{\theta}\right)(1-\beta) f\left(\overline{\theta}\,|\underline{\eta}\right) \end{array} \right\} = (1-\alpha) \tag{45}$$

From (20), $k(0)$ satisfies

$$\left(V_x\left(x^*\left(\overline{\theta}\right),\underline{\eta}\right) - C_x\left(x^*\left(\overline{\theta}\right),\underline{\eta},\overline{\theta}\right)\right)(1-\beta) f\left(\overline{\theta}\,|\underline{\eta}\right) - (1-\alpha) = -\beta(1-\alpha) + k(0).$$

Substituting from (45), I have

$$\Gamma_\pi(0) = -\left(V_x\left(x^*\left(\overline{\theta}\right),\overline{\eta}\right) - C_x\left(x^*\left(\overline{\theta}\right),\overline{\eta},\overline{\theta}\right)\right)\beta f\left(\overline{\theta}\,|\overline{\eta}\right).$$

Hence, $\Gamma_\pi(0) > 0$ ($\Gamma_\pi(0) \leq 0$) and thus $\pi^* > 0$ ($\pi^* = 0$) if and only if $x^*\left(\overline{\theta}\right)$ is larger (smaller) than the first-best level of $x^{fb}\left(\overline{\theta},\overline{\eta}\right)$ that solves

$$V_x\left(x^{fb}\left(\overline{\theta},\overline{\eta}\right),\overline{\eta}\right) - C_x\left(x^{fb}\left(\overline{\theta},\overline{\eta}\right),\overline{\eta},\overline{\theta}\right) = 0.$$

In the sequel, I write $x^*\left(\overline{\theta}\right) = x^*\left(\overline{\theta};\beta\right)$ to make the dependence on β explicit. I now show that the conditions for $\pi^* > 0$ are satisfied for small enough β if and only if $x^*\left(\overline{\theta};\underline{\eta}\right)$, defined in (7) as the optimal production for type $\overline{\theta}$ when the quality capacity is known to be $\underline{\eta}$, satisfies

$$x^*\left(\overline{\theta};\underline{\eta}\right) > x^{fb}\left(\overline{\theta},\overline{\eta}\right).$$

Using (26), I can compute how $x^*\left(\overline{\theta};\beta\right)$ varies with β:

$$\frac{dx^*\left(\overline{\theta};\beta\right)}{d\beta} = \frac{-\left[V_x\left(x^*\left(\overline{\theta}\right),\overline{\eta}\right) - C_x\left(x^*\left(\overline{\theta}\right),\overline{\eta},\overline{\theta}\right)\right] f\left(\overline{\theta}\,|\overline{\eta}\right) + \left[V_x\left(x^*\left(\overline{\theta}\right),\underline{\eta}\right) - C_x\left(x^*\left(\overline{\theta}\right),\underline{\eta},\overline{\theta}\right)\right] f\left(\overline{\theta}\,|\underline{\eta}\right)}{B_{xx}\left(\overline{x}(\theta),\overline{\eta},\theta,\overline{\eta}\right)\beta f\left(\theta\,|\overline{\eta}\right) + B_{xx}\left(\underline{x}(\theta),\underline{\eta},\theta,\underline{\eta}\right)(1-\beta) f\left(\theta\,|\underline{\eta}\right)}.$$

The denominator of this expression is negative. I now show that the numerator is positive, so that the overall expression is negative.

By submodularity,

$$V_x\left(x^*\left(\overline{\theta}\right),\overline{\eta}\right) - C_x\left(x^*\left(\overline{\theta}\right),\overline{\eta},\overline{\theta}\right) < V_x\left(x^*\left(\overline{\theta}\right),\underline{\eta}\right) - C_x\left(x^*\left(\overline{\theta}\right),\underline{\eta},\overline{\theta}\right). \tag{46}$$

Moreover, from condition (20), I must have

$$V_x\left(x^*\left(\overline{\theta}\right),\underline{\eta}\right)-C_x\left(x^*\left(\overline{\theta}\right),\underline{\eta},\overline{\theta}\right)>0.$$

Multiplying (46) through by $f\left(\overline{\theta}\,|\overline{\eta}\right)$, I obtain

$$\left(V_x\left(x^*\left(\overline{\theta}\right),\overline{\eta}\right)-C_x\left(x^*\left(\overline{\theta}\right),\overline{\eta},\overline{\theta}\right)\right)f\left(\overline{\theta}\,|\overline{\eta}\right)<\left(V_x\left(x^*\left(\overline{\theta}\right),\underline{\eta}\right)-C_x\left(x^*\left(\overline{\theta}\right),\underline{\eta},\overline{\theta}\right)\right)f\left(\overline{\theta}\,|\overline{\eta}\right).$$

