Adverse selection without single crossing*

Christoph Schottmüller†

September 19, 2011

Abstract

Screening models are used to analyze contracting in many subfields of economics like regulation, labor economics, monopoly pricing, taxation or finance. Most models assume single crossing. This simplifies the analysis as local incentive compatibility is in this case sufficient for global incentive compatibility. If single crossing is violated, global incentive compatibility constraints have to be taken into account. This paper studies monotone solutions in a model where single crossing is violated.

It is shown that local and non-local incentive constraints distort the solution in opposite directions. Therefore, the optimal decision might involve distortions above as well as below the first best decision. Furthermore, the well known "no distortion at the top" property does not necessarily hold. Sufficient conditions for monotonicity and continuity of the solution and an algorithm to obtain such a solution are derived.

Some results can be readily applied. For example, overinsurance, i.e. insurance levels above first best as in so called "Cadillac" insurance plans, can be rationalized. In a non-linear pricing framework, the model also provides an explanation for marginal prices below marginal costs as observed in flat rate offers.

Keywords: adverse selection, single crossing, Spence-Mirrlees condition, global incentive compatibility

JEL classification: D82, D86

*I want to thank Jan Boone and Bruno Jullien for interesting discussions and numerous suggestions. I have also benefited from comments by Matthias Lang, François Salanié and seminar participants at the Toulouse School of Economics and Tilburg University.

†Department of Economics, University of Tilburg, P.O. Box 90153, 5000 LE, Tilburg, The Netherlands; Tilec, CentER; E-mail: christophschottmueller [at] googlemail.com

1. Introduction

Adverse selection models[1] are among the most used microeconomic models since their introduction by Akerlof (1970). The main feature of these models is that one (or more) agents have private information which is relevant for transactions with other players. This private information can be the efficiency of a firm in models of regulation (Baron and Myerson, 1982; Laffont and Tirole, 1987), the productivity of a worker in labor market (Guasch and Weiss, 1981) as well as in optimal taxation models (Mirrlees, 1971), the risk of an accident in insurance models (Stiglitz, 1977) or the willingness to pay for a product in models of monopoly pricing (Mussa and Rosen, 1978) and auctions (Myerson, 1981).

Two standard research questions typically emerge in this kind of models: What will be the market outcome, e.g. the optimal contract? How does the presence of asymmetric information affect welfare and the distribution of the social surplus? Generally speaking, a menu emerges as optimal contract, i.e. several options are offered and the player who has private information will choose his preferred option. The chosen option will normally not be what a benevolent planner with complete information would assign. Hence, informational distortions exist and will reduce welfare. The reason in a nutshell is that the agent reveals (some of) his private information by his choice. This will not be costless for the principal who designs the menu: The agent receives an informational rent. By distorting the menu away from first best, the principal can reduce this informational rent to his own benefit.

In the regulation example, a regulator will want a more efficient firm to produce a higher quantity than a less efficient firm. But an efficient firm could claim to be inefficient and choose the (low quantity) option intended for an inefficient firm from the menu. Since the firm is efficient, it would make a profit by "misrepresenting". By distorting the quantity intended for an inefficient firm, the regulator can make such misrepresenting less profitable for an efficient firm. Hence, the regulator can save on the informational rents of an efficient firm by distorting the menu option intended for an inefficient firm away from first best.

Single crossing[2] is a technical assumption which is usually made in hidden information models. In one dimensional models, single crossing states that types[3] can be ordered according to their marginal rate of substitution between monetary transfers and the decision, e.g. produced quantity in the regulation example above. With the usual quasilinear preferences, single crossing is equivalent to a type ordering according to marginal utilities.

In the regulation example above, the firm's cost function depends on quantity and type. Single

[1] Adverse selection models are sometimes also referred to as hidden information or screening models.

[2] The single crossing property is also referred to as Spence-Mirrlees condition or sorting condition.

[3] A "type" is an agent with a specific private information attribute, see Harsanyi (1967). In the regulation example types correspond to cost functions of the firm.

crossing means that higher types have lower marginal costs (for any admissible quantity). Single crossing is violated if such an ordering is impossible, e.g. a higher type has lower marginal costs for high quantities but higher marginal costs for low quantities.

In a non-linear pricing framework, single crossing would mean that higher types have a higher marginal utility at every possible quantity level. Now think of fixed line internet access. Heavy internet users will certainly have a higher marginal utility from the fifth gigabyte of data than light users. If heavy users, however, also own smartphones with internet access (and light users do not), light users will probably have a higher willingness to pay for the first 50 megabyte: They cannot switch to their mobile devices to check emails etc.. Hence, single crossing would be violated.

This paper analyzes an adverse selection model in which single crossing is violated. Agents have quasilinear preferences and a one-dimensional type. The setting allows for a one time violation of single crossing; e.g. for a given quantity, marginal costs are first in- and then decreasing in type. Without single crossing, local incentive compatibility does no longer guarantee global incentive compatibility. Therefore, non-local incentive compatibility constraints have to be taken into account. The paper analyzes monotone solutions in this setup, e.g. situations in which higher types produce higher quantities under the optimal contract. Sufficient conditions for the existence of a monotone solution and an algorithm to calculate such a solution are presented.

With single crossing, there is no distortion at the top and the distortion for all types goes in the same direction, e.g. all types produce a quantity which is weakly below their first best quantity. If single crossing is violated, both results no longer hold. The reason is that binding non-local incentive constraints will counteract the normal distortion stemming from local incentive compatibility and rent extraction motives. A rough intuition for this result is the following: With single crossing, distortions occur because the principal wants to lower the agent's informational rent. If a non-local incentive constraint is violated, a certain type's rent at "his contract" is too low compared with another type's contract. To satisfy his non-local incentive constraint, his rent has to be increased. Reducing the normal distortion (or even distorting the decision in the opposite direction) will result in such an increase.

The following section gives several examples of settings in which single crossing is violated. Then the literature is reviewed and the formal model is introduced. Section 4 also states a sufficient condition for the existence of a monotone solution. Given existence, one can turn to analyzing the solution. Section 5 introduces necessary conditions which have to hold at types where non-local incentive constraints are binding. The core of the paper are the sections 6 and 7: The former characterizes monotone solutions while the latter focuses on the special case of monotone and continuous solutions. An explanation why the no-distortion-at-the-top property is not always satisfied follows. Before concluding, I discuss some assumptions and point out differences with solutions obtained in related problems in the literature. Most proofs are relegated to the appendix.

2. Examples

This section illustrates why single crossing is violated in a number of reasonable economic settings. A common theme of the examples is that there are more than one input/option/relevant characteristic. It is then a priori not clear (and sometimes even unreasonable) that a higher type is "better" on all dimensions. But this is exactly what single crossing would require.

Example 1: two factor production. Take a setting where a firm or government has to contract with the provider of a good (input or public good/infrastructure etc.). If the principal is a government, this setting is mathematically equivalent to incentive regulation (compare for example Laffont and Tirole (1993)). Assume now that production uses variable production factors in fixed proportions. Costs of these factors can be proportional to output, e.g. energy costs and unskilled labor, while other factors increase costs convexly in quantity, e.g. skilled labor (due to search costs) and machine utilization. Type indexes the possible production technologies and denotes which of these two groups of factors is used more efficiently by the firm. A cost function representing this setting could be[4]

$$c(q, \theta) = \theta q + \frac{q^2}{\theta} + \gamma(\theta)$$

where $\gamma(\theta)$ are (possibly type dependent) fixed costs. Whether marginal costs $c_q(q, \theta) = \theta + \frac{2q}{\theta}$ are in- or decreasing in type depends on q. Put differently, the cross partial derivative $c_{q\theta}(q, \theta) = 1 - 2q/\theta^2$ can change sign and therefore single crossing is violated. The idea is simple: For low quantities, the linear part of the cost function dominates marginal costs and therefore high types have higher marginal costs. For high quantities, the convex part of the cost function is more relevant and therefore high types have lower marginal costs.

In practice, type could represent whether a firm uses a labor intensive or capital intensive production technology. A labor intensive production technology requires especially unskilled labor which can be hired at a constant market wage (linear part). A capital intensive technology requires less but more skilled employees. Finding them is increasingly difficult and results therefore in convexly increasing costs.

A second interpretation of the cost function above could apply in the case of environmental regulation. Let the principal be a government designing a subsidy scheme to reduce emissions. The decision

[4] The alternative cost function $c(q, \theta) = \theta q + (1 - \theta)q^2 + \gamma(\theta)$ also violates single crossing.

q is the amount of emission reduction. Reducing emissions can be achieved by lowering the content of a dirty input in favor of a more expensive clean input. This is a linear cost. Alternatively, the emission reduction can be obtained by filtering and other changes in the production process. This second option becomes increasingly costly the more one has to rely on it. Hence, the convex part of the cost function. The government does not know the firm's production technology which is its type θ. Depending on the production technology, it is easier for the firm to filter or to substitute inputs.

It should be mentioned that the cost function in this example can be viewed as a simplified version of the flexible fixed cost quadratic cost function suggested by Baumol et al. (1982). Beard et al. (1991) estimate such a cost function for savings and loans associations. Interestingly, they allow for two unobservable types of production technology in their estimation. In table 5, Beard et al. (1991) report estimated costs for the two types ("mixtures" in their language) at different quantity levels. If one interprets cost differences between the output levels as marginal costs, it turns out that mixture 1 has lower marginal costs at low output levels but higher marginal costs at high output levels. Hence, single crossing is violated.

Example 2: hiring talent and productivity. This example is in the context of compensation of workers.[5] The principal is the owner of a firm and the agent a worker the firm wants to hire. For the quality of the worker talent and effort are relevant, e.g. talent is what the worker produces in a regular working time like the 40 hours week and effort is the additional time he is willing to invest. Assume the worker creates value $q = e\theta + T$ where T is his talent, e is the unobservable effort and θ is his type. The owner of the firm observes a public signal, e.g. education, which is a mix of talent and productivity (he does not observe T and θ directly). To be precise, assume that the signal is $\sigma = \theta * T$. Given this signal, a more productive worker will have lower talent and vice versa. The production function of the manager for a given signal is $q = e\theta + \sigma/\theta$ where q is the quantity/value produced by the worker. If costs of effort are e^2 and the worker's preferences are quasilinear in money, his utility function can be written as

$$u(q, \theta) = w - \frac{(q - \sigma/\theta)^2}{\theta^2} \tag{1}$$

where w is wage. It is easy to check that single crossing is violated. The intuition is that a low type can produce a low output q without much effort just within the regular working time. Hence, his marginal costs of effort (and therefore of q) are low. A high type already has to exert some effort to reach the same output level and therefore his marginal costs of effort (and q) are higher. Note here that the contract is conditional on education, i.e. given σ a more productive type will be less talented. For high output, where effort of both types is substantial, higher types have lower marginal costs since

[5] A similar example can be found in Araujo and Moreira (2010).

they are more productive.

