Multi-Dimensional Screening with a One-Dimensional Allocation Space.

Raymond Deneckere and Sergei Severinov *

First Version: June 26, 2004
This Version: March 9, 2009

Abstract

We develop a general method for solving multi-dimensional screening problems in which the 'physical' allocation space is one-dimensional, and provide necessary and sufficient conditions for the existence of 'exclusion' in the optimal mechanism. We illustrate the application of our method to an example with quadratic utility and uniformly distributed types. Interestingly, the optimal solution exhibits discontinuity along the boundary of the region between exclusion and non-exclusion for a large set of parameter values.

JEL Nos: C72, D82
Keywords: multi-dimensional screening, private information.

1 Introduction

This paper studies a screening problem in which the type space is multi-dimensional and the allocation space is one-dimensional. Such problems are common in economics, for two distinct reasons.

First, in many important economic environments agents typically differ along several dimensions on which there is private information. In the area of price discrimination, consumers differ both in demand intensity (intercept of demand) and price sensitivity (slope of demand). For example, high demand consumers can be price insensitive (because they are rich) or price sensitive (because they are poor and have large families). Similarly, an industrial customer's valuation for an input may depend both on the technology this firm uses to process the input, and the demand for the final product. Additionally, firms often have available multiple socioeconomic data that are imperfectly correlated with customers' purchase patterns. In other areas, multi-dimensionality of types is also prevalent. In insurance, customers differ both in risk aversion and the probability of having an accident. In labour taxation, the government may wish to differentially treat individuals who have low ability and those who have a high

*Department of Economics, University of Wisconsin, Madison, and Department of Economics, University of British Columbia and the University of Essex, respectively. We thank Ken Judd, Preston McAfee, and Bob Wilson for helpful comments and suggestions. Email: `rjdeneck@wisc.edu`

preference for leisure. And in the regulation of monopolies, the regulatory agency may wish to allow a different regulatory price and access charge for firms that have a high cost than for firms that have a low demand.

Secondly, in many of these screening environments, the principal cannot discriminate between agents along more than one dimension. In price discrimination, firms can often differentially treat customers only by purchase quantity. For non-durable consumption goods there may be no opportunity to differentiate by quality, so quantity becomes the only instrument. Examples include soft drinks (which come in various sizes), residential electricity, and public transportation. On the other hand, for many consumer durables customers only purchase one unit, so then the only available dimension for discrimination becomes quality. Frequently, there is only one (or at least one dominant) dimension of quality, such as the speed of a microprocessor or internet connection, or the number of megapixels in a camera. In auctions, there is often only one unit offered for sale, and the single dimension then becomes the probability of obtaining the object. In areas other than price discrimination, the allocation space if often also one-dimensional. In insurance markets, the allocation consists of the amount of coverage, in labour taxation the instrument is the tax rate, and in regulation it is the regulatory price.

Despite the existence of a rather voluminous literature on screening, relatively little is known about the type of problem we study. There are several reasons for this. First, as we will demonstrate, one of the dominant current approaches, the method of demand profiles (pioneered by Goldman, Leland and Sibley (1984), and further popularized by Brown and Sibley (1986) and, most forcefully, by Wilson (1993)) fails to adequately solve the problem. The difficulty with the demand profile method is that it requires that the derived marginal price schedule intersect a customer's demand schedule from below. In the one-dimensional type case, this is assured by the condition that marginal valuation is increasing in type ($u_{q\theta} > 0$), and that the assignment of quantities to types is nondecreasing (ensured by a monotonic inverse hazard rate, or by ironing). In the multi-dimensional case, no such sufficient condition is known. Furthermore, crossing from below is hard to ensure, because demand curves vary both in slope and intercept - sufficient variation in the intercept will thus necessarily lead to a violation of the required condition. As a consequence, the allocation will fail to be incentive compatible: the quantity assigned to customers whose demand curve intersects the tariff from above will correspond to a local minimum rather than a global maximum of their surplus maximization problem. Many of the worked out examples in the literature, such as the linear quadratic one studied in Wilson (1993, p. 196), therefore involve tariffs that are not incentive compatible.

A closely related approach is proposed in McAfee and McMillan (1988). These authors introduce a condition termed "Generalized Single Crossing" which ensures that any solution satisfying the first and second order conditions of the agent's surplus maximization problem is globally incentive compatible. Generalized Single Crossing implies that iso-price curves are linear in the type space, thereby permitting a reduction to a one-dimensional screening problem. McAfee and McMillan's contribution is considerable, but suffers from a number of drawbacks. First, the limitation to *linear* iso-price curves is significant in our context. Second, their approach implicitly assumes that in equilibrium all agent types along an iso-price line will participate. Unfortunately, as our analysis will reveal, this assumption is often violated.[1]

Lewis and Sappington (1988) adopt the Generalized Single Crossing assumption, but instead of formulating the problem in terms of demand profiles use the direct method pioneered by Mussa and Rosen (1978), leading to an objective based on virtual utility functions. Because it is based upon McAfee and McMillan's method for reducing the problem to a one-dimensional screening problem, this approach suffers from the same drawbacks. In addition, Lewis and Sappington's analysis assumes that in equilibrium there is no exclusion. They do not provide conditions for exclusion to be absent, and unfortunately, as we will show, exclusion is rather prevalent. In particular, in the context of nonlinear pricing, absence of exclusion requires the aggregate demand curve to be perfectly inelastic at the seller's marginal cost of production.[2] Finally, Lewis and Sappington implicitly assume that the slope of iso-price curves in type space is constant, which (in the nonlinear pricing interpretation) can happen only if the slope of the agent's demand function is independent of type.[3]

Rochet and Stole (2001) develop the direct method for arbitrary multi-dimensional screening problems. Their approach has two drawbacks. First, because the problem is not reduced to a one-dimensional screening problem, the associated first-order conditions require the solution of a partial differential equation, which cannot be solved analytically, except in very special cases. Secondly, because the direct approach only imposes the local incentive compatibility constraints, the solution typically violates the conditions for global incentive compatibility. A general method for solving our type of problem therefore remains lacking.

Lastly, a solution method for our problem has recently become available for the special case where the agent's utility function is linear in type. This was made possible by two breakthroughs in the analysis of multi-dimensional screening problems. First, Rochet and Choné (1998) developed a "sweeping" procedure (analogous to ironing for the one-dimensional case), which adjusts the solution derived by the direct method so as to ensure global incentive compatibility. Rochet and Choné's method requires that the dimension of the type space and allocation space coincide. However, by interpreting the coefficients on consumer types as artificial goods in the utility function, Basov (2001) was able to transform the problem from one where the number of consumer characteristics exceeds the dimension of the physical allocation space to one where the two dimensions coincide. While ingenious, this approach also has several drawbacks. It requires agents' utility functions to be linear in type, which is great for applications such as auctions, but quite limiting in the current context. The method also necessitates the solution of a partial differential equation, which generally can be solved only numerically. Finally, sweeping is a complicated procedure which does lend not itself to analytical solutions.

It is fair to conclude that because of all these issues, our type of screening problem has hitherto remained inaccessible to most economists, and therefore failed to generate interesting

[1] Properly taking into account the agent's participation constraint changes the integrand of principal's objective function in an essential way: rather than depending only on the allocation $q(t_1)$ and its derivative $q'(t_1)$ it now also depends on $q(t_1')$ for all $t_1' > t_1$. As a consequence, McAfee and McMillan's formulation of the problem can no longer be solved by the method of calculus of variation.

[2] Armstrong (1999) already pointed out this deficiency, but was unable to either provide neccessary and sufficient conditions for exclusion to occur, or solve the associated multi-dimensional screening problem.

[3] As a consequence, Lewis and Sappington's characterization of an optimal mechanism (Proposition 1, p. 447) is generally incorrect. However, it does hold for the special example studied in Section 5 of their paper.

practical applications.

Our paper contains several methodological contributions. First, by correctly characterizing the isoquants, the set of agent types that consume the same quantity, we are able to reduce the multi-dimensional screening problem to a one-dimensional optimal control problem, whose solution is governed by an ordinary differential equation. Our solution method is therefore accessible to most economists, and generates analytical solutions. Second, we formulate the multi-dimensional screening problem as one of assigning agent types and tariff to the one-dimensional allocation. This approach is not only natural here, underscoring the one-dimensional nature of the principal's optimization problem, but also avoids some of the difficulties associated with discontinuities in the quantity allocation as a function of types that typically arise in our problem (see the discussion in the next paragraph). Our method also handles bunching in a straightforward and transparent way, without any need to resort to "ironing" or "sweeping". Finally, we present a novel condition, termed Single Crossing of Demand (SCD), which ensures that the solution to the principal's relaxed problem is globally incentive compatible.

The solution to our multi-dimensional screening problem exhibits several interesting properties. First, it may or may not be optimal to exclude some consumer types from consumption. The result that it can be optimal to have full consumer participation contrasts with established wisdom, which holds that when the type space is multi-dimensional, exclusion is generic (Armstrong (1996), Basov (2005)). Second the optimal quantity allocation is discontinuous at the boundary between the region of exclusion (where the optimal quantity is zero) and the region of non-exclusion (where the optimal quantity is generally bounded away from zero). Finally, and perhaps most surprisingly, we find that there can be a bunching of quantities allocated to a type located on the boundary between exclusion and non-exclusion, i.e. there can be a discontinuity of quantity as a function of type. The consumer type on which the quantities are bunched is then indifferent between all quantities in the bunch.

The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 introduces the SCD condition, and characterizes the associated implementable allocations. Section 4 uses this characterization to reformulate the principals' problem to a one-dimensional screening problem in which the density of types is endogenous. Section 5 solves the associated optimal control problem, and presents necessary and sufficient conditions for exclusion to occur. Section 6 studies an example with linear quadratic utility and uniformly distributed types. Section 7 contains the conclusion.

2 The Model

A monopolist supplier of a single good faces a population of consumers. Consumers are distinguished by a two dimensional preference parameter $t = (\alpha, \theta)$, which is private information. The limitation to two dimensions of uncertainty is made here for ease of exposition and compactness in notation. With minor modifications, our results generalize to higher dimensions (for details, see Deneckere and Severinov (2009a). When consuming a quantity $q \in \mathbf{R}_+$ of the good, acquired at cost p, a consumer of type t receives net utility $u(q, t) - p$. Consumers' reservation utilities are equal to zero.

The distribution function $F(t)$ of consumer types in the population is common knowledge. We assume that $F(.)$ is twice continuously differentiable function, with density function $f(\theta) > 0$, and a rectangular support $[a, b] \times [c, d]$. Renormalizing, we can without loss of generality take the support to be $[0, 1] \times [0, 1]$.

We assume that the firm's marginal and average cost of production is constant at the level $c > 0$. To handle the case in which the monopolist's aggregate cost $C(Q)$ is an increasing function of aggregate output $Q = \int q(t)f(t)dt$ we would need one extra step. Precisely, for any given constant marginal cost level c, our model would predict the corresponding aggregate output level Q selected by the firm. Equilibrium then obtains whenever $C'(Q) = c$.

We maintain the following assumptions on preferences throughout the paper:

Assumption 1 *The function $u(q, \alpha, \theta) : \mathbf{R}_+ \times [0, 1]^2$ is of class C^3. Furthermore,*
(i) $u(0, \alpha, \theta) = 0$ for all $(\theta, \alpha) \in [0, 1]^2$;
(ii) $u_q(q, \alpha, \theta) > 0$, $u_\theta(q, \alpha, \theta) > 0$ and $u_\alpha(q, \alpha, \theta) > 0$, for all $q > 0$ and $(\alpha, \theta) \in [0, 1]^2$;
(iii) $u_{\theta q}(q, \alpha, \theta) > 0$, $u_{\alpha q}(q, \alpha, \theta) > 0$, for all $(\alpha, \theta) \in (0, 1]^2$ and $q > 0$;
(iv) $u_{qq}(q, \alpha, \theta) < 0$ for all θ, α and q.

Assumption 1 is fairly standard. Part (iii) requires consumer's utility functions to be supermodular. Part (iv) ensures that consumers' demand functions are downward sloping.

We also make extensive use of a novel assumption, specific to the higher-dimensional type space, which we term "Single-Crossing of Demand":

Assumption 2 *(SCD) $\frac{d}{dq}\frac{u_{q\alpha}}{u_{q\theta}} > 0$ for all $q > 0$.*

The economic interpretation of Assumption 2 is that the inverse demand functions can intersect at most once, as the next lemma demonstrates.

Lemma 1 *Suppose Assumption 2 holds and $\alpha' > \alpha$. Then $u_q(q, \alpha', \theta') = u_q(q, \alpha, \theta)$ implies $u_{qq}(q, \alpha', \theta') > u_{qq}(q, \alpha, \theta)$.*

Assumption 2 should not be confused with the single-crossing condition in one-dimensional screening problems, which guarantees that consumers' indifference curves in (q, t) space intersect at most once. In fact, the latter condition is extremely restrictive, as it implies that consumers' demand curves do not intersect at all, i.e. can be ranked. In the next section, we will show that Assumption 2 has many important consequences: it implies that isoquants in (θ, α) space cannot intersect, and must "fan out".

The monopolist's problem is to select a nonlinear pricing schedule, i.e. a lower semi-continuous function $P : \mathbf{R}_+ \to \mathbf{R}_+$, so as to maximize her profits subject to consumer maximization.[4] Consumer maximization requires that if a consumer of type t consumes the quantity q, then the following incentive compatibility and individual rationality constraints must be satisfied:

$$\begin{aligned} u(q, t) - P(q) &\geq u(q', t) - P(q'), \text{ for all } q' &(1)\\ u(q, t) - P(q) &\geq 0 &(2)\end{aligned}$$

[4]The Taxation principle implies that the monopolist cannot gain by offering more complicated screening mechanisms. Below, we show that the schedule $T(\cdot)$ may be taken to be continuous, so the assumption of l.s.c. is without loss of generality.

Let $T^*(q)$ denote the set of all types t for which (1) and (2) hold. Also let $T^{**}(q)$ be the set of all types who consume q or more when facing the pricing schedule P, i.e. $T^{**}(q) = \cup_{q' \geq q} T^*(q')$. Finally, let

$$\omega(q) = 1 - \int_{T^{**}(q)} f(t)dt \tag{3}$$

denote the proportion of types who consume q or less when faced with the pricing schedule P. Then the monopolist's problem is to solve

$$\max_{P(\cdot)} \int_0^{\infty} (P(q) - c) d\omega(T^{**}(q)) \tag{4}$$

This formulation of the monopolist's problem underscores its fundamental one-dimensional nature, but comes at the cost of characterizing the potentially quite complicated structure of the set $T^{**}(q)$. Our next section shows that Assumption 2 lends a relatively simple structure to this set.

3 Characterization of Isoquants

Let us start with an example which demonstrates the difficulties that must be overcome in characterizing the set $T^{**}(q)$. This example also illustrates where the previous literature has gone wrong.

Example 1: Let $u(q, \alpha, \theta) = \theta q - \frac{b-\alpha}{2} q^2$, with $b < \frac{3}{2}$. Also let $c = 0$, and let F be the uniform distribution on the unit square. Following Wilson (1993) define the demand profile $N(p, q)$ as the fraction of consumers in the population whose demand price u_q exceeds p. A simple calculation yields:

$$N(p,q) = \begin{cases} \frac{1}{2q}\{(1-p-(b-1)q)^2 - (1-p-bq)^2\}, & \text{if } p + bq \leq 1 \\ \frac{1}{2q}(1-p-(b-1)q)^2, & \text{if } p + bq \geq 1. \end{cases}$$

According to the demand profile approach, $N(p, q)$ represents the demand schedule for quantity increment q. Thus for each increment q, the monopolist should select the price for increment q, $p(q) = P'(q)$ to solve

$$\max_p \{(p - c)N(p, q)$$

Performing this maximization gives

$$p(q) = \begin{cases} \frac{1}{2} - \frac{1}{4}(2b-1)q, & \text{if } q \leq \frac{2}{2b+1} \\ \frac{1}{3}(1-(b-1)q), & \text{if } q \geq \frac{2}{2b+1}. \end{cases} .$$

resulting in the rtariff $P(q) = \int_0^q p(z)dz$

$$P(q) = \begin{cases} \frac{1}{2}q + \left(\frac{1}{8} - \frac{b}{4}\right) q^2, & \text{if } q \leq \frac{2}{2b+1} \\ \frac{1}{6(2b+1)} + \frac{q}{3} - \frac{b-1}{6} q^2, & \text{if } q \geq \frac{2}{2b+1}. \end{cases} .$$

For this approach to be correct, every consumer type whose demand price equals $p(q)$ should also be willing to purchase all increments $q' < q$ and not purchase any increments $q' > q$. This will be the case if the iso-price curves in type space, defined by the equation $u_q(q,t) = p(q)$, do not intersect, for then every consumer type t will have only one solution to the first order condition associated with her surplus maximization problem $\max_q\{u(q,t) - P(q)\}$.[5]

Let us therefore examine the iso-price curves associated with the schedule P. Solving the equation $\theta - (b-\alpha)q = p(q)$ yields

$$\theta(q,\alpha) = \begin{cases} \frac{1}{2} + \frac{1}{4}(2b+1-4\alpha)q, & \text{if } q \leq \frac{2}{2b+1} \\ \frac{1}{3} + \frac{1}{3}(2b+1-3\alpha)q, & \text{if } q \geq \frac{2}{2b+1}. \end{cases}$$

Figure 1 illustrates these iso-price curves. All iso-price curves are straight lines. For $q \in [0, \frac{2}{2b+1}]$, iso-price lines go through the point $(\alpha,\theta) = (\frac{2b+1}{4}, \frac{1}{2})$, rotating up form a flat line at the level $q = 0$ to the the quantity $q = \frac{2}{2b+1}$, where the northwest corner point $(\alpha,\theta) = (0,1)$ is reached. For $q \geq \frac{2}{2b+1}$, all iso-price lines rotate up through the point $(\alpha,\theta) = (\frac{2b+1}{3}, \frac{1}{3})$, until the quantity $q = \frac{1}{b-1}$ is reached, when the north-east corner point $(\alpha,\theta) = (1,1)$ is hit. This means that any point (α,θ) in the interior of triangle the Δ defined by the inequalities $\theta \leq 1/2$, $\alpha \geq \frac{2b+1}{4}$, and $\theta \geq \frac{1+2b-2\alpha}{1+2b}$ is the intersection point of an iso-price line from the region $q < \frac{2}{2b+1}$ and an iso-price line from the region $q > \frac{2}{2b+1}$. The objective function of such a type therefore has two stationary points, one at a quantity $q_-(\alpha,\theta) < \frac{2}{2b+1}$ and one at a quantity $q\_(\alpha,\theta) > \frac{2}{2b+1}$. It is easy to see that q_- corresponds to a local minimum, and q_+ to a local maximum.

The presence of a local minimum to the consumer's objective function has two immediate consequences. First, the demand profile approach, in which consumers are presented with marginal price schedules $p(q)$, is no longer equivalent to the original approach, where consumers are presented with a nonlinear tariff $P(q)$. Indeed, any consumer in the above mentioned triangle would be unwilling to purchase any quantity increment in the interval $[0, q_-]$, whereas they might purchase this increment when presented with the nonlinear pricing schedule P. Secondly, and more damagingly, the quantity q_+ may no longer be a global maximum to the consumer's optimization problem. Since the only other candidate for an optimum occurs at $q = 0$, this raises the important issue of whether all consumer types who are purchasing increment q_+ under the marginal schedule $p(q)$ would be willing to participate in the mechanism. As indicated above, this is not an issue for consumer types with $\theta \geq \frac{1}{2}$, since iso-price lines do not cross for such types. For consumers types in the triangle Δ, only $q > \frac{2}{2b+1}$ can be a maximum, and for such q we have

$$u(q,\alpha,\theta(q,\alpha)) - P(q) = \frac{1}{6}\left((1+2b)q^2 - \frac{1}{1+2b}\right) - \frac{\alpha}{2}q^2$$

[5] More formally, consider any type t on the iso-price curve at the quantity q, i.e. $u_q(q,t) - p(q) = 0$. Since iso-price lines do not cross, iso-price curves at quantities $q' > q$ will lie to the northeast of the iso-price curve at quantity q, and iso-price curves at quantities $q' < q$ will lie to the southwest of the iso-price curve at quantity q. It then follows from assumption 1(iii) that $u_q(q',t) - p(q') > 0$ for $q' < q$, and $u_q(q',t) - p(q') > 0$ for $q' > q$. Consequently, type t's objective function is strictly quasiconcave, implying that q is a global maximum.

Setting this expression equal to zero yields

$$\begin{aligned}\underline{\alpha}(q) &= \frac{1+2b}{3} - \frac{1}{3(1+2b)q^2} \\ \underline{\theta}(q) &= \frac{1+(1+2b)q}{3(1+2b)q}\end{aligned}$$

These equations trace out a strictly decreasing a curve in type space, which may equivalently be expressed as

$$\underline{\theta}(\alpha) = \frac{1}{3} + \frac{\sqrt{(1+2b)(1+2b-3\alpha)}}{3(2b+1)}$$

Note that the participation constraint is violated for all types in Δ that lie below the curve $\underline{\theta}$. As a consequence, the demand profile approach necessarily fails whenever when $b < \frac{3}{2}$.

The crossing of iso-price lines also implies that the methods of McAfee and McMillan and Lewis and Sappington are flawed. Indeed, since a consumer type can lie on two distinct iso-price lines $u_q(q,t) - p(q) = 0$, merely being located on an iso-price line generally cannot identify the quantity purchased by a consumer.

There are several approaches to resolving these difficulties. One could try to identify conditions under which iso-price curves never cross. This is a useful approach, and we pursue it elsewhere (Deneckere and Severinov, 2009b). The main drawback of this approach is that it fails to solve some of the most rudimentary examples, such as the one presented above. For this reason, the present paper concentrates on the more difficult question of solving the multi-dimensional screening problem when iso-price curves are allowed to intersect.

Since consumer participation is problematic under these conditions, let us start by ensuring that the non-exclusion region takes on a simple form. The main ingredient that will make this possible is the supermodularity assumption 1(iii). Let s denote the consumer's equilibrium surplus under tariff $P(\cdot)$, and let Q^* the associated argmax correspondence:

$$\begin{aligned}s(t) &= \max_q \{u(q,t) - P(q)\} \\ Q^*(t) &= \arg\max_q \{u(q,t) - P(q)\}\end{aligned}$$

Then we have:

Lemma 2 *Suppose Assumption 1 holds. Then s is an absolutely continuous function, and the envelope condition $\nabla s(t) = \nabla_t u(q,t)$ holds for a.e. t. Furthermore, $Q^*(t)$ is a non-empty closed-valued u.h.c. correspondence, and every selection of $Q^*(t)$ is an increasing function.*[6]

Let $q(t) = \max Q^*(t)$. An immediate consequence of Lemma 2 is that if type t is willing to participate, i.e. $q(t) > 0$, then any type t' to the northeast of t is also willing to participate, i.e. has $s(t') > 0$. As a consequence, the boundary between the region of participation and non-participation,

$$\underline{\theta}(\alpha) = \inf\{\theta | s(\alpha, \theta) > 0\}$$

[6] More precisely, if q' is a selection from Q^* then $q'(\theta', \alpha') \geq q(\theta, \alpha)$ whenever $\theta' \geq \theta$ and $\alpha' \geq \alpha$.

is well defined, and is downward sloping:[7],[8]

Lemma 3 *The function $\underline{\theta}(\alpha)$ is continuous, non-increasing in α, and strictly decreasing in α whenever $q(\alpha, \underline{\theta}(\alpha)) > 0$ and $\underline{\theta}(\alpha) > 0$.*

Lemma 3 and 2 jointly imply that if $q(\alpha, \underline{\theta}(\alpha)) = 0$, and $\alpha > 0$ then it is possible to have $\underline{\theta}(\alpha') = \underline{\theta}(\alpha')$ for all $\alpha' < \alpha$, i.e. the lower boundary may have a flat segment $[0, \alpha]$. Note that at the right endpoint of such a flat initial segment, there can be an upward jump in $q(\alpha, \underline{\theta}(\alpha))$, as is illustrated in Example 1. Flat sections can also be present because $\underline{\theta}$ hits the lower or upper boundaries of the type space.

