Does the Monopoly Need to Exclude?

Sergei Severinov and Raymond Deneckere *

First Version: November 25, 2002
This Version:

Abstract

We examine the optimal selling strategy of a monopolist facing consumers who have privately known demands and whose abilities to misrepresent their preferences may be limited. We derive the optimal mechanism in this case and characterize its properties. In contrast to standard environments, Revelation or Taxation Principles do not hold here, and the choice of a game form affects the performance of the mechanism. We show that consumers who have better abilities to misrepresent information benefit from the presence of consumers whose lack such abilities. We also demonstrate that whenever the fraction of the latter groups of consumers is positive, there will be no exclusion: the firm will supply a positive quantity of the good to all consumers whose valuation exceeds the cost of production.

JEL Nos: C72, D82

Keywords: mechanism design, screening, honesty,

1 Introduction.

The nature and qualitative properties of optimal selling strategies for a profit-maximizing monopolist have been thoroughly explored by many authors. The relevant literature contains detailed analysis of a broad range of selling mechanisms, marketing techniques and pricing schemes, such as different forms of price discrimination, bundling and tying (see, for example, Tirole (1988) and Wilson (1993)), and encompasses a variety of environments. The most ubiquitous situation is one where the monopolist faces a population of heterogeneous consumers who have private information about their preferences. It is well-known that the optimal mechanism in this case can be implemented via a simple non-linear pricing schedule (e.g. Maskin and Riley (1987)). This is the essence of the Taxation Principle.

In practice, however, firms possessing significant market power often do not only use non-linear pricing techniques, but in addition also rely on a significant amount of communication and interaction with each customer. Firms in many industries, e.g. car dealerships, insurance companies and airlines, try to obtain information regarding income, occupation, demographic

---

*Fuqua School of Business, Duke University and Department of Economics, University of Wisconsin, Madison. We thank seminar participants at the University of Wisconsin, UCLA, CalTech, Winter 2001 North American Econometric Society Meetings, and 2001 Canadian Economic Theory Conference for helpful comments and suggestions. We are responsible for all errors. email: severin@ssc.wisc.edu
status, as well as the tastes and habits of their customers before making a sale to them.\footnote{For example, insurance companies price car insurance on the basis of self-reported consumer characteristics, such as the percentage of travel on a vehicle that is attributable to driving to and from work, that are largely unverifiable, and indeed appear to remain unverified.} This information is clearly related to the customer’s willingness to pay for the good. It is often collected through voluntary questionnaires, although sometimes the consumers may be requested to provide evidence supporting their claims.

The evidence shows that firms (in particular, car dealerships and airlines) use such information to offer the same goods or service to different individuals at different prices depending on the customer’s responses.\footnote{Price discrimination based upon personal interviews by salespeople in care dealerships is so rampant that in the 1990 launch of its new Saturn brand GM decided to take advantage of the public perception of unfair price differences by offering a “no haggle” pricing policy. According to this policy, all dealers in the same geographic area must charge the same price for identical cars.} In internet commerce, it is becoming more common for the prices quoted by internet stores to depend on the path which the person has used to access the site. The path itself, i.e. the history of the customer’s visits to other sites and her responses to questionnaires along the way, contain information about the customer’s preferences.\footnote{According to a recent McKinsey report (McKinsey & Co. 2000), companies with an on-line presence can use a multitude of sources to help determine a customer’s price sensitivity, such as “clickstream” information about the customer’s current on-line session, or... customer buying histories tracked in databases or stored in ‘cookies’ created on customers’ computers. As an example, the report mentions that Ford is tracking individual customer history and behavior on the Internet, enabling the company to finely target tailored promotions at specific customers. As a result, Ford expects to significantly improve the yield on the nearly $10 billion it spends annually on promotional pricing.}

Obviously, selling mechanisms that offer identical products or services at different prices depending on the information provided by the consumer would not be feasible if all consumers could easily imitate one another: consumers would infer that their responses affect the price, and therefore provide answers signalling that their willingness to pay for the good is low.

Although environments in which individuals can costlessly and effortlessly manipulate and misrepresent their private information are standard in the economic literature, they clearly represent an extreme point in the spectrum of various possibilities. In fact, it is likely that at least some individuals have limited abilities to misrepresent their true types and imitate the behavior of others. There are several reasons for this.

At the most basic level, some consumers may not understand whether or how their behavior will affect their subsequent terms of trade. One may think about such consumers as naive or possessing bounded rationality. For example, in internet commerce, many consumers are unaware that “cookies” allow manufacturers and retailers to monitor customers’ behavior, not only on their own site, but also on competitors’ sites. Savvy consumers employ different strategies to avoid being charged higher prices on the basis of their browsing history and/or past purchases, such as periodically erasing the cookies on their computer or using different computers to get price quotes.\footnote{We have witnessed an instance of an airline internet pricing in which two tickets for the same itinerary were offered at a higher per-ticket price when purchased in one transaction than if the same tickets were purchased in two separate transactions. Apparently, the airline perceived that two people travelling together could be a couple who have a higher willingness to pay. Yet, a savvy couple could get a lower price by purchasing the two tickets in two simultaneous transactions.}
Secondly, an individual may be unable or unwilling to misrepresent her information if she or he is naturally averse to lying. For some individuals, the act of lying may be associated with stress or discomfort (“blushing”, “feeling wrong”) causing a disutility. This may be due to psychological or ethical reasons. Erard and Feinstein (1994) argue that “some taxpayers appear to be inherently honest, willing to bear their full tax burden even when faced with financial incentives to underreport their income. Evidence for such inherently honest taxpayers derives not just from casual introspection; it is also supported by econometric evidence and survey findings...” Alger and Ma (1998) advocate a similar view. They maintain that some physicians have stronger ethical views and are not able to exaggerate the medical problems of a patient when requesting coverage from an HMO, while other physicians are willing to do so. Experimental evidence confirms that a nonnegligible portion of the population choose not to lie regarding their private information, even though lying increases their monetary payoff. Thirdly, in some environments messages may be supported by submission of credible or verifiable claims. Indeed, individuals are often asked to support their statements with some form of evidence. For example, telecommunication firms provide discounts to households that can credibly verify that their incomes are below a certain threshold. The government requires taxpayers to justify deductions on their tax return forms with receipts or other forms of evidence. Clearly, the failure of an individual to provide evidence which is known to be available to certain types can serve as a proof that the given individual if of a different type. This makes imitation harder for the individuals who lack the technology, skills or evidence to support their claims in a credible way. At the same time, other individuals who have better skills or access to the technologies may be able to imitate other types without any difficulty.
Lippman and Seppi (1995) study an environment where messages can be supported by ‘credible claims.’ Green and Laffont (1986) consider a situation in which the set of types that an agent can imitate depends on the agent’s own type. This situation also admits interpretation in the spirit of the ‘credible claims’ model. Che and Gale (2000) study an optimal mechanism for selling a good to a buyer who may be budget constrained. In this case the ability of a buyer to misrepresent her willingness to pay would be limited, if the seller could ask her to post a bond and thus credibly disclose some information about her budget.

Finally, misrepresenting the truth may require costly physical actions. Lacker and Weinberg (1989) argue that ‘there are many instances in which lying about the state of nature requires more than simply sending a false signal regarding one’s private information. Often, costly actions must be taken to lend credence to the signals being sent.’ For example, a consumer may have to hide her assets or move to a less affluent neighborhood if her report of a low wealth and/or income is to be credible.

The main goal of this paper is to examine how the presence of consumers, who have limited ability to misrepresent their private information and imitate others, affects the optimal selling mechanism of a monopolist and, in particular, whether or how the monopolist can extract the private information of such consumers at little or zero cost. The presence of such consumers is incorporated into a standard screening model in the simplest possible way: we assume that a certain fraction of consumers always provide true information about their willingness to pay for the good when asked to report it. For brevity, such consumers will be referred to as ‘honest.’ (Alternatively, one can refer to these consumers as naive or boundedly rational.) All other consumers who can misrepresent their valuations costlessly will be referred to as ‘strategic.’ Note that whether a consumer is ‘honest’ or ‘strategic’ is not an observable characteristic, because a ‘strategic’ consumer can easily imitate an ‘honest’ one. So, the firm cannot simply segment the market into two parts, i.e. the third-degree price discrimination is not feasible.

We derive the optimal selling mechanism for such environment, and characterize its properties. Our analysis consists of two steps. First, we derive an optimal game form for the selling mechanism. Second, we characterize the unique optimal allocation profile implementable via the optimal game form.

In a standard environment where all consumers are ‘strategic’ and choose a report maximizing their expected payoffs, the choice of a game form has no real significance. This conclusion stems from the Revelation Principle (or, the Taxation Principle). However, the Revelation Principle does not hold here. The mechanism designer can typically take advantage of the fact that different consumers have different sets of feasible messages, and elicit more information about types by constructing a game form where some types submit non-truthful reports in equilibrium.\footnote{The reporting of valuations need not be understood literally. It can be seen as a reduced form representing the ultimate results of such activities as submission of evidence or actions directed at concealing one’s willingness and ability to pay for the good.}

\footnote{The failure of the Revelation Principle in a general class of environments that includes the one studied here has been demonstrated by Green and Laffont (1986). In (Deneckere and Severinov 2001), we argue that an extended version of the Revelation Principle applies in such environments, but the choice of the game form remains important.}
We establish that the following game form which we call a ‘password mechanism,’ is optimal in our case. First, a consumer is asked to report her valuation. Then, depending on her report, she is either offered a specific quantity/transfer pair, or is given a menu of quantity/transfer pairs to choose from. The optimality of the ‘password’ mechanism stems from the fact that it allows the principal to implement allocation profiles satisfying a minimal set of incentive constraints. In particular, no incentive constraints of the ‘honest’ consumers have to be satisfied.

Using this mechanism, we derive the unique optimal allocation profile (i.e. the set of quantity/transfer pairs- one for each consumer type) for an arbitrary fraction of ‘honest’ consumers in the population and characterize its properties.

Overall, the presence of ‘honest’ consumers tends to increase the efficiency of the optimal quantity schedule: the quantity consumed by each type is less distorted relative to the first-best than in the standard case with no ‘honest’ consumers. However, the quantity distortions do not disappear even as the fraction of ‘honest’ consumers converges to 1. Not surprisingly, all ‘honest’ consumers are held at their reservation utility level. At the same time, the surplus earned by the strategic ‘consumers’ increases as the fraction of the ‘honest’ consumers get larger. So, the ‘strategic’ consumers benefit from the presence of the ‘honest’ ones in the population.

The most surprising qualitative property of the optimal allocation profile is the absence of exclusion. We demonstrate that all consumers who have a positive valuation for the good end up consuming a positive quantity, no matter how small the fraction of the ‘honest’ consumers may be.

In contrast, exclusion appears to be a very robust qualitative feature of the optimal screening mechanisms in the standard case with no ‘honest’ types: a profit-maximizing monopoly will choose not to sell to consumers whose willingness to pay for the good is low, but is nevertheless above the production cost. (see for example Maskin and Riley (1987)). This result also extends to settings with multidimensional private information where a monopolist sells multiple products to the consumers (see Armstrong (1996) and Chone and Rochet (1998)).

The intuition behind the optimality of exclusion is well-understood. If some low valuation consumers were offered a positive quantity, it would enhance the overall efficiency. Yet, the firm would have to leave a larger surplus to the higher-valuation consumers in order to prevent them from imitating the low-valuation ones. The latter effect is more significant when a consumer’s valuation is not too high above the production cost. Therefore, it is optimal for the firm not to sell to such consumers at all.

The prospect of exclusion is troubling and may be unacceptable from a social point of view, especially when it concerns such vitally important areas as telecommunication, energy or transportation. It provides a strong argument in support of government regulation of monopolistic industries. Indeed, it is easy to come up with simple examples where the monopolist’s optimal strategy is to exclude a significant proportion of consumers.\textsuperscript{12}

However, our no-exclusion result suggests that such concerns may not be well-founded, as even unregulated monopolists may find it optimal to serve all consumers whose willingness to

\textsuperscript{12}For example, if a firm with production costs equal to $q^2/2$ faces a population of consumers whose valuation for the good is equal to $\theta q$ where $\theta$ is consumer’s private information and is distributed uniformly in the population, then it is optimal for the firm to exclude all consumers whose valuations are below 1/2.
pay for the good is above the marginal cost.

Technically, the characterization of the optimal allocation profile in our environment belongs to the class of multidimensional screening problems. Thus our paper also contributes to this literature.

2 Model and Main Results.

A monopolistic supplier is facing a consumer randomly drawn from a population. The consumer has private information about her preferences. Specifically, she gets utility \( u(q, \theta) \) from consuming quantity \( q \) of the good supplied by the firm. The preference parameter \( \theta \) is consumer’s private information and takes values in the interval \([0, 1]\). \( \theta \) is distributed on \([0, 1]\) in the population from which the consumer is drawn. The probability distribution \( F(\theta) \) is common knowledge. The firm incurs the cost equal to \( c(q) \) when it supplies quantity \( q \) to the consumer.

Equivalently, consider a market in which the firm is facing a continuum of consumers each of whom has an inelastic unit demand for the good. Consumers are heterogeneous in their valuations for quality: a consumer’s valuation for a unit of the good is equal to \( q \) where \( q \) is consumer’s privately known valuation and \( q \) is quality chosen by the firm. The firm’s cost of production is additive: the marginal cost of producing a unit of quality \( q \) is constant and equal to \( c(q) \). It is easy to see that these two formulations are equivalent. We also assume that consumer’s reservation utility level is 0.

We assume that \( u(q, \theta) \) and \( c(q) \) are twice continuously differentiable, \( u(q, \theta) - c(q) \) and \( u(q, \theta) - c(q) - \frac{1-F(\theta)}{F(\theta)} u_\theta(q, \theta) \) are concave in \( q \) \( \forall \theta \in [0, 1] \), and \( F(\theta) \) satisfies the monotone likelihood ratio property. We also assume that \( c(0) = 0, \ c'(0) = 0, \ u(0, \theta) = 0 \ \forall \theta \in [0, 1], \ u_q(q, \theta) > 0 \ \forall \theta \in (0, 1] \) and \( q \geq 0 \) and \( u_\theta(q, \theta) > 0 \ \forall \theta \in (0, 1] \) and \( q > 0 \). Also, \( \exists Q > 0 \) s.t. \( \forall q > Q \) and \( \theta \in [0, 1] \) \( u(q, \theta) - c(q) < 0, \ u_{q\theta}(q, \theta) > 0 \ \forall q \geq 0 \) and \( \theta \in [0, 1] \). These assumptions imply that there exists a unique positive \( q^*(\theta) = \arg\max u(q, \theta) - c(q) \) and \( q^*(\theta) \) is increasing in \( \theta \).

To this standard model we add an additional assumption that a fraction \( \gamma \in (0, 1) \) of consumers are ‘honest’. An ‘honest’ consumer is not able or not willing to misrepresent her valuation, and truthfully reveals it when asked to report it. The rest of the consumers behave in a standard fashion: they can and will always misrepresent their type if this allows them to obtain a larger surplus. We refer to such consumers as ‘strategic.’ Whether a consumer is ‘honest’ or ‘strategic’ is not observable, since a ‘strategic’ consumer can imitate an ‘honest’ one. We assume that whether a consumer is ‘honest’ or ‘strategic’ is independent of her valuation.

Our goal is to understand how the presence of the ‘honest’ consumers affects the optimal selling mechanism. The characterization of the optimal mechanism will involve two steps. First, we will design an optimal game form for the mechanism. Second, we will derive an allocation profile that maximizes the firm’s expected profits among all allocation profiles that are implementable via the chosen game form.

To define an allocation profile, let \( q(\theta) \ (g(\theta)) \) be the quantity obtained by a ‘strategic’ (‘honest’) consumer with valuation \( \theta \), and \( t^s(\theta) \ (t^e(\theta)) \) be the transfer paid by her. Then an
allocation profile is a collection of vectors \((q(\theta), t^s(\theta), g(\theta), t^r(\theta)) \forall \theta \in [0, 1]\). An allocation profile is implementable via a game form if it is optimal for each consumer type to choose a strategy giving her an allocation corresponding to her true type, i.e. a ‘strategic’ (‘honest’) consumer with valuation \(\theta\) chooses a strategy that gives her an allocation \(q(\theta), t^s(\theta) (g(\theta), t^r(\theta))\).

Consider the choice of a game form. As a preliminary step, note that a mechanism that uses only a non-linear pricing schedule is not optimal in this environment, because a consumer’s choice from such schedule will only depend on her valuation. Hence, the firm would not be able to differentiate ‘honest’ consumers from ‘strategic,’ and exploit the presence of the former.

Next, consider a class of mechanisms which are called ‘direct’ in a standard environment: a consumer is asked to report her valuation and is assigned an allocation (quantity/quality and transfer) based on her report. Since ‘honest’ consumers always report their valuations truthfully, while ‘strategic’ consumers choose reports maximizing their payoffs, the firm would face a choice between two alternatives. First, it could offer an allocation profile which keeps the consumers reporting truthfully at their reservation utility levels. This strategy allows to extract full surplus from ‘honest’ consumers. However, the ‘strategic’ consumers faced with such an allocation profile will choose to underreport their valuations, which would reduce the efficiency of the mechanism and the firm’s expected profits.

Alternatively, the firm can extract a larger surplus from ‘strategic’ consumers by offering an allocation profile which makes reporting the true valuation incentive compatible. However, incentive compatibility comes at a cost: the firm has to leave some surplus to all consumers who report their valuations truthfully, including the ‘honest’ ones.\(^\text{13}\)

In fact, we will show that the firm can do better than in any such quasi-direct mechanism.\(^\text{14}\) It can choose a game form which allows to distinguish the ‘honest’ consumers from the ‘strategic’ ones without leaving any surplus to the former, and also induce the ‘strategic’ ones to make self-selecting choices in the mechanism.

Generally, a game form is optimal if it allows to implement the largest set of allocation profiles or, equivalently, if in any alternative game form the set of incentive constraints which have to be imposed on an implementable allocation profile is (weakly) larger. Since a ‘strategic’ consumer can always imitate any other type, an allocation profile cannot be implemented via any game form unless it satisfies all incentive constraints of ‘strategic’ consumers. Hence, a game form is optimal if an allocation profile implementable via it has to satisfy only the incentive constraints of ‘strategic’ consumers.

Consider the following “Password” Mechanism:

\(^\text{13}\) Alger and Ma (1998) study how this tradeoff is resolved in a related model with two possible valuations.

\(^\text{14}\) A direct mechanism in our environment is one where the consumer is asked to report both her valuation and whether she is ‘honest’ or ‘strategic.’ Recall that ‘honest’ consumers are the ones who have limited ability to misrepresent their valuations. Thus, an ‘honest’ person can claim to be able to misrepresent the truth. We believe that this view is appropriate, since the question whether the person is ‘honest’ can be viewed as hypothetical, so that in response to it an ‘honest’ person can also give a hypothetical answer. In this context, direct mechanisms perform no better than the quasi-direct mechanisms, where the consumers are only asked to report their valuations.

Alternatively, if an ‘honest’ consumer could not claim that she is capable of misrepresentation, the optimal allocation profile in the direct mechanisms would be identical to the one derived in the next section of this paper.
Stage 1. A consumer reports her valuation.

Stage 2. (a) If the reported valuation \( \hat{\theta} \) is strictly greater than 0 (the lowest valuation), then the consumer is assigned the allocation \( g(\hat{\theta}), t^*(\hat{\theta}) \).