Example 3: common agency. As already mentioned in Martimort and Stole (2009), violations of single crossing can arise if more than one principal contract with the same agent. Interestingly, the utility function itself will satisfy single crossing (for a fixed decision with the other principal) and the violation of single crossing results from the existence of multiple principals. This example tries to convey the idea in a simplified setup.

The source of hidden information in this example is the inability of firms to know the exact preferences of a customer. A firm cannot observe the preferences of a customer but it can engage in non-linear pricing, i.e. second degree price discrimination.

Say, consumers can buy two goods which are imperfect substitutes: Good A is sold only by firm A while good B is available on a perfectly competitive market at a constant per unit price p^B.[6] For concreteness, let the demand function for good B of a type θ consumer be

$$q^B(q^A,\theta) = \theta(\beta - p^B - \delta q^A) \tag{2}$$

which means that type rotates the inverse demand function outwards. The following quadratic utility function yields such a demand function:

$$u(q^A,q^B,\theta) = \alpha q^A + \beta q^B - \frac{\gamma}{2\theta}{q^A}^2 - \frac{1}{2\theta}{q^B}^2 - \delta q^A q^B - p^B q^B - p^A(q^A)$$

Firm A faces consumers buying product B according to (2). By plugging (2) into the utility function, one can obtain utility as a function of q^A and θ alone, i.e. $v(q^A,\theta) = u(q^A,q^B(q^A,\theta),\theta)$. This is the utility function firm A has to take into account in its profit maximization problem. Because consumers buy also product B, single crossing is violated:

$$v_{q^A\theta}(q^A,\theta) = u_{q^A\theta}(q^A,q^B(q^A,\theta),\theta) + u_{q^B\theta}(q^A,q^B(q^A,\theta),\theta)\frac{\partial q^B(q^A,\theta)}{\partial q^A} = q^A\left(\frac{\gamma}{\theta^2}+\delta^2\right) - \delta(\beta - p^B)$$

Clearly, $v_{q^A\theta}$ is negative for low q^A and positive for high q^A. The reason for the violation of single crossing is that high type consumers have, on the one hand, a higher marginal willingness to pay because of their high type (that is the $\frac{\gamma}{2\theta}{q^A}^2$ term in the utility function). On the other hand, a high type buys more of product B which reduces his willingness to pay for product A as the two goods are substitutes.

Example 4: health insurance. This example is worked out in Boone and Schottmüller (2011) and therefore only sketched here. In health insurance, it is empirically documented that people with high health risks have often little insurance coverage. This cannot be explained by a standard insurance

[6] See Martimort and Stole (2009) for a model in which the second good is also offered by a strategically acting principal.

model with single crossing. Boone and Schottmüller (2011) point out that an empirically observed correlation between health risk and income might lead to a violation of single crossing.

In a nutshell, assume that risk is private information but that high risk agents are poorer. At full coverage, that is indemnity insurance without deductible, wealth does not matter and high risk agents will (in expectation) consume more care. Therefore, their marginal willingness to pay for coverage is higher. Now think of a situation without insurance coverage: As health care is a normal good, poor, high risk agents will consume less care when falling ill: They cannot afford care. Because they utilize less, their marginal willingness to pay for insurance coverage is less than the one of a rich, low risk agent. Consequently, it depends on the coverage level whether higher types are willing to pay more or less for a marginal unit of insurance coverage; single crossing is violated.

Example 5: insurance with mean variance utility. An agent faces a risk of losing a (money equivalent) amount D with probability θ where θ is private information. His preferences are given by the mean variance utility function

$$u(q,\theta) = \theta(w - (1-q)D) + (1-\theta)w - p - 1/2r\theta(1-\theta)(1-q)^2D^2$$

where p is the insurance premium of an insurance covering fraction q of the loss, w is initial wealth and $r > 0$ is a measure of risk aversion. The cross derivative $u_{q\theta} = D + (1-q)rD^2(1-2\theta)$. If $\theta > 1/2$ and $rD > 1$, the cross derivative can change sign depending on q. Hence, single crossing is violated.

The intuition is that for $\theta > 1/2$ a higher risk also implies less variance. Consequently, a higher type is on the one hand more eager to buy insurance because he has a higher risk on the other hand he is less eager to buy insurance because there is less variance in his payoffs. At full coverage, i.e. for $q = 1$, the payoff variance is zero and the latter effect is no longer present. For lower coverage levels, however, it might dominate.

3. Literature

The standard screening model with single crossing is well known and explained in many textbooks, see for example Fudenberg and Tirole (1991) or Bolton and Dewatripont (2005). Surprisingly, the literature on violations of single crossing in screening models remains relatively scarce.

Some insights have been gained for discrete type insurance models with perfect competition among principals. Several papers analyze settings where private information has two dimensions and can take either a high or a low value in each dimension, i.e there are 2×2 types. In Smart (2000), the two dimensions are risk and risk aversion while in Wambach (2000) they are wealth and risk. Netzer and Scheuer (2010) model an additional labor supply decision and the two dimensions are productivity and risk. All three papers share a pooling result, i.e. if single crossing is violated two of the four types can be pooled. Boone and Schottmüller (2011) show that with imperfect competition among principals there can even be an order reversal: Types with higher risk can have more but also less insurance coverage if single crossing is violated.

My paper will analyze a model with a continuum of types and one principal. As I will illustrate in the next section, the main technical difficulty caused by a violation of single crossing are non-locally binding incentive constraints. In discrete type models one can take all incentive constraints explicitly into account. This is quite difficult in a continuous type model since a continuum of constraints exist. Indeed the main technical challenge is to handle those constraints. Also some additional qualitative results emerge from the continuous type model, e.g. distortion above as well as below first best and distortion at the top.

Araujo and Moreira (2010) characterize in a continuous type framework (inversely) U-shaped solutions in a setup where single crossing is not satisfied. In these solutions, some contracts are chosen by two types ("discrete pooling"). It turns out that in (inversely) U-shaped solutions non-local incentive constraints are only binding between types choosing the same contract from the menu. My paper complements their work by characterizing monotone solutions in the same model. The main technical difference is that non-local incentive constraints can bind between types choosing different options from the menu. The solution in Araujo and Moreira (2010) features either a discontinuity or a bunching interval. My paper shows that this is not the case for monotone solutions and therefore not a necessary implication of a violation of single crossing.

Violations of single crossing are also related to the literature on multidimensional screening, see Armstrong (1996) and Rochet and Chone (1998) for seminal contributions and Rochet and Stole (2003) for a survey. As pointed out in the survey, "the problems arise not because of multiple dimensionality itself, but because of a commonly associated lack of exogenous type-ordering in multiple-dimensional environments." A violation of single crossing also conveys a lack of type-ordering. To make the relationship clear, think of a multidimensional, discrete type model. Clearly, one can reassign types to a one-dimensional parameter but this reassigned type will regularly not satisfy single crossing. Consequently, an applied researcher will often have the choice between a multidimensional type model or a one-dimensional type model violating single crossing. My paper provides tools to make the latter way feasible.

The paper also relates to work relaxing the basic assumptions of the textbook model. Jullien (2000) allows for type dependent participation constraints while Hellwig (2010) analyzes the case of irregular type distributions, i.e. distributions with mass points and zero densities. In section 9 the solution obtained with a violation of single crossing will be compared with the solutions obtained in those papers.

4. Model

There is a one-dimensional decision in a principal agent relationship which is denoted by $q \in \mathbb{R}_+$. Furthermore, there is a monetary transfer $t \in \mathbb{R}$. The agent's utility is $\pi = t - c(q, \theta)$ where $\theta \in \Theta \equiv [\underline{\theta}, \bar{\theta}] \subset \mathbb{R}$ is the type of the agent which is his private information. The function $c(q, \theta)$ is assumed to be three times continuously differentiable with $c_q > 0$, $c_{qq} > 0$, $c_\theta < 0$.[7] The principal's utility is $u(q, \theta) - t$ and is two times continuously differentiable with $u_q > 0$ and $u_{qq} \leq 0$. The principal has the prior distribution $F(\theta)$ with continuous density $f(\theta) > 0$ for all $\theta \in [\underline{\theta}, \bar{\theta}]$.

For example, the principal could be the regulator of a natural monopolist and q could be the quality (or quantity) of service provided. The regulator might maximize expected consumer surplus which could be $q - p$ where p is the price paid. The natural monopolist would have cost function $c(q, \theta)$ and maximize profits. A higher type would correspond to a more efficient firm in the sense that its costs are lower than the costs of a lower type.

By the revelation principle, any general mechanism can also be implemented by a direct revelation mechanism in which the agent truthfully reports his type. The task is to design a menu $q(\theta)$, implemented by transfers $t(\theta)$, which is individually rational (ir) and incentive compatible (ic) for the agent and maximizes the principal's objective under these two constraints.

Faced with a menu $(q(\theta), t(\theta))$, a type θ agent will maximize $t(\hat{\theta}) - c(q(\hat{\theta}), \theta)$ over his type announcement $\hat{\theta}$. If an implementable menu $(q(\theta), t(\theta))$ leads to rents/profits $\pi(\theta)$, the envelope theorem and truthful revelation therefore require $\pi_\theta(\theta) = -c_\theta(q(\theta), \theta)$.