Let us now turn to the structure of isoquants associated with quantities $q > 0$. The two main ingredients that allow us to characterize these isoquants are the supermodularity of the agent's payoff function and SCD.

Lemma 4 *Suppose Assumption 2 holds, and $q_1 \in Q^*(\theta_1, \alpha_1)$. Then for all (θ_2, α_2) s.t. $\alpha_1 > \alpha_2$ and $u_q(q_1, \theta_2, \alpha_2) = u_q(q_1, \theta_1, \alpha_1)$, one has $q_1 \in Q^*(\theta_2, \alpha_2)$.*

Lemma 4 says that for any point (α, θ) in the participation region, the portion of the iso-price curve at the quantity $q(\alpha, \theta)$ that lies to the northwest of the point (α, θ) belongs to the set the set $T^*(q(\alpha, \theta))$. Accordingly, for any point (α, θ) in the participation region, define $I(\alpha, \theta)$ to be portion of the iso-price curve through the point (α, θ) at the quantity $q(\alpha, \theta)$ for which $q(\alpha, \theta)$ is an optimal choice, i.e.

$$I(\alpha, \theta) = \{(\alpha', \theta') : u_q(q(\alpha, \theta), \alpha', \theta') = u_q(q(\alpha, \theta), \alpha, \theta) \text{ and } q(\alpha, \theta) \in Q^*(\alpha', \theta')\}$$

We will refer to the set $I(\alpha, \theta)$ as an isoquant at the quantity $q(\alpha, \theta)$.[9] Our next lemma shows that isoquants associated with quantities $q < q(\alpha, \theta)$ cannot intersect the isoquant $I(\alpha, \theta)$ at any point other than (α, θ):

Lemma 5 *Suppose Assumption 2 and 1(iii) hold. Let (α, θ) be any point in the participation region. Then $q \in Q^*(\alpha, \theta')$ and $q < q(\alpha, \theta)$ imply $\theta' \leq \theta$. Furthermore, if $\theta' = \theta$, then neither $I(\alpha, \theta)$ nor $I(q, \alpha, \theta)$ contain a point (α'', θ'') with $\alpha'' > \alpha$.*

In addition, except when $q(\theta, \alpha) = 0$, the isoquant $I(\alpha, \theta)$ cannot intersect the lower boundary at any point to the northwest of the point (θ, α):

Lemma 6 *Suppose that Assumption 2 holds. Let $q(\theta, \alpha) > 0$. Then $I(\theta, \alpha)$ does not intersect the boundary $\underline{\theta}(\cdot)$ at any point to the northwest of (θ, α).*

[7] We adopt the convention that the infimum of the empty set equals 1.

[8] It follows from Lemma 2 that $q(\alpha, \theta) = 0$ for all $\theta < \underline{\theta}(\alpha)$ and that $q(\alpha, \theta) > 0$ for all $\theta \geq \underline{\theta}(\alpha)$. Thus we may equivalently define $\underline{\theta}(\alpha) = \inf\{\theta | q(\alpha, \theta) > 0\}$.

[9] For (α, θ) where $Q^*(\alpha, \theta)$ is multi-valued, which is a set of measure zero in type space, there may be isoquants emanating from (α, θ) at quantities $q \in Q^*(\alpha, \theta)$ s.t. $q < q(\alpha, \theta)$. We will denote those isoquants by $I(q, \alpha, \theta)$.

In general, this is as far as a characterization of isoquants can go. Indeed, isoquants may emanate from the interior of the participation region, and may touch in the interior of that region, as the next example demonstrates:

Example 2: Let $u(q,\alpha,\theta) = \theta q - (1-\frac{\alpha}{2})q^2$. Consider the following tariff:

$$P(q) = \begin{cases} \frac{1}{9} & \text{if } q \leq \frac{1}{3} \\ \frac{2}{27} & \text{if } \frac{1}{3} < q \leq \frac{2}{3} \\ \infty & \text{if } q > \frac{2}{3} \end{cases}$$

Then the lower boundary consists of two parts:

$$\underline{\theta}(\alpha) = \begin{cases} \frac{19}{36} - \frac{\alpha}{18} & \text{if } \alpha \leq \frac{1}{2} \\ \frac{5}{9} - \frac{\alpha}{9} & \text{if } \alpha \geq \frac{1}{2} \end{cases}$$

Consumers types along the curve

$$\theta_{12}(\alpha) = \frac{7}{12} - \frac{\alpha}{6},\ \alpha \leq \frac{1}{2}$$

are indifferent between consuming $q_1 = \frac{1}{9}$ and $q_2 = \frac{2}{9}$. All isoquants associated with the quantity $q_1 = \frac{1}{9}$ emanating from the portion of the lower boundary with $\alpha \leq \frac{1}{2}$ end up along the curve $\alpha = 0$, and so do all isoquants associated with the quantity $q_2 = \frac{2}{9}$ emanating from the portion of the lower boundary with $\alpha \geq \frac{1}{2}$. All remaining isoquants at q_1 and q_2 emanate from the curve θ_{12}.

To simplify and complete our description of isoquants, we will therefore make one additional assumption:

Assumption 3 *(Continuity) The allocation $q(\alpha,\theta)$ is continuous on the interior of the participation region.*

Our justification for making this assumption is the following:

Lemma 7 *Suppose the tariff $P(\cdot)$ solves problem (4). Then the associated allocation satisfies Assumption 3.*

Under Assumption 3 no isoquants can emanate from the interior of the participation region. As a consequence, isoquants are entirely determined by the behavior of the allocation $q(\alpha,\theta)$ along the curve

$$L = \{(\alpha, \underline{\theta}(\alpha)) : 0 \leq \alpha \leq 1\} \cup \{(1,\theta) : \theta \geq \underline{\theta}(1)\},$$

which traces out the lower boundary and the right boundary of the participation region.

Theorem 1 *(Necessity) Suppose Assumptions 1, 2 and 3 hold, and suppose that* $\sup \frac{u_\alpha}{u_\theta}(q,\alpha,\theta) < \infty$. *Then for any tariff $P(\cdot)$ there exists $(\widehat{\alpha},\widehat{\theta})$ with either $\widehat{\alpha} = 0$ or $\widehat{\theta} = 1$, and two non-decreasing functions $\underline{q}(\cdot) : [0,1] \to \mathbf{R}_+$ and $\bar{q}(\cdot) : [\widehat{\theta},1] \to \mathbf{R}_+$ with $\underline{q}(1) = \bar{q}(\widehat{\theta})$, such that* *(i) The lower boundary $\underline{\theta}(\alpha)$ is an absolutely continuous function, satisfying $\underline{\theta}(\alpha) > 0$ for all $\alpha \leq \widehat{\alpha}$ and $\underline{\theta}(\alpha) = 0$ for all $\alpha \in [\widehat{\alpha}, 1)$. Furthermore, for almost every α s.t. $0 < \underline{\theta}(\alpha) < 1$ we have*

$$\frac{d\underline{\theta}}{d\alpha} = -\frac{u_\alpha}{u_\theta}(q(\alpha, \underline{\theta}(\alpha)), \alpha, \underline{\theta}(\alpha)) \tag{5}$$

(ii) $\underline{q}(\alpha) = \min Q^*(\alpha, \underline{\theta}(\alpha))$ for every $\alpha \in [0,1]$, and $\bar{q}(\theta) = \min Q^*(1, \theta)$ for every $\theta \in [\widehat{\theta}, 1]$.

(iii) $P(q)$ is an absolutely continuous function, with

$$P(q) = \begin{cases} u(q, \alpha, \underline{\theta}(\alpha)) & \text{if } q \in [\underline{q}(0), \underline{q}(\widehat{\alpha})] \\ u(\underline{q}(\widehat{\alpha}), \widehat{\alpha}, \underline{\theta}(\widehat{\alpha})) + \int_{\underline{q}(\widehat{\alpha})}^{q} u_q(z, \underline{\theta}(\alpha(z)), \alpha(z)) dz & \text{if } q \in [\underline{q}(\widehat{\alpha}), \underline{q}(1)] \\ P(\bar{q}(\widehat{\theta})) + \int_{\bar{q}(\widehat{\theta})}^{q} u_q(z, 1, \theta(z)) dz & \text{if } q \in [\underline{q}(1), \bar{q}(1)] \end{cases} \tag{6}$$

If, in addition Assumption 3 holds, then:

(iv) $Q^*(\alpha, \underline{\theta}(\alpha))$ is convex-valued for every $\alpha \in [0,1]$, and $Q^*(1, \theta)$ is convex valued for every $\theta \in [\widehat{\theta}, 1]$.

(v) $I(\alpha, \theta) \cap L$ is a singleton, for all (α, θ) such that $s(\alpha, \theta) > 0$.

Part (i) of Theorem 1 links the lower boundary to the allocation along the lower boundary. Part (ii) shows that the allocation is nondecreasing along the boundary L. Part (iii) links the tariff $P(\cdot)$ to the allocation along along the boundary L. Part (iv) says that any point $(\alpha, \theta) \in L$ where the allocation jumps up, type (α, θ) is indifferent between all quantities in the interval $[\min Q^*(\alpha, \theta), \max Q^*(\alpha, \theta)]$. Part (v) asserts that every isoquant in the participation region must emanate from a unique point in L. The final result of this section provides a converse to Theorem 1:

Theorem 2 *(Sufficiency) Suppose Assumptions 1, 2 hold, and $\sup \frac{u_\alpha}{u_\theta}(q, \alpha, \theta) < \infty$. Suppose we are given $(\widehat{\alpha}, \widehat{\theta})$ with either $\widehat{\alpha} = 0$ or $\widehat{\theta} = 1$, and two non-decreasing uppersemicontinuous functions $\underline{q}(\cdot) : [0,1] \to \mathbf{R}_+$ and $\bar{q}(\cdot) : [\widehat{\theta}, 1] \to \mathbf{R}_+$ with $\underline{q}(1) = \bar{q}(\widehat{\theta})$. Let*

(i) $\underline{\theta}(\alpha) = \min\{\widehat{\theta} + \int_{\alpha}^{\widehat{\alpha}} \frac{u_\alpha}{u_\theta}(\underline{q}(a), a, \underline{\theta}(a)) da\}$ for $\alpha \leq \widehat{\alpha}$, and $\underline{\theta}(\alpha) = 0$ for $\alpha > \widehat{\alpha}$

(ii) $q(\alpha, \theta) = q$ for any $q > 0$ s.t. $u_q(q, \alpha, \theta) = u_q(q, \alpha(q), \theta(q))$, where

$$(\alpha(q), \theta(q)) = \begin{cases} (\alpha, \underline{\theta}(\alpha)) & \text{if } \alpha = \sup\{a : \underline{q}(a) < q\}, \text{ and } q \leq \underline{q}(1) \\ (1, \theta) & \text{if } \theta = \sup\{\tau : \bar{q}(\tau) < q\}, \text{ and } q > \underline{q}(1) \end{cases} \tag{7}$$

and $q = 0$, otherwise

(iii)

$$P(q) = \begin{cases} u(q, \alpha(q), \theta(q)) & \text{if } q \leq \underline{q}(\widehat{\alpha}) \\ u(\underline{q}(\widehat{\alpha}), \widehat{\alpha}, \underline{\theta}(\widehat{\alpha})) + \int_{\underline{q}(\widehat{\alpha})}^{q} u_q(z, \alpha(z), \theta(z)) dz & \text{if } \underline{q}(\widehat{\alpha}) < q \leq \bar{q}(1) \\ \infty & \text{if } q > \bar{q}(1) \end{cases}$$

Then the allocation $q(\cdot)$ is incentive compatible for the tariff $P(\cdot)$, and satisfies Assumption 3. Furthermore, $\underline{\theta}$ is the lower boundary associated with the tariff $P(\cdot)$.

Theorems 1 and 2 imply that the set $T^{**}(q)$ takes on a very simple form: it is the region of type space bounded by the upper envelope of the lowest isoquant associated with the quantity q (emanating from the point $(\alpha(q), \theta(q))$ and the lower boundary $\underline{\theta}$.

Before proceeding with the solution of problem (4) we need to address one more technical issue. Note that flat sections in the allocation along the boundary L of the participation region are associated with discontinuities in the function $(\alpha(\cdot), \theta(\cdot))$. Such discontinuities generate discontinuities in the set $T^{**}(q)$, and hence the measure $\mu(q)$. This complicates the solution of (4).

For simplicity, we will henceforth assume that the function $(\alpha(q), \theta(q))$ is absolutely continuous on the interval $(\underline{q}(0), \bar{q}(1)]$. This approach is justified by the following reasoning. Suppose there exists a solution to problem (4) when $(\alpha(\cdot), \theta(\cdot))$ is restricted to be absolutely continuous. Because absolutely continuous functions are dense in the set of measurable functions, such a solution must also be a solution to the unrestricted problem. Our analysis below will identify conditions under which the restricted problem has such a solution. These existence conditions automatically identify circumstances under which "ironing" of the allocation along L will not be needed. When ironing is necessary, a more complicated approach using impulse control is appropriate. We pursue this route elsewhere (Deneckere and Severinov, 2009c).

4 The Reformulated Problem

In this section, we use Theorem 2 to reformulate the monopolist's problem (4) into an optimal control problem. Suppose therefore that we are given a point $(\widehat{\alpha}, \widehat{\theta})$ with either $\widehat{\alpha} = 1$ or $\widehat{\theta} = 0$, and two strictly increasing functions $\underline{q} : [0, 1] \to \mathbf{R}_+$ and $\bar{q} : [\widehat{\theta}, 1] \to \mathbf{R}_+$, satisfying $\bar{q}(\widehat{\theta}) = \underline{q}(1)$. Let $(\alpha(q), \theta(q))$ be defined by (7). Also, for any point (α, θ) let the function $\sigma(q, \alpha, \theta, \cdot)$ represent the iso-price curve through (α, θ) at the quantity q. Thus σ solves the equation

$$u_q(q, \sigma, a) = u_q(q, \theta, \alpha).$$

We may now state:

Theorem 3 *Suppose Assumptions 1 and 2 hold, and suppose that* $(\alpha(\cdot), \theta(\cdot)) : [0, \bar{q}(1)] \to \mathbf{R}_+^2$ *is piecewise continuously differentiable. Define the functions*

$$h(q, \alpha, \theta, \alpha', \theta') = \int_0^{\alpha} f(\sigma(q, \alpha, \theta, a), a)[\sigma_q(q, \alpha, \theta, a) + \sigma_\theta(q, \alpha, \theta, a)\theta' + \sigma_\alpha(q, \alpha, \theta, a)\alpha']da$$

and

$$H(q, \alpha, \theta) = \int_0^1 \int_{\sigma(q,\alpha,\theta,a)}^{\infty} f(a, \theta) d\theta da$$

Then the monopolist's profits are given by

$$\int_0^{\underline{q}(\widehat{\alpha})} u(q,\alpha(q),\theta(q))h(q,\alpha(q),\theta(q),\alpha'(q),\theta'(q))dq$$
$$+u(\underline{q}(\widehat{\alpha}),\widehat{\alpha},\widehat{\theta})H(\underline{q}(\widehat{\alpha}),\widehat{\alpha},\widehat{\theta})+\int_{\underline{q}(\widehat{\alpha})}^{\bar{q}(1)} H(q,\alpha(q),\theta(q))u_q(q,\alpha(q),\theta(q))dq$$

Theorem 3 says that the monopolist's profits can be split into two parts, one part that depends only upon the allocation $(\alpha(q),\theta(q))$ for $q\le\underline{q}(\widehat{\alpha})$, and another part that depends only upon the allocation $(\alpha(q),\theta(q))$ for $q\ge\underline{q}(\widehat{\alpha})$. Over the first interval the tariff $P(q)$ equals the gross utility of the type $(\alpha(q),\theta(q))$ along the lower boundary that consumes q. The number of types from which $P(q)$ is collected is given by $h(q,\alpha(q),\theta(q),\alpha'(q),\theta'(q))$, which represents the density of types located on the isoquant through the point $(\alpha(q),\theta(q))$. Over the second interval, profits equals the sum of two terms. First, from each type that consumes more than $\underline{q}(\widehat{\alpha})$, of which there are $H(\underline{q}(\widehat{\alpha}),\widehat{\alpha},\widehat{\theta})=1-\mu(\underline{q}(\widehat{\alpha}))$, the monopolist collects the price paid by type $(\widehat{\alpha},\widehat{\theta})$, i.e. $u(\underline{q}(\widehat{\alpha}),\widehat{\alpha},\widehat{\theta})$. Second, the monopolist collects the marginal price $P'(q)=u_q(q,\alpha(q),\theta(q))$ from each type that consumes more than q, of which there are $H(q,\alpha(q),\theta(q))=1-\mu(q)$.

A consequence of Theorem 3 is that the monopolist's optimization problem can be split into three stages. First, given $\widehat{q}\in\mathbf{R}_+$ and $(\widehat{\alpha},\widehat{\theta})\in[0,1]^2$ such that either $\widehat{\alpha}=1$ or $\widehat{\theta}=0$, solve the following problem: Select $\underline{q}(1)$ and $\bar{q}(1)\in\mathbf{R}_+$, and functions $\alpha\in Lip([\widehat{q},\underline{q}(1)])$ and $\theta\in Lip([\underline{q}(1),\bar{q}(1)])$ to maximize

$$W(\widehat{q},\widehat{\alpha},\widehat{\theta})=u(\widehat{q},\widehat{\alpha},\widehat{\theta})H(\widehat{q},\widehat{\alpha},\widehat{\theta})+\int_{\widehat{q}}^{\bar{q}(1)} H(q,\alpha(q),\theta(q))u_q(q,\alpha(q),\theta(q))dq \tag{8}$$

subject to the constraints

$$\widehat{q}\le\underline{q}(1)\le\bar{q}(1)$$
$$\alpha(q)=1,\text{ for }q\ge\underline{q}(1)$$
$$\theta(q)=0,\text{ for }q\in(\widehat{q},\underline{q}(1)]$$
$$\theta(\bar{q}(1))=1$$
$$\alpha'(q),\theta'(q)\ge0$$

In the second stage, select $\underline{q}(0)$, and functions $\alpha\in Lip([0,\widehat{q}])$ and $\theta\in Lip([0,\widehat{q}])$ to maximize

$$V(\widehat{q},\widehat{\alpha},\widehat{\theta})=\int_{\underline{q}(0)}^{\widehat{q}} u(q,\alpha(q),\theta(q))h(q,\alpha(q),\theta(q),\alpha'(q),\theta'(q))dq+W(\widehat{q},\widehat{\alpha},\widehat{\theta}) \tag{9}$$

subject to the constraints

$$\alpha(\underline{q}(0))\ge0,\ \alpha(\widehat{q})=\widehat{\alpha},\text{ and }\theta(\widehat{q})=\widehat{\theta}$$
$$\alpha'(q)\ge0$$
$$\theta'(q)=-\frac{u_\alpha}{u_\theta}(q,\theta(q),\alpha(q))\alpha'(q)$$

Finally, in the last stage, select $\widehat{q}\in\mathbf{R}_+$, and $(\widehat{\alpha},\widehat{\theta})\in[0,1]^2$ such that either $\widehat{\alpha}=1$ or $\widehat{\theta}=0$ to maximize $V(\widehat{q},\widehat{\alpha},\widehat{\theta})$.

5 Solving the Control Problems

In this section, we solve the optimal control problems (8) and (9).

5.1 Solving the Variational Problem (8)

First, let us consider the calculus of variations problem (8), with fixed left hand and right hand boundaries, fixed initial "time" $\widehat{q}$, and free right hand "time" $\bar{q}(1)$. Our next theorem describes the solution to this problem

Theorem 4 *Suppose that the functions*

$$\begin{aligned}\phi(q,\theta) &= u_q(q,1,\theta)H_\theta(q,1,\theta)+u_{\theta q}(q,1,\theta)H(q,1,\theta)\\ \varkappa(q,\alpha) &= u_q(q,\alpha,0)H_\alpha(q,\alpha,0)+u_{\alpha q}(q,\alpha,0)H(q,\alpha,0)\end{aligned}$$

are increasing in q and decreasing in α and θ. Let $\alpha^{\varkappa}(q)$ denote the solution in α to the equation $\varkappa(q,\alpha)=0$, and let $\theta^{\phi}(q)$ denote the solution to the equation $\phi(q,\theta)=0$.

Then if $\widehat{\alpha}=1$, we have $\theta^{\phi}(q)\le\widehat{\theta}$ and the solution $\theta(q)$ to problem (8) satisfies $\theta(q)=\max\{\theta^{\phi}(q),\widehat{\theta}\}$. If $\widehat{\alpha}<1$ then $\underline{q}(1)$ solves $\phi(q,0)=0$, and $\alpha^{\varkappa}(q)<\widehat{\alpha}$. Furthermore, over the interval $(\widehat{q},\underline{q}(1)]$ the solution $\alpha(q)$ to problem (8) satisfies $\alpha(q)=\max\{\alpha^{\varkappa}(q),\widehat{\alpha}\}$, and over the interval $[\underline{q}(1),\bar{q}(1)]$ it satisfies $\theta(q)=\theta^{\phi}(q)$.

Under the conditions of Theorem 4 the monotonicity constraints $\alpha'(q)\ge 0$ and $\theta'(q)\ge 0$ can be ignored, and problem (8) is solved by pointwise maximization under the integrand. The conditions $\varkappa(q,\alpha)=0$ and $\phi(q,\theta)=0$ are a multi-dimensional version of a condition familiar from the one-dimensional type case, that at the optimum marginal virtual surplus must be equal to zero.[10], [11] Using the facts that $\lim_{q\to\bar{q}(1)}H(q,q,\theta(q)=0$ and $\lim_{q\to\bar{q}(1)}H_\theta(q,q,\theta(q)<0$, we also obtain the familiar condition that the allocation of the "top" type $(1,1)$ must be undistorted.

$$u_q(\bar{q}(1),1,1)=0.$$

When the monotonicity constraints are binding, we must associate Lagrange multipliers $\lambda(q)\ge 0$ and $\nu(q)\ge 0$ with the constraints $\alpha'(q)\ge 0$ and $\theta'(q)\ge 0$. Our next result describes how to obtain an 'ironed' solution:

Corollary: The solution to (8) satisfies $\lambda(q)\alpha'(q)=0$ and $\nu(q)\theta'(q)=0$. Over any interval in $(\widehat{q},\underline{q}(1)]$ on which $\lambda(q)>0$ we have $\lambda'(q)=\varkappa(q,\alpha(q))$, and over any interval in $[\underline{q}(1),\bar{q}(1)]$ on which $\nu(q)>0$ we have $\nu'(q)=\phi(q,\theta(q))$.

We can use Theorem 4 to establish necessary and sufficient conditions for the absence of exclusion in the optimal screening mechanism:

[10] More explicitly, letting t denote the type parameter in the one-dimensional screening model, and letting $F(t)$ denote its distribution function, the optimality condition is $u_qF'+u_{qt}(1-F(t))=0$.

[11] Note that when the monopolist selects $\widehat{\theta}<\theta(\widehat{q})$ or $\widehat{\alpha}<\alpha(\widehat{q})$, there is a non-empty right neighborhood of $\widehat{q}$ over which all isoquants emanate from $(\widehat{\alpha},\widehat{\theta})$. We shall show in Theorem 7 below that it is never optimal to do so.

