A Solution to a Class of Multi-Dimensional Screening Problems: Isoquants and Clustering*

Raymond Deneckere† and Sergei Severinov‡

This Version: June 8, 2017

Abstract

We develop a general method for solving screening problems with multi-dimensional types and one-dimensional 'physical' allocation space. Our method is based on characterizing and computing the isoquants, the sets of types who are allocated the same quantity (or quality) of the good, and then assigning the quantities optimally along the endogenously chosen boundary of the set of types who get positive quantities in the optimal mechanism. The optimal mechanism exhibits a number of qualitative properties that distinguish this setting from the one dimensional case. In particular, the optimal allocation exhibits a discontinuity along the boundary of the region of excluded types and also, for a set of parameters of positive measure, exhibits clustering, a situation in which an interval of quantities is optimal for a single consumer type. We illustrate the application of our method to an example with uniformly distributed types.
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

*We thank Ken Judd, Preston McAfee, and Bob Wilson, and participants in numerous conferences and seminars for helpful comments and suggestions. A standard disclaimer applies.

†Department of Economics, University of Wisconsin, Madison. Email: rjdeneck@wisc.edu.

‡Vancouver School of Economics, University of British Columbia, Vancouver, and Higher School of Economics, Moscow. Email: sseverinov@gmail.com.

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 used to process the input, and the demand for the final product. 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 a low ability and those who have a high preference for leisure. In the sphere of regulation, the regulatory agency may wish to apply a different regulatory regime to 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 based on purchased quantity. In particular, 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 is often also one-dimensional. In insurance markets, the allocation consists of the amount of coverage, in labor taxation the instrument is the tax rate, and in regulation it is the regulated price.

The analysis of this problem in our paper delivers several methodological contributions. First, by correctly characterizing the isoquants, the sets 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 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 the discontinuities in the quantity allocation as a function of types that typically arise in our problem (as discussed in the next paragraph). Our method also handles bunching in a straightforward and transparent way, without any need to resort to "ironing" technique. Finally, we present a novel condition, that we call Single Crossing of Demand (SCD). It 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, 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). Second, 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 regions. The consumer type at which the quantities are bunched is then indifferent between all quantities in the bunch, and at this type there is another discontinuity of quantity as a function of type. This bunching is different from the traditional phenomenon in one-dimensional screening problems, where multiple types are assigned the same quantity. For this reason, we call this new phenomenon "clustering."

The rest of the paper is organized as follows. In Section 2 we review related literature and methods to solve multidimensional screening problems. In section 3 we present the model and describe the isoquants method. In section 4 we reformulate and solve our problem as that of assignment of types to quantities on the boundary of the set of "active" types and the determination of this boundary. Section 5 characterizes the qualitative properties of the optimal mechanism. Section 6 studies an example with uniformly distributed types and linear-power utility function. Section 7 contains the conclusions. The proofs are relegated to the Appendix.

2 Literature Review

Despite the existence of a large literature on screening, much remains unknown 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 the marginal valuation of quantity/quality of the good is increasing in type, and that the assignment of quantities to types is nondecreasing (ensured by a monotonic hazard rate condition, or achieved by ironing). In the multi-dimensional case, no such sufficient condition is known. Furthermore, it is hard to ensure that the marginal price schedule crosses the demand curve from below, because demand curves vary both in slope and intercept, so a sufficient variation in the intercept will 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 below will correspond to a local minimum rather than a global maximum of their surplus maximization problem. Some examples worked out in the literature, such as the linear quadratic one studied in Wilson (1993, p. 196), therefore involve tariffs that are not incentive compatible. To illustrate this, consider the following example.

Example 1 Suppose that a monopolistic seller of a good faces a consumer with utility function $u(q, \alpha, \theta) = \theta q - \frac{b-\alpha}{2} q^2$, where q is the quantity of the good, (α, θ) is a private consumer type distributed uniformly over the unit square, and b is a constant satisfying $b < \frac{3}{2}$. The seller has zero cost of production. We are interested in the optimal pricing strategy or, equivalently, the optimal screening mechanism.

Following Wilson (1993) define the demand profile $R(p,q)$ as the fraction of consumers in the population whose demand price for quantity increment q, u_q, exceeds p. A simple calculation yields:

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

According to the demand profile approach, $R(p,q)$ represents the demand for quantity increment q. Thus for the quantity increment q, monopolist should charge the price $p(q)$ that solves the following problem:

$$\max_p \{(p-c)R(p,q)\}$$

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 tariff $P(q) = \int_0^q p(z)dz$

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

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 $u_q(q,\alpha,\theta) = p(q)$, do not intersect, for then every consumer type (α,θ) will have only one solution to the first-order condition associated with her surplus maximization problem $\max_q\{u(q,t) - P(q)\}$.[1]

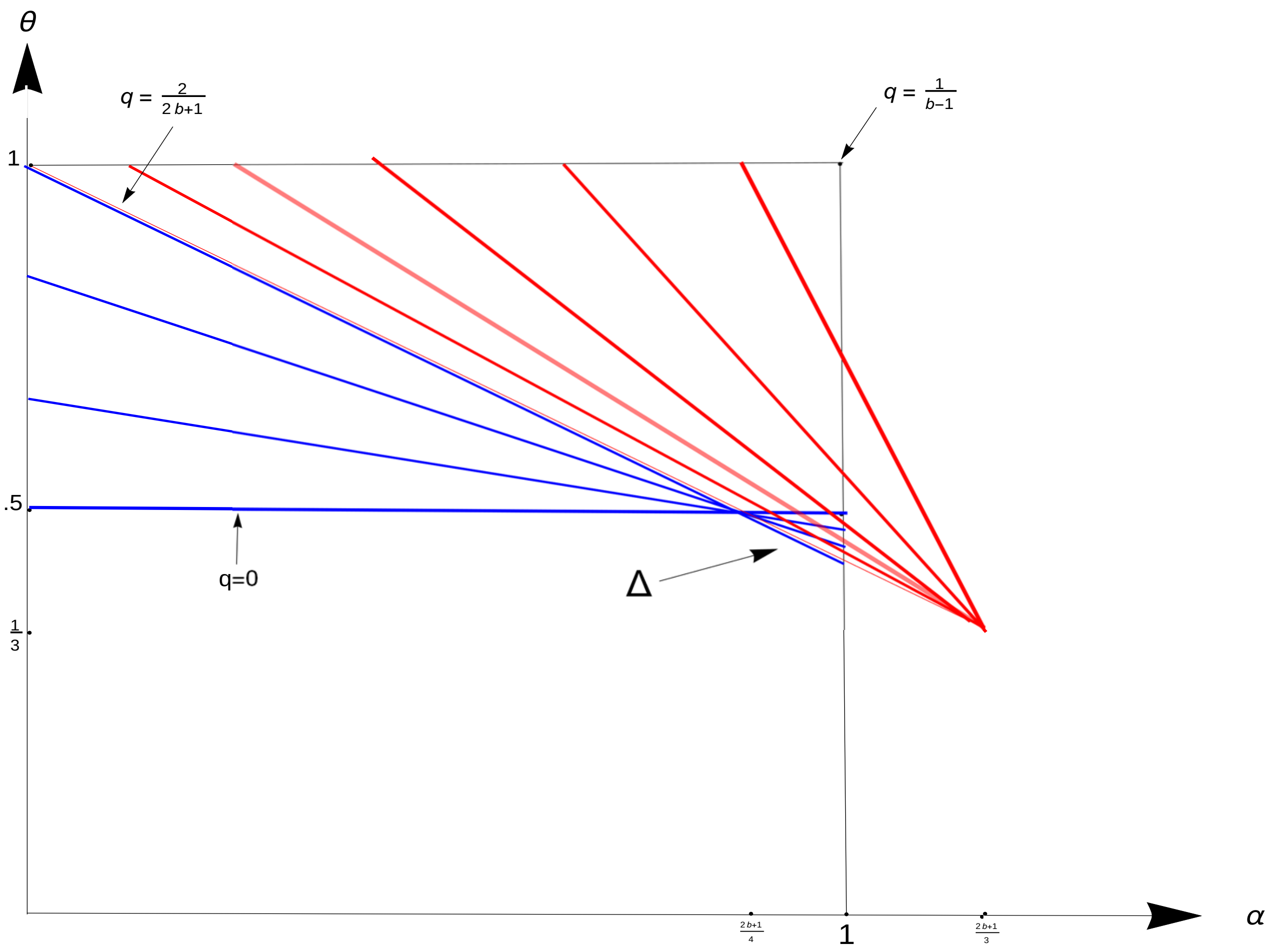
Let us therefore examine the iso-price curves associated with the schedule P. Solving the equation $\theta - (b-\alpha)q = p(q)$ yields

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

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 from a flat line at the level $q = 0$ to 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,

[1]More formally, consider any type (α,θ) on the iso-price curve at the quantity q, i.e. $u_q(q,\alpha,\theta) - p(q) = 0$. 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 (iv) that $u_q(q',\alpha,\theta) - p(q') > 0$ for $q' < q$, and $u_q(q',\alpha,\theta) - p(q') < 0$ for $q' > q$. Consequently, type (α,θ)'s objective function is strictly quasiconcave, implying that q is a global maximum.

Figure 1: Iso Price Curves



when the north-east corner point $(\alpha, \theta) = (1, 1)$ is hit. This means that any point (α, θ) in the interior of the triangle Δ defined by the inequalities $\frac{1+2b-2\alpha}{1+2b} \leq \theta \leq 1/2$ and $\alpha \leq 1$ 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 of 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

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

Setting this expression equal to zero traces out a curve:

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

The participation constraint is violated for all types in Δ that lie below the curve $\theta_L(\alpha)$. As a consequence, the demand profile approach necessarily fails when $b < \frac{3}{2}$.

McAfee and McMillan (1988) develop a different approach. 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 an important limitation: 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.[2]

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 drawback. 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.[3]

The crossing of iso-price lines demonstrated in Example 1 also implies that the methods of McAfee and McMillan and Lewis and Sappington have an inherent flaw. Indeed, since a consumer type can lie on two distinct iso-price lines $u_q(q, \alpha, \theta) - p(q) = 0$, merely being located on an iso-price line generally cannot identify the quantity purchased by a consumer.

[2]Properly taking into account the agent's participation constraint changes the integrand of principal's objective function in an essential way. As a consequence, McAfee and McMillan's formulation of the problem can no longer be solved by the method of calculus of variation.

[3]Armstrong (1996) already pointed out this deficiency.

Rochet and Stole (2003) advance the state of the art considerably by developing the direct method for arbitrary multi-dimensional screening problems. However, because their solution method does not reduce the dimensionality of the 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. More importantly, because the direct approach only imposes the local incentive compatibility constraints, the solution typically violates the conditions for global incentive compatibility.

More recently, a general solution method for our problem has become available in 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 useful practical applications.

Several other papers are related to our work. Laffont, Maskin, Rochet (1987) were the first to tackle the problem we are studying, albeit in a special case. They analyze the quadratic utility, uniformly distributed problem considered in Example 1, under the assumption that $b \geq \frac{3}{2}$. They develop a change of variables technique that expresses one of the utility parameters as a function of the control variable, thereby obtaining a single parameter objective. However, their analysis relies upon an endogenous assumption on the optimal tariff. We revisit this example in Section 6 below (see Theorem 10). We find that with $b \geq \frac{3}{2}$ the demand profile method is valid, because there is then little relative variability in the slope of consumers' demand functions. In other words, we are then sufficiently close to a one dimensional screening problem in which there is significant variation in the demand intercepts only.

Armstrong (1996) reduces the multi-dimensional screening problem to a one-dimensional one, at the cost of a very strong separability condition on the indirect utility function, and severe restrictions on the distribution of consumer characteristics. His most significant contribution is to show that there is a robust sense in which exclusion is much more prevalent in the multi-dimensional case than in the one-dimensional case.

Rochet and Stole (2002) consider a model in which the consumer's value of the outside option is private information. In principle, this yields a multi-dimensional screening problem in which the value of the outside option is another dimension of type. Using a clever transformation, these authors are able to reduce this problem to a variant of the standard uni-dimensional screening problem. They derive the interesting conclusion that there is either no distortion or bundling at the bottom of the distribution of consumer valuations.

Finally, Jehiel, Moldovanu and Stacchetti (1996) consider optimal auctions when a trader's purchase exerts a negative externality upon all non-traders. The value of the externality then becomes a second dimension of bidders' types. They establish that screening on the externality is not feasible, so that isoquants do not depend upon this second dimension of type. Jehiel, Moldovanu and Stacchetti (1999) also study an auction with externalities and multidimensional types. Converting multidimensional types into bids, they derive differential equations that characterize bidding strategies and iso-bid curves in the type space. Allowing for externalities is an interesting extension of the standard model, but the analysis is limited to utilities that are linear in type.

3 The Model and Characterization of Isoquants

A monopolist supplier of a single good faces a population of consumers. Consumers are distinguished by a two dimensional preference parameter (α, θ), which is private information. When consuming a quantity $q \in \mathbf{R}_+$ of the good, acquired at cost p, a consumer of type (α, θ) receives net utility $u(q, \alpha, \theta) - p$. Consumers' reservation utilities are equal to zero.

The distribution function $F(\alpha, \theta)$ of consumer types in the population is common knowledge. We assume that $F(.)$ is a twice continuously differentiable function, with density function $f(\alpha, \theta) > 0$, 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 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* $(\alpha, \theta) \in [0, 1]^2$; $u(q, \alpha, 0) \leq 0$ *for all* $q \in \mathbf{R}_+$, $\alpha \in [0, 1]$;
(ii) $u_q(0, \alpha, \theta) \geq 0$ *for all* (α, θ), *with strict inequality whenever* $\theta > 0$; $u_{qq}(q, \alpha, \theta) < 0$ *and* $\lim_{q \to \infty} u_q(q, \alpha, \theta) < 0$ *for all* (q, α, θ);
(iii) $u_\alpha(q, \alpha, \theta) > 0$ *and* $u_\theta(q, \alpha, \theta) > 0$ *for all* (α, θ) *and* $q > 0$;
(iv) $u_{q\theta}(q, \alpha, \theta) > 0$ *and* $u_{q\alpha}(q, \alpha, \theta) \geq 0$ *for all* (q, α, θ); $u_{q\alpha}(q, \alpha, \theta) > 0$ *for all* (α, θ) *and* $q > 0$; $\lim_{q \to 0} \frac{u_{\theta q}(q, \alpha, \theta)}{u_{\alpha q}(q, \alpha, \theta)} = \infty$ *for all* $(\alpha, \theta) \in (0, 1]^2$.

Assumption 1 is fairly standard. Part (i) says that a consumer receives zero utility from consuming nothing. Furthermore, it implies that consumer types with a sufficiently low value of the parameter θ do not value consumption, and hence in equilibrium will be excluded from purchasing. Part (ii) ensures that consumers' inverse demand functions are downward sloping, with bounded and non-zero intercepts on both axes. As a consequence, the first-best quantity $q^{FB}(\alpha,\theta) = \arg\max_{q\in\mathbf{R}_+} u(q,\alpha,\theta)$ is bounded for all α and θ. Part (iv) requires consumer's utility functions to be supermodular. In particular, the assumption $u_{q\theta}(0,\alpha,\theta) > 0$ says that the inverse demand intercept increases as θ increases, and therefore imparts θ the interpretation of an intercept parameter. Meanwhile, the assumption $\lim_{q\to 0} \frac{u_{\theta q}(q,\alpha,\theta)}{u_{\alpha q}(q,\alpha,\theta)} = \infty$ lends α the interpretation of a slope parameter.

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 interpretation of Assumption 2 is that 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 significantly more restrictive, as it implies that consumers' demand curves do not intersect at all, i.e. can be ranked. As we will show, Assumption 2 has many important consequences. In particular, it implies that isoquants in (α,θ) space -which we will define below- cannot intersect, and must "fan out."

We assume that the firm's production cost is zero. This is a normalizing assumption, as we can allow for any convex cost of production $C(q)$ and then subtract if from the utility function. That is, one should view the utility function $u(q,\alpha,\theta)$ as a net surplus obtained by subtracting the production cost from the buyer's utility.

By the Revelation Principle, the monopolist's problem can be stated as a choice of a direct mechanism $(q(\alpha,\theta),t(\alpha,\theta))$, where $q(\alpha,\theta)$ is the quantity assigned to type (α,θ) and $t(\alpha,\theta)$ is the transfer that this type pays to the firm, which maximizes the firm's expected profits subject to the consumer's incentive and individual rationality constraints.[4] Formally, this problem can be stated as follows:

$$\begin{gathered}\max \textstyle\int_{[0,1]^2} t(\alpha,\theta)dF(\alpha,\theta) \qquad (1)\\ u(q(\alpha,\theta),\alpha,\theta) - t(\alpha,\theta) \geq u(q(\alpha',\theta'),\alpha,\theta) - t(\alpha',\theta') \quad \text{for all } (\alpha,\theta),\ (\alpha',\theta') \in [0,1]^2 \\ u(q(\alpha,\theta),\alpha,\theta) - t(\alpha,\theta) \geq 0 \quad \text{for all } (\alpha,\theta)\ \in [0,1]^2\end{gathered}$$

The solution to this problem exists by standard arguments, since $u(.)$ is continuous and bounded in (α,θ).

[4]According to the Taxation principle, we can view this problem equivalently as the choice of an optimal tariff $P(q)$.

Since the dimension of the type space in our problem is greater than the dimension of the quantity space, it is natural to conjecture that in the optimal mechanism the same quantity q will be assigned to several types, rather than a single type. This is, indeed, the case. Moreover, characterizing the sets of types that consume the same quantity -isoquants, as we call them- will be an important component of our solution method. The second component of our solution method involves characterizing the endogenous boundary of the set of types who are assigned a positive quantity in the mechanism and showing that the quantity allocation on this boundary uniquely determines the rest of the mechanism. The remainder of this section develops this approach.

Let Ω_+ denote the participation region, i.e. the set of all types that consume a positive quantity in the mechanism. Formally, $\Omega_+ \equiv \{(\alpha, \theta) | q(\alpha, \theta) > 0\}$. The complement of Ω_+ is the exclusion region, which contains all types who consume a zero quantity in the mechanism. Our first step is to characterize the lower boundary, $\underline{\theta}(\alpha)$ between the participation region Ω_+ and the exclusion region. This boundary is defined as follows:[5]

$$\underline{\theta}(\alpha) \equiv \inf\{\theta | \theta \geq 0,\ q(\alpha, \theta) > 0\}$$

The following Lemma provides a characterization of the lower boundary $\underline{\theta}(\alpha)$:

Lemma 2 *Suppose that Assumptions 1 and 2 hold. Then in the optimal mechanism we have:*
(i) The lower boundary $\underline{\theta}(\alpha)$ is continuous and decreasing in α, strictly so when $q(\alpha, \underline{\theta}(\alpha)) > 0$. For all $\alpha \in [0, 1)$, $\underline{\theta}(\alpha) > 0$ and $u(q(\alpha, \underline{\theta}(\alpha)), \alpha, \underline{\theta}(\alpha)) - t(\alpha, \underline{\theta}(\alpha)) = 0$.
(ii) For almost all α s.t. $\underline{\theta}(\alpha) < 1$, we have

$$\frac{d\underline{\theta}}{d\alpha} = -\frac{u_\alpha}{u_\theta}(q(\alpha, \underline{\theta}(\alpha)), \alpha, \underline{\theta}(\alpha)). \tag{2}$$

(iii) If $\sup \frac{u_\alpha}{u_\theta}(q, \alpha, \theta) < \infty$, then $\underline{\theta}(\alpha)$ is absolutely continuous.

The characterizations of the lower boundary $\underline{\theta}(\alpha)$ is important because, as we show below, the quantity assignment on the union of the lower boundary and right boundary $\{(1, \theta) | \theta \geq \underline{\theta}(1)\}$, fully determines the mechanism everywhere in the type space. In particular, note that according to part (iii) of the Lemma the lower boundary and the right boundary meet at the point $(1, \underline{\theta}(1))$.

We now set out to characterize the isoquants - the sets of types assigned the same quantity in the mechanism. The next Lemma provides a first step in this direction.

Lemma 3 *Suppose Assumption 2 holds. Let $q_1 > 0$ be an optimal quantity choice for type (α_1, θ_1) in the mechanism, and suppose that*

$$u_q(q_1, \alpha_2, \theta_2) = u_q(q_1, \alpha_1, \theta_1) \text{ for some } \quad (\alpha_2, \theta_2) \quad \text{s.t. } \alpha_1 > \alpha_2.$$

Then type (α_2, θ_2) has a unique optimal quantity in the mechanism, q_1. Thus we have $q(\alpha_2, \theta_2) = q_1$.

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

Lemma 3 shows how one can start from any type (α_1, θ_1) for which the quantity q_1 is optimal, and identify an interval of types with lower values of the parameter α for which q_1 is the unique optimum and which therefore lie on the same isoquant. Accordingly, let us define:

$$I(q, \alpha, \theta) = \{(\alpha', \theta') : \; u_q(q, \alpha', \theta') = u_q(q, \alpha, \theta), \; \alpha' < \alpha \; \} \tag{3}$$

We will refer to the set $I(q, \alpha, \theta)$ as the left q-isoquant from (α, θ), or simply q-isoquant.

For later reference, a convenient way to parameterize an isoquant $I(q, \alpha, \theta)$ is by introducing a function $\sigma(q, \alpha, \theta, a)$ defined as a solution in σ to the following equation for $a \in [0, \alpha]$:

$$\begin{aligned} & u_q(q, \sigma, a) = u_q(q, \theta, \alpha), \;\; \text{if} \;\; u_q(q, 1, a) \geq u_q(q, \theta, \alpha) \geq u_q(q, 0, a) \\ & \sigma = 1, \;\; \text{if} \;\; u_q(q, 1, a) < u_q(q, \theta, \alpha) \\ & \sigma = 0, \;\; \text{if} \;\; u_q(q, 0, a) > u_q(q, \theta, \alpha) \end{aligned} \tag{4}$$

Next, let us define the set L as the union of the lower boundary and the right boundary of the participation region:

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

Lemma 3 suggests that all isoquants in the participation region emanate from the boundary L. Showing this, however, requires some effort, in particular, because multiple isoquants may emanate from the same type $(\alpha, \theta) \in L$ that has multiple optimal quantities.

For this reason, we will define a correspondence $Q^* : L \rightrightarrows \mathbf{R}_+$ assigning quantities along L as the upper-hemicontinuous closure of quantities assigned by the mechanism in the participation region:

$$Q^*(\alpha, \theta) = \{q : q = \lim_{n \to \infty} q(\alpha_n, \theta_n), \text{ for some sequence } \{(\alpha_n, \theta_n)\} \subset \Omega_+ \text{ s.t. } (\alpha_n, \theta_n) \to (\alpha, \theta)\}$$

Building in part on Lemma 3, we establish the following important properties of $Q^*(.)$:

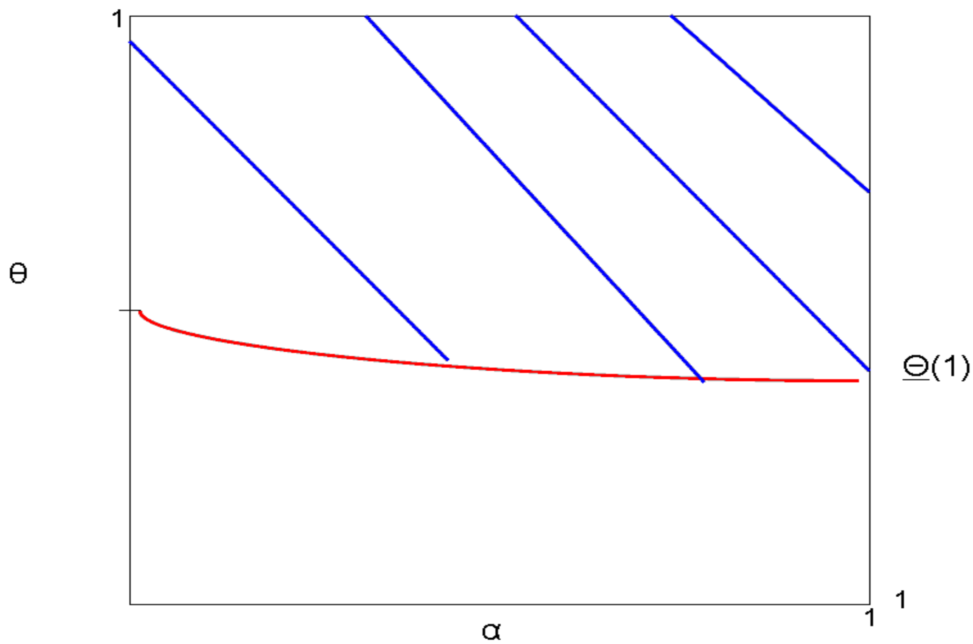
Lemma 4 *In an optimal mechanism,*
(i) The allocation $q(\alpha, \theta)$ is continuous on the interior of Ω_+.
(ii) Any $q \in Q^*(\alpha, \theta)$ is an optimal quantity choice for type $(\alpha, \theta) \in L$.
(iii) The correspondence $Q^*(\alpha, \theta)$ is increasing along L, [6] *upper-hemicontinuous, closed- and convex-valued on L.*

For almost all $(\alpha, \theta) \in L$, $Q^*(\alpha, \theta)$ is a singleton.
(iv) For any $(\alpha_1, \theta_1) \in L$ and $q_1 > 0$ s.t. $q_1 \in Q^*(\alpha_1, \theta_1)$ we have: $I(q_1, \alpha_1, \theta_1) \cap L = (\alpha_1, \theta_1)$.