Incentive compatibility of a decision $q(\theta)$ requires in general for any $\theta, \hat{\theta} \in \Theta$

$$\Phi(\theta, \hat{\theta}) \equiv \pi(\theta) - \pi(\hat{\theta}) - c(q(\hat{\theta}), \hat{\theta}) + c(q(\hat{\theta}), \theta) \geq 0. \tag{IC}$$

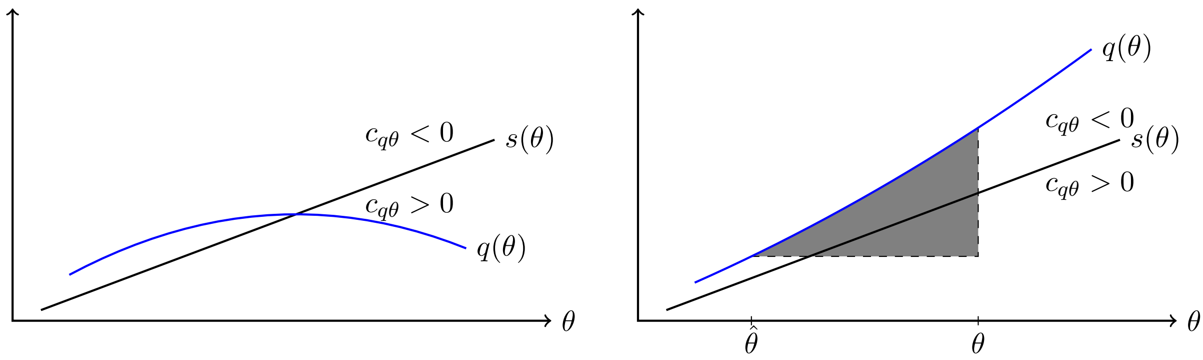
Using the envelope condition above, $\Phi(\theta, \hat{\theta})$ can be rewritten as

$$\Phi(\theta, \hat{\theta}) = \int_\theta^{\hat{\theta}} c_\theta(q(t), t) - c_t(q(\hat{\theta}), t)\, dt = -\int_\theta^{\hat{\theta}} \int_{q(t)}^{q(\hat{\theta})} c_{q\theta}(s, t)\, ds\, dt.$$

Consequently, (IC) is equivalent to

$$-\int_\theta^{\hat{\theta}} \int_{q(t)}^{q(\hat{\theta})} c_{q\theta}(s, t)\, ds\, dt \geq 0. \tag{IC'}$$

[7] For the case where c_θ can change sign (but single crossing is satisfied) see Jullien (2000).



(a) inverse U-shape solution

(b) monotone solution

Figure 1: possible solution shapes

Single crossing in this model is equivalent to $c_{q\theta}(q, \theta)$ not changing sign for any value of q and θ. But then incentive compatibility in (IC') boils down to a simple monotonicity condition on $q(\theta)$ (plus the envelope condition): If $c_{q\theta} < 0$, then inequality (IC') will hold whenever $q(\theta)$ is monotonically increasing. If however $c_{q\theta}$ can change sign, this is no longer true. It remains true that $q(\theta)$ has to be increasing (decreasing) at θ if $c_{q\theta}(q(\theta), \theta) < (>) 0$. Otherwise, (IC') would be violated for types close enough to θ. But this no longer implies global incentive compatibility for two arbitrary types θ and $\hat{\theta}$.

This paper focusses on a one-time violation of single crossing also used by Araujo and Moreira (2010): It is assumed that $c_{q\theta}$ changes sign only once for a given q (or a given θ). More precisely, I assume $c_{q\theta\theta} > 0$ and $c_{qq\theta} < 0$. Hence, there exists a strictly increasing function $s(\theta)$ such that $c_{q\theta}(s(\theta), \theta) = 0$. The assumption on third derivatives are normally made to ensure concavity of the objective function and monotonicity of the decision, see for example Fudenberg and Tirole (1991). Here, however, they provide some structure on the way single crossing is violated.

Araujo and Moreira (2010) find necessary conditions for the case where the solution is inversely U-shaped, see figure 1a.[8] Note that distinct types are assigned the same decision. Consequently, they have to get the same transfer as well and the incentive compatibility constraint has to be binding between those types. It turns out that non-local incentive compatibility constraints are *only* binding between such discretely pooled types.

The focus of my paper will be on the case where the optimal decision is monotone.

Although $c_{q\theta}(q(\theta), \theta) < 0$ for all θ, the violation of single crossing still plays a role in monotone solutions. It follows from (IC') that one can represent incentive compatibility as an integral over the shaded area in figure 1b: If the integral of $c_{q\theta}$ over this shaded area is negative, incentive compatibility

[8] The figure is more schematic than reflecting the solution in Araujo and Moreira (2010): They show that the inversely U-shaped solution typically displays a bunching interval or a discontinuity.

is satisfied for θ and $\hat{\theta}$. Hence, the part where $c_{q\theta} > 0$ plays a role though the solution does not pass it.

The intuition is the following: Take two types θ and $\hat{\theta}$ with $\theta > \hat{\theta}$. Type $\hat{\theta}$ is assigned a transfer decision pair $(\hat{t}, \hat{q})$ and likewise θ has pair (t, q) with $q > \hat{q}$. When deciding whether he should misrepresent, type θ will compare the transfer difference $t - \hat{t}$ with the cost difference $c(q, \theta) - c(\hat{q}, \theta)$. Note that the transfer difference does not depend on type while the cost difference does. With single crossing, the cost difference is decreasing in type. If a type $\theta' \in (\hat{\theta}, \theta)$ with $q' \in (\hat{\theta}, q)$ is introduced, it follows that $c(q, \theta) - c(\hat{q}, \theta) < c(q, \theta) - c(q', \theta) + c(q', \theta') - c(\hat{q}, \theta')$. On the other hand, for transfers $t - \hat{t} = t - t' + t' - \hat{t}$ holds. Therefore, incentive compatibility between θ and $\hat{\theta}$ is implied by incentive compatibility between θ and θ' as well as between θ' and $\hat{\theta}$. Evidently, local incentive compatibility implies non-local incentive compatibility because single crossing implies that the cost difference is decreasing in type. Without the single crossing assumption, the cost difference $c(q, \theta) - c(\hat{q}, \theta)$ is not necessarily decreasing in type and therefore local incentive constraints are not necessarily more demanding than non-local ones.

Before turning to the analysis of the solution, some definitions and one assumption is needed. I define the *first best* solution as the solution to

$$\max_{q(\theta)} u(q(\theta), \theta) - c(q(\theta), \theta)$$

which would be the optimal decision if the principal observed the agent's type. As a second reference point, it is useful to look at the *relaxed program*. This is the program taking only local incentive compatibility into account:

$$\begin{aligned} \max_{q(\theta)} \ & \int_{\underline{\theta}}^{\bar{\theta}} \{u(q(\theta), \theta) - c(q(\theta), \theta) - \pi(\theta)\} f(\theta) \, d\theta \\ s.t.: \ & \pi_\theta(\theta) = -c_\theta(q(\theta), \theta) \\ & q_\theta(\theta) c_{q\theta}(q(\theta), \theta) \leq 0 \\ & \pi(\theta) \geq 0 \end{aligned} \tag{RP}$$

The first and second constraint are the local incentive compatibility constraints. More specifically, the first constraint is a first order condition for incentive compatibility and the second constraint is the so called monotonicity constraint. The third constraint is the participation constraint which will bind only for $\underline{\theta}$ by the assumption $c_\theta < 0$. I will call the solution of (RP) the *relaxed solution* and denote it by $q^r(\theta)$.

Since this paper focuses on the violation of single crossing in monotone solutions, the following assumption is made:

Assumption 1. *The relaxed program is strictly concave in $q(\theta)$ and the relaxed solution is strictly monotonically increasing and strictly above $s(\theta)$.*

Put differently, I assume that the monotonicity constraint does not bind and the relaxed solution is fully characterized by the first order condition. It is easy to show that $u_{qq} \leq 0$ and $c_{qq} \geq 0$ are sufficient for concavity. For strict monotonicity and $q^r(\theta) > s(\theta)$, the following assumptions would be sufficient: $u_{q\theta} \geq 0$, $q^{fb}(\theta) > s(\theta)$ and the commonly made monotone hazard rate assumption, i.e. $f(\theta)/(1 - F(\theta))$ non-decreasing in θ.

Under assumption 1, it is routine to verify that the relaxed solution is characterized by the first order condition

$$\{u_q(q(\theta), \theta) - c_q(q(\theta), \theta)\}f(\theta) + (1 - F(\theta))c_{q\theta}(q(\theta), \theta) = 0. \tag{3}$$

Since $q^r(\theta) > s(\theta)$, it follows that $c_{q\theta}(q^r(\theta), \theta) < 0$. Therefore, (3) implies that $q^r(\theta) \leq q^{fb}(\theta)$ where the inequality is strict for all types but $\bar{\theta}$.

As already indicated, solutions can be monotone or inversely U-shaped (or even crossing $s(\theta)$ with a discontinuous jump). It is therefore useful to have a sufficient condition under which the solution is monotone. To get such a sufficient condition, a technical condition has to be added to assumption 1.

To state this technical condition some "mirror images" have to be defined: Take a decision q below $s(\theta)$ and consider mirroring this decision in two ways: First, mirror it along $s(\theta)$ such that $\int_q^{q^s} c_{q\theta}(x, \theta)\, dx = 0$ where $q^s(q, \theta)$ is the implicitly defined mirror image. Second, mirror q along the relaxed solution q^r such that $\{u(q, \theta) - c(q, \theta)\}f(\theta) + (1 - F(\theta))c_{q\theta}(q(\theta), \theta)$ is the same for q and its mirror image $q^v(q, \theta)$. Since $c_\theta(q, \theta)$ and (RP) are concave in q, the two mirror images are well defined. Last define $q^f(\theta) < s(\theta)$ such that $q^s(q^f(\theta), \theta) = q^r(\theta)$, i.e. $q^f(\theta)$ is a kind of mirror image of the relaxed solution along $s(\theta)$.[9]

Proposition 1. *If $q^v(q, \theta) \geq q^s(q, \theta)$ for all $q \in [0, q^f(\theta)]$ and all $\theta \in [\underline{\theta}, \bar{\theta}]$, then any decision function $q(\theta)$ which imposes decisions below $s(\theta)$ for some type is dominated by the following changed decision*

$$q^c(\theta) = \begin{cases} q(\theta) & \text{if } q(\theta) \geq s(\theta) \\ q^s(q(\theta), \theta) & \text{else} \end{cases}$$

combined with transfers such that $\pi^c_\theta = -c_\theta(q^c(\theta), \theta)$.