Theorem 5 *Suppose the conditions of Theorem 4 hold. Then a necessary and sufficient condition for the absence of exclusion is that there exist $\mathring{\alpha} > 0$ such that*

$$-u_q(0,\alpha,0)\int_0^{\alpha}\frac{f(a,0)}{u_{q\theta}(q,a,0)}da + H(0,\alpha,0) = 0 \tag{10}$$

and such that $u_{q\alpha}(0,\alpha,0) = 0$ for all $\alpha \in [0,\mathring{\alpha}]$.

Theorem 5 sheds considerable light on how multi-dimensionality in customer types affects the monopolist's incentive to exclude some customers from the market. To this effect, let us start by giving an economic interpretation to equation (10). Consider the aggregate demand curve for the first increment, $N(p,0) = \#\{t : u_q(0,t) \geq p\}$. Let $\alpha(p)$ be such that the demand price of type $(\alpha(p),0)$ equals p, i.e. $u_q(0,\alpha(p),0) = p$. Then we have $N(p,0) = H(0,\alpha(p),0)$. Thus equation (10) says that the marginal price for the first increment must maximize the profits from that increment:

$$p\frac{\partial N}{\partial p}(p,0) + N(p,0) = 0$$

Exclusion will occur if and only if at the monopoly price for this increment some consumers decide not to purchase the increment, i.e. if there are some types $\alpha < \mathring{\alpha}$ for which $u_q(0,\alpha,0) < u_q(0,\mathring{\alpha},0)$. The condition $u_{q\alpha}(0,\alpha,0) = 0$ for all $\alpha \in [0,\mathring{\alpha}]$ rules this out.

As in the one-dimensional type case, absence of exclusion requires the demand curve for the first increment to be perfectly inelastic at a price equal to marginal cost. Indeed, if there is no "gap" between the lowest demand price for the first increment and marginal cost, i.e. if $u_q(0,\alpha,0) = 0$, then equation (5) can hold only if $\int_0^{\alpha} f(a,0)da = \infty$, i.e. if and only if $\frac{\partial N}{\partial p}(0,0) = \infty$. On the other hand, if the gap between the lowest demand price for the first increment and marginal cost is sufficiently large, then like in the one-dimensional type case there can be no exclusion, provided $u_{q\alpha}(q,\alpha,0) = 0$ for all $\alpha \in [0,\mathring{\alpha}]$. Our next example shows that this can indeed happen.

Example 3: Let $u(q,\alpha,\theta) = \frac{1+\theta+k}{2}q - \frac{b-\alpha}{2}q^2$ for some $k \geq 0$ and $b \geq 1$, let $c = 0$, and let $f(\alpha,\theta) = 1$ for all $(\alpha,\theta) \in [0,1]^2$. Then (10) becomes $-(1+k)\alpha + 1 = 0$, so we have $\mathring{\alpha} = \frac{1}{1+k} \in (0,1]$.

Armstrong (1996) has argued forcefully that exclusion necessarily occurs types are multi-dimensional. Since our other assumptions are consistent with those put forth by Armstrong,[12] Theorem 5 indicates that Armstrong's conclusion is specific to cases where the allocation space and the type space have the same dimensionality. Nevertheless, our theorem also demonstrates that there is a sense in which non-exclusion is harder to obtain when the type space is multi-dimensional. In considering raising the marginal price for the first increment above the level where all consumers are included, the monopolist trades off the extra dollar gained on all existing customers (measured by the term $H(0,\alpha,0)$) against the loss in revenue caused by some consumers dropping out of the market (measured by the term in (10)). The number of lost customers is measured (roughly) by the length of the isoquant emanating from the point

[12] In particular, our type space is a strictly convex set with non-empty interior. Also, the utility function in Example 3 is convex and homogeneous of degree one in types.

$(\mathring{\alpha}, 0)$. If we had $\mathring{\alpha} = 0$, then the number of customers dropping out would be negligible, and exclusion would always pay. This is essentially the effect identified by Armstrong. On the other hand, if $\mathring{\alpha} > 0$, then for no customer to be excluded at the price $u_q(q, \mathring{\alpha}, 0)$ the isoquant through $(\mathring{\alpha}, 0)$ at the quantity $q = 0$ must be flat. If it were the case that $u_{q\alpha}(0, \alpha, \theta) > 0$ for all (α, θ) then exclusion would necessarily occur.

5.2 Solving the Optimal Control Problem (9)

To solve (9), we formulate it as the following optimal control problem: Find piecewise continuous control functions $v(q)$ and $w(q)$, and associated continuous and piecewise differentiable state variables $\underline{\theta}(q)$ and $\alpha(q)$, defined on the "time" interval $[\underline{q}(0), \widehat{q}]$, that solve

$$W_0(\widehat{q}, \widehat{\alpha}, \widehat{\theta}) = \max_{(v(\cdot), w(\cdot)) \in U, \underline{q}(0)} \int_{\underline{q}(0)}^{\widehat{q}} u(q, \alpha(q), \underline{\theta}(q)) h(q, \alpha(q), \underline{\theta}(q), v, w) dq, \tag{11}$$

subject to the state evolution equations

$$\begin{aligned} \dot{\alpha} &= v, \\ \dot{\underline{\theta}} &= w, \end{aligned} \tag{12}$$

initial conditions $\alpha(0) \geq 0$, terminal conditions $\alpha(\widehat{q}) = \widehat{\alpha}$ and $\underline{\theta}(\widehat{q}) = \widehat{\theta}$, and control variable restrictions

$$(v, w) \in U(q, \alpha, \underline{\theta}) \equiv \{(v, w) : v \geq 0 \text{ and } w + vg(q, \underline{\theta}, \alpha) = 0\}$$

To solve this control problem, let

$$F(q, \alpha, \underline{\theta}, v, w) = u(q, \alpha, \underline{\theta}) h(q, \alpha, \underline{\theta}, v, w)$$

and form the Hamiltonian

$$H(q, \alpha, \underline{\theta}, v, w, \lambda, \mu) = F(q, \alpha, \underline{\theta}, v, w) + \mu v + \lambda w$$

According to Pontryagin's maximum principle, the optimal controls (v, w) must maximize H over the set U, i.e. must be stationary points of the Lagrangian

$$L(q, \alpha, \underline{\theta}, v, w, \lambda, \mu, \xi, \chi) = F(q, \alpha, \underline{\theta}, v, w) + \mu v + \lambda w + \xi(w + vg(q, \alpha, \underline{\theta})) + \chi v$$

and the costates variables μ and λ must satisfy the evolution equations

$$\begin{aligned} \dot{\mu} &= -\frac{\partial L}{\partial \alpha} \\ \dot{\lambda} &= -\frac{\partial L}{\partial \underline{\theta}} \end{aligned}$$

Our next theorem summarizes the solution, using the following notation:

$$\begin{aligned} h_0(q, \alpha, \underline{\theta}) &= \int_{\underline{\alpha}(q, \alpha, \underline{\theta})}^{\alpha} f(\sigma(q, \alpha, \underline{\theta}, a), a) \sigma_q(q, \sigma(q, \alpha, \underline{\theta}, a), a) da \\ h_1(q, \alpha, \underline{\theta}) &= \int_{\underline{\alpha}(q, \alpha, \underline{\theta})}^{\alpha} f(\sigma(q, \alpha, \underline{\theta}, a), a) \sigma_{\underline{\theta}}(q, \sigma(q, \alpha, \underline{\theta}, a), a) da \\ h_2(q, \underline{\theta}, \alpha) &= \int_{\underline{\alpha}(q, \alpha, \underline{\theta})}^{\alpha} f(\sigma(q, \alpha, \underline{\theta}, a), a) \sigma_\alpha(q, \sigma(q, \alpha, \underline{\theta}, a), a) da. \end{aligned} \tag{13}$$

and

$$\psi(q,\alpha,\underline{\theta}) = \frac{u_q u_\theta}{u_{q\theta}}(q,\alpha,\underline{\theta})h_1(q,\alpha,\underline{\theta})$$

Theorem 6 *(i) Over any nondegenerate interval on which α is strictly increasing, and hence $\underline{\theta}$ is strictly decreasing, the solution to (11) satisfies*

$$\begin{aligned}
\mu(q) &= -uh_2 + \psi g \\
\lambda(q) &= -uh_1 + \psi \\
\xi(q) &= -\psi \\
\chi(q) &= 0
\end{aligned}$$

On such an interval, the state variables α and $\underline{\theta}$ satisfy the pair of differential equations

$$\begin{aligned}
\alpha'(q) &= \frac{u_\theta h_0 - u_q h_1 + \psi_q}{uf\sigma_{\underline{\theta}} + \psi g_\theta + \psi_{\underline{\theta}} g - \psi_\alpha - u_\theta(h_2 - gh_1)} \\
\underline{\theta}'(q) &= -g\alpha'.
\end{aligned} \tag{14}$$

(ii) Over any nondegenerate interval on which α (and hence) $\underline{\theta}$ are constant, we have

$$\begin{aligned}
\dot{\xi} &= u_\theta h_0 - u_q h_1, \text{ and} \\
\dot{\chi} &= -(\psi + \xi) g_q \\
\dot{\mu} &= -\frac{\partial}{\partial\alpha} uh_0 \\
\dot{\lambda} &= -\frac{\partial}{\partial\underline{\theta}} uh_0.
\end{aligned}$$

(iii) The functions $\lambda(q)$, $\mu(q)$ and $\xi(q)$ are continuous, except possibly at a point q^{**} where $\sigma(q,\alpha(q),\underline{\theta}(q),0) = 1$.

(iii) The solution satisfies $\alpha(\underline{q}(0))\underline{q}(0) = 0$.

Theorem 6 tells us how to construct the solution $(\alpha(q),\underline{\theta}(q))$, for $q \in [0,\widehat{q}]$. Starting from the point $(\widehat{\alpha},\widehat{\theta})$ run the differential equation system (14) backwards, for as along as $\alpha' > 0$. Remain at any point $(\alpha,\underline{\theta})$ for any interval of q for which the constraint $\alpha'(q) \geq 0$ is violated, then resume running (14) backwards, and continue this process until either $\alpha(q) = 0$ or $q = 0$.

The reason for a potential discontinuity in the costate variables is that function $h(q,\alpha,\underline{\theta},v,w) = h_0(q,\alpha,\underline{\theta}) + h_1(q,\alpha,\underline{\theta})v + h_2(q,\alpha,\underline{\theta})w$, and hence the objective $F(q,\alpha,\underline{\theta},v,w)$, is not continuously differentiable at a point q^{**} at which the isoquant trough $(\alpha(q),\,\underline{\theta}(q))$ hits the northwest corner point $(0,1)$. The reason for this can be gleaned from equation (13): $\underline{\alpha}(q,\alpha,\underline{\theta}) > 0$ for any $q > q^{**}$, but $\underline{\alpha}(q,\alpha,\underline{\theta}) = 0$ for any $q \leq q^{**}$. The derivatives of $\underline{\alpha}$, and hence of the function h, will therefore exhibit a discontinuity at $q = q^{**}$.[13]

Our next lemma states some mild regularity conditions which yields more structure to the solution of (11):

[13] Optimal control theory requires the objective function to be continuously differentiable in the state. To handle the discontinuity at q^{**}, we split the single 'time' interval $[\underline{q}(0), q^{**}]$ into to two separate intervals, $[\underline{q}(0), q^{**}]$ and $[q^{**}, \bar{q}(1)]$.

Lemma 8 *(i) Suppose that* $u - \frac{u_\theta u_q}{u_{q\theta}} > 0$ *for all* $q > 0$ *and* t. *Then* $\alpha(\underline{q}(0)) > 0$ *and* $\underline{q}(0) = 0$.

(ii) Suppose that $\frac{\partial}{\partial q}\{u_\theta h_0 - u_q h_1 + \psi_q\} < 0$. *Then there is at most one interval on which the constraint* $\alpha' \geq 0$ *is binding. If there is such an interval, then it contains* $q = 0$.

Lemma 8 reveals some important qualitative properties of the optimal mechanism. First, whenever $u - \frac{u_\theta u_q}{u_{q\theta}} > 0$ for all $q > 0$, the monopolist finds it optimal to sell quantities arbitrarily close to zero to some consumer types. The required condition holds for most commonly specified utility functions, and is satisfied whenever $u - \frac{u_\theta u_q}{u_{q\theta}}$ is strictly increasing in q. A sufficient condition for the latter property is that $u_{qq\theta} \geq 0$.[14] Second, over any interval $[q_-, q_+]$ on which $\alpha'(q)$ is constant, all isoquants associated with $q \in [q_-, q_+]$ emanate from the same point $(\breve{\alpha}, \breve{\theta})$ on the lower boundary. In other words, there is a discontinuity in the allocation assigned to types on the lower boundary at the point $(\breve{\alpha}, \breve{\theta})$. Unlike in the one-dimensional type case, such discontinuities are not associated with gaps in the consumption schedule.[15] Our regularity condition in (ii) ensures that there is at most one point on the lower boundary where this can happen, in which case it is necessary that $q_- = 0$.

5.3 Transversality Conditions for $(\widehat{\alpha}, \widehat{\theta})$ and $\widehat{q}$.

It remains to determine the transversality conditions for $(\widehat{\alpha}, \widehat{\theta})$ and $\widehat{q}$. Recall that there are only two free variables, since either $\widehat{\alpha} = 1$ or $\widehat{\theta} = 0$. Our first result establishes that at the optimum there is a one to one relationship between $(\widehat{\alpha}, \widehat{\theta})$ and $\widehat{q}$, effectively reducing our optimization problem to the determination of a single parameter.

Theorem 7 *Suppose that the functions* $\phi(q, \theta)$ *and* $\varkappa(q, \alpha)$ *are increasing in* q *and decreasing in* α *and* θ. *Then at the optimum,* $\widehat{\theta} = \theta^{\phi}(\widehat{q})$ *whenever* $\widehat{\alpha} = 1$, *and* $\widehat{\alpha} = \alpha^{\varkappa}(\widehat{q})$ *whenever* $\widehat{\theta} = 0$.

Define q^0 to be the unique solution to the equation $\phi(q, 0) = 0$. According to Theorem 7 whenever $\widehat{q} \geq q^0$ it must be that $\widehat{\theta} = \theta^{\phi}(\widehat{q})$, and whenever $\widehat{q} \leq q^0$we have $\widehat{\alpha} = \alpha^{\varkappa}(\widehat{q})$. It remains to determine the optimal value of $\widehat{q}$. Our next theorem addresses this issue, using the notation $\mu_-(q^{**}) = \lim_{q \uparrow q^{**}} \mu(q)$ and $\mu_+(q^{**}) = \lim_{q \downarrow q^{**}} \mu(q)$, and similarly for $\lambda_-(q^{**})$ and $\lambda_+(q^{**})$:

Theorem 8 *Suppose that the functions* $\phi(q, \theta)$ *and* $\varkappa(q, \alpha)$ *are increasing in* q *and decreasing in* α. *Then:*

(i) If $\underline{\alpha}(\widehat{q}, \widehat{\alpha}, \widehat{\theta}) > 0$ *and* $\alpha'(\widehat{q}) > 0$ *the following transversality condition must hold:*

$$[\mu_+(q^{**}) - \mu_-(q^{**})]\frac{d\alpha^{**}}{d\widehat{q}} + [\lambda_+(q^{**}) - \lambda_-(q^{**})]\frac{d\theta^{**}}{d\widehat{q}} = 0;$$

[14] We have $\frac{\partial}{\partial q}(u_{q\theta}u - u_\theta u_q) = u_{qq\theta}u - u_\theta u_{qq} > 0$ whenever $u_{qq\theta} > \frac{u_\theta u_{qq}}{u_\theta}$.

[15] In the one-dimensional type case, if there is an interval of $[q_-, q_+]$ on which $t(q)$ is constant, then no consumer other than $t(q_-)$ purchases quantities in the interval $[q_-, q_+]$.

(ii) If $\underline{\alpha}(\widehat{q},\widehat{\alpha},\widehat{\theta}) = 0$, or if $\alpha'(q) = 0$ for all $q \leq \widehat{q}$, the following transversality condition must hold:[16]

$$\begin{aligned}
\lambda(\widehat{q}) &= \psi(\widehat{q},1,\theta^{\phi}(\widehat{q})) - u(\widehat{q},1,\theta^{\phi}(\widehat{q}))h_1(\widehat{q},1,\theta^{\phi}(\widehat{q})), \text{ if } \theta^{\phi}(\widehat{q}) > 0 \\
\mu(\widehat{q}) &= \psi(\widehat{q},\alpha^{\varkappa}(\widehat{q}),0)g(\widehat{q},\alpha^{\varkappa}(\widehat{q}),0) - u(\widehat{q},\alpha^{\varkappa}(\widehat{q}),0)h_2(\widehat{q},\alpha^{\varkappa}(\widehat{q}),0), \text{ if } \alpha^{\varkappa}(\widehat{q}) < 1
\end{aligned}$$

To interpret Theorem 8 note that in case (i) the transversality condition only has bite if $\alpha'(q) > 0$ in a right neighborhood of q^{**} but $\alpha'(q) = 0$ in a left neighborhood of q^{**}. Theorem 8 therefore suggests that at the optimum, whenever $\widehat{q} > 0$ or $\underline{\alpha}(\widehat{q},\widehat{\alpha},\widehat{\theta}) > 0$, there will be an interval of quantities (a left neighborhood of q^{**} in case (i), and a left neighborhood of $\widehat{q}$ in case (ii)) for which isoquants emanate from the same point (the point $(\alpha^{**},\theta^{**})$ in case (i), and the point $(\widehat{\alpha},\widehat{\theta})$ in case (ii)). Our next lemma simplifies the task of applying Theorem 8, under an additional regularity condition:

Lemma 9 *Suppose that $\frac{\partial}{\partial q}\{u_\theta h_0 - u_q h_1 + \psi_q\} < 0$, and suppose that $\alpha'(q) = 0$. Then $\lambda(q) = -\int_0^q \frac{\partial(uh_0)}{\partial\underline{\theta}}(z,\alpha(q),\underline{\theta}(q))dz$ and $\mu(q) = -\int_0^q \frac{\partial(uh_0)}{\partial\alpha}(z,\alpha(q),\underline{\theta}(q))dz$.*

We can use Theorem 8 to derive necessary conditions for the demand profile approach to yield the correct optimal screening mechanism:

Theorem 9 *Suppose that the functions $\phi(q,\theta)$ and $\varkappa(q,\alpha)$ are increasing in q and decreasing in α and θ. Then for the demand profile approach to yield the optimal screening mechanism it is necessary and sufficient that $\widehat{q} = 0$ in the optimal mechanism.*

The conditions of Theorem 9 are extremely stringent, as our example below will illustrate.

6 A linear-quadratic example

In this section, we derive an explicit solution for a parametrically specified example. Let

$$u(q,\alpha,\theta) = \theta q - \frac{b-\alpha}{2}q^2 \tag{15}$$

where $b \geq 1$. Furthermore, let (α,θ) be uniformly distributed on the unit square $I = [0,1]\times[0,1]$:

$$f(\alpha,\theta) = 1 \text{ for all } (\alpha,\theta) \in I. \tag{16}$$

Note that since $u_q(0,\alpha,0) = 0$, by Theorem 5 there will always be exclusion in the optimal mechanism. The solution to this example takes on a different qualitative form depending upon whether $b \geq \frac{3}{2}$ or $b < \frac{3}{2}$. We start with the case $b \geq \frac{3}{2}$, which was previously analyzed by Laffont, Maskin and Rochet (1987).

Theorem 10 *The optimal screening mechanism for the linear-quadratic uniformly distributed example (15)-(16) with $b \geq 3/2$ is as follows:*

[16] The proof of the Theorem states a more complicated condition that must hold if $\alpha'(\widehat{q}) = 0$, but $\alpha'(q) > 0$ for some $q < \widehat{q}$.

Let $q^* = \frac{2}{2b+1}$, $\theta^* = \frac{2b-1}{2b+1}$, and $\bar{q} = \frac{1}{b-1}$. Then $\alpha(q) = 1$ for all $q \in [0, \bar{q}]$ and

$$\theta(q) = \begin{cases} \frac{1+2(b-1)q}{3}, & \text{for } q \in [q^*, \bar{q}] \\ \frac{1+(b-\frac{3}{2})q}{2}, & \text{for } q \in [0, q^*]. \end{cases}$$

Thus the optimal nonlinear tariff is given by:

$$P(q) = \begin{cases} \frac{1}{6(2b+1)} + \frac{q(2-(b-1)q)}{6}, & \text{for } q \in [q^*, \bar{q}] \\ \frac{q}{8}(4-(2b-1)q), & \text{for } q \in [0, q^*]. \end{cases}$$

When $b \geq \frac{3}{2}$, we have $q^{**} = \widehat{q} = 0$ so the region associated with 9 is empty. All isoquants therefore emanate from the portion of right hand boundary $\alpha = 1$ above $\theta = \frac{1}{2}$, i.e. the interval of points $\{(1, \theta) : \theta \in [\frac{1}{2}, 1]\}$. Note in particular that the isoquant associated with $q = 0$ is a flat line segment at $\theta = \frac{1}{2}$, i.e. the collection of points $\{(\alpha, \frac{1}{2}) : \alpha \in [0, 1]\}$. Figure 2 illustrates the isoquants for this case. None of the iso-price lines associated with this mechanism intersect each other in the type space. As a consequence, the demand profile approach properly identifies the optimal mechanism. Since α varies from $b-1$ to b, large values of b are associated with low variability in the slope parameter. Thus, one way to interpret this result is that when the uncertainty is (sufficiently) close to one dimensional, the demand profile approach is valid. We now turn to the significantly more complicated case where $b < 3/2$.

Theorem 11 *The optimal screening mechanism for the linear-quadratic uniformly distributed example (15)-(16) with $b < 3/2$ is as follows:*

Let $\alpha^{**} = \frac{2b}{3}$, $\theta^{**} = 1 - \frac{2bq^{**}}{3}$, q^{**} the unique non-negative root to the equation

$$(1 + bq - bq^2(\frac{3}{2} + 2b))^2 = (1 - bq)(1 + bq - bq^2(\frac{5}{2} + b))^3, \tag{17}$$

and

$$\widehat{\theta} = \frac{1 + \sqrt{(1-bq)(4(1+bq) - 2bq^2(5+2b)}}{3},$$

and $\widehat{q} = \frac{3\widehat{\theta}-1}{2(b-1)}$. Then

$$\alpha(q) = \begin{cases} \alpha^{**}, & \textit{for } q \in [0, q^{**}] \\ c_0 + \frac{c_1}{27}(2 - \sqrt{1 - \frac{6}{c_1 q}})(1 + \sqrt{1 - \frac{6}{c_1 q}})^2, & \textit{for } q \in [q^{**}, \widehat{q}] \\ 1, & \textit{for } q \in [\widehat{q},\, \bar{q}] \end{cases}$$

and

$$\theta(q) = \begin{cases} \theta^{**}, & \textit{for } q \in [0, q^{**}] \\ \frac{2}{3} - \frac{1}{3}\sqrt{1 - \frac{6}{c_1 q}}, & \textit{for } q \in [q^{**}, \widehat{q}] \\ 1, & \textit{for } q \in [\widehat{q},\, \bar{q}] \end{cases}$$

where the constants c_0 *and* c_1 *are related to* $\widehat{\theta}$ *as follows:*

$$c_1 = \frac{4(b-1)}{(1-\widehat{\theta})(3\widehat{\theta}-1)^2}, \text{ and} \tag{18}$$

$$c_0 = \frac{\widehat{\theta}^2(5+4b)-\widehat{\theta}(2+4b)+1}{(3\widehat{\theta}-1)^2} < 0 \tag{19}$$

Thus the optimal nonlinear tariff is given by:

$$P(q) = \begin{cases} u(q, \alpha(q), \theta(q))\,, & \text{for } q \in [0, \widehat{q}] \\ \frac{(b-1)\widehat{q}^2}{3} + \frac{q}{6}(2-(b-1)q), & \text{for } q \in [0, q^*]. \end{cases}$$

For $b < \frac{3}{2}$, in the optimal screening mechanism the isoquants for $q \in [0, q^{**}]$ all emanate from the point $(\alpha^{**}, \theta^{**})$ on the lower boundary. In particular, for $q = 0$ the isoquant is the flat segment at the level $\theta = \theta^{**}$ with $\alpha = \alpha^{**}$, i.e. the collection of points $\{(\alpha, \theta^{**}) : \alpha \in [0, \alpha^{**}]\}$. For $q \in [q^{**}, \widehat{q}]$ the lower boundary is strictly decreasing, and given by the equation $\alpha = c_0 + c_1\theta(1-\theta)^2$. For this segment of q values there is a unique isoquant associated with every point on the lower boundary. Note that since all types (α, θ^{**}) along the lower boundary with $\alpha \leq \alpha^{**}$ are assigned a quantity 0, and since all types along the lower boundary with $\theta > \theta^{**}$ are assigned a quantity $q \geq q^{**}$, there is a discontinuity in the optimal quantity assignment along the lower boundary. Finally, for $q \geq \widehat{q}$, all isoquants emanate from the portion of the right hand boundary $\alpha = 1$ with $\theta \geq \widehat{\theta}$. Figure 3 illustrates the lower boundary and the isoquants for the case $b < \frac{3}{2}$. It is important to observe that while the isoquants associated with the optimal mechanism never intersect in the interior of the participation region, the corresponding price lines would intersect in the region of non-participation. Thus in accordance with Theorem 9 for every value of the parameter b with $b < 3/2$, the demand profile is incapable of correctly identifying the optimal mechanism.