The next Theorem establishes that an incentive compatible individually rational mechanism induces such boundary L and quantity correspondence $Q^*(.)$ along it that every point inside the participation region Ω_+ lies on an isoquant emanating from L and the transfer associated with any quantity offered in this mechanism can be backed out from $Q^*(.)$.

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

Figure 2: Typical Isoquant Map.



Theorem 1 *Consider some incentive compatible individually rational mechanism* $(q(.), t(.))$. *Then for every* $(\alpha, \theta) \in \Omega_+$ *there exists a unique* $(\alpha', \theta') \in L$ *such that* $q(\alpha, \theta) \in Q^*(\alpha', \theta')$. *Furthermore, let* $\theta^m(q) = \max\{\theta | q \in Q^*(1, \theta)\}$. *We have:*

$$t(\alpha, \theta) = \begin{cases} u(q(\alpha, \theta), \alpha', \theta'), \; \text{if } \alpha' < 1, \\ u(\min Q^*(1, \underline{\theta}(1)), 1, \underline{\theta}(1)) + \int_{\min Q^*(1,\underline{\theta}(1))}^{q(\alpha,\theta)} u_q(q, 1, \theta^m(q))dq \;\; \text{if } \alpha' = 1 \end{cases}$$

Theorem 1 shows that every point in the participation region lies on one isoquant emanating from a point on the boundary L, and links the transfer function $t(.)$ in a mechanism to the allocation along the boundary L. A typical isoquant map is depicted in Figure 2. Thus, we reach a key conclusion that all isoquants, as well as quantity allocations and the transfers over the whole domain are completely determined by the optimal quantity correspondence $Q^*(\alpha, \theta)$ along the boundary L. This suggests that a mechanism is completely defined by the quantity allocation along the boundary L as well as the location of this boundary.

It turns out that the optimal mechanism often has points on the lower boundary from which multiple isoquants emanate. At such points, the quantity allocation along the lower boundary will be discontinuous. So formulating our problem in terms of a quantity allocation along L would require complicated methods of impulse control.

To circumvent this problem, we instead formulate the problem as selecting an assignment of types to quantities along L. Formally, let us introduce a notion of an admissible 5-tuple $(q_0, \hat{q}, \bar{q}, \alpha(.), \theta(.))$ where $q_0, \hat{q}, \bar{q} \in \mathbf{R}_+$ are such that $q_0 \leq \hat{q} \leq \bar{q}$, and $\alpha(\cdot)$ and $\theta(\cdot)$ are two functions from $[q_0, \bar{q}]$ to $[0, 1]$. The admissibility requirements are that $\alpha(\cdot)$ is nondecreasing and absolutely continuous with $\alpha(q) = 1$ on $[\hat{q}, \bar{q}]$. Meanwhile, $\theta(\cdot)$ is nonincreasing on $[q_0, \hat{q}]$ where it satisfies the differential equation $\theta'(q) = -\alpha'(q)\frac{u_\alpha(q,\alpha(q),\theta(q))}{u_\theta(q,\alpha(q),\theta(q))}$, and is nondecreasing and absolutely continuous on $[\hat{q}, \bar{q}]$ with $\theta(\bar{q}) = 1$. Finally, if $q_0 > 0$ then we either have $\alpha(q_0) = 0$ or $\theta(q_0) = 1$.[7]

To describe how an admissible 5-tuple induces an incentive compatible mechanism, we also need to operate with quantity assignments to types induced by this 5-tuple. Particularly, the induced quantity assignment $q(\alpha)$ is a correspondence mapping $[0, 1]$ into $\mathbf{R}_+$ and is given by: $q(\alpha) = [q_l(\alpha), q_u(\alpha))]$ where $q_l(\alpha) = \inf\{q \geq q_0 : \alpha(q) \geq \alpha\}$ and $q_u(\alpha) = \sup\{|q \geq q_0 : \alpha(q) \leq \alpha\}]$ for $\alpha \in [0, 1]$. Note that $q(\alpha)$ is a singleton everywhere except such points α'' that $\alpha(q) = \alpha^{''}$ for $q \in [q_u, q_h]$. In this case $q(\alpha'') = [q_u, q_h]$. The set of such α'' is at most countable, so $q(\alpha)$ is a singleton almost everywhere on $[0, 1]$. Note that $q(.)$ is convex-valued, closed and continuous because $\alpha(q)$ is increasing.

The lower boundary induced by the 5-tuple then equals $\underline{\theta}(\alpha)) = \theta(q(\alpha))$ for all $\alpha \in [0, 1]$ whenever $\theta(q_0) < 1$, and $\underline{\theta}(\alpha)) = \theta(q(\alpha))$ for all $\alpha \in [\alpha(q_0), 1]$, otherwise. Observe that $\theta(q(\alpha))$ is well-defined. This is immediate when $q(\alpha)$ is a singleton. On the other hand, when $q(\alpha)$ is an interval $[q_1, q_2]$, then $\alpha'(q) = 0$ and hence by construction $\theta'(q) = 0$ on $[q_1, q_2]$. Therefore, we can take any value $q(\alpha)$ as an argument in $\theta(.)$ in the definition of $\underline{\theta}(\alpha))$. Also, note that, whenever $\underline{\theta}(\alpha) < 1$, it satisfies the differential equation $\frac{d\underline{\theta}(\alpha)}{d\alpha} = -\frac{u_\alpha(q(\alpha),\alpha,\underline{\theta}(\alpha))}{u_\theta(q(\alpha),\alpha,\underline{\theta}(\alpha))}$ with initial condition $\underline{\theta}(1) = \theta(\hat{q})$.

Similarly, for $\theta \in [\theta(\hat{q}), 1]$, the quantity allocation $q(\theta)$ induced by the 5-tuple satisfies $q(\theta) = [q_l(\theta), q_u(\theta))]$ where $q_l(\theta) = \inf\{q \geq \hat{q}|\theta(q) \geq \theta\}$ and $q_u(\theta) = \sup\{q \geq \hat{q}|\theta(q) \leq \theta\}$. Since $\theta(.)$ is increasing on $[\hat{q}, \bar{q}]$, $q(\theta)$ is a closed, continuous, and convex-valued correspondence. It is also a singleton for almost all $\theta \in [\theta(\hat{q}), 1]$.

Now we can define:

Definition 1 *A mechanism (q, t) is said to be induced by an admissible 5-tuple $(q_0, \hat{q}, \bar{q}, \alpha(.), \theta(.))$ if this mechanism satisfies the following conditions:*
(i) The lower boundary in the mechanism is $\underline{\theta}(\alpha) = \theta(q(\alpha))$ for all $\alpha \in [0, 1]$ whenever $\theta(q_0) < 1$, and $\underline{\theta}(\alpha)) = \theta(q(\alpha)$ for all $\alpha \in [\alpha(q_0), 1]$, otherwise.
(ii) The quantity allocation $q(\alpha, \underline{\theta}(\alpha))$ along the lower boundary $\underline{\theta}(\alpha)$ is such that $q(\alpha, \underline{\theta}(\alpha)) = q(\alpha)$.
The quantity allocation along the right boundary, $q(1, \theta)$, is defined by $q(1, \theta) = q(\theta)$ for $\theta \in [\theta(\hat{q}), 1]$.

[7] If $q_0 > 0$ and we had $\alpha(q_0) > 0$ and $\theta(q_0) < 1$, then any type on the isoquant emanating from $(\alpha(q_0), \theta(q_0))$, other than $(\alpha(q_0), \theta(q_0))$ itself, would necessarily obtains a strictly positive net surplus. This would mean that the lower boundary extends to the left of this isoquant, i.e. the leftmost point on the lower boundary then cannot be $(\alpha(q_0), \theta(q_0))$.

(iii) The tariff $P(q)$ associated with quantity $q \in [q_0, \bar{q}]$ is defined as follows.

$$P(q) = \begin{cases} u(q, \alpha(q), \underline{\theta}(\alpha(q))) \text{ if } q \in [q_0, \hat{q}], \\ u(\hat{q}, 1, \theta(\hat{q})) + \int_{\hat{q}}^{q} u_q(z, 1, \theta(z))dz \text{ if } q \in (\hat{q}, \bar{q}] \end{cases} \quad (5)$$

(iv) $t(\alpha, \underline{\theta}(\alpha)) = u(q_u(\alpha), \alpha, \underline{\theta}(\alpha))$ *for* $\alpha \in [0,1]$, *and* $t(1, \theta) = u(\hat{q}, 1, \theta(\hat{q})) + \int_{\hat{q}}^{\sup q(\theta)} u_q(q, \theta(q), 1)dq$ *for* $\theta \in [\theta(\hat{q}), 1]$.

(v) For every (α, θ) *s.t.* $\alpha < 1$ *and* $\theta > \underline{\theta}(\alpha)$ *there exists either:*

(a) $\breve{q} \in [q_0, \hat{q}]$ *and* $\breve{\alpha} \in [0,1]$ *such that* $\breve{q} \in [q_l(\breve{\alpha}), q_u(\breve{\alpha})]$ *and* $u_q(\breve{q}, \alpha, \theta) = u_q(\breve{q}, \breve{\alpha}, \underline{\theta}(\breve{\alpha}))$. *In this case,* $q(\alpha, \theta) = \breve{q}$, *and* $t(\alpha, \theta) = P(\breve{q})$;

(b) or $\theta^\dagger \in [\hat{\theta}, 1]$ *and* $q^\dagger \in [\hat{q}, \bar{q}]$ *s.t.* $q^\dagger \in [q_l(\theta^\dagger), q_u(\theta^\dagger))]$ *and* $u_q(q^\dagger, \alpha, \theta) = u_q(q^\dagger, 1, \theta^\dagger)$. *In this case,* $q(\alpha, \theta) = q^\dagger$ *and* $t(\alpha, \theta) = P(q^\dagger)$.

Then we have:

Theorem 2 *Suppose Assumptions 1 and 2 hold, and* $\sup \frac{u_\alpha}{u_\theta}(q, \alpha, \theta) < \infty$ *for all* $(\alpha, \theta) \in [0,1]$ *and* $\bar{q} \in [0, q^{FB}(1,1)]$, *where* $q^{FB}(1,1) = \arg\max_q u(q, 1, 1)$.

Then the mechanism (q, t) *induced by an admissible 5-tuple* $(q_0, \hat{q}, \bar{q}, \alpha(.), \theta(.))$ *is well-defined, incentive compatible and individually rational.*

Theorem 2 together with Definition 1 show how to construct an incentive compatible individually rational mechanism from an admissible 5-tuple $(q_0, \hat{q}, \bar{q}, \alpha(.), \theta(.))$.

In particular, the incentive and individual rationality constraints along the boundary L allow us to solve for the transfer associated with each quantity in this mechanism. Finally, the quantity allocation at any point (α, θ) s.t. $\alpha < 1$ and $\theta > \underline{\theta}(\alpha)$ is uniquely defined by the isoquant to which (α, θ) belongs according to Lemma 3.

4 The Reformulated Problem

Taken together, Theorems 1 and 2 establish a one-to-one relationship between the set of incentive compatible individually rational direct mechanisms, on the one hand, and, on the other hand, the set of admissible 5-tuples $(q_0, \hat{q}, \bar{q}, \alpha(.), \theta(.))$. In this section, we will use this isomorphism to reformulate the optimal design problem (1) as an optimal choice of an admissible 5-tuple and characterize the optimal mechanism.

We will henceforth assume that the functions $\alpha(.)$ and $\theta(.))$ are piecewise continuously differentiable on $[q^0, \hat{q}]$ and $[\hat{q}, \bar{q}]$, respectively. Because piecewise continuously differentiable functions are dense in the set of measurable functions, such a solution must also be a solution on the domain of measurable functions.

To state our mechanism design problem as a choice of the optimal 5-tuple $(q_0, \hat{q}, \bar{q}, \alpha(.), \theta(.))$, we need to determine the seller's expected revenue associated with it. We can compute the probability measure on the set of quantities induced by such 5-tuple as follows. By Theorem 2 the set of types assigned quantities that exceed q is given by:

$$\{(a,\theta)\in[0,1]^2 : a\le\alpha(q),\ \theta\ge\sigma(q,\alpha(q),\theta(q),a)\}\cup\{(a,\theta)\in[0,1]^2 : a\ge\alpha(q),\ \theta\ge\underline{\theta}(a)\}$$

Therefore, the probability measure of this set of types is

$$H(q,\alpha(q),\theta(q)) = \int_0^{\alpha(q)}\int_{\sigma(q,\alpha(q),\theta(q),a)}^1 f(a,\theta)d\theta da + \int_{\alpha(q)}^1\int_{\underline{\theta}(a)}^1 f(a,\theta)d\theta da = \int_0^1\int_{\max\{\sigma(q,\alpha(q),\theta(q),a),\underline{\theta}(a)\}}^1 f(a,\theta)d\theta da \tag{6}$$

Correspondingly, the probability measure of the set of types assigned quantities no larger than q equals $1-H(q,\alpha(q),\theta(q))$. The points of discontinuity of $\alpha(q)$ correspond to atoms of the probability distribution $1-H(q,\alpha(q),\theta(q))$. Particularly, the size of an atom at a quantity $\tilde{q}$ is equal to $\lim_{q\uparrow\tilde{q}} H(q,\alpha(q),\theta(q)) - H(\tilde{q},\alpha(\tilde{q}),\theta(\tilde{q}))$.

When $\alpha(.)$ and $\theta(.)$ are differentiable, the density of of types assigned the quantity q can be computed as follows:

$$h(q,\alpha(q),\theta(q),\alpha'(q),\theta'(q)) = \int_0^\alpha f(\sigma(q,\alpha,\theta,a),a)[\sigma_q(q,\alpha,\theta,a)+\sigma_\theta(q,\alpha,\theta,a)\theta'+\sigma_\alpha(q,\alpha,\theta,a)\alpha']da \tag{7}$$

where, from equation 4, we have:

$$\sigma_q(q,\alpha,\theta,a) = \frac{u_{qq}(q,\alpha,\theta)-u_{qq}(q,a,\sigma)}{u_{q\theta}(q,a,\sigma)} \tag{8}$$

$$\sigma_\theta(q,\alpha,\theta,a) = \frac{u_{q\theta}(q,\alpha,\theta)}{u_{q\theta}(q,a,\sigma)} \tag{9}$$

$$\sigma_\alpha(q,\alpha,\theta,a) = \frac{u_{q\alpha}(q,\alpha,\theta)}{u_{q\theta}(q,a,\sigma)} \tag{10}$$

Then the seller's expected profit in the mechanism is equal to

$$ER = \int_{q_0}^{\bar{q}} P(q)d\left(1-H(q,\alpha(q),\theta(q)\right) \tag{11}$$

where $P(q)$ is the tariff in expression (5) in Theorem 2. Using $(\alpha(q),\theta(q))$ in (5) yields:

$$P(q) = t(\alpha(q),\theta(q)) = \left\{\begin{array}{cc} u(q,\alpha(q),\theta(q)), & \text{for all } q\in[q_0,\hat{q}] \\ u(\hat{q},\alpha(\hat{q}),\theta(\hat{q})) + \int_{\hat{q}}^q u_q(z,\alpha(z),\theta(z))dz, & \text{for all } q\in[\hat{q},\bar{q}] \end{array}\right.,$$

Finally, substituting this expression for $P(q)$ into (11) and then integrating the second integral of the first expression by parts and using $H(\bar{q},\alpha(\bar{q}),\theta(\bar{q}))=0$ yields:

$$\int_{q_0}^{\hat{q}} u(q,\alpha(q),\theta(q))h(q,\alpha(q),\theta(q),\alpha'(q),\theta'(q))dq + \int_{\hat{q}}^{\bar{q}}\left\{u(\hat{q},1,\theta(\hat{q}))+\int_{\hat{q}}^q u_q(z,1,\theta(z))dz\right\}d(1-H(q,1,\theta(q))) =$$
$$\int_{q_0}^{\hat{q}} u(q,\alpha(q),\theta(q))h(q,\alpha(q),\theta(q),\alpha'(q),\theta'(q))dq + u(\hat{q},1,\theta(\hat{q}))H(\hat{q},1,\theta(\hat{q})) + \int_{\hat{q}}^{\bar{q}} H(q,\alpha(q),\theta(q))u_q(q,\alpha(q),\theta(q))dq \tag{12}$$

Equation (12) highlights that the monopolist's profits consists of two parts. The first part depends only upon the allocation $(\alpha(q), \theta(q))$ for $q \leq \hat{q}$ on the lower boundary. As the types on the lower boundary earn zero surplus, the associated transfer is equal to the gross utility of the type $(\alpha(q), \theta(q))$. The set of types from which this transfer is collected is the isoquant through the point $(\alpha(q), \theta(q))$, and hence the associated density of q is $h(q, \alpha(q), \theta(q), \alpha'(q), \theta'(q))$.

The second part of the monopolist's expected profit comes from the quantities $q \in [\hat{q}, \bar{q}]$. From each type that consumes more than $\hat{q}$ the monopolist, first, collects $u(\hat{q}, 1, \theta(\hat{q}))$, the transfer paid by type $(1, \theta(\hat{q}))$. The probability measure of these types is $H(\hat{q}, 1, \theta(\hat{q}))$. Second, the monopolist collects the marginal price $u_q(q, \alpha(q), \theta(q))$ from each type that consumes more than q, of which there are $H(q, \alpha(q), \theta(q))$.

Inspecting (12) one can see immediately that the monopolist's optimization problem can be split into the following three subproblems.

Subproblem (i). For fixed $\hat{q} \in \mathbf{R}_+$ and $\widehat{\theta} \in [0, 1]$ choose $q_0 \in \mathbf{R}_+$ and functions $\alpha(q)$ and $\theta(q)$ to solve

$$W(\widehat{q}, \widehat{\theta}) = \max \int_{q_0}^{\hat{q}} u(q, \alpha(q), \theta(q)) h(q, \alpha(q), \theta(q), \alpha'(q), \theta'(q)) dq \tag{13}$$

subject to the following constraints:

$$\begin{aligned} &\alpha(q_0) \geq 0, \alpha(\widehat{q}) = 1, \ \theta(\widehat{q}) = \widehat{\theta} \\ &\alpha'(q) \geq 0 \quad \theta'(q) = -\frac{u_\alpha(q, \theta(q), \alpha(q))}{u_\theta(q, \theta(q), \alpha(q))} \alpha'(q) \end{aligned} \tag{14}$$

Subproblem (ii). Given $\widehat{q} \in \mathbf{R}_+$ and $\widehat{\theta} \in [0, 1]$, choose $\bar{q} \in \mathbf{R}_+$ s.t. $\hat{q} \leq \bar{q}$ and a nondecreasing functions $\theta(.)$ to solve:

$$Z(\widehat{q}, \widehat{\theta}) = \max \int_{\widehat{q}}^{\bar{q}} H(q, 1, \theta(q)) u_q(q, 1, \theta(q)) dq \tag{15}$$

subject to the constraint $\theta(\bar{q}) = 1$.

Subproblem (iii). Choose $\widehat{q} \in \mathbf{R}_+$ and $\widehat{\theta} \in [0, 1]$ to solve:

$$V(\widehat{q}, \widehat{\theta}) = \max_{\widehat{q}, \widehat{\theta}} W(\widehat{q}, \widehat{\theta}) + u(\widehat{q}, 1, \widehat{\theta}) H(\widehat{q}, 1, \widehat{\theta}) + Z(\widehat{q}, 1, \widehat{\theta}) \tag{16}$$

4.1 Solution to Subproblem (i)

We will first state Subproblem (i) as an optimal control problem. To simplify the notation, let us rewrite the expression (7) for the density $h(q)$ as follows:

$$h(q, \alpha, \theta, \alpha', \theta') = h_0(q, \alpha, \theta) + \left(h_2(q, \alpha, \theta) - \frac{u_\alpha}{u_\theta}(q, \alpha, \theta) h_1(q, \alpha, \theta) \right) \alpha'(q). \tag{17}$$

where:

$$
\begin{aligned}
h_0(q,\alpha,\theta) &= \int_{\underline{\alpha}(q,\alpha,\theta)}^{\alpha} f(\sigma(q,\alpha,\theta,a),a)\sigma_q(q,\sigma(q,\alpha,\theta,a),a)da && (18)\\
h_1(q,\alpha,\theta) &= \int_{\underline{\alpha}(q,\alpha,\theta)}^{\alpha} f(\sigma(q,\alpha,\theta,a),a)\sigma_\theta(q,\sigma(q,\alpha,\theta,a),a)da && (19)\\
h_2(q,\theta,\alpha) &= \int_{\underline{\alpha}(q,\alpha,\theta)}^{\alpha} f(\sigma(q,\alpha,\theta,a),a)\sigma_\alpha(q,\sigma(q,\alpha,\theta,a),a)da. && (20)
\end{aligned}
$$

and where $\underline{\alpha}(q,\alpha,\theta)$ is the solution in a to the equation $\sigma(q,\alpha,\theta,a)=1$, with $\sigma(.)$ given by equation (4). That is $u_q(q,\alpha,\theta)=u_q(q,\underline{\alpha}(q,\alpha,\theta),1)$ if such a solution exists, and $\underline{\alpha}(q,\alpha,\theta)=0$ otherwise i.e., if $u_q(q,\alpha,\theta)\le u_q(q,0,1)$. (In the latter case, there exists $\theta'\in[0,1)$ such that $u_q(q,\alpha,\theta)=u_q(q,0,\theta')$).

Substituting (17) into the objective (13), we can now form the Hamiltonian for subproblem (i) as follows:

$$
J(q,\alpha,\theta,\alpha',\mu,\lambda)=uh+\mu\alpha'-\lambda\frac{u_\alpha}{u_\theta}\alpha'=uh_0+\left(u\left(h_2-\frac{u_\alpha}{u_\theta}h_1\right)+\left(\mu-\lambda\frac{u_\alpha}{u_\theta}\right)\right)\alpha', \quad (21)
$$

where μ and λ are the multipliers on the state evolution equations for α and θ, respectively, and α' is the control variable.

The linearity of the Hamiltonian (21) in the control variable α' creates certain technical difficulties for solving subproblem (i), as it implies that α' cannot be solved for directly from the standard first-order conditions of optimality. However, Pontryagin's Maximum principle still applies and requires that the optimal control $\alpha'\ge 0$ maximizes the Hamiltonian (21). Particularly, let

$$
S(q,\alpha(q),\theta(q),\mu(q),\lambda(q))=u\left(h_2-\frac{u_\alpha}{u_\theta}h_1\right)+\left(\mu-\lambda\frac{u_\alpha}{u_\theta}\right) \quad (22)
$$

The function $S(q,\alpha(q),\theta(q),\mu(q),\lambda(q))$ is called the *switching function*. Note that it can never be strictly positive, since then the value of the objective would be infinite. Optimality requires the following "switching conditions" to hold:

$$
\begin{aligned}
S(q,\alpha(q),\theta(q),\mu(q),\lambda(q)) &< 0\Rightarrow\alpha'=0\\
S(q,\alpha(q),\theta(q),\mu(q),\lambda(q)) &= 0\Rightarrow\alpha'\ge 0
\end{aligned}
$$

An interval of q on which S vanishes ($S=0$) is called a *singular arc*. On a singular arc, the optimality conditions do not pin down the value of the optimal control α'. As a consequence, such problems of singular control are notoriously difficult to solve. Only a few solutions have been developed up to now, most notably Merton (1969)'s celebrated portfolio choice problem in finance, and trajectory optimization in aeronautics (see e.g. Bryson and Ho (1975) Ch. 8).

The approach we follow here is to recover the optimal control α' along a singular arc by differentiating the identity $S=0$ with respect to q until the control variable appears in a non-trivial way, and then solve for it.

An interval of q on which $S < 0$ is a *nonsingular arc*. As pointed above, $\alpha'(q) = 0$ for all q on a non-singular arc. Pontryagin's Maximum principle yields the remaining optimality conditions along such an arc.

This still leaves the difficult problem of finding where to join singular and nonsingular arcs. A point q where singular and a nonsingular arcs meet is called a *junction point*. As is apparent from the switching conditions, at a junction point the optimal control may be discontinuous.