Proof. see appendix

Note that the imposed condition is automatically satisfied for q close to $q^f(\theta)$ by assumption 1. Hence, the condition roughly states that $q^s(q, \theta)$ is not much steeper in q than $q^v(q, \theta)$. This holds,

[9] If no $q^f(\theta) \geq 0$ exists, take $q^f(\theta) = 0$.

for example, true if $\{u(q,\theta) - c(q,\theta)\}f(\theta) + (1 - F(\theta))c_\theta(q(\theta),\theta)$ and $c_\theta(\cdot)$ are both symmetric in q, i.e. if $\{u(q^r(\theta) - \Delta,\theta) - c(q^r(\theta) - \Delta,\theta)\}f(\theta) + (1 - F(\theta))c_\theta(q^r(\theta) - \Delta,\theta) = \{u(q^r(\theta) + \Delta,\theta) - c(q^r(\theta) + \Delta,\theta)\}f(\theta) + (1 - F(\theta))c_\theta(q^r(\theta) + \Delta,\theta)$ and $c_\theta(s(\theta) - \Delta,\theta) = c_\theta(s(\theta) + \Delta,\theta)$ for any Δ as then $q^v_q(q,\theta) = q^s_q(q,\theta) = -1$.

In short, proposition 1 says that under the condition $q^v(q,\theta) \geq q^s(q,\theta)$ the optimal decision is monotone. This is not exactly true as proposition 1 does not establish existence of an optimal solution. Appendix C closes this loophole by showing that a solution exists.

Given that $q^v(q,\theta) \geq q^s(q,\theta)$ is sufficient but not necessary for a monotone solution, this condition will not be used in the remainder of the paper where monotone solutions are characterized.

5. Necessary conditions

This section presents necessary conditions which have to be met whenever a non-local incentive constraint is binding. Since these conditions are only a slight generalization of those presented in Araujo and Moreira (2010), the presentation will be brief and more intuitive than formal.

Take an optimal decision schedule $q(\theta)$ and let transfers be determined by local incentive compatibility, i.e. such that $\pi_\theta(\theta) = -c_\theta(q(\theta),\theta)$ and $\pi(\underline{\theta}) = 0$. Furthermore, suppose that IC is binding for two types θ and $\hat{\theta}$, i.e. $\Phi(\theta,\hat{\theta}) = 0$. By incentive compatibility, $\Phi(\cdot)$ has to be non-negative for all types and therefore $(\theta,\hat{\theta}) \in argmin_{(s,t)}\Phi(s,t)$.

Given that $\pi(\cdot)$ and $c(\cdot)$ are differentiable, the first order condition with respect to θ has to hold:[10]

$$\frac{\partial\Phi(\theta,\hat{\theta})}{\partial\theta} = -c_\theta(q(\theta),\theta) + c_\theta(q(\hat{\theta}),\theta) \leq 0 \qquad \text{with “=” if } \theta < \bar{\theta} \tag{C1}$$

In the same way the first order condition for $\hat{\theta}$ is derived:

$$\frac{\partial\Phi(\theta,\hat{\theta})}{\partial\hat{\theta}} = q_\theta(\hat{\theta})\left(-c_q(q(\hat{\theta}),\hat{\theta}) + c_q(q(\hat{\theta}),\theta)\right) \geq 0 \qquad \text{with “=” if } \hat{\theta} > \underline{\theta} \tag{C2}$$

Hence, $\hat{\theta}$ is either bunched or marginal costs of θ and $\hat{\theta}$ are equal at $q(\hat{\theta})$.

The interpretation of these two conditions is straightforward. Recall that $\pi_\theta(\theta) = -c_\theta(q(\theta),\theta)$ while $c_\theta(q(\hat{\theta}),\theta)$ is how profits of misrepresenting as $\hat{\theta}$ change in the misrepresenting type θ. Then condition (C1) says that profits $\pi(\theta)$ should change in type in the same way as misrepresentation-profits change in type. For a graphical interpretation, it is worthwhile to rewrite (C1) as

$$\int_{q(\hat{\theta})}^{q(\theta)} c_{q\theta}(q,\theta)\,dq = 0 \tag{C1'}$$

[10] It turns out that non-local incentive compatibility constraints are only downward binding, see lemma 1. For this reason as well as notational convenience, I ignore the possibilities $\Phi(\theta,\bar{\theta}) = 0$ and $\Phi(\underline{\theta},\hat{\theta}) = 0$ already here.

which means that the right hand side boundary of the shaded area in figure 1b is zero when weighted with $c_{q\theta}$. If the integral above was positive and $\Phi(\theta,\hat{\theta})=0$, then incentive compatibility would be violated for $\theta+\varepsilon$ and $\hat{\theta}$ as $\Phi(\theta+\varepsilon,\hat{\theta}) \approx \Phi(\theta,\hat{\theta}) - \varepsilon \int_{q(\hat{\theta})}^{q(\theta)} c_{q\theta}(q,\theta)\,dq$, i.e. the "shaded area" for $\theta+\varepsilon$ would be the same plus some area having the "wrong" sign.

If the integral above is negative, the same applies accordingly for $\theta-\varepsilon$, i.e. $\Phi(\theta-\varepsilon,\hat{\theta}) \approx \Phi(\theta,\hat{\theta}) + \varepsilon \int_{q(\hat{\theta})}^{q(\theta)} c_{q\theta}(q,\theta)\,dq$.

The second condition simply says that either $\hat{\theta}$ is bunched with other types or also the weighted lower boundary of the shaded area in figure 1b is zero, i.e.

$$\int_{\hat{\theta}}^{\theta} c_{qt}(q(\hat{\theta}),t)\,dt = 0. \tag{C2'}$$

Again, figure 1b illustrates the idea. If the integral was positive, incentive compatibility would be violated between θ and $\hat{\theta}-\varepsilon$ as $\Phi(\theta,\hat{\theta}-\varepsilon) \approx \Phi(\theta,\hat{\theta}) - \varepsilon q_{\theta}(\hat{\theta}) \int_{\hat{\theta}}^{\theta} c_{qt}(q(\hat{\theta}),t)\,dt$.

The graphical interpretation also allows to quickly generalize these conditions at points of discontinuity and bunching. This situation is depicted in figure 2. Assume $\Phi(\theta,\hat{\theta}_i)=0$ for $i=1,2$. To keep incentive compatibility for types close to θ, $\hat{\theta}_1$ and $\hat{\theta}_2$ the following conditions have to hold:[11]

- $\int_{q(\hat{\theta}_i)}^{q^-(\theta)} c_{q\theta}(q,\theta)\,dq \geq 0$ as otherwise $\Phi(\theta-\varepsilon,\hat{\theta}_i) < 0$
- $\int_{q(\hat{\theta}_i)}^{q^+(\theta)} c_{q\theta}(q,\theta)\,dq \leq 0$ as otherwise $\Phi(\theta+\varepsilon,\hat{\theta}_i) < 0$
- $\int_{\hat{\theta}_1}^{\theta} c_{qt}(q(\hat{\theta}),t)\,dt \leq 0$ as otherwise $\Phi(\theta,\hat{\theta}_1-\varepsilon) < 0$
- $\int_{\hat{\theta}_2}^{\theta} c_{qt}(q(\hat{\theta}),t)\,dt \geq 0$ as otherwise $\Phi(\theta,\hat{\theta}_1+\varepsilon) < 0$

Given (C1) and (C2), one can use variational calculus to derive a third necessary condition for types at which the incentive constraint binds. While (C1) and (C2) are purely driven by incentive compatibility, this third condition will be derived from the principal's optimization. The idea is to perturb the optimal decision around θ and $\hat{\theta}$ such that the two necessary conditions are still satisfied. For an optimal decision the derivative of the principal's virtual valuation with respect to the perturbation parameter has to be zero. The method differs only slightly from the one used in Araujo and Moreira (2010) for discretely pooled types and therefore the steps are relegated to appendix A. The following variational condition results:

$$\frac{[u_q(q(\theta),\theta)-c_q(q(\theta),\theta)]f(\theta)}{c_{q\theta}(q(\theta),\theta)} + 1 - F(\theta) = \frac{[u_q(q(\hat{\theta}),\hat{\theta})-c_q(q(\hat{\theta}),\hat{\theta})]f(\hat{\theta})}{c_{q\theta}(q(\hat{\theta}),\hat{\theta})} + 1 - F(\hat{\theta}) \tag{C3}$$

The interpretation of this condition will become clearer later on.

[11] I use the superscript "–" ("+") to indicate limits from below (above).

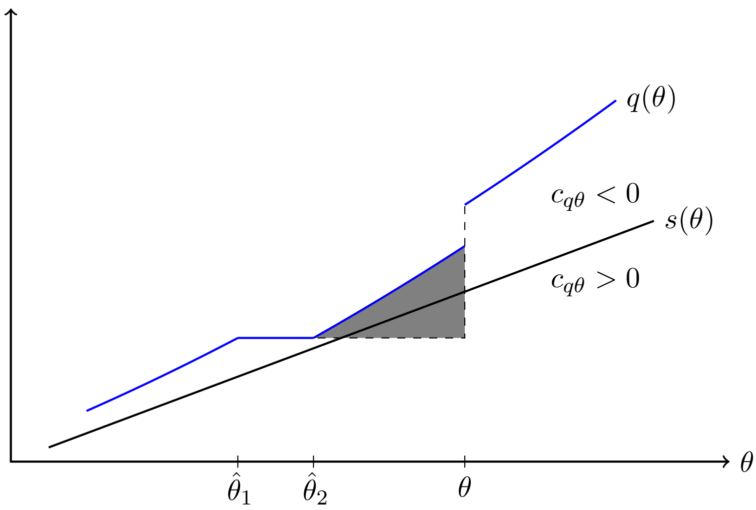


Figure 2: necessary conditions at discontinuity

6. Monotone solution

The remainder of the paper deals with the characterization of monotone solutions. As pointed out before, the main difficulties are non-locally binding incentive constraints. The following two lemmata show that only a certain subset of non-local incentive constraints can be binding. Lemma 1 implies that incentive constraints cannot be upward binding in monotone solutions. Put differently, no type will be indifferent between the contract designated for him and the contract of a higher type. The only possible way a non-local incentive constraint can be binding is downward, i.e. a type might be indifferent between his contract and the contract of a lower type.