7 Conclusion

In this paper, we have shown that the traditional method for identifying an optimal screening mechanism, the demand profile approach, generally fails when there is multi-dimensional uncertainty. Only under rather extreme conditions on the type distribution, essentially reducing the problem to one with single dimensional uncertainty, will the chosen mechanism be optimal. We identified the reason for this failure: with multi-dimensional uncertainty, a consumer's demand schedule must generally intersect the optimal marginal price schedule multiple times, thereby wreaking havoc with the global incentive compatibility requirement.

We introduced a novel condition, termed single crossing of demand (SCD), under which global incentive compatibility can nevertheless be assured. This condition guarantees that if a quantity $q > 0$ solves the surplus maximization problem of an agent of type (α, θ), then q must also be a global optimum for any type on the portion of the iso-price curve at the quantity q going through the point (α, θ) that lies to the northwest of this point. As a consequence, isoquants are the portions of isoprice curves that lie above a lower boundary defined by the individual rationality constraint.

Correct identification of these isoquants then allowes us to reduce the problem to a one-dimensional screening problem, all be it a rather complicated one. We were able to reduce the resulting optimization problem to an optimal control problem, and identify its solution. We also illustrated application of our methodology to an example with quadratic demand and uniformly distributed types.

Our methodology has already identified some relatively robust properties of optimal screening mechanism with multidimensional types. In particular, the allocation to an agent may be discontinuous in type along the boundary of the participation region. We also showed that the optimal mechanism may or may not exclude some types from participation. We hope that our paper will stimulate new research into several of the applications cited in the introduction.

While the present analysis was confined to the case where the (physical) allocation space is one-dimensional, our approach should prove useful in analyzing more general screening problems in which the dimensionality of the type space exceeds the dimensionality of the allocation space.

References

[1]

Armstrong, M. (1996), “Multiproduct Nonlinear Pricing,” *Econometrica*, 64, pp. 51-75.

Armstrong, M. (1999), “Optimal Regulation with Unknown Demand and Cost” Journal of Economic Theory, 84, pp. 196-215.

Ausubel and Deneckere (1993), “A Generalized Theorem of the Maximum” Economic Theory, 3, pp. 99-107.

Basov, S. (2003), Multidimensional Screening, Springer-Verlag.

Basov, S. (2001), “Hamiltonian Approach to Multidimensional Screening,” *Journal of Mathematical Economics*, 36, 77-94

Brown and Sibley (1986), The Theory of Public Utility Pricing, Cambridge University Press, Cambridge.

Carlier, G., (2001), “A General Existence Result for the Principal-agent Problem with Adverse Selection,” *Journal of Mathematical Economics,* 35(1), 129-150.

Deneckere, R. and S. Severinov (2009a), “Multi-Dimensional Screening: The Case n>2,” mimeo, University of Wisconsin-Madison.

Deneckere, R. and S. Severinov (2009b), “A Sufficient Condition for the Validity of the Demand Profile Approach to Multi-Dimensional Screening,” mimeo, University of Wisconsin-Madison.

Deneckere, R. and S. Severinov (2009b), “Ironing in Multi-Dimensional Screening with a One-Dimensional Allocation Space,” mimeo, University of Wisconsin-Madison.

Goldman, M., H. Leland and D. Sibley (1984), “Optimal Nonuniform Pricing,” Review of Economic Studies, 51, pp.305-320.

Laffont, J-J., E. Maskin and J.-C. Rochet (1987), “Optimal Non-Linear Pricing with Two-Dimensional Characteristics,” in T. Groves, R. Radner and S. Reiter, eds., “Information, Incentives and Economic Mechanisms,” University of Minnesota Press, Minneapolis, pp. 256-266.

Lewis, T. R. and D. M. Sappington, (1988), “Regulating a Monopolist with Unknown Demand and Cost Functions,” *Rand Journal of Economics*, 19, 438-457.

Milgrom, P. and I. Segal (2002), ‘Envelope Theorems for Arbitrary Choce Sets’, *Econometrica*, 70(2), pp. 583-601.

McAfee, R.P. and J. McMillan (1988), 'Multidimensional Incentive Compatibility and Mechanism Design', *Journal of Economic Theory*, 46, pp. 335-354.

Rochet, J. C., (1987), "A Necessary and Sufficient Condition for Rationalizability Quasi-Linear Context," *Journal of Mathematical Economics*, 16, 191-200.

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

J.-C. Rochet and L. Stole (2001), "The Economics of Multidimensional Screening" mimeo, University of Chicago.

Wilson, R. (1993), "Non-Linear Pricing," Oxford University Press.

8 Appendix A

Proof of Lemma 1: Observe that, for a fixed q the relation $u_q(q,\theta',\alpha') = u_q(q,\theta,\alpha)$ implicitly defines a function $\widetilde{\theta}(\alpha)$. Note that $u_{q\theta}(q,\widetilde{\theta}(\alpha),\alpha)\frac{d\widetilde{\theta}}{d\alpha} + u_{q\alpha}(q,\widetilde{\theta}(\alpha),\alpha) = 0$. Hence $u_{qq}(q,\theta',\alpha') - u_{qq}(q,\theta,\alpha) = \int_\alpha^{\alpha'}[u_{qq\theta}(q,\widetilde{\theta}(a),a)\frac{d\widetilde{\theta}}{d\alpha} + u_{qq\alpha}(q,\widetilde{\theta}(a),a)]da = \int_\alpha^{\alpha'}[-u_{qq\theta}\frac{u_{q\alpha}}{u_{q\theta}} + u_{qq\alpha}]da > 0$, proving the desired result. *Q.E.D.*

Proof of Lemma 2: That Q^* is a non-empty closed-valued u.h.c. correspondence, and that s is a continuous function follows from the Generalized Theorem of the Maximum (Ausubel and Deneckere, 1993). The Generalized Envelope Theorem (Milgrom and Segal, 2002) implies that the envelope condition $\nabla s(t) = \nabla_t u(q(t),t)$ holds a.e. Monotonicity of Q^* follows because u is strictly supermodular. *Q.E.D.*

Proof of Lemma 3: We start by showing that $s(\alpha,\theta) > 0$ for all $\theta > \underline{\theta}(\alpha)$. We also argue that $q(\alpha,\theta) = 0$ for all $\theta \leq \underline{\theta}(\alpha)$ and $q(\alpha,\theta) > 0$ for all $\theta \geq \underline{\theta}(\alpha)$. By definition of $\underline{\theta}(\alpha)$, we have $s(\alpha,\theta) = 0$ for all $\theta \leq \underline{\theta}(\alpha)$. It then follows from the envelope theorem, and Assumption 1(ii) that $q(\alpha,\theta) = 0$ for all $\theta < \underline{\theta}(\alpha)$. The monotonicity of Q^* implies that for all $\theta > \underline{\theta}(\alpha)$ we have $q(\alpha,\theta) > 0$. Application of the envelope theorem then yields $s(\alpha,\theta) = \int_{\underline{\theta}(\alpha)}^{\theta} u_\theta(q(\alpha,\theta'),\theta')d\theta'$ > 0 for all $\theta > \underline{\theta}(\alpha)$.

Next let us show that $\alpha' < \alpha$ implies $\underline{\theta}(\alpha') \geq \underline{\theta}(\alpha)$. Since $q(\theta,\alpha) = 0$ for all $\theta < \underline{\theta}(\alpha)$, it follows from monotonicity of q in α implies that $q(\theta,\alpha') = 0$ for all $\alpha' < \alpha$ and all $\theta < \underline{\theta}(\alpha)$, implying the required inequality.

Finally, let us show that $\underline{\theta}(\alpha') < \underline{\theta}(\alpha)$ whenever $\alpha' > \alpha$ and $q(\alpha,\underline{\theta}(\alpha)) > 0$. Monotonicity of q yields $q(a,\underline{\theta}(\alpha)) \geq q(\alpha,\underline{\theta}(\alpha)) > 0$ for all $a \geq \alpha$. From the the envelope theorem, we have $s(\alpha',\underline{\theta}(\alpha)) = s(\alpha,\underline{\theta}(\alpha)) + \int_\alpha^{\alpha'} u_\alpha(q(a,\underline{\theta}(\alpha)),a,\underline{\theta}(\alpha))da > s(\alpha,\underline{\theta}(\alpha))$ for all $\alpha' > \alpha$. Continuity of s then implies that $\underline{\theta}(\alpha') = \inf\{\theta | s(\theta,\alpha') > 0\} < \underline{\theta}(\alpha)$. *Q.E.D.*

Proof of Lemma 4: If $q_1 \in Q^*(\alpha_1,\theta_1)$ then $u(q_1,\alpha_1,\theta_1) - T(q_1) \geq u(q,\alpha_1,\theta_1) - T(q)$, for all q. Rearranging, we have $T(q) \geq u(q,\alpha_1,\theta_1) - u(q_1,\alpha_1,\theta_1) + T(q_1)$, for all q. We now claim that $u(q,\alpha_1,\theta_1) - u(q_1,\alpha_1,\theta_1) \geq u(q,\alpha_2,\theta_2) - u(q_1,\alpha_2,\theta_2)$, for all q and (α_2,θ_2) satisfying the hypothesis of the Lemma. It then follows from the previous inequality that $T(q) \geq u(q,\alpha_2,\theta_2) - u(q_1,\alpha_2,\theta_2) + T(q_1)$, for all q so that $q_1 \in Q^*(\alpha_2,\theta_2)$.

To prove the claim, we need to show that $u(q,\alpha_1,\theta_1) - u(q,\alpha_2,\theta_2) \geq u(q_1,\alpha_1,\theta_1) - u(q_1,\alpha_2,\theta_2)$, for all $q \in R_+$, i.e. that q_1 minimizes $\psi(q) \equiv u(q,\alpha_1,\theta_1) - u(q,\alpha_2,\theta_2)$. Now by assumption q_1 is a stationary point of ψ, and SCD implies that q_1 is a strict local minimum (see Lemma 1). In fact, single crossing implies that there is no other point $q \in R_+$ such that $\psi'(q) = 0$, i.e. that q_1 is a global minimum of ψ, as was to be demonstrated. *Q.E.D.*

Proof of Lemma 5: By monotonicity of Q^*, if $\theta' > \theta$ we would have $q \geq q(\alpha,\theta)$, a contradiction. Next, for any $\alpha' \geq \alpha$ let us define $\widetilde{\theta}(\alpha')$ so that type $(\alpha',\widetilde{\theta}(\alpha'))$ is indifferent between

q and $q(\alpha,\theta)$. Thus $\widetilde{\theta}(\alpha')$ solves the equation

$$u(q(\alpha,\theta),\alpha',\widetilde{\theta}) - u(q,\alpha',\widetilde{\theta}) = P(q(\alpha,\theta)) - P(q).$$

By the implicit function theorem, we have

$$\frac{d\widetilde{\theta}}{d\alpha'} = -\frac{u_\alpha(q(\alpha,\theta),\alpha',\widetilde{\theta}) - u_\alpha(q,\alpha',\widetilde{\theta})}{u_\theta(q(\alpha,\theta),\alpha',\widetilde{\theta}) - u_\theta(q,\alpha',\widetilde{\theta})} = -\frac{\int_q^{q(\alpha,\theta)} u_{q\alpha}(z,\alpha',\widetilde{\theta})dz}{\int_q^{q(\alpha,\theta)} u_{q\theta}(z,\alpha',\widetilde{\theta})dz}$$

It follows from Assumption 2 that

$$-\frac{u_{q\alpha}(q(\alpha,\theta),\alpha',\widetilde{\theta})}{u_{q\alpha}(q(\alpha,\theta),\alpha',\widetilde{\theta})} < \frac{d\widetilde{\theta}}{d\alpha'} < -\frac{u_{q\alpha}(q,\alpha',\widetilde{\theta})}{u_{q\alpha}(q,\alpha',\widetilde{\theta})}$$

Thus the iso-price line through the point (α,θ) at the quantity q is flatter than the curve $\widetilde{\theta}$, which is in turn flatter than the iso-price line through the point (α,θ) at the quantity $q(\alpha,\theta)$. It follows that for $\alpha' > \alpha$ types along the iso-price line through the point (α,θ) at the quantity q strictly prefer $q(\alpha,\theta)$ to q, so such types cannot belong to $I(q,\alpha,\theta)$. Similarly, types along the iso-price line through the point (α,θ) at the quantity $q(\alpha,\theta)$ with $\alpha' > \alpha$ strictly prefer q to $q(\alpha,\theta)$, so such types cannot belong to $I(\alpha,\theta)$. *Q.E.D.*

Proof of Lemma 6: We first show that Assumption 2 implies that $\frac{u_{q\alpha}}{u_{q\theta}} - \frac{u_\alpha}{u_\theta} > 0$ whenever $q > 0$. For fixed (θ,α) define $\varphi(q) = \frac{u_{q\alpha}}{u_{q\theta}}(q,\alpha,\theta) - \frac{u_\alpha}{u_\theta}(q,\alpha,\theta)$, so that $\varphi'(q) = \frac{d}{dq}\frac{u_{q\alpha}}{u_{q\theta}} - \varphi(q)(\frac{u_{q\theta}}{u_\theta})$. It then follows from Assumption 2 that for any $q > 0$ s.t. $\varphi(q) \leq 0$ we have $\varphi'(q) > 0$. Thus $\varphi(q) \leq 0$ for some $q > 0$ would imply $\phi(q') < \phi(q)$ for all $q' < q$, and in particular also that $\lim_{q'\to 0}\varphi(q') < 0$. But since u_α and u_θ both converge to zero as $q \to 0$, it follows from l'Hospital's rule that $\lim_{q\to 0}\varphi(q) = 0$, a contradiction. We conclude that we must have $\varphi(q) > 0$ whenever $q > 0$.

Next, let (α,θ) be s.t. $q(\alpha,\theta) > 0$. Let $(a,\widetilde{\theta}(a))$ parameterize the isoquant $I(\alpha,\theta)$, i.e. $\widetilde{\theta}(a)$ solves $u_q(q(\alpha,\theta),a,\widetilde{\theta}(a)) = u_q(q(\alpha,\theta),\alpha,\theta)$. Then we have

$$\widetilde{\theta}'(a) = -\frac{u_{q\alpha}(q(\alpha,\theta),a,\widetilde{\theta}(a))}{u_{q\theta}(q(\alpha,\theta),a,\widetilde{\theta}(a))} < 0.$$

By the envelope theorem $s((\alpha',\widetilde{\theta}(\alpha')) = s(\alpha,\theta) + \int_\alpha^{\alpha'}\{u_\theta(q(\alpha,\theta),a,\widetilde{\theta}(a))\widetilde{\theta}'(a) + u_\alpha(q(\alpha,\theta),a,\widetilde{\theta}(a))\}da$. Since $s(\alpha,\theta) \geq 0$, and since $q(\alpha,\theta) > 0$ and Assumption 2 imply that $\frac{u_\alpha}{u_\theta} - \frac{u_{q\alpha}}{u_{q\theta}} = -\varphi < 0$, we have $s((\alpha',\widetilde{\theta}(\alpha')) > 0$ for all $\alpha' < \alpha$. It follows that $\widetilde{\theta}(\alpha') > \underline{\theta}(a)$ for all $\alpha' < \alpha$. *Q.E.D.*

Proof of Lemma 3: It follows from the results of Rochet and Stole (2003) and Basov (2001) that the optimal allocation $q(\alpha,\theta)$ must satisfy an elliptical partial differential equation. It is well-known that solutions to elliptical partial differential equations on a domain with a piecewise smooth boundary are continuous on the interior of that domain. *Q.E.D.*

Proof of Theorem 1: (i) To prove absolute continuity of $\underline{\theta}(\alpha)$, observe that incentive compatibility implies

$$\begin{aligned} s(\alpha,\underline{\theta}(\alpha)) &= u(q(\alpha,\underline{\theta}(\alpha)),\alpha,\underline{\theta}(\alpha)) - T(q(\alpha,\underline{\theta}(\alpha))) \leq u(q(\alpha',\underline{\theta}(\alpha')),\alpha,\underline{\theta}(\alpha)) - T(q(\alpha',\underline{\theta}(\alpha'))) \\ s(\alpha',\underline{\theta}(\alpha')) &= u(q(\alpha',\underline{\theta}(\alpha')),\alpha',\underline{\theta}(\alpha')) - T(q(\alpha',\underline{\theta}(\alpha')) \leq u(q(\alpha,\underline{\theta}(\alpha)),\alpha',\underline{\theta}(\alpha')) - T(q(\alpha,\underline{\theta}(\alpha)) \end{aligned}$$

Consequently, we have

$$u(q(\alpha',\underline{\theta}(\alpha')),\alpha,\underline{\theta}(\alpha)) - u(q(\alpha',\underline{\theta}(\alpha')),\alpha',\underline{\theta}(\alpha')) \leq s(\alpha,\underline{\theta}(\alpha)) - s(\alpha',\underline{\theta}(\alpha')) \leq u(q(\alpha,\underline{\theta}(\alpha)),\alpha,\underline{\theta}(\alpha)) - u(q(\alpha,\underline{\theta}(\alpha)),$$

Using the mean value theorem, and the fact that $s(\alpha,\underline{\theta}(\alpha)) = s(\alpha',\underline{\theta}(\alpha')) = 0$, we obtain

$$\begin{aligned} u_\theta(q(\alpha',\underline{\theta}(\alpha')),\alpha_0,\underline{\theta}(\alpha_0))(\underline{\theta}(\alpha)-\underline{\theta}(\alpha')) + u_\alpha(q(\alpha',\underline{\theta}(\alpha')),\alpha_0,\underline{\theta}(\alpha_0))(\alpha-\alpha') &\leq 0 \quad (20) \\ u_\theta(q(\alpha,\underline{\theta}(\alpha)),\alpha_1,\underline{\theta}(\alpha_1))(\underline{\theta}(\alpha)-\underline{\theta}(\alpha')) + u_\alpha(q(\alpha,\underline{\theta}(\alpha)),\alpha_1,\underline{\theta}(\alpha_1))(\alpha-\alpha') &\geq 0 \quad (21) \end{aligned}$$

for some α_0 and α_1 between α and α'. If $\alpha < \alpha'$ then inequalities (20) and (21) imply

$$-\frac{u_\alpha}{u_\theta}(q(\alpha',\underline{\theta}(\alpha')),\alpha_0,\underline{\theta}(\alpha_0)) \leq \frac{\underline{\theta}(\alpha)-\underline{\theta}(\alpha')}{\alpha-\alpha'} \leq -\frac{u_\alpha}{u_\theta}(q(\alpha,\underline{\theta}(\alpha)),\alpha_1,\underline{\theta}(\alpha_1))$$

If $\alpha' < \alpha$ then inequalities (20) and (21) imply

$$-\frac{u_\alpha}{u_\theta}(q(\alpha,\underline{\theta}(\alpha)),\alpha_1,\underline{\theta}(\alpha_1)) \leq \frac{\underline{\theta}(\alpha)-\underline{\theta}(\alpha')}{\alpha-\alpha'} \leq -\frac{u_\alpha}{u_\theta}(q(\alpha',\underline{\theta}(\alpha')),\alpha_0,\underline{\theta}(\alpha_0))$$

Define $L = \max_{(q,\alpha,\theta)} \frac{u_\alpha}{u_\theta}(q,\alpha,\theta) < \infty$. The previous two inequalities show that $\underline{\theta}(\alpha)$ is Lipschitz continuous with Lipschitz constant L. It follows that $\underline{\theta}(\alpha)$ is absolutely continuous. Furthermore, at all continuity points of $q(\alpha,\underline{\theta}(\alpha))$ (which in part (ii) we shall show excludes all but at most a countable set of α) taking limits in the above two inequalities yields

$$\underline{\theta}'(\alpha) = -\frac{u_\alpha}{u_\theta}(q(\alpha,\underline{\theta}(\alpha)),\alpha,\underline{\theta}(\alpha)).$$

(ii) First, we shall argue that the correspondence Q^* is nondecreasing along L. To this effect, define an artificial type $\lambda \in [0, 2-\widehat{\theta}]$ along L such that $\lambda = \alpha$ if $\alpha < 1$ and $\lambda = 1+(\theta-\widehat{\theta})$ if $\alpha = 1$. Lemma 6 implies that $u_{q\lambda} = u_{q\theta}\underline{\theta}'(\alpha) + u_{q\alpha} > 0$ for $\lambda < 1$ and $q > 0$, and $u_{q\lambda} = u_{q\alpha} > 0$ for $\lambda \geq 1$. This supermodularity implies that every selection from $Q^*(\lambda)$ must be non-decreasing. Hence $Q^*(\lambda)$ is single-valued for almost all λ, and any selection from $Q^*(\lambda)$ is a non-decreasing function. Since changing the allocation on a set of measure zero of λ does not alter the monopolist's expected profits, we may select $\underline{q}(\alpha) = \min Q^*(\alpha,\underline{\theta}(\alpha))$ and $\underline{q}(\theta) = \min Q^*(1,\theta)$.

(iii) It follows from incentive compatibility that if $q \in Q^*(\alpha,\underline{\theta}(\alpha))$ and $q' \in Q^*(\alpha',\underline{\theta}(\alpha'))$ then:

$$u(q,\alpha,\underline{\theta}(\alpha)) - u(q',\alpha,\underline{\theta}(\alpha)) \leq P(q) - P(q') \leq u(q,\alpha',\underline{\theta}(\alpha')) - u(q',\alpha',\underline{\theta}(\alpha'))$$

Using the mean value theorem, we obtain:

$$u_q(z_0, \underline{\theta}(\alpha), \alpha)(q - q') \leq P(q) - P(q') \leq u_q(z_1, \underline{\theta}(\alpha'), \alpha')(q - q') \tag{22}$$

for some z_0 and $z_1$between q and q'. Let $M = \max_{(\theta,\alpha)\in[0,1]^2} u_q(0, \theta, \alpha)$. Then it follows from (22) that P is Lipschitz continuous with Lipschitz constant M, and hence absolutely continuous.