(b) If the reported valuation is 0 (the lowest valuation), then the consumer is given a choice from the menu which includes \( (q(\theta), t^s(\theta)) \) \( \forall \theta \in [0, 1] \) and \( g(0), t^*(0) \).

**Theorem 1** The ‘password’ mechanism is optimal, i.e. any allocation profile implementable via another game form can also be implemented via the ‘password’ mechanism.\(^{15}\)

**Proof:** Consider an allocation profile \( (q(\theta), g(\theta), t^s(\theta), t^*(\theta)) \) implementable via some game form \( \Gamma \). Let us show that it is also implementable (or a.e. implementable) via the ‘password’ mechanism.

Since a ‘strategic’ consumer can imitate any other type in any game form including \( \Gamma \), this allocation profile has satisfy all incentive constraints of all the ‘strategic’ consumers (or, at least, the set of measure 1 of the ‘strategic’ consumers), i.e. the allocation \( (q(\theta), t^s(\theta)) \) has to provide more surplus for a consumer with valuation \( \theta \) than the allocation \( (q(\theta'), t^s(\theta')) \) \( \forall \theta' \neq \theta \), or the allocation \( (g(\theta'), t^*(\theta')) \) \( \forall \theta' \in [0, 1] \). This implies that the ‘strategic’ consumers will not imitate other types if the principal offers \( (q(\theta), g(\theta), t^s(\theta), t^*(\theta)) \) in the ‘password’ mechanism.

Further, in stage 1 of the ‘password’ mechanism the only feasible reporting strategy for an ‘honest’ consumer with valuation \( \theta > 0 \) is to report her true valuation. So, an ‘honest’ consumer with valuation \( \theta \) cannot deviate from the allocation \( (g(\theta), t^*(\theta)) \).

Thus, an allocation profile satisfying all incentive constraints of the ‘strategic’ consumers is almost always implementable via the ‘password mechanism’, i.e. there is at most one consumer type - an ‘honest’ consumer with valuation 0 who gets access to the menu \( \{q(\theta'), t^s(\theta')| \theta' \in [0, 1]\} \) - who may choose to deviate from an allocation designed for her.\(^{16}\) Q.E.D.

The ‘password’ mechanism combines an implementation method based on the Revelation Principle (reporting the valuation) with an implementation method based on the Taxation Principle (choosing from a menu). Intuitively, the ‘password’ mechanism is optimal because it allows to ignore all incentive constraints of ‘honest’ consumers. An ‘honest’ consumer cannot get access to the menu designed for strategic consumers, because she needs to report valuation 0 to do so. Thus, the report of the lowest valuation \( \theta = 0 \) can be viewed as a ‘password’ necessary to access the menu.

In the first stage of the password mechanism almost all ‘strategic’ consumers misrepresent their valuations. This raises the question whether there exists an optimal incentive compatible mechanism where the valuations are reported truthfully. The answer to this question is

---

\(^{15}\)Formally, we show that an allocation profile that satisfies incentive constraints of the ‘strategic’ consumers is *almost everywhere* (a.e.) implementable via ‘password’ mechanism. That is, the set of types who choose strategies giving them the allocations corresponding to their true types is of measure 1. But, any two allocation profiles which differ only on a set of types of measure zero are associated with the same expected profits for the firm. So, there is no loss in considering game forms that guarantee implementation only almost everywhere.

\(^{16}\)As will be shown below, the optimal allocation profile is such that, although we do not impose any incentive constraints on the allocation \( (g(0), t^*(0)) \), the ‘honest’ consumer with valuation 0 would not choose any allocation from \( \{q(\theta'), t^s(\theta')| \theta' \in [0, 1]\} \), instead. So, in fact, the optimal allocation profile is implemented everywhere. We discuss a.e. implementation here for the proof to be formally correct.
negative. Green and Laffont (1986) have demonstrated that the Revelation Principle fails, i.e. incentive compatible direct mechanism may be suboptimal, in environments where some agents are not able to send certain message and, in particular, to misrepresent themselves as certain other types. Intuitively, inducing some types to lie can help to eliminate incentive constraints of other types, thereby increasing the set of implementable allocation profiles.

In a companion paper (Deneckere and Severinov 2001), we develop a general approach to implementation in a broad class of environments where sending certain messages can be costly for some types. This class includes the environments considered by Green and Laffont (1986), as well as the one studied in this paper. In particular, one can regard ‘honest’ consumers as the ones who incur a very high cost when they misrepresent their valuations. Deneckere and Severinov (2001) establish an Extended Revelation Principle which allows to characterize the set of implementable social choice functions in these environments via mechanisms in which the agents are asked to send more than one message. Potentially, the agents may be asked to send all messages that they are capable of sending. In equilibrium, these messages truthfully reveal the agent’s type, i.e. her preference parameter as well as her communication abilities. Therefore, this result can be viewed as an extension of the Revelation Principle.\textsuperscript{17}

Theorem 1 implies that the optimal allocation profile can be derived as a solution to the firm’s profit maximization problem subject to the incentive constraints of the ‘strategic’ consumers and the individual rationality constraints of all types. This is done in the following section. This problem is significantly more complex than in the standard environment without ‘honest’ consumers, because the solution has to satisfy a larger set of incentive constraints. This is related to the fact that our problem is two-dimensional: the consumers differ in their valuations, as well as with regards to being ‘honest’ or ‘strategic.’

The literature on non-linear pricing with multidimensional private information (e.g. Wilson (1993), Armstrong (1996), Chone and Rochet (1998)) points out that identifying the set of binding constraints is the key step towards characterizing the optimal mechanism.

In a key step of our analysis, we demonstrate that there are either one or two binding incentive constraints per ‘strategic’ consumer, depending on her valuation (see Lemma 5). First, as in the standard framework without ‘honest’ consumers, downward incentive constraints between ‘adjacent’ ‘strategic’ types are binding in the optimal mechanism. Consequently, the net payoff (informational rent) $U(\theta)$ of a ‘strategic’ consumer with valuation $\theta$ in the optimal mechanism is equal to $U(0) + \int_0^\theta u_\theta(q(s), s)ds$. (We also establish that $U(0) = 0$.)

The inability of an ‘honest’ consumer to lie implies that her first and the second messages ($\hat{\theta}_1$ and $\hat{\theta}_2$) must be the same and equal to her true valuation. So, her allocation $(g(\hat{\theta}_1), t'\tilde{\theta}(\hat{\theta}_1))$ need not satisfy any incentive constraints. A ‘strategic’ consumer can ‘prove’ that she is ‘strategic’ by reporting $\hat{\theta}_1$ and $\tilde{\theta}_2$ s.t. $\hat{\theta}_1 \neq \tilde{\theta}_2$. Her first report $\hat{\theta}_1$ will be equal to the her true valuation if it is incentive compatible, i.e. $(q(\hat{\theta}_1), t'\tilde{\theta}(\hat{\theta}_1))$ provides a bigger surplus to her than any other allocation.

Thus, an allocation profile satisfying all incentive constraints of the ‘strategic’ consumers is implementable via Extended Direct Mechanism, and so this mechanism is also optimal. We call it an Extended Direct Mechanism, because the first message indicates the consumer’s true valuation and the second message demonstrates her communication abilities (i.e. whether she is or not able to misrepresent her valuation).

\textsuperscript{17}In our case, the following Extended Direct Mechanism can be used to characterize the whole set of implementable allocation profiles:

The consumer is asked to report her valuation twice. Let $(\hat{\theta}_1, \hat{\theta}_2)$ denote her report. If $\hat{\theta}_1 = \hat{\theta}_2$, $(\hat{\theta}_1 \neq \hat{\theta}_2)$ then she is assigned an allocation $(g(\hat{\theta}_1), t'\tilde{\theta}(\hat{\theta}_1))$ designed for an ‘honest’ (‘strategic’) consumer with valuation $\hat{\theta}_1$.

The inability of an ‘honest’ consumer to lie implies that her first and the second messages ($\hat{\theta}_1$ and $\hat{\theta}_2$) must be the same and equal to her true valuation. So, her allocation $(g(\hat{\theta}_1), t'\tilde{\theta}(\hat{\theta}_1))$ need not satisfy any incentive constraints. A ‘strategic’ consumer can ‘prove’ that she is ‘strategic’ by reporting $\hat{\theta}_1$ and $\tilde{\theta}_2$ s.t. $\hat{\theta}_1 \neq \tilde{\theta}_2$. Her first report $\hat{\theta}_1$ will be equal to the her true valuation if it is incentive compatible, i.e. $(q(\hat{\theta}_1), t'\tilde{\theta}(\hat{\theta}_1))$ provides a bigger surplus to her than any other allocation.

Thus, an allocation profile satisfying all incentive constraints of the ‘strategic’ consumers is implementable via Extended Direct Mechanism, and so this mechanism is also optimal. We call it an Extended Direct Mechanism, because the first message indicates the consumer’s true valuation and the second message demonstrates her communication abilities (i.e. whether she is or not able to misrepresent her valuation).
Second, for a strategic type with valuation $\theta$ there is exactly one other incentive constraint that may be binding: the incentive constraint between this type and an ‘honest’ consumer with valuation $r(\theta)$ that satisfies $U(\theta) = u(q(\theta), \theta) - u(q(\theta), r(\theta))$. Lemma 5 establishes that an allocation profile is incentive compatible if for a ‘strategic’ consumer with any valuation $\theta$ the two mentioned incentive constraints are satisfied.

This incentive constraint reduction allows us to identify the link between the quantity schedules $q(.)$ and $g(.)$ in the optimal mechanism. Specifically, we show that the optimal schedule $g(.)$ is uniquely determined by the quantity schedule $q(.)$. These findings allow us to reduce the dimensionality of our problem and make it amenable to the optimal control methods.

Let us point out a few properties of the optimal allocation profile. First, an ‘honest’ consumer earns zero surplus irrespective of her valuation, because information regarding her valuation is elicited for free. Yet, this does not imply that an ‘honest’ consumer is always assigned an efficient quantity $q^*(\theta)$, because an ‘honest’ consumer with valuation $\theta$ cannot be assigned a higher quantity than a ‘strategic’ consumer with valuation $\theta'$ s.t. $r(\theta') = \theta$. This is a simple consequence of incentive compatibility. (In fact, our no-exclusion result implies that the set of ‘honest’ consumers for whom this constraint is binding is non-empty.)

So, if the firm attempts to reduce the informational rent $U(\theta)$ earned by some strategic consumer, it will need to reduce not only the quantities assigned to ‘strategic’ consumers with valuations less than $\theta$ (as in the standard case), but also the quantity allocations $g(r(\theta'))$ assigned to an ‘honest’ consumers with valuation $r(\theta')$ for some $\theta' < \theta$.

This well-known tradeoff between efficiency of an allocation profile and the informational rents implies that in the optimal mechanism both quantity schedules $q(\theta) \forall \theta \in [0, 1]$ and $g(\theta) \forall \theta \in [0, r(\theta)]$ will be distorted downwards below the efficient level. However, because the efficiency losses are exacerbated by the fact that the firm has to reduce the quantities assigned to the ‘honest,’ the balance of the tradeoff between higher efficiency and lower informational rents has to shift towards higher efficiency. Indeed, in the special case the solution to which is characterized explicitly in the paper, $q(.)$ is closer to the efficient level than $q^{st}(\theta)$ (the optimal quantity in the standard case), and $g(.)$ is even closer. This implies that $U(\theta) > U^{st}(\theta)$. Thus, the ‘strategic’ consumers benefit from the presence of the ‘honest’ ones. The former are paid more to prevent them from imitating the latter.

A particular manifestation of the shifting tradeoff between efficiency and informational rents is the following no-exclusion result.

**Theorem 2** For any $\alpha > 0$, the optimal allocation profile $q(\theta), g(\theta), t^s(\theta), t^*(\theta)$ is such that $q(\theta) > 0$ and $g(\theta) > 0 \forall \theta \in [0, 1]$.

**Proof:** see the Appendix.

This result stands in contrast with the standard case where the optimality of exclusion is a very robust property. As shown by Armstrong (1996), exclusion is also generic in the multidimensional non-linear pricing model.

Absence of exclusion in our mechanism is explained by two factors. First, in lemma 4 we establish the following ‘common cutoff’ property which follows from a simple analysis of incentive compatibility: if ‘strategic’ consumers with valuations below some threshold level $\theta > 0$ are assigned zero quantity, then so are the ‘honest’ consumers with valuations below $\theta$. 
Second, since the firm does not leave any surplus to the ‘honest’ consumers, rationing them is costly. Particularly, if the firm raises $q(\theta)$ on $[0, \theta]$ above zero by some small $\epsilon > 0$, then the profits that it collects from ‘strategic’ consumers decrease. However, the firm can now set $g(\theta) > 0$ on this interval, and collect extra profits from ‘honest’ consumers. We prove Theorem 2 by showing that the extra profits collected from the ‘honest’ consumers is of higher order of magnitude than the loss of profits collected from the ‘strategic’ consumers.

Below, we characterize the optimal allocation profile explicitly in a special but common case.

\textbf{Theorem 3} Suppose that $u(q, \theta) = \theta q$, $c(q) = \frac{q^2}{2}$ and $F(.)$ is uniform. Then the optimal allocation profile $(q(\theta), g(\theta), t^*(\theta), t^*(\theta))$ is given below:

$t^*(\theta) = \theta g(\theta) - \int_0^1 q(s) ds$, \hspace{1em} $t^*(\theta) = \theta g(\theta)$.

If $\alpha \neq 4$, then

\begin{align*}
&\forall \theta \in \left[0, \frac{2}{3} + \frac{1}{3(\sqrt{1 + 2\alpha} + 1)} \right], \hspace{1em} q(\theta) \text{ is uniquely defined by the following two conditions:} \\
&(i) \hspace{1em} q(\theta) \in \left[0, \frac{1}{3} + \frac{2}{3(\sqrt{1 + 2\alpha} + 1)} \right] \\
&(ii) \hspace{1em} \frac{1 - \alpha}{2 - \alpha/2} q(\theta) + \frac{1 + \frac{\alpha}{2(\sqrt{1 + 2\alpha} + 1)}}{(2 - \alpha/2) \left( \frac{1}{3} + \frac{2}{3(\sqrt{1 + 2\alpha} + 1)} \right)} \sqrt{\frac{2\alpha + 1}{\sqrt{\frac{2\alpha + 1}{3}}}} q(\theta)^{\frac{\sqrt{2\alpha + 1}}{2} - 1} = \theta \\
&\forall \theta \in \left[\frac{2}{3} + \frac{1}{3(\sqrt{1 + 2\alpha} + 1)}, 1 \right], \hspace{1em} q(\theta) = 2\theta - 1 \\
&\forall \theta \in \left[0, \frac{1}{3} + \frac{2}{3(\sqrt{1 + 2\alpha} + 1)} \right], \hspace{1em} g(\theta) \text{ is uniquely defined by the following two conditions:} \\
&(i) \hspace{1em} g(\theta) \in \left[0, \frac{1}{3} + \frac{2}{3(\sqrt{1 + 2\alpha} + 1)} \right] \\
&(ii) \hspace{1em} \frac{1 - \alpha}{4 - \alpha} g(\theta) + \frac{3\sqrt{\frac{2\alpha + 1}{\sqrt{\frac{2\alpha + 1}{3}}}}}{4 - \alpha} \left( \frac{\sqrt{1 + 2\alpha + 1}}{\sqrt{1 + 2\alpha + 3}} \right)^{\frac{\sqrt{2\alpha + 1}}{2} - 1} g(\theta)^{\frac{\sqrt{2\alpha + 1}}{2} - 1} = \theta \\
&\forall \theta \in \left[\frac{1}{3} + \frac{2}{3(\sqrt{1 + 2\alpha} + 1)}, 1 \right], \hspace{1em} g(\theta) = \theta
\end{align*}

If $\alpha = 4$, then

\begin{align*}
&\forall \theta \in [0, 3/4], \hspace{1em} q(\theta) \text{ is uniquely defined by the following two conditions:} \\
&(i) \hspace{1em} q(\theta) \in [0, 1/2] \hspace{1em} (ii) \hspace{1em} (3/2 + \log(1/2))q - q \log(q) = \theta \\
&\forall \theta \in [3/4, 1], \hspace{1em} q(\theta) = 2\theta - 1 \\
&\forall \theta \in [0, 1/2], \hspace{1em} g(\theta) \text{ is uniquely defined by the following two conditions:} \\
&(i) \hspace{1em} g(\theta) \in [0, 1/2] \hspace{1em} (ii) \hspace{1em} (1 + \log(1/2)/2)g - g \log(g)/2 = \theta \\
&\forall \theta \in [1/2, 1], \hspace{1em} g(\theta) = \theta
\end{align*}
For each $\alpha > 0$, $q(\theta)$ and $g(\theta)$ are characterized indirectly on the lower part of the range of valuations the optimal quantity schedules, i.e. $\theta$ is expressed as a function of $q$ and $g$ respectively. In the proof, we demonstrate that these functions are invertible, and so $q(\theta)$ and $g(\theta)$ are well-defined.

The two-part nature of the optimal schedules $q(\theta)$ and $g(\theta)$ is explained by the fact that the incentive constraint between a ‘strategic’ consumer with valuation $\theta$ and an ‘honest’ consumer with valuation $r(\theta)$ is binding (not binding) if the strategic consumer’s valuation belongs to the interval $[0, \tilde{\theta}]$. We establish that the optimal free boundary where the change of regime occurs is $\tilde{\theta} = \frac{2}{3} + \frac{1}{3(1+2\alpha^2+1)}$. The structure of binding incentive constraints is depicted in Figure 3.

The following Properties of the optimal schedules $q(\theta)$ and $g(\theta)$ follow immediately from Theorem 3:

1. $\forall \alpha > 0$ $q(\theta)$ and $g(\theta)$ are continuous, increasing and convex on $[0, 1]$.
   - $q(\theta)$ is strictly convex on $[0, \frac{2}{3} + \frac{1}{3(1+2\alpha^2+1)}]$.
   - $g(\theta)$ is strictly convex on $[0, \frac{1}{3} + \frac{2}{3(1+2\alpha^2+1)}]$.

2. $q^{st}(\theta) \leq q(\theta) < g(\theta) \leq \theta \\forall \theta \in (0, 1)$ where $q^{st}(\theta)$ is the optimal quantity in the standard case without ‘honest’ consumers.
   - $q^{st}(\theta) < q(\theta)$ iff $\theta \in (0, \frac{2}{3} + \frac{1}{3(1+2\alpha^2+1)})$.
   - $g(\theta) < \theta$ iff $\theta \in (0, \frac{1}{3} + \frac{2}{3(1+2\alpha^2+1)})$.

   Since in the first-best solution $q(\theta) = q(\theta) = \theta$, all quantity allocations are distorted downwards, except $g(\theta)$ on the interval $[\frac{1}{3} + \frac{2}{3(1+2\alpha^2+1)}, 1]$.

3. Smooth pasting: the left-hand and right-hand side derivatives of $q(\theta)$ at $\frac{2}{3} + \frac{1}{3(1+2\alpha^2+1)}$ are equal to 2.