Lemma 1. *If $q(\theta) \geq s(\theta)$ and $q(\theta)$ is locally incentive compatible, then no type wants to (non-locally) misrepresent upwards.*

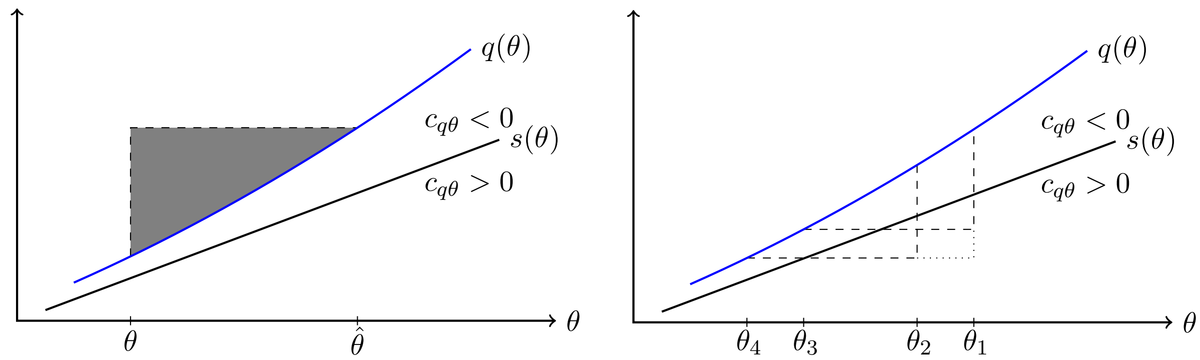
Proof. Recall that local incentive compatibility requires monotonicity of $q(\theta)$, i.e. $q(\theta)$ has to be monotonically increasing as $q(\theta) \geq s(\theta)$. Now take $\hat{\theta} > \theta$. Incentive compatibility requires

$$\Phi(\theta, \hat{\theta}) \equiv \pi(\theta) - \pi(\hat{\theta}) - c(q(\hat{\theta}), \hat{\theta}) + c(q(\hat{\theta}), \theta) \geq 0 \tag{4}$$

Because of local incentive compatibility, this can be rewritten as

$$\int_{\theta}^{\hat{\theta}} c_t(q(t), t) - c_t(q(\hat{\theta}), t)\, dt = -\int_{\theta}^{\hat{\theta}} \int_{q(t)}^{q(\hat{\theta})} c_{qt}(s, t)\, ds\, dt \geq 0$$

But the last inequality holds automatically since $q(\theta) \geq s(\theta)$ and $q_\theta(\theta) \geq 0$. This implies that the integrand is non-positive for all (s, t) in question. Figure 3a gives a graphical representation of this fact. $\square$



(a) no upwards binding

(b) no overlap

Figure 3: non-binding constraints

The intuition for lemma 1 is the same as in models with single crossing. A higher decision increases the costs for higher types less than for lower types. For a low type, this holds true for all decisions above his own. Local incentive compatibility induces transfer differences making higher types indifferent between their decision and a marginally higher decision. A lower type will face the same transfer differences but higher cost differences when opting for a higher decision. Therefore, local incentive compatibility of higher types implies that low types do not want to misrepresent upwards non-locally.

The following lemma puts more structure on the ways incentive compatibility constraints can bind. It states that binding non-local incentive constraints cannot overlap. Before, stating the lemma one remark on wording: I say a non-local incentive constraint binds *from* θ *to* $\hat{\theta}$ if $\Phi(\theta, \hat{\theta}) = 0$.

Lemma 2. *Assume the solution is monotone. If the non-local incentive constraint binds from* θ *to* $\hat{\theta}$*, it cannot bind from any* $\theta' \in [\hat{\theta}, \theta)$ *to any* $\hat{\theta}' \notin [\hat{\theta}, \theta)$*. Neither can it bind for any* $\hat{\theta}'' \in (\hat{\theta}, \theta]$ *and* $\theta'' \notin (\hat{\theta}, \theta)$*. (assuming that not all relevant types are bunched)*

Proof. The proof is by contradiction. Suppose, contrary to the lemma, there are types $\theta_1 > \theta_2 \geq \theta_3 > \theta_4$ with $\Phi(\theta_1, \theta_3) = 0$ and $\Phi(\theta_2, \theta_4) = 0$. Then the incentive constraint between θ_1 and θ_4 will be

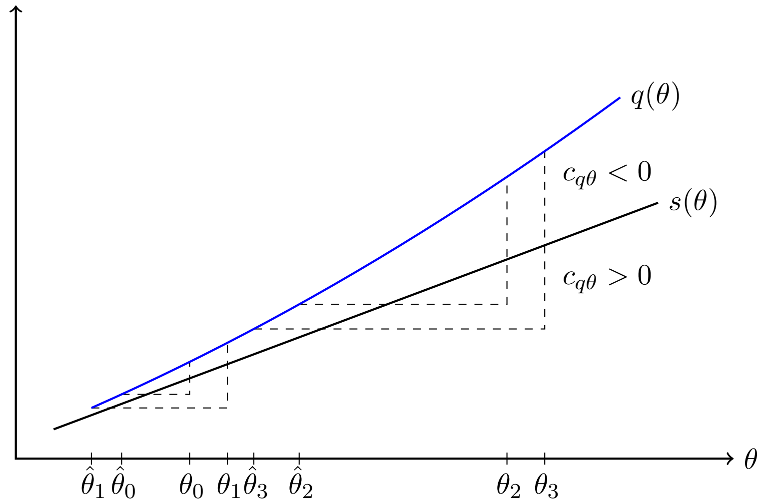


Figure 4: how incentive constraints can bind

violated, i.e. $\Phi(\theta_1, \theta_4) < 0$:

$$\begin{aligned}
\Phi(\theta_1, \theta_4) &= -\int_{\theta_4}^{\theta_1} \int_{q(\theta_4)}^{q(t)} c_{qt}(s,t)\, ds\, dt \\
&= -\int_{\theta_4}^{\theta_2} \int_{q(\theta_4)}^{q(t)} c_{qt}(s,t)\, ds\, dt - \int_{\theta_2}^{\theta_1} \int_{q(\theta_4)}^{q(\theta_3)} c_{qt}(s,t)\, ds\, dt - \int_{\theta_2}^{\theta_1} \int_{q(\theta_3)}^{q(t)} c_{qt}(s,t)\, ds\, dt \\
&= -\int_{\theta_4}^{\theta_2} \int_{q(\theta_4)}^{q(t)} c_{qt}(s,t)\, ds\, dt - \int_{\theta_2}^{\theta_1} \int_{q(\theta_4)}^{q(\theta_3)} c_{qt}(s,t)\, ds\, dt \\
&\quad + \int_{\theta_3}^{\theta_2} \int_{q(\theta_3)}^{q(t)} c_{qt}(s,t)\, ds\, dt - \int_{\theta_3}^{\theta_1} \int_{q(\theta_3)}^{q(t)} c_{qt}(s,t)\, ds\, dt \\
&= -\Phi(\theta_2, \theta_3) - \int_{\theta_2}^{\theta_1} \int_{q(\theta_4)}^{q(\theta_3)} c_{qt}(s,t)\, ds\, dt < 0
\end{aligned}$$

The first and second equality are simple splitting up the integral steps (and can readily be seen in figure 3b), the third uses the fact that $\Phi(\theta_1, \theta_3) = \Phi(\theta_2, \theta_4) = 0$ and the last inequality follows from the incentive compatibility between θ_2 and θ_3 as well as the following idea: By the binding constraint between θ_2 and θ_4 and the fact that θ_2 is interior, $\int_{q(\theta_4)}^{q^-(\theta_2)} c_{s\theta}(s, \theta_2)\, ds \geq 0$ holds by C1 (with equality if $q(\theta)$ is continuous at θ_2). By the monotonicity of $q(\cdot)$, $q(\theta_3) \leq q^-(\theta_2)$ and therefore $\int_{q(\theta_4)}^{q(\theta_3)} c_{s\theta}(s, \theta_2)\, ds \geq 0$ (see figure 3b). The inequality above follows then from $c_{q\theta\theta} \geq 0$. □

As a special case, i.e. with $\theta_2 = \theta_3$, the preceding lemma includes the following: If θ is indifferent between his and $\hat{\theta}$'s contract, i.e. $\Phi(\theta, \hat{\theta}) = 0$, then no other type θ' is indifferent between his contract and θ's contract, i.e. $\Phi(\theta', \theta) > 0$ for all $\theta' \in \Theta \setminus \theta$. Put differently, incentive compatibility can bind non-locally from a type or to a type but not both. Figure 4 summarizes the two previous lemmata by showing how non-local incentive compatibility constraints can bind in a monotone solution.

One of the contributions of this paper is that a violation of single crossing can affect the solution without leading to irregularities, i.e. discontinuities or bunching. The following lemma shows that some irregularities can be ruled out on the grounds of incentive compatibility alone.

Lemma 3. *Assume a non local incentive constraint binds from θ to $\hat{\theta}$, i.e. $\Phi(\theta, \hat{\theta}) = 0$. The decision is continuous at $\hat{\theta}$ if $\hat{\theta}$ is not the boundary type of a bunching interval. Furthermore, θ cannot be bunched if the decision is continuous at θ and $\theta < \bar{\theta}$.*

Proof. see appendix

After these technical results, it is possible to obtain a qualitative result of practical importance. If the solution is monotone, non-local incentive compatibility might require "distortions" that are unusual: *With single crossing,* local incentive constraints are downward binding. This explains why the relaxed solution is below the first best decision. With single crossing, a high type has lower marginal costs than a low type. By distorting the low type's decision downward, the cost advantage of the high type is reduced, i.e. the low type's decision becomes less attractive. Consequently, the rent paid to the high type can be lower without inducing misrepresentation. *Without single crossing,* it is no longer clear that a high type has lower marginal costs than a low type at the low type's decision. Figure 1b, for example, illustrates that $\int_{\hat{\theta}}^{\theta} c_{q\theta}(q(\hat{\theta}), t)\, dt = c_q(q(\hat{\theta}), \theta) - c_q(q(\hat{\theta}), \hat{\theta})$ could be positive. Therefore, making the low type's contract unattractive might require increasing the low type's decision. Informational distortion from local and non-local incentive constraints will then go in opposite directions. The following proposition shows that this indeed the case.