If $q \in [\underline{q}(0), \underline{q}(\widehat{\alpha})]$, then $q \in Q^*(\alpha, \underline{\theta}(\alpha))$ for some α. Hence we have $s(\alpha, \underline{\theta}(\alpha)) = 0$, implying $u(q, \alpha, \underline{\theta}(\alpha)) - P(q) = 0$.

Next, suppose that $\widehat{\alpha} < 1$, and $q \in [\underline{q}(\widehat{\alpha}), \underline{q}(1)]$. First, consider any $\alpha \in [\widehat{\alpha}, 1]$ at which $\underline{q}(\alpha)$ is discontinuous, so that $Q^*(\alpha, \underline{\theta}(\alpha))$ is multi-valued. Let $q_1 = \min Q^*(\underline{\theta}(\alpha), \alpha)$ and $q_2 = \max Q^*(\underline{\theta}(\alpha), \alpha)$. By (ii) we have $u(q, \alpha, \underline{\theta}(\alpha)) - P(q) = u(q_1, \alpha, \underline{\theta}(\alpha)) - P(q_1)$ for all $q \in [q_1, q_2]$, implying $u_q(q, \alpha, \underline{\theta}(\alpha)) = P'(q)$ for all $q \in (q_1, q_2)$. Next, consider any $\alpha \in [\widehat{\alpha}, 1]$ at which $\underline{q}(\alpha)$ is continuous and strictly increasing. Dividing (22) by $(q' - q)$ and taking limits as $\alpha' \downarrow \alpha$, it follows that $P'(q) = u_q(q, \alpha, \underline{\theta}(\alpha))$. We conclude that the Stieltjes integral $P(q) = u(\underline{q}(\widehat{\alpha}), \widehat{\alpha}, \underline{\theta}(\widehat{\alpha})) + \int_{\underline{q}(\widehat{\alpha})}^{q} u_q(z, \underline{\theta}(\alpha(z)), \alpha(z))dz$ holds. For $q \in (\underline{q}(1), \bar{q}(1)]$ the argument is analogous.

(iv) We show that Assumption 3 implies that $Q^*(\lambda)$ is convex for every λ. Suppose that $(\alpha, \theta) \in L$ is such that q_1, $q_2 \in Q^*(\alpha, \theta)$ with $q_1 < q_2$, and let $q \in (q_1, q_2)$. For $\theta' \in (\theta, \theta + \varepsilon)$, let α_1 and α_2 be such that $(\alpha_1, \theta') \in I(q_1, \alpha, \theta)$ and $(\alpha_2, \theta') \in I(q_2, \alpha, \theta)$; such values of α exists by Lemma 4. By Assumption 3 there exists $\alpha'(\theta') \in (\alpha_1, \alpha_2)$ such that $q = q(\alpha'(\theta'), \theta')$. Since this is true for all $\theta' \in (\theta, \theta + \varepsilon)$, it follows from u.h.c. of the correspondence Q^* that $q \in Q^*(\alpha, \theta)$.

(v) It follows from (ii) that the correspondence $w : L \to R$ defined by $w(\alpha, \theta) = \{u_q(q, \alpha, \theta) : q \in Q^*(\alpha, \theta)\}$ is convex valued and hence that $w(L)$ is a closed interval. Furthermore, we claim that for each $(\alpha, \theta) \notin L$ s.t. $s(\alpha, \theta) > 0$, there exists some $(\alpha', \theta') \in L$ and $q' \in Q^*(\alpha', \theta')$ such that $u_q(q', \alpha, \theta) = u_q(q', \alpha', \theta')$. Indeed, since $u_{q\theta} > 0$, we have $u_q(q(\alpha, \underline{\theta}(\alpha)), \alpha, \underline{\theta}(\alpha)) < u_q(q(\alpha, \underline{\theta}(\alpha)), \alpha, \theta)$, and since $u_{q\alpha} > 0$, we have $u_q(q(1, \theta), 1, \theta)) > u_q(q(1, \theta), \alpha, \theta)$. Hence there exists (α', θ') on the segment of L connecting $(\alpha, \underline{\theta}(\alpha))$ to $(1, \theta)$ and $q' \in Q^*(\alpha', \theta')$ such that $u_q(q', \alpha, \theta) = u_q(q', \alpha', \theta')$. Finally, (α, θ) can lie on at most one isoquant emanating from L, for otherwise isoquants would intersect in the interior of the participation region, violating Lemma 5. *Q.E.D.*

Proof of Theorem 2: First, we establish that the allocation $q(\alpha, \theta)$ is incentive compatible along L. It follows from (ii) that $q(\alpha, \underline{\theta}(\alpha)) = \underline{q}(\alpha)$ for all $\alpha \in [0, 1]$, and $q(1, \theta) = \bar{q}(\theta)$ for all $\theta \in [\widehat{\theta}, 1]$. Hence the allocation is nondecreasing along L. Since $u_{q\lambda} > 0$, it follows that the allocation is incentive compatible along L. Lemma 4 then implies that $q(\alpha, \theta)$ is incentive compatible for all (α, θ) in the participation region. It remains to be shown that $Q^*(\alpha, \theta) = \{0\}$ for all (α, θ) such that $\theta < \underline{\theta}(\alpha)$. Note that for any $q > 0$, we have $u(q, \theta, \alpha) - P(q) < u(q, \underline{\theta}(\alpha), \alpha) - P(q) \leq s(\underline{\theta}(\alpha), \alpha)) = 0$. Thus for any such type it is uniquely optimal to select $q = 0$. That (iii) holds follows from the proof of part (iii) of Theorem 1.. *Q.E.D.*

Proof of Theorem 3: It follows from Theorem 2 that

$$T^{**}(q) = \{(\alpha, \theta) \in [0,1]^2 : \theta \geq \underline{\theta}(\alpha) \text{ and } \theta \geq \sigma(q, \alpha(q), \theta(q), \alpha)\}$$

Using (3) then yields

$$\mu(q) = 1 - \int_0^1 \int_{\max\{\underline{\theta}(a), \sigma(q,\alpha(q),\theta(q),a)}^{\infty} f(a,\theta) d\theta da,$$

Thus over the interval $[0, \underline{q}(\widehat{\alpha})]$ is absolutely continuous, and

$$\mu'(q) = \int_0^{\alpha(q)} f(\sigma(q, \alpha(q), \theta(q), a) \frac{d}{dq} \sigma(q, \alpha(q), \theta(q), a) da = h(q, \alpha(q), \theta(q), \alpha'(q), \theta'(q)).$$

Moreover, over the interval $[\underline{q}(\widehat{\alpha}), \bar{q}(1)]$ we have

$$\mu(q) = 1 - \int_0^1 \int_{\sigma(q,\alpha(q),\theta(q),a)}^{\infty} f(a,\theta) d\theta da = 1 - H(q, \alpha(q), \theta(q)).$$

Using (6), we may re-write the monopolist's profits as

$$\int_0^{\underline{q}(\widehat{\alpha})} u(q, \alpha(q), \theta(q)) \mu'(q) dq + \int_{\underline{q}(\widehat{\alpha})}^{\bar{q}(1)} \left\{ u(\underline{q}(\widehat{\alpha}), \widehat{\alpha}, \widehat{\theta}) + \int_{\underline{q}(\widehat{\alpha})}^{q} u_q(z, \alpha(z), \theta(z)) dz \right\} d\mu(q)$$

Integrating by parts, and using the fact that $\mu(\bar{q}(1)) = 1$, the second integral may be re-written as

$$H(\underline{q}(\widehat{\alpha}), \widehat{\alpha}, \widehat{\theta}) u(\underline{q}(\widehat{\alpha}), \widehat{\alpha}, \widehat{\theta}) + \int_{\underline{q}(\widehat{\alpha})}^{\bar{q}(1)} H(q, \alpha(q), \theta(q)) u_q(q, \alpha(q), \theta(q)) dq$$

Q.E.D.

Remark: We may compute

$$\sigma_q(q, \alpha, \theta, a) = \frac{u_{qq(}q, a, \sigma(q, \alpha, \theta, a)) - u_{qq}(q, \alpha, \theta)}{u_{q\theta}(q, a, \sigma(q, \alpha, \theta, a))}$$

$$\sigma_\alpha(q, \alpha, \theta, a) = \frac{u_{q\alpha}(q, \alpha, \theta)}{u_{q\theta}(q, a, \sigma(q, \alpha, \theta, a))}$$

$$\sigma_\theta(q, \alpha, \theta, a) = \frac{u_{q\theta}(q, \alpha, \theta)}{u_{q\theta}(q, a, \sigma(q, \alpha, \theta, a))}$$

It follows from Lemma 1 that $\sigma_q(q, \alpha(q), \theta(q), a) > 0$. Also, using (5) and the fact that $u_{q\alpha} - \frac{u_\alpha}{u_\theta} u_{q\theta} > 0$ (see the proof of Lemma 6) we have

$$\sigma_\theta \theta'(q) + \sigma_\alpha \alpha'(q) = \frac{u_{q\alpha}(q, \alpha(q), \theta(q)) - \frac{u_\alpha}{u_\theta}(q, \alpha(q), \theta(q)) u_{q\theta}(q, \alpha(q), \theta(q))}{u_{q\theta}(q, a, \sigma(q, \alpha(q), \theta(q)), a))} > 0$$

We conclude that $\frac{d}{dq}\sigma(q,\alpha(q),\theta(q),a) > 0$.
Proof of Theorem 4: Define

$$F(q,\alpha,\theta) = u_q(q,\alpha,\theta)H(q,\alpha,\theta).$$

First, let us consider the case where $\widehat{\alpha} = 1$. The first order condition for (8) then becomes

$$F_\theta - \frac{d}{dq}F_{\theta'} = \phi(q,\theta) = 0$$

If $\phi_q > 0$ and $\phi_\theta < 0$, then application of the implicit function theorem to the equation $\phi(q,\theta) = 0$ establishes the existence and strict monotonicity of the function $\theta^\phi(q)$. If $\theta^\phi(\widehat{q}) > \widehat{\theta}$ then the boundary condition $\theta(\widehat{q}) = \widehat{\theta}$ is necessarily violated. If $\theta^\phi(\widehat{q}) < \widehat{\theta}$, then we must set $\theta(q) = \widehat{\theta}$ for all q such that $\theta^\phi(\widehat{q}) \le \widehat{\theta}$.

Next, let us consider the case where $\widehat{\alpha} < 1$ so that $\widehat{\theta} = 0$. Analogously to the proof in the previous paragraph, we may establish the following. First, we must have $\theta^\phi(\underline{q}(1)) \le 0$, and that $\theta(q) = \max\{\theta^\phi(q), 0\}$ for all $q \in [\widehat{q}, \underline{q}(1)]$. Second, we must have $\alpha^\varkappa(\widehat{q}) \le \widehat{\alpha}$ and $\alpha^\varkappa(\underline{q}(1)) \ge 1$, and on the interval $[\widehat{q}, \underline{q}(1)]$ we have $\alpha(q) = \max\{\alpha^\varkappa(q), \widehat{\alpha}\}$ whenever $\alpha^\varkappa(q) \le 1$, and $\alpha(q) = 1$, otherwise. It remains to be shown that $\underline{q}(1)$ must be chosen so that $\alpha^\varkappa(\underline{q}(1)) = 1$, or equivalently that $\theta^\phi(\underline{q}(1)) = 0$. We shall establish below that $\phi(\widetilde{q}, 0) = 0$ if and only if $\varkappa(\widetilde{q}, 1) = 0$. This will imply the desired result, since $\theta^\phi(\underline{q}(1)) \le 0$ requires $\underline{q}(1) \le \widetilde{q}$, and since $\underline{q}(1) < \widetilde{q}$ would imply $\alpha^\varkappa(\underline{q}(1)) < 1$ and hence $\alpha(\underline{q}(1)) < 1$.

We will now show that $H_\theta(\widetilde{q},1,0) = \frac{u_{q\theta}}{u_{q\alpha}}(\widetilde{q},1,0)H_\alpha(\widetilde{q},1,0)$, so that $\phi(\widetilde{q},0) = \frac{u_{q\theta}}{u_{q\alpha}}(\widetilde{q},1,0)\varkappa(\widetilde{q},1)$, implying the desired result. Observe that

$$H(q,1,\theta) = \int_{\alpha_-(q,\theta)}^1 \int_{\sigma(q,1,\theta,a)}^1 f(a,\theta')d\theta' da \tag{23}$$

where $\alpha_-(q,\theta)$ is the solution in a to the equation $\sigma(q,1,\theta,a) = 1$ if a nonnegative such a solution exists, and $\alpha_-(q,\alpha) = 0$, otherwise. Hence

$$H_\theta(q,1,\theta) = -\int_{\alpha_-(q,\theta)}^1 f(a,\sigma(q,1,\theta,a)\sigma_\theta(q,1,\theta,a)da \tag{24}$$

Observe also that

$$H(q,\alpha,0) = \int_{\alpha_-(q,\theta)}^\alpha \int_{\sigma(q,1,\theta,a)}^1 f(a,\theta')d\theta' da + \int_\alpha^1 \int_0^1 f(a,\theta)d\theta da \tag{25}$$

where, with some abuse in notation, $\alpha_-(q,\alpha)$ is the solution in a to the equation $\sigma(q,\alpha,0,a) = 1$, if a nonnegative such a solution exists, and $\alpha_-(q,\alpha) = 0$, otherwise. Hence

$$H_\alpha(q,\alpha,0) = -\int_{\alpha_-(q,\theta)}^1 f(a,\sigma(q,\alpha,0,a)\sigma_\alpha(q,\alpha,0,a)da \tag{26}$$

Finally, application of the Implicit Function Theorem to the defining equation $u_q(q,a,\sigma) = u_q(q,\alpha,\theta)$ yields

$$\sigma_\alpha(q,\alpha,\theta,a) = \frac{u_{q\alpha}(q,\alpha,\theta)}{u_{q\theta}(q,a,\sigma)} \tag{27}$$

$$\sigma_\theta(q,\alpha,\theta,a) = \frac{u_{q\theta}(q,\alpha,\theta)}{u_{q\theta}(q,a,\sigma)} \tag{28}$$

Combining (24), (26), (27) and (28) then yields $H_\theta(\widetilde{q},1,0) = \frac{u_{q\theta}}{u_{q\alpha}}(\widetilde{q},1,0)H_\alpha(\widetilde{q},1,0)$. *Q.E.D.*

Proof of Corollary 5.1: On the interval $[\underline{q}(1),\bar{q}(1)]$ form the Lagrangian

$$L(q,1,\theta,\theta') = F(q,1,\theta) + \lambda\theta'$$

The first-order conditions associated with this variational problem are:

$$\begin{aligned} L_\theta - \frac{d}{dq}L_{\theta'} &= \phi(q,\theta) - \lambda' = 0 \\ \lambda(q)\theta'(q) &= 0 \\ \lambda(q) &\geq 0 \\ \theta'(q) &\geq 0 \end{aligned} \tag{29}$$

In addition, the transversality condition for the free 'terminal time' $\bar{q}(1)$ is :

$$L - L_{\theta'}\theta' = \phi(\bar{q}(1),1) = 0$$

Since $H(\bar{q}(1),1,1) = 0$, and since $H_\theta(\bar{q}(1),1,1) < 0$, the transversality condition yields $u_q(\bar{q}(1),1,1) = 0$. From (29), we have $\lambda'(q) = F_\alpha(q,\alpha(q),0) = \varkappa(q,\alpha(q))$. The proof for $\nu'(q)$ is analogous. *Q.E.D.*

Proof of Theorem 5: As a preliminary to the proof, note that (23), (24 and (28) imply $\phi(q,\theta) = u_{q\theta}(q,1,\theta)v(q,\theta)$, where

$$v(q,\theta) = -u_q(q,1,\theta)\int_{\alpha_-(q,\theta)}^{1} \frac{f(a,\sigma(q,1,\theta,a)}{u_{q\theta}(q,a,\sigma(q,1,\theta,a))}da + H(q,1,\theta)$$

Also, it follows from (25), (26) and(27) that $\varkappa(q,\alpha) = u_{q\alpha}(q,\alpha,0)\rho(q,\alpha)$, where

$$\varrho(q,\alpha) = -u_q(q,\alpha,0)\int_{\alpha_-(q,\alpha)}^{\alpha} \frac{f(a,\sigma(q,\alpha,0,a))}{u_{q\theta}(q,a,\sigma(q,\alpha,0,a))}da + H(q,\alpha,0)$$

First, we argue necessity of condition (10). Suppose first that $\rho(0,\mathring{\alpha}) = 0$ for some $\mathring{\alpha} \in (0,1]$ but $u_{q\alpha}(0,\alpha,0) > 0$ for some $\alpha \in (0,\mathring{\alpha})$. Since $\sigma_\alpha(0,\alpha,0,a) = u_{q\alpha}(0,\alpha,0)/u_{q\theta}(0,a,\sigma(0,\alpha,0,a))$, it then follows that $\sigma(0,\alpha,0,0) > 0$. Thus a positive measure of types must receive a zero quantity. Next, suppose that $u_{q\alpha}(0,\alpha,0) = 0$ for all α, but there exists no $\mathring{\alpha} \in (0,1]$ such that $\rho(0,\alpha)=0$. Since $\rho(0,0)=1$, we must then have $\rho(0,\alpha)>0$ for all α. Because $\phi(0,0)=u_{q\theta}(0,1,0)\rho(0,1)>0$, and because by assumption we have $\phi_\theta<0$, it then follows that $\phi(0,\theta(0))=0$ for some $\theta(0)>0$. Therefore all consumer types (α,θ) with $\theta<\theta(0)$ are excluded.

Next, we will argue that the if the conditions of the Lemma hold, then the associated mechanism is optimal, and no consumer is excluded. Let the monopolist select $(\widehat{\alpha},\widehat{\theta})=(\mathring{\alpha},0)$ and $\widehat{q}=0$. With these choices, it follows from Theorem 4 that $\underline{q}(1)$ is the solution to the equation $\varkappa(q,1)=0$, that on the interval $[0,\underline{q}(1)]$ the solution $\alpha(q)$ solves $\varkappa(q,\alpha)=0$, and that on the interval $[\underline{q}(1),\bar{q}(1)$ the solution $\theta(q)$ solves $\phi(q,\theta)=0$. This mechanism has no exclusion, because the isoquant emanating from the point $(\mathring{\alpha},0)$ is flat, i.e. $\sigma_a(0,\mathring{\alpha},0,a)=-\frac{u_{q\alpha}(0,a,0)}{u_{q\theta}(0,a,0)}=0$. *Q.E.D.*

As a preliminary to the proof of Theorem 6 we prove the following lemma:

Lemma 10 *We have*

$$h_\alpha-\frac{d}{dq}h_{\alpha'} = f(\underline{\theta},\alpha)\sigma_{\underline{\theta}}(q,\alpha,\underline{\theta},a)\underline{\theta}'(q) \tag{30}$$

$$h_{\underline{\theta}}-\frac{d}{dq}h_{\underline{\theta}'} = -f(\underline{\theta},\alpha)\sigma_{\underline{\theta}}(q,\alpha,\underline{\theta},a)\alpha'(q) \tag{31}$$

Proof : Since

$$h(q,\alpha,\underline{\theta},\alpha',\underline{\theta}')=\int_{\underline{\alpha}(q,\underline{\theta},\alpha)}^{\alpha} f(\sigma(q,\alpha,\underline{\theta},a),a)\left\{\sigma_q+\sigma_{\underline{\theta}}\underline{\theta}'+\sigma_\alpha\alpha'\right\}da. \tag{32}$$

we have

$$h_{\underline{\theta}}=-\underline{\alpha}_{\underline{\theta}}fs_q|_{a=\underline{\alpha}}+\int_{\underline{\alpha}(q)}^{\alpha} f_\theta\sigma_{\underline{\theta}}s_q da+\int_{\underline{\alpha}(q)}^{\alpha} f\left\{\sigma_{q\underline{\theta}}+\sigma_{\underline{\theta}\underline{\theta}}\theta'+\sigma_{\alpha\underline{\theta}}\alpha'\right\}da,$$

$$h_{\underline{\theta}'}(q,\underline{\theta},\alpha,\underline{\theta}',\alpha')=\int_{\underline{\alpha}(q,\underline{\theta},\alpha)}^{\alpha} f(\sigma(q,\alpha,\underline{\theta},a),a)\sigma_{\underline{\theta}}(q,\alpha,\underline{\theta},a)da,$$

and

$$\frac{d}{dq}h_{\underline{\theta}'}=\alpha'f\sigma_{\underline{\theta}}\,|_{a=\alpha}-\left\{\underline{\alpha}_q+\underline{\alpha}_{\underline{\theta}}\underline{\theta}'+\underline{\alpha}_\alpha\alpha'\right\}f\sigma_{\underline{\theta}}\,|_{a=\underline{\alpha}}+\int_{\underline{\alpha}(q)}^{\alpha} f_\theta s_q\sigma_{\underline{\theta}}da+\int_{\underline{\alpha}(q)}^{\alpha} f\left\{\sigma_{q\underline{\theta}}+\sigma_{\underline{\theta}\underline{\theta}}\underline{\theta}'+\sigma_{\alpha\underline{\theta}}\alpha'\right\}$$

so

$$h_{\underline{\theta}}-\frac{d}{dq}h_{\underline{\theta}'}=-\alpha'f\sigma_{\underline{\theta}}\,|_{a=\alpha}+\left\{[\underline{\alpha}_q+\underline{\alpha}_\alpha\alpha']\sigma_{\underline{\theta}}-\underline{\alpha}_\theta[\sigma_q+\sigma_\alpha\alpha']\right\}f.$$

Now

$$[\underline{\alpha}_q+\underline{\alpha}_\alpha\alpha']\sigma_{\underline{\theta}}-\underline{\alpha}_\theta[\sigma_q+\sigma_\alpha\alpha'=0.$$

This is obvious over any interval of q where $\underline{\alpha}=0$. Over any interval where $\underline{\alpha}>0$ it is defined as the solution in a to the equation $\sigma(q,\underline{\theta},\alpha,a)=1$. Thus we have $\sigma_q+\sigma_a\underline{\alpha}_q=0$ and

$\sigma_{\underline{\theta}} + \sigma_a \underline{\alpha}_{\underline{\theta}} = 0$, implying $\underline{\alpha}_q \sigma_{\underline{\theta}} - \underline{\alpha}_{\underline{\theta}} \sigma_q = 0$. Furthermore, we have $\sigma_\alpha + \sigma_a \underline{\alpha}_\alpha = 0$, so $\underline{\alpha}_\alpha \sigma_{\underline{\theta}} - \underline{\alpha}_{\underline{\theta}} \sigma_\alpha = 0$. Thus (31) holds. Next, we have

$$h_\alpha = f s_q \mid_{a=\alpha} - \underline{\alpha}_\alpha f s_q \mid_{a=\underline{\alpha}} + \int_{\underline{\alpha}(q)}^{\alpha} f_\theta \sigma_\alpha s_q da + \int_{\underline{\alpha}(q)}^{\alpha} f \left\{ \sigma_{q\alpha} + \sigma_{\underline{\theta}\alpha} \underline{\theta}' + \sigma_{\alpha\alpha} \alpha' \right\} da,$$

$$h_{\alpha'} = \int_{\underline{\alpha}(q,\underline{\theta},\alpha)}^{\alpha} f(\sigma(q, \underline{\theta}, \alpha, a)) \sigma_\alpha da, \text{ and}$$

$$\frac{d}{dq} h_{\alpha'} = f \sigma_\alpha \alpha' \mid_{a=\alpha} - f \sigma_\alpha \underline{\alpha}' \mid_{a=\underline{\alpha}} + \int_{\underline{\alpha}(q)}^{\alpha} f_\theta s_q \sigma_\alpha da + \int_{\underline{\alpha}(q)}^{\alpha} f \left\{ \sigma_{\alpha q} + \sigma_{\underline{\theta}\alpha} \underline{\theta}' + \sigma_{\alpha\alpha} \alpha' \right\} da.$$