Figures 1 and 2 illustrate the solution for three different values of $\alpha$: 10, 1 and 0.02. When $\alpha = 1$, i.e. half of the population is ‘honest’ and the other half is ‘strategic,’ the solution has the following closed form:

$$q(\theta) = \begin{cases} 
\sqrt{3^3} \sqrt{3} (\sqrt{3}-1) \sqrt{3+1} \theta^{3+1} & \text{if } \theta \in [0, \frac{1}{3-\sqrt{3}}] \\
2\theta - 1 & \text{if } \theta \in [\frac{1}{3-\sqrt{3}}, 1]
\end{cases}$$

$$g(\theta) = \begin{cases} 
\sqrt{3^3} \sqrt{3} \theta^{3+1} \left( g\left(\frac{\theta}{\sqrt{3}}-1\right) \right) & \text{if } \theta \in [0, \frac{1}{\sqrt{3}}] \\
\theta & \text{if } \theta \in [\frac{1}{\sqrt{3}}, 1]
\end{cases}$$

These Properties confirm that the presence of ‘honest’ consumers causes a shift of the tradeoff between higher efficiency and lower informational rents towards higher efficiency, and that the ‘strategic’ consumers benefit from the presence of the ‘honest’ ones.

12
Figure 1: Optimal quantity schedule $q(.)$ of the ‘strategic’ consumers.
Figure 2: Optimal quantity schedule $g(.)$ of the ‘honest’ consumers.
The informational rent $U(\theta)$ of a ‘strategic’ consumer with valuation $\theta > \bar{\theta} = \frac{2}{3} + \frac{1}{3(\sqrt{1 + 2\alpha_2 + 1})}$ is large enough that it exceeds the payoff which this consumer could get by imitating an ‘honest’ consumer with valuation $r(\theta)$, even if the latter was assigned the first-best quantity. Therefore, ‘honest’ consumers with valuations above $r(\bar{\theta})$ are assigned efficient quantities $g(\theta) = \theta$, while $q(\theta)$ for $\theta > \bar{\theta}$ is determined by the same informational rent vs. efficiency tradeoff as in the standard case. Hence $q(\theta) = q^{st}(\theta)$ on this interval. Nevertheless, because $U(\theta) = \int_0^\theta q(s) ds$, there is a sort of a ‘domino effect’ in the informational rents of ‘strategic’ consumers: because ‘strategic’ consumers with lower valuations are assigned larger quantities and are paid more not to imitate ‘honest’ types, $U(\theta)$ is strictly above $U^{st}(\theta)$ $\forall \theta \in (0,1]$.

By varying $\alpha$, the ratio of the ‘honest’ consumers to ‘strategic’ in the population, we establish a number of interesting comparative statics results described in the following corollary. Let us explicitly incorporate the dependence of the solution on $\alpha$ by using the notation $q(.,\alpha)$, $g(.,\alpha)$, $U(.,\alpha)$ and $\bar{\theta}(\alpha)$ for the quantity schedules, the informational rent and the threshold value.

**Corollary 1** For all $\alpha_1, \alpha_2$ s.t. $\alpha_1 > \alpha_2 > 0$:

(i) There exists a unique $\theta_c \in (0, \frac{2}{3} + \frac{1}{3(\sqrt{1 + 2\alpha_1 + 1})})$ (which depends on $\alpha_1$ and $\alpha_2$) s.t. $q(\theta_c, \alpha_1) = q(\theta_c, \alpha_2)$, $q(\theta, \alpha_1) > q(\theta, \alpha_2)$ $\forall \theta \in (0, \theta_c)$, and $q(\theta, \alpha_1) < q(\theta, \alpha_2)$ $\forall \theta \in (\theta_c, \frac{2}{3} + \frac{1}{3(\sqrt{1 + 2\alpha_2 + 1})})$.

(ii) $U(\theta, \alpha_1) > U(\theta, \alpha_2)$ $\forall \theta \in (0,1]$.

(iii) $g(\theta, \alpha_1) > g(\theta, \alpha_2)$ $\forall \theta \in (0, \frac{1}{3} + \frac{2}{3(\sqrt{1 + 2\alpha_2 + 1})})$.

Parts (ii) and (iii) are easy to understand. Clearly, as the proportion of the ‘honest’ consumers increases, the benefit to the firm from increasing $g(.)$ towards an efficient level

---

Note: $\alpha = \frac{\gamma}{1-\gamma}$ where $\gamma \in [0,1]$ is a fraction of ‘honest’ consumers in the population.
and extracting more surplus from the ‘honest’ consumers becomes larger than the cost of an associated increase in informational rents \( U(\cdot) \) paid to ‘strategic’ consumers whose fraction has decreased.

The intuition behind part (i) is similar, but more complex. Since \( g(\cdot, \alpha) \) and hence \( U(\cdot, \alpha) \) increase in \( \alpha \), and \( U(\theta, \alpha) = \int_0^\theta q(s, \alpha)ds \), \( q(\theta, \alpha) \) must also be increasing in \( \alpha \), at least for small \( \theta \). Intuitively, an increase in the proportion of the strategic consumers causes a ‘ripple’ effect in the form of higher \( q(\cdot) \). However, this effect disappears when \( \theta \) is high. This happens because the informational rent \( U(\theta, \alpha) \) of a ‘strategic’ consumer with valuation \( \theta > \theta(\alpha) \) is sufficiently large that this consumer strictly prefers not to imitate an ‘honest’ consumer even when all ‘honest’ consumers with valuations that exceed \( r(\theta(\alpha), \alpha) \) are assigned an efficient quantity. Therefore, if \( \theta > \theta(\alpha) \), it is no longer optimal for the firm to raise \( q(\theta, \alpha) \) above the quantity schedule optimal in the standard case.

Since the firm can set \( g(\cdot, \alpha) \) efficiently on the interval \([r(\theta(\alpha), \alpha), 1] \), \( \theta(\cdot) \) is decreasing in \( \alpha \). But then we have \( q(\theta(\alpha_1), \alpha) < q(\theta(\alpha_1), \alpha_2) \) for \( \alpha_2 < \alpha_1 \). So, \( q(\theta, \alpha) \) must be decreasing in \( \alpha \) when \( \theta \) is sufficiently high.

Essentially, when \( \alpha \) is large, then \( q(\cdot) \) is front-loaded: it is high when \( \theta \) is small, and relatively low when \( \alpha \) is large. The opposite is true when \( \alpha \) is small.

### 3 Proof of the Main Result.

Theorem 1 implies that an allocation profile is implementable if it satisfies the incentive constraints of all ‘strategic’ types, the individual rationality constraints of all ‘strategic’ and ‘honest’ types and the feasibility constraints \( q(\theta) \geq 0 \), \( g(\theta) \geq 0 \ \forall \theta \in [0, 1] \). Therefore, the firm’s problem can be stated as follows:

\[
\max_{(q(\theta) \geq 0, t^*(\theta), g(\theta) \geq 0, t^*(\theta))} \int_0^1 (t^*(\theta) - c(q(\theta))) f(\theta)d\theta + \alpha \int_0^1 (t^*(\theta) - c(g(\theta))) f(\theta)d\theta \quad (1)
\]

\[
u(q(\theta), \theta) - t^*(\theta) \geq u(q(\theta'), \theta) - t^*(\theta') \quad \forall \theta, \theta' \in [0, 1] \quad (2)
\]

\[
u(q(\theta), \theta) - t^*(\theta) \geq u(g(\theta'), \theta) - t^*(\theta') \quad \forall \theta, \theta' \in [0, 1] \quad (3)
\]

\[
u(q(\theta), \theta) - t^*(\theta) \geq 0 \quad \forall \theta \in [0, 1] \quad (4)
\]

\[
u(g(\theta), \theta) - t^*(\theta) \geq 0 \quad \forall \theta \in [0, 1] \quad (5)
\]

The presence of the second set of incentive constraints (3) for the ‘strategic’ consumers illustrates the multidimensional nature of our problem, and explains why the standard approach based on replacing the whole set of incentive constraints with a single differential equation does not work here. To derive a solution, we will transform Problem 1 into an equivalent optimal control problem. For this, we need to establish a number of preliminary results.

Let us establish several restrictions that can without loss of generality be imposed on the set of implementable allocation profiles. First, we need to consider only such allocation profiles where the individual rationality constraint (5) of an ‘honest’ consumer is binding for all \( \theta \in [0, 1] \). Otherwise, the value of 1 can be increased by raising the corresponding transfer \( t^*(\theta) \) without violating any incentive or individual rationality constraints. Hence, we can substitute \( u(g(\theta), \theta) \) for \( t^*(\theta) \) in (1) and eliminate (5).
Further, we will say that quantity schedules $q(\cdot)$ and $g(\cdot)$ are admissible if there exist such transfer functions $t^s(\cdot)$ and $t^r(\cdot)$ that the allocation profile $(q(\cdot), t^s(\cdot), g(\cdot), t^r(\cdot))$ satisfies (2)-(5) (i.e. is implementable). Then (2) implies that a feasible quantity schedule $q(\cdot)$ must be nondecreasing. The following lemma allows to restrict the set of admissible schedules further:

**Lemma 1** Without loss of generality, we can impose the following restrictions on the set of admissible quantity schedules in Problem 1:

(i) $g(\theta)$ is nondecreasing and satisfies $g(\theta) \leq q^*(\theta) \forall \theta \in [0, 1]$;

(ii) $q(0) = 0$ and $q(1) = q^*(1)$.

**Proof:** See the Appendix.

Lemma 1 has several implications. Since $q(\theta)$ is non-decreasing and bounded, it is Riemann integrable by Theorem 6.9, p.126 in Rudin (1976) and a.e. differentiable by Theorem 3, p.100 in Royden (1987). Hence, $u(q(\theta), \theta)$ is also bounded, Riemann integrable and a.e. differentiable. Let $U(\theta) \equiv u(q(\theta), \theta) - t^s(\theta)$. Incentive constraints (2) imply that $U(\theta)$ is increasing and satisfies the following inequalities $\forall \theta, \theta' \in [0, 1]$:

$$u(q(\theta'), \theta) - u(q(\theta'), \theta') \leq U(\theta) - U(\theta') \leq u(q(\theta), \theta) - u(q(\theta), \theta')$$

By the intermediate value theorem, $\exists \lambda_1, \lambda_2 \in [0, 1]$ s.t. $u_\theta(q(\theta'), \lambda_1 \theta + (1 - \lambda_1)\theta')(\theta - \theta') \leq U(\theta) - U(\theta') \leq u_\theta(q(\theta), \lambda_2 \theta + (1 - \lambda_2)\theta')(\theta - \theta')$. Since, $u_\theta(q(\theta'), \theta) \leq \max_{\theta \in [0, 1]} u_\theta(q^*(1), \theta) < \infty$, $U(\theta)$ is absolutely continuous. Therefore by Theorem 14, p.110 in Royden (1987), $U'(\theta) = u_\theta(q(\theta), \theta)$ and $U(\cdot)$ is equal to the integral of its derivative, i.e. $U(\theta) - U(0) = \int_0^\theta u_\theta(q(s), s)ds$.

It is easy to show that in the optimal mechanism $U(0) \leq u(q^*(1), 1)$. Therefore, $t^s(\theta) = u(q(\theta), \theta) - \int_0^\theta u_\theta(q(s), s)ds - U(0)$ is bounded on $[0, 1]$. Then the existence of a solution to Problem 1 follows because the objective (1) is a continuous functional, $q(\cdot), g(\cdot)$ and $t^s(\cdot)$ are bounded, and by a direct check we can show that the set of allocation profiles satisfying (2)-(5) is closed and convex, and is therefore, compact.

Let us use $t^s(\theta) = u(q(\theta), \theta) - \int_0^\theta q(s)ds - U(0)$ to substitute $t^s(\theta)$ out of (1) and eliminate (2). Then integrating by parts, we obtain that Problem 1 is equivalent to the following one:

$$\max_{q(\theta) \geq 0, g(\theta) \geq 0, U(0) \geq 0} -U(0) + \int_0^1 \left(u(q(\theta), \theta) - c(q(\theta)) - u_\theta(q(\theta), \theta) \frac{1 - F'(\theta)}{f(\theta)}\right) f(\theta)d\theta + \alpha \int_0^1 (u(q(\theta), \theta) - c(g(\theta))) f(\theta)d\theta$$

subject to: \hspace{1cm} (i) $q(\theta)$ is nondecreasing

(ii) ICT$(\theta, \theta')$: $U(\theta) \equiv U(0) + \int_0^\theta u_\theta(q(s), s)ds \geq u(g(\theta'), \theta) - u(g(\theta'), \theta') \forall \theta, \theta' \in [0, 1]$

We can now establish several additional properties of a solution. This is done in the following lemmas which are proved in the appendix:

---

19 This equality implies that, as in the standard case, only downwards incentive constraints between ‘adjacent’ types are binding among the ‘strategic’ consumers whenever $q(\cdot)$ is strictly increasing.
Lemma 2. Optimal quantity schedule \( q(\theta) \) is continuous.

Lemma 3. In the optimal mechanism \( U(0) = 0 \).

Lemma 4. Common cutoff. Optimal quantity schedules are such that \( \forall \theta \in [0, 1) q(\theta) = 0 \) if and only if \( g(\theta) = 0 \).

By lemmas (1)-(3), we can without loss of generality impose the following additional restrictions on the domain in Problem 6: (iii) \( q(.) \) is continuous, (iv) \( q(0) = 0 \), (v) \( g(\theta) \leq q(1) = q^*(1) \), (vi) \( U(0) = 0 \).

In the next lemma we establish that the family of incentive constraints \( ICT(\theta, \theta') \) in (8) can be replaced with a simpler set of constraints.

Lemma 5. Define \( r(\theta) \) as a unique solution to \( U(\theta) = u(q(\theta), \theta) - u(q(\theta), r(\theta)) \forall \theta \in [0, 1] \) s.t. \( q(\theta) > 0 \), and let \( r(\theta) = \theta \) if \( q(\theta) = 0 \).

If \( q(.) \) is nondecreasing, then \( r(\theta) \) is well-defined and nonnegative. Additionally, if \( q(.) \) is also continuous, \( g(\theta) \leq q(1) = q^*(1) \), and \( U(0) = 0 \), then:

(i) \( r(.) \) is continuous, increasing and strictly increasing at \( \theta \) if \( q(.) \) is strictly increasing at \( \theta \). Furthermore, \( r(0) = 0 \).

(ii) \( U(\theta) \geq u(g(\theta'), \theta) - u(g(\theta'), \theta') \forall \theta, \theta' \in [0, 1] \) if and only if \( q(\theta) \geq g(r(\theta)) \forall \theta \in [0, 1] \).

Since \( r(\theta) \) is continuous and nondecreasing and \( r(0) = 0 \), \( r(.) \) maps \([0, 1]\) onto \([0, r(1)]\).

The inverse image \( r^{-1}(\theta) \) from \([0, r(1)]\) is unique if \( q(.) \) is strictly increasing at \( \theta \) s.t. \( r(\theta) = \theta \).

However, even if \( r^{-1}(\theta) \) is not unique, \( q(r^{-1}(\theta)) \) is unique. This follows from (40) in the proof of the Lemma. Lemma 5 allows to establish the following important result:

Lemma 6. Fix a nondecreasing continuous quantity schedule \( q(\theta) \) s.t. \( q(1) = q^*(1) \), and set \( U(0) = 0 \). Then the optimal quantity schedule \( g(\theta) \) that maximizes (6) subject to (8) is given by:

\[
\begin{align*}
g(\theta) &= \min\{q^*(\theta), q(r^{-1}(\theta))\} \quad \text{if} \quad \theta \leq r(1) \quad \text{(9)} \\
g(\theta) &= q^*(\theta) \quad \text{if} \quad \theta > r(1)
\end{align*}
\]

Proof: By lemma 1, we can without loss of generality impose the restriction \( g(\theta) \leq q^*(1) \).

So, by (ii) of lemma 5, the family \( ICT(\theta, \theta') \) of incentive constraints in (8) can be replaced with the following family: \( q(\theta) \geq g(r(\theta)) \forall \theta \in [0, 1] \). Then the result follows because the integrand \( u(g(\theta), \theta) - c(g(\theta)) \) of the second integral in (6) is strictly concave in \( g(\theta) \) and is strictly increasing (decreasing) in \( g(\theta) \) if \( g(\theta) < q^*(\theta) (g(\theta) > q^*(\theta)) \).

Q.E.D.

Lemma 6 allows to reduce the dimensionality of our problem. According to (9), \( q(.) \) is completely determined by \( q(.) \), and so \( q(.) \) remains the only choice variable. Imposing the additional constraints (iii) — (vi), and using (9), we conclude that Problem (6) is equivalent
to the following one:

$$\max_{q(\theta) \geq 0} \int_0^1 \left( u(q(\theta), \theta) - c(q(\theta)) - u_\theta(q(\theta), \theta) \frac{1-F(\theta)}{f(\theta)} \right) f(\theta) d\theta + \alpha \int_{r(1)}\left( u(\min\{q^*(\tilde{\theta}), q(r^{-1}(\tilde{\theta}))\}, \tilde{\theta}) - c(\min\{q^*(\tilde{\theta}), q(r^{-1}(\tilde{\theta}))\}) \right) f(\theta) d\tilde{\theta}$$

subject to: $q(.)$ is nondecreasing, continuous, $q(0) = 0$, $q(1) = 1$.

Now we are ready to prove our no-exclusion result Theorem 2. Let $\tilde{\theta} \equiv \inf\{\theta | q(\theta) > 0\}$. The proof in the appendix demonstrates that $\tilde{\theta} = 0$.

Our next step is to reformulate this problem in order to apply standard methods of optimal control. For this, we will need to use the derivative $q'(\cdot)$ as the control variable, and so $q'(\cdot)$ has to be piecewise continuous or, equivalently, $q(.)$ has to belong to the space $C_p([0,1])$ of piecewise smooth (continuous and piecewise continuously differentiable) functions on $[0,1]$. So far, we have established that $q(.)$ must be continuous, i.e. $q(.) \in C([0,1])$, and that $q'(\cdot)$ exists almost everywhere. Nevertheless, the next lemma demonstrates that we can without loss of generality assume that $q(.) \in C_p^1([0,1])$.

**Lemma 7** If $q(\theta)$ is a solution to maximization problem (10) on the domain $C_p^1([0,1])$, then $q(\theta)$ also maximizes (10) on the domain $C([0,1])$.

**Proof:** see the Appendix.

Next, let us make a change of variables $\tilde{\theta} = r(\theta)$ in the second term of (10). By Lemma 5, $r(\theta)$ is continuous, increasing and bounded on $[0,1]$. Therefore, it is Riemann integrable and the change of variables is legitimate.