Multiplying and dividing the right-hand side by $f\left(\overline{\theta}\,|\underline{\eta}\right)$, I have

$$\left(V_x\left(x^*\left(\overline{\theta}\right),\overline{\eta}\right)-C_x\left(x^*\left(\overline{\theta}\right),\overline{\eta},\overline{\theta}\right)\right)f\left(\overline{\theta}\,|\overline{\eta}\right)<\left(V_x\left(x^*\left(\overline{\theta}\right),\underline{\eta}\right)-C_x\left(x^*\left(\overline{\theta}\right),\underline{\eta},\overline{\theta}\right)\right)\frac{f\left(\overline{\theta}\,|\overline{\eta}\right)}{f\left(\overline{\theta}\,|\underline{\eta}\right)}f\left(\overline{\theta}\,|\underline{\eta}\right)$$

I note that $\frac{f(\overline{\theta}|\overline{\eta})}{f(\overline{\theta}|\underline{\eta})}\geq 1$ by the monotonicity of the reversed hazard rate in η. Hence, $\frac{dx^*(\overline{\theta};\beta)}{d\beta}<0$, as claimed. Thus, for $\beta=0$, $x^*\left(\overline{\theta}\right)$ is largest and satisfies

$$V_x\left(x^*\left(\overline{\theta}\right),\underline{\eta}\right)-C_x\left(x^*\left(\overline{\theta}\right),\underline{\eta},\overline{\theta}\right)=\frac{(1-\alpha)}{f\left(\overline{\theta}\,|\underline{\eta}\right)},$$

which is exactly condition (26) that determines $x^*\left(\overline{\theta};\underline{\eta}\right)$. ■

Proof of Proposition 8. Totally differentiating (27) with respect to x, q and θ, and using $C(x,q,\theta)=K+c(x,\theta)+k(x,q)$, I get

$$\begin{aligned}
&\left(V_{xx}\left(x^*(\theta),q(\theta)\right)-C_{xx}\left(x^*(\theta),q(\theta),\theta\right)-(1-\alpha)\,c_{xx\theta}\left(x^*(\theta),\theta\right)\frac{F(\theta)}{f(\theta)}\right)dx^*\\
&+\left(V_{xq}\left(x^*(\theta),q(\theta)\right)-k_{xq}\left(x^*(\theta),q(\theta)\right)\right)dq^*\\
=&\left(c_{x\theta}\left(x^*(\theta),\theta\right)+(1-\alpha)\,c_{x\theta\theta}\left(x^*(\theta),\theta\right)\frac{F(\theta)}{f(\theta)}+(1-\alpha)\,c_{x\theta}\left(x^*(\theta),\theta\right)\frac{\partial}{\partial\theta}\frac{F(\theta)}{f(\theta)}\right)d\theta.
\end{aligned}$$

Totally differentiating (28) with respect to x and q, I get

$$\left(V_{qx}\left(x^*,q^*\right)-C_{qx}\left(x^*,q^*,\theta\right)\right)dx^*+\left(V_{qq}\left(x^*,q^*\right)-C_{qq}\left(x^*,q^*,\theta\right)\right)dq^*=0.$$

Rearranging, I have

$$dq^*=-\frac{\left(V_{qx}\left(x^*,q^*\right)-C_{qx}\left(x^*,q^*,\theta\right)\right)}{V_{qq}\left(x^*,q^*\right)-C_{qq}\left(x^*,q^*,\theta\right)}dx^*. \tag{47}$$

Substituting from (47) for dq, using Young's theorem, the additive separability of the cost function, and rearranging, I get

$$\frac{dx^*}{d\theta}=\frac{c_{x\theta}\left(x^*(\theta),\theta\right)+(1-\alpha)\,c_{x\theta\theta}\left(x^*(\theta),\theta\right)\frac{F(\theta)}{f(\theta)}+(1-\alpha)\,c_{x\theta}\left(x^*(\theta),\theta\right)\frac{\partial}{\partial\theta}\frac{F(\theta)}{f(\theta)}}{\left(V_{xx}\left(x^*,q^*\right)-C_{xx}\left(x^*,q^*,\theta\right)-(1-\alpha)\,c_{xx\theta}\left(x^*,\theta\right)\frac{F(\theta|\eta)}{f(\theta|\eta)}-\frac{\left(V_{qx}(x^*,q^*)-k_{qx}(x^*,q^*)\right)^2}{\left(V_{qq}(x^*,q^*)-k_{qq}(x^*,q^*)\right)}\right)}.$$

Hence, the solution is incentive compatible. ■

REFERENCES

Armstrong, M., "Multiproduct Nonlinear Pricing", *Econometrica* **64**, 1996, 51-76.