Hence

$$\begin{aligned} h_\alpha - \frac{d}{dq} h_{\alpha'} &= \left\{ s_q - \sigma_\alpha \alpha' \right\} f \mid_{a=\alpha} + f \left\{ \sigma_\alpha \underline{\alpha}' - \underline{\alpha}_\alpha s_q \right\} \mid_{a=\underline{\alpha}} \\ &= \left\{ \sigma_q + \sigma_{\underline{\theta}} \underline{\theta}' \right\} f \mid_{a=\alpha} - f \underline{\alpha}_\alpha \left\{ \sigma_{\underline{\alpha}} \underline{\alpha}' + s_q \right\} \mid_{a=\underline{\alpha}} \\ &= \left\{ \sigma_q + \sigma_{\underline{\theta}} \underline{\theta}' \right\} f \mid_{a=\alpha} \end{aligned}$$

where we used $s_q = \sigma_q + \sigma_{\underline{\theta}} \underline{\theta}' + \sigma_\alpha \alpha'$ and the fact that $s(q, \underline{\alpha}(q)) = 1$ implies $s_q + \sigma_{\underline{\alpha}} \underline{\alpha}' = 0$. The identity $\sigma(q, \underline{\theta}, \alpha, \alpha) = \underline{\theta}$ implies $\sigma_q(q, \underline{\theta}, \alpha, \alpha)$, so we conclude that (30) holds. *Q.E.D.*

Proof of Theorem 6: We will start by showing that the solution to problem 11 satisfies

$$\xi(q) = -\psi(q, \alpha, \underline{\theta}) - \frac{\dot{\chi}}{g_q}. \tag{33}$$

The derivation of this equation is long and intricate. Maximization of the Lagrangian L w.r.t. v and w yields

$$\begin{aligned} \frac{\partial F}{\partial \alpha'}(q, \alpha, \underline{\theta}, \alpha', \underline{\theta}') + \mu + \xi g + \chi &= 0 \\ \frac{\partial F}{\partial \underline{\theta}'}(q, \alpha, \underline{\theta}, \alpha', \underline{\theta}') + \lambda + \xi &= 0 \end{aligned}$$

and so we have

$$\begin{aligned} \mu &= -\frac{\partial F}{\partial \alpha'}(q, \alpha, \underline{\theta}, \alpha', \underline{\theta}') - \xi g - \chi \\ \lambda &= -\frac{\partial F}{\partial \underline{\theta}'}(q, \alpha, \underline{\theta}, \alpha', \underline{\theta}') - \xi \end{aligned} \tag{34}$$

Differentiating (34) w.r.t. q then yields

$$\begin{aligned} \dot{\mu} + \dot{\chi} + \frac{d}{dq} \xi g &= -\frac{d}{dq} \frac{\partial F}{\partial \alpha'} \\ \dot{\lambda} + \dot{\xi} &= -\frac{d}{dq} \frac{\partial F}{\partial \underline{\theta}'} \end{aligned} \tag{35}$$

Multiplying the second equation in (35) by g, and subtracting from the first yields

$$\dot{\mu} + \dot{\chi} + \xi \dot{g} - \dot{\lambda} g = -\frac{d}{dq}\frac{\partial F}{\partial \alpha'} + g\frac{d}{dq}\frac{\partial F}{\partial \underline{\theta}'}$$

Substituting in the costate equations

$$\begin{aligned} \dot{\mu} &= -\frac{\partial F}{\partial \alpha} - \xi g_{\alpha} \alpha' \\ \dot{\lambda} &= -\frac{\partial F}{\partial \underline{\theta}} - \xi g_{\underline{\theta}} \alpha', \end{aligned} \tag{36}$$

and simplifying, we obtain

$$\dot{\chi} + \xi g_q = (\frac{\partial F}{\partial \alpha} - \frac{d}{dq}\frac{\partial F}{\partial \alpha'}) - g(\frac{\partial F}{\partial \underline{\theta}} - \frac{d}{dq}\frac{\partial F}{\partial \underline{\theta}'}) \tag{37}$$

Now

$$\begin{aligned} & (\frac{\partial F}{\partial \alpha} - \frac{d}{dq}\frac{\partial F}{\partial \alpha'}) - g(\frac{\partial F}{\partial \underline{\theta}} - \frac{d}{dq}\frac{\partial F}{\partial \underline{\theta}'}) \\ = \;& (u_\alpha h + u h_\alpha) - \frac{d}{dq}(u h_v) - g(u_\theta h + u h_\theta) + g\frac{d}{dq}(u h_w) \\ = \;& h(u_\alpha - g u_\theta) + u(h_\alpha - \frac{d}{dq}h_{\alpha'}) + u_q h_{\alpha'} - g u(h_\theta - \frac{d}{dq}h_{\theta'}) + g u_q h_{\theta'} \\ = \;& u(h_\alpha - \frac{d}{dq}h_{\alpha'}) - g u(h_\theta - \frac{d}{dq}h_{\theta'}) - u_q(h_{\alpha'} - g h_{\theta'}) \\ = \;& u f \sigma_{\underline{\theta}}(\underline{\theta}' + g\alpha') - u_q(h_{\alpha'} - g h_{\theta'}) \\ = \;& -u_q(h_{\alpha'} - g h_{\theta'}) \\ = \;& -u_q \int_{\underline{\alpha}(q,\underline{\theta},\alpha)}^{\alpha} f(\sigma(q,\alpha,\underline{\theta},a),a) \left\{ \sigma_\alpha(q,\alpha,\underline{\theta},a) - \sigma_{\underline{\theta}}(q,\alpha,\underline{\theta},a) g \right\} da \\ = \;& -u_q \left\{ \frac{u_{q\alpha}}{u_{q\theta}} - \frac{u_\alpha}{u_\theta} \right\} h_1 \end{aligned} \tag{38}$$

where the fourth equality follows from (30) and (31), and the last equality follows from the fact that

$$\begin{aligned} \sigma_{\underline{\theta}}(q,\underline{\theta},\alpha,a) &= \frac{u_{q\theta}(q,\underline{\theta},\alpha)}{u_{q\theta}(q,\sigma,a)} \\ \sigma_{\alpha}(q,\underline{\theta},\alpha,a) &= \frac{u_{q\alpha}(q,\underline{\theta},\alpha)}{u_{q\theta}(q,\sigma,a)}. \end{aligned}$$

Now we may compute

$$g_q = \frac{d}{dq}\frac{u_\alpha}{u_\theta} = \frac{u_{\alpha q}}{u_\theta} - \frac{u_{\theta q} u_\alpha}{u_\theta^2} = \frac{u_{q\theta}}{u_\theta}\{\frac{u_{q\alpha}}{u_{q\theta}} - \frac{u_\alpha}{u_\theta}\}. \tag{39}$$

Substituting (39) and (38) into (37) then yields the desired formula for $\xi(q)$.

(i) Over any non-degenerate interval on which $\alpha(q)$ is strictly increasing, we have $\chi(q) = 0$ and hence $\dot{\chi}(q) = 0$, so (33) becomes

$$\xi(q) = -\psi(q, \alpha, \underline{\theta}). \tag{40}$$

Furthermore, substituting this expression for ξ in (**??**) we obtain

$$\begin{aligned} \mu(q) &= -uh_2 + \psi g \\ \lambda(q) &= -uh_1 + \psi \end{aligned}$$

Next, we derive the expression for $\alpha'(q)$. Equating the two expressions for $\dot{\lambda}$ in (35) and (36) yields

$$(\frac{\partial F}{\partial \underline{\theta}} - \frac{d}{dq}\frac{\partial F}{\partial \underline{\theta}'}) + \xi g_\theta \alpha' - \dot{\xi} = 0$$

Using

$$\frac{\partial F}{\partial \underline{\theta}} - \frac{d}{dq}\frac{\partial F}{\partial \underline{\theta}'} = u_\theta h + u(h_{\underline{\theta}} - \frac{d}{dq} h_{\underline{\theta}'}) - u_q h_{\underline{\theta}'}$$

and (31) then yields

$$u_\theta h - uf\sigma_{\underline{\theta}}\alpha' + \xi g_\theta \alpha' - u_q h_1 - \dot{\xi} = 0 \tag{41}$$

Finally, since $\xi(q) = -\psi(q, \alpha, \underline{\theta})$, we have

$$\dot{\xi} = -\psi_q - \psi_{\underline{\theta}}\underline{\theta}' - \psi_\alpha \alpha'. \tag{42}$$

Substituting (40), (42) and $h = h_0 + h_1\underline{\theta}' + h_2\alpha'$ into (41), and solving for α' then yields

$$\{u_\theta h_0 - u_q h_1 + \psi_q\} - \alpha'\{uf\sigma_{\underline{\theta}} + \psi g_\theta + \psi_{\underline{\theta}} g - \psi_\alpha - u_\theta(h_2 - gh_1)\} = 0$$

(ii) Over any interval on which the constraint $\alpha' \geq 0$ binds, we have

$$\dot{\chi} = -(\psi(q, \alpha, \underline{\theta}) + \xi(q))g_q \tag{43}$$

Furthermore, it follows from (41) that

$$u_\theta h_0 - u_q h_1 - \dot{\xi} = 0 \tag{44}$$

(iii) Let $\widetilde{H}(q) = H(q, \alpha(q), \underline{\theta}(q), v(q), w(q), \lambda(q), \mu(q))$. The transversality condition associated with the free left hand 'time' $\underline{q}(0)$ is $\widetilde{H}(\underline{q}(0) = 0$. Since $F(q, \alpha, \underline{\theta}, v, w) = u(q, \alpha, \underline{\theta})\{h_0(q, \alpha, \underline{\theta}) + h_1(q, \alpha, \underline{\theta})v + h_2(q, \alpha, \underline{\theta})w\}$, it follows from (34) that $\widetilde{H}(q) = u(q, \alpha(q), \underline{\theta}(q))h_0(q, \alpha(q), \underline{\theta}(q))$. Hence we must either have $u(\underline{q}(0), \alpha(\underline{q}(0)), \underline{\theta}(\underline{q}(0))) = 0$, implying $\underline{q}(0) = 0$, or $h_0(\underline{q}(0), \alpha(\underline{q}(0)), \underline{\theta}(\underline{q}(0))) = 0$ implying $\alpha(\underline{q}(0)) = 0$. Q.E.D.

Proof of Lemma 8: (i) First, we will prove that $\alpha(\underline{q}(0)) > 0$. If $\underline{\alpha}(q, \alpha(q), \underline{\theta}(q)) > 0$ for all q in a right neighborhood of $\underline{q}(0)$, then the result is immediate, since $\alpha(\underline{q}(0)) > \underline{\alpha}(q, \alpha(q), \underline{\theta}(q)) \geq 0$. Hence we may assume that there exists $\varepsilon > 0$ such that $\underline{\alpha}(q, \alpha(q), \underline{\theta}(q)) = 0$ for all

$q \in [\underline{q}(0), \underline{q}(0)+\varepsilon)$. We shall argue that in this case there exists $\breve{\alpha} > 0$ such that for any $q \in (\underline{q}(0), \underline{q}(0)+\varepsilon)$ and $\alpha < \breve{\alpha}$ have $\alpha'(q) < 0$, thereby implying that $\alpha(\underline{q}(0)) \geq \breve{\alpha}$.

To prove the existence of such an $\breve{\alpha} > 0$, we must first compute the numerator and denominator of α' in some more detail. We claim that:

$$\begin{aligned}\text{Numerator (14)} &= u_\theta \int_0^\alpha f(a,\sigma)\frac{2u_{qq}(q,\alpha,\underline{\theta}) - u_{qq}(q,a,\sigma)}{u_{q\theta}(q,a,\sigma)}da + \\ &\quad u_\theta u_q \int_0^\alpha \frac{f_\theta u_{q\theta} - f u_{q\theta\theta}}{u_{q\theta}^2}\frac{u_{qq}(q,\underline{\theta},\alpha) - u_{qq}(q,\sigma,a)}{u_{q\theta}(q,\sigma,a)}da\end{aligned} \tag{45}$$

$$\begin{aligned}\text{Denominator (14)} &= \left(u - \frac{u_\theta u_q}{u_{q\theta}}\right) f + 2(u_{q\theta}u_\alpha - u_{q\alpha}u_\theta)\int_0^\alpha \frac{f(a,\sigma)}{u_{q\theta}(q,a,\sigma)}da \\ &\quad + u_q(u_{q\theta}u_\alpha - u_{q\alpha}u_\theta)\int_0^\alpha \frac{f_\theta u_{q\theta} - f u_{q\theta\theta}}{u_{q\theta}^3}(q,a,\sigma)da\end{aligned} \tag{46}$$

To establish (45), observe that

$$\sigma_q(q,\alpha,\underline{\theta},a) = \frac{u_{qq}(q,\alpha,\underline{\theta}) - u_{qq}(q,a,\sigma)}{u_{q\theta}(q,a,\sigma)}$$

and

$$\sigma_{\underline{\theta}}(q,\alpha,\underline{\theta},a) = \frac{u_{q\theta}(q,\alpha,\underline{\theta})}{u_{q\theta}(q,a,\sigma)}$$

$\underline{\alpha}(q,\alpha(q),\underline{\theta}(q)) = 0$ imply that

$$\begin{aligned} h_0(q,\underline{\theta}(q),\alpha(q)) &= \int_0^\alpha f(\sigma(q,\alpha,\underline{\theta},a),a)\frac{u_{qq}(q,\alpha,\underline{\theta}) - u_{qq}(q,a,\sigma(q,\alpha,\underline{\theta},a))}{u_{q\theta}(q,a,\sigma(q,\alpha,\underline{\theta},a))}da && (47) \\ h_1(q,\underline{\theta}(q),\alpha(q)) &= \int_0^\alpha f(\sigma(q,\alpha,\underline{\theta},a),a)\frac{u_{q\theta}(q,\alpha,\underline{\theta})}{u_{q\theta}(q,a,\sigma(q,\alpha,\underline{\theta},a))}da && (48)\end{aligned}$$

Furthermore, since

$$\psi(q,\alpha,\underline{\theta}) = u_q(q,\alpha,\underline{\theta})u_\theta(q,\alpha,\underline{\theta})\int_0^\alpha \frac{f(a,\sigma)}{u_{q\theta}(q,a,\sigma)}da$$

we have

$$\psi_q = \left\{\begin{array}{c} u_{qq}(q,\alpha,\underline{\theta})u_\theta(q,\alpha,\underline{\theta})\int_{\underline{\alpha}(q,\alpha,\underline{\theta})}^\alpha \frac{f(a,\sigma)}{u_{q\theta}(q,a,\sigma)}da \\ +u_q(q,\alpha,\underline{\theta})u_{q\theta}(q,\alpha,\underline{\theta})\int_{\underline{\alpha}(q,\alpha,\underline{\theta})}^\alpha \frac{f(a,\sigma)}{u_{q\theta}(q,a,\sigma)}da \\ +u_q(q,\alpha,\underline{\theta})u_\theta(q,\alpha,\underline{\theta})\int_0^\alpha \frac{f_\theta u_{q\theta} - f u_{q\theta\theta}}{u_{q\theta}^2}\frac{u_{qq}(q,\underline{\theta},\alpha) - u_{qq}(q,\sigma,a)}{u_{q\theta}(q,\sigma,a)}da \end{array}\right\} \tag{49}$$

Using (47), (48) and (49) then yields (45). To establish (46) note that:

$$\psi g_\theta = u_q(u_{\alpha\theta} - \frac{u_\alpha}{u_\theta}u_{\theta\theta})\int_0^\alpha \frac{f(a,\sigma)}{u_{q\theta}(q,a,\sigma)}da$$

$$\psi_{\underline{\theta}} g = (u_{q\theta}u_\theta + u_{\theta\theta}u_q)g\int_0^\alpha \frac{f(\sigma,a)}{u_{q\theta}(q,\sigma,a)}da + u_q u_\theta g\int_0^\alpha \frac{f_\theta(a,\sigma)u_{q\theta}(q,a,\sigma) - f(a,\sigma)u_{q\theta\theta}(q,a,\sigma)}{u_{q\theta}^2(q,a,\sigma)}\sigma_{\underline{\theta}} da$$

$$-\psi_\alpha = \left\{ \begin{array}{c} -\frac{u_\theta u_q}{u_{q\theta}} f - (u_{q\alpha}u_\theta + u_{\theta\alpha}u_q)\int_0^\alpha \frac{f(\sigma,a)}{u_{q\theta}(q,\sigma,a)}da \\ -u_q u_\theta \int_0^\alpha \frac{f_\theta(\sigma,a)u_{q\theta}(q,\sigma,a) - f(\sigma,a)u_{q\theta\theta}(q,\sigma,a)}{u_{q\theta}^2(q,\sigma,a)}\sigma_\alpha da \end{array} \right\}$$

$$-u_\theta(h_2 - gh_1) = -u_\theta(u_{q\alpha} - \frac{u_\alpha}{u_\theta}u_{q\theta})\int_0^\alpha \frac{f(a,\sigma)}{u_{q\theta}(q,a,\sigma)}da.$$

Combining these four terms then yields (46).

Now consider any $q \in (\underline{q}(0), \underline{q}(0) + \varepsilon)$. Since

$$\lim_{\alpha \to 0}\{u_{qq}(q,\alpha,\underline{\theta}) - u_{qq}(q,a,\sigma)\} = 0$$

and since $u_{qq}(q,\alpha,\underline{\theta})$ it follows that (45) < 0 for α sufficiently small. Also since $\left(u - \frac{u_\theta u_q}{u_{q\theta}}\right) > 0$, it follows that (46) > 0 for α sufficiently small. Hence $\alpha'(q) < 0$ for all α sufficiently small. It follows that α must remain bounded away from zero.

(ii) Let us consider any interval over which $\alpha' = 0$. Then

$$\begin{aligned} \dot{\xi} &= u_\theta h_0 - u_q h_1, \text{ and} \\ \dot{\chi} &= -(\psi + \xi)g_q. \end{aligned}$$

Differentiating $\dot{\chi}$ and using (**??**) then gives

$$\begin{aligned} \chi'' &= -\{\psi + \xi\}g_{qq} - \{\psi_q + \xi'\}g_q \\ \chi''' &= -\{\psi + \xi\}g_{qqq} - 2\{\psi_q + \xi'\}g_{qq} - \{\psi_{qq} + \xi'')g_q \end{aligned} \tag{50}$$

Let $\widetilde{q}$ be the left endpoint of an interval over which $\alpha' = 0$, and suppose that contrary to the claim we had $\widetilde{q} > 0$. Since the constraint $\alpha' \geq 0$ is not binding at $\widetilde{q}$, we have $\xi(\widetilde{q}) = -\psi(\widetilde{q}, \underline{\theta}(\widetilde{q}), \alpha(\widetilde{q}))$. Furthermore, since $\alpha'(\widetilde{q}) = 0$ we have $\psi_q = -u_\theta h_0 + u_q h_1$, and so $\psi_q + \xi'(\widetilde{q}) = 0$. We therefore have $\chi(\widetilde{q}) = \chi'(\widetilde{q}) = \chi''(\widetilde{q}) = 0$, and $\chi'''(\widetilde{q})) = -\{\psi_{qq} + \xi'')g_q$. From (44) we have $\psi_{qq} + \xi'' = \frac{\partial}{\partial q}\{u_\theta h_0 - u_q h_1 + \psi_q\} < 0$. Since $g_q > 0$, we conclude that $\chi'''(\widetilde{q}) > 0$. A Taylor series expansion then yields $\mu(q) = \frac{1}{6}\mu'''(\widetilde{q})(q - \widetilde{q})^3 < 0$ for q in a left neighborhood of $\widetilde{q}$, contradicting that the constraint $\alpha' \geq 0$ is binding over this neighborhood. We conclude that $\widetilde{q} = 0$. Q.E.D.

Proof of Theorem 7: There are several cases to be considered, depending upon whether $\widehat{\alpha} = 1$ or $\widehat{\theta} = 0$, whether $\underline{\alpha}(\widehat{q}, \widehat{\alpha}, \widehat{\theta}) > 0$ or $\underline{\alpha}(\widehat{q}, \widehat{\alpha}, \widehat{\theta}) = 0$, and whether $\underline{\alpha}(0, \alpha(0), \theta(0)) > 0$ or $\underline{\alpha}(0, \alpha(0), \theta(0)) = 0$. Here, we will treat the case where $\widehat{\alpha} = 1$, $\underline{\alpha}(\widehat{q}, \widehat{\alpha}, \widehat{\theta}) > 0$, and $\underline{\alpha}(0, \alpha(0), \theta(0)) > 0$. The proof for the other cases is analogous.

First, let us establish that there is a unique value $q = q^{**}$ such that $\sigma(q, \alpha(q), \theta(q)) = 1$. Applying the the implicit function theorem to the equation $u_q(q, a, \sigma) = u_q(q, \alpha, \theta)$ yields:

$$
\begin{aligned}
\sigma_q(q,\alpha,\theta,a) &= \frac{u_{qq}(q,\alpha,\underline{\theta}) - u_{qq}(q,a,\sigma)}{u_{q\theta}(q,a,\sigma)} \\
\sigma_\alpha(q,\alpha,\theta,a) &= \frac{u_{q\alpha}(q,\alpha,\underline{\theta})}{u_{q\theta}(q,a,\sigma)} \\
\sigma_\alpha(q,\alpha,\theta,a) &= \frac{u_{q\theta}(q,\alpha,\underline{\theta})}{u_{q\theta}(q,a,\sigma)}
\end{aligned}
$$

Now

$$
\begin{aligned}
u_{qq}(q,\alpha(q),\underline{\theta}(q)) - u_{qq}(q,0,\sigma) &= \int_0^{\alpha(q)} \frac{d}{da} u_{qq}(q,a,\sigma(q,\alpha(q),\underline{\theta}(q),a)da \\
&= \int_0^{\alpha(q)} \left(u_{qq\alpha} - u_{qq\theta}\sigma_a\right) da \\
&= \int_0^{\alpha(q)} \left(u_{qq\alpha} - u_{qq\theta}\frac{u_{q\alpha}}{u_{q\theta}}\right) da > 0
\end{aligned}
$$

were the final equality follows from (SCD). Hence $\sigma_q > 0$. In the proof of Lemma 6 we also established that $u_{q\alpha} - \frac{u_\alpha}{u_\theta} u_{q\theta} > 0$ whenever $q > 0$. It then follows that

$$
\begin{aligned}
\frac{d}{dq}\sigma(q,\alpha(q),\underline{\theta}(q),0) &= \sigma_q + (\sigma_\alpha - g\sigma_\theta)\alpha' \\
&= \sigma_q + \frac{\alpha'(q)}{u_{q\theta}(q,\sigma,0)}(u_{q\alpha} - \frac{u_\alpha}{u_\theta}u_{q\theta})(q,\alpha(q),\underline{\theta}(q)) > 0
\end{aligned}
$$

Therefore q^{**} is uniquely determined.