Proposition 2. *If the optimal decision is monotone, it will be above the relaxed solution, i.e. if $q(\theta)$ monotonically increasing, then $q(\theta) \geq q^r(\theta)$.*

Proof. see appendix

The previous proposition highlights how violations of non-local ic are dealt with under monotone solutions. This can also be illustrated with figure 1b. Incentive compatibility is violated if the grey area weighted by $c_{q\theta}$ is positive. To satisfy incentive compatibility one can raise q for all types between $\hat{\theta}$ and θ. The additional grey area features $c_{q\theta} < 0$ and therefore the incentive problem is mitigated.

One noteworthy point is that the incentive constraint is mainly relaxed by increasing q for types at which the incentive constraint is non-binding; i.e. if ic is binding from θ' to $\hat{\theta}'$, it is less $q(\theta')$ and $q(\hat{\theta}')$ that has to be increased but q for the types between $\hat{\theta}'$ and θ'. To see the intuition, recall that $\pi_\theta(\theta) = -c_\theta(q(\theta), \theta)$ and that $c_{q\theta}(q(\theta), \theta) < 0$. Therefore, increasing q will raise the slope of the rent function $\pi(\theta)$. Increasing q for types in $(\hat{\theta}', \theta')$ will therefore increase the rent of θ' at his assigned menu point. Obviously, the non-local incentive constraint is relaxed.

The last paragraph illustrates that non-local incentive constraints are potentially difficult to handle: The decision of a type is not only influenced by the incentive constraints binding for him but also by binding incentive constraints of other types. The following theorem structures this intuition and characterizes the solution.

Theorem 1. *A monotone solution is characterized by the equation*

$$(u_q(q(\theta),\theta) - c_q(q(\theta),\theta))f(\theta) + (1 - F(\theta))c_{q\theta}(q(\theta),\theta) = \eta(\theta)c_{q\theta}(q(\theta),\theta) \tag{5}$$

where $\eta(\theta)$ is a non-negative function with the following properties:

- *$\eta(\theta)$ is constant on each interval of types for which non-local incentive constraints are not binding and the decision is strictly increasing.*
- *$\eta(\theta)$ is non-decreasing at types $\hat{\theta}$ to which non-local incentive constraints are binding whenever $\hat{\theta}$ is not bunched.*
- *$\eta(\theta)$ is non-increasing at types from which non-local incentive constraints are binding.*
- *$\eta(\bar{\theta})$ is zero if no non-local incentive constraint is binding from $\bar{\theta}$.*
- *$\eta(\underline{\theta})$ is zero if no non-local incentive constraint is binding to $\underline{\theta}$.*

Proof. see appendix

Before giving an intuitive interpretation to $\eta(\theta)$, let me briefly sketch the idea behind the proof of the theorem. Given the solution $q(\theta)$, one can simply define $\eta(\theta)$ by (5). The properties of $\eta(\theta)$ are derived by showing that $q(\theta)$ could be changed in a way that (i) is incentive compatible and (ii) increases the principal's payoff if these properties were not satisfied. Figure 5 shows feasible changes when a non-local incentive constraint is binding from θ' to $\hat{\theta}'$. Increasing the decision for types slightly below θ' will relax (or not affect) binding non-local incentive constraints. Since this change relaxes the incentive constraints from types above θ' to types below θ', it is then feasible to assign types slightly above θ' a lower decision, see figure 5. Note that lemma 2 is essential for feasibility as it assures that no non-local incentive constraint is binding *to* types slightly above θ'. It can then be shown that such a feasible change would increase the principal's payoff if $\eta(\theta)$ was increasing at θ'. At $\hat{\theta}$, a different change in the decision is feasible, see figure 5, which can be used to show that $\eta(\theta)$ cannot be decreasing at $\hat{\theta}$. At types where non-local incentive constraints are lax, both kind of changes are feasible and consequently $\eta(\theta)$ has to be constant.

The properties of $\eta(\theta)$ have an intuitive interpretation. The left hand side of (5) measures by how much the principal's payoff is changed when marginally increasing $q(\theta)$. Marginally increasing $q(\theta)$ will also relax all non-local incentive constraints binding from types $\theta' > \theta$ to types $\hat{\theta}' < \theta$, see figure 1b. As these incentive constraints can be expressed as integrals over $c_{q\theta}$ (see equation (IC')), the "amount" by

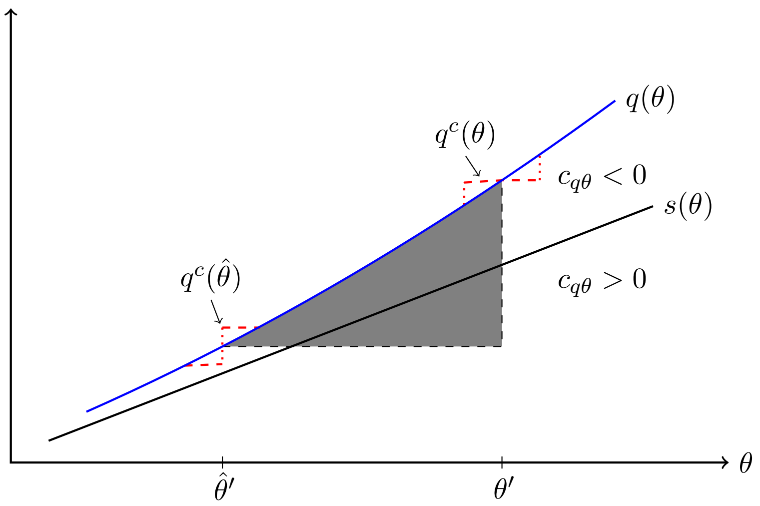


Figure 5: feasible changes

which those non-local incentive constraints are relaxed is given by $c_{q\theta}(q(\theta), \theta)$ which can be found on the right hand side of (5). Consequently, $\eta(\theta)$ could be interpreted as the shadow value of all non-local incentive constraints binding from types $\theta' > \theta$ to types $\hat{\theta}' < \theta$. These binding constraints are the same for all types in an interval of types for which non-local incentive constraints are lax, see figure 4. This explains the first property of $\eta(\theta)$.

The other properties can also be explained by the shadow value interpretation of $\eta(\theta)$. If a non-local incentive constraint is binding to a type $\hat{\theta}$, then there are more non-local incentive constraints binding "over" $\hat{\theta} + \varepsilon$ than "over" $\hat{\theta} - \varepsilon$.[12] Consequently, the shadow value of non-local incentive constraints binding over a type has to be higher for $\hat{\theta} + \varepsilon$ than for $\hat{\theta} - \varepsilon$. Put differently, increasing $q(\hat{\theta} + \varepsilon)$ relaxes more non-local incentive constraints than increasing $q(\hat{\theta} - \varepsilon)$.

Also the last two properties are straightforward: Increasing the decision of the boundary types does not affect non-local incentive constraints of other types.

Furthermore, the interpretation as shadow value provides some intuition for the necessary condition (C3) which basically says that $\eta(\theta) = \eta(\hat{\theta})$ when a non-local incentive constraint is binding from θ to $\hat{\theta}$. This makes sense in light of lemma 2. Because there is no overlap in binding incentive constraints, the non-local incentive constraints binding over θ are the same as the ones binding over $\hat{\theta}$. Consequently, the shadow value of relaxing those constraints is the same for the two types.

Theorem 1 establishes what happens at types where non-local incentive constraints are binding (or lax). Here I want to argue that non-local incentive constraints are typically binding from and to intervals of types. Put differently, there are intervals $[\theta_0, \theta_1]$ and $[\hat{\theta}_1, \hat{\theta}_0]$ such that a non-local incentive

[12] With binding "over" θ I mean binding from a type $\theta' > \theta$ to a type $\hat{\theta} < \theta$.

constraint is binding from each $\theta' \in [\theta_0, \theta_1]$ to some $\hat{\theta}' \in [\hat{\theta}_1, \hat{\theta}_0]$. From theorem 1, it follows that $\eta(\theta') = \eta(\hat{\theta}')$ and $\eta(\theta)$ is increasing (decreasing) on $[\hat{\theta}_1, \hat{\theta}_0]$ (on $[\theta_0, \theta_1]$). The intuition for this structure is the following: Take types θ' and $\hat{\theta}'$ such that a non-local incentive constraint between θ' and $\hat{\theta}'$ is violated under the relaxed solution. Proposition 2 indicates that the decision of the types between $\hat{\theta}$ and θ' is increased to establish incentive compatibility. The usual optimization intuition suggests that it should be optimal to increase the decision for all those types by "the same amount."[13] However, this is not possible because of incentive compatibility constraints: Clearly, the decision of types $\theta' - \varepsilon$ cannot be increased discretely because of the monotonicity constraint at θ'. Lemma 3 establishes that the monotonicity constraint cannot even be binding for θ' as then the non-local constraint from $\theta' - \varepsilon$ to $\hat{\theta}'$ would be violated. Lemma 3 also makes clear that the decision should not jump at $\hat{\theta}'$ as otherwise the non-local constraint from θ' to $\hat{\theta}' + \varepsilon$ would be violated. One could now conjecture that non-local incentive constraints are binding from θ' not only to $\hat{\theta}'$ but also to slightly higher types and–with the same logic–from types slightly below θ' to $\hat{\theta}'$. However, it is not difficult to show that the incentive constraint between $\theta' - \varepsilon$ and $\hat{\theta}' + \varepsilon$ would be violated in this case. Consequently, one is left with the interval structure described above where non-local incentive constraints are binding from types slightly below θ' to types slightly above $\hat{\theta}'$.

The following lemma takes another perspective on the structure by establishing that non-local incentive constraints cannot bind at a finite number of interior types. With the additional properties established in the lemma, one should indeed expect the set of types where non-local incentive constraints bind to contain an interval.[14]

Lemma 4. *If the optimal solution is monotone and the relaxed solution is not implementable, non-local incentive constraints cannot bind only from a finite number of interior types to a finite number of interior types. Even stronger, the set of types from (to) which non-local incentive constraints bind cannot consist of isolated interior types.*[15]

The solution can be chosen such that (i) the set of types from which non-local incentive constraints are binding is closed and (ii) the set of types to which non-local incentive constraints are binding is closed.

Proof. see appendix

[13] Theorem 1 confirms this intuition by establishing that $\eta(\theta)$ is constant at types where non-local incentive constraints are lax.

[14] Strictly speaking, the lemma leaves the option that non-local incentive constraints are binding at a Cantor set of interior types. As the following results do not depend on this artificial looking case, I will ignore this possibility and speak of intervals in the remainder of the paper.