Observe that $r(\theta)$ is a.e. differentiable and piecewise continuous. Specifically,

$$r'(\theta) = \begin{cases} \frac{q'(\theta)(u_q(q(\theta), \theta) - u_\theta(q(\theta), r(\theta)))}{u_\theta(q(\theta), r(\theta))} & \text{if } q > 0 \\ 1 & \text{if } q = 0 \end{cases}$$

The piecewise continuity of $r'(\theta)$ follows from (11) and the fact that $q'(\theta)$ is piecewise continuous. Note that $r'(\theta)$ could be discontinuous at $\tilde{\theta} \equiv \inf\{\theta | q(\theta) > 0\}$. However, since $0 \leq r(\theta) \leq \theta$ and $r(\theta) = \tilde{\theta}$, the right-hand side limit $r'(\tilde{\theta} + 0)$ exists and does not exceed 1, so $r'(\cdot)$ is Riemann integrable. After the change of variables $\tilde{\theta} = r(\theta)$ the second term in (10) becomes:

$$\alpha \int_0^{r(1)} \left( u(\min\{q^*(r(\theta)), q(\theta)\}, r(\theta)) - c(\min\{q^*(r(\theta)), q(\theta)\}) \right) f(r(\theta)) r'(\theta) d\theta$$
Using (11), we obtain that Problem (10) is equivalent to the following one:

\[
\max_{q(\cdot)} \int_{0}^{1} \left( u(q(\theta), \theta) - c(q(\theta)) - u_\theta(q(\theta), \theta) \frac{1 - F'(\theta)}{f(\theta)} \right) f(\theta) d\theta + \alpha \int_{r(1)}^{1} \left( u(q^*(\theta), \theta) - c(q^*(\theta)) \right) f(\theta) d\theta + \alpha \int_{0}^{r(1)} \left( u(\min\{q^*(r(\theta)), q(\theta)\}, r(\theta)) - c(\min\{q^*(r(\theta)), q(\theta)\}) \right) f(r(\theta)) \frac{q'(\theta)(u_q(q(\theta), r(\theta)) - u_\theta(q(\theta), r(\theta)))}{u_\theta(q(\theta), r(\theta))} d\theta
\]

subject to: \[r'(\theta) = \frac{q'(\theta)(u_q(q(\theta), \theta) - u_\theta(q(\theta), r(\theta)))}{u_\theta(q(\theta), r(\theta))}, \quad q(0) = 0, \quad r(0) = 0, \quad q(1) = q^*(1), \quad q'(\theta) \geq 0 \quad (13)\]

Observe that (12) and (13) is an optimal control problem with control variable \(q'(\theta)\), two state variables \(q(\theta)\) and \(r(\theta)\), and a restriction on the control \(q'(\cdot) \geq 0\). It has ‘scrap value’ \(S(r_1) = \alpha \int_{r(1)}^{1} (u(q^*(\theta), \theta) - c(q^*(\theta))) f(\theta) d\theta\) at \(\theta = 1\).

The existence of a solution to this problem follows from Filippov-Cesari theorem (see Seierstad and Sydsæter (1987)). Pontryagin’s Maximum principle provides a standard method of solution to this problem. The Hamiltonian of the problem is given by:

\[
H(q, r, \lambda, \delta, \theta) = \left( u(q, \theta) - c(q) - u_\theta(q, \theta) \frac{1 - F'(\theta)}{f(\theta)} \right) f(\theta) + \alpha \left( u(\min\{q^*(r), q\}, r) - c(\min\{q^*(r), q\}) \right) f(r) \frac{q'(u_q(q, \theta) - u_\theta(q, r))}{u_\theta(q, r)} \quad (14)
\]

where costate variables \(\lambda(\theta)\) and \(\delta(\theta)\) are associated with the laws of motion of \(q(\theta)\) and \(r(\theta)\) respectively. Incorporating the constraint \(q'(\theta) \geq 0\), we get the following Lagrangian:

\[
\mathcal{L} = H(q, r, \lambda, \delta, \theta) + \tau q' \quad (15)
\]

where \(\tau \geq 0\) and \(\tau q' = 0\). The transversality conditions on the costate variable \(\delta\) is \(\delta(1) = \frac{dS(r_1)}{dr_1} = -\alpha(u(q^*(r(1)), r(1)) - c(q^*(r(1)))) f(r(1))\). Necessary conditions for the optimum are given by the following first-order conditions: \(-\lambda'(\theta) = \frac{\partial H}{\partial q'}, -\delta'(\theta) = \frac{\partial H}{\partial q}\) and \(\frac{\partial \mathcal{L}}{\partial q'} = 0\).

We will solve the problem explicitly in the special case where \(u(q, \theta + \theta q, c(q) = \frac{q^2}{2}, \) and \(F(.)\) is uniform. In this case, the Lagrangian is given by:

\[
\mathcal{L}(q, r, \lambda, \delta, \theta) = (2\theta - 1)q - \frac{q^2}{2} + \alpha \left( r \min\{r, q\} - \frac{(\min\{r, q\})^2}{2} \right) \frac{q'(\theta - r)}{q} + \lambda q' + \delta q'(\theta - r) + \tau q' \quad (16)
\]

The transversality conditions on the costate variable \(\delta\) becomes \(\delta(1) = -\frac{c(1)^2}{2}\).

Unfortunately, the problem does not satisfy any standard sufficient conditions (Arrow’s or Mangasarian’s). So, our proof will proceed as follows. We will derive the necessary conditions for optimality, and show that there is a unique schedule \(q(.)\) that satisfies them. The necessary first-order conditions will be derived by using Pontryagin’s Maximum Principle. These
conditions depend on whether \( q(\theta) > r(\theta) \) (case 1), \( q(\theta) < r(\theta) \) (case 2) or \( q(\theta) = r(\theta) \) (case 3), and whether the constraint \( q'(\theta) \geq 0 \) is binding or not.

**Case 1:** \( q(\theta) > r(\theta) \). The necessary first-order conditions are:

\[
\begin{align*}
-\lambda'(\theta) &= \frac{\partial H}{\partial q} = (2\theta - 1) - q - \frac{r^2 q'(\theta - r)}{2} - \frac{\delta q'(\theta - r)}{q^2} \quad (17) \\
-\delta'(\theta) &= \frac{\partial H}{\partial r} = \alpha r \frac{q'(\theta - r)}{q} - \left( \alpha \frac{r^2}{2} + \delta \right) \frac{q'}{q} \\
0 &= \frac{\partial L}{\partial q'} = \left( \alpha \frac{r^2}{2} + \delta \right) \frac{\theta - r}{q} + \lambda + \tau \\ 
\end{align*}
\]

At first, consider \( \theta \) s.t. \( q'(\theta) > 0 \). Then \( \tau = 0 \). Integrating equation (18) we get: \( \alpha \frac{r^2}{2} + \delta = kq \) where \( k \) is a constant of integration. Then (19) implies that \( \lambda = -k(\theta - r) \).

Substituting this in (17) we obtain \( q = 2\theta - 1 - k \), and so \( \delta(\theta) = -\alpha \frac{r^2(\theta)}{2} + k(2\theta - 1 - k) \), \( r(\theta) = \frac{\theta^2 + c}{\theta} \) where \( c \) is another constant of integration. The constants \( k \) and \( c \) will be determined below.

Next, let us examine the possibility that \( q'(\theta) = 0 \). We will demonstrate that in Case 1 there does not exist an interval \([\theta_1, \theta_2]\) s.t. \( q'(\theta) = 0 \) \( \forall \theta \in (\theta_1, \theta_2) \), and \( q'(\theta) > 0 \) \( \forall \theta \in (\theta_1 - \epsilon_1, \theta_1) \cup (\theta_2, \theta_2 + \epsilon_2) \) for some \( \epsilon_1, \epsilon_2 > 0 \).

The proof is by contradiction. So, suppose otherwise. Inspecting the first-order conditions it is easy to see that \( \tau(\theta) \) is continuous on \([0, 1]\). Therefore, \( \tau(\theta_1) = \tau(\theta_2) = 0 \). By (18), \( \delta'(\theta) = 0 \) \( \forall \theta \in [\theta_1, \theta_2] \). Integrating (17) with respect to \( \theta \) and substituting the result into (19) we obtain that \( \forall \theta \in [\theta_1, \theta_2] \)

\[
\tau(\theta) - \tau(\theta_1) = (\theta - \theta_1)(\theta + \theta_1 - 1 - q - \alpha r^2/2q - \delta/q)
\]

Since \( q + \alpha r^2/2q + \delta/q \) is constant on \([\theta_1, \theta_2]\), \( \tau(\theta_1) = 0 \) and \( \tau(\theta) \geq 0 \), we have \( 2\theta_1 \geq 1 + q + \alpha r^2/2q + \delta/q \). So, \( \tau(\theta_2) > \tau(\theta_1) \). Contradiction.

**Case 2:** \( q(\theta) < r(\theta) \). The necessary first-order conditions are:

\[
\begin{align*}
-\lambda'(\theta) &= \frac{\partial H}{\partial q} = (2\theta - 1) - q - \alpha \frac{q'(\theta - r)}{2} - \frac{\delta q'(\theta - r)}{q^2} \quad (20) \\
-\delta'(\theta) &= \frac{\partial H}{\partial r} = \alpha q'(\theta + q/2 - 2r) - \frac{\delta q'}{q} \\
0 &= \frac{\partial L}{\partial q'} = \alpha(r - q/2)(\theta - r) + \delta \frac{\theta - r}{q} + \lambda + \tau \\ 
\end{align*}
\]

First, consider \( \theta \) s.t. \( q'(\theta) > 0 \). Combining (20) and (21) we get:

\[
- \left( \lambda(\theta) + \delta(\theta) \frac{\theta - r}{q} \right)' = - \alpha \frac{q'(\theta - r)}{2} + 2\theta - 1 - q + \alpha q'(\theta + q/2 - 2r) \frac{\theta - r}{q} - \frac{\delta(\theta)}{q} \quad (23)
\]

On the other hand, differentiating (22) we get:

\[
- \left( \lambda(\theta) + \delta(\theta) \frac{\theta - r}{q} \right)' = \alpha \left( r - q/2 - \frac{q'(\theta - r)}{2} + q'(\theta + q/2 - 2r) \frac{\theta - r}{q} \right) \quad (24)
\]
Equating the right-hand sides of (23) and (24) we get:
\[ \delta = q(2\theta - 1 - q - \alpha(r - q/2)) \]  
(25)
Differentiating (25) we obtain:
\[ \delta' = q'(2\theta - 1 - q - \alpha(r - q/2)) + q\left(2 - q' - \alpha\frac{q'\theta - r}{q} + \alpha q'/2\right) = 2q + q'(2\theta - 1 - 2q - \alpha(\theta - q)) \]  
(26)
Finally, equate the right-hand sides of (21) and (26) and use (25) to substitute \( \delta(\theta) \) out to get:
\[ q'(\alpha r + (1-\alpha)q) = 2q \]  
(27)
Thus, we have a system of differential equations consisting of (27) and ‘the law of motion’:
\[ r' = \frac{q'(\theta - r)}{q} \]  
(28)
This system has multiple solutions. We return to it below once we determine the appropriate boundary conditions for it.

Next, consider the possibility that \( q'(\theta) = 0 \). Using a proof similar to the one in Case 1, we can show that in Case 2 there does not exist an interval \([\theta_1, \theta_2] \) s.t. \( q'(\theta) = 0 \) \( \forall \theta \in (\theta_1, \theta_2) \), and \( q'(\theta) > 0 \) \( \forall \theta \in (\theta_1 - \epsilon_1, \theta_1) \cup (\theta_2, \theta_2 + \epsilon_2) \) for some \( \epsilon_1, \epsilon_2 > 0 \).

**Case 3:** \( q(\theta) = r(\theta) \). The necessary first-order conditions are as follows:
\[ -\chi'(\theta) = \frac{\partial H}{\partial q} = 2\theta - 1 - q - \alpha\frac{q'\theta - r}{2} - \delta \frac{q'\theta - r}{q^2} \]  
(29)
\[ -\delta'(\theta) = \frac{\partial H}{\partial r} = -q'\left(\alpha\theta - 3/2\alpha r - \frac{\delta}{q}\right) \]  
(30)
\[ 0 = \frac{\partial L}{\partial q'} = \alpha \frac{r(\theta - r)}{2} + \lambda + \delta \frac{\theta - r}{q} + \tau \]  
(31)
Let us now consider whether there may exist an interval \([\theta_1, \theta_2] \) s.t. \( q(\theta) = r(\theta) \) \( \forall \theta \in [\theta_1, \theta_2] \). Clearly, we must also have \( q'(\theta) = r'(\theta) \) on this interval. Depending of the value of \( q'(\theta) \) we will distinguish two subcases.

**Subcase (3a):** \( q'(\theta) = r'(\theta) \neq 0 \) a.e. on \([\theta_1, \theta_2] \). Then (11) implies that \( r(\theta) = q(\theta) = \theta/2 \), so that \( q'(\theta) = r'(\theta) = 1/2 \) and \( \tau = 0 \). Substituting these in (29)-(31) we obtain the following system of three equations:
\[ \chi'(\theta) = 1 + \alpha \theta/8 + \delta/\theta \]
\[ \delta'(\theta) = \alpha \theta/8 - \delta/\theta \]
\[ \alpha \theta^2/8 + \lambda + \delta = 0 \]

\[ ^{20}\text{We can derive 27 in an alternative way as an Euler equation of the calculus of variation problem} \]
\[ \max \int_0^1 F(U, q, q', \theta) d\theta \text{ where } F(U, q, q', \theta) = (2\theta - 1)q - q^2 + \alpha \left(\theta - \frac{\theta}{\theta} \min\{\theta - \frac{\theta}{\theta}, q\} - \left(\frac{\min(\theta - \frac{\theta}{\theta}, q)^2}{2}\right) \right) \frac{\theta}{\theta}. \]
Since \( U' = q \), the Euler equation for this problem is \( F_U \frac{dU}{d\theta} + \frac{dF}{q} = 0 \). Note that \( F(U, q, q', \theta) \) is equal to the integrand of the optimal control problem (12).
Differentiating the last equation and using the first and the second equations we obtain the equality \( a\theta^2/2 + 1 = 0 \) which does not hold.

Subcase (3b): \( q'(\theta) = r'(\theta) = 0 \) a.e. on \([1, \theta_2]\).

Note that such interval \([1, \theta_2]\) cannot exist if \(q(\theta)\) is strictly increasing on \((\theta_1 - \epsilon_1, \theta_1)\) and on \((\theta_2, \theta_2 + \epsilon_2)\) for some \(\epsilon_1, \epsilon_2 > 0\). The proof is identical to the one provided in Case 1.

Let us summarize our results so far. First, we have established that if \(q(\theta)\) is strictly increasing on some interval, then a.e. in this interval either \(q(\theta) > r(\theta)\) (Case 1) or \(q(\theta) < r(\theta)\) (Case 2). Second, we have shown that there does not exist an interval \([1, \theta_2]\) s.t. \(q'(\theta) = 0 \forall \theta \in (\theta_1, \theta_2)\), and \(q'(\theta) > 0 \forall \theta \in (\theta_1 - \epsilon_1, \theta_1) \cup (\theta_2, \theta_2 + \epsilon_2)\) for some \(\epsilon_1, \epsilon_2 > 0\). Therefore, if the quantity schedule \(q(.)\) is ‘flat’ (i.e. \(q'(.) = 0\)) on some interval, then this interval must include one of the end-points 0 or 1.

To proceed further, we need to determine more precisely when the solution is in Case 1 and Case 2. This is done in the following lemma which also establishes the ‘no exclusion’ result. Recall that \(\theta = \sup \{\theta | q(\theta) = 0\}\). We have:

**Lemma 8** A solution to problem (12) has the following properties: (i) \(\theta = 0\) or \(\theta = 1\) s.t. \(q(\theta) < r(\theta)\) \forall \theta \in (0, 1]\) (Case 2), and \(q(\theta) > r(\theta)\) \forall \theta \in [1, \theta_2]\) (Case 1).

**Proof of lemma 8:** The lemma will be proved in several steps.

**Step 1.** It is not optimal to set \(q(\theta) = 0\) \forall \theta \in [0, 1]. For if otherwise, then the value of the optimal control problem in (12) is 0. However, its value is strictly positive if the firm sets \(q(\theta) = g(\theta) = 1/2 \forall \theta \in [1/2, 1]\).

**Step 2.** A ‘flat’ interval cannot include 1, i.e. if the quantity schedule \(q(.)\) is constant on \([\theta^f, 1]\) for some \(\theta^f < 1\), then \(q(.)\) is not optimal.

Suppose otherwise. Since \(q(1) = 1\), then \(q(\theta) = 1 \forall \theta \in [\theta^f, 1]\), and so \(r(.)\) is constant on this interval. Since \(q(.)\) is continuous and \(q(0) = 0\), \(\exists \theta^d s.t. q(\theta^d) = \theta^d\) and \(q(\theta) > r \forall \theta \in (\theta^d, 1)\).

Let us show that the value of the objective can be strictly increased by replacing the quantity schedule \(q(.)\) with modified quantity schedule \(q^m(.)\) s.t. \(q^m(\theta) = q(\theta)\) for \(\theta \in [0, \theta^d]\), and \(q^m(\theta) = \theta\) for \(\theta \in [\theta^d, 1]\).

For this, we need the following: **Result 1:** \(r(\theta) \leq \theta \forall \theta\).

Suppose otherwise, i.e. \(\exists \theta_1 s.t. r(\theta_1) > \theta_1\). Since \(r(0) = 0\), there exists \(\theta_2 \in (0, \theta_1)\) s.t. \(r(\theta_2) = \frac{r(\theta_1) + \theta_1}{2}\). Note that \(\theta_2 < r(\theta_2)\), and so \(q(\theta_2) > 0\). Therefore, \(r'(\theta) = \frac{q'(\theta)(\theta - r(\theta))}{q(\theta)} < 0 \forall \theta \in [\theta_2, \theta_1]\) s.t. \(r(\theta) > \theta\). This implies that \(r(\theta_1) \leq \frac{r(\theta_1) + \theta_1}{2} < r(\theta_1)\). Contradiction.

Consider the change in the value of the objective (12) after we modify \(q(.)\) on the interval \([\theta^d, 1]\). Result 1 implies that \(\forall \theta \in (\theta^d, 1]\), the solution is in Case 1 under both the original and the modified quantity schedule. It is easy to see that the sum of the second and the third term in (12) does not change after this modification, while the first term changes by \(\int_1^\theta (2\theta - 1)(q^m(\theta) - q(\theta)) - (\theta^d(\theta - r(\theta)) - \theta^2 - q(\theta)^2)d\theta\), which is strictly positive because \(2\theta - 1 < \theta < q(\theta) \forall \theta \in (\theta^d, 1)\).

**Step 3.** Suppose that \(q(\theta_1) > r(\theta_1)\) for some \(\theta_1 \in (0, 1)\). Then \(\exists \theta_2 s.t. q(\theta_2) < r(\theta_2)\).

The proof is by contradiction, so assume otherwise. Let \(\theta = \inf \{\theta | \theta \in [0, 1], q(\theta) > r(\theta)\}\).
If \( \overline{\theta} > 0 \), then \( q(\theta) = r(\theta) \forall \theta \in [0, \overline{\theta}] \), i.e., the solution is in Case 3. We have shown above that this may only be possible if \( q'(\theta) = r'(\theta) = 0 \forall \theta \in [0, \overline{\theta}] \). Since \( q(0) = 0 \), this implies that \( q(\theta) = 0 \) and therefore \( r(\theta) = \theta \forall \theta \in [0, \overline{\theta}] \). Contradiction.

If \( \overline{\theta} = 0 \), then \( q(\theta) > r(\theta) \forall \theta \in [0, \overline{\theta}] \). Then the analysis of Case 1 shows that on this interval \( q(\theta) = 2\theta \). By result 1, \( r(\theta) \leq \theta \forall \theta \in [0, 1] \). So, the solution cannot enter the region corresponding to Case 2 where \( q(\theta) < r(\theta) \). Then the terminal condition \( q(1) = 1 \) can only be attained if \( q(\theta) = 2\theta \forall \theta \in [0, 1/2] \) and \( q(\theta) = 1 \forall \theta \in [1/2, 1] \). But by Step 2 this is not optimal.