To state the main result of this subsection, let us define:

$$N(q,\alpha,\theta) = u_\theta \int_{\underline{\alpha}(q,\alpha,\theta))}^{\alpha} f(a,\sigma)\frac{2u_{qq}(q,\alpha,\theta) - u_{qq}(q,a,\sigma))}{u_{q\theta}(q,a,\sigma)}da - u_q u_\theta \frac{\partial \underline{\alpha}}{\partial q}\frac{f(\sigma(q,\alpha,\theta,\underline{\alpha}(q,\alpha,\theta)),\underline{\alpha}(q,\alpha,\theta))}{u_{q\theta}(q,\underline{\alpha}(q,\alpha,\theta),\sigma(q,\alpha,\theta,\underline{\alpha}(q,\alpha,\theta)))} + u_q u_\theta \int_{\underline{\alpha}(q,\alpha,\theta))}^{\alpha}\left(f_\theta(a,\sigma)u_{q\theta}(q,\alpha,\theta) - f(a,\sigma)u_{qq\theta}(q,\alpha,\theta) - f(a,\sigma)u_{q\theta\theta}(q,a,\sigma)\frac{u_{q\theta}(q,\alpha,\theta)}{u_{q\theta}(q,a,\sigma)}\right)\frac{1}{u_{q\theta}(q,a,\sigma(q,\alpha,\theta))^2}da \tag{23}$$

$$D(q,\alpha,\theta) = \left(u - \frac{u_\theta u_q}{u_{q\theta}}\right)f(\alpha,\theta) + (u_{q\theta}u_\alpha - u_{q\alpha}u_\theta)\left(2\int_{\underline{\alpha}(q,\alpha,\theta)}^{\alpha}\frac{f(a,\sigma)}{u_{q\theta}(q,a,\sigma)}da + u_q\int_{\underline{\alpha}(q,\alpha,\theta)}^{\alpha}\frac{f_\theta u_{q\theta} - f u_{q\theta\theta}}{u_{q\theta}^3}(q,a,\sigma)da\right) + u_q\left(\frac{\partial\underline{\alpha}}{\partial\alpha}u_\theta - \frac{\partial\underline{\alpha}}{\partial\theta}u_\alpha\right)\frac{f(\sigma(q,\alpha,\theta,\underline{\alpha}(q,\alpha,\theta)),\underline{\alpha}(q,\alpha,\theta))}{u_{q\theta}(q,\underline{\alpha}(q,\alpha,\theta),\sigma(q,\alpha,\theta,\underline{\alpha}(q,\alpha,\theta)))} \tag{24}$$

where

$$\frac{\partial\underline{\alpha}}{\partial q} = 0,\ \frac{\partial\underline{\alpha}}{\partial\alpha} = 0,\ \frac{\partial\underline{\alpha}}{\partial\theta} = 0,\ \text{if } \underline{\alpha} = 0$$
$$\frac{\partial\underline{\alpha}}{\partial q} = \frac{u_{qq}(q,\alpha,\theta) - u_{qq}(q,\underline{\alpha},1)}{u_{q\alpha}(q,\underline{\alpha},1)},\ \frac{\partial\underline{\alpha}}{\partial\alpha} = \frac{u_{q\alpha}(q,\alpha,\theta)}{u_{q\alpha}(q,\underline{\alpha},1)},\ \frac{\partial\underline{\alpha}}{\partial\theta} = \frac{u_{q\theta}(q,\alpha,\theta)}{u_{q\alpha}(q,\underline{\alpha},1)},\ \text{if } \underline{\alpha} > 0$$

Theorem 3 *The solution to the maximization problem (13) has the following properties:*
(i) Over any interval where $\alpha(q)$ is strictly increasing and hence $\theta(q)$ is strictly decreasing we have:

$$\alpha'(q) = \frac{N(q,\alpha(q),\theta(q))}{D(q,\alpha(q),\theta(q))} \tag{25}$$
$$\theta'(q) = -\frac{u_\alpha}{u_\theta}\alpha'. \tag{26}$$
$$\lambda(q) = \left(\frac{u_q u_\theta}{u_{q\theta}} - u\right)h_1 \tag{27}$$
$$\mu(q) = \frac{u_q u_\alpha}{u_{q\theta}}h_1 - uh_2 \tag{28}$$

(ii) Over any interval on which α, and hence θ, are constant, we have:

$$\dot{\mu} = -u_\alpha h_0 - u\frac{\partial h_0}{\partial\alpha} \tag{29}$$
$$\dot{\lambda} = -u_\theta h_0 - u\frac{\partial h_0}{\partial\theta} \tag{30}$$

(iii) The functions $\mu(q)$ and $\lambda(q)$ are continuous.
(iv) We have: $\alpha(q_0)q_0 = 0$.

Theorem 3 provides the optimal solution on every singular and non-singular arc. It also gives a partial answer regarding the location of such arcs. Particularly, by part (iii) of this Theorem, the junction points between singular and non-singular arcs must be chosen so that the costate functions μ and λ remain continuous throughout. Below, we will explore this property further to provide a more detailed characterization of the solution.

The following two Lemmas characterize the properties of the solution to subproblem (i) which turn out to be useful for constructing the overall solution to our problem. Recall that $D(q, \alpha(q), \theta(q))$ is the denominator of (25) and is given by (24).

Lemma 5 ***(Generalized Legendre-Clebsch):*** *The solution to the subproblem (i) is such that* $D(q, \alpha(q), \theta(q)) \leq 0$ *and* $N(q, \alpha(q), \theta(q)) \leq 0$ *along any optimal singular arc.*

Using Lemma 5 we can establish the following property of the optimal q_0:

Lemma 6 *Suppose that* $u - \frac{u_\theta u_q}{u_{q\theta}} > 0$ *for all* (q, α, θ) *s.t.* $q > 0$*. Then* $q_0 = 0$*.*

Note that the condition that $u - \frac{u_\theta u_q}{u_{q\theta}} > 0$ for all (q, α, θ) holds for most commonly specified utility functions. In particular, it 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 $u_{qq\theta} \geq 0$ since $\frac{\partial}{\partial q}(u - \frac{u_\theta u_q}{u_{q\theta}}) = -\frac{u_{qq} u_\theta}{u_{q\theta}} + \frac{u_{qq\theta} u_q u_\theta}{u_{q\theta}^2}$. We will maintain the assumption of Lemma 6 henceforth.

Finally, we characterize the regions where the solution to subproblem (i) consists of a singular arc and where it consists of a non-singular arc. For this purpose, let q^* be such that $\sigma(q^*, \alpha(q^*), \theta(q^*), 0) = 1$. In words, the isoquant corresponding to q^*, $I(q^*, \alpha(q^*), \theta(q^*))$, hits the "northwest" corner $(0, 1)$ of the type space. Consequently, the density function $h(q, \alpha(q), \theta(q), \alpha'(q))$ is discontinuous at q^*. To see this, consider the lower limit of the integrals in (18)-(20), $\underline{\alpha}(q, \alpha(q), \theta(q))$. It is not continuously differentiable at q^* since its total derivative from the left is zero, while its right-hand side derivative is strictly positive. For this reason, we need to use the methods of nonsmooth optimal control to handle our problem. In particular, we rely on (Ioffe and Rockafellar 1996) to establish the continuity of the Lagrange multipliers in part (iii) of Theorem 3 above.

The following assumption imposes regularity conditions on $N(q, \alpha, \theta)$:

Assumption 3 *(i) If* $N(q, \alpha, \theta) \geq 0$*, then* $N_q(q, \alpha, \theta) \geq 0$*, for all* (q, α, θ)*;*
(ii) $N_{q\alpha}(q, \alpha, \theta) < 0$ *at* $q = 0$*.*

Under this assumption, the solution to subproblem (i) takes on a particularly simple form:

Theorem 4 *Suppose that the Assumption in Lemma 6 and Assumption 3 hold, and suppose that* $q^* < \widehat{q}$*. Then the solution to subproblem (i) is a nonsingular arc on* $[0, q^*]$ *and a singular arc on* $(q^*, \widehat{q}]$*.*

Theorem 4 says that the optimal solution on the interval $[0, q^*]$ is a nonsingular, and thus the isoquants for all quantities $q \in [0, q^*]$ emanate from the single point $(\alpha(q^*), \theta(q^*))$. Hence, the type $(\alpha(q^*), \theta(q^*))$ is indifferent between all quantities $q \in [0, q^*]$. In other words, we have a clustering of quantities at the point $(\alpha(q^*), \theta(q^*))$.

This implies that we have a discontinuity in the allocation assigned to types on the lower boundary at the point $(\alpha(q^*), \theta(q^*))$. Unlike in the one-dimensional type case, this discontinuity is not associated with gaps in the consumption schedule.

4.2 Solution to subproblem (ii)

Next, let us consider maximization subproblem (ii). It has fixed initial "time" $\widehat{q}$, free terminal "time" $\bar{q}$, and fixed initial and terminal boundaries $\hat{\theta}$ and 1, respectively, and the monotonicity constraint that $\theta(.)$ is nondecreasing. We incorporate the latter by forming a Lagrangian which includes the constraint $\theta'(q) \geq 0$ with Lagrange multiplier $\delta(q)$ associated with it:

$$\max \int_{\widehat{q}}^{\bar{q}} u_q(q, 1, \theta) H(q, 1, \theta) + \delta\theta' dq \tag{31}$$

Let $\phi(q, \theta)$ be the derivative of the integrand of the original objective w.r.t. θ i.e.,

$$\phi(q, \theta) \quad = \quad u_q(q, 1, \theta) H_\theta(q, 1, \theta) + u_{\theta q}(q, 1, \theta) H(q, 1, \theta) \tag{32}$$

Also, let $\theta^\phi(q)$ be the solution to $\phi(q, \theta) = 0$, when such exists; $\theta^\phi(q) = 1$ if $\phi(q, \theta) > 0$ for all $\theta \in [0, 1]$; $\theta^\phi(q) = 0$ if $\phi(q, \theta) < 0$ for all $\theta \in [0, 1]$. Observe that the condition $\phi(q, \theta) = 0$ is a multi-dimensional version of a condition familiar from the one-dimensional problem, that at the optimum marginal virtual surplus must be equal to zero.[8] The solution to subproblem (ii) is characterized in the following Theorem:

Theorem 5 *The solution to subproblem (ii) is as follows.*

1. *The optimal $\bar{q}$ solves $u_q(\bar{q}, 1, 1) = 0$.*
2. *If $\phi(.)$ is increasing in q and decreasing in θ on $[\widehat{q}, \bar{q}]$, then $\theta(q) = \max\{\theta^\phi(q), \widehat{\theta}\}$.*
3. *If the condition that $\phi(.)$ is increasing in q and decreasing in θ on $[\widehat{q}, \bar{q}]$ does not hold, then:*
   (a) Over any interval in $[\widehat{q}, \bar{q}]$ on which $\theta'(q) > 0$, we have $\phi(q, \theta(q)) = 0$.
   (b) Over any interval in $[\widehat{q}, \bar{q}]$ on which $\theta'(q) = 0$, we have $\delta(q) \geq 0$ and $\delta'(q) = \phi(q, \theta(q))$.

[8] Letting t denote the type in the one-dimensional screening problem and $F(t)$ denote its distribution function, the optimality condition is $u_q F' + u_{qt}(1 - F(t)) = 0$.

Theorem 5 shows that, if $\phi(.)$ is increasing in q and decreasing in θ, then the constraint $\theta'(q) \geq 0$ can be ignored, and subproblem (ii) is solved by pointwise maximization under the integrand.

When $\widehat{\theta} > \theta^{\phi}(\widehat{q})$, then there is a non-empty right neighborhood of $\widehat{q}$ over which all isoquants emanate from $(1, \widehat{\theta})$. We shall show in Theorem 6 that this cannot be optimal.

Since $\lim_{q \to \bar{q}(1)} H(q, 1, \theta(q)) = 0$ and $\lim_{q \to \bar{q}(1)} H_{\theta}(q, 1, \theta(q)) < 0$, we also obtain the familiar condition that the allocation of the "top" type $(1, 1)$ is undistorted i.e.,

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

If the condition that $\phi(.)$ is increasing in q and decreasing in θ does not hold, then the monotonicity constraints may be binding, and part (3) of Theorem 5 describes how the solution is obtained in this case.

4.3 Solution to subproblem (iii)

Next, we can combine the solutions to subproblems (i) and (ii) to characterize $\widehat{\theta}$ and $\widehat{q}$.

Theorem 4 implies that q^* must be a junction point, i.e. $S(q^*) = 0$. In conjunction with the fact that the isoquant emanating from the point (α^*, θ^*) must go through the northwest boundary point $(0, 1)$, the continuity of the Lagrange multipliers at q^*, and the singularity of the solution on the interval $(q^*, \widehat{q})$, this yields four equations in the four unknowns $(\alpha^*, \theta^*, q^*, \widehat{q})$ (see the first four equations in Theorem 7 below). As a consequence, subproblem (iii) has only one remaining variable to optimize over, i.e. $\widehat{\theta}$. Our next Theorem performs this optimization.

Theorem 6 *Suppose that the function $\phi(q, \theta)$ is increasing in q and decreasing in θ. Then at the optimum, $\widehat{\theta} = \theta^{\phi}(\widehat{q})$.*

We are now in a position to describe the overall solution to our problem whenever $q^* < \widehat{q}$:

Theorem 7 *Suppose that the assumption in Lemma 6 and Assumption 3 hold, and suppose that $q^* < \widehat{q}$. Then in the unique solution to problem (13) the triple $(q^*, \alpha^*, \theta^*)$ is characterized by the following system of equations:*

$$\begin{aligned} S(q^*) &= 0 \\ \sigma(q^*, \alpha^*, \theta^*, 1) &= 0 \\ \left(\frac{u_q u_\theta}{u_{q\theta}} - u \right) h_1 &= - \int_{q^*}^{\widehat{q}} \left(u_\theta h_0 - u \frac{\partial h_0}{\partial \theta} \right) dq = \lambda(q^*). \end{aligned} \tag{33}$$

Furthermore, the pair $(\widehat{q}, \widehat{\theta})$ satisfies:

$$\begin{aligned} \int_{q^*}^{\widehat{q}} \alpha'(q) dq &= 1 - \alpha^* \\ \widehat{\theta} &= \theta^{\phi}(\widehat{q}), \end{aligned}$$

On the interval $[0, q^*]$*, we have* $\alpha(q) = \alpha^*$ *and* $\theta(q) = \theta^*$*. On the interval* $[q^*, \widehat{q}]$ *the functions* $\alpha(q)$ *and* $\theta(q)$ *satisfy (25) and (26), and on the interval* $[\widehat{q}, \bar{q}]$ *we have* $\alpha(q) = 1$ *and* $\theta(q) = \theta^{\phi}(q)$*.*

When $q^* \geq \widehat{q}$, the region associated with subproblem (i) is empty, and so $\alpha(q) = 1$ for all q. Hence we have:

Theorem 8 *Suppose that the assumption in Lemma 6 and Assumption 3 hold, and suppose that* $q^* \geq \widehat{q}$*. Then* $\widehat{q} = 0$ *and for all* $q \in [0, \bar{q}]$ *we have* $\alpha(q) = 1$ *and* $\theta(q) = \theta^{\phi}(q)$*.*

Using this result, we can characterize necessary and sufficient conditions for the demand profile approach to yield the correct optimal screening mechanism:

Theorem 9 *Suppose that the function* $\phi(q, \theta)$ *is increasing in* q *and decreasing in* 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 quite stringent, as the example below will illustrate.

5 Optimal Mechanism - An Example

In this section, we derive an explicit solution in a special but quite prominent case where (α, θ) is uniformly distributed on the unit square $[0, 1]^2$ and the utility function is given by:

$$u(q, \alpha, \theta) = \theta q - \frac{b - \alpha}{2} q^{\gamma} \tag{34}$$

where $b > 1$ and $\gamma > 1$.

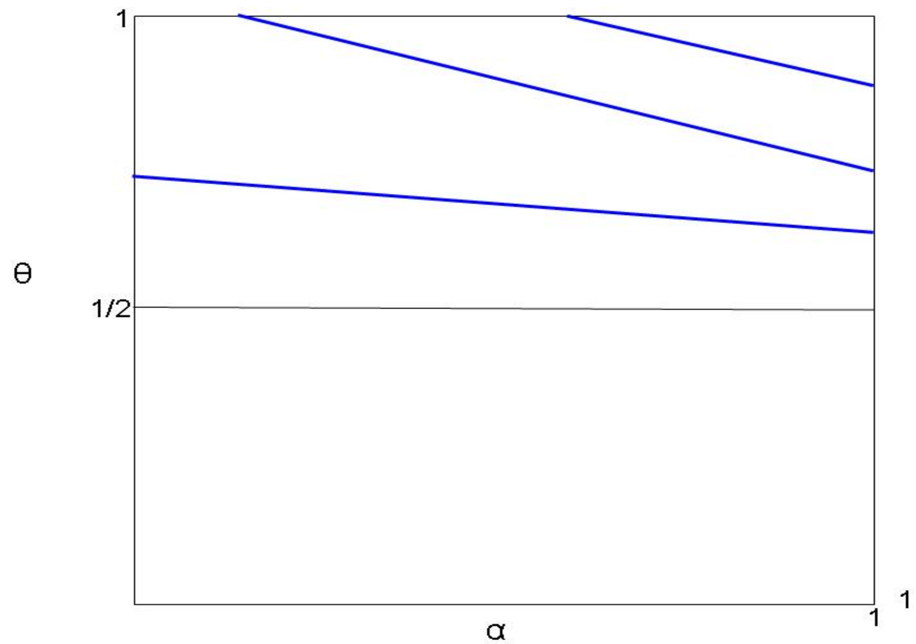
Importantly, the optimal mechanism in this case has qualitatively different forms depending upon whether $b \geq \frac{3}{2}$ or $b < \frac{3}{2}$.

We start with the case $b \geq \frac{3}{2}$, which was previously analyzed by Laffont, Maskin, Rochet (1987) under the assumption that $\gamma = 2$ (i.e. in the quadratic utility case) and for general values of gamma by Basov (2001), pp. 161-166. The main qualitative properties of the optimal mechanism in this case are as follows: (1) the lower boundary is flat and is given by $\theta = \frac{1}{2}$; (2) all positive isoquants emanate from the portion of the right-hand boundary $\alpha = 1$ above $\theta = \frac{1}{2}$. In particular, the isoquant associated with $q = 0$ is a flat line segment at $\theta = \frac{1}{2}$ given by $\{(\alpha, \frac{1}{2}) : \alpha \in [0, 1]\}$. So, the region associated with subproblem (i) in expression (13) is empty i.e., $\hat{q} = 0$. According to Theorem 9 the demand profile approach correctly identifies the optimal mechanism in this case.

Theorem 10 [9] *If the utility function is given by (34) with* $b \geq \frac{3}{2}$ *and* $F(.)$ *is uniform on a unit square, then the optimal screening mechanism is as follows:* $\alpha(q) = 1$ *for all* q*,* $q^* = \hat{q} = \left(\frac{4}{(2b+1)\gamma}\right)^{\frac{1}{\gamma - 1}}$*,*

[9] The proofs of Theorems 10 and 11 are available in the online Appendix at *http://www.severinov.com/mdimsupp*.

Figure 3: Isoquants in the case $b \geq \frac{3}{2}$.



$\theta^* = \frac{2b-1}{2b+1}$, $\bar{q} = \left(\frac{1}{b-1}\right)^{\frac{1}{\gamma-1}}$, *and the optimal quantity assignment on the right boundary is as follows:*

$$q = \begin{cases} \left(\frac{2\theta-1}{\gamma(\frac{2b-3}{4})}\right)^{\frac{1}{\gamma-1}}, & \text{for } \theta \in [\frac{1}{2}, \frac{2b-1}{2b+1}] \\ \left(\frac{3\theta-1}{\gamma(b-1)}\right)^{\frac{1}{\gamma-1}}, & \text{for } \theta \in [\frac{2b-1}{2b+1}, 1]. \end{cases}$$

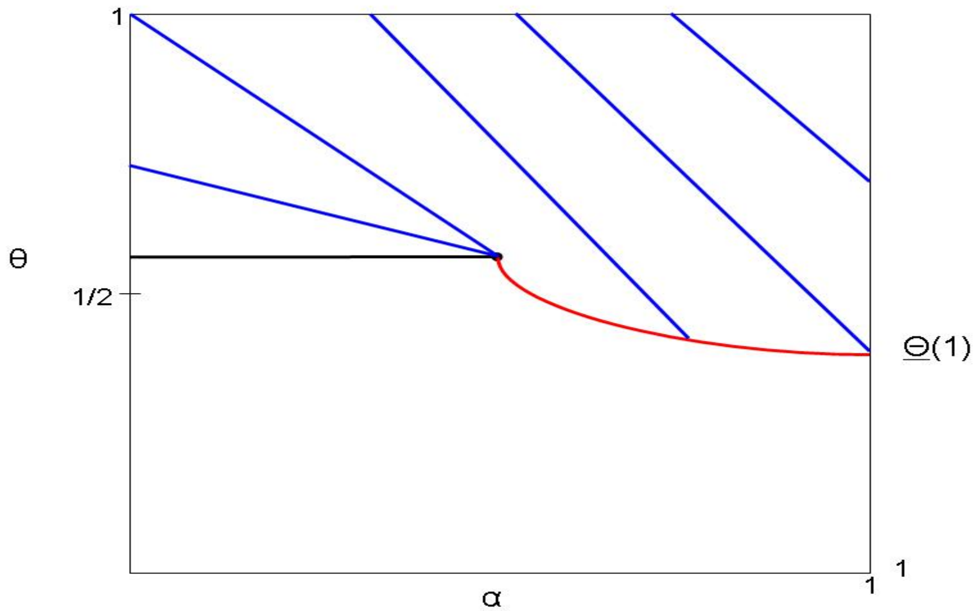
The corresponding optimal tariff is given by:

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

Figure 3 depicts the isoquants in this case for $\gamma = 2$ (quadratic utility function). Note that none of the iso-price/isoquant lines intersect each other in the type space. As a consequence, the demand profile approach properly identifies the optimal mechanism. Since the slope of the marginal utility varies from $b-1$ to b, large values of b are associated with low relative variability, $\frac{1}{b}$, in the slope of the marginal utility. 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 complex and interesting case where $b < 3/2$.

Figure 4: Isoquants in the case $b < \frac{3}{2}$.



Theorem 11 *If the utility function is given by (34) with $b < \frac{3}{2}$ and $F(.)$ is uniform on a unit square, then the optimal screening mechanism is as follows: $\bar{q} = \left(\frac{1}{b-1}\right)^{\frac{1}{\gamma-1}}$. The quantities q^* and $\hat{q}$ are uniquely defined by the following two equations:*

$$b(q^*)^\gamma(2 - b\gamma(q^*)^{\gamma-1}) = (b-1)\hat{q}^\gamma(2 - (b-1)\gamma\hat{q}^{\gamma-1})$$

$$\left(\frac{1}{2} - \frac{b}{3}\right)\left(\frac{b\gamma(q^*)^\gamma}{3}(2 - b\gamma(q^*)^{\gamma-1})\right)^{\gamma-1} = \int_{\frac{1+\gamma(b-1)\hat{q}^{\gamma-1}}{3}}^{1-\frac{b\gamma(q^*)^{\gamma-1}}{3}} ((1-\theta)(3\theta-1))^{\gamma-1}\, d\theta$$

All isoquants for quantities in the interval $[0, q^*]$ emanate from the point (α^*, θ^*), where $\alpha^* = \frac{2b}{3}$ and $\theta^* \equiv \theta(q^*) = 1 - \frac{b\gamma(q^*)^{\gamma-1}}{3}$, so that $\theta(q) = \theta^*$ and $\alpha(q) = \frac{2b}{3}$ for all $q \in [0, q^*]$.

For $q \in [q^*, \hat{q}]$, the optimal $\theta(q)$ and $\alpha(q)$ along the lower boundary are given by:

$$\theta(q) = \frac{2 - \sqrt{1 - \frac{b\gamma(q^*)^\gamma(2-b\gamma(q^*)^{\gamma-1})}{q}}}{3} \text{ for all } q \in [q^*, \hat{q}]$$

$$\alpha(q) = 2\left(\frac{b\gamma(q^*)^\gamma}{3}(2 - b\gamma(q^*)^{\gamma-1})\right)^{\gamma-1} \int\limits_{\frac{2-\sqrt{1-\frac{b\gamma(q^*)^\gamma(2-b\gamma(q^*)^{\gamma-1})}{q}}}{3}}^{1-\frac{b\gamma(q^*)^{\gamma-1}}{3}} ((1-\theta)(3\theta-1))^{\gamma-1}\, d\theta \tag{35}$$

For $q \in [\hat{q}, \bar{q}]$, the optimal $\theta(q)$ and $\alpha(q)$ along the right boundary are such that $\alpha(q) = 1$ and

$$\theta(q) = \frac{1 + \gamma(b-1)q^{\gamma-1}}{3}.$$

The associated optimal nonlinear tariff is given by:

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

With $b < \frac{3}{2}$, the isoquants for all $q \in [0, q^*]$ emanate from the point (α^*, θ^*) on the lower boundary. In particular, the isoquant for $q = 0$ is the flat segment at the level $\theta = \theta^*$ with $\alpha \in [0, \frac{2b}{3}]$, i.e. the collection of points $\{(\alpha, \theta^*) : \alpha \in [0, \frac{2b}{3}]\}$. For $q \in [q^*, \widehat{q}]$ the lower boundary is strictly decreasing, and there is a unique isoquant emanating from every point on the lower boundary. Since all types (α, θ^*) along the lower boundary with $\alpha \in [0, \frac{2b}{3}]$ 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. This happens because the optimal solution exhibits a clustering of quantities at the type (α^*, θ^*).