Armstrong, M. ,"Optimal Regulation with Unknown Demand and Cost Functions", *Journal of*

Economic Theory **99**, 1999, 196-215.

Armstrong, M. and J. Rochet ,"Multi-dimensional Screening: A User's Guide", *European Economic Review* **43**, 1999, 959-979.

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

Beaudry, P., C. Blackorby and D. Szalay, "Taxes and Employment Subsidies in Optimal Redistribution Programs ", *American Economic Review* **99**(1), 2009, 216-242.

Blackorby, C. and D. Szalay, "Regulating a Monopolist with unknown costs and unknown quality capacity", *The Warwick Economic Research Papers Series wp 858* , 2008

Che, Y. and I. Gale, "Standard Auctions with Financially Constrained Bidders", *Review of Economic Studies* **65**, 1998, 1-21.

Che, Y. and I. Gale, "The Optimal Mechanism for Selling to a Budget-Constrained Buyer", *Journal of Economic Theory* **92**, 2000, 198-233.

Fudenberg, D. and J. Tirole, *Game Theory*, MIT Press, Cambridge, MA., 1991.

Iyengar, G. and Kumar, A. "Optimal Procurement Mechanisms for Divisible Goods with Capacitated Suppliers" working paper Columbia University, 2006.

Jehiel, P., Moldovanu, B, and E. Stacchetti, "Multidimensional Mechanism Design for Auctions with Externalities", *Journal of Economic Theory* **85**, 1999, 258-293.

Kamien, M., and N. Schwartz, "Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management", *Elsevier Science Publishers B.V. North Holland*, 1981

Laffont, J., E. Maskin, and J. Rochet, "Optimal Nonlinear Pricing with Two-Dimensional Characteristics", in *Information, Incentives and Economic Mechanisms* edited by T. Groves, R. Radner, and S. Reiter, Minneapolis University of Minnesota Press, 1987, 256-266.

Lewis, T. and D. Sappington, "Regulating a Monopolist with Unknown Demand and Cost Function", *Rand Journal of Economics* **19**, 1988, 438-457.

McAfee, R.P. and J. MacMillan, "Multidimensional Incentive Compatibility and Mechanism Design", *Journal of Economic Theory* **46**, 1988, 335-354.

Malakow, A. and R. Vohra, "An Optimal Auction for Capacitated Bidders - a Network Perspective", mimeo University of North Carolina and Northwestern University, 2005a.

Malakow, A. and R. Vohra, "Optimal Auctions for Asymmetrically Budget Constrained Bidders", mimeo University of North Carolina and Northwestern University, 2005b.

Matthews, S. and Moore, J. H., Optimal Provision of Quality and Warranties: An Exploration in the Theory of Multidimensional Screening, *Econometrica*, **55**, 1997, 441-467.

Milgrom, P., and R. Weber, "A Theory of Auctions and Competitive Bidding", *Econometrica*, **50**, 1982, 1089-1122.

Mirrlees, J.A., An exploration of the Theory of Optimal Taxation, *The Review of Economic Studies*, **38**, 1971, 175-208.

Mirrlees, J.A., Information and Incentives: The Economics of Carrots and Sticks, *Economic Journal*, **107**, 1997, 1311-1329.

Rochet, J.-C. and Choné, P., Ironing, Sweeping, and Multidimensional Screening, *Econometrica*, **66**, 1998, 783-826.

Rochet, J.-C. and L. A. Stole, "The Economics of Multidimensional Screening" Chapter 5 in *Advances in Economics and Econometrics, Theory and Applications, Eighth World Congress, Vol. I* edited by M. Dewatripont, L. P. Hansen, and S. Turnovsky, Cambridge University Press, Cambridge, 2003, 150-197.

Seyerstad, K. and A. Sydsaeter "Optimal Control Theory with Economic Applications; Second reprint", *North Holland,* 1999

Shaked, M. and Shantikumar, J.G., "Stochastic Orders", *Springer, New York,* 2007

Topkis, D. M., "Minimizing a Submodular function on a Lattice, *Operations Research* **26**, 1978, 305-321

Topkis, D. M., Supermodularity and Complementarity, *Frontiers of Economic Research; Princeton University Press* 1998.

Vives, X., *Oligopoly Pricing: Old Ideas and New Tools* The MIT Press, Cambridge, Massachusetts, 1999.

Wilson, R., *Nonlinear Pricing* Oxford University Press, Oxford, 1993.