**Step 4.** Suppose that \( q(\theta_1) > r(\theta_1) \) for some \( \theta_1 \in (0, 1) \). Then \( q(\theta) > r(\theta) \forall \theta \in (\theta_1, 1] \).

Recall that \( \overline{\theta} = \sup \{ \theta | q(\theta) = 0 \} \). By step 3, \( \exists \overline{\theta} < \theta_1 \) s.t. \( q(\theta) < r(\theta) \forall \theta \in (\overline{\theta}, \theta_1) \) - the solution is in Case 2, and \( q(\theta) > r(\theta) \forall \theta \in \left( \overline{\theta}, \theta_1 \right] \) - the solution is in Case 1.

By continuity of the optimal quantity schedule \( q(\cdot), q(\overline{\theta}) = r(\overline{\theta}) \), and \( q'(\overline{\theta}) > r'(\overline{\theta}) = q'(\overline{\theta}) \frac{\overline{\theta} - r(\overline{\theta})}{q(\overline{\theta})} \). Thus, \( \overline{\theta} < r(\overline{\theta}) + q(\overline{\theta}) \). Recall from the analysis of Case 1 that \( q'(\theta) = 2 \) when \( q(\theta) > r(\theta) \). Hence, \( \theta < r(\theta) + q(\theta) \) and \( q'(\theta) > r'(\theta) \forall \theta \in (\overline{\theta}, \theta_1) \). So, the solution cannot switch to Case 1 at \( \overline{\theta} \), implying that \( \overline{\theta} = 1 \).

**Summary of Steps 1-4.** \( \exists \overline{\theta}, \overline{\theta} < 0 \) s.t. \( q(\theta) = 0 \forall \theta \in [0, \overline{\theta}], q(\theta) < r(\theta) \forall \theta \in (\overline{\theta}, 1] \), and \( q(\theta) > r(\theta) \forall \theta \in (\overline{\theta}, 1] \). Note that we have not yet ruled out the case that \( \overline{\theta} = 1 \).

**Step 5.** No exclusion: \( \overline{\theta} = 0 \).

On \( (\overline{\theta}, \overline{\theta}) \), the solution is in Case 2, so \( q(\cdot) \) and \( r(\cdot) \) are determined by the pair of differential equations (27) and (28) with initial conditions \( q(\overline{\theta}) = 0 \) and \( r(\overline{\theta}) = \overline{\theta} \). Let us focus on (27). Since \( q(\cdot) \) and \( r(\cdot) \) are nonnegative and nondecreasing in \( \theta, \forall \theta \in (0, \overline{\theta}) \):

\[
q'(\theta) = \frac{2q(\theta)}{\alpha r(\theta) + (1 - \alpha)q(\theta)} \leq \frac{2q(\theta)}{\alpha \overline{\theta}}
\]

Pick \( \theta \in (\overline{\theta}, \overline{\theta} + \frac{\alpha \overline{\theta}}{2}) \). Integrating, we get \( q(\theta) \leq \frac{2}{\alpha \overline{\theta}} \int_{\overline{\theta}}^{\theta} q(s)ds \). The latter inequality implies that \( q(\theta) = 0 \), contradicting the definition of \( \overline{\theta} = \sup \{ \theta | q(\theta) = 0 \} \).

**Step 6.** \( \overline{\theta} < 1 \).

We will demonstrate that any solution to (27) and (28) with initial conditions \( r(0) = q(0) = 0 \) is such that \( \lim_{\theta \to 1} r(\theta) < 1 \), which in combination with \( q(1) = 1 \) implies that the optimal schedule must switch to Case 1 at some point in \( (0, 1) \).

Suppose otherwise, i.e. \( \lim_{\theta \to 1} r(\theta) = 1 \). Note that by **Result 1**, \( r(\theta) \leq \theta \forall \theta \in [0, 1] \). Choose \( \theta^* \) s.t. \( r(\theta^*) \geq 3/4 \) and \( q(\theta^*) \geq 3/4 \). Combining (27) and (28), we obtain that \( r'(\theta) \leq 2/3 \forall \theta \in [\theta^*, 1] \). So, \( r(1) \leq r(\theta^*) + 2/3(1 - \theta^*) \leq \theta^* + 2/3(1 - \theta^*) < 1 \). \( Q.E.D. \)

Note the following immediate implication of Lemma 8. Since \( q(1) = 1 \), continuity requires that on the interval \( [\overline{\theta}, 1] \), \( q(\theta) = 2\theta - 1 \). So, \( q(\overline{\theta}) = r(\overline{\theta}) = 2\overline{\theta} - 1 \).

Finally, to find \( \overline{\theta} \) and determine the solution on the interval \( (0, \overline{\theta}) \), we need the following lemma.

**Lemma 9** There exists at most one solution \((q(\cdot), r(\cdot))\) to the system of differential equations (27) and (28) which satisfies the following conditions: (i) \( q(0) = r(0) = 0 \) (ii) \( \exists \overline{\theta} \in [1/2, 1] \) such that \( q(\overline{\theta}) = r(\overline{\theta}) = 2\overline{\theta} - 1 \), and \( q(\theta) < r(\theta) \forall \theta \in (0, \overline{\theta}) \).
Proof: Suppose that there exist two different solutions \((q_1(.), r_1(.))\) and \((q_2(.), r_2(.))\) satisfying the conditions on the lemma. Let \(q(\theta) = r(\theta) = 2\theta_i - 1\) for \(i \in \{1, 2\}\), and assume without loss of generality that \(\theta_2 > \theta_1\). We will establish a contradiction in a number of steps.

**Step 1.** \(q_2(\theta_1) > q_1(\theta_1)\).

Suppose otherwise, i.e. \(q_2(\theta_1) \leq q_1(\theta_1) = 2\theta_1 - 1\). By (27), \(q_2'(\theta) < 2 \forall \theta \in (\theta_1, \theta_2)\). So, \(q_2(\theta_2) - q_2(\theta_1) < 2(\theta_2 - \theta_1)\) which contradicts the fact that \(q(\theta_2) = 2\theta_2 - 1\).

**Step 2.** \(q_2(\theta) > q_1(\theta)\) and \(r_2(\theta) > r_1(\theta) \forall \theta \in (0, \theta_1)\).

We know that \(r_2(\theta_1) > q_2(\theta_1) > q_1(\theta_1) = r_1(\theta_1)\). Let us choose any \(\theta_3 \in (0, \theta_1)\) s.t. \(q_2(\theta_3) > q_1(\theta)\) and \(r_2(\theta) > r_1(\theta) \forall \theta \in (\theta_3, \theta_1)\). Such \(\theta_3\) exists by continuity of the solutions to (27) and (28). To prove this step, it is sufficient to show that \(q_2(\theta_3) > q_1(\theta_3)\) and \(r_2(\theta_3) > r_1(\theta_3)\).

By continuity, we only need to rule out \(q_2(\theta) = q_1(\theta)\) and \(r_2(\theta) = r_1(\theta)\) for \(\theta \in (0, \theta_1)\), because this would imply that \(q_2(\theta) = q_1(\theta)\) and \(r_2(\theta) = r_1(\theta)\) for \(\theta \in (0, \theta_1)\).

So, suppose that \(q_2(\theta_3) = q_1(\theta_3)\). We cannot have \(r_2(\theta_3) = r_1(\theta_3)\) at the same time, because otherwise we would have \(q(\theta) = q_2(\theta)\) and \(r_1(\theta) = r_2(\theta) \forall \theta \in (0, \theta_1)\). So, \(r_2(\theta_3) > r_1(\theta_3)\).

Then, by (27) \(q_1'(\theta_3) > q_2'(\theta_3)\), and therefore \(q_1(\theta) > q_2(\theta) \forall \theta \in (\theta_3, \theta_3 + \epsilon_3)\) for some \(\epsilon_3 > 0\). Contradiction.

Now suppose that \(r_2(\theta_3) = r_1(\theta_3)\). Again, we can rule out \(q_2(\theta_3) = q_1(\theta_3)\), because otherwise the two solutions would be identical. So, \(q_2(\theta_3) > q_1(\theta_3)\). Combining (27) and (28) we obtain:

\[
r' = \frac{2(\theta - r)}{ar + (1 - \alpha)q}
\]

By (32), \(r_1'(\theta_3) > r_2'(\theta_3)\), and therefore \(r_1(\theta) > r_2(\theta) \forall \theta \in (\theta_3, \theta_3 + \epsilon_4)\) for some \(\epsilon_4 > 0\). Contradiction.

**Step 3.** If \(\lim_{\theta \to 0} \theta_1(\theta) = 0\), then \(\lim_{\theta \to 0} \theta_2(\theta) > 0\).

Step 2 and equation (32) imply that \(r_2'(\theta) < r_1'(\theta) \forall \theta \in (0, \theta_1)\). Since \(r_2(\theta_1) > r_1(\theta_1)\), Step 3 follows.

Thus, we have shown that two different pairs \((q_1(.), r_1(.))\) and \((q_2(.), r_2(.))\) cannot satisfy the conditions of the lemma.

Finally, let us derive the solution to (27) and (28) that satisfies the conditions of Lemma 9. We conjecture that \(r(.\)) is a linear function of \(\theta\) and \(q(.\)), i.e. \(r(\theta) = a\theta + bq(\theta)\) for some constants \(a, b\). Then \(r'(\theta) = a + bq'(\theta)\). Substituting these into (27) and (28) and rearranging we obtain:

\[
\theta(aa^2 + 2a - 2) = -q(\theta)(4b + a(1 - \alpha + ab))
\]

If \(q(\theta)\) is a linear function of \(\theta\), then from (27) and (28) we obtain the following solution:

\[
q(\theta) = \theta^{3 + a^2} \quad \text{and} \quad r(\theta) = \theta^{2/3}. \quad \text{So,} \quad q(\theta) > r(\theta), \quad \text{and we can rule out this solution.}
\]

Thus, suppose that (33) holds as an identity, i.e.

\[
\alpha a^2 + 2a - 2 = 0 \quad \text{and} \quad (1 - \alpha)a + 4b + ab\alpha = 0.
\]
Solving for the coefficients $a$ and $b$ yields:

$$a = \frac{-1 \pm \sqrt{1 + 2\alpha}}{\alpha}$$

$$b = \frac{1 - \alpha}{4 + \alpha a}$$

We will choose the positive root for $a$, so that $b = \frac{1 - \alpha}{\sqrt{1 + 2\alpha}}$. By computation we can show that $a < 1$ and $a + b < 1 \forall \alpha > 0$, so $r(\theta) = a\theta + bq(\theta) < \theta$.

Let $u(\theta) = \frac{q(\theta)}{\theta}$. Then using (27) we obtain:

$$q' = \theta u' + u = \frac{2q}{\alpha r + (1 - \alpha)q} = \frac{2u}{\alpha a + (1 - \alpha + \alpha b)u}$$

Rearranging, we obtain:

$$\frac{d\theta}{\theta} = \frac{du}{u} \left( \frac{c_0 + c_1 u}{c_2 - c_1 u} \right), \quad (34)$$

where $c_0 = \alpha a = \sqrt{1 + 2\alpha} - 1$, $c_1 = 1 - \alpha + \alpha b = \frac{4(1 - \alpha)}{3 + \sqrt{1 + 2\alpha}}$, and $c_2 = 2 - \alpha a = 3 - \sqrt{1 + 2\alpha}$.

Using the change of variable $v = \ln u_2$, we can rewrite (34) as:

$$\frac{d\theta}{\theta} = dv \frac{c_0 + c_1 e^v}{c_2 - c_1 e^v}.$$  

When $\alpha \neq 4$ so that $c_2 \neq 0$, we can integrate both sides of this equation to get:

$$\ln \theta = c' + \frac{c_0}{c_2} v - \frac{c_0 + c_2}{c_2} \ln(|c_2 - c_1 e^v|),$$

where $c'$ is a constant of integration. Exponentiating both sides and using $c_0 + c_2 = 2$ we get:

$$\theta = e^{c'} u^{c_0/c_2} (|c_2 - c_1 u|)^{-\frac{2}{c_2}},$$

Substituting $u(\theta) = \frac{q(\theta)}{\theta}$, and simplifying we get:

$$((2 - \alpha/2)\theta - (1 - \alpha)q)^2 = k(\alpha)q^{\sqrt{1+2\alpha}-1} \quad (35)$$

where $k(\alpha) = \frac{(3+\sqrt{1+2\alpha})^2}{16}e^{c'(3-\sqrt{1+2\alpha})}$.

When $\alpha = 4$, (34) can be rewritten as

$$\frac{d\theta}{\theta} = \frac{du}{u} \left( \frac{1 - u}{u} \right)$$

This equation can be solved directly to yield:

$$\theta = k(4)q - q \ln(q) \quad (36)$$
Both (35) and (36) characterize \( q(\cdot) \) as an implicit function of \( \theta \) for a given value of \( \alpha \). We will check that these functions can be inverted to obtain a solution to our problem. From now on, we will explicitly recognize the dependence of the solution on the parameter \( \alpha \). In particular, we will use the notation \( q(\theta, \alpha) \), \( r(\theta, \alpha) \) and \( \overline{\theta}(\alpha) \) to describe the optimal schedules and the switch point for a given value of \( \alpha \). Still, we will sometimes omit the argument \( \alpha \) when there is no risk of confusion.

To compute the switch point \( \overline{\theta}(\alpha) \) note that it has to satisfy \( q(\overline{\theta}(\alpha), \alpha) = r(\overline{\theta}(\alpha), \alpha) = 2\overline{\theta}(\alpha) - 1 \). Since \( r(\theta, \alpha) = a\theta + bq(\theta, \alpha) \), we have:

\[
r(\overline{\theta}(\alpha), \alpha) = \frac{\sqrt{1+2\alpha} - 1}{\alpha} \overline{\theta}(\alpha) - \frac{1 - \alpha}{3} \frac{\sqrt{1+2\alpha} - 1}{\alpha} (2\overline{\theta}(\alpha) - 1) = 2\overline{\theta}(\alpha) - 1
\]

Solving, we get \( \overline{\theta}(\alpha) = \frac{2}{3} + \frac{1}{3(\sqrt{1+2\alpha}+1)} \), and so \( q(\overline{\theta}(\alpha), \alpha) = r(\overline{\theta}(\alpha), \alpha) = \frac{1}{3} + \frac{2}{3(\sqrt{1+2\alpha}+1)} \).

Note that \( \overline{\theta}(\alpha) \) is decreasing in \( \alpha \). It converges to 2/3 as \( \alpha \) increases to infinity (almost all consumers are ‘honest’), and converges to 5/6 as \( \alpha \) decreases to 0 (almost all consumers are strategic).

To specify \( k(\alpha) \), substitute the derived values of \( \overline{\theta}(\alpha), q(\overline{\theta}(\alpha)) \) into (35). Then for \( \alpha \neq 4 \):

\[
(2 - \alpha/2) \left( \frac{2}{3} + \frac{1}{3(\sqrt{1+2\alpha}+1)} \right) - (1 - \alpha) \left( \frac{1}{3} + \frac{2}{3(\sqrt{1+2\alpha}+1)} \right) = k(\alpha) \left( \frac{1}{3} + \frac{2}{3(\sqrt{1+2\alpha}+1)} \right)^{1/2} = \frac{1 + \frac{2}{2(\sqrt{1+2\alpha}+1)}}{\left( \frac{1}{3} + \frac{2}{3(\sqrt{1+2\alpha}+1)} \right)^{1/2}}
\]

Solving this equation, we obtain:

\[
k(\alpha) = \frac{1}{\sqrt{1 + 2\alpha} - 1 - 1/2}
\]

Note that \( k(\alpha) \) is increasing in \( \alpha \). For \( \alpha = 4 \), we substitute \( \overline{\theta}(4) = 3/4 \), \( q(\overline{\theta}(4)) = 2\overline{\theta}(4) - 1 = 1/2 \) into (36) to get

\[
k(4) = 3/2 + \ln(1/2)
\]

Next, let us show that (35) and (36) can be inverted and so we can compute \( q(\theta, \alpha) \) \( \forall \theta \in [0, \overline{\theta}(\alpha)] \). Start with \( \alpha \neq 4 \). We claim that (35) implies that \( (2 - \alpha/2) \theta - (1 - \alpha)q(\theta, \alpha) > 0 \). Clearly, this is true for \( \alpha \leq 1 \) since \( q(\theta, \alpha) \leq \theta \). Further, \( (2 - \alpha/2)\overline{\theta}(\alpha) - (1 - \alpha)q(\overline{\theta}(\alpha), \alpha) = 1 + \frac{\alpha}{2(\sqrt{1+2\alpha}+1)} > 0 \). So, if \( \exists \theta_1 > 0 \) s.t. \( (2 - \alpha/2)\theta_1 - (1 - \alpha)q(\theta_1, \alpha) < 0 \), then by continuity \( \exists \theta_2 \in (\theta_1, \overline{\theta}(\alpha)) \) satisfying \( (2 - \alpha/2)\theta_2 - (1 - \alpha)q(\theta_2, \alpha) = 0 \). But since \( q(\theta_2, \alpha) > 0 \), (35) cannot hold at \( \theta_2 \). Thus, \( q(\theta, \alpha) \) is a solution to:

\[
\theta = \frac{(1 - \alpha)q + (1 + \frac{\alpha}{2(\sqrt{1+2\alpha}+1)})}{(2 - \alpha/2) \left( \frac{1}{3} + \frac{2}{3(\sqrt{1+2\alpha}+1)} \right)^{1/2}}
\]

Since (37) may have multiple roots, we need to establish that \( q(\alpha, \theta) \) is well-defined. This will be done by establishing that on the interval \( [0, \frac{2}{3} + \frac{1}{3(\sqrt{1+2\alpha}+1)}] \) (37) admits a unique
Proof of Corollary 1: 

Part (i). 

Observe that 

To show that 

After the substitution, we get: 

To substitute 

Finally, let us derive 

Since (37) is obtained as a solution to (27), 

is strictly concave in 

whenever 

(r(q, α) − q > 0, and, in particular, on the interval 

Therefore, 

is strictly increasing on 

By a similar argument, we can derive 

g(θ, 4) as a solution to (36) and show that it is continuous, increasing and convex on 

Finally, let us derive 

Since 

is strictly increasing 

whenever 

(r(q, α) − q > 0, and, in particular, on the interval 

Therefore, 

is strictly increasing, concave and continuous on 

By a similar argument, we can derive 

g(θ, 4) as a solution to (36) and show that it is continuous, increasing and convex on 

To substitute 

out in this expression in the case 

To substitute 

out in this expression in the case 

After the substitution, we get: 

To show that 

is well-defined by (38) and is continuous, increasing and strictly convex on 

we can use an argument similar to the one which we used to establish that these properties hold for 

We can compute 

g(θ, 4) in a similar way. Q.E.D.

Proof of Corollary 1: Fix some 

and 

s.t. 

Part (i). Instead of 

it is more convenient to operate with its inverse - the function 

given by (37). Since 

we need to show that 

when 

is sufficiently small, and that there exists a unique point of intersection 

s.t. 