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 4 illustrates the lower boundary and the isoquants for the case $b < \frac{3}{2}$. Importantly, while the isoquants associated with the optimal mechanism never intersect in the interior of the participation region, they do intersect in the endogenously derived region of non-participation. The consequence of this is that, as we have shown in Section 2, for $\gamma = 2$ the optimal nonlinear price schedule $P(q)$ differs from the one identified by the demand profile approach. So, according to Theorem 9, the demand profile approach is incapable of correctly identifying the optimal mechanism whenever $b < 3/2$.

6 Conclusions

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 isoquant 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 allows 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 an application of our methodology with an example in which demand is a power function and types are uniformly distributed.

Our methodology has identified some relatively robust properties of optimal screening mechanism with multidimensional types. In particular, the allocation may be discontinuous in agent's type along the boundary of the participation region and exhibits a clustering of quantities at a particular type along the lower boundary. We expect that these findings should stimulate new research into several of the applications cited in the introduction.

While the present analysis deals with 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

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, L. and R.Deneckere (1993), "A Generalized Theorem of the Maximum" Economic Theory, 3, pp. 99-107.

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

Basov, S. (2005), "Multidimensional Screening," Springer-Verlag.

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

Bryson, A. and Y-C. Ho (1975), "Applied Optimal Control," Hemisphere Publishing Corporation, Washington D.C.

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

Ioffe, A., and T. Rockafellar (1996), "The Euler and Weierstrass Conditions for Nonsmooth Variational Problems," Calculus of Variations and Partial Differential Equations, 4, pp 59-87.

Jehiel, P., B. Moldovanu and E. Stacchetti (1996), "How (Not) to Sell Nuclear Weapons," American Economic Review, 86(4), 814-829.

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

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.

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

Merton, R (1969), 'Lifetime Portfolio Selection under Uncertainty: the Continuous-Time Case," *The Review of Economics and Statistics*, 51(3), pp. 247-257.

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 (2002), "Nonlinear Pricing with Random Participation," *Review of Economic Studies,* 69 (1), 277-311.

J.-C. Rochet and L. Stole (2003), "The Economics of Multidimensional Screening," in *Advances in Economic Theory, 7th World Congress,* Dewatripont, M. and L. Hansen, editors, Cambridge University Press.

Seierstad, A. and K. Sydsaeter (2002), "Optimal Control Theory with Economic Applications," North-Holland.

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

7 Appendix

In the proofs, we will make use of the following Lemma:

Lemma 7 *Suppose that Assumption 2 holds. Then for any* $q > 0$, $\frac{u_{q\alpha}(q,\alpha,\theta)}{u_{q\theta}(q,\alpha,\theta)} - \frac{u_{\alpha}(q,\alpha,\theta)}{u_{\theta}(q,\alpha,\theta)} > 0.$

Proof: Fix $(\alpha, \theta) \in [0, 1]^2$ and define $\varphi(q) = \frac{u_{q\alpha}}{u_{q\theta}}(q, \alpha, \theta) - \frac{u_{\alpha}}{u_{\theta}}(q, \alpha, \theta)$. Then, $\varphi'(q) = \frac{d}{dq}\left(\frac{u_{q\alpha}}{u_{q\theta}}\right) - \varphi(q)\frac{u_{q\theta}}{u_{\theta}}$. Assumption 2 implies that for any $q > 0$ s.t. $\varphi(q) \leq 0$ we have $\varphi'(q) > 0$. Thus, if $\varphi(q) \leq 0$ for some $q > 0$, then $\varphi(q') < \varphi(q)$ for all $q' < q$, and so $\lim_{q' \to 0} \varphi(q') < 0$. But since by Assumption 1(i) both u_{α} and u_{θ} converge to zero as $q \to 0$, it follows from l'Hospital's rule that $\lim_{q \to 0} \varphi(q) = 0$, a contradiction. Hence, $\varphi(q) > 0$ if $q > 0$. *Q.E.D.*

Proof of Lemma 1: Observe that, for fixed α, θ and $q > 0$, the equation $u_q(q, \alpha', \theta') = u_q(q, \alpha, \theta)$ implicitly defines θ' as a function of α': $\theta' = \widetilde{\theta}(\alpha'|q)$. We will omit the dependence of $\widetilde{\theta}(\alpha'|q)$ on the parameter q whenever the value of q is clear from the context. By the Implicit Function Theorem, this function is well-defined, with $u_{q\theta}(q, \alpha, \widetilde{\theta}(\alpha))\frac{d\widetilde{\theta}}{d\alpha} + u_{q\alpha}(q, \alpha, \widetilde{\theta}(\alpha)) = 0$ Hence $u_{qq}(q, \alpha', \theta') - u_{qq}(q, \alpha, \theta) = \int_{\alpha}^{\alpha'} [u_{qq\theta}(q, a, \widetilde{\theta}(a))\frac{d\widetilde{\theta}}{d\alpha} + u_{qq\alpha}(q, a, \widetilde{\theta}(a))]da = \int_{\alpha}^{\alpha'} [-u_{qq\theta}\frac{u_{q\alpha}}{u_{q\theta}} + u_{qq\alpha}]da > 0$, where the inequality follows from Assumption 2. This proves the desired result. *Q.E.D.*

Proof of Lemma 2: (i) To establish that $\underline{\theta}(\alpha)$ is monotonically decreasing recall that $u(.)$ is supermodular in (α, θ). So, if $q(\alpha, \theta) > 0$ for some (α, θ) and $\alpha' > \alpha$, then $q(\alpha', \theta) > 0$. Hence, $\underline{\theta}(\alpha') \leq \underline{\theta}(\alpha)$.

Next, to show that $\underline{\theta}(.)$ is continuous and is also strictly decreasing at α if $q(\alpha, \underline{\theta}(\alpha)) > 0$ and $\underline{\theta}(\alpha) > 0$, consider the "net payoff" function $s(\alpha, \theta) = u(q(\alpha, \theta), \alpha, \theta) - t(\alpha, \theta)$. Note that $s(\alpha, \theta)$ is continuous in (α, θ) because $u(q, \alpha, \theta)$ is continuous. In the optimal mechanism $s(\alpha, \theta) = 0$ if $q(\alpha, \theta) = 0$, for otherwise the firm can increase its profits by setting to zero the transfer paid by the types who get zero quantity. Hence, $s(\alpha, \underline{\theta}(\alpha)) = 0$ if $\underline{\theta}(\alpha) > 0$, for otherwise the mechanism would not be incentive compatible because the type $(\alpha, \underline{\theta}(\alpha) - \epsilon)$ would prefer to imitate the type $(\alpha, \underline{\theta}(\alpha))$, when $\epsilon > 0$ is sufficiently small .

Then continuity of $\underline{\theta}(.)$ follows immediately from the continuity of $s(\alpha, \theta)$ in (α, θ). Now consider $\alpha \in [0, 1)$ such that $\underline{\theta}(\alpha) > 0$ and $q(\alpha, \underline{\theta}(\alpha)) > 0$. Since $u(q, \alpha, \theta)$ is strictly increasing in α when $q > 0$, it follows that $s(\alpha', \underline{\theta}(\alpha)) > 0$ for any $\alpha' > \alpha$. Hence, $\underline{\theta}(\alpha') < \underline{\theta}(\alpha)$.

The proof that $\underline{\theta}(\alpha) > 0$ for all $\alpha \in [0, 1)$ is by contradiction, so suppose that $\underline{\theta}(\alpha) = 0$ for some $\alpha < 1$. Lemma 4(iii) below establishes that $q(.)$ is increasing in α along L, so we must have $q(\alpha, 0) > 0$ for all $\alpha \in (\hat{\alpha}, 1]$. Let $\hat{\alpha} = \min\{\alpha | \underline{\theta}(\alpha) = 0\}$. Then we have $\hat{\alpha} < 1$. Since $q(.)$ is increasing in α along L, we must have $q(\alpha, 0) > 0$ for all $\alpha \in (\hat{\alpha}, 1]$. But this cannot be optimal since $u(q, \alpha, 0) < 0$ so the transfer $t(\alpha, 0)$ must be nonpositive. But then the seller can strictly increase her profits by raising the tariff corresponding to all positive $q(\alpha, 0)$ to some positive level ϵ. Hence, it cannot be optimal to set $\underline{\theta}(\alpha) = 0$ for some $\alpha < 1$.

(ii) To establish equation (2), consider some $\alpha, \alpha' \in [0, 1]$ s.t. $\alpha > \alpha'$. Since $s(\alpha, \underline{\theta}(\alpha)) = s(\alpha', \underline{\theta}(\alpha')) = 0$ and the mechanism is incentive compatible, we have:

$$\begin{aligned} 0 &= u(q(\alpha, \underline{\theta}(\alpha)), \alpha, \underline{\theta}(\alpha)) - t(\alpha, \underline{\theta}(\alpha)) \geq u(q(\alpha', \underline{\theta}(\alpha')), \alpha, \underline{\theta}(\alpha)) - t(\alpha', \underline{\theta}(\alpha')) \\ 0 &= u(q(\alpha', \underline{\theta}(\alpha')), \alpha', \underline{\theta}(\alpha')) - t(\alpha', \underline{\theta}(\alpha')) \geq u(q(\alpha, \underline{\theta}(\alpha)), \alpha', \underline{\theta}(\alpha')) - t(\alpha, \underline{\theta}(\alpha)) \end{aligned}$$

By Lemma 4(iii) below, $q(\alpha, \underline{\theta}(\alpha)) \geq q(\alpha', \underline{\theta}(\alpha'))$, and hence $t(\alpha, \underline{\theta}(\alpha)) \geq t(\alpha', \underline{\theta}(\alpha'))$. Consequently, we have

$$u(q(\alpha, \underline{\theta}(\alpha)), \alpha, \underline{\theta}(\alpha)) - u(q(\alpha, \underline{\theta}(\alpha)), \alpha', \underline{\theta}(\alpha')) \geq 0 \geq u(q(\alpha', \underline{\theta}(\alpha')), \alpha, \underline{\theta}(\alpha)) - u(q(\alpha', \underline{\theta}(\alpha')), \alpha', \underline{\theta}(\alpha'))$$

Using the mean value theorem, we can rewrite the above as follows:

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

for some α_0 and α_1 s.t. $\alpha_0, \alpha_1 \in [\alpha', \alpha]$. The last two inequalities can be rewritten as follows:

$$-\frac{u_\alpha}{u_\theta}(q(\alpha, \underline{\theta}(\alpha)), \alpha_1, \underline{\theta}(\alpha_1)) \leq \frac{\underline{\theta}(\alpha) - \underline{\theta}(\alpha')}{\alpha - \alpha'} \leq -\frac{u_\alpha}{u_\theta}(q(\alpha', \underline{\theta}(\alpha')), \alpha_0, \underline{\theta}(\alpha_0)) \tag{36}$$

Since $\underline{\theta}(\alpha)$ is monotonically decreasing, it is differentiable almost everywhere. Taking the limits in (36), yields $\underline{\theta}'(\alpha) = -\frac{u_\alpha}{u_\theta}(q(\alpha, \underline{\theta}(\alpha)), \alpha, \underline{\theta}(\alpha))$ at any continuity point of $q(\alpha, \underline{\theta}(\alpha))$ i.e., (2) holds.

(iii) Let $B = \max_{(q,\alpha,\theta)} \frac{u_\alpha}{u_\theta}(q, \alpha, \theta) < \infty$. From (36) it follows that the function $\underline{\theta}(\alpha)$ is Lipschitz continuous with Lipschitz constant B, and hence it is absolutely continuous. *Q.E.D.*

Proof of Lemma 3: Let t_1 be the transfer associated with quantity q_1 in the mechanism (i.e., there is a type $(\tilde{\alpha}, \tilde{\theta})$ s.t. $q_1 = q(\tilde{\alpha}, \tilde{\theta})$, $t_1 = t(\tilde{\alpha}, \tilde{\theta})$). Since q_1 is an optimal quantity for type (α_1, θ_1), $u(q_1, \alpha_1, \theta_1) - t_1 \geq u(q(\alpha', \theta'), \alpha_1, \theta_1) - t(\alpha', \theta')$ for all (α', θ'). Rearranging, we have

$$t(\alpha', \theta') - t_1 \geq u(q(\alpha', \theta'), \alpha_1, \theta_1) - u(q_1, \alpha_1, \theta_1) \tag{37}$$

Next, let us show that $u(q,\alpha_1,\theta_1) - u(q,\alpha_2,\theta_2)$ has a unique global minimum at $q = q_1$. First, by assumption of the Lemma, $u_q(q_1,\alpha_1,\theta_1) - u_q(q_1,\alpha_2,\theta_2) = 0$. Further, for any $q'' \in (0,q_1)$, we have: $u_q(q'',\alpha_1,\theta_1) - u_q(q'',\alpha_2,\theta_2) = \int_{\alpha_2}^{\alpha_1} u_{q\alpha}(q'',a,\widetilde{\theta}(a|q_1)) + u_{q\theta}(q'',a,\widetilde{\theta}(a|q_1))\frac{d\widetilde{\theta}(a|q_1)}{d\alpha}da = \int_{\alpha_2}^{\alpha_1} u_{q\alpha}(q'',a,\widetilde{\theta}(a|q_1)) - u_{q\theta}(q'',a,\widetilde{\theta}(a|q_1))\frac{u_{q\alpha}(q_1,a,\widetilde{\theta}(a|q_1))}{u_{q\theta}(q_1,a,\widetilde{\theta}(a|q_1))}da =$
$\int_{\alpha_2}^{\alpha_1} u_{q\theta}(q'',a,\widetilde{\theta}(a|q_1))\left(\frac{u_{q\alpha}(q'',a,\widetilde{\theta}(a|q_1))}{u_{q\theta}(q'',a,\widetilde{\theta}(a|q_1))} - \frac{u_{q\alpha}(q_1,a,\widetilde{\theta}(a|q_1))}{u_{q\theta}(q_1,a,\widetilde{\theta}(a|q_1))}\right)da < 0$ where the last inequality follows from Assumption 2 (SCD) because $q'' < q_1$. Similarly, $u_q(q'',\alpha_1,\theta_1) - u_q(q'',\alpha_2,\theta_2) > 0$ if $q'' > q_1$.

So, $u(q'',\alpha_1,\theta_1) - u(q'',\alpha_2,\theta_2)$ is strictly decreasing at any $q'' \in (0,q_1)$ and is strictly increasing at $q'' > q_1$, and hence it reaches a unique global minimum at any $q = q_1$. Then combining $u(q_1,\alpha_1,\theta_1) - u(q_1,\alpha_2,\theta_2) < u(q(\alpha',\theta'),\alpha_1,\theta_1) - u(q(\alpha',\theta'),\alpha_2,\theta_2)$ with inequality (37), we obtain:

$$t(\alpha',\theta') - t_1 > u(q(\alpha',\theta'),\alpha_2,\theta_2) - u(q_1,\alpha_2,\theta_2)$$

Since $(q(\alpha',\theta'),t(\alpha',\theta'))$ was chosen arbitrarily, the pair (q_1,t_1) is the unique optimal choice for type (α_2,θ_2), and so $q(\alpha_2,\theta_2) = q_1$. *Q.E.D.*

Proof of Lemma 4:

(i) Rochet and Stole (2003) and Basov (2005, Theorem 191) have shown that the optimal allocation $q(\alpha,\theta)$ must satisfy an elliptical partial differential equation in the interior of the region where optimal quantities are strictly positive. It is well-known that solutions to elliptical partial differential equations on a domain with a piecewise smooth boundary are continuous. Since these conditions hold in our case, in the optimal mechanism $q(\alpha,\theta)$ is continuous at all (α,θ) s.t. $\theta > \underline{\theta}(\alpha)$.

(ii) By part (i), the set of positive quantities assigned in the mechanism, $\{q(\alpha,\theta) : (\alpha,\theta) \in \Omega_+\}$, is an interval. Without loss of generality, we may assume that this interval is closed. Hence if $q \in Q^*(\alpha,\theta)$ for some $(\alpha,\theta) \in L$, then q is available for type (α,θ) to select, i.e. there exists (α',θ') such that $q(\alpha',\theta') = q$. It remains to be shown that q is an optimal choice for type (α,θ). Suppose to the contrary that there existed (α'',θ'') such that $u(q(\alpha'',\theta''),\alpha,\theta) - t(\alpha'',\theta'') > u(q(\alpha',\theta'),\alpha,\theta) - t(\alpha',\theta')$. Let $\{\alpha_n,\theta_n,q_n\}$ be a sequence in the definition of $Q^*(\alpha,\theta)$ such that $(\alpha_n,\theta_n) \to (\alpha,\theta)$ and $q_n \to q$. Then for sufficiently large n we would have $u(q(\alpha'',\theta''),\alpha_n,\theta_n) - t(\alpha'',\theta'') > u(q(\alpha_n,\theta_n),\alpha_n,\theta_n) - t(\alpha_n,\theta_n)$, contradicting that $(q(\alpha_n,\theta_n),t(\alpha_n,\theta_n))$ is an optimal choice for type (α_n,θ_n).

To show that $Q^*(\alpha,\theta)$ is increasing along L, as we move from $(0,\underline{\theta}(0))$ to $(1,1)$, let us first show that $Q^*(\alpha,\underline{\theta}(\alpha))$ is increasing in α . Suppose to the contrary that there existed $\alpha_1,\alpha_2 \in [0,1]$, $\alpha_1 > \alpha_2$, with $q_i \in Q^*(\alpha_i,\underline{\theta}(\alpha_i))$ for $i \in \{1,2\}$ s.t. $q_1 < q_2$. Consider the isoquant $I(q_1,\alpha_1,\underline{\theta}(\alpha_1))$ and some $(\alpha',\theta') \in I(q_1,\alpha_1,\underline{\theta}(\alpha_1)) \setminus (\alpha_1,\underline{\theta}(\alpha_1))$ By Lemma 3, $q_1 = q(\alpha',\theta')$.

Without loss of generality the allocation $(0,0)$ is offered in the optimal mechanism. So the fact that 0 is not an optimal quantity choice for type (α',θ') implies that $s(\alpha',\theta') > 0$, and hence that $\theta' > \underline{\theta}(\alpha')$. Since (α',θ') is an arbitrary point in $I(q_1,\alpha_1,\underline{\theta}(\alpha_1)) \setminus (\alpha_1,\underline{\theta}(\alpha_1))$, it follows that $I(q_1,\alpha_1,\underline{\theta}(\alpha_1)) \cap \{(\alpha,\underline{\theta}(\alpha))|\alpha \in [0,1]\} = (\alpha_1,\underline{\theta}(\alpha_1))$. Combining the latter fact with Assumption 1 (iii) we conclude that there exists $\alpha'' \in (\alpha_2,\alpha_1)$ s.t. $(\alpha'',\underline{\theta}(\alpha_2)) \in I(q_1,\alpha_1,\underline{\theta}(\alpha_1))$ and thus that

$q_1 = q(\alpha'', \underline{\theta}(\alpha_2))$. But since $q_1 < q_2 = q(\alpha_2, \underline{\theta}(\alpha_2))$, this contradicts that the allocation $q(\cdot, \underline{\theta}(\alpha_2))$ is increasing.

A similar argument establishes that $Q^*(1, \theta)$ is increasing in θ. It follows that Q^* is increasing along L.

Because Q^* is monotonically increasing along L and bounded, it follows that except at countably many points, $Q^*(\alpha, \underline{\theta}(\alpha))$ is a singleton, i.e. $Q^*(\alpha, \underline{\theta}(\alpha)) = q(\alpha, \underline{\theta}(\alpha))$, and $q(\alpha, \underline{\theta}(\alpha))$ is continuous.

(iii) By part (i), the set of positive quantities assigned in the mechanism, $\{q(\alpha, \theta) : (\alpha, \theta) \in \Omega_+\}$, is an interval. Without loss of generality, we may assume that this interval is closed. Hence if $q \in Q^*(\alpha, \theta)$ for some $(\alpha, \theta) \in L$, then q is available for type (α, θ) to select, i.e. there exists (α', θ') such that $q(\alpha', \theta') = q$. It remains to be shown that q is an optimal choice for type (α, θ). Suppose to the contrary that there existed (α'', θ'') such that $u(q(\alpha'', \theta''), \alpha, \theta) - t(\alpha'', \theta'') > u(q(\alpha', \theta'), \alpha, \theta) - t(\alpha', \theta')$. Let $\{\alpha_n, \theta_n, q_n\}$ be a sequence in the definition of $Q^*(\alpha, \theta)$ such that $(\alpha_n, \theta_n) \to (\alpha, \theta)$ and $q_n \to q$. Then for sufficiently large n we would have $u(q(\alpha'', \theta''), \alpha_n, \theta_n) - t(\alpha'', \theta'') > u(q(\alpha_n, \theta_n), \alpha_n, \theta_n) - t(\alpha_n, \theta_n)$, contradicting that $(q(\alpha_n, \theta_n), t(\alpha_n, \theta_n))$ is an optimal choice for type (α_n, θ_n).

Next, let us establish upper hemi-continuity of Q^* along L. Let $\{\alpha_n, \theta_n, q_n\}$ be a sequence such that for each n we have $(\alpha_n, \theta_n) \in L$, $q_n \in Q^*(\alpha_n, \theta_n)$, and such that $(\alpha_n, \theta_n) \to (\alpha, \theta) \in L$ and $q_n \to q$. We need to show that $q \in Q^*(\alpha, \theta)$. For each n we may select $(\alpha'_n, \theta'_n) \in \Omega_+$ such that $((\alpha'_n, \theta'_n), q(\alpha'_n, \theta'_n))$ is within distance $\frac{1}{n}$ from $(\alpha_n, \theta_n, q_n)$. It follows that the sequence $\{(\alpha'_n, \theta'_n), q(\alpha'_n, \theta'_n)\}$ converges to (α, θ, q) and satisfies the requirements in the definition of $Q^*(\alpha, \theta)$. Thus $q \in Q^*(\alpha, \theta)$, and Q^* is u.h.c.

The proof that $Q^*(\alpha, \theta)$ is a closed set follows along similar lines. We next establish that $Q^*(\alpha, \theta)$ is convex-valued.

First we will show that $Q^*(0, \underline{\theta}(0))$ is a singleton, and hence a convex set. Let $\{\alpha_n, \theta_n\} \subset \Omega_+$ be a sequence s.t. $(\alpha_n, \theta_n) \to (0, \underline{\theta}(0))$. By part (iii) we may assume that $Q^*(\alpha_n, \underline{\theta}(\alpha_n))$ is single-valued. The monotonicity of $q(\cdot, \cdot)$ then implies that $q(\alpha_n, \theta_n) \geq Q^*(\alpha_n, \underline{\theta}(\alpha_n))$, and hence that $q = \lim_{n \to \infty} q(\alpha_n, \theta_n) \geq \lim_{n \to \infty} Q^*(\alpha_n, \underline{\theta}(\alpha_n))$. The u.h.c. and monotonicity of Q^* imply that $\lim_{n \to \infty} Q^*(\alpha_n, \underline{\theta}(\alpha_n))$ is the largest element of $Q^*(0, \underline{\theta}(0))$. Since q is an arbitrary element of $Q^*(0, \underline{\theta}(0))$, it follows that $q = \lim_{n \to \infty} Q^*(\alpha_n, \underline{\theta}(\alpha_n))$, and hence that $Q^*(0, \underline{\theta}(0))$ is single-valued. Using a similar argument, we may show that at any $(\alpha, \theta) \in L$ where $\theta = 1$ the correspondence $Q^*(\alpha, \theta)$ is single-valued.