[15] Isolated means here that for each type θ from (to) which a non-local incentive constraint binds, there exists a neighborhood of θ in which non-local incentive constraints are lax for all types but θ.

Some of the properties of $\eta(\theta)$ in theorem 1 hold only at types where the decision is strictly increasing. The reason is that, the way (5) is written, $\eta(\theta)$ captures not only the effect of non-local incentive constraints but also the effect of the monotonicity constraint. If one wants to avoid this cluttering of effects, it is straightforward to introduce a monotonicity parameter $\nu(\theta)$ which captures the effect of the monotonicity constraint. In this case it is easy to see that the properties of $\eta(\theta)$ described in theorem 1 extend also to bunched types. Instead of (5) the solution would then be characterized by

$$\nu_\theta(\theta) = (u_q(q(\theta), \theta) - c_q(q(\theta), \theta))f(\theta) + (1 - F(\theta) - \eta(\theta))c_{q\theta}(q(\theta), \theta)$$

where $\nu(\theta)q_\theta(\theta) = 0$ for all $\theta \in \Theta$, i.e. $\nu(\theta)$ corresponds to the Lagrange parameter of the monotonicity constraint. If the start and ending type of a bunching interval are denoted by θ_s^b and θ_e^b, then obviously $\int_{\theta_s^b}^{\theta_e^b} \nu_\theta(\theta)\, d\theta = 0$. As described in the existing literature on ironing, see Guesnerie and Laffont (1984) or the exposition in Fudenberg and Tirole (1991), the bunching interval is characterized by this last condition and the endpoint conditions $\nu(\theta_s^b) = \nu(\theta_e^b) = 0$. The following lemma formalizes the discussion of the last paragraph.

Lemma 5. *If types in the interval (θ_s^b, θ_e^b) are bunched in the optimal solution, then there exists a function $\eta(\theta)$ which satisfies the properties of theorem 1 also for bunched types. In particular, $\eta(\theta)$ is non-decreasing on (θ_s^b, θ_e^b) and constant if no non-local incentive constraint binds to the bunched types. Furthermore, $\eta(\theta)$ satisfies (i) $\eta(\theta) = \eta(\hat{\theta})$ if $\Phi(\theta, \hat{\theta}) = 0$ and (C1') as well as (C2') hold, (ii) $\int_{\theta_s^b}^{\theta_e^b} \nu_\theta(\theta)\, d\theta = 0$ with $\nu_\theta(\theta)$ defined as above.*

Proof. see appendix

7. Continuous solutions

This section has two goals: First, to provide sufficient conditions under which a monotone solution is continuous and, second, to introduce an algorithm for determining such a continuous solution.

The first sufficient condition for continuity is loosely based on the idea of having a one-to-one relationship between η and q for a given type θ; i.e. the idea that for a given type θ and $\eta(\theta) > 0$, equation (5) yields a unique solution for q. The condition in the proposition ensures this and also ascertains that this relationship is monotonic, i.e. a higher $\eta(\theta)$ results in a higher q.

Proposition 3. *A monotone solution is continuous if*

$$\frac{u_{qq}(q, \theta) - c_{qq}(q, \theta)}{c_{qq\theta}(q, \theta)} > \frac{u_q(q, \theta) - c_q(q, \theta)}{c_{q\theta}(q, \theta)} \tag{6}$$

holds for all types and all $q \geq q^{fb}(\theta)$.[16]

Proof. see appendix

Hence, if the social objective $u(q, \theta) - c(q, \theta)$ is concave enough or if the cross derivative $c_{q\theta}(q, \theta)$ is in absolute value large enough (at the first best decision), the optimal decision will be continuous. Take for example the cost function in example 1 in section 4 and assume that $u(q, \theta) = \beta q$. It turns out that (6) is equivalent to the condition for $q^{fb}(\theta) > s(\theta)$, i.e. $\beta > 2\bar{\theta}$.[17]

The following proposition gives an alternative condition under which the optimal solution is below the first best decision. Having a solution below first best turns out to be sufficient for continuity and strict monotonicity of the solution (under a standard monotone hazard rate assumption). This is in itself remarkable. As the relaxed solution is below first best, one should expect the solution to be below first best whenever non-local incentive constraints are not violated "too much" by the relaxed solution. Hence, there is a broad class of problems in which the solution will be strictly monotone and continuous. Furthermore, the proof of the following proposition shows that the property holds also locally. That is, if the decision is below first best on some interval (θ_1, θ_2), then the decision will be strictly monotone and continuous on (θ_1, θ_2).

Before stating the proposition some additional notation is needed. Define $q^m(\theta)$ such that $c_\theta(q^{fb}(\theta), \theta) = c_\theta(q^m(\theta), \theta)$. Hence, $q^m(\theta)$ is a mirror image of $q^{fb}(\theta)$ along $s(\theta)$ with respect to $c_\theta(q, \theta)$.

Proposition 4. *Assume that $q^m(\theta)$ is non-decreasing and that there is no distortion at the top.*[18] *Then the optimal solution is below first best and continuous. The optimal solution is strictly increasing at all types where it is below first best if $f(\theta)/(1 - F(\theta))$ is non-decreasing and $u_{q\theta} \geq 0$.*

Proof. see appendix

One example for a class of function where $q^m(\theta)$ is increasing are cost functions of the form $c(q, \theta) = \theta q + \phi(q - \alpha\theta) + \gamma(\theta)$ where $\phi(\cdot)$ is a function of which the first three derivatives are positive.[19] Any increasing and concave benefit function $u(q, \theta)$ with $u_{q\theta} = 0$ and $q^{fb}(\theta) > s(\theta)$ yields an increasing $q^m(\theta)$.

Note that in many applications $u_{q\theta} = 0$ will hold. For example, in regulation models, labor market models and monopoly pricing, this property will typically hold because the principal's utility depends only on the decision and the transfer and not directly on the agent's type.

[16] Obviously, it is enough if the condition holds for all $q \in [q^{fb}(\theta), \bar{q}]$ where $\bar{q}$ is defined as in appendix C.

[17] In fact, this also holds true if q^2 in the cost function is replaced by any increasing and convex function.

[18] See the following section for a simple sufficient condition for no distortion at the top.

[19] The interpretation of this cost function is that there is a "normal scale" of $\alpha\theta$. Producing above this normal scale is increasingly costly. Type reflects a tradeoff between the size of the normal scale and marginal cost when producing within the normal scale.

Now it is time to turn to the issue of calculating a solution. In principle, the solution is already described by (5), the properties of $\eta(\theta)$ and the necessary conditions C1, C2 and C3. If a non-local incentive constraint binds from a type θ, the three necessary conditions could be used to determine $\hat{\theta}$, $q(\theta)$ and $q(\hat{\theta})$ (assuming that there is a unique solution). If non-local incentive constraints are lax at a type θ, (5) can be used to calculate $q(\theta)$ where $\eta(\theta)$ equals $\eta(\hat{\theta}')$ with $\hat{\theta}'$ being defined as the next lower type to which a non-local incentive constraint is binding. While nothing is wrong with this description, it might be burdensome to calculate a solution in this way. Hence, a more structured alternative to obtain a continuous solution might be helpful. This alternative will also give some additional insights into the logic behind the solution. The algorithm is based on the following proposition.

Proposition 5. *Define $\Phi^{\eta}(\theta,\hat{\theta})$ as $\Phi(\theta,\hat{\theta})$ under $\tilde{q}(\theta)$ where $\tilde{q}(\theta)$ is derived from*

$$\{u_q(q,\theta)-c_q(q,\theta)\}f(\theta)+(1-F(\theta)-\eta)c_{q\theta}(q,\theta)=0.$$

If the incentive constraint binds between θ' and $\hat{\theta}'$ in a continuous solution $q(\theta)$, then $(\theta',\hat{\theta}')$ minimize $\Phi^{\eta}(\theta,\hat{\theta})$ on $[\hat{\theta}',\theta']$ where $\eta=\eta(\theta')=\eta(\hat{\theta}')$. Furthermore, $\Phi^{\eta}(\theta',\hat{\theta}')<\Phi^{\eta}(\theta'',\hat{\theta}'')$ for any $\theta''>\theta'$ and $\hat{\theta}''<\hat{\theta}'$.

Proof. see appendix

To get a feeling for this proposition take $\eta=0$. Then $\tilde{q}(\theta)=q^r(\theta)$. Denote the global minimizer of $\Phi^0(\theta,\hat{\theta})$ by $(\theta^r,\hat{\theta}^r)$. Although a little extra work is needed, the following result follows almost directly from proposition 5:

Corollary 1. *If the relaxed solution is not implementable, the non-local incentive constraint from θ^r to $\hat{\theta}^r$ will bind in the optimal decision. If one of the two types (both) is interior, his (their) optimal decision is the relaxed decision; i.e. $q(\theta)=q^r(\theta)$ or (and) $q(\hat{\theta})=q^r(\hat{\theta})$ respectively.*

Proof. see appendix

The proposition then says that a similar logic applies for all pairs $(\theta',\hat{\theta}')$ at which incentive compatibility is binding: One only has to replace $q^r(\theta)$ in the corollary by the corresponding $\tilde{q}(\theta)$. This $\tilde{q}$ is the decision that would result if all types had the same $\eta(\theta)$ and this $\eta(\theta)$ would equal $\eta(\theta')$ in the optimal decision.

The last proposition in connection with theorem 1 gives a method for determining $q(\theta)$.

Solve (5) for q as a function of type θ and η. Plugging this $q(\theta,\eta)$ into $\Phi(\cdot)$ yields a function $\Phi^{\eta}(\theta,\hat{\theta})$ which can be minimized over θ and $\hat{\theta}$ yielding $\theta(\eta)$ and $\hat{\theta}(\eta)$ as minimizers. There could be several pairs $(\theta(\eta),\hat{\theta}(\eta))$ locally minimizing $\Phi^{\eta}(\theta,\hat{\theta})$. Relevant is each pair $(\theta,\hat{\theta})$ (i) that globally minimizes $\Phi^{\eta}(\cdot)$ on the interval $[\hat{\theta},\theta]$, (ii) for which no $\Phi^{\eta}(\cdot)$ minimizer $(\theta',\hat{\theta}')$ with $\theta'>\theta$, $\hat{\theta}'<\hat{\theta}$ and $\Phi^{\eta}(\theta',\hat{\theta}')<\Phi^{\eta}(\theta,\hat{\theta})$ exists. For now, assume there is only one such relevant pair.