Step 1. 

exists 

s.t. 

and 

∀q ∈ (0, q).
When both $\alpha_1 \neq 4$ and $\alpha_2 \neq 4$, then by (37) $\theta(q, \alpha_1) < \theta(q, \alpha_2)$ iff

$$\frac{(1 - \alpha_1)}{2 - \alpha_1/2} q + \frac{k(\alpha_1)^{1/2}}{2 - \alpha_1/2} \frac{\sqrt{1 + \alpha_1 - 1}}{2} < \frac{(1 - \alpha_2)}{2 - \alpha_2/2} q + \frac{k(\alpha_2)^{1/2}}{2 - \alpha_2/2} \frac{\sqrt{1 + \alpha_2 - 1}}{2}$$

Dividing both sides of this inequality by $\frac{\sqrt{1 + \alpha_2 - 1}}{2}$ and rearranging we obtain an equivalent inequality:

$$\frac{k(\alpha_1)^{1/2}}{2 - \alpha_1/2} q \sqrt{1 + \alpha_1 - 1} < \left( \frac{(1 - \alpha_2)}{2 - \alpha_2/2} - \frac{(1 - \alpha_1)}{2 - \alpha_1/2} \right) q + \frac{k(\alpha_2)^{1/2}}{2 - \alpha_2/2} \frac{\sqrt{1 + \alpha_2 - 1}}{2}$$

Let $q$ go to zero. Then the left-hand of the above inequality converges to zero. If $\alpha_2 > 2$, then the first term on the right-hand side converges to plus infinity because $\frac{(1 - \alpha_2)}{2 - \alpha_2/2} > \frac{(1 - \alpha_1)}{2 - \alpha_1/2}$, while the second term is bounded. If $\alpha_2 < 2$, then the first term on the right-hand side converges to zero, while the second remains is a positive and constant. So, the inequality holds when $q$ is sufficiently small.

If $\alpha_2 = 2$, then we need to show that

$$\frac{(1 - \alpha_1)}{2 - \alpha_1/2} q + \frac{k(\alpha_1)^{1/2}}{2 - \alpha_1/2} \sqrt{1 + \alpha_1 - 1} \leq q \left( \frac{3}{2} + \log(1/2) - \log(q) \right)$$

It is easy to see that this inequality holds for small $q$ after we divide both sides of it by $q$, and then let $q$ go to zero.

If $\alpha_1 = 2$, then we need to show that

$$q \left( \frac{3}{2} + \log(1/2) - \log(q) \right) < \frac{(1 - \alpha_2)}{2 - \alpha_2/2} q + \frac{k(\alpha_2)^{1/2}}{2 - \alpha_2/2} \frac{\sqrt{1 + \alpha_2 - 1}}{2}$$

To see that this inequality holds for small $q$, we divide both sides of it by $q \frac{\sqrt{1 + \alpha_2 - 1}}{2}$, and let $q$ go to zero.

**Step 2. Existence of an intersection point.** $\exists q_i \in [0, \frac{1}{3} + \frac{2}{(1 + 2\alpha_1 + 1)^2}]$, $\epsilon_1 > 0$ and $\epsilon_2 > 0$ s.t. $\theta(q_1, \alpha_1) = \theta(q_2, \alpha_2)$, $\theta(q_1, \alpha_1) > \theta(q_2, \alpha_2)$ $\forall q \in (q_1, q_1 + \epsilon_1)$, and $\theta(q_1, \alpha_1) < \theta(q_2, \alpha_2)$ $\forall q \in (q_1, q_1 - \epsilon_2)$.

The existence of an intersection point $q_i$ follows from Step 1 and from the fact that

$$\theta(\frac{1}{3} + \frac{2}{(1 + 2\alpha_1 + 1)^2}, \alpha_1) = \frac{2}{3(1 + 2\alpha_1 + 1)} > \theta(\frac{1}{3} + \frac{2}{(1 + 2\alpha_1 + 1)^2}, \alpha_2).$$

**Step 3.** If $\theta_{qq}(q_1, \alpha_1) \leq \theta_{qq}(q_2, \alpha_2)$, then $\theta_{qq}(q_2, \alpha_1) < \theta_{qq}(q_2, \alpha_2)$ $\forall q > q_1$, where $\theta_{qq}(q, \alpha)$ is the second derivative of $\theta(q, \alpha)$ with respect to $q$. This step follows by simple computation.

**Step 4.** Uniqueness of an intersection point on $[0, \frac{1}{3} + \frac{2}{(1 + 2\alpha_1 + 1)^2}]$.

Let $q_r$ be the smallest $q$ s.t. $\theta(q_r, \alpha_1) = \theta(q_r, \alpha_2)$. Then by Step 1 $\theta(q_r, \alpha_1) < \theta(q_r, \alpha_2)$ $\forall q \in (0, q_r)$, and so $\theta(q_r, \alpha_1) \geq \theta(q_r, \alpha_2)$. To finalize the proof, consider two cases.

**Case 1.** $\exists q_1 > 0$ s.t. $\theta(q_1, \alpha_1) > \theta(q_1, \alpha_2)$ $\forall q \in (q_1, q_1 + \epsilon_1)$. Then $\theta(q_1, \alpha_1) > \theta(q_1, \alpha_2)$. If $\exists q_r \in (q_1, \frac{1}{3} + \frac{2}{(1 + 2\alpha_1 + 1)^2})$ s.t. $\theta(q_r, \alpha_1) = \theta(q_r, \alpha_2)$, let us choose the smallest such $q_r$. So, $\theta(q, \alpha_1) > \theta(q, \alpha_2)$ for $q \in (q_r, q_r)$. Therefore, $\theta(q_r, \alpha_1) \leq \theta(q_r, \alpha_2)$. But since $\theta(q_r, \alpha_1) > \theta(q_r, \alpha_2)$
\(\theta_q(q_i, \alpha_2)\), we conclude that \(\exists q_m \in [q_i, q_r)\) s.t. \(\theta_{qq}(q_m, \alpha_1) \leq \theta_{qq}(q_m, \alpha_2)\). But then by Step 3 \(\theta_{qq}(q, \alpha_1) < \theta_{qq}(q, \alpha_2)\) \(\forall q > q_m\).

Consequently, \(\theta_q(q_1, \alpha_1) < \theta_q(q_2, \alpha_2)\) \(\forall q > q_r\), and so \(\theta(q, \alpha_1) < \theta(q, \alpha_2)\) \(\forall q > q_r\). But this contradicts the fact that \(\theta\left(\frac{1}{3} + \frac{2}{3(1+2q_1+1)}, \alpha_1\right) = \theta\left(\frac{1}{3} + \frac{2}{3(1+2q_1+1)}, \alpha_2\right)\).

Case 2. \(\exists \epsilon_1 > 0\) s.t. \(\theta(q, \alpha_1) \leq \theta(q_1, \alpha_2)\) \(\forall q \in (q_i, q_1 + \epsilon_1)\). Then \(\theta_q(q_1, \alpha_1) = \theta_q(q_1, \alpha_2)\) and \(\theta_{qq}(q_1, \alpha_1) < \theta_{qq}(q_1, \alpha_2)\), and so by Step 3, \(\theta_{qq}(q, \alpha_1) < \theta_{qq}(q, \alpha_2)\) \(\forall q \geq q_1\). Hence, \(\theta_q(q_1, \alpha_1) < \theta_q(q_2, \alpha_2)\) and, consequently, \(\theta(q, \alpha_1) < \theta(q, \alpha_2)\) \(\forall q \geq q_1\). Again, this contradicts the fact that \(\theta\left(\frac{1}{3} + \frac{2}{3(1+2q_1+1)}, \alpha_1\right) = \theta\left(\frac{1}{3} + \frac{2}{3(1+2q_1+1)}, \alpha_2\right)\).

Part (ii). Recall that \(U(\theta, \alpha) = \theta q(\theta, \alpha) - c(\theta, \alpha)\) is the total surplus of a ‘strategic’ consumer with valuation \(\theta\). We have established that \(U(\theta, \alpha) = \int_0^1 q(s, \alpha)ds\).

Since \(q(\theta, \alpha_1) > q(\theta, \alpha_2)\) \(\forall \theta \in (0, \theta_c(\alpha_1, \alpha_2)\), we also have \(U(\theta, \alpha_1) > U(\theta, \alpha_2)\) \(\forall \theta \in (0, \theta_c(\alpha_1, \alpha_2)\).

Let us show that \(U(\theta(\alpha_1, \alpha_2), \alpha_1) > U(\theta(\alpha_2, \alpha_2), \alpha_2)\). Note that \(\theta(\alpha_1) = \frac{2}{3} + \frac{1}{3(1+2q_1+1)}\) and \(\theta(\alpha_2) = \frac{1}{3} + \frac{2}{3(1+2q_1+1)}\). Combining these inequalities and invoking Lemma 9, we conclude that \(g\left(\frac{1}{3} + \frac{2}{3(1+2q_1+1)}, \alpha_1\right) = g\left(\frac{1}{3} + \frac{2}{3(1+2q_1+1)}, \alpha_2\right)\).

By definition, \(U(\theta(\alpha_1, \alpha_2), \alpha_1) = \left(\theta(\alpha_2) - \frac{1}{3} - \frac{2}{3(1+2q_1+1)}\right) \left(\frac{1}{3} + \frac{2}{3(1+2q_1+1)}\right)\). At the same time, by Lemma 8, \(U(\theta(\alpha_2, \alpha_2), \alpha_1) = \left(\theta(\alpha_2) - \frac{1}{3} - \frac{2}{3(1+2q_1+1)}\right) g\left(\frac{1}{3} + \frac{2}{3(1+2q_1+1)}, \alpha_1\right) = \left(\theta(\alpha_2) - \frac{1}{3} - \frac{2}{3(1+2q_1+1)}\right) \left(\frac{1}{3} + \frac{2}{3(1+2q_1+1)}\right)\). So, \(U(\theta(\alpha_1, \alpha_2), \alpha_1) > U(\theta(\alpha_2, \alpha_2), \alpha_2)\).

Further, \(U(\theta, \alpha_1) > U(\theta, \alpha_2)\) \(\forall \theta \in [\theta_c(\alpha_1, \alpha_2), \theta(\alpha_2)]\), because \(U_\theta(\theta(\alpha_1, \alpha_2), \alpha_1) = q(\theta, \alpha_1) < q(\theta, \alpha_2) = U_\theta(\theta(\alpha_2, \alpha_2), \alpha_2)\) on this interval. Finally, \(U(\theta, \alpha_1) > U(\theta, \alpha_2)\) \(\forall \theta \in [\theta(\alpha_2), 1]\) because \(U_\theta(\theta(\alpha_1, \alpha_2), \alpha_1) = q(\theta, \alpha_1) = 2\theta_1 - 1 = q(\theta, \alpha_2) = U_\theta(\theta(\alpha_2, \alpha_2), \alpha_2)\) on this interval.

Part (iii).

Step 1. \(\exists \theta_m > 0\) s.t. \(\forall \theta \in (0, \theta_m)\) \(g(\theta, \alpha_1) > g(\theta, \alpha_2)\).

Using an argument identical to the one in Step 1 of Part (i) in this proof, we can prove an equivalent result - if \(g\) is small enough, then \(\theta(g, \alpha_1) < \theta(g, \alpha_2)\) where \(\theta(g, \alpha)\) is given by (38).

Step 2. \(g(\theta, \alpha_1) > g(\theta, \alpha_2)\) \(\forall \theta \in (\frac{1}{3} + \frac{2}{3(1+2q_1+1)}, \frac{1}{3} + \frac{2}{3(1+2q_1+1)})\).

Note that \(\theta(\alpha_1) = \frac{2}{3} + \frac{1}{3(1+2q_1+1)}\) and \(\theta(\alpha_2) = \frac{1}{3} + \frac{2}{3(1+2q_1+1)}\). Therefore, \(g(\theta, \alpha_1) > \forall \theta \in (\frac{2}{3} + \frac{1}{3(1+2q_1+1)})\). At the same time, \(g(\theta, \alpha_2) < \forall \theta \in (\frac{1}{3} + \frac{2}{3(1+2q_1+1)})\) because \(r(\theta(\alpha_1, \alpha_2), \alpha_2) = \frac{1}{3} + \frac{2}{3(1+2q_1+1)}\).

Step 3. \(\exists \theta \in (0, \frac{1}{3} + \frac{2}{3(1+2q_1+1)})\) s.t. \(\theta(\alpha_1) \leq \theta(\alpha_2)\), then \(\exists \theta_1, \theta_2 \in (\theta, \frac{1}{3} + \frac{2}{3(1+2q_1+1)})\), \(\theta_1 \leq \theta_2\) and \(\delta > 0\), s.t. (i) \(g(\theta_1, \alpha_1) = g(\theta_1, \alpha_2)\) and \(g(\theta_1, \alpha_1) > g(\theta_1, \alpha_2)\) \(\forall \theta \in (\theta_1 - \delta, \theta_1)\); (ii) \(g(\theta_2, \alpha_1) = g(\theta_2, \alpha_2)\) and \(g(\theta_1, \alpha_1) > g(\theta_2, \alpha_2)\) \(\forall \theta \in (\theta_2, \theta_2 + \delta)\). Consequently, \(g(\theta(\alpha_1, \alpha_1) \leq g(\theta(\alpha_1, \alpha_2) \leq g(\theta(\theta_2, \alpha_2)\).

The proof of this step is obvious.

Step 4. \(g(\theta, \alpha) \geq \frac{g(\theta, \alpha)}{r(\theta, \alpha) - \delta} = \frac{g(\theta, \alpha)}{r(\theta, \alpha) - \delta} \forall \theta \in (0, \frac{1}{3} + \frac{2}{3(1+2q_1+1)})\).

To see this, differentiate \(g(\theta, \alpha) = \frac{q(\theta, \alpha)}{r(\theta, \alpha) - \delta}\) and use (28) to make a substitution.
Step 5. \( r^{-1}(\theta_1, \alpha_1) > r^{-1}(\theta_1, \alpha_2) \).

Since \( g(\theta_1, \alpha_1) = g(\theta_1, \alpha_2) \) and \( g(\theta_1, \alpha_1) \leq g(\theta_1, \alpha_2) \), Step 4 implies that \( r^{-1}(\theta_1, \alpha_1) \geq r^{-1}(\theta_1, \alpha_2) \).

This inequality must be strict, i.e. \( r^{-1}(\theta_1, \alpha_1) \neq r^{-1}(\theta_1, \alpha_2) \). To see this, rewrite \( r(\theta, \alpha) = \theta - \frac{U(\theta, \alpha)}{q(\theta, \alpha)} \) as \( \theta = r^{-1}(\theta, \alpha) - \frac{U(r^{-1}(\theta, \alpha), \alpha)}{q(r^{-1}(\theta, \alpha), \alpha)} \). Therefore, if \( r^{-1}(\theta_1, \alpha_1) = r^{-1}(\theta_1, \alpha_2) \), then we have:

\[
\frac{U(r^{-1}(\theta_1, \alpha_1), \alpha_1)}{q(r^{-1}(\theta_1, \alpha_1), \alpha_1)} = \frac{U(r^{-1}(\theta_1, \alpha_2), \alpha_2)}{q(r^{-1}(\theta_1, \alpha_2), \alpha_2)}.
\]

But \( q(r^{-1}(\theta_1, \alpha_1), \alpha_1) = q(\theta_1, \alpha_1) = q(\theta_1, \alpha_2) = q(r^{-1}(\theta_1, \alpha_2), \alpha_2) \), yet \( U(r^{-1}(\theta_1, \alpha_1), \alpha_1) > U(r^{-1}(\theta_1, \alpha_2), \alpha_2) \) as established above. Contradiction.

Step 6. \( r^{-1}(\theta_1, \alpha_2) > \theta_c(\alpha_1, \alpha_2) \).

Note that \( q(r^{-1}(\theta_1, \alpha_1), \alpha_1) = g(\theta_1, \alpha_1) = g(\theta_1, \alpha_2) = q(r^{-1}(\theta_1, \alpha_2), \alpha_2) \). But by Step 5, \( r^{-1}(\theta_1, \alpha_1) > r^{-1}(\theta_1, \alpha_2) \). So, since \( q(\theta, \alpha) \) is strictly increasing in \( \theta \), \( q(r^{-1}(\theta_1, \alpha_2), \alpha_1) < q(r^{-1}(\theta_1, \alpha_2), \alpha_2) \), and hence \( r^{-1}(\theta_1, \alpha_2) > \theta_c(\alpha_1, \alpha_2) \).

Step 7. \( r^{-1}(\theta_2, \alpha_2) \geq r^{-1}(\theta_2, \alpha_1) \).

To see this, combine \( g(\theta_2, \alpha_1) \equiv \frac{q(r^{-1}(\theta_2, \alpha_1), \alpha_1)}{r^{-1}(\theta_2, \alpha_1) - \theta_2} \geq g(\theta_1, \alpha_2) \equiv \frac{q(r^{-1}(\theta_1, \alpha_2), \alpha_2)}{r^{-1}(\theta_1, \alpha_2) - \theta_2} \) with the fact that \( q(r^{-1}(\theta_2, \alpha_2), \alpha_1) \equiv q(\theta_2, \alpha_1) = g(\theta_2, \alpha_2) \equiv q(r^{-1}(\theta_2, \alpha_2), \alpha_2) \).

Step 8. Since \( r^{-1}(\theta_2, \alpha_2) \geq r^{-1}(\theta_2, \alpha_1) > \theta_c(\theta_2, \alpha_2) \), it follows that \( g(\theta_2, \alpha_2) \equiv q(r^{-1}(\theta_2, \alpha_2), \alpha_2) > g(\theta_2, \alpha_1) \equiv q(r^{-1}(\theta_2, \alpha_1), \alpha_1) \). However, by assumption \( g(\theta_2, \alpha_2) = g(\theta_2, \alpha_1) \). This contradiction implies that \( g(\theta, \alpha_1) > g(\theta, \alpha_2) \) \( \forall \theta \in (0, \frac{1}{3} + \frac{2}{3(\sqrt{1 + 2\alpha_2} + 1)}) \).

Q.E.D.

4 Appendix

Proof of lemma 1:

We will demonstrate that, if an admissible quantity schedule does not satisfy the condition in part (i) (part (ii)), then there exists an alternative admissible quantity schedule that satisfies this condition and also guarantees a higher value of the objective (1).

Part (i). Consider an admissible schedule \( s.t. \ g(\theta) > q^*(\theta) \) for some \( \theta \in [0, 1] \). Then the value of the integrand in (1) can be increased by the following modification: set \( g(\theta) = q^*(\theta) \) and reduce the transfer \( t^*(\theta) \) appropriately to make (5) binding. It is easy to see that this modification does not violate any incentive constraints in (3). So, the modified schedule is also admissible.

Now, suppose that \( g(\theta_2) \geq g(\theta_1) \) for some \( \theta_2 < \theta_1 \). Then \( u(g(\theta_2), \theta) - u(g(\theta_2), \theta_2) > u(g(\theta_1), \theta) - u(g(\theta_1), \theta_1) \) \( \forall \theta > \theta_2 \). Hence, \( U(\theta) > u(g(\theta_1), \theta) - u(g(\theta_1), \theta_1) \) \( \forall \theta \in [0, 1] \). So, if \( g(\theta_1) < q^*(\theta_1) \), let us increase \( g(\theta_1) \) by some \( \epsilon \) s.t. \( q^*(\theta_1) - g(\theta_1) \geq \epsilon > 0 \). This modification does not violate any incentive constraints, and causes an increase in the value of (1). If \( g(\theta_1) = q^*(\theta_1) \), then \( g(\theta_2) > q^*(\theta_2) \) which contradicts the result that \( g(\theta) \leq q^*(\theta) \).