Next, we show that for any $(\alpha, \theta) \in L$ s.t. either $\theta < 1$ or $\alpha > 0$, the set $Q^*(\alpha, \theta)$ is convex. Let $q_1, q_2 \in Q^*(\alpha, \theta)$ with $q_1 < q_2$. Fix some $q \in (q_1, q_2)$. Choose some $\varepsilon > 0$ sufficiently small that there exist $(\alpha_1, \theta_1) \in I(q_1, \alpha, \theta)$ and $(\alpha_2, \theta_2) \in I(q_2, \alpha, \theta)$ s.t. the distance between (α, θ) and both (α_1, θ_1) and (α_2, θ_2) equals ε. By Lemma 3, we have $q(\alpha_1, \theta_1) = q_1$ and $q(\alpha_2, \theta_2) = q_2$. Further, by continuity of $q(\cdot, \cdot)$, there exists (α_3, θ_3) on the line segment connecting (α_1, θ_1) and (α_2, θ_2) such that $q(\alpha_3, \theta_3) = q$. Since this is true for all $\varepsilon > 0$, it follows from the definition of the correspondence $Q^*(.)$ that $q \in Q^*(\alpha, \theta)$, so $Q^*(.)$ is convex-valued.

(iv) Let $\alpha_2 < \alpha_1$ so that $\theta_2 > \theta_1$. Suppose that $(\alpha_2, \theta_2) \in I(q_1, \alpha_1, \theta_1)$. By Lemma 3, $q(\alpha_2, \theta_2) = q_1$, and so $s(\alpha_2, \theta_2) = u(q_1, \alpha_2, \theta_2) - t(\alpha_2, \theta_2)$. By Lemma 2, $s(\alpha, \underline{\theta}(\alpha)) = 0$ for any sufficiently small α.

So, to establish that $(\alpha_2, \theta_2) \notin L$, it suffices to show that $s(\alpha_2, \theta_2) > 0$ which, in turn, would follow if $u(q_1, \alpha_2, \theta_2) - u(q_1, \alpha_1, \theta_1) > 0$. To establish the latter, note that for any $\alpha \in (\alpha_2, \alpha_1)$ there exists $\sigma(\alpha)$ such that $u_q(q_1, \alpha, \sigma(\alpha)) = u_q(q_1, \alpha_1, \theta_1)$, and hence $(\alpha, \sigma(\alpha)) \in I(q_1, \alpha_1, \theta_1))$. Differentiating, we obtain: $\sigma'(\alpha) = -\frac{u_{q\alpha}(q_1,\sigma(\alpha),\alpha)}{u_{q\theta}(q_1,\sigma(\alpha),\alpha)} < 0$. Then we have: $u(q_1, \alpha_2, \theta_2) - u(q_1, \alpha_1, \theta_1) =$

$$-\int_{\alpha_2}^{\alpha_1} [u_\theta(q_1, \sigma(\alpha), \alpha)\sigma'(\alpha) + u_\alpha(q_1, \sigma(\alpha), \alpha)]d\alpha = \int_{\alpha_2}^{\alpha_1} u_\theta(q_1, \sigma(\alpha), \alpha)\frac{u_{q\alpha}(q_1, \sigma(\alpha), \alpha)}{u_{q\theta}(q_1, \sigma(\alpha), \alpha)} - u_\alpha(q_1, \sigma(\alpha), \alpha)d\alpha > 0$$

where the inequality follows from Lemma 7 and the fact that $q_1 > 0$. *Q.E.D.*

Proof of Theorem 1:

Part (iii) of Lemma 4 implies 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. So the image of $w(.)$, $w(L)$, is a closed interval.

Now consider some $(\alpha, \theta) \notin L$ s.t. $q(\alpha, \theta) > 0$. Since $\theta > \underline{\theta}(\alpha)$ and $u_{q\theta} > 0$, we have $u_q(q(\alpha, \underline{\theta}(\alpha)), \alpha, \underline{\theta}(\alpha)) < u_q(q(\alpha, \underline{\theta}(\alpha)), \alpha, \theta)$. Also, $u_q(q(1, \theta), 1, \theta)) > u_q(q(1, \theta), \alpha, \theta)$ because $u_{q\alpha} > 0$. Therefore, since $w(L)$ is a closed interval, there exists $(\alpha', \theta') \in L$ s.t. $\alpha' \geq \alpha$, $\theta' < \theta$ and $u_q(q', \alpha, \theta) = u_q(q', \alpha', \theta')$ for some $q' \in Q^*(\alpha', \theta')$. That is, $(\alpha, \theta) \in I(q', \alpha', \theta')$. So, by Lemma 3, $q(\alpha, \theta) = q'$.

Next, let us show that (α, θ) cannot lie on more than one isoquant emanating from L. First, Lemma 3 implies that (α, θ) cannot belong to two isoquants $I(q', \alpha', \theta')$ and $I(q'', \alpha'', \theta'')$ such that $q' \neq q''$. So, it remains to rule out the following case: $(\alpha, \theta) \in I(q', \alpha', \theta') \cap I(q', \alpha'', \theta'')$, with $(\alpha', \theta') \in L$, $(\alpha'', \theta'') \in L$, and $q' \in Q^*(\alpha', \theta') \cap Q^*(\alpha'', \theta'')$. The proof is by contradiction, so suppose otherwise. Then, $u_q(q', \alpha, \theta) = u_q(q', \alpha', \theta') = u_q(q', \alpha'', \theta'')$.

However, Lemmas 2 and 7 imply that for any α s.t. $\underline{\theta}(\alpha) > 0$, we have:

$$\frac{du_q(q', \alpha, \underline{\theta}(\alpha))}{d\alpha} = \left(u_{q\alpha} - u_{q\theta}\frac{u_\alpha}{u_\theta}\right)(q', \alpha, \underline{\theta}(\alpha)) > 0 \tag{38}$$

Finally, if $\alpha = 1$, then $\frac{du_q(q',1,\theta)}{d\theta} = u_{q\theta}(q', 1, \theta) > 0$. So, for fixed $q = q'$, $u_q(q', .)$ increases along the lower boundary as we increase α and then up the right boundary as we raise θ. Hence, we cannot have $u_q(q', \alpha', \theta') \neq u_q(q', \alpha'', \theta'')$, which establishes that (α, θ) lies on a unique isoquant emanating from L.

Now let us derive the transfer $t(\alpha, \theta)$. If $q(\alpha, \theta) \in Q^*(\alpha', \theta')$ such that $\alpha' < 1$, then $\theta' = \underline{\theta}(\alpha') > 0$ and by Lemma 2 we have: $u(q(\alpha, \theta), \alpha', \theta') = t(\alpha, \theta)$.

Now, suppose that $q(\alpha, \theta) \in Q^*(1, \theta')$. To complete the proof of the Theorem we need to show that

$$t(\alpha, \theta) \equiv P(q(\alpha, \theta)) = u(\min Q^*(1, \underline{\theta}(1)), 1, \underline{\theta}(1)) + \int_{\min Q^*(1,\underline{\theta}(1))}^{q(\alpha,\theta)} u_q(q, 1, \theta^m(q))dq. \tag{39}$$

Since by Lemma 4 the correspondence $Q^*(1, \theta)$ is u.h.c., increasing, closed and convex-valued, for any $q \in [\min Q^*(1, \underline{\theta}(1)), \max Q^*(1, 1)]$, there exists θ such that $q \in Q^*(1, \theta)$. Accordingly, $\theta^m(q) = \max\{\theta | q \in Q^*(1, \theta)\}$ in (39) exists and is well-defined.

Consider any q that is not a discontinuity point of $\theta^m(\cdot)$. Note that since $\theta^m(\cdot)$ is increasing, this excludes at most countably many q. Now pick $q'' \neq q$. Since the mechanism $(q(.), t(.))$ is incentive compatible, we have:

$$u(q'', 1, \theta^m(q)) - u(q, 1, \theta^m(q)) \leq P(q'') - P(q) \leq u(q'', 1, \theta^m(q'')) - u(q, 1, \theta^m(q''))$$

Since $\theta^m(q'') \to \theta^m(q)$, the above inequality implies that $P(\cdot)$ is absolutely continuous, so that $P'(q) = u_q(q, 1, \theta^m(q))$, and (39) holds. *Q.E.D.*

Proof of Theorem 2: The proof proceeds in three steps. First, we establish that the mechanism $(q(.), t(.))$ is incentive compatible and individually rational along the boundary L. Second, we show that the mechanism is incentive compatible and individually rational for every type (α, θ) inside the participation region. Third, we will argue that any type in the non-participation region prefers the outside option $(q = 0, t = 0)$ to any other available pair (q, t) from the mechanism.

To begin with, observe that the boundary $\underline{\theta}(\alpha)$ in the mechanism $(q(.), t(.))$ is uniquely defined in Definition 1. Since $q(\alpha, \underline{\theta}(\alpha)) = q(\alpha)$, it is increasing and continuous in α almost everywhere in $[0, 1]$. Similarly, it follows that $q(\theta, 1)$ is increasing and continuous in θ almost everywhere on $[\hat{\theta}, 1]$.

Next, let us show that the mechanism $(q(.), t(.))$ is incentive compatible at every point (α, θ) on the boundary L. Then for every $q' \in [q_0, \bar{q}]$ we must have:

$$u(q(\alpha, \theta), \alpha, \theta) - t(\alpha, \theta) \equiv u(q(\alpha, \theta), \alpha, \theta) - P(q(\alpha, \theta)) \geq u(q', \alpha, \theta) - P(q'). \tag{40}$$

To establish (40), we need to consider several cases.

(a) $\alpha < 1$ and $q' \in [q_0, \hat{q}]$.

In this case, $\theta = \underline{\theta}(\alpha)$, $P(q(\alpha, \underline{\theta}(\alpha))) = u(q(\alpha, \underline{\theta}(\alpha)), \alpha, \underline{\theta}(\alpha)))$ and $P(q') = u(q', \alpha'(q'), \underline{\theta}(\alpha(q')))$. Therefore, (40) becomes: $u(q', \alpha, \underline{\theta}(\alpha)) - u(q', \alpha'(q'), \underline{\theta}(\alpha(q'))) \leq 0$. Let us show that this inequality holds for $q' > q(\alpha, \underline{\theta}(\alpha))$. The proof for $q' < q(\alpha, \underline{\theta}(\alpha))$ is similar. We have:

$$u(q', \alpha, \underline{\theta}(\alpha)) - u(q', \alpha', \underline{\theta}(\alpha')) = \int_{\alpha}^{\alpha'} u_\alpha(q', a, \underline{\theta}(a)) + u_\theta(q', a, \underline{\theta}(a))\frac{d\underline{\theta}}{da}da =$$

$$\int_{\alpha}^{\alpha'} u_\alpha(q', a, \underline{\theta}(a)) - u_\theta(q', a, \underline{\theta}(a)))\frac{u_\alpha(q(a, \underline{\theta}(a)), a, \underline{\theta}(a))}{u_\theta(q(a, \underline{\theta}(a)), a, \underline{\theta}(a))}da \leq 0$$

The last inequality follows from Lemma 7 and the fact that $q(a, \underline{\theta}(a)) \leq q'$ for all $a \leq a'$.

(b) $\alpha = 1$, $q' \in [\hat{q}, \bar{q}]$. Let us consider the case $q' > q(1, \theta)$. The case $q' < q(1, \theta)$ is similar and therefore omitted. Then (40) can be rewritten as follows:

$$\int_{q(1,\theta)}^{q'} u_q(q, 1, \theta)dq \leq \int_{q(1,\theta)}^{q'} u_q(q, 1, \theta(q))dq$$

This inequality holds because $\theta(.)$ is increasing and therefore $\theta(q) \geq \theta$ for all $q \geq q(1, \theta)$.

The remaining cases are:

(c) (α,θ) are such that $\alpha = 1$, $\theta \geq \hat{\theta}$, and $q' \in [0,\hat{q})$.

(d) (α,θ) are such that $\alpha < 1$, $\theta = \underline{\theta}(\alpha)$, and $q' \in (\hat{q},\bar{q}]$.

The proof for these cases follows from cases (a)-(b) and a simple monotonicity argument. So, since the proofs are similar, we will provide the proof for case (c) only. Then we have:

$$u(q(1,\theta),1,\theta) - P(q(1,\theta)) \geq u(q(1,\hat{\theta}),1,\theta) - P(q(1,\hat{\theta})) = u(q(1,\hat{\theta}),1,\theta) - u(q(1,\hat{\theta}),1,\hat{\theta})$$
$$+ u(q(1,\hat{\theta}),1,\hat{\theta}) - P(q(1,\hat{\theta})) > u(q',1,\theta) - u(q',1,\hat{\theta}) + u(q',1,\hat{\theta}) - P(q') = u(q',1,\theta) - P(q')$$

where the first inequality holds by case (b), the first equality is an identity, the second inequality holds by case (a) and because $u_{q\theta} \geq 0$ and $q(1,\hat{\theta}) > q'$, and the last equality is an identity. This completes the proof of incentive compatibility of our mechanism along L.

Next let us show the induced mechanism (q,t) is incentive compatible for any (α,θ) in the interior of Ω_+ i.e., s.t. $\alpha < 1$ and $\theta > \underline{\theta}(\alpha)$. Recall that the correspondences $q(\alpha)$ and $q(\theta)$ are upper-hemicontinuous, convex and closed. Therefore, the correspondence $v : L \to R$ defined by $v(\alpha,\theta) = \{u_q(q,\alpha,\theta) : q \in q(\alpha)\}$ if $\alpha < 1$ and $v(1,\theta) = \{u_q(q,1,\theta) : q \in q(\theta)\}$ for $\theta \in [\hat{\theta},1]$, is convex-valued and its image, $v(L)$, is a closed interval.

Further, $u_q(q(\alpha,\underline{\theta}(\alpha)),\alpha,\underline{\theta}(\alpha)) < u_q(q(\alpha,\underline{\theta}(\alpha)),\alpha,\theta)$ for $\theta > \underline{\theta}(\alpha)$ since $u_{q\theta} > 0$. Also, $u_q(q(1,\theta),1,\theta)) > u_q(q(1,\theta),\alpha,\theta)$ for $\alpha < 1$ because $u_{q\alpha} > 0$. Therefore, since $v(L)$ is a closed interval, there exists $(\alpha'',\theta'') \in L$ s.t. $\alpha'' \geq \alpha$, $\theta'' < \theta$ and $u_q(q'',\alpha,\theta) = u_q(q'',\alpha'',\theta'')$ where $q'' \in q(\alpha'')$ if $\alpha'' < 1$ and $q'' \in q(\theta'')$ if $\alpha'' = 1$. That is, $(\alpha,\theta) \in I(q'',\alpha'',\theta'')$. So, by Lemma 3, $(q'',P(q''))$ is the optimal choice for the type (α,θ). Thus, our mechanism is incentive compatible inside the participation region Ω_+.

Finally, the unique optimal choice of a type (α,θ) such that $\theta < \underline{\theta}(\alpha)$ is her outside option $(q = 0, t = 0)$ because for any $q > 0$, $u(q,\theta,\alpha) - P(q) < u(q,\underline{\theta}(\alpha),\alpha) - P(q) \leq 0$. *Q.E.D.*

Proof of Theorem 3: The proof of the Theorem proceeds through a number of steps.

Step 1. In this step we establish parts (ii) and (iii) of the Theorem.

By Pontryagin's Maximum Principle the costate equations for the Hamiltonian (21) are:

$$\dot{\mu} = -\frac{\partial J}{\partial \alpha} = -\frac{\partial (uh)}{\partial \alpha} - \frac{\partial(\mu - \lambda\frac{u_\alpha}{u_\theta})}{\partial \alpha}\alpha' = -\frac{\partial (uh_0)}{\partial \alpha} - \frac{\partial (u(h_2 - \frac{u_\alpha}{u_\theta}h_1)}{\partial \alpha}\alpha' + \lambda\frac{\partial\frac{u_\alpha}{u_\theta}}{\partial \alpha}\alpha' =$$
$$- u\frac{\partial\left(h_0 + (h_2 - \frac{u_\alpha}{u_\theta}h_1)\alpha'\right))}{\partial \alpha} - u_\alpha\left(h_0 + (h_2 - \frac{u_\alpha}{u_\theta}h_1)\alpha'\right) + \lambda\frac{\partial\frac{u_\alpha}{u_\theta}}{\partial \alpha}\alpha' = -u\frac{\partial h}{\partial \alpha} - u_\alpha h + \lambda\frac{\partial\frac{u_\alpha}{u_\theta}}{\partial \alpha}\alpha' \quad (41)$$

$$\dot{\lambda} = -\frac{\partial J}{\partial \theta} = -\frac{\partial (uh)}{\partial \theta} - \frac{\partial(\mu - \lambda\frac{u_\alpha}{u_\theta})}{\partial \theta}\alpha' = -\frac{\partial (uh_0)}{\partial \theta} - \frac{\partial (u(h_2 - \frac{u_\alpha}{u_\theta}h_1)}{\partial \theta}\alpha' + \lambda\frac{\partial\frac{u_\alpha}{u_\theta}}{\partial \theta}\alpha' =$$
$$- u\frac{\partial\left(h_0 + (h_2 - \frac{u_\alpha}{u_\theta}h_1)\alpha'\right))}{\partial \theta} - u_\theta\left(h_0 + (h_2 - \frac{u_\alpha}{u_\theta}h_1)\alpha'\right) + \lambda\frac{\partial\frac{u_\alpha}{u_\theta}}{\partial \theta}\alpha' = -u\frac{\partial h}{\partial \theta} - u_\theta h + \lambda\frac{\partial\frac{u_\alpha}{u_\theta}}{\partial \theta}\alpha' \quad (42)$$

Note that the last equality in both (41) and (42) holds because $h = h_0 + h_2 - \frac{u_\alpha}{u_\theta} h_1$. Then part (ii) of the Theorem i.e., equations (29) and (30), follow immediately from (41) and (42), respectively, when we set $\alpha' = 0$.

The continuity of the costate variables μ and λ i.e., part (iii) of the Theorem, also follows from Pontryagin's Maximum Principle.

Step 2. In this step we establish part (iv) of the Theorem i.e. $\alpha(q_0)q_0 = 0$.

It is enough to consider the case $q_0 > 0$. When $q_0 > 0$, the transversality condition associated with free left time q_0 is $J(q_0, \alpha(q_0), \theta(q_0), \alpha'(q_0), \mu(q_0), \lambda(q_0)) = 0$. The lineariry of the Hamiltonian J in α' implies that $S\alpha' = 0$ for all q. So by equation (21) we have: $J(q_0, \alpha(q_0), \theta(q_0), \alpha'(q_0), \mu(q_0), \lambda(q_0)) = u(q_0, \alpha(q_0), \theta(q_0))h_0(q_0, \alpha(q_0), \theta(q_0)) = 0$. Inspection of the definition of h_0 in equation (18) establishes that $uh_0 = 0$ is equivalent to either $q_0 = 0$ or $\alpha(q_0) = 0$.

The remaining Steps 3-6 of the proof establish part (i) of the Theorem.

Step 3. In this step we show that:

$$\dot{S} = \frac{\partial \frac{u_\alpha}{u_\theta}}{\partial q} \left(\left(\frac{u_q u_\theta}{u_{q\theta}} - u \right) h_1 - \lambda \right) \tag{43}$$

Differentiating the expression (22) for $S \equiv \frac{\partial J}{\partial \alpha'}$ yields:

$$\begin{aligned}
\dot{S} &= \frac{d}{dq}\frac{\partial J}{\partial \alpha'} = \frac{d\left(u(h_2 - \frac{u_\alpha}{u_\theta}h_1)\right)}{dq} + \dot{\mu} - \dot{\lambda}\frac{u_\alpha}{u_\theta} - \lambda\frac{d\left(\frac{u_\alpha}{u_\theta}\right)}{dq} = \\
&\frac{d\left(u(h_2 - \frac{u_\alpha}{u_\theta}h_1)\right)}{dq} - \frac{\partial(uh_0)}{\partial\alpha} - \frac{\partial(u(h_2 - \frac{u_\alpha}{u_\theta}h_1)}{\partial\alpha}\alpha' + \frac{u_\alpha}{u_\theta}\frac{\partial(uh_0)}{\partial\theta} + \frac{u_\alpha}{u_\theta}\frac{\partial(u(h_2 - \frac{u_\alpha}{u_\theta}h_1)}{\partial\theta}\alpha' - \lambda\frac{\partial\left(\frac{u_\alpha}{u_\theta}\right)}{\partial q} \\
&= \frac{\partial(u(h_2 - \frac{u_\alpha}{u_\theta}h_1))}{\partial q} - \frac{\partial(uh_0)}{\partial\alpha} + \frac{u_\alpha}{u_\theta}\frac{\partial(uh_0)}{\partial\theta} - \lambda\frac{\partial\left(\frac{u_\alpha}{u_\theta}\right)}{\partial q} = \frac{\partial(u(h_2 - \frac{u_\alpha}{u_\theta}h_1))}{\partial q} - u\left(\frac{\partial h_0}{\partial\alpha} - \frac{u_\alpha}{u_\theta}\frac{\partial h_0}{\partial\theta}\right) - \lambda\frac{\partial\left(\frac{u_\alpha}{u_\theta}\right)}{\partial q} \\
&= u_q(h_2 - \frac{u_\alpha}{u_\theta}h_1) - \frac{\partial\frac{u_\alpha}{u_\theta}}{\partial q}(uh_1 + \lambda) = \frac{\partial\frac{u_\alpha}{u_\theta}}{\partial q}\left(\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right)h_1 - \lambda\right)
\end{aligned} \tag{44}$$

where the third equality is obtained by substituting (41) and (42) and using $\frac{d\left(\frac{u_\alpha}{u_\theta}\right)}{dq} = \frac{\partial\left(\frac{u_\alpha}{u_\theta}\right)}{\partial q} + \frac{\partial\left(\frac{u_\alpha}{u_\theta}\right)}{\partial\alpha}\alpha' - \frac{\partial\left(\frac{u_\alpha}{u_\theta}\right)}{\partial\theta}\frac{u_\alpha}{u_\theta}\alpha'$ to cancel terms. The fourth equality holds because $\frac{d\left(u(h_2 - \frac{u_\alpha}{u_\theta}h_1)\right)}{dq} = \frac{\partial\left(u(h_2 - \frac{u_\alpha}{u_\theta}h_1)\right)}{\partial q} + \frac{\partial\left(u(h_2 - \frac{u_\alpha}{u_\theta}h_1)\right)}{\partial\alpha}\alpha' - \frac{u_\alpha}{u_\theta}\frac{\partial\left(u(h_2 - \frac{u_\alpha}{u_\theta}h_1)\right)}{\partial\theta}\alpha'$. The fifth equality holds because $\frac{\partial(uh_0)}{\partial\alpha} - \frac{u_\alpha}{u_\theta}\frac{\partial(uh_0)}{\partial\theta} = u\left(\frac{\partial h_0}{\partial\alpha} - \frac{u_\alpha}{u_\theta}\frac{\partial h_0}{\partial\theta}\right)$. The sixth equality follows from $\frac{\partial h_2}{\partial q} = \frac{\partial h_0}{\partial\alpha}$ and $\frac{\partial h_1}{\partial q} = \frac{\partial h_0}{\partial\theta}$. The seventh (last) equality in (44) holds because $u_q(h_2 - \frac{u_\alpha}{u_\theta}h_1) = \frac{\partial\left(\frac{u_\alpha}{u_\theta}\right)}{\partial q}\frac{u_q u_\theta}{u_{q\theta}}h_1$.