Defining $\alpha^{**} = \alpha(q^{**})$, $\theta^{**} = \theta(q^{**})$ and

$$
\begin{aligned}
W_1(q^{**},\alpha^{**},\theta^{**}) &= \int_0^{q^{**}} u(q,\alpha(q),\theta(q))h(q,\alpha(q),\theta(q),\alpha'(q),\theta'(q))dq \\
W_2(q^{**},\alpha^{**},\theta^{**},\widehat{q},\widehat{\theta}) &= \int_{q^{**}}^{\widehat{q}} u(q,\alpha(q),\theta(q))h(q,\alpha(q),\theta(q),\alpha'(q),\theta'(q))dq
\end{aligned}
$$

we may re-write (9) as

$$
V(\widehat{q},1,\widehat{\theta}) = W_1(q^{**},\alpha^{**},\theta^{**}) + W_2(q^{**},\alpha^{**},\theta^{**},\widehat{q},\widehat{\theta}) + W(\widehat{q},1,\widehat{\theta})
$$

where it is understood that q^{**}, α^{**} and θ^{**} are functions of $(\widehat{q},\widehat{\theta})$. The partial derivatives of the value functions W_1 and W_2 are given by (Seierstad and Sydsaeter, p. 213):

$$
\frac{\partial W_1}{\partial q^{**}} = u(q^{**},\alpha^{**},\theta^{**})h_0(q^{**},\alpha^{**},\theta^{**})
$$

$$
\frac{\partial W_1}{\partial \alpha^{**}} = -\mu_-(q^{**});\ \frac{\partial W_1}{\partial \theta^{**}} = -\lambda_-(q^{**})
$$

$$\frac{\partial W_2}{\partial q^{**}} = -u(q^{**}, \alpha^{**}, \theta^{**})h_0(q^{**}, \alpha^{**}, \theta^{**})$$

$$\frac{\partial W_2}{\partial \alpha^{**}} = \mu_+(q^{**}); \ \frac{\partial W_2}{\partial \theta^{**}} = -\lambda_+(q^{**})$$

$$\frac{\partial W_2}{\partial \widehat{q}} = u(\widehat{q}, 1, \widehat{\theta})h_0(\widehat{q}, 1, \widehat{\theta}); \ \frac{\partial W_2}{\partial \widehat{\theta}} = -\lambda(\widehat{q})$$

Furthermore, since

$$W(\widehat{q}, 1, \widehat{\theta}) = u(\widehat{q}, 1, \widehat{\theta})H(\widehat{q}, 1, \widehat{\theta}) + \int_{\widehat{q}}^{q^{\phi}(\widehat{\theta})} H(q, 1, \widehat{\theta})u_q(q, 1, \widehat{\theta})dq + \int_{q^{\phi}(\widehat{\theta})}^{\bar{q}(1)} H(q, 1, \theta(q))u_q(q, 1, \theta(q))dq$$

where $q^{\phi}(\widehat{\theta})$ is the unique solution to the equation $\theta^{\phi}(q) = \widehat{\theta}$, we have

$$\frac{\partial W}{\partial \widehat{q}} = u(\widehat{q}, 1, \widehat{\theta})H_q(\widehat{q}, 1, \widehat{\theta})$$

$$\frac{\partial W}{\partial \widehat{\theta}} = u_\theta(\widehat{q}, 1, \widehat{\theta})H(\widehat{q}, 1, \widehat{\theta}) + u(\widehat{q}, 1, \widehat{\theta})H_\theta(\widehat{q}, 1, \widehat{\theta}) + \int_{\widehat{q}}^{q^{\phi}(\widehat{\theta})} \phi(q, \widehat{\theta})dq$$

Observe that since $\phi_\theta < 0$ is decreasing in θ, the value of the integral in the previous equation is negative. Also, because

$$H(\widehat{q}, 1, \widehat{\theta}) = \int_{\underline{\alpha}(\widehat{q}, 1, \widehat{\theta})}^{1} \int_{\sigma(\widehat{q}, 1, \widehat{\theta}, a)}^{1} f(a, \theta)d\theta da$$

we have

$$H_\theta(\widehat{q}, 1, \widehat{\theta}) = -\int_{\underline{\alpha}(\widehat{q}, 1, \widehat{\theta})}^{1} f(a, \sigma((\widehat{q}, 1, \widehat{\theta}, a))\sigma_\theta(q, 1, \theta, a)da = -h_1(\widehat{q}, 1, \widehat{\theta})$$

and

$$H_q(\widehat{q}, 1, \widehat{\theta}) = -\int_{\underline{\alpha}(\widehat{q}, 1, \widehat{\theta})}^{1} f(a, \sigma((\widehat{q}, 1, \widehat{\theta}, a))\sigma_q(\widehat{q}, 1, \widehat{\theta}, a)da = -h_0(\widehat{q}, 1, \widehat{\theta})$$

Suppose now that contrary to the statement of the theorem we have $\widehat{\theta} > \theta^{\phi}(\widehat{q})$. Consider the following perturbation: lower $\widehat{\theta}$ by the amount $d\widehat{\theta}$, and adjust $\widehat{q}$ by an amount $d\widehat{q} < 0$ so as to keep α^{**} and θ^{**} (and hence q^{**}) unchanged. Since W_1 remains the same, the total effect on the monopolist's profit then equals

$$\left(\frac{\partial W_2}{\partial \widehat{q}} + \frac{\partial W}{\partial \widehat{q}}\right) d\widehat{q} + \left(\frac{\partial W_2}{\partial \widehat{\theta}} + \frac{\partial W}{\partial \widehat{\theta}}\right) d\widehat{\theta} \tag{51}$$

Now

$$\frac{\partial W_2}{\partial \widehat{q}} + \frac{\partial W}{\partial \widehat{q}} = 0 \tag{52}$$

We also have

$$\begin{aligned}\frac{\partial W_2}{\partial \widehat{\theta}} + \frac{\partial W}{\partial \widehat{\theta}} &= -\lambda(\widehat{q}) + u_\theta(\widehat{q},1,\widehat{\theta})H(\widehat{q},1,\widehat{\theta}) + u(\widehat{q},1,\widehat{\theta})H_\theta(\widehat{q},1,\widehat{\theta}) + \int_{\widehat{q}}^{q^\phi(\widehat{\theta})} \phi(q,\widehat{\theta})dq \quad (53)\\ &= \psi(\widehat{q}) + u_\theta(\widehat{q},1,\widehat{\theta})H(\widehat{q},1,\widehat{\theta}) + \int_{\widehat{q}}^{q^\phi(\widehat{\theta})} \phi(q,\widehat{\theta})dq \\ &\leq \int_{\widehat{q}}^{q^\phi(\widehat{\theta})} \phi(q,\widehat{\theta})dq < 0\end{aligned}$$

where the second equality follows from we used $\lambda(q) = -uh_1 + \psi$, and the penultimate inequality follows because $\phi(q,\theta)$ is decreasing in θ yields:

$$\phi(\widehat{q},\widehat{\theta}) = u_q(\widehat{q},1,\widehat{\theta})H_\theta(\widehat{q},1,\widehat{\theta}) + u_{\theta q}(\widehat{q},1,\widehat{\theta})H(\widehat{q},1,\widehat{\theta}) \leq \phi(\widehat{q},\theta^\phi(\widehat{q})) = 0$$

and so

$$u_\theta(\widehat{q},1,\widehat{\theta})H(\widehat{q},1,\widehat{\theta}) \leq \frac{u_q u_\theta h_1}{u_{q\theta}}(\widehat{q},1,\widehat{\theta}) = \psi(\widehat{q},1,\widehat{\theta})$$

Combining (51), (52) and (53) established that the perturbation is profitable, contradicting the optimality of $(\widehat{q},\widehat{\theta})$. We conclude that optimality requires $\phi(\widehat{q},\widehat{\theta}) = 0$. Q.E.D.

Proof of Theorem 8: (i) There are two cases to be considered: $\widehat{\alpha} = 1$ and $\widehat{\theta} = 0$. We give a proof for the case $\widehat{\alpha} = 1$. The proof for the case $\widehat{\theta} = 0$ is analogous. According to Theorem 7 we have $\widehat{\theta} = \theta^\phi(\widehat{q})$. Hence

$$W(\widehat{q},\theta^\phi(\widehat{q})) = u(\widehat{q},1,\theta^\phi(\widehat{q}))H(\widehat{q},1,\theta^\phi(\widehat{q})) + \int_{\widehat{q}}^{\bar{q}(1)} H(q,1,\theta^\phi(q))u_q(q,1,\theta^\phi(q))dq$$

Thus we have

$$\begin{aligned}\frac{d}{d\widehat{q}}W(\widehat{q},\theta^\phi(\widehat{q})) &= uH_q + (u_\theta H + uH_\theta)\frac{d\theta^\phi}{dq}\\ &= -uh_0 + (\frac{u_\theta u_q}{u_{q\theta}} - u)h_1\frac{d\theta^\phi}{dq}\end{aligned}$$

where the final equality follows from $\phi(\widehat{q},\widehat{\theta}) = u_qH_\theta + u_{\theta q}H = 0$, and the definition $\psi = \frac{u_q u_\theta}{u_{q\theta}}h_1$. Furthermore, we have

$$\frac{\partial W_2}{\partial q} + \frac{\partial W_2}{\partial \theta}\frac{d\theta^\phi}{dq} = uh_0 - \lambda\frac{d\theta^\phi}{dq}$$

Hence

$$\left(\frac{\partial W_2}{\partial q} + \frac{\partial W_2}{\partial \theta}\frac{d\theta^\phi}{dq}\right) + \frac{d}{d\widehat{q}}W(\widehat{q},\theta^\phi(\widehat{q})) = (\psi - uh_1 - \lambda)\frac{d\theta^\phi}{dq} = 0 \quad (54)$$

where the ultimate equality follows from the assumption that $\alpha'(\widehat{q}) > 0$ and the expression for λ in Theorem 6.

Next, observe that

$$\frac{\partial W_1}{\partial q^{**}} + \frac{\partial W_2}{\partial q^{**}} = 0$$

and

$$\begin{aligned}\frac{\partial W_1}{\partial \alpha^{**}} + \frac{\partial W_2}{\partial \alpha^{**}} &= \mu_+(q^{**}) - \mu_-(q^{**}) \\ \frac{\partial W_1}{\partial \theta^{**}} + \frac{\partial W_2}{\partial \theta^{**}} &= \lambda_+(q^{**}) - \lambda_-(q^{**})\end{aligned}$$

Since at the optimum we must have

$$\frac{d}{d\widehat{q}} V(\widehat{q}, 1, \theta^{\phi}(\widehat{q})) = 0$$

it follows that we must have

$$[\mu_+(q^{**}) - \mu_-(q^{**})]\frac{d\alpha^{**}}{d\widehat{q}} + [\lambda_+(q^{**}) - \lambda_-(q^{**})]\frac{d\theta^{**}}{d\widehat{q}} = 0$$

(ii) Again, there are two cases to be considered, $\widehat{\alpha} = 1$ and $\widehat{\theta} = 0$. We give a proof for the case $\widehat{\alpha} = 1$. The proof for the case $\widehat{\theta} = 0$ is analogous. If $\alpha'(\widehat{q}) = 0$ then (54) still holds, but we no longer have $\lambda(\widehat{q}) = \psi - uh_1$. Hence if $\underline{\alpha}(\widehat{q}, \widehat{\alpha}, \widehat{\theta}) \leq 0$ then

$$\frac{d}{d\widehat{q}} V(\widehat{q}, 1, \theta^{\phi}(\widehat{q})) = (-\psi - uh_1 - \lambda)\frac{d\theta^{\phi}}{dq}$$

It follows that we must have $\lambda(\widehat{q}) = \psi - uh_1$. If $\underline{\alpha}(0, \alpha^{**}(0), \underline{\theta}(0)) > 0$, then $V(\widehat{q}, 1, \theta^{\phi}(\widehat{q})) = W_2(q^{**}, \alpha^{**}, \theta^{**}, \widehat{q}, \theta^{\phi}(\widehat{q})) + W(\widehat{q}, \theta^{\phi}(\widehat{q})$ we have

$$\frac{d}{d\widehat{q}} V(\widehat{q}, 1, \theta^{\phi}(\widehat{q})) = (\psi - uh_1 - \lambda)\frac{d\theta^{\phi}}{dq} + \mu(0)\frac{d\alpha^{**}}{d\widehat{q}} + \lambda(0)\frac{d\theta^{**}}{d\widehat{q}} - uh_0\frac{dq^{**}}{d\widehat{q}}$$

The desired conclusion then follows from the transversality condition for problem (11): $\mu(0) = \lambda(0) = 0$, and the fact that $u(0, \alpha, \theta) = 0$ for all (α, θ). We are therefore left with the case where $\underline{\alpha}(\widehat{q}, \widehat{\alpha}, \widehat{\theta}) > 0$ and $\underline{\alpha}(0, \alpha^{**}(0), \underline{\theta}(0)) = 0$. We then have

$$\frac{d}{d\widehat{q}} V(\widehat{q}, 1, \theta^{\phi}(\widehat{q})) = (\psi - uh_1 - \lambda)|_{q=\widehat{q}}\frac{d\theta^{\phi}}{dq}(\widehat{q}) + [\mu_+(q^{**}) - \mu_-(q^{**})]\frac{d\alpha^{**}}{d\widehat{q}} + [\lambda_+(q^{**}) - \lambda_-(q^{**})]\frac{d\theta^{**}}{d\widehat{q}}$$

The desired result then follows because $\alpha'(\widehat{q}) = 0$ and the regularity assumption $\frac{\partial}{\partial q}\{u_\theta h_0 - u_q h_1 + \psi_q\} < 0$ imply $\alpha'(q) = 0$ for all $q \in [0, \widehat{q}]$. In that case μ and λ are continuous on the interval $[0, \widehat{q}]$, and so the last two terms in the above expression equal zero. Q.E.D.

Proof of Theorem 9: First, let us establish that it is necessary that $\alpha'(q) = 0$ for all $q \leq \widehat{q}$. Suppose instead that in the optimal mechanism there existed an interval $[q_-, q_+]$ of $q < \widehat{q}$ on which $\alpha'(q) > 0$. Then for any $q \in [q_-, q_+)$ we have $\underline{\theta} > 0$. It follows that the iso-price line $\sigma(q, \alpha(q), \underline{\theta}(q), a)$ trough the point $(\alpha(q), \underline{\theta}(q))$ at the level q contains points (those with coordinates $a \in (\alpha, \alpha(\widehat{q})])$ which violate the individual rationality condition. Types $(\sigma(q, \alpha(q), \underline{\theta}(q), a), a)$ with $a \in (\alpha, \alpha(\widehat{q})]$ will therefore not consume the the increment q, or any of the increments $z < q$, as is assumed in the demand profile approach.

Next, we establish the necessity of $\widehat{q} = 0$. Suppose to the contrary that we had $\widehat{q} > 0$. Let us now assume that $\theta^{\phi}(0) > 0$; and entirely analogous argument treats the case where $\theta^{\phi}(q) = 0$ for some $q > 0$. It follows from Theorem 7 that $\widehat{\theta} = \theta^{\phi}(\widehat{q})$, and so we have $\phi(\widehat{q}, \widehat{\theta}) = 0$. Furthermore, since ϕ is decreasing in q, we have $\phi(q, \widehat{\theta}) > 0$ for all $q < \widehat{q}$, and so

$$u_q(q, 1, \widehat{\theta}) H_\theta(q, \widehat{\theta}) + u_{q\theta}(q, 1, \widehat{\theta}) H(q, \widehat{\theta}) > 0. \tag{55}$$

Now recall that $N(p, q)$ is be the measure of types (α, θ) for whom $u_q(q, \alpha, \theta) \geq p$. Thus, letting $\widetilde{\theta}(p, q)$ be the solution to $u_q(q, 1, \theta) = p$, we have $N(p, q) = H(q, \widetilde{\theta}(p, q))$. The optimality condition for the problem $\max_p pN(p, q)$ can thus be written as

$$N(p, q) + p \frac{\partial N}{\partial p}(p, q) = 0, \tag{56}$$

or equivalently that

$$u_{q\theta}(q, 1, \theta) H(q, \theta) + u_q(q, 1, \theta) H_\theta(q, \theta) = 0 \text{ at } \theta = \widetilde{\theta}(p, q). \tag{57}$$

It follows from (55), (57) and the fact that ϕ is increasing in θ that $\widehat{\theta} > \widetilde{\theta}(p, q)$. Consequently, the optimal mechanism must differ from the mechanism selected by the demand profile approach.

Next, let us establish sufficiency. If $\widehat{q} = 0$, then in the optimal mechanism we have $\phi(q, \theta^{\phi}(q)) = 0$ for all $q \in [0, \bar{q}(1)]$, implying that (56) holds at $p = u_q(q, 1, \theta^{\phi}(q))$. Furthermore, the monotonicity of ϕ in θ implies that there is no $\theta \neq \theta^{\phi}(q)$ for which (57) holds, so $p = u_q(q, 1, \theta^{\phi}(q))$ is a global optimizer of (56). We conclude that the demand profile approach identifies the optimal mechanism. *Q.E.D.*

9 Appendix B

In this appendix, we prove Theorems 10 and 11. Observe that $u - \frac{u_\theta u_q}{u_{q\theta}} = \frac{(b-\alpha)}{2}q^2 > 0$ for all $q > 0$. It then follows from Lemma 8(i) that $\underline{q}(0) = 0$. We start by analyzing the solution to (8).

Lemma 11 *Consider the problem of quadratic utility and uniformly distributed types, (15)-(16). Let* $q^* = \frac{2}{2b+1}$, $\theta^* = \frac{2b-1}{2b+1}$, *and* $\bar{q} = \frac{1}{b-1}$. *Then when* $\widehat{q} \geq q^*$ *the solution to problem (8) equals:*

$$\theta^\phi(q) = \frac{1+2(b-1)q}{3}, \text{ for } q \in [\widehat{q}, \bar{q}]. \tag{58}$$

When $\widehat{q} < q^*$ *the solution to problem (**??**) is as follows. For* $b \geq \frac{3}{2}$ *we have*

$$\theta^\phi(q) = \begin{cases} \frac{1+2(b-1)q}{3}, & \text{for } q \in [q^*, \bar{q}] \\ \frac{1+(b-\frac{3}{2})q}{2}, & \text{for } q \in [\widehat{q}, q^*]. \end{cases}$$

For $b < 3/2$ *we have*

$$\theta^\phi(q) = \begin{cases} \frac{1+2(b-1)q}{3}, & \text{for } q \in [q^*, \bar{q}] \\ \theta^*, & \text{for } q \in [\widehat{q}, q^*]. \end{cases}$$

Proof: It follows from the definition of σ and (15) that

$$\sigma(q, \theta, 1, a) = \theta + (1-a)q.$$

Let q^* be the solution to the equation $\underline{\alpha}(q, 1, \theta^\phi(q)) = 0$. Whenever $q \geq q^*$, $\underline{\alpha}$ is the solution in a to the equation $\sigma(q, \theta, 1, a) = 1$, so we have:

$$\underline{\alpha}(q, \theta) = 1 - \frac{1-\theta}{q}. \tag{59}$$

Thus

$$\begin{aligned} H(q,\theta) &= \int_{\underline{\alpha}(q)}^{1} \int_{\sigma(q,\theta,1,a)}^{1} dt da \\ &= \frac{1}{2}\frac{(1-\theta)^2}{q} \end{aligned} \tag{60}$$

and hence $\phi(q, \theta) = 0$ yields

$$\theta^\phi(q) = \frac{1+2(b-1)q}{3}, \text{ for } q \in [\widehat{q}, \bar{q}].$$

Here $\bar{q}$ is determined by the condition $\theta^\phi(\bar{q}) = 1$, i.e.

$$\bar{q} = \frac{1}{b-1}.$$

Substituting the expression for $\theta^{\phi}(q)$ into (59) then yields the values for q^{**} and θ^{**}.

For $q \le q^*$ we have $\underline{\alpha}(q, 1, \theta^{\phi}(q)) = 0$, and so

$$\begin{aligned} H(q, \theta) &= \int_0^1 \int_{\sigma(q,\theta,\alpha)}^{\sigma(q^*,\theta^*,\alpha)} dt d\alpha + H(q^*, \theta(q^*)) \\ &= 1 - \theta - \frac{1}{2} q \end{aligned} \tag{61}$$

Thus for $q \le q^*$ the condition $\phi(q, \theta) = 0$ yields:

$$\theta^{\phi}(q) = \frac{1 + (b - \frac{3}{2})q}{2}, \text{ for } q \in [\widehat{q}, q^*]. \tag{62}$$

whenever this expression is non-decreasing in q, i.e. whenever $b \ge \frac{3}{2}$. When $b < \frac{3}{2}$, the constraint $\theta'(q) \ge 0$ is binding, and so we have $\theta^{\phi}(q) = \theta^*$ for all $q \le q^*$. Q.E.D.

Lemma 11 has an important immediate implication: if $b \ge \frac{3}{2}$ then $\widehat{\theta} = \theta^{\phi}(\widehat{q}) \ge \theta^{\phi}(0) = \frac{1}{2}$, and if $b < \frac{3}{2}$ then $\widehat{\theta} = \theta^* \ge \frac{1}{3}$. Hence regardless of the value of b, we will never be in a case where $\widehat{\alpha} < 1$.

Next, we investigate the solution to (9). We start by characterizing the solution over the region $q \in [q^{**}, \widehat{q}]$, when $\widehat{q}$ is not so chosen so large that the constraint $\alpha'(q)$ becomes binding at a value of $q \in (q^{**}, \widehat{q}]$:

Lemma 12 *Suppose $q^* \le \widehat{q} < q^m$, where $q^m = \frac{(3\theta^m - 1)}{2(b-1)}$, and θ^m is the unique real root in the interval $[\theta^*, \frac{2}{3}]$ to the equation*

$$-27(1-\theta)(1 - (2+4b)\theta + (5+4b)\theta^2) = 1.$$

Then the interval $[q^{**}, \widehat{q}]$ is non-empty and the solution to (9) satisfies $\alpha'(q) > 0$ for all $q \in [q^{**}, \widehat{q}]$. It is given by

$$\alpha(q) = c_0 + \frac{c_1}{27}(2 - \sqrt{1 - \frac{6}{c_1 q}})(1 + \sqrt{1 - \frac{6}{c_1 q}})^2 \tag{63}$$

$$\underline{\theta}(q) = \frac{2}{3} - \frac{1}{3}\sqrt{1 - \frac{6}{c_1 q}} \tag{64}$$

where the constants c_0 and c_1 chosen so as to satisfy the two boundary conditions at $q = \widehat{q}$, i.e. $\underline{\theta}(\widehat{q}) = \theta^{\phi}(\widehat{q})$ and $\alpha(\widehat{q}) = 1$:

$$\begin{aligned} c_1 &= \frac{4(b-1)}{(1-\widehat{\theta})(3\widehat{\theta}-1)^2}, \text{ and} \\ c_0 &= \frac{\widehat{\theta}^2(5+4b) - \widehat{\theta}(2+4b) + 1}{(3\widehat{\theta}-1)^2} < 0 \end{aligned}$$

Also, the initial terminal surface constraint $\sigma(q^{**}, \alpha(q^{**}), \underline{\theta}(q^{**}))$ yields:

$$\theta^{**} = 1 - (-\frac{2c_0}{c_1})^{\frac{1}{3}} \tag{65}$$

$$\alpha^{**} = 3c_0 + c_1(-\frac{2c_0}{c_1})^{\frac{2}{3}} \tag{66}$$

$$q^{**} = \frac{(-\frac{2c_0}{c_1})^{\frac{1}{3}}}{3c_0 + c_1(-\frac{2c_0}{c_1})^{\frac{2}{3}}} \tag{67}$$

Proof: In order for the middle region $[q^{**}, \widehat{q}]$ not to be empty, the indifference curve emanating from the point $(\alpha, \theta) = (1, \widehat{\theta})$ must intersect the upper boundary $\theta = 1$ at some $a > 0$, i.e. we must have $\underline{\alpha}(\widehat{q}, 1, \widehat{\theta}) > 0$. Since $\underline{\alpha}(q, \alpha, \underline{\theta})$ is the solution in a to the equation $\sigma(q, \alpha, \underline{\theta}, a) = 1$, and since $\sigma(q, \alpha, \theta, a) = \theta + (\alpha - a)q$, we have

$$\underline{\alpha}(q, \alpha, \underline{\theta}) = \alpha - \frac{1 - \underline{\theta}}{q}. \tag{68}$$

The requirement $\underline{\alpha}(\widehat{q}, 1, \widehat{\theta}) > 0$ can therefore be rewritten as $\widehat{\theta} \geq 1 - \widehat{q}$. By Theorem 7 it must be that $\widehat{\theta} = \theta^{\phi}(\widehat{q})$. Using (58) and simplifying then yields $\widehat{q} \geq q^*$.