Part (ii). Fix an admissible schedule \( s.t. \ g(\theta) > q^*(1) \) for some \( \theta \in [0, 1] \). Let \( \tilde{\theta} = \inf\{\theta|q(\theta) > q^*(1)\} \). Since \( q(.) \) is non-decreasing, \( q(\theta) > q^*(1) \) \( \forall \theta \in (\tilde{\theta}, 1] \). Consider a modified quantity/transfer schedule \( (\tilde{q}(\theta), \tilde{t}^*(\theta)) \) s.t. \( \tilde{q}(\theta) = q(\theta) \), \( \tilde{t}^*(\theta) = t^*(\theta) \) if \( \theta \in [0, \tilde{\theta}] \), and \( \tilde{q}(\theta) = q^*(1), \tilde{t}^*(\theta) = t^*(\theta) + u(q^*(1), \theta) - u(g(\theta), \theta) \) if \( \theta \in [\tilde{\theta}, 1] \). It is easy to check that \( (\tilde{q}(\theta), \tilde{t}^*(\theta)) \) satisfies all incentive constraints. In particular, all constraints in (3) hold because \( g(\theta) \leq q^*(1) \) \( \forall \theta \in [0, 1] \).
The new quantity schedule \( \tilde{q}(\theta) \) generates a larger total surplus. Furthermore, any ‘strategic’ consumer with valuation \( \theta \geq \tilde{\theta} \) gets a lower payoff than prior to this modification, while the payoffs earned by the other consumer types do not change. So, the firm’s expected profit increases.

Now suppose that \( q(1) = \mu < q^*(1) \). Since \( q(\theta) \) is nondecreasing, \( q(\theta) \leq \mu \ \forall \theta \in [0, 1) \). Let \( \theta_m \) be well-defined by the following equality: \( u(q^*(\theta_m), \theta_m) - c(q^*(\theta_m)) = u(\mu, 1) - c(\mu) \). Since \( \mu < q^*(1), \theta_m < 1 \) and \( q^*(\theta_m) > \mu \). So, \( \theta_m > \theta_\mu \) where \( \theta_\mu \) satisfies \( q^*(\theta_\mu) = \mu \). Therefore, \( u(q^*(\theta_m), \theta_m) - c(q^*(\theta_m)) > u(\mu, \theta) - c(\mu) > u(q(\theta), \theta) - c(q(\theta)) \ \forall \theta > \theta_m \).

Let us show that the firm’s expected profit goes up when it offers a modified schedule \( (\tilde{q}(\theta), \tilde{\ell}(\theta)) \) which differs from \( (q(\theta), \ell^*(\theta)) \) on the interval \([\theta_m, 1]\) as follows: \( \tilde{q}(\theta) = q^*(\theta_m) \), \( \tilde{\ell}(\theta) = \ell^*(\theta_m) + u(q^*(\theta_m), \theta_m) - u(q(\theta_m), \theta_m) \). Note that all incentive constraints remain satisfied because all ‘strategic’ consumers’ with valuations in \([\theta_m, 1]\) earn a higher surplus than before this modification.

To see that the firm’s expected profits increases, note that it earns the same profit from selling to a consumer with valuation \( \theta \leq \theta_m \). Also, simple computation demonstrates that the firm’s expected profit from selling to a consumer with valuation \( \theta > \theta_m \) changes by:

\[
u(q^*(\theta_m), \theta_m) - c(q^*(\theta_m)) - u(q(\theta), \theta) - c(q(\theta)) - \int_{\theta_m}^{\theta} u_\theta(q(s), s) ds > 0 \]

The proof that it is optimal to set \( q(0) = 0 \) can be provided along similar lines. \( Q.E.D. \)

**Proof of lemma 2:**

Since the schedule \( q(.) \) is non-decreasing, by Theorems 4.29 and 4.30, p.96 in Rudin (1976) it has at most countably many points of discontinuity on \([0, 1]\), and both the left-hand and the right-hand limits exist at all discontinuity points of \( q(.) \).

Suppose that the optimal quantity schedule \( q(.) \) is discontinuous at \( x \in (0, 1) \). Let \( q(x^-) \) and \( q(x^+) \) be, respectively, the left-hand and the right-hand limits of \( q(.) \) at \( x \). Consequently, we have \( q(x^-) = q(x^+) - 2\delta \) for some \( \delta > 0 \).

Let \( G(q, \theta) = u(q, \theta) - c(q) - u_\theta(q, \theta) \frac{1-F(\theta)}{f(\theta)} \) and \( \Delta(x) = q(x^+) - q(x^-) \). We will consider two different cases.

Case 1: \( G(q(x^-), x) < G(q(x^+, x)) \). By continuity of \( G(q, \theta) \) and \( f(\theta), \exists \varepsilon > 0 \) s.t. \( \forall \theta \in (x-\varepsilon, x), G(q(\theta)+\Delta(x), \theta) > G(q(\theta), \theta) \). Then let us replace the schedule \( q(\theta) \) with modified quantity schedule \( \tilde{q}(\theta) \) s.t. \( \tilde{q}(\theta) = q(\theta) \forall \theta \in [0, x-\varepsilon) \cup (x, 1], \tilde{q}(\theta) = q(\theta)+\Delta(x) \forall \theta \in (x-\varepsilon, x) \) and \( \tilde{q}(x) = q(x^+) \). Note that \( \tilde{q}(\theta) \) is increasing in \( \theta \), and all incentive constraints in (8) still hold because \( U(\theta) \geq U(\theta) \forall \theta \in [0, 1] \). At the same time, the value of the objective 6 increases.

Case 2: \( G(q(x^-), x) \geq G(q(x^+, x)) \). By concavity of \( G(q, \theta) \) in \( q, G((x^-)+q(x^+)/2, x) > G(q(x^-), x)/2 + G(q(x^+), x)/2 \). Furthermore, by continuity of \( G(q, \theta) \) and \( f(\theta), \exists \varepsilon > 0 \) s.t. \( \forall \theta \in (x-\varepsilon, x), G((q^-)+q(x^+)/2, \theta)f(\theta) + G((q(x^-)+q(x^+)/2, \theta + \varepsilon)f(\theta + \varepsilon) > G(q(\theta), \theta)f(\theta) + G(q(\theta + \varepsilon), \theta + \varepsilon)f(\theta + \varepsilon) \).

So, let \( \tilde{q}(\theta) = q(\theta) \forall \theta \in [0, x-\varepsilon) \cup [x+\varepsilon, 1] \) and \( \tilde{q}(\theta) = (q(x^-)+q(x^+)/2) \forall \theta \in (x-\varepsilon, x+\varepsilon) \). Note that \( \tilde{q}(\theta) \) is increasing in \( \theta \).

If \( u_\theta((q^-)+q(x^+)/2, x) > u_\theta(q(x^-), x)/2 + u_\theta(q(x^+), x)/2 \) then \( \varepsilon \) can be chosen small enough that \( \forall \theta \in (x-\varepsilon, x), u_\theta((q(x^-)+q(x^+)/2, \theta)f(\theta) + u_\theta(q(x^-)+q(x^+)/2, \theta + \varepsilon)f(\theta + \varepsilon) > u_\theta(q(\theta), \theta)f(\theta) + u_\theta(q(\theta + \varepsilon), \theta + \varepsilon)f(\theta + \varepsilon) \). So under the quantity schedule \( \tilde{q}(\theta) \),
\[ \tilde{U}(\theta) \equiv \int_0^\theta u_\theta(q(s), s)ds \text{ is greater than } U(\theta) \equiv \int_0^\theta u_\theta(q(s), s)ds \quad \forall \theta \in [0, 1]. \] The value of \( 6 \) changes by:

\[ \int_{x-\varepsilon}^{x+\varepsilon} (G((q(x-)+q(x+))/2, \theta) - G(q(\theta), \theta)) f(\theta)d\theta > 0 \]

If \( u_\theta((q(x-)+q(x+))/2, x) \leq u_\theta(q(\theta), x)/2 + u_\theta(q(x+), x)/2 \), then it is possible that \( \Delta U(x+\varepsilon) = \int_{x-\varepsilon}^{x+\varepsilon} u_\theta(q(s), s) - u_\theta(q(\theta), s)ds < 0 \). In this case, \( \forall \theta \in (x-\varepsilon, x+\varepsilon) \) set \( \tilde{q}(\theta) = \tilde{q} \) s.t. \( \int_{x-\varepsilon}^{x+\varepsilon} u_\theta(\tilde{q}(s), s) - u_\theta(q(\theta), s)ds = 0 \). If \( \varepsilon \) is sufficiently small, then the value of the problem 6 changes approximately by:

\[ \int_{x-\varepsilon}^{x+\varepsilon} (u(\tilde{g}, \theta) - c(\tilde{g}) - u(q(\theta), \theta) + c(q(\theta))) f(\theta)d\theta > 0 \]

The inequality holds by concavity of \( u(\theta, q) - c(q) \). \textit{Q.E.D.}

\textbf{Proof of lemma 3:} Suppose that in the optimal mechanism \( U(0) = \underline{u} > 0 \). Consider set \( Z \subset \Theta \) s.t. \( \theta \in Z \) iff

\[ \underline{u} + \int_0^\theta u_\theta(q(x), x)dx = \sup_{\theta' \in [0, 1]} u(g(\theta'), \theta) - u(g(\theta'), \theta') \tag{39} \]

The set \( Z \) is non-empty, because otherwise the firm could reduce \( U(0) \) and hence increase its expected profits without violating any of the incentive constraints in (8). Let \( \hat{\theta} \) be the minimal element of \( Z \). \( \hat{\theta} \) exists because both the left-hand side and the right-hand side of (39) are continuous in \( \theta \).

Define \( U_\theta(\theta) \equiv \sup_{\theta' \in [0, 1]} u(g(\theta'), \theta) - u(g(\theta'), \theta') \). Note that \( U_\theta(\theta) \) is continuous and strictly increasing in \( \theta \). Since \( g(\theta) \leq q^*(1) \forall \theta \), \( |U_\theta(\theta_1) - U_\theta(\theta_2)| \leq |\theta_1 - \theta_2| \max_{\theta \in [0, 1]} u(\theta, q^*(1)) \), \( U_\theta(\theta) \) is absolutely continuous. Hence, it is almost everywhere differentiable and has finite left-hand and right-hand derivatives for all \( \theta \in [0, 1] \).

The following lemma will be useful below:

\textbf{Lemma 10} Suppose that \( U_\theta(\theta_2) - U_\theta(\theta_1) > u(g(\theta_2), \theta_2) - u(g(\theta_1), \theta_1) \) for some \( \theta_1, \theta_2 \) s.t. \( \theta_2 > \theta_1 \) and \( g \), then \( U_\theta(\theta_4) - U_\theta(\theta_3) > u(g(\theta_4), \theta_4) - u(g(\theta_3), \theta_3) \forall \theta_4, \theta_3 \) s.t. \( \theta_4 > \theta_3 \geq \theta_2 \).

\textbf{Proof:} Suppose that sequence \( \{\theta_n\}_{n=1}^\infty \) is s.t. \( \lim_{n \to \infty} u(g(\theta_n), \theta_2) - u(g(\theta_n), \theta_n) = U_\theta(\theta_2) \) and \( \lim_{n \to \infty} g(\theta_n) = \tilde{g}_2 \). (Such a sequence exists because \( g(\theta) \in [0, q^*(1)] \forall \theta \in [0, 1], \) and any sequence in a compact set has a converging subsequence.) Then \( \tilde{g}_2 > g \). For suppose not, i.e. \( \tilde{g}_2 \leq g \). Then

\[ U_\theta(\theta_2) - U_\theta(\theta_1) \leq \lim_{n \to \infty} u(g(\theta_n), \theta_2) - u(g(\theta_n), \theta_1) = u(\tilde{g}_2, \theta_2) - u(\tilde{g}_2, \theta_1) < u(g(\theta_2), \theta_2) - u(g(\theta_1), \theta_1) \]

Contradiction.

Next, consider a sequence \( \{\theta_m\}_{m=1}^\infty \) s.t. \( \lim_{m \to \infty} u(g(\theta_m), \theta_3) - u(g(\theta_m), \theta_m) = U_\theta(\theta_3) \) and \( \lim_{m \to \infty} g(\theta_m) = \tilde{g}_3 \). Then \( \tilde{g}_3 > g \). Again, suppose otherwise i.e. \( \tilde{g}_3 \leq g \). We have:

\[ U_\theta(\theta_4) = \lim_{m \to \infty} u(g(\theta_m), \theta_3) - u(g(\theta_m), \theta_m) \geq \lim_{n \to \infty} u(g(\theta_n), \theta_3) - u(g(\theta_n), \theta_n) \]

33
Since $\tilde{g}_3 < \tilde{g}_2$ by assumption, it follows that $\exists N, M$ s.t. $\forall n \geq N$ and $m \geq M$, $g(\theta_m) < g(\theta_n)$, and so $u(g(\theta_m), \theta_2) - u(g(\theta_n), \theta_2) > u(g(\theta_m), \theta_3) - u(g(\theta_n), \theta_2)$.

But then $\lim_{m \to \infty} u(g(\theta_m), \theta_2) - u(g(\theta_n), \theta_2) > \lim_{m \to \infty} u(g(\theta_n), \theta_2) = \sup_{\theta' \in [0, 1]} u(g(\theta_2), \theta') - u(g(\theta'), \theta')$. Contradiction. Finally note that

$$U_r(\theta_4) - U_r(\theta_3) \geq \lim_{m \to \infty} u(g(\theta_m), \theta_4) - u(g(\theta_m), \theta_3) = u(\tilde{g}_3, \theta_4) - u(\tilde{g}_3, \theta_1) > u(g, \theta_4) - u(g, \theta_3)$$

$Q.E.D.$

Now, let us demonstrate that the firm can strictly increase its expected profits by offering a modified quantity schedule $\tilde{q}(\theta)$ and setting $U(0) = 0$.

Prior to defining $\tilde{q}(\theta)$ we need to complete a number of intermediate steps. Let $g(\theta -)$ denote the left-hand limit of $g(.)$ at $\theta$ (this limit exists since $g(.)$ is increasing and bounded), and let $\theta^m = \min\{\theta, \sup\{\theta | g(\theta -) \leq q(\theta)\}\}$. Then for $\theta \in [0, \theta^m]$ define:

$$V(\theta) = \int_0^\theta u_\theta(\max\{g(s), q(s)\}, s)ds + \int_\theta^{\hat{\theta}} u_\theta(\max\{g(q(\theta^-), q(s)\}, s)ds$$

We will show that $\exists \theta_0 \in [0, \theta^m]$ s.t. $V(\theta_0) = \theta + \int_0^{\hat{\theta}} u_\theta(q(s), s)ds$. Note that $V(\theta)$ is continuous in $\theta_0$ and $V(0) = \int_0^{\hat{\theta}} u_\theta(q(s), s)ds < \theta + \int_0^{\hat{\theta}} u_\theta(q(s), s)ds$.

Next, let us establish that $V(\theta^m) \geq U_r(\theta)$. Note that $V(\theta^m) = \int_0^\theta u_\theta(\max\{g(s), q(s)\}, s)ds + u(q(\hat{\theta}), \hat{\theta}) - u(q(\hat{\theta}), \theta^m)$. Since $g(.)$ is nondecreasing, $\int_0^\theta u_\theta(\max\{g(s), q(s)\}, s)ds \geq U_r(\theta)$ for $\theta \in [0, \theta^m]$. Hence, by continuity of $V(\theta)$, $\exists \theta_0 \in [0, \theta^m]$ s.t. $V(\theta_0) = U_r(\theta)$.

Define $\tilde{q}(\theta) = \max\{g(\theta), q(\theta)\} \forall \theta \in [0, \theta_0)$, $\tilde{q}(\theta) = \max\{g(\theta_0^-), q(\theta)\} \forall \theta \in [\theta_0, \tilde{\theta}]$. Then clearly, $\tilde{U}(\theta) = U_r(\theta) \forall \theta \in [0, \theta_0]$. Also, $\tilde{U}(\theta) = U_r(\theta) \forall \theta \in [\theta_0, \tilde{\theta}]$ where $\tilde{\theta} = \sup\{\theta | q(\theta^-) < g(\theta_0^-)\}$. Note that $\theta_0 \leq \tilde{\theta}$.

Suppose that $\exists \theta_1 \in (\theta_0, \tilde{\theta})$ s.t. $\tilde{U}(\theta_1) < U_r(\theta_1)$. Then we have $\tilde{U}(\theta_1) - U_r(\theta_1) < U_r(\theta_i) - U_r(\theta_0)$. Note that $q(\theta) = g(\theta_0^-)$ and so $\tilde{U}'(\theta) = u_\theta(g(\theta_0^-), q) \forall \theta \in [\theta_0, \tilde{\theta}]$. So, by lemma 10 $U_r(\theta_1) - U_r(\theta_1) < U_r(\theta_1) - U_r(\theta_1)$, i.e. $U_r(\theta_1) > U_r(\theta_1)$. Contradiction.

When the firm implements quantity schedule $\tilde{q}(\theta)$ rather than $q(\theta)$, sets $U_0 = 0$, and does not modify $g(.)$, the change in the firm’s expected profits is equal to:

$$\theta + \int_0^1 (u(\tilde{q}(\theta), \theta) - u(q(\theta), \theta))f(\theta)d\theta - \int_0^1 (u_2(q(\theta), \theta) - u_2(q(\theta), \theta))f(\theta)(1 - F(\theta))d\theta = \int_0^1 (u(\tilde{q}(\theta), \theta) - u(q(\theta), \theta))f(\theta)d\theta + \int_0^1 (u_2(\tilde{q}(\theta), \theta) - u_2(q(\theta), \theta))f(\theta)F(\theta)d\theta > 0$$

The equality follows from the fact that $\tilde{U}(1) = U(1)$. The inequality follows because both terms in the expression following the equality signs are positive. The second term is positive
because \( q(\theta) \geq q(\theta) \forall \theta \in [0, 1] \). The first term is positive, because in addition, whenever \( q(\theta) > q(\theta), \) \( q(\theta) \leq g(\theta) \leq q^*(1) \), and so, since \( u(q, \theta) \) is quasiconcave in \( q \), \( u(q(\theta), \theta) - u(q(\theta), \theta) > 0 \). Q.E.D.

**Proof of lemma 4:**
Suppose that \( \exists \theta \text{ s.t. } q(\theta) = 0 \) and \( g(\theta) > 0 \). By continuity of \( q(\theta), \exists \theta' > \theta \text{ s.t. } q(\theta') < g(\theta) \) and \( q(\theta) \) is nondecreasing, \( U(\theta) = \int_0^\theta u_\theta(q(s), s)ds < u(g(\theta), \theta') - u(g(\theta), \theta) \) i.e., \( ICT(\theta', \theta) \) in (8) fails.

Next, suppose that \( \exists \theta > 0 \text{ s.t. } q(\theta) = 0 \) and \( q(\theta) > 0 \). By Lemma 1 \( \tilde{\theta} > 0 \). Since \( g(\cdot) \) is nondecreasing and \( q(\cdot) \) is continuous, \( \exists \varepsilon > 0 \text{ s.t. } q(\theta) > 0 = g(\theta) \forall \theta \in [\tilde{\theta} - \varepsilon, \tilde{\theta}] \).

But then without violating any incentive constraints in (8), the firm can increase its profits by setting \( g(\theta) = \min\{q^*(\theta), q(\theta)\} > 0 \) and \( t'(\theta) = u(g(\theta), \theta) \forall \theta \in [\tilde{\theta} - \varepsilon, \tilde{\theta}] \). Q.E.D.