Step 4. In order to compute $\ddot{S}$, in this step we establish the following intermediate result:

$$\frac{d}{dq}h_1 - h_\theta = f(\theta, \alpha)\alpha' + \frac{\partial\left(\frac{u_\alpha}{u_\theta}\right)}{\partial\theta}h_1\alpha' \tag{45}$$

Taking a partial derivative of $h(q,\alpha,\theta,\alpha',\theta')$ in (17) yields:

$$h_\theta = \int_{\underline{\alpha}(q,\alpha,\theta)}^{\alpha} f_\theta \sigma_\theta \left(\sigma_q - \frac{u_\alpha}{u_\theta}\sigma_\theta \alpha' + \sigma_\alpha \alpha'\right) + f\left(\sigma_{q\theta} - \frac{u_\alpha}{u_\theta}\sigma_{\theta\theta}\alpha' + \sigma_{\alpha\theta}\alpha'\right) - f\frac{\partial\left(\frac{u_\alpha}{u_\theta}\right)}{\partial\theta}\sigma_\theta\alpha' da$$
$$- \underline{\alpha}_\theta f\left(\sigma_q - \frac{u_\alpha}{u_\theta}\sigma_\theta\alpha' + \sigma_\alpha\alpha'\right)|_{a=\underline{\alpha},\theta=\sigma(q,\alpha,\theta,\underline{\theta})} \tag{46}$$

On the other hand, fully differentiating (19) with respect to q we obtain:

$$\frac{dh_1}{dq} = \int_{\underline{\alpha}(q,\alpha,\theta)}^{\alpha} f_\theta\left(\sigma_q - \frac{u_\alpha}{u_\theta}\sigma_\theta\alpha' + \sigma_\alpha\alpha'\right)\sigma_\theta + f\left(\sigma_{q\theta} - \frac{u_\alpha}{u_\theta}\sigma_{\theta\theta}\alpha' + \sigma_{\alpha\theta}\alpha'\right) da$$
$$+ \alpha' f\sigma_\theta\,|_{a=\alpha} - \left(\underline{\alpha}_q - \underline{\alpha}_\theta\frac{u_\alpha}{u_\theta}\alpha' + \underline{\alpha}_\alpha\alpha'\right) f\sigma_\theta\,|_{a=\underline{\alpha},\theta=\sigma(q,\alpha,\theta,\underline{\theta})} \tag{47}$$

Combining (46) and (47) yields:

$$\frac{dh_1}{dq} - h_\theta = \alpha' f\sigma_\theta\,|_{a=\alpha} + f(\sigma,a)\left(\underline{\alpha}_\theta\sigma_q - \underline{\alpha}_q\sigma_\theta + \left(\underline{\alpha}_\theta\sigma_\alpha - \underline{\alpha}_\alpha\sigma_\theta\right)\alpha'\right)|_{a=\underline{\alpha},\theta=\sigma(q,\alpha,\theta,\underline{\theta})} + \frac{\partial\left(\frac{u_\alpha}{u_\theta}\right)}{\partial\theta}\int_{\underline{\alpha}(q,\alpha,\theta)}^{\alpha} f\sigma_\theta\alpha' da \tag{48}$$

Equation (9) implies that $\sigma_\theta\,|_{a=\alpha}= 1$. Comparing (48) and (45) reveals that (45) holds if and only if

$$\underline{\alpha}_\theta\sigma_q - \underline{\alpha}_q\sigma_\theta + \left(\underline{\alpha}_\theta\sigma_\alpha - \underline{\alpha}_\alpha\sigma_\theta\right)\alpha'_{a=\underline{\alpha}} = 0 \tag{49}$$

Recall that $\underline{\alpha} = 0$ if $\sigma(q,\alpha,\theta,0) < 1$, and otherwise $\underline{\alpha}$ solves the equation $\sigma(q,\theta,\alpha,a) = 1$ in a. So, (49) holds for all (q,α,θ) such that $\underline{\alpha} = 0$, because then $\underline{\alpha}_q(.) = \underline{\alpha}_\alpha(.) = \underline{\alpha}_\theta(.) = 0$.

Next, if $\underline{\alpha}(q,\alpha,\theta) > 0$, then its partial derivatives can be computed from the equation $\sigma(q,\theta,\alpha,\underline{\alpha}) = 1$. In particular, $\sigma_q(q,\theta,\alpha,\underline{\alpha})+\sigma_a(q,\theta,\alpha,\underline{\alpha})\underline{\alpha}_q = 0$, $\sigma_\theta(q,\theta,\alpha,\underline{\alpha})+\sigma_a(q,\theta,\alpha,\underline{\alpha})\underline{\alpha}_\theta = 0$, and $\sigma_\alpha(q,\theta,\alpha,\underline{\alpha})+\sigma_a(q,\theta,\alpha,\underline{\alpha})\underline{\alpha}_\alpha = 0$. The first two of these equations imply that $\underline{\alpha}_\theta\sigma_q - \underline{\alpha}_q\sigma_\theta = 0$ and the second and the third equations imply that $\underline{\alpha}_\theta\sigma_\alpha - \underline{\alpha}_\alpha\sigma_\theta = 0$. Thus, (49) also holds in this case.

Step 5. In this step, we compute $\ddot{S}$. Start by fully differentiating (43) with respect to q to obtain:

$$\ddot{S} = \frac{d\left(\frac{\partial\frac{u_\alpha}{u_\theta}}{\partial q}\right)}{dq}\left(\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right)h_1 - \lambda\right) + \frac{\partial\frac{u_\alpha}{u_\theta}}{\partial q}\frac{d\left(\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right)h_1 - \lambda\right)}{dq} \tag{50}$$

Next, let us consider the second term of (50), $\frac{d\left(\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right)h_1 - \lambda\right)}{dq}$. We have:

$$\frac{d\left(\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right)h_1\right)}{dq} = \frac{\partial\left(\frac{u_q u_\theta}{u_{q\theta}}h_1\right)}{\partial q} + \left(\frac{\partial}{\partial\alpha} - \frac{u_\alpha}{u_\theta}\frac{\partial}{\partial\theta}\right)\left(\frac{u_q u_\theta}{u_{q\theta}}h_1\right)\alpha' - u_q h_1 - u\frac{dh_1}{dq} =$$
$$\frac{\partial\left(\frac{u_q u_\theta}{u_{q\theta}}h_1\right)}{\partial q} + \left(\frac{\partial}{\partial\alpha} - \frac{u_\alpha}{u_\theta}\frac{\partial}{\partial\theta}\right)\left(\frac{u_q u_\theta}{u_{q\theta}}h_1\right)\alpha' - u_q h_1 - u\left(h_\theta + f(\theta,\alpha)\alpha' + \frac{\partial\left(\frac{u_\alpha}{u_\theta}\right)}{\partial\theta}h_1\alpha'\right) \tag{51}$$

The first equality in (51) is obtained by differentiating, substituting $-\frac{u_\alpha}{u_\theta}\alpha'$ for θ' and using $\frac{du}{dq} = u_q + u_\alpha\alpha' + u_\theta\theta' = u_q$. To obtain the second equality, we use equation (45) from Step 4.

Combining (51) with the expression for $\dot\lambda$ in (42) and rearranging terms yields:

$$\frac{d\left(\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right)h_1 - \lambda\right)}{dq} = \frac{\partial\left(\frac{u_q u_\theta}{u_{q\theta}}h_1\right)}{\partial q} - u_q h_1 + u_\theta h_0$$
$$+\left(\left(\frac{\partial}{\partial\alpha} - \frac{u_\alpha}{u_\theta}\frac{\partial}{\partial\theta}\right)\left(\frac{u_q u_\theta}{u_{q\theta}}h_1\right) - uf(\theta,\alpha) + u_\theta(h_2 - \frac{u_\alpha}{u_\theta}h_1) - (\lambda + uh_1)\frac{\partial\frac{u_\alpha}{u_\theta}}{\partial\theta}\right)\alpha' \tag{52}$$

Expanding the first three terms in (52) and comparing the result to the definition of $N(q,\alpha,\theta)$ in (23 we obtain:

$$\frac{\partial\left(\frac{u_q u_\theta}{u_{q\theta}}h_1\right)}{\partial q} - u_q h_1 + u_\theta h_0 = N(q,\alpha,\theta) \tag{53}$$

Relying on the above results, we can now compute $\ddot S$. On a non-singular arc we have $\alpha' = 0$. Substituting $\alpha' = 0$, (52) and (53) into (50 we obtain that on a non-singular arc:

$$\ddot S = \frac{d\left(\frac{\partial\frac{u_\alpha}{u_\theta}}{\partial q}\right)}{dq}\left(\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right)h_1 - \lambda\right) + \frac{\partial\frac{u_\alpha}{u_\theta}}{\partial q}N(q,\alpha,\theta) \tag{54}$$

On a singular arc we have $\dot S = 0$, which by (43) implies that $\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right)h_1 - \lambda = 0$. So the first term in (50) is equal to zero. Further, substituting $\lambda = \left(\frac{u_q u_\theta}{u_{q\theta}} - u\right)h_1$ into the multiplier of α' in (52), expanding all terms and comparing the result to the definition of $D(q,\alpha,\theta)$ in (24) we obtain that the multiplier of α' in (52) is equal to:

$$\left(\frac{\partial}{\partial\alpha} - \frac{u_\alpha}{u_\theta}\frac{\partial}{\partial\theta}\right)\left(\frac{u_q u_\theta}{u_{q\theta}}h_1\right) - uf(\theta,\alpha) + u_\theta h_2 - u_\alpha h_1 - \frac{u_q u_\theta}{u_{q\theta}}h_1\frac{\partial\frac{u_\alpha}{u_\theta}}{\partial\theta} = -D(q,\alpha,\theta) \tag{55}$$

Using (52), (53) and (55) in (50) then yields:

$$\ddot S = \frac{\partial\frac{u_\alpha}{u_\theta}}{\partial q}\left(N(q,\alpha,\theta) - D(q,\alpha,\theta)\alpha'\right) \tag{56}$$

Step 6. In this step we complete the proof of part (i) of the Theorem.

Note that, if $\alpha(q)$ is strictly increasing, then q must belong to a singular arc where $S = \dot S = \ddot S = 0$. Indeed, we know that $S \le 0$ for all q and setting $\alpha' > 0$ at q s.t. $S(q) < 0$ is strictly suboptimal. Then setting (56) to zero yields the differential equation (25). Equation (26) is identical to the differential equation for θ' in (14).

Further, setting (43) to zero yields (27). Finally, using (27) in $S = u\left(h_2 - \frac{u_\alpha}{u_\theta}h_1\right) + \mu - \lambda\frac{u_\alpha}{u_\theta} = 0$ to zero yields (28). Q.E.D.

Proof of Lemma 5: Let $J_{\alpha'}$ be the partial derivative of the Hamiltonian (21) with respect to α'. The Generalized Legendre Clebsch condition requires that if p is the smallest number such that $\frac{d^{2p}J_{\alpha'}}{dq^{2p}} \neq 0$ at some point on the optimal singular arc, then: $(-1)^p \frac{d^{2p}J_{\alpha'}}{dq^{2p}} \leq 0$. In our case, $p = 1$, and $\frac{d^{2p}J_{\alpha'}}{dq^{2p}} = -D$. So we must have $D \leq 0$.

Equation (25) and the fact that $\alpha'(q)$ is nondecreasing along a singular arc imply that *Q.E.D.*

Proof of Lemma 6: Suppose to the contrary that $q_0 > 0$. By Theorem 3 (iv), it follows that $\alpha(q_0) = 0$. Then for all $\varepsilon > 0$ there exists $q \in [q_0, q_0 + \varepsilon)$ such that $\alpha(q) > 0$. It follows that a right neighborhood of q_0 belongs to a singular arc where $D(q, \alpha, \theta) \leq 0$ by Lemma 5.

To establish a contradiction with Lemma 5, let us show that $D(q, \alpha, \theta) > 0$ for all sufficiently small α. To see this consider the definition of $D(q, \alpha, \theta)$ in equation (24). First, note that when α is sufficiently small $\underline{\alpha}(q, \alpha, \theta) = 0$ and so $\frac{\partial \underline{\alpha}(q,\alpha,\theta)}{\partial \alpha} = \frac{\partial \underline{\alpha}(q,\alpha,\theta)}{\partial \theta} = 0$. At the same time, $\left(u - \frac{u_\theta u_q}{u_{q\theta}}\right) > 0$ by the Assumption of the Lemma and $u_{q\theta}$ is bounded away from zero. So, when α is sufficiently small, i.e. for all $q \in (q_0, \varepsilon)$ for some $\varepsilon > 0$, the sign of $D(q, \alpha, \theta)$ is the same as the sign of $\left(u - \frac{u_\theta u_q}{u_{q\theta}}\right)$, and hence $D(q, \alpha, \theta) > 0$. *Q.E.D.*

Proof of Theorem 4:

The proof of the Theorem proceeds through four Lemmas and treats the intervals $[0, q^*)$ and $(q^*, \widehat{q}]$ separately. The reason for the latter is that our arguments rely on Assumption 3 and, in particular, on the continuity of N. While N can be continuous on each of the intervals $[0, q^*)$ and $(q^*, \widehat{q}]$, it cannot be globally continuous on $[0, \widehat{q}]$, because N exhibits a downward discontinuity at $q = q^*$.

This discontinuity arises because the partial derivatives of h_0, h_1 and h_2 with respect to q are discontinuous at q^*. Indeed, for all q s.t. $q < q^*$ the function $\underline{\alpha}(q, \alpha, \theta)$ is equal to zero, and hence does not change with q. But at q s.t. $q > q^*$, $\underline{\alpha}(q, \alpha, \theta)$ is strictly increasing in q. So, N exhibits a downward discontinuity at $q = q^*$.

The solution to subproblem (i) can have two types of junction points. The first one is associated with a transition from a nonsingular arc to a singular arc. Let us call this a type I junction point. The other one is associated with a transition from a singular arc to a nonsingular arc. We call this a type II junction point.

In the first two Lemmas we establish limits on the number of possible junction points and, in the end, show that there can be at most one junction point on the interval $[0, \hat{q}]$, and if such exists, then it must be a junction point of type I.

Lemma 8 *There exists $\varepsilon > 0$ such that $[0, \varepsilon]$ belongs to a nonsingular arc.*

Proof: Suppose to the contrary that $[0, \varepsilon]$ belongs to a singular arc for some $\varepsilon > 0$. We will show that in this case $\alpha(q) \to 0$ as $q \to 0$, contradicting Lemma 6. To establish this claim rewrite equation

(25) as follows:

$$\alpha' = \frac{1}{q}\kappa(q,\alpha,\theta) \text{ where } \kappa(q,\alpha,\theta) = \frac{(N/q)}{(D/q^2)}. \tag{57}$$

Next, let us show that in a neighborhood of $q = 0$, $|D/q^2|$ is bounded from above and $|N/q|$ is bounded from below.

For sufficiently small $q > 0$, $\underline{\alpha}(q,\alpha,\theta) = 0$. Using this in the definition of $D(.)$ in (24)yields:

$$D(q,\alpha,\theta) = \left(u - \frac{u_\theta u_q}{u_{q\theta}}\right) f(\alpha,\theta) + (u_{q\theta}u_\alpha - u_{q\alpha}u_\theta)\left\{2\int_0^\alpha \frac{f(a,\sigma)}{u_{q\theta}(q,a,\sigma)}da + u_q\int_0^\alpha \frac{f_\theta u_{q\theta} - f u_{q\theta\theta}}{u_{q\theta}^3}(q,a,\sigma)da\right\} \tag{58}$$

Observe that $u_{q\theta}$ is bounded away from zero on the compact set $[0, q^*(1,1)] \times [0,1]^2$. Since the numerators of both integrands in braces in (58) are continuous, they are bounded above on this same compact set. It follows that the multiplier of the term $(u_{q\theta}u_\alpha - u_{q\alpha}u_\theta)$ in (58) is bounded above.

Since $u(0,\alpha,\theta) = u_\alpha(0,\alpha,\theta) = u_\theta(0,\alpha,\theta) = 0$, a second order Taylor series expansion in q around $q = 0$ yields:

$$u - \frac{u_\theta u_q}{u_{q\theta}} = q^2\left(-u_{qq} - \frac{u_q u_{qq\theta}}{u_{q\theta}}\right)|_{(0,\alpha,\theta)} + o(q^2)$$
$$u_{q\theta}u_\alpha - u_{q\alpha}u_\theta = q^2(u_{qq\theta}u_{q\alpha} - u_{qq\alpha}u_{q\theta})|_{(0,\alpha,\theta)} + o(q^2)$$

Applying these Taylor series results in (58), and using the boundedness of the multiplier of the term $(u_{q\theta}u_\alpha - u_{q\alpha}u_\theta)$ in (58) establishes that $|D/q^2|$ is bounded from above in a neighborhood of $q = 0$.

Next, let us establish that the function $|N(q,\theta,\alpha)/q|$ is bounded from below in the neighborhood of $q = 0$. For this, it will suffice to show that $N_q(0,\alpha,\theta)$ is bounded away from zero since $\lim_{q\to 0}\frac{N}{q}(q,\alpha,\theta) = N_q(0,\alpha,\theta)$.

Since by assumption $[0,\varepsilon]$ belong to a singular arc, by Lemma 5 we have $N(q,\alpha(q),\theta(q)) \le 0$ for all $q \in (0,\varepsilon)$. We also have:

$$N(q,\alpha,\theta) = u_\theta\left(h_1\frac{\partial}{\partial q}\left(\frac{u_q}{u_{q\theta}}\right) + \frac{u_q}{u_{q\theta}}\frac{\partial}{\partial q}h_1 + h_0\right). \tag{59}$$

The term in brackets in (59) is equal to zero at $\alpha = 0$. Together with $u_\theta(0,\theta,\alpha) = 0$, this implies that $N_q(0,0,\theta(0)) = 0$. By Lemma 6, $\alpha(0) > 0$. Part (ii) of Assumption 3 ($N_{q\alpha} < 0$ at $q = 0$) then implies that $N_q(0,\alpha(0),\theta(0)) < 0$. So by continuity, $N_q(q,\alpha(q),\theta(q))$ is bounded away from zero in a neighborhood of $q = 0$.

Finally, since in a neighborhood of $q = 0$, $|D/q^2|$ remains bounded from above and $|N/q|$ is bounded from below, there exists $\eta > 0$ such that $\kappa(q,\alpha,\theta) \ge \eta$ over this neighborhood. Hence over this neighborhood we have $\alpha' \ge \frac{\eta}{q}$, implying that for fixed $q_1 > 0$ in this neighborhood and all $q \in (0,q_1)$,

$$\alpha(q) \le \alpha(q_1) - \eta\ln\left(\frac{q_1}{q}\right).$$

Thus $\alpha(q) = 0$ for some $q > 0$, thereby establishing the required contradiction. *Q.E.D.*

The next step in the proof of Theorem 4 shows that there is no type I junction point in $(0, q^*)$. So this interval belongs to a non-singular arc.

Lemma 9 *In the solution to subproblem (i), the interval* $[0, q^*)$ *belongs to a nonsingular arc.*

Proof of Lemma 9: We claim that $S(0) = \dot{S}(0) = 0$. Indeed, the transversality conditions at $q = 0$ are $\lambda(0) = \mu(0) = 0$. Also, $u(0, \alpha, \theta) = 0$ for all (α, θ). So we have $S(0) = u(h_2 - gh_1) + (\mu - \lambda g) = 0$. Now since $g_q > 0$, it follows from equation (43) that the sign of $\dot{S}$ is the same as the sign of $\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right) h_1 - \lambda$. Now, $\dot{S}(0) = 0$ follows from $\lambda(0) = 0$ and $u_\theta(0, \alpha, \theta) = 0$, where the latter holds because $u(0, \alpha, \theta) = 0$ for all (α, θ).

By Lemma 8, the solution is nonsingular in a neighborhood of $q = 0$, and so $S(q) < 0$ on $(0, \varepsilon]$ for some $\varepsilon > 0$. Suppose now that $\widetilde{q} \in (\varepsilon, q^*)$ is the smallest type I junction point so that $[0, \widetilde{q}]$ is a nonsingular arc. There are now two possibilities.

First suppose that there exists $q \in (0, \widetilde{q}]$ such that $N(q, \alpha(q), \theta(q)) > 0$. Assumption 3 then implies that $N(q', \alpha(q'), \theta(q')) > 0$ for all $q' \in (\widetilde{q}, q^*]$. But since the solution is singular in a right neighborhood of $\widetilde{q}$, it follows from equation (25) and Lemma 5 that over this neighborhood we must have $N(q, \alpha(q), \theta(q)) \leq 0$. Contradiction.

Hence suppose that we have $N(q, \alpha(q), \theta(q)) \leq 0$ for all $q \in (0, \widetilde{q})$. We claim then that over this interval we have $\dot{S} \leq 0$. Indeed, equation (43) and the fact that $g_q > 0$ imply that over the interval $(0, \widetilde{q})$ the sign of $\dot{S}$ is the same as the sign of $\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right) h_1 - \lambda$. By equation (52) on this interval we have:

$$\frac{d}{dq}\left(\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right) h_1 - \lambda\right) = \frac{\partial \frac{u_q u_\theta}{u_{q\theta}}}{\partial q} + u_\theta h_0 - u_q h_1 = N \leq 0$$

Since at $q = 0$ we have $\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right) h_1 - \lambda = 0$, it follows that $\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right) h_1 - \lambda \leq 0$ and hence $\dot{S} \leq 0$ on $[0, \widetilde{q})$. Because $S(\varepsilon) < 0$, we have $S(\widetilde{q}) < 0$, contradicting that $\widetilde{q}$ is a type I junction point.

We conclude that there exists no type I junction point on the interval $[0, q^*]$, and hence that this interval belongs to a nonsingular arc. *Q.E.D.*

Lemma 10 *The solution to subproblem (i) is such that there is no type II junction point on the interval* $(q^*, \widehat{q})$.

Proof of Lemma 10:

The proof of the Lemma involves two steps. First, let us establish that if $q^\# \in (q^*, \widehat{q})$ is a type II junction point, then there is no type I junction point in $(q^\#, \widehat{q})$.

Suppose to the contrary that q' is the smallest type I junction point in $(q^\#, \widehat{q})$. Since some left neighborhood of $q^\#$ and some right neighborhood of q' belong to singular arcs, equation (25) and

Lemma 5 taken together imply that $N(q^{\#}, \alpha(q^{\#}), \theta(q^{\#})) \leq 0$ and $N(q', \alpha(q'), \theta(q')) \leq 0$. Furthermore, since the interval $(q^{\#}, q')$ is a singular arc, we have $\alpha(q^{\#}) = \alpha(q')$ and $\theta(q^{\#})) = \theta(q')$. We now claim that $N(q, \alpha(q^{\#}), \theta(q^{\#})) \leq 0$ for all $q \in (q^{\#}, q')$.

Indeed, suppose to the contrary that there existed some $q \in (q^{\#}, q')$ such that $N(q, \alpha(q^{\#}), \theta(q^{\#})) > 0$. Assumption 3 then implies that $N(q'', \alpha(q^{\#}), \theta(q^{\#})) > 0$ for all $q'' \in (q^{\#}, q')$. But this contradicts that we must have $N(q', \alpha(q'), \theta(q')) \leq 0$, thereby establishing the claim.

We shall now prove that over the interval $(q^{\#}, q')$ we have $\dot{S} \leq 0$. Indeed, equation (43) and the fact that $g_q > 0$ imply that over the interval $(q^{\#}, q')$ the sign of $\dot{S}$ is the same as the sign of $\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right) h_1 - \lambda$. Because the solution is singular in a left neighborhood of $q^{\#}$, Theorem 3 (equation (27)) implies $(\frac{u_q u_\theta}{u_{q\theta}} - u)h_1 - \lambda = 0$ at $q = q^{\#}$. Furthermore, as shown in the proof of Theorem 3 (equation (52)) for all $q \in (q^{\#}, q')$ we have:

$$\frac{\partial}{\partial q}\left(\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right) h_1 - \lambda\right) = \frac{\partial \frac{u_q u_\theta}{u_{q\theta}}}{\partial q} + u_\theta h_0 - u_q h_1 = N \leq 0$$

Hence $\dot{S} \leq 0$ on $(q^{\#}, q')$.

Now since the solution is nonsingular on $(q^{\#}, q')$, there exists $\varepsilon \in (0, q' - q^{\#})$ such that $S(q^{\#} + \varepsilon) < 0$. Because $\dot{S} < 0$ on $(q^{\#}, q')$, it follows that $S(q') \leq S(q^{\#} + \varepsilon) < 0$, contradicting that q' is a type I junction point.

Thus, if $q^{\#} \in (q^*, \widehat{q})$ is a type II junction point, then the interval $(q^{\#}, \widehat{q})$ must lie in a non-singular arc, and so $\alpha(q^{\#}) = \alpha(\widehat{q})$. But Theorem 6 then implies that $q^{\#} = \widehat{q}$, so there is no type II junction point and hence the solution must be singular on the whole interval $(q^*, \widehat{q})$. This completes the proof of the Lemma. *Q.E.D.*

Lemma 11 *The solution to subproblem (i) is such that there is no type I junction point on the interval* $(q^*, \widehat{q})$.

Proof of Lemma 11: By Lemma 10, if there exists such a junction point $q^{\#}$, then because the solution is singular on $(q^{\#}, \widehat{q})$, we must have $N \leq 0$ at $q = \widehat{q}$. Assumption 3 then implies that $N \leq 0$ for all $q \leq q^{\#}$. As shown in the proof of Lemma 10, this implies that over this interval we have $\dot{S} \leq 0$. It follows from $S(\varepsilon) < 0$ for some $\varepsilon > 0$ that we have $S(q^{\#}) < 0$, contradicting that $q^{\#}$ is a type I junction point. We conclude that we must have $q^{\#} = q^*$. This concludes the proof of the Lemma and of the Theorem. *Q.E.D.*

Proof of Theorem 5:

Recall that the Lagrangian for subproblem (ii) in (31) is given by:

$$\max \int_{\widehat{q}}^{\bar{q}} u_q(q, 1, \theta) H(q, 1, \theta) + \delta\theta' dq$$

subject to $\theta(\hat{q}) \geq \hat{\theta}$, $\hat{q} \leq \bar{q}$, $\theta(\bar{q}) = 1$.