Next, let us derive the functional form for (14) over the region $[q^{**}, \widehat{q}]$. Using (68) we may compute

$$\begin{aligned} h(q, \alpha, \underline{\theta}, \underline{\theta}', \alpha') &= \int_{\underline{\alpha}(q,\underline{\theta},\alpha)}^{\alpha} \left\{\sigma_q + \sigma_{\underline{\theta}}\underline{\theta}' + \sigma_\alpha \alpha'\right\} da \\ &= \frac{1}{2}\left(\frac{1 - \underline{\theta}}{q}\right)^2 + \left(\frac{1 - \underline{\theta}}{q}\right)(\underline{\theta}' + q\alpha') \end{aligned}$$

Thus

$$\begin{aligned} h_0(q, \alpha, \underline{\theta}) &= \frac{1}{2}\left(\frac{1 - \underline{\theta}}{q}\right)^2 \\ h_1(q, \alpha, \underline{\theta}) &= \left(\frac{1 - \underline{\theta}}{q}\right) = \alpha \\ h_2(q, \alpha, \underline{\theta}) &= 1 - \underline{\theta} \end{aligned}$$

Also

$$\begin{aligned} \psi(q, \alpha, \underline{\theta}) &= \frac{u_q u_\theta}{u_{q\theta}}(q, \alpha, \underline{\theta}) h_1(q, \alpha, \underline{\theta}) \\ &= (1 - \underline{\theta})(\underline{\theta} - (b - \alpha)q) \end{aligned}$$

and so (14) becomes

$$\alpha'(q) = \frac{(1 - \underline{\theta})(1 - 3\underline{\theta})}{q^2(3\underline{\theta} - 2)} \tag{69}$$

Let us now proceed to solve (69). Separating the variables produces

$$\frac{dq}{q} = -\frac{2(2 - 3\underline{\theta})}{(1 - \underline{\theta})(3\underline{\theta} - 1)} d\underline{\theta}$$

Upon integrating both sides, we obtain:

$$Log(q) = -Log(1 - 4\underline{\theta} + 3\underline{\theta}^2)$$

and so

$$(1 - \underline{\theta})(3\underline{\theta} - 1) = \frac{k}{q}, \tag{70}$$

for some constant $k > 0$. Solving for $\underline{\theta}$, and setting $k = 2/c_1$ then produces (64). The boundary condition $k = \frac{2}{c_1}$ is derived as follows. Evaluating (70) at $q = \widehat{q}$ yields $k = (1 - \widehat{\theta})(3\widehat{\theta} - 1)\widehat{q}$, where $\widehat{q} = \frac{3\widehat{\theta}-1}{2(b-1)}$. It follows that $k = \frac{(1-\widehat{\theta})(3\widehat{\theta}-1)^2}{2(b-1)} = \frac{2}{c_1}$.

Let us now proceed to derive (63). Define the function $\widetilde{\alpha}(\theta)$ implicitly through $\alpha(q) = \widetilde{\alpha}(\underline{\theta}(q))$. Then $\alpha'(q) = \widetilde{\alpha}'(\underline{\theta}(q))\underline{\theta}'(q)$ so

$$\begin{aligned} \widetilde{\alpha}'(\underline{\theta}(q)) &= -\frac{2}{q} \\ \widetilde{\alpha}''(\underline{\theta}(q)) &= \frac{2}{q^2\underline{\theta}'(q)} \end{aligned} \tag{71}$$

Then using (69) to substitute for $\underline{\theta}'(q)$ yields

$$\frac{\widetilde{\alpha}''}{\widetilde{\alpha}'} = \frac{2(2 - 3\underline{\theta})}{(1 - \underline{\theta})(3\underline{\theta} - 1)}$$

which may be integrated to obtain

$$\widetilde{\alpha}'(\theta) = -c_1(1 - \theta)(3\theta - 1). \tag{72}$$

Integrating once more then produces

$$\widetilde{\alpha}(\theta) = c_0 + c_1\theta(1 - \theta)^2. \tag{73}$$

Setting $\theta = \underline{\theta}(q)$, and using (64) then produces(63).

To find θ^{**}, invert (71) to yield

$$q(\theta) = -\frac{2}{\widetilde{\alpha}'(\theta)} = \frac{2}{c_1(1 - \theta)(3\theta - 1)} \tag{74}$$

Now the initial boundary condition $\sigma(q^{**}, \underline{\theta}(q^{**}), \alpha(q^{**}), 0) = 1$ can be rewritten as

$$\theta + q(\theta)\widetilde{\alpha}(\theta) = 1 \tag{75}$$

Substituting (74) and (73) into (75) and solving yields (65). Substituting (65) into (73) and simplifying produces (66). Finally, it follows from (75) that $q^{**} = (1 - \theta^{**})/\alpha^{**}$, which yields (67). Next, let us derive the equations fro c_0 and c_1. From (75) we have

$$c_1 = \frac{2}{\widehat{q}(1 - \widehat{\theta})(3\widehat{\theta} - 1)}$$

Using the equation $\widehat{q} = \frac{3\widehat{\theta}-1}{2(b-1)}$ then yields the required formula for c_1. It follows from (73) that $c_0 + c_1\widehat{\theta}(1-\widehat{\theta})^2 = 1$. Solving for c_0, and substituting in the formula for c_1 then yields

$$c_0 = 1 - \frac{4(b-1)\widehat{\theta}(1-\widehat{\theta})}{(3\widehat{\theta}-1)^2}$$

It remains to be shown that $\alpha'(q) > 0$ for all $q \in [q^{**}, \widehat{q}]$. From (69), this will be true provided $\frac{1}{3} \leq \underline{\theta}(q) < \frac{2}{3}$. Since $\underline{\theta}$ is decreasing in q, to establish these inequalities it suffices to show that $\widehat{\theta} \geq \frac{1}{3}$ and $\theta^{**} \leq \frac{2}{3}$. Now $\widehat{\theta} = \theta^{\phi}(\widehat{q}) \geq \theta^{\phi}(q^*) = \theta^*$; the inequality $\widehat{\theta} \geq \frac{1}{3}$ then follows because θ^* is increasing in b, and because $\theta^* = \frac{1}{3}$ at $b = 1$. Form (65) the inequality $\theta^{**} \leq \frac{2}{3}$ holds if and only if $(-\frac{2c_0}{c_1}) \geq \frac{1}{27}$. Using (19) and (18), and rearranging yields $-27(1-\theta)(1-(2+4b)\theta + (5+4b)\theta^2) \geq 1$, which holds if and only if $\widehat{\theta} \leq \theta^m$. Q.E.D.

Our next lemma characterizes the solution over the interval $[0, q^{**}]$ when $\widehat{q} \geq q^*$, and over the interval $[0, \widehat{q}]$ when $\widehat{q} \leq q^*$.

Lemma 13 *(i) Suppose that $\widehat{q} \geq q^*$. Then the solution to (9) satisfies $\alpha'(q) = 0$ for all $q \in [0, q^{**})$. Thus we have $\alpha(q) = \alpha^{**}$ and $\theta(q) = \theta^{**}$.*

(ii) Suppose that $\widehat{q} \leq q^*$. Then the solution to (9) satisfies $\alpha'(q) = 0$ for all $q \in [0, \widehat{q})$. Thus we have $\alpha(q) = 1$, and $\theta(q) = \widehat{\theta}$.

Proof: Theorem 6 is applicable to these regions, but since $\underline{\alpha}(q, \underline{\theta}(q), \alpha(q)) = 0$ for all $q \leq q^{**}$, the formulae ψ and h need to be adjusted. We may compute:

$$\begin{aligned} h(q, \underline{\theta}, \alpha, \underline{\theta}', \alpha') &= \int_0^{\alpha} f(\sigma(q, \underline{\theta}, \alpha, a), a) \frac{d}{dq} \sigma(q, \underline{\theta}(q), \alpha(q), a) da \\ &= \frac{1}{2}\alpha^2 + \alpha(\underline{\theta}' + q\alpha') \end{aligned} \tag{76}$$

and

$$\begin{aligned} \psi(q, \underline{\theta}, \alpha) &= -u_q(q, \underline{\theta}, \alpha) u_{\theta}(q, \underline{\theta}, \alpha) \int_0^{\alpha} \frac{f(\sigma(q, \underline{\theta}, \alpha, a), a)}{u_{q\theta}(q, \sigma(q, \underline{\theta}, \alpha, a), a)} da \\ &= -q(\underline{\theta} - (b-\alpha)q)\alpha \end{aligned}$$

Using these formulae, we (14) becomes

$$\alpha'(q) = \frac{\alpha(3\alpha - 2b)}{(b - 3\alpha)q} \tag{77}$$

(i) We divide the analysis into three cases. First, consider the case where $\alpha^{**} > \frac{2b}{3}$. Suppose that contrary to the statement of the Lemma there existed a $q^+ < q^*$ such that $\alpha(q^+) < \alpha^{**}$. Without loss of generality, we may assume that $\alpha(q^+) \geq \frac{2b}{3}$. We now claim that for any $q \in (q^+, q^{**})$ it must be that $\alpha'(q) = 0$, yielding an immediate contradiction to the assumption that $\alpha(q^+) < \alpha^{**}$. Indeed, if we had $\alpha'(q) > 0$, then from Theorem 6 equation

(77) would hold. But (77) yields $\alpha'(q) \leq 0$ whenever $\alpha(q) \in [\frac{2b}{3}, \alpha^{**})$, a contradiction to the assumption that $\alpha'(q) > 0$. We conclude that whenever $\alpha^{**} > \frac{2b}{3}$ we must have $\alpha(q) = \alpha^{**}$ for all $q \in [0, q^{**}]$.

Next, consider the case where $\alpha^{**} \leq \frac{b}{3}$. Suppose that contrary to the statement of the lemma there existed a $q^- < q^*$ such that $\alpha(q^-) < \alpha^{**}$. Let $q^+ = \min\{z : \alpha(z) = \alpha^{**}\}$. We now claim that for any $q \in (q^-, q^+)$ we must have $\alpha'(q) = 0$, contradicting the assumption that $\alpha(q^-) < \alpha^{**}$. To establish the claim, suppose to the contrary that we had $\alpha'(q) > 0$ for some $q \in (q^-, q^+)$. Then equation (77) must hold. But since $\alpha(q) < \frac{b}{3}$, it follows from (77) that $\alpha'(q) < 0$, a contradiction. We conclude that whenever $\alpha^{**} \leq \frac{b}{3}$ we have $\alpha(q) = \alpha^{**}$ for all $q \in [0, q^{**}]$.

Finally, consider the case where $\alpha^{**} \in (\frac{b}{3}, \frac{2b}{3}]$. First, we claim the constraint $\alpha'(q) \geq 0$ must be binding somewhere on the interval $[0, q^{**})$. Indeed, if this were not the case, then from Theorem 6 equation (77) would hold for all $q \in [0, q^{**})$. Separating by variables, this equation can be rewritten as

$$\frac{dq}{q} = \frac{1}{2} d \ln(\alpha(2b - 3\alpha))$$

Integrating both sides yields

$$\alpha(q) = \frac{b}{3} + \frac{1}{3}\sqrt{b^2 - 3c_2^2 q^{-2}} \tag{78}$$

or equivalently that

$$q = \frac{c_2}{\sqrt{\alpha(2b - 3\alpha)}} \tag{79}$$

The constant c_2 is chosen so that the boundary condition $\alpha(q^{**}) = \alpha^{**}$ is satisfied, i.e.

$$c_2 = q^{**}\sqrt{\alpha^{**}(2b - 3\alpha^{**})}$$

If $\alpha^{**} = \frac{2b}{3}$, then $c_2 = 0$ and so (78) yields $\alpha(q) = \frac{2b}{3}$ for all q, contradicting the presumption that $\alpha'(q) > 0$ for all $q \in [0, q^{**})$. If $\alpha^{**} \in (\frac{b}{3}, \frac{2b}{3})$ it follows from (78) that $\alpha(q) \geq \frac{b}{3}$ for all q, and so from (79) we obtain a bound

$$q \geq \frac{c_2}{b\sqrt{3}}$$

below which (77) no longer has a solution, also contradicting the presumption that $\alpha'(q) > 0$ for all $q \in [0, q^{**})$. This establishes the claim that we cannot have $\alpha'(q) > 0$ for all $q \in [0, q^{**}]$. The same proof also establishes that there cannot exist a neighborhood of $q = 0$ on which the constraint $\alpha'(q) \geq 0$ is never binding.

Next, suppose that contrary to the statement of the lemma, there existed an interval $(q^-, q^+) \subset (0, q^{**})$ such that $\alpha'(q) > 0$ for all $q \in (q^-, q^+)$. Without loss of generality, let q_- $= \inf\{z : \alpha'(q) > 0 \text{ for all } q \in (q, q^+)\}$. Then since $\mu(q) = 0$ for all (q^-, q^+), we would have $\mu(q^-) = \mu'(q^-) = 0$. By the previous claim, there exists a sequence $\{q_n\}$ with $q_n \uparrow q^-$ such that $\mu(q_n) > 0$. By Theorem 6, we have $\mu'(q_n) = (\psi - \lambda) g_q|_{q=q_n}$. Taking limits as $n \to \infty$ it follows that $\mu'(q^-) = (\psi - \lambda) g_q|_{q=q^-} = 0$. Now $g_q = \frac{1}{2} > 0$, so we must have $(\psi - \lambda)|_{q=q^-} = 0$.

Furthermore, we may calculate that at $q = q_n$:

$$\begin{aligned}\mu''(q) &= (\psi_q - \lambda')g_q + (\psi - \lambda)\frac{d}{dq}g_q \\ &= (\psi_q - u_\theta h_0 - u_q h_1)g_q + (\psi - \lambda)\frac{d}{dq}g_q\end{aligned}$$

Therefore

$$\begin{aligned}\mu''(q^-) &= (\psi_q - u_\theta h_0 - u_q h_1)g_q|_{q=q^-} \\ &= \frac{q\alpha}{4}(2b - 3\alpha)|_{q=q^-}\end{aligned}$$

Since $\alpha(q^-) < \alpha^{**} \leq \frac{2b}{3}$ we therefore have $\mu''(q^-) > 0$. But $\mu(q^-) = 0$ and $\mu''(q^-) > 0$ implies that there exists $\varepsilon > 0$ such that $\mu(q) < 0$ for all $q \in (q^- - \varepsilon, q^-)$. This contradicts that $\mu(q_n) > 0$ for all n. We conclude that we must have $\alpha'(q) = 0$ for all $q \in [0, q^{**}]$.

(ii) The proof for this case is analogous to the proof for case (i). *Q.E.D.*

Proof of Theorem 11: According to Lemma 11 we must have $\widehat{\theta} \geq \theta^*$. Since Theorem 7 we must have $\widehat{\theta} = \theta^\phi(\widehat{q})$, it follows that we must have $\widehat{q} \geq q^*$. Lemmas 11, 12 and 13 then characterize the solution as a function of the parameter $\widehat{q}$, as $\widehat{q}$ ranges over the interval $[q^*, q^m]$. We shall prove that over this range the profit function $V(\widehat{q}, 1, \theta^\phi(\widehat{q}))$ is strictly quasiconcave in $\widehat{q}$, attaining a maximum at the unique value of $\widehat{q}$ for which $\alpha^{**}(\widehat{q}) = \frac{2b}{3}$:[17]

$$\frac{d}{d\widehat{q}}V(\widehat{q}, 1, \theta^\phi(\widehat{q})) \lesseqgtr 0 \text{ as } \alpha^{**} \gtreqless \frac{2b}{3} \tag{80}$$

From the proof of Theorem 8(i), we have

$$\frac{d}{d\widehat{q}}V(\widehat{q}, 1, \theta^\phi(\widehat{q})) = [\mu_+(q^{**}) - \mu_-(q^{**})]\frac{d\alpha^{**}}{d\widehat{q}} + [\lambda_+(q^{**}) - \lambda_-(q^{**})]\frac{d\theta^{**}}{d\widehat{q}}$$

Let us first calculate $\mu_-(q^{**})$ and $\lambda_-(q^{**})$. Since 11 has free left hand endpoints $\alpha(0)$ and $\underline{\theta}(0)$, it must be the case that $\mu(0) = \lambda(0) = 0$. Furthermore, since by Lemma 13 the constraint $\alpha'(q) \geq 0$ is binding on the interval $[0, q^{**})$, it follows from Theorem 6 that $\dot{\mu} = -\frac{\partial}{\partial\alpha}uh_0$ and $\dot{\lambda} = -\frac{\partial}{\partial\underline{\theta}}uh_0$ on $[0, q^{**})$. Furthermore, the proof of Lemma 13 also established that $u(q, \alpha(q), \underline{\theta}(q)) = u(q, \alpha^{**}, \theta^{**})$ and $h_0(q, \alpha(q), \underline{\theta}(q)) =. \frac{1}{2}(\alpha^{**})^2$. It follows that

$$\begin{aligned}\dot{\lambda} &= -\frac{1}{2}(\alpha^{**})^2 q \\ \dot{\mu} &= -\frac{q\alpha^{**}}{4}(4\theta^{**} + q(3\alpha^{**} - 2b))\end{aligned}$$

Integrating then yields

$$\lambda_-(q^{**}) = -\frac{1}{2}(\alpha^{**})^2 \int_0^{q^{**}} q dq = -(\frac{\alpha^{**}q^{**}}{2})^2$$

[17]It is straightfowrward to characterize the solution when $\widehat{q} > q^m$, and to show that $\frac{d}{d\widehat{q}}V(\widehat{q}, 1, \theta^\phi(\widehat{q})) < 0$ for $q > q^m$. For the sake of brevisty, we omit the details here.

$$\mu_-(q^{**}) = -\frac{\alpha^{**}(q^{**})^2}{12}((2b-3\alpha^{**})q^{**} - 6\theta^{**})$$

Next, since $\alpha'(q) > 0$ for all $q \in (q^{**}, \widehat{q})$, it follows from Theorem 6 that

$$\begin{aligned}\lambda_+(q^{**}) &= -uh_1 + \psi \\ \mu_+(q^{**}) &= -uh_2 + \psi g\end{aligned}$$

By Lemma 12 we have $h_1(q^{**}, \alpha^{**}, \theta^{**}) = \alpha^{**}$ and $h_2(q^{**}, \alpha^{**}, \theta^{**}) = 1 - \theta^{**} = \alpha^{**}q^{**}$. Furthermore

$$\begin{aligned}\psi(q^{**}, \alpha^{**}, \theta^{**}) &= \frac{u_q u_\theta}{u_{q\theta}}(q^{**}, \alpha^{**}, \theta^{**})h_1(q^{**}, \alpha^{**}, \theta^{**}) \\ &= \alpha^{**}q^{**}(\theta^{**} - (b-\alpha^{**})q^{**})\end{aligned}$$

Hence

$$\lambda_+(q^{**}) = -\alpha^{**}(q^{**})^2\frac{(b-\alpha^{**})}{2}$$

$$\mu_+(q^{**}) = -\frac{\alpha^{**}(q^{**})^2}{2}\theta^{**}$$

and so

$$\lambda_+(q^{**}) - \lambda_-(q^{**}) = \frac{\alpha^{**}(q^{**})^2}{4}(3\alpha^{**} - 2b)$$

$$\mu_+(q^{**}) - \mu_-(q^{**}) = \frac{\alpha^{**}(q^{**})^3}{12}(2b - 3\alpha^{**})$$

Now over the considered range, we have $\frac{d\alpha^{**}}{d\widehat{q}} < 0$ and $\frac{d\theta^{**}}{d\widehat{q}} > 0$, so (80) follows. Furthermore, since $\frac{d\alpha^{**}}{d\widehat{q}} < 0$, (80) implies that

$$\frac{d}{d\widehat{q}}V(\widehat{q}, 1, \theta^{\phi}(\widehat{q})) \gtreqless 0 \text{ as } \widehat{q} \lesseqgtr \mathring{q}$$

where $\mathring{q}$ solves (17). Thus $V(\widehat{q}, 1, \theta^{\phi}(\widehat{q}))$ is strictly quasiconcave in $\widehat{q}$ over the range $[q^{**}, q^m]$, attaining a unique maximum at $\widehat{q} = \mathring{q}$. Q.E.D.

Proof of Theorem 10: .For the case $b \geq \frac{3}{2}$, Lemmas 11, and 13 characterize the solution as a function of the parameter $\widehat{q}$, as $\widehat{q}$ ranges over the interval $[0, q^*]$. We shall first prove that over this range the profit function $V(\widehat{q}, 1, \theta^{\phi}(\widehat{q}))$ is strictly quasiconcave in $\widehat{q}$, attaining a maximum at the unique value of $\widehat{q} = 0$:[18]

$$\frac{d}{d\widehat{q}}V(\widehat{q}, 1, \theta^{\phi}(\widehat{q})) \leq 0, \text{ with equality only if } \widehat{q} = 0 \tag{81}$$

[18] It is straightfowrward to characterize the solution when $\widehat{q} > q^*$, and to show that $\frac{d}{d\widehat{q}}V(\widehat{q}, 1, \theta^{\phi}(\widehat{q})) < 0$ for $q > q^*$. For the sake of brevisty, we omit the details.

From the proof of Theorem 8(ii), we have

$$\frac{d}{d\widehat{q}}V(\widehat{q},1,\theta^{\phi}(\widehat{q})) = (\psi - uh_1 - \lambda)\frac{d\theta^{\phi}}{dq}$$

Now from the proof of Lemma 13 we have

$$h_1(\widehat{q},1,\widehat{\theta}) = 1$$

and so

$$\psi(\widehat{q},1,\widehat{\theta}) = \frac{u_q u_\theta}{u_{q\theta}}(\widehat{q},1,\widehat{\theta}) = \widehat{q}^2(\widehat{\theta} - (b-1)\widehat{q})$$

$$\begin{aligned}(\psi - uh_1)(\widehat{q},1,\widehat{\theta}) &= \widehat{q}(\widehat{\theta} - (b-1)\widehat{q}) - \widehat{q}(\widehat{\theta} - \frac{b-1}{2}\widehat{q}) \\ &= -\widehat{q}^2\frac{b-1}{2}\end{aligned}$$

Now since 11 has free left hand endpoint $\underline{\theta}(0)$, it must be the case that $\lambda(0) = 0$. At the same time, it follows from Theorem 6 that $\dot{\lambda} = -\frac{\partial}{\partial\underline{\theta}}uh_0$ on $[0,\widehat{q})$. Also, the proof of Lemma 13 established that $u(q,\alpha(q),\underline{\theta}(q)) = u(q,1,\widehat{\theta})$ and $h_0(q,\alpha(q),\underline{\theta}(q)) =. \frac{1}{2}$. Hence we have

$$\dot{\lambda} = -\frac{q}{2}$$

and so $\lambda(\widehat{q}) = -\frac{\widehat{q}^2}{4}$. We conclude that

$$\frac{d}{d\widehat{q}}V(\widehat{q},1,\theta^{\phi}(\widehat{q})) = -\frac{\widehat{q}^2}{4}(2b-3) \leq 0,$$

with equality only if $\widehat{q} = 0$. Q.E.D.