**Proof of lemma 5:**
(i) Since \( q(\theta) \) is nondecreasing, \( U(\theta) = \int_0^\theta u_\theta(q(s), s)ds \leq u(q(\theta), \theta) \), and so \( r(\theta) \geq 0 \). It is immediate from the definition of \( r(\theta) \) that \( r(\theta) \leq \theta \). Let \( \theta_2 > \theta_1 \) and \( q(\theta_1) > 0 \). Then, using the definition of \( r(\theta) \) and the fact that \( q(\cdot) \) is nondecreasing we have:

\[
\begin{align*}
u(q(\theta_2), r(\theta_2)) - u(q(\theta_1), r(\theta_1)) & = u(q(\theta_2), \theta_2) - u(q(\theta_1), \theta_1) - \int_{\theta_1}^{\theta_2} u_\theta(q(s), s)ds \\
& \geq u(q(\theta_2), \theta_1) - u(q(\theta_1), \theta_1)
\end{align*}
\]

(40)

The first equality implies that \( r(\theta_2) = r(\theta) \) if \( q(\theta_2) = q(\theta_1) \), while the inequality in combination with the fact that \( u_{\theta q}(q(\theta) > 0 \) implies that \( r(\theta_2) > r(\theta) \) if \( q(\theta_2) > q(\theta_1) \).

If \( q(\theta) > 0 \), then the continuity of \( r(\theta) \) at \( \theta \) follows from the continuity of \( U(\theta), q(\theta) \) and continuity of \( u(q, s) \) in both arguments. Since \( q(\cdot) \) is nondecreasing, \( \forall \theta \leq \theta \equiv \max\{\theta' | q(\theta') = 0\} q(\theta) = 0 \) and so by definition \( r(\theta) = \theta \). Hence \( r(\theta) \) is continuous on \( [0, \theta] \).

To establish the continuity of \( r(\theta) \) at \( \tilde{\theta} \), consider any \( \theta > \tilde{\theta} \) and note the following:

\[
\begin{align*}
\int_{\theta}^{\tilde{\theta}} u_\theta(q(\theta), s)ds & = u(q(\theta), \theta) - u(q(\theta), r(\theta)) = \int_{\tilde{\theta}}^{\theta} u_\theta(q(s), s)ds \\
& \leq \int_{\theta}^{\tilde{\theta}} u_\theta(q(\theta), s)ds
\end{align*}
\]

The last inequality can hold only if \( r(\theta) \geq \tilde{\theta} \). Since \( r(\theta) \leq \theta \) \( \forall \theta \), we conclude that the right-hand limit of \( r(\theta) \) at \( \tilde{\theta} \) is equal to \( \tilde{\theta} \).

(ii) Suppose that \( ICT(\theta, \theta') \) in (8) holds \( \forall \theta, \theta' \in [0, 1] \).

If \( U(\theta) = 0 \), then \( q(\theta) = 0 \) \( \forall \theta' < \theta \). So by continuity \( q(\theta) = 0 \), and hence \( r(\theta) = \theta \). By Lemma 4 \( g(\theta) = 0 \), so \( g(r(\theta)) = g(\theta) = q(\theta) = 0 \).

If \( U(\theta) > 0 \), then using the definition and the fact that \( ICT(\theta, r(\theta)) \) holds, we have:

\[
\begin{align*}
u(\theta, q(\theta)) - u(r(\theta), q(\theta)) & = U(\theta) \geq u(\theta, g(r(\theta))) - u(r(\theta), g(r(\theta)))
\end{align*}
\]

Then the single-crossing implies that \( q(\theta) \geq g(r(\theta)) \).

Now suppose that \( q(\theta) \geq g(r(\theta)) \) \( \forall \theta \in [0, 1] \). Fix any pair \((\theta, \tilde{\theta})\). Since \( r(\theta) \) is continuous, nondecreasing and \( r(0) = 0 \), either \( \tilde{\theta} = r(\theta) \) for some \( \theta \in [0, 1] \) or \( \tilde{\theta} > r(1) \). If \( \tilde{\theta} = r(\theta) \) for
some \( \tilde{\theta} \in [0,1] \),\(^{21}\) then by assumption \( g(\tilde{\theta}) \leq q(\tilde{\theta}) \). Take any \( \theta \geq \tilde{\theta} \). Then:

\[
U(\theta) = \int_{\tilde{\theta}}^{\theta} u_\theta(q(s),s)ds + U(\tilde{\theta}) \geq u(\tilde{\theta},q(\tilde{\theta})) - u(\tilde{\theta},q(\tilde{\theta})) + u(\tilde{\theta},q(\tilde{\theta})) - u(\tilde{\theta},q(\tilde{\theta})) + u(\tilde{\theta},g(\tilde{\theta})) - u(\tilde{\theta},g(\tilde{\theta}))
\]

The first inequality follows from the fact that \( q(.) \) is nondecreasing and the definition of \( r(\tilde{\theta}) \), while the second inequality follows from single-crossing and \( q(\tilde{\theta}) \geq g(\tilde{\theta}) \). So, \( ICT(\theta, \tilde{\theta}) \) holds. If \( \tilde{\theta} < \theta < \tilde{\theta} \), then \( q(\tilde{\theta}) \leq q(\tilde{\theta}) \). Then \( ICT(\theta, \tilde{\theta}) \) holds because

\[
U(\theta) = U(\tilde{\theta}) - \int_{\tilde{\theta}}^{\theta} u_\theta(q(s),s)ds \geq u(\tilde{\theta},q(\tilde{\theta})) - u(\tilde{\theta},q(\tilde{\theta})) - (u(\tilde{\theta},q(\tilde{\theta})) - u(\tilde{\theta},q(\tilde{\theta}))) \geq u(\tilde{\theta},g(\tilde{\theta})) - u(\tilde{\theta},g(\tilde{\theta}))
\]

Finally, suppose that \( \tilde{\theta} > r(1) \). Since \( g(\tilde{\theta}) \leq g^*(1) = q(1) \), for \( \theta > \tilde{\theta} \) we have:

\[
U(\theta) = U(1) - \int_{\tilde{\theta}}^{1} u_\theta(q(s),s)ds \geq u(q(1),1) - u(q(1),1) - (u(q(1),1) - u(q(1),1)) = u(q(1),1) - u(q(1),1) \geq u(g(\tilde{\theta}),\theta) - u(g(\tilde{\theta}),\theta) > u(g(\tilde{\theta}),\theta) - u(g(\tilde{\theta}),\tilde{\theta})
\]

Thus, \( ICT(\theta, \tilde{\theta}) \) also holds in this case. \( QED. \)

**Proof of Theorem 2:** Suppose that \( \tilde{\theta} \equiv \inf\{\theta | q(\theta) > 0\} > 0 \). We will show that there exists \( \epsilon \in (0, \tilde{\theta}/4) \), s.t. the firm can strictly increase its profits by replacing the quantity schedule \( q(\theta) \) with \( \tilde{q}(\theta) = \max\{\epsilon, q(\theta)\} \ \forall \theta \in [0,1]. \)

Since \( q(\theta) \) is continuous and nondecreasing, for any sufficiently small \( \epsilon > 0 \) there exists a unique \( \tilde{\theta}(\epsilon) \) s.t. \( q(\tilde{\theta}(\epsilon)) = \epsilon \) and \( q(\tilde{\theta}) \geq \epsilon \ \forall \theta \in [\tilde{\theta}(\epsilon), 1]. \)\(^{22}\) Under the modified schedule \( \tilde{q}(\theta), \) 'strategic' consumer obtains surplus \( \tilde{U}(\theta) = u(\epsilon, \theta) \ \forall \theta \in [0, \tilde{\theta}(\epsilon)] \), and \( \tilde{U}(\theta) = u(\epsilon, \tilde{\theta}(\epsilon)) + \int_{\tilde{\theta}(\epsilon)}^{\theta} u_\theta(q(s),s)ds \ \forall \theta \in [\tilde{\theta}(\epsilon), 1]. \)

Define \( \tilde{r}(\theta) \) as the solution to \( \tilde{U}(\theta) = u(\tilde{q}(\theta),\theta) - u(\tilde{q}(\theta),\tilde{r}(\theta)) \). By lemma 9 \( g(\theta) = \min\{q^*(\theta),g(q^{-1}(\theta))\} \) and \( \tilde{g}(\theta) = \min\{q^*(\theta),\tilde{g}(q^{-1}(\theta))\} \). Also, define \( G(q,\theta) = u(q,\theta) - c(q) - u_\theta(q,\theta) \frac{1-F(\theta)}{f(\theta)} \).

Then (6) implies that after this modification the firm’s expected profit changes by:

\[
\Delta(\epsilon) = \int_{0}^{1} (G(\tilde{q}(\theta),\theta) - G(q(\theta),\theta)) f(\theta)d\theta + \alpha \int_{0}^{1} (u(\tilde{g}(\theta),\theta) - c(\tilde{g}(\theta)) - u(g(\theta),\theta) - c(g(\theta)))) f(\theta)d\theta
\]

We will show that \( \Delta(\epsilon) > 0 \) if \( \epsilon \) is sufficiently small, because the second term is positive and is of higher order than \( \epsilon \), while the first term may be negative, but is at most of order \( \epsilon \).

To establish the claim regarding the first term, pick some \( \epsilon > 0 \) and define \( \omega = \max_{\theta \in [0,\tilde{\theta}(\epsilon)]} \frac{\partial G(q(\theta))}{\partial q} \). Note that \( \omega < \infty \), and by Weierstrass Theorem, \( \forall \epsilon \leq 2 \)

\[
G(\tilde{q}(\theta),\theta) - G(q(\theta),\theta) \leq \omega(\tilde{q}(\theta) - q(\theta)) \leq \omega \epsilon, \text{ and so } \int_{0}^{1} (G(\tilde{q}(\theta),\theta) - G(q(\theta),\theta)) f(\theta)d\theta \leq \omega \epsilon.
\]

\(^{21}\)\( \tilde{\theta} \) need not be unique. However, if \( \tilde{\theta} = r(\tilde{\theta}_1) = r(\tilde{\theta}_2) \), then as established above \( q(\tilde{\theta}_1) = q(\tilde{\theta}_2) \).

\(^{22}\)If such \( \tilde{\theta}(\epsilon) \) fails to exist for all \( \epsilon > 0 \), then \( q(\theta) = 0 \) everywhere which is suboptimal.
Now, let us focus on the second term. If $\theta \geq \hat{\theta}(\epsilon)$, then $q(\theta) = \tilde{q}(\theta)$ and $U(\theta) < \hat{U}(\theta)$, so $\bar{r}(\theta) < r(\theta)$. Also, $\forall \theta \in [0, \hat{\theta}(\epsilon)] \bar{r}(\theta) = 0$, while $r(\theta) = \theta \forall \theta \in [0, \hat{\theta}]$ and $r(\theta) \geq \hat{\theta} \forall \theta \in (\hat{\theta}, \hat{\theta}(\epsilon)]$. So, $\bar{r}(\theta) < r(\theta) \forall \theta 
exists (0, 1)$. Lemma 9 implies that $q^*(\theta) \geq \tilde{g}(\theta) \geq g(\theta) \forall \theta \in [0, 1]$. Therefore, since $u(q(\theta)) \leq c(q)$ is concave, $u(\tilde{g}(\theta) \theta - c(\tilde{g}(\theta))) \geq u(g(\theta) \theta - c(g(\theta))) \forall \theta \in [0, 1]$. By Lemma 4 $g(\theta) = 0 \forall \theta \in [0, \hat{\theta}]$. So,

$$\int_{\theta}^{0} (u(\tilde{g}(\theta), \theta - c(\tilde{g}(\theta))) - (u(g(\theta), \theta - c(g(\theta)))) f(\theta)d\theta \geq \int_{\theta}^{0} (u(\tilde{g}(\theta), \theta - c(\tilde{g}(\theta))) f(\theta)d\theta$$

Let us show that $\tilde{g}(\theta) = q(\bar{r}^{-1}(\theta)) \leq q^*(\theta) \forall \theta \in [\theta/2, \hat{\theta}]$. First, let us establish that $\lim_{\epsilon \to 0} \bar{r}^{-1}(\theta) = \theta_{23}$. For suppose, to the contrary, that there exists a sequence $\epsilon_n$, $\lim_{n \to \infty} \epsilon_n = 0$, and $\eta > 0$ s.t. $\theta_n = \bar{r}^{-1}(\theta, \epsilon_n)$ (we explicitly incorporate the dependence of $\bar{r}^{-1}(\theta)$ on $\epsilon$) and $\lim_{n \to \infty} \theta_n \geq \theta + \eta$. Let $\theta_l$ denote the limit of a converging subsequence of $\theta_n$. Obviously, $\theta_l \geq \theta + \eta$.

Note that $\bar{U}(\theta_n, \epsilon_n) = u(q(\theta_n), \theta_n) - u(q(\theta_n), \theta_n)$. Since $\bar{U}(\theta_n, \epsilon_n)$ converges to $U(\theta)$ uniformly as $\epsilon$ converges to zero, and $U(\theta)$ is continuous, we have $U(\theta_l) = u(q(\theta_l), \theta_l) - u(q(\theta_l), \theta_l)$. On the other hand, $U(\theta_l) = \int_{\theta}^{0} u_\theta(q(s), s) ds < u(q(\theta_l), \theta_l) - u(q(\theta_l), \theta_l)$ where the inequality follows from continuity of $q(\theta)$. This contradiction implies that $\lim_{\epsilon \to 0} \bar{r}^{-1}(\theta) \leq \theta$. Since $\bar{r}^{-1}(\theta) > \theta \forall \epsilon > 0$, we conclude that $\lim_{\epsilon \to 0} \bar{r}^{-1}(\theta) = \theta$.

Let us now fix some $\psi \in (0, q^*(\theta/2))$. Combining $\lim_{\epsilon \to 0} \bar{r}^{-1}(\theta) = \theta$ with the continuity of $q(\theta)$ we obtain that $\exists \epsilon > 0$ s.t. $\forall \epsilon < \epsilon g(\bar{r}^{-1}(\theta)) = q^*(\theta/2) - \psi$, and so $\bar{g}(\theta) = q(\bar{r}^{-1}(\theta)) \leq q^*(\theta/2) - \psi \forall \theta \in [\theta/2, \hat{\theta}]$.

Let $\zeta = \min_{\theta \in [\theta/2, \hat{\theta}], \epsilon \in [0, q^*(\theta/2) - \psi]} u_\theta(q, \theta) - c(q)$ and $f = \min_{\theta \in [\theta/2, \hat{\theta}]} f(\theta)$. Note that $\zeta > 0$ and $\hat{f} > 0$. Then $\int_{\theta}^{\theta} \int_{\theta}^{\theta} (u(\bar{g}(\theta), \theta) - c(\bar{g}(\theta))) f(\theta)d\theta \geq \zeta \hat{f} \int_{\theta}^{\theta} \bar{g}(\theta)d\theta$.

Next we establish a lower bound on $\bar{g}(\theta)$ for $\theta \in [\theta, \hat{\theta}]$. Let $m = \min_{\theta \in [0, 1], \epsilon \in [0, q^*(\theta/2)]} u_\theta(q, \theta)$ and $M = \max_{\theta \in [0, 1], \epsilon \in [0, q^*(\theta/2)]} u_\theta(q, \theta)$. Our assumptions on $u(q, \theta)$ imply that $0 < m \leq M < \infty$. Then

$$\bar{U}(\bar{r}^{-1}(\theta)) = u(\bar{g}(\theta), \bar{r}^{-1}(\theta)) - u(\bar{g}(\theta), \theta) = \int_{\theta}^{\bar{r}^{-1}(\theta)} u_\theta(\bar{g}(\theta), s) ds = \int_{\theta}^{\bar{r}^{-1}(\theta)} \int_{\theta}^{\bar{g}(\theta)} u_\theta(q, s) dq ds \leq M \bar{g}(\theta)(\bar{r}^{-1}(\theta) - \theta) \tag{42}$$

On the other hand, since $\bar{r}^{-1}(\theta) > \theta$,

$$\bar{U}(\bar{r}^{-1}(\theta)) = \bar{U}(\theta) = \int_{\theta}^{\theta} u_\theta(\epsilon, s) ds = \int_{\theta}^{\theta} \int_{\theta}^{\theta} u_\theta(q, s) ds dq \geq m \theta \tag{43}$$

Combining (42) and (43), we obtain that $\bar{g}(\theta) \geq \frac{m \theta}{M(\bar{r}^{-1}(\theta) - \theta)}$. Collecting these results together, we have:

$$\int_{\theta}^{\theta} (u(\bar{g}(\theta), \theta) - c(\bar{g}(\theta))) \frac{d\theta}{\bar{r}^{-1}(\theta) - \theta} \geq \int_{\theta}^{\theta} \frac{d\theta}{\bar{r}^{-1}(\theta) - \theta}$$

$^{23}$Note that, as shown in Lemma 5, $q(\bar{r}^{-1}(\theta))$ is well-defined, but the preimage $\bar{r}^{-1}(\theta)$ may be an interval. However, all the arguments in this paper apply to every element in the preimage $\bar{r}^{-1}(\theta)$ i.e. each $\theta'$ s.t. $\bar{r}(\theta') = \theta$. For simplicity, we will continue to use notation $\bar{r}^{-1}(\theta)$ which we set equal to $\max\{\theta'|\bar{r}(\theta') = \theta\}$. 

37
To complete the proof, we will show that \( \int_{\theta/2}^{\theta} \frac{d\theta}{\tilde{r}^{-1}(\theta) - \theta} \) increases to \( \infty \) as \( \epsilon \) converges to 0. Fix some \( \rho \in [0, \theta/2] \). Since \( \lim_{\epsilon \to 0} \tilde{r}^{-1}(\theta) = \theta \) \( \forall \theta \in [\theta/2, \theta] \), by Lebesgue’s dominated convergence theorem, \[
\lim_{\epsilon \to 0} \int_{\theta/2}^{\theta} \frac{d\theta}{\tilde{r}^{-1}(\theta) - \theta} = \int_{\theta/2}^{\theta} \frac{d\theta}{\theta - \theta} = \log(\theta/2) - \log(\rho) \]
Note that \( \lim_{\rho \to 0} \log(\rho) = -\infty \), which proves the desired result. \( Q.E.D. \)

**Proof of lemma 7:**

Suppose that \( q(\theta) \) is a solution to Problem (10) on the domain \( C^1_p([0,1]) \), but there exists an admissible schedule \( \hat{q}(\theta) \in C([0,1]) \setminus C^1_p([0,1]) \) s.t. the objective function in (10) takes a strictly higher value under \( \hat{q}(\theta) \) than under \( q(\theta) \).

By the Stone-Weierstrass theorem, the space of continuously differentiable functions \( C^1([0,1]) \), which is a subspace of \( C^1_p([0,1]) \) is dense in \( C([0,1]) \). Therefore, \( C^1_p([0,1]) \) is dense in \( C([0,1]) \). So, there exists a sequence \( \hat{q}_n(\theta) \in C^1_p([0,1]) \) converging to \( \hat{q}(\theta) \) in the \( sup \)-norm. The objective function (10) is continuous in the \( sup \)-norm. Therefore, \( \exists N > 0 \) s.t. \( \forall n \geq N \) (10) takes a strictly higher value under \( \hat{q}_n(\theta) \) than under \( q(\theta) \). This contradicts the hypothesis that \( q(\theta) \) is a solution on \( C^1_p([0,1]) \). \( Q.E.D. \)

**References**