The Euler condition for this maximization problem is:

$$\frac{\partial\left(u_q(q,1,\theta)H(q,1,\theta)+\delta\theta'\right)}{\partial\theta} - \frac{d\frac{\partial(u_q(q,1,\theta)H(q,1,\theta)+\delta\theta')}{\partial\theta'}}{dq} = \phi(q,\theta) - \delta' = 0 \tag{60}$$

The complementary slackness condition is $\delta(q)\theta'(q) = 0$ and the boundary condition is $\delta(q) \geq 0$. From $\delta(q)\theta'(q) = 0$ it follows that $\delta(q) = 0$, on any interval $[q_1, q_2] \subseteq [\underline{q}, \bar{q}]$ where $\theta'(q) > 0$, and hence $\delta'(q) = 0$ on this interval. Using the latter in (60) yields the first-order condition on $[q_1, q_2]$:

$$\phi(q,\theta) = u_q(q,1,\theta)H_\theta(q,1,\theta) + u_{q\theta}(q,1,\theta)H(q,1,\theta) = 0 \tag{61}$$

Let $\theta^\phi(q)$ be the solution to (61), when such exists. Also, let $\theta^\phi(q) = 1$ if $\phi(q,\theta) > 0$ for all $\theta \in [0,1]$ and let $\theta^\phi(q) = 0$ if $\phi(q,\theta) < 0$ for all $\theta \in [0,1]$.

Let us first consider the case when the conditions $\phi_q > 0$ and $\phi_\theta < 0$ hold. Then $\theta^\phi(q)$ is increasing in θ. If $\theta^\phi(\widehat{q}) \geq \widehat{\theta}$, then the boundary condition $\theta(\widehat{q}) \geq \widehat{\theta}$ is non-binding and it is optimal to set $\theta(q) = \theta^\phi(q)$ for all $q \in [\widehat{q}, \bar{q})$. On the other hand, if $\theta^\phi(\widehat{q}) < \widehat{\theta}$, then we have to set $\theta(q) = \widehat{\theta}$ for q s.t. $\theta^\phi(q) \leq \widehat{\theta}$ and $\theta(q) = \theta^\phi(q)$ for q s.t. $\theta^\phi(q) > \widehat{\theta}$. Also, the boundary condition requires that $\theta(\bar{q}) = 1$. To summarize, the optimal solution in this case is $\theta(q) = \max\{\theta^\phi(q), \widehat{\theta}\}$ for $q \in [\hat{q}, \bar{q})$ and $\theta(\bar{q}) = 1$.

If the conditions $\phi_q > 0$ and $\phi_\theta < 0$ do not hold, then $\theta^\phi(q)$ may not be increasing for all $q \in [\widehat{q}, \bar{q}]$. In this case, the Euler condition (60) implies that the optimal θ has to satisfy conditions (a) and (b) in part (3) of the Theorem.

Finally, the transversality condition (see Seierstad and Sydsaeter (2002), p.32-33) for the free 'terminal time' $\bar{q}$ is $\frac{\partial(u_q(\bar{q},1,1)H(\bar{q},1,1)+\delta(\bar{q})\theta'(\bar{q}))}{\partial\theta'} = \delta(\bar{q}) \leq 0$. Combining this inequality with the boundary condition $\delta(\bar{q}) \geq 0$ yields $\delta(\bar{q}) = 0$. Since $\delta(q) \geq 0$ for all $q \in [\underline{q}, \bar{q}]$, it follows that $\delta'(\bar{q}) \leq 0$. Hence, by (60), we have $\phi(\bar{q}, \theta(\bar{q})) \leq 0$.

If $\phi(\bar{q}, \theta(\bar{q})) = 0$ we are done. Indeed, $\phi(\bar{q}, \theta(\bar{q})) \equiv u_{q\theta}(\bar{q},1,1)H(\bar{q},1,1) + u_q(\bar{q},1,1)H_\theta(\bar{q},1,1) = 0$. Since $H(\bar{q},1,1) = 0$ and $H_\theta(\bar{q},1,1) < 0$, the latter is equivalent to $u_q(\bar{q},1,1) = 0$.

Now, let us rule out $\phi(\bar{q}, \theta(\bar{q})) < 0$, which is equivalent to $\bar{q} > q^*(1,1)$. In this case, we can modify the tentative solution $\theta(q)$ as follows. First, let the new terminal "time" be $q^*(1,1)$. Second, replace the assignment $\theta(q)$ with the assignment $\tilde{\theta}(q)$ defined as follows: $\tilde{\theta}(q) = \theta(q)$ for all q, $q < q^*(1,1)$, and $\tilde{\theta}(q^*(1,1)) = 1$. Then $\tilde{\theta}(q)$ is increasing. Also, the value of the objective under $\tilde{\theta}(.)$ is greater than under $\theta(.)$ since the integrand of the objective (15) is negative for all $q \in [q^*(1,1), \bar{q}]$. This rules out $\phi(\bar{q}, q) < 0$. *Q.E.D.*

Proof of Theorem 6:

Suppose that, contrary to the statement of the Theorem, $\widehat{\theta} > \theta^\phi(\widehat{q})$. We will show that marginally lowering $\widehat{\theta}$ raises the value function of the problem, $V(\widehat{q}, 1, \widehat{\theta})$, which, according to (13), is given by:

$$V(\widehat{q},1,\widehat{\theta}) = W(\widehat{q},1,\widehat{\theta}) + u(\widehat{q},1,\widehat{\theta})H(\widehat{q},1,\widehat{\theta}) + Z(\widehat{q},1,\widehat{\theta})$$

The function $W(\widehat{q}, 1, \widehat{\theta})$ is given by (16). By (Seierstad and Sydsaeter 2002), (p. 213) we have: $\frac{\partial W}{\partial \widehat{\theta}} = -\lambda(\widehat{q})$.

Also, recall that $Z(\widehat{q}, 1, \widehat{\theta}) = \int_{\widehat{q}}^{q^{\phi}(\widehat{\theta})} H(q, 1, \widehat{\theta}) u_q(q, 1, \widehat{\theta}) dq + \int_{q^{\phi}(\widehat{\theta})}^{\bar{q}(1)} H(q, 1, \theta(q)) u_q(q, 1, \theta(q)) dq$, where $q^{\phi}(\widehat{\theta})$ is the unique solution to the equation $\theta^{\phi}(q) = \widehat{\theta}$. Differentiation of $Z(.)$ yields $\frac{\partial Z}{\partial \widehat{\theta}} = \int_{\widehat{q}}^{q^{\phi}(\widehat{\theta})} \phi(q, \widehat{\theta}) dq$, where $\phi(q, \widehat{\theta}) = u_q(q, 1, \widehat{\theta}) H_\theta(q, 1, \widehat{\theta}) + u_{\theta q}(q, 1, \widehat{\theta}) H(q, 1, \widehat{\theta})$.

Combining the above, we obtain:

$$\begin{aligned} \frac{\partial V}{\partial \widehat{\theta}}(\widehat{q}, 1, \widehat{\theta}) &= -\lambda(\widehat{q}) + u_\theta(\widehat{q}, 1, \widehat{\theta}) H(\widehat{q}, 1, \widehat{\theta}) + u(\widehat{q}, 1, \widehat{\theta}) H_\theta(\widehat{q}, 1, \widehat{\theta}) + \int_{\widehat{q}}^{q^{\phi}(\widehat{\theta})} \phi(q, \widehat{\theta}) dq \\ &= -\frac{u_q u_\theta h_1}{u_{q\theta}}(\widehat{q}, 1, \widehat{\theta}) + u_\theta(\widehat{q}, 1, \widehat{\theta}) H(\widehat{q}, 1, \widehat{\theta}) + \int_{\widehat{q}}^{q^{\phi}(\widehat{\theta})} \phi(q, \widehat{\theta}) dq \leq \int_{\widehat{q}}^{q^{\phi}(\widehat{\theta})} \phi(q, \widehat{\theta}) dq < 0 \end{aligned}$$

The second equality follows because $\lambda = \left(\frac{u_q u_\theta}{u_{q\theta}} - u \right) h_1$, and because by definition, we have

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

The penultimate inequality follows because $\phi(q, \theta)$ is decreasing in θ which yields:

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

which also implies that $u_\theta(\widehat{q}, 1, \widehat{\theta}) H(\widehat{q}, 1, \widehat{\theta}) \leq -\frac{u_q u_\theta}{u_{q\theta}}(\widehat{q}, 1, \widehat{\theta})) H_\theta(\widehat{q}, 1, \widehat{\theta})$.

It follows that lowering $\hat{\theta}$ increases the value of $V(\widehat{q}, 1, \widehat{\theta})$, so we cannot have $\widehat{\theta} > \theta^{\phi}(\widehat{q})$ in an optimal solution to (16).

Next, suppose that contrary to the statement of the theorem, $\widehat{\theta} < \theta^{\phi}(\widehat{q})$. Then since by Theorem 5 any type $(1, \theta)$ with $\theta \in [\widehat{\theta}, \theta^{\phi}(\widehat{q})]$ is assigned the quantity $\widehat{q}$, and pays the transfer associated with this quantity, we have:

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

Since $V(\widehat{q}, 1, \widehat{\theta}) = W(\widehat{q}, 1, \widehat{\theta}) + u(\widehat{q}, 1, \widehat{\theta}) H(\widehat{q}, 1, \widehat{\theta}) + Z(\widehat{q}, 1, \widehat{\theta})$, by differentiating we obtain:

$$\begin{aligned} \frac{\partial V}{\partial \widehat{\theta}}(\widehat{q}, 1, \widehat{\theta}) &= -\lambda(\widehat{q}) + u_\theta(\widehat{q}, 1, \widehat{\theta}) H(\widehat{q}, 1, \widehat{\theta}) + u(\widehat{q}, 1, \widehat{\theta}) H_\theta(\widehat{q}, 1, \widehat{\theta}) \\ &= -\frac{u_q u_\theta h_1}{u_{q\theta}}(\widehat{q}, 1, \widehat{\theta}) + u_\theta(\widehat{q}, 1, \widehat{\theta}) H(\widehat{q}, 1, \widehat{\theta}) > 0 \end{aligned}$$

The second equality follows because $\lambda = \left(\frac{u_q u_\theta}{u_{q\theta}} - u \right) h_1$ and $H_\theta(\widehat{q}, 1, \widehat{\theta}) = -h_1(\widehat{q}, 1, \widehat{\theta})$. The final inequality follows because $\widehat{\theta} > \theta^{\phi}(\widehat{q})$ implies $\phi(\widehat{q}, \widehat{\theta}) > \phi(\widehat{q}, \theta^{\phi}(\widehat{q})) = 0$, and so $H(\widehat{q}, 1, \widehat{\theta}) > \frac{u_q h_1}{u_{\theta q}}(\widehat{q}, 1, \widehat{\theta})$.

Thus, raising $\widehat{\theta}$ marginally is profitable , and so we cannot have $\widehat{\theta} < \theta^{\phi}(\widehat{q})$ in the optimal mechanism. Hence, we must have $\widehat{\theta} = \theta^{\phi}(\widehat{q})$. *Q.E.D.*

Proof of Theorem 7: Since q^* is a type I junction point, we have $S(q^*) = 0$. Furthermore, it follows from the definition of q^* that we have $\sigma(q^*, \alpha^*, \theta^*, 1) = 0$. According to Theorem 3, the Lagrange multiplier λ must be continuous at the junction point q^*, On the one hand, the fact that $\lambda(0) = 0$, and equation (30) then imply, and $\lambda(q^*) = \int_{q^*}^{\widehat{q}} \left(u\frac{\partial h_0}{\partial \theta} - u_\theta h_0 \right) dq$. On the other hand, because q^* is a Junction point of type 1, equation (27) then implies $\lambda(q^*) = \left(\frac{u_q u_\theta}{u_{q\theta}} - u \right) h_1$, so (33) holds.

It remains to be argued that μ is also continuous at q^*. Observe that according to equation (22) we have:

$$0 = S(q^*) - S(0) = u\left(h_2 - \frac{u_\alpha}{u_\theta}h_1\right) + \left(\mu - \lambda\frac{u_\alpha}{u_\theta}\right)$$

The continuity of S and λ then imply the continuity of μ, as indicated by equation (29).

Secondly, since $\alpha(\widehat{q}) = 1$, it follows that $\int_{q^*}^{\widehat{q}} \alpha'(q)dq = 1 - \alpha^*$, where $\alpha'(q)$ is given by (25). Finally, Theorem 6 implies that $\widehat{\theta} = \theta^\phi(q)$. *Q.E.D.*

Proof of Theorem 8: If $q^* \geq \widehat{q}$, then the region associated with subproblem (i) is empty, and so $\alpha(q) = 1$ for all q. By Theorem 6, we must have $\widehat{\theta} = \theta^\phi(q)$ for all q. Furthermore, Lemma 6 implies that $\widehat{q} = 0$. *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 $\theta > 0$. It follows that the isoquant $\sigma(q, \alpha(q), \theta(q), a)$ through the point $(\alpha(q), \theta(q))$ at the level q contains points (those with coordinates $a \in (\alpha, 1]$) which violate the individual rationality condition. Types $(\sigma(q, \alpha(q), \theta(q), a), a)$ with $a \in (\alpha, 1]$ will therefore not consume 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 $\theta^\phi(0) > 0$. An analogous argument treats the case where $\theta^\phi(q) = 0$ for some $q > 0$. It follows from Theorem 6 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

$$u_q(q, 1, \widehat{\theta})H_\theta(q, \widehat{\theta}) + u_{q\theta}(q, 1, \widehat{\theta})H(q, \widehat{\theta}) > 0. \tag{62}$$

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

$$N(p, q) + p\frac{\partial N}{\partial p}(p, q) = 0, \tag{63}$$

or equivalently that

$$u_{q\theta}(q, 1, \theta)H(q, \theta) + u_q(q, 1, \theta)H_\theta(q, \theta) = 0 \text{ at } \theta = \widetilde{\theta}(p, q). \tag{64}$$

It follows from (62), (64) 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.

Now 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 (63) 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 (64) holds, so $p = u_q(q, 1, \theta^{\phi}(q))$ is a global optimizer of (63), and so the demand profile approach identifies the optimal mechanism. *Q.E.D.*

Online Appendix (Not for Publication)
Proofs of Theorems 10 and 11.

In this appendix, we prove Theorems 10 and 11.

We start by considering the solution to Subproblem (i) in (15) for an arbitrary fixed $\widehat{q} > 0$.

First, observe that $u - \frac{u_\theta u_q}{u_{q\theta}} = \frac{(b-\alpha)}{2}(\gamma-1)q^\gamma > 0$ for all $q > 0$. It then follows from Lemma 6(i) that $q_0 = 0$.

Recall that (α^*, θ^*) is a point on the lower boundary such that an isoquant $I(q^*, \alpha^*, \theta^*)$ emanating from it hits the corner $(\alpha = 0, \theta = 1)$.

Lemma 12 *Suppose that* $u(q, \theta, \alpha) = \theta q - \frac{b-\alpha}{2} q^\gamma$, *and* $F(\theta, \alpha)$ *is uniform on* $[0, 1]^2$. *Then in the optimal solution* $\alpha^* = \frac{2b}{3}$ *and the interval* $[0, q^*]$ *forms a non-singular arc where* $\alpha' = \theta' = 0$, *while the interval* $[q^*, \widehat{q}]$ *forms a singular arc where* $\alpha' > 0$.

Proof of Lemma 12:

Step 1. Preliminary Computations.

Note the following simple results:

$$\sigma(q, \alpha, \theta, a) = \theta + \frac{\gamma}{2}(\alpha - a)q^{\gamma-1} \tag{65}$$

$$\sigma_\theta(q, \alpha, \theta, a) = 1 \tag{66}$$

$$\sigma_\alpha(q, \alpha, \theta, a) = \frac{\gamma}{2} q^{\gamma-1} \tag{67}$$

$$\sigma_q(q, \alpha, \theta, a) = \frac{\gamma(\gamma-1)}{2} q^{\gamma-2} \tag{68}$$

$$-\frac{d\theta}{d\alpha} = g \equiv \frac{u_\alpha}{u_\theta} = \frac{q^{\gamma-1}}{2} \tag{69}$$

First, consider $q \in [0, q^*]$. Then $\underline{\alpha} = 0$, and so we have:

$$h_0(q) = \frac{\alpha^2 \gamma(\gamma-1)}{4} q^{\gamma-2} \tag{70}$$

$$h_1(q) = \alpha \tag{71}$$

$$h_2(q) = \frac{\alpha\gamma}{2} q^{\gamma-1} \tag{72}$$

Using (70)-(72), we obtain:

$$D(q, \alpha, \theta) = \frac{\gamma-1}{2}\left(b - 3\alpha\right) q^\gamma \tag{73}$$

$$N(q, \alpha, \theta) = \frac{\alpha\gamma(\gamma-1)\left(\frac{3\alpha}{2} - b\right)}{2} q^{\gamma-1} \tag{74}$$

Combining (73) and (74) yields:

$$\alpha' = \frac{N}{D} = \frac{\gamma\alpha\left(\frac{3\alpha}{2} - b\right)}{q(b - 3\alpha)} \tag{75}$$

Next, consider $q \in [q^*, \widehat{q}]$. In this case, $\underline{\alpha} = \alpha - \frac{2(1-\theta)}{\gamma q^{\gamma-1}}$, and so we have:

$$h_0(q) = \frac{(1-\theta)^2(\gamma-1)}{\gamma} q^{-\gamma} \tag{76}$$

$$h_1(q) = \frac{2(1-\theta)}{\gamma q^{\gamma-1}} \tag{77}$$

$$h_2(q) = (1-\theta) \tag{78}$$

Using (76)-(78), we obtain that for $q \in [q^*, \widehat{q}]$:

$$D(q, \alpha, \theta) = (3\theta - 2)\left(1 - \frac{1}{\gamma}\right) q \tag{79}$$

$$N(q, \alpha, \theta) = (1-\theta)(1-3\theta)\frac{\gamma-1}{\gamma} q^{1-\gamma} \tag{80}$$

Combining (79) and (80), we obtain for $q \in [q^*, \widehat{q}]$:

$$\alpha'(q) = \frac{N}{D} = \frac{(1-\theta)(1-3\theta)}{q^{\gamma}(3\theta-2)} \tag{81}$$

Step 2. Let us show that $\alpha'(q) = 0$ on some neighborhood $[0, \epsilon)$, with $\epsilon > 0$.

For, suppose not and we have $\alpha'(q) > 0$ on some neighborhood of $q = 0$. By Step 1, $\alpha'(q)$ is given by equation (75). So, we must have $\alpha \in [\frac{b}{3}, \frac{2b}{3}]$. Applying a separation of variables in equation (75) yields:

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

Integrating both sides yields:

$$q = \frac{c}{\alpha(2b - 3\alpha)^{\frac{1}{\gamma}}} \tag{82}$$

where c is some positive constant. Note that the expression under the root in the denominator of (82), $\alpha(2b-3\alpha)$ is increasing in α on $[\frac{b}{3}, \frac{2b}{3}]$. Therefore, for all $\alpha \in [\frac{b}{3}, \frac{2b}{3}]$, $q \geq \frac{c_2}{\left(\frac{b}{3}\right)^{\frac{1}{\gamma}}}$, where $c_2 = 3^{(1-\frac{1}{\gamma})}bc > 0$.

So, we must have $\alpha'(q) = 0$ for all $q \in \left[0, \frac{c_2}{\sqrt[1/\gamma]{\frac{b}{3}}}\right)$.

Step 3. The interval $[0, q^*]$ belongs to a non-singular arc with $\alpha'(q) = 0$ on this interval.

By step 2, there exists $q_s \in (0, q^*]$ such that $[0, q_s]$ is a non-singular arc. We will show that $q_s = q^*$. The proof is by contradiction. So, suppose not i.e., $q_s < q^*$. Then $\alpha'(q) > 0$ on $(q_s, q_s + \epsilon)$

for some ϵ such that $q^* - q_s > \epsilon > 0$, so $\alpha'(q)$ is given by equation (75) on this interval. Therefore, $\frac{b}{3} \le \alpha(q_s) < \frac{2b}{3}$, for otherwise by (75) we would have $\alpha'(q) < 0$ for some $q \in (q_s, q_s + \epsilon)$.

It remains to show that we cannot have $\alpha(q_s) \in [\frac{b}{3}, \frac{2b}{3})$. To this end, we will establish that $S(q_s) < 0$ given that $[0, q_s]$ is a non-singular arc, $q_s \le q^*$, and $\alpha(q) \in [\frac{b}{3}, \frac{2b}{3})$. This would imply that such q_s cannot be a junction point since this will contradict the continuity of $S(.)$ at the junction point between a non-singular and singular arcs.

By definition, $S = u(h_2 - gh_1) + (\mu - \lambda g)$. Using (69) and (71)-(72), we obtain:

$$u(h_2 - gh_1) = (\theta q - \frac{b-\alpha}{2}q^\gamma)(\frac{\alpha\gamma}{2}q^{\gamma-1} - \frac{q^{\gamma-1}}{2}\alpha) = (\theta q - \frac{b-\alpha}{2}q^\gamma)\frac{\alpha(\gamma-1)q^{\gamma-1}}{2} \quad (83)$$

Also, by Theorem 3 on the interval $[q_{b/3}, \min\{q_{2b/3}, q^*\})$ we have: $\dot{\mu} = -u_\alpha h_0 - u\frac{\partial h_0}{\partial \alpha}$, $\dot{\lambda} = -u_\theta h_0 - u\frac{\partial h_0}{\partial \theta}$. Since $\mu(0) = 0$ and $\lambda(0) = 0$, we have:

$$\mu(q) = \int_0^q \mu'(x)dx = \int_0^q -u_\alpha h_0 - u\frac{\partial h_0}{\partial \alpha}dx = \int_0^q -\frac{x^{2(\gamma-1)}\gamma(\gamma-1)\alpha^2}{8} - \frac{\alpha\gamma(\gamma-1)x^{\gamma-2}}{2}(\theta x - \frac{b-\alpha}{2}x^\gamma)dx$$
$$= -\frac{\alpha\theta(\gamma-1)q^\gamma}{2} + \frac{\alpha\gamma(\gamma-1)(b - \frac{3}{2}\alpha)q^{2\gamma-1}}{4(2\gamma-1)} \quad (84)$$
$$\lambda(q) = \int_0^q \lambda'(x)dx = \int_0^q -u_\theta h_0 - u\frac{\partial h_0}{\partial \theta}dx = \int_0^q -\frac{\alpha^2\gamma(\gamma-1)}{4}x^{\gamma-1}dx = -\frac{\alpha^2(\gamma-1)q^\gamma}{4} \quad (85)$$

Combining (83), (84) and (85) yields:

$$S = u(h_2 - gh_1) + \mu - \lambda g =$$
$$(\theta q - \frac{b-\alpha}{2}q^\gamma)\frac{\alpha(\gamma-1)q^{\gamma-1}}{2} - \frac{\alpha\theta(\gamma-1)q^\gamma}{2} + \frac{\alpha\gamma(\gamma-1)(b - \frac{3}{2}\alpha(q))q^{2\gamma-1}}{4(2\gamma-1)} + \frac{q^{\gamma-1}}{2}\frac{\alpha^2(\gamma-1)q^\gamma}{4} =$$
$$-\frac{b-\alpha}{2}q^\gamma\frac{\alpha(\gamma-1)q^{\gamma-1}}{2} + \frac{\alpha\gamma(\gamma-1)(b - \frac{3}{2}\alpha)q^{2\gamma-1}}{4(2\gamma-1)} + \frac{q^{\gamma-1}}{2}\frac{\alpha^2(\gamma-1)q^\gamma}{4} =$$
$$\left(-1 + \frac{\gamma}{2\gamma-1}\right)\frac{\alpha(\gamma-1)(b - \frac{3}{2}\alpha(q))q^{2\gamma-1}}{4} = \frac{\alpha(\gamma-1)^2(\frac{3}{2}\alpha - b)q^{2\gamma-1}}{4(2\gamma-1)} \quad (86)$$

Note that (86) is negative if $\alpha < \frac{2b}{3}$ and is positive if $\alpha > \frac{2b}{3}$ So, we cannot have $\alpha(q) < \frac{2b}{3}$ for $q < q^*$.

Step 4. The interval $[0, q^*]$ belongs to a non-singular arc with $\alpha'(q) = 0$ for all $q \in [0, q^*]$. Also, $\alpha^* \ge \frac{2b}{3}$.

(i) First, suppose that there is a $q \in [0, q^*]$ s.t. $\alpha'(q) > 0$ and $\alpha(q) < \frac{b}{3}$. Then the denominator of equation (75) is positive i.e. $D(q) > 0$ which contradicts Lemma 5.

(ii) Next, suppose that there is a $q \in [0, q^*]$ s.t. $\alpha'(q) > 0$ and $\alpha(q) > \frac{2b}{3}$. Then by (75), we have $\alpha'(q) < 0$ which contradicts the fact that $\alpha'(q) \ge 0$ for all $q \in [0, q^*]$.

(iii) Finally, suppose that there is a $q \in [0, q^*]$ s.t. $\alpha'(q) > 0$ and $\alpha(q) \in [\frac{b}{3}, \frac{2b}{3})$. In this case, we can show using the previous arguments that $S(q) < 0$ ruling out this possibility.

Step 5. Let $[0, q_n)$ be the non-singular arc (so that $\alpha'(q) = 0$ for all $q \in [0, q_n)$). Then $q_n = q^*$ and $\alpha^* = \frac{2b}{3}$, $\theta^* = 1 - \alpha^* \frac{\gamma}{2} (q^*)^{\gamma-1}$.

Proof: The previous steps have established that $q_n \geq q^*$. So we only need to rule out $q_n > q^*$. To this end, let us compute the value of $S(q_n)$. By definition, $S(q_n) = u(h_2 - gh_1) + \mu - g\lambda$. Since $\theta'(q) = \alpha'(q)$ for $q \in [0, q_n]$, for brevity throughout this step we use θ to denote $\theta(q)$ and α to denote $\alpha(q)$, $q \in [0, q_n]$.

By Theorem 3, on the non-singular arc $[0, q_n)$, $\dot{\mu} = -u_\alpha h_0 - u\frac{\partial h_0}{\partial \alpha}$ $\dot{\lambda} = -u_\theta h_0 - u\frac{\partial h_0}{\partial \theta}$, and also $\mu(0) = 0$ and $\lambda(0) = 0$. Therefore, we have:

$$\mu(q_n) = \int_0^{q_n} \mu'(q)dq = \int_0^{q_n} -u_\alpha h_0 - u\frac{\partial h_0}{\partial \alpha}dq = \int_0^{q^*} -u_\alpha h_0 - u\frac{\partial h_0}{\partial \alpha}dq + \int_{q^*}^{q_n} -u_\alpha h_0 - u\frac{\partial h_0}{\partial \alpha}dq =$$
$$\int_0^{q^*} -\frac{x^{2(\gamma-1)}\gamma(\gamma-1)\alpha^2}{8} - \frac{\alpha\gamma(\gamma-1)x^{\gamma-2}}{2}(\theta x - \frac{b-\alpha}{2}x^\gamma)dx + \int_{q^*}^{q_n} -\frac{x^\gamma}{2}\frac{(1-\theta)^2(\gamma-1)}{\gamma}x^{-\gamma}dx =$$
$$-\frac{\alpha\theta(\gamma-1)(q^*)^\gamma}{2} + \frac{\alpha\gamma(\gamma-1)(b-\frac{3}{2}\alpha)(q^*)^{2\gamma-1}}{4(2\gamma-1)} - \frac{(1-\theta)^2(\gamma-1)}{2\gamma}(q_n - q^*). \tag{87}$$

$$\lambda(q_n) = \int_0^{q_n} \lambda'(q)dq = \int_0^{q_n} -u_\theta h_0 - u\frac{\partial h_0}{\partial \theta}dq = \int_0^{q^*} -u_\theta h_0 - u\frac{\partial h_0}{\partial \theta}dq + \int_{q^*}^{q_n} -u_\theta h_0 - u\frac{\partial h_0}{\partial \theta}dq =$$
$$\int_0^{q} -\frac{\alpha^2\gamma(\gamma-1)}{4}x^{\gamma-1}dx + \int_{q^*}^{q_n} -\frac{(1-\theta)^2(\gamma-1)}{\gamma}x^{1-\gamma} - (\theta x - \frac{b-\alpha}{2}x^\gamma)\frac{-2(1-\theta)(\gamma-1)}{\gamma}x^{-\gamma}dx$$
$$= -\frac{\alpha^2(\gamma-1)(q^*)^\gamma}{4} - \frac{(b-\alpha)(1-\theta)(\gamma-1)}{\gamma}(q_n - q^*) + \frac{(3\theta-1)(1-\theta)(\gamma-1)}{\gamma}\int_{q_n}^{q^*} q^{1-\gamma}dq \tag{88}$$

Using (69), (71)-(72), (77)-(78), (87) and (88), we can now compute $S(q_n)$ for $q \geq q^*$:

$$
\begin{aligned}
S(q_n) &= u(h_2 - gh_1) + \mu - g\lambda = (\theta q_n - \frac{b-\alpha}{2} q_n^{\gamma})(1-\theta)\frac{\gamma-1}{\gamma} \\
&- \frac{(1-\theta)\theta(\gamma-1)q^*}{\gamma} + \frac{(1-\theta)^2(\gamma-1)(b-\frac{3}{2}\alpha)q^*}{\alpha\gamma(2\gamma-1)} - \frac{(1-\theta)^2(\gamma-1)}{2\gamma}(q_n - q^*) \\
&+ \frac{\alpha^2(\gamma-1)(q^*)^{\gamma}}{8} q_n^{\gamma-1} + \frac{(b-\alpha)(1-\theta)(\gamma-1)}{2\gamma}(q_n - q^*) q_n^{\gamma-1} - \frac{(3\theta-1)(1-\theta)(\gamma-1)}{2\gamma}\int_{q^*}^{q_n}\left(\frac{q_n}{q}\right)^{\gamma-1} dq = \\
&\frac{(3\theta-1)(1-\theta)(\gamma-1)}{2\gamma}(q_n - q^*) - \frac{b-\alpha}{2} q_n^{\gamma}(1-\theta)\frac{\gamma-1}{\gamma} + \frac{(1-\theta)^2(\gamma-1)(b-\frac{3}{2}\alpha)q^*}{\alpha\gamma(2\gamma-1)} \\
&+ \frac{\alpha^2(\gamma-1)(q^*)^{\gamma}}{8} q_n^{\gamma-1} + \frac{(b-\alpha)(1-\theta)(\gamma-1)}{2\gamma}(q_n - q^*) q_n^{\gamma-1} - \frac{(3\theta-1)(1-\theta)(\gamma-1)}{2\gamma}\int_{q^*}^{q_n}\left(\frac{q_n}{q}\right)^{\gamma-1} dq = \\
&\frac{(3\theta-1)(1-\theta)(\gamma-1)}{2\gamma}(q_n - q^*) - \frac{(b-\alpha)(1-\theta)(\gamma-1)}{2\gamma} q^* q_n^{\gamma-1} + \frac{(1-\theta)^2(\gamma-1)(b-\frac{3}{2}\alpha)q^*}{\alpha\gamma(2\gamma-1)} \\
&+ \frac{\alpha^2(\gamma-1)(q^*)^{\gamma}}{8} q_n^{\gamma-1} - \frac{(3\theta-1)(1-\theta)(\gamma-1)}{2\gamma}\int_{q^*}^{q_n}\left(\frac{q_n}{q}\right)^{\gamma-1} dq = \\
&- \frac{(3\theta-1)(1-\theta)(\gamma-1)}{2\gamma}\int_{q^*}^{q_n}\left(\left(\frac{q_n}{q}\right)^{\gamma-1} - 1\right) dq - \frac{(b-\frac{3}{2}\alpha)(1-\theta)(\gamma-1)}{2\gamma} q^* q_n^{\gamma-1} + \frac{(1-\theta)^2(\gamma-1)(b-\frac{3}{2}\alpha)q^*}{\alpha\gamma(2\gamma-1)} \\
&- \frac{(3\theta-1)(1-\theta)(\gamma-1)}{2\gamma}\int_{q^*}^{q_n}\left(\left(\frac{q_n}{q}\right)^{\gamma-1} - 1\right) dq - \frac{(b-\frac{3}{2}\alpha)(1-\theta)^2(\gamma-1)}{\alpha\gamma} q^*\left(\frac{1}{\gamma} - \frac{1}{2\gamma-1}\right) \quad (89)
\end{aligned}
$$

Note that to obtain the last equality we use $\theta = 1 - \alpha\frac{\gamma}{2}(q^*)^{\gamma-1}$.

If $\alpha^* > \frac{2b}{3}$, then using $q_n \geq q^*$ in the last equality of (89) yields $S(q_n) < 0$. But this contradicts the fact that $S(q_n) = 0$ since q_n is a junction point between a non-singular and singular arcs and hence we must have $S(q_n) = 0$. So, $\alpha^* = \frac{2b}{3}$. Then, we also must have $q_n = q^*$, because otherwise $S(q_n) < 0$.

It remains to check that the solution with $\alpha^* = \frac{2b}{3}$ and $q_n = q^*$ is consistent with the continuity of the Lagrange multipliers, particularly at the junction point q_n.

Since q_n is a junction point, there exists $z > 0$ such that $[q_n, q_n + z]$ in a singular arc. By Theorem 3, on a singular arc $\lambda(q) = \left(\frac{u_q u_\theta}{u_{q\theta}} - u\right) h_1$ and $\mu(q) = \frac{u_q u_\alpha}{u_{q\theta}} h_1 - uh_2$. Therefore, we must have:

$$
\begin{aligned}
\int_0^{q_n} \mu'(q)dq &= \int_0^{q_n} -u_\alpha h_0 - u\frac{\partial h_0}{\partial \alpha} dq = \mu(q_n) = \frac{u_q u_\alpha}{u_{q\theta}} h_1 - uh_2, \\
\int_0^{q_n} \lambda'(q)dq &= \int_0^{q_n} -u_\theta h_0 - u\frac{\partial h_0}{\partial \theta} dq = \lambda(q_n) = \left(\frac{u_q u_\theta}{u_{q\theta}} - u\right) h_1.
\end{aligned}
$$

Using (69) and (77)-(78) we may compute:

$$\frac{u_q u_\alpha}{u_{q\theta}} h_1 - u h_2 = (1-\theta)\left(\frac{2\theta}{\gamma} q^{2-\gamma} - (b-\alpha) q\right)\frac{q^{\gamma-1}}{2} - (1-\theta)\left(\theta q - (b-\alpha)\frac{q^\gamma}{2}\right) = -(1-\theta)\theta\frac{\gamma-1}{\gamma} q_n, \tag{90}$$

$$\left(\frac{u_q u_\theta}{u_{q\theta}} - u\right) h_1 = (1-\theta)\left(\frac{2\theta}{\gamma} q^{2-\gamma} - (b-\alpha) q\right) - \left(\theta q - (b-\alpha)\frac{q^\gamma}{2}\right)\frac{2(1-\theta)}{\gamma q^{\gamma-1}} = -(1-\theta)(b-\alpha)\frac{\gamma-1}{\gamma} q_n \tag{91}$$

To check the continuity of μ equate (87) and (90) to obtain:

$$-(1-\theta)\theta\frac{\gamma-1}{\gamma} q_n = -\frac{\alpha\theta(\gamma-1)(q^*)^\gamma}{2} + \frac{\alpha\gamma(\gamma-1)(b-\frac{3}{2}\alpha)(q^*)^{2\gamma-1}}{4(2\gamma-1)} - \frac{(1-\theta)^2(\gamma-1)}{2\gamma}(q_n - q^*). \tag{92}$$

Now, substituting $1 = \theta + \frac{\alpha\gamma}{2}(q^*)^{\gamma-1}$. into the first and second terms on the right-hand side of (92) yields:

$$-(1-\theta)\theta\frac{\gamma-1}{\gamma} q_n = -\frac{(1-\theta)\theta(\gamma-1)q^*}{\gamma} + \frac{(1-\theta)^2(\gamma-1)(b-\frac{3}{2}\alpha)q^*}{\alpha\gamma(2\gamma-1)} - \frac{(1-\theta)^2(\gamma-1)}{2\gamma}(q_n - q^*). \tag{93}$$

which can be further simplified as follows:

$$q_n = q^* + \frac{(2b-3\alpha)}{\alpha(2\gamma-1)} q^*. \tag{94}$$

By inspection, equality (94) holds when $\alpha = \frac{2b}{3}$ and $q_n = q^*$.

Now let us check the continuity of the Lagrange multiplier λ at q_n. Equating (88) and (91) we obtain:

$$(1-\theta)(b-\alpha)\frac{\gamma-1}{\gamma} q_n = \frac{\alpha^2(\gamma-1)(q^*)^\gamma}{4} + \frac{(b-\alpha)(1-\theta)(\gamma-1)}{\gamma}(q_n - q^*) - \frac{(3\theta-1)(1-\theta)(\gamma-1)}{\gamma}\int_{q_n}^{q^*} q^{1-\gamma} dq$$

Using the equation $1-\theta = \frac{\alpha\gamma}{2}(q^*)^{\gamma-1}$ and cancelling the term on the left-hand side with the first part of the expansion of the second term on the right-hand side yields:

$$0 = \frac{(b-\frac{3\alpha}{2})(1-\theta)(\gamma-1)}{\gamma} q^* + \frac{(3\theta-1)(1-\theta)(\gamma-1)}{\gamma}\int_{q_n}^{q^*} q^{1-\gamma} dq \tag{95}$$

Obviously, (95) holds when $\alpha = \frac{2b}{3}$ and $q_n = q^*$, which completes the proof. *Q.E.D.*

Next, we consider the solution to Subproblem (ii) in (15) for an arbitrary fixed $\widehat{q} > 0$.

Lemma 13 *Let* $u(q,\theta,\alpha) = \theta q - \frac{b-\alpha}{2} q^\gamma$ *and suppose that the types are distributed uniformly over* $[0,1]^2$*. Let* $\bar{q} = \left(\frac{2}{\gamma(b-1)}\right)^{\frac{1}{\gamma-1}}$*, and* $\tilde{q} = \left(\frac{4}{\gamma(2b+1)}\right)^{\frac{1}{\gamma-1}}$*. If* $b \le \frac{3}{2}$*, then the solution to subproblem (ii) in (15) is as follows:*

$$\theta^\phi(q) = \frac{1+\gamma(b-1)q^{\gamma-1}}{3} \quad \text{for } \; q \in [\widehat{q}, \bar{q}]. \tag{96}$$

If $b > 3/2$*, then the solution to the problem (15) is given by:*

$$\theta^\phi(q) = \begin{cases} \frac{1+\gamma(b-1)q^{\gamma-1}}{3}, & \text{if } q \in [\tilde{q}, \bar{q}] \\ \frac{1+\gamma q^{\gamma-1}\left(\frac{2b-3}{4}\right)}{2}, & \text{if } q \in [0, \tilde{q}]. \end{cases} \tag{97}$$

Proof: By Theorem (6), the solution to the problem (15) is found by setting (32) to zero and solving that equation i.e., $\phi(q,\theta) \equiv u_q(q,1,\theta)H_\theta(q,1,\theta) + u_{\theta q}(q,1,\theta)H(q,1,\theta) = 0$.

First, let us compute $H(q,1,\theta) \equiv \int_{\underline{\alpha}(q,1,\theta)}^1 \int_{\sigma(q,\theta,1,a)}^1 f(t,a)dtda$. Recall that $\underline{\alpha}(q,1,\theta) = 1 - \frac{2(1-\theta)}{\gamma q^{\gamma-1}} > 0$ if $\theta > 1 - \frac{\gamma q^{\gamma-1}}{2}$ and $\underline{\alpha}(q,1,\theta) = 0$ otherwise, while $\sigma(q,\theta,1,a) = \min\{\theta + \frac{\gamma(1-a)}{2}q^{\gamma-1}, 1\}$ with $\sigma(q,\theta,1,a) = \theta + \frac{\gamma(1-a)}{2}q^{\gamma-1}$ for all $a \in [\underline{\alpha}(q,1,\theta), 1]$.

So, when $\underline{\alpha}(q,1,\theta) = 1 - \frac{2(1-\theta)}{\gamma q^{\gamma-1}} > 0$ we may compute:

$$\begin{aligned} H(q,1,\theta) &= \int_{\underline{\alpha}(q)}^1 \int_{\sigma(q,\theta,1,a)}^1 dtda = \int_{1-\frac{2(1-\theta)}{\gamma q^{\gamma-1}}}^1 \int_{\theta+\frac{\gamma(1-a)}{2}q^{\gamma-1}}^1 dtda = \int_{1-\frac{2(1-\theta)}{\gamma q^{\gamma-1}}}^1 1-\theta-\frac{\gamma(1-a)}{2}q^{\gamma-1}da \\ &= \frac{(1-\theta)^2}{\gamma q^{\gamma-1}} \end{aligned} \tag{98}$$

Then using (98) in the equation $\phi(q,\theta) \equiv u_q(q,1,\theta)H_\theta(q,1,\theta) + u_{\theta q}(q,1,\theta)H(q,1,\theta) = 0$ and solving yields:

$$\theta^\phi(q) = \frac{1+\gamma(b-1)q^{\gamma-1}}{3}. \tag{99}$$

Next, suppose $\underline{\alpha}(q,1,\theta) = 0$. Then we have:

$$H(q,1,\theta) = \int_0^1 \int_{\sigma(q,\theta,1,a)}^1 dtda = \int_0^1 \int_{\theta+\frac{\gamma(1-a)}{2}q^{\gamma-1}}^1 dtda = \int_0^1 1-\theta-\frac{\gamma(1-a)}{2}q^{\gamma-1}da = 1-\theta-\frac{\gamma}{4}q^{\gamma-1} \tag{100}$$

Then solving $\phi(q,\theta) \equiv u_q(q,1,\theta)H_\theta(q,1,\theta) + u_{\theta q}(q,1,\theta)H(q,1,\theta) = 0$ with (100) substituted in, yields:

$$\theta^\phi(q) = \frac{1+\gamma q^{\gamma-1}\left(\frac{2b-3}{4}\right)}{2}. \tag{101}$$

It remains to determine the intervals on which (99) and (101) hold respectively. First, note that a simple monotonicity argument shows that, if (99) applies at q_1, then it applies at $q_2 > q_1$. The highest q for which (99) applies, $\bar{q}$, is implicitly and uniquely defined by setting $\theta^\phi(\bar{q}) = 1$, which yields $\bar{q} = \left(\frac{2}{\gamma(b-1)}\right)^{\frac{1}{\gamma-1}}$. Further, Lemma 12 establishes that $\alpha^* = \frac{2b}{3} \le 1$ when $b \le \frac{3}{2}$. So, in this case $\underline{\alpha}(q,1,\theta^\phi(q)) \ge 0$ for all $q \ge \widehat{q}$, and hence (99) applies for all $q \in [\widehat{q}, \bar{q}]$.

If $b > \frac{2}{3}$, then $\alpha^* = 1$, and $\underline{\alpha}(q, 1, \theta^\phi(q)) = 0$ for $q \in [\widehat{q}, \tilde{q}]$, where $\tilde{q} = \left(\frac{4}{\gamma(2b+1)}\right)^{\frac{1}{\gamma-1}}$ is the solution to the equation $\underline{\alpha}(q, 1, \theta^\phi(q)) = 1 - \frac{2(1-\theta^\phi(q^{\gamma-1}))}{\gamma q} = 0$ for q. So, (101) applies for all $q \in [\widehat{q}, \tilde{q}]$, and (99) applies for all $q \in [\tilde{q}, \bar{q}]$. Q.E.D.

The following Lemma characterizes the solution to subproblem (i) for the case $b < \frac{3}{2}$ on its unique singular arc $[q^*, \widehat{q}]$ where $\alpha' < 0$.

Lemma 14 *Suppose that* $u = q\theta - \frac{b-\alpha}{2} q^\gamma$, $b < \frac{3}{2}$, *and* F *is uniform on* $[0, 1]^2$. *The solution to subproblem (i) on its unique singular arc* $[q^*, \widehat{q}]$ *is as follows:*

$$\theta^* \equiv \theta(q^*) = 1 - \frac{b\gamma(q^*)^{\gamma-1}}{3} \tag{102}$$
$$\widehat{\alpha} \equiv \alpha(\widehat{q}) = 1$$

The quantities q^* *and* $\widehat{q}$ *are uniquely defined as a solution to the following two equations:*

$$b(q^*)^\gamma (2 - b\gamma(q^*)^{\gamma-1}) = (b-1)\widehat{q}^\gamma (2 - (b-1)\gamma\widehat{q}^{\gamma-1}) \tag{103}$$

$$\left(\frac{1}{2} - \frac{b}{3}\right)\left(\frac{b\gamma(q^*)^\gamma}{3}(2 - b\gamma(q^*)^{\gamma-1})\right)^{\gamma-1} = \int_{\frac{1+\gamma(b-1)\widehat{q}^{\gamma-1}}{3}}^{1-\frac{b\gamma(q^*)^{\gamma-1}}{3}} ((1-\theta)(3\theta-1))^{\gamma-1}\, d\theta \tag{104}$$

Also,

$$\theta(q) = \frac{2 - \sqrt{1 - \frac{b\gamma(q^*)^\gamma(2-b\gamma(q^*)^{\gamma-1})}{q}}}{3} \text{ for all } q \in [q^*, \widehat{q}] \tag{105}$$

$$\alpha(q) = 2\left(\frac{b\gamma(q^*)^\gamma}{3}(2 - b\gamma(q^*)^{\gamma-1})\right)^{\gamma-1} \int_{\frac{2-\sqrt{1-\frac{b\gamma(q^*)^\gamma(2-b\gamma(q^*)^{\gamma-1})}{q}}}{3}}^{1-\frac{b\gamma(q^*)^{\gamma-1}}{3}} ((1-\theta)(3\theta-1))^{\gamma-1}\, d\theta \tag{106}$$

Proof of Lemma 14: By Lemma 12, $\alpha^* = \frac{2b}{3}$. Substituting this into equation $u_q(q^*, \theta^*, \alpha^*) = u_q(q^*, 0, 1)$ i.e., $\theta^* + \frac{b\gamma(q^*)^{\gamma-1}}{3} = 1$ (via which the triplet $(q^*, \theta^*, \alpha^*)$ is defined) gives us (102).

Further, Lemma 13 implies that, if $b \geq \frac{3}{2}$, then $\widehat{\theta} = \theta^\phi(\widehat{q}) \geq \theta^\phi(0) = \frac{1}{2}$, and if $b < \frac{3}{2}$ then $\widehat{\theta} = \tilde{\theta} \geq \frac{1}{3}$. Hence regardless of the value of b, we have $\widehat{\alpha} = 1$.

Next, combining (69) and (81) yields for all $q \in [q^*, \widehat{q}]$:

$$\theta'(q) = \frac{d\theta}{d\alpha}\alpha'(q) = -\frac{1}{2}\frac{(1-\theta)(1-3\theta)}{q(3\theta-2)}$$

which can be rearranged as follows: $\frac{(4-6\theta)d\theta}{(1-\theta)(1-3\theta)} = \frac{dq}{q}$. Integrating the above equation yields that for all $q \in [q^*, \widehat{q}]$ and some constant $k > 0$ we have:

$$q = \frac{k}{(1-\theta)(1-3\theta)}. \tag{107}$$

Evaluating (107) at q^* and making use of $\alpha^* = \frac{2b}{3}$ and (102) yields $k = \frac{b\gamma}{3}(q^*)^\gamma(2 - b\gamma(q^*)^{\gamma-1})$. Solving (107) for θ and using $k = \frac{b\gamma}{3}(q^*)^\gamma(2 - b\gamma(q^*)^{\gamma-1})$ yields (105).

Next, since $\theta(q)$ is continuous at $\widehat{q}$, the boundary between the domains of subproblems (i) and (ii), the value of (105) at $\widehat{q}$ must be equal to the first expression of (97) at $\widehat{q}$, $\frac{1+\gamma(b-1)\widehat{q}^{\gamma-1}}{3}$, yielding (103).

To compute (104), note that using (69) we get: $1 - \frac{2b}{3} = \widehat{\alpha} - \alpha^* = \int_{\theta^*}^{\widehat{\theta}} \alpha'(\theta)d\theta = \int_{\theta^*}^{\widehat{\theta}} -\frac{2}{q^{\gamma-1}}d\theta = \int_{\widehat{\theta}}^{\theta^*} \frac{2}{q^{\gamma-1}}d\theta$. Substituting (107) and $k = \frac{b\gamma}{3}(q^*)^\gamma(2 - b\gamma(q^*)^{\gamma-1})$ into the last equation yields (104).

Finally, let us compute $\alpha(q)$. We have: $\alpha(q) - \frac{2b}{3} = \alpha(q) - \alpha^* = \int_{\theta^*}^{\theta(q)} \alpha'(\theta)d\theta = \int_{\theta^*}^{\theta(q)} -\frac{2}{q^{\gamma-1}}d\theta = \int_{\theta(q)}^{\theta^*} \frac{2}{q^{\gamma-1}}d\theta = 2\left(\frac{b\gamma(q^*)^\gamma}{3}(2 - b\gamma(q^*)^{\gamma-1})\right)^{\gamma-1} \int_{\frac{2-\sqrt{1-\frac{b\gamma(q^*)^\gamma(2-b\gamma(q^*)^{\gamma-1})}{q}}}{3}}^{1-\frac{b\gamma(q^*)^{\gamma-1}}{3}} ((1-\theta)(3\theta-1))^{\gamma-1}\,d\theta.$ *Q.E.D.*