Under the optimal decision, the constraint will bind from $\theta(\eta)$ to $\hat{\theta}(\eta)$ for all $\eta \in [0, \bar{\eta}]$ where $\bar{\eta}$ is determined by $\Phi^{\eta}(\theta(\eta), \hat{\theta}(\eta)) = 0$. The optimal decision for types θ where the constraint binds is given by $q(\theta, \eta)$ where η is such that $\theta = \theta(\eta)$. Types for which the constraint does not bind can be sorted into two categories: First, types θ such that non-local incentive constraints do not bind from any type above θ to any type below θ. These types simply have $q(\theta) = q^r(\theta)$. Second, types θ such that the constraint is binding from some $\theta' > \theta$ to some $\hat{\theta}' < \theta$. These types have $\eta(\theta)$ equal to $\eta(\inf\{\theta' : \Phi(\theta', \hat{\theta}') = 0 \text{ with } \theta' > \theta > \hat{\theta}'\})$, i.e. their η is the same as the one of the next lowest type to which a non-local incentive constraint binds. Their $q(\theta)$ is then $q(\theta, \eta(\theta))$.

One remark on the possibility that several relevant pairs $(\theta(\eta), \hat{\theta}(\eta))$ exist. For example, say there exist the pairs $(\theta_1(\eta), \hat{\theta}_1(\eta))$ and $(\theta_2(\eta), \hat{\theta}_2(\eta))$ both satisfying (i) and (ii) above. The non-local incentive constraint could in this case bind from an interval $[\theta_0, \theta_1]$ to the interval $[\hat{\theta}_1, \hat{\theta}_0]$ as well as from the interval $[\theta_2, \theta_3]$ to the interval $[\hat{\theta}_3, \hat{\theta}_2]$ where $\hat{\theta}_1 < \hat{\theta}_0 < \theta_0 < \theta_1 < \hat{\theta}_3 < \hat{\theta}_2 < \theta_2 < \theta_3$; see figure 4 for an illustration. Indeed one has to be a bit more precise in this case: There will be different $\bar{\eta}$ for the two "brackets" of binding incentive constraints. In this case $\eta(\theta)$ will not be single peaked. Hence, the algorithm will then be applied to the two brackets separately and nothing else changes.

A second remark has to be made with regard to bunching. Some types might have an ironed out solution. This solution is then not $q(\theta, \eta(\theta))$ as described above but an ironed out version of it. The condition for determining $\bar{\eta}$, i.e. $\Phi^{\eta}(\theta(\eta), \hat{\theta}(\eta)) = 0$ has to hold for the ironed out decision whenever ironing is relevant. If the monotone hazard rate holds and $u_{q\theta} \geq 0$, one does not have to worry about ironing as long as $\eta \leq 1 - F(\theta(\eta))$: This implies $q(\theta) \leq q^{fb}(\theta)$ for all types for which bunching could have been possible and the decision will be strictly increasing (see the proof of proposition 4).

The algorithm is illustrated with a numerical example in the following section.

8. Distortion at the top

If the non-local incentive constraint binds from $\bar{\theta}$, something unusual can happen. Recall that the necessary condition (C1) might hold with inequality at $\theta = \bar{\theta}$. It is therefore possible that non-local incentive constraints bind from $\bar{\theta}$ to several non-bunched $\hat{\theta}$ even if the solution is continuous. Note that this is impossible for interior types: For a given $q(\theta)$, (C1) and (C2) will uniquely determine $\hat{\theta}$ and $q(\hat{\theta})$.

Now consider the case where the non-local incentive constraint binds not only to several but to a mass of types $\hat{\theta}$ (or to $\underline{\theta}$ as will be shown below). Then the shadow value of the constraint $\eta(\theta)$ will be strictly positive and bounded away from 0 for types slightly below $\bar{\theta}$. Hence, these types have a decision $q(\theta)$ which is at least ε away from their relaxed decision $q^r(\theta)$ for some $\varepsilon > 0$. Obviously, the same has then to apply for $\bar{\theta}$ because of the monotonicity constraint. Put differently, $\eta(\bar{\theta}) > 0$ and therefore $q(\bar{\theta})$ is distorted: There is distortion at the top.

The algorithm described above works also in this situation. The minimizer $\theta(\eta)$ will then be the boundary type $\bar{\theta}$. The decision of $\bar{\theta}$ and his shadow value are determined by the highest $\hat{\theta}$ to which his non-local incentive constraint binds. At this $\hat{\theta}$ also condition (C1) holds with equality (if $\hat{\theta}$ is above $\underline{\theta}$).

It should be pointed out that distortion at the top is a generic property. Put differently, there will still be distortion at the top if, for example, the distribution of types is slightly perturbed. By proposition 5, distortion at the top implies that $\bar{\theta}$ will minimize $\Phi^\eta(\theta, \hat{\theta})$ for all $\eta < \tilde{\eta}$ for some $\tilde{\eta} > 0$. $\Phi^\eta(\theta, \hat{\theta})$ is continuous in $q(\theta, \eta)$ which in turn is continuous in the density $f(\theta)$. Therefore, $\bar{\theta}$ will remain global minimizer of $\Phi^\eta(\theta, \hat{\theta})$ under minor perturbations of the density. Consequently, distortion at the top has to be generic by proposition 5.

A natural question is whether there is a sufficient condition for no distortion at the top. Indeed corollary 1 allows to formulate such a condition. If $\bar{\theta}$ is not the global minimizer of $\Phi^r(\theta, \hat{\theta})$ where $\Phi^r(\cdot)$ is $\Phi(\cdot)$ under the relaxed solution $q^r(\cdot)$, then non local incentive constraints cannot bind from $\bar{\theta}$. Therefore, the relaxed decision is optimal for $\bar{\theta}$ implying that $q(\bar{\theta}) = q^{fb}(\bar{\theta})$.

Another sufficient condition for no distortion at the top can be formulated using (C1): $\int_0^{q^{fb}(\bar{\theta})} c_{q\theta}(q, \bar{\theta})\, dq \leq 0$ is sufficient since (C1) cannot hold with inequality.

To illustrate the distortion at the top result and also the algorithm introduced in the previous section, consider the following numerical example which is inspired by example 1 in section 2.[20]

The cost function is given by $c(q, \theta) = \theta q + \frac{q^2}{\theta} - \frac{\theta}{3}$. The principal's valuation function is $u(q) = \frac{8q}{5}$. Furthermore, I assume that types are distributed on $[1/4, 3/4]$ according to a triangular density with a "cushion" (to prevent $f(\theta) = 0$). I use the density $f(\theta) = 4/5(8\theta - 2)$. Recall from section 7 that with these parameter values the sufficient condition in proposition 3 is met. The solution will therefore be continuous.

The first order condition for the relaxed solution is

$$\left(\frac{8}{5} - \theta - \frac{2q}{\theta}\right) * \frac{4}{5}(8\theta - 2) + \frac{33 + 64\theta - 144\theta}{40}\left(1 - \frac{2q}{\theta^2}\right) = 0$$

which leads to the relaxed solution

$$q^r(\theta) = \frac{-347\theta^2 + 1660\theta^3 - 2444\theta^4}{330 + 1440\theta^2}.$$

[20] A *Mathematica* notebook with detailed calculations can be found under `https://www.sites.google.com/site/christophschottmueller/jmp`.

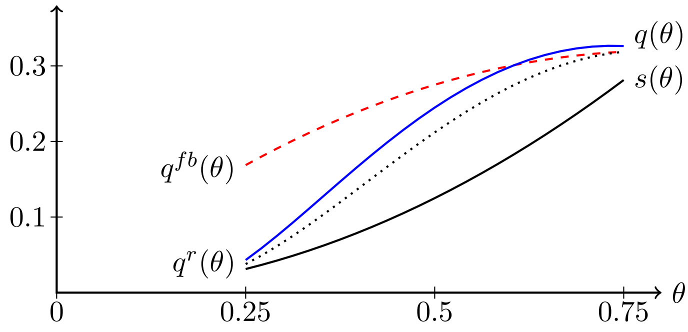


Figure 6: numerical example 1

To use the algorithm, $q(\theta,\eta)$ has to be calculated. In this example

$$q(\theta,\eta)=\frac{-2160\theta^4+2944\theta^3-347\theta^2-200\eta\theta^2}{330-400\eta+1440\theta^2}.$$

$\Phi^\eta(\theta,\hat{\theta})$ can be numerically minimized. The result is that $\bar{\theta}$ and $\underline{\theta}$ minimize $\Phi^\eta(\theta,\hat{\theta})$ for all $\eta \leq \bar{\eta} \approx 0.47298$. This means that a non-local incentive constraint is only binding from $\bar{\theta}$ to $\underline{\theta}$ and $\eta(\theta) = 0.47298$ for all types. Consequently, there is distortion at the top and the optimal decision is $q(\theta) = q(\theta, \bar{\eta})$ or

$$q(\theta)=\frac{-\frac{110399}{1250}\theta^2+\frac{2944}{5}\theta^3-432\theta^4}{\frac{17601}{625}+288\theta^2}.$$

Graphically, figure 6 shows that $q(\theta)$ (upper solid line) is above $q^r(\theta)$ (dotted line) for all types and that $q(\theta)$ is above $q^{fb}(\theta)$ (dashed line) for high types.

9. Discussion

This section discusses assumptions and compares the monotone solution with the solution of the standard screening model with single crossing and some related papers.

First, I want to discuss the assumptions on third derivatives, i.e. $c_{qq\theta} < 0$ and $c_{q\theta\theta} > 0$. The fact that these derivatives do not change sign ensures that the cross derivative $c_{q\theta}$ changes sign only once for any given θ (or q). While this property is admittedly important for the analysis, it is immaterial which sign the third derivatives have (as long as the sign is the same for all relevant decisions and types). To illustrate this (and also to show an example where the monotonicity constraint binds) consider the following version of example 2:[21] Types are distributed uniformly on $[2,3]$ and the principal's objective is the expected value of $q(\theta)-t(\theta)$. The agent's utility is given by

$$\pi(q,\theta)=t(\theta)-\frac{(q-\theta/\sigma)^2}{\theta^2}+\gamma(3-\theta).$$

[21] A *Mathematica* notebook with detailed calculations can be found under `https://www.sites.google.com/site/christophschottmueller/jmp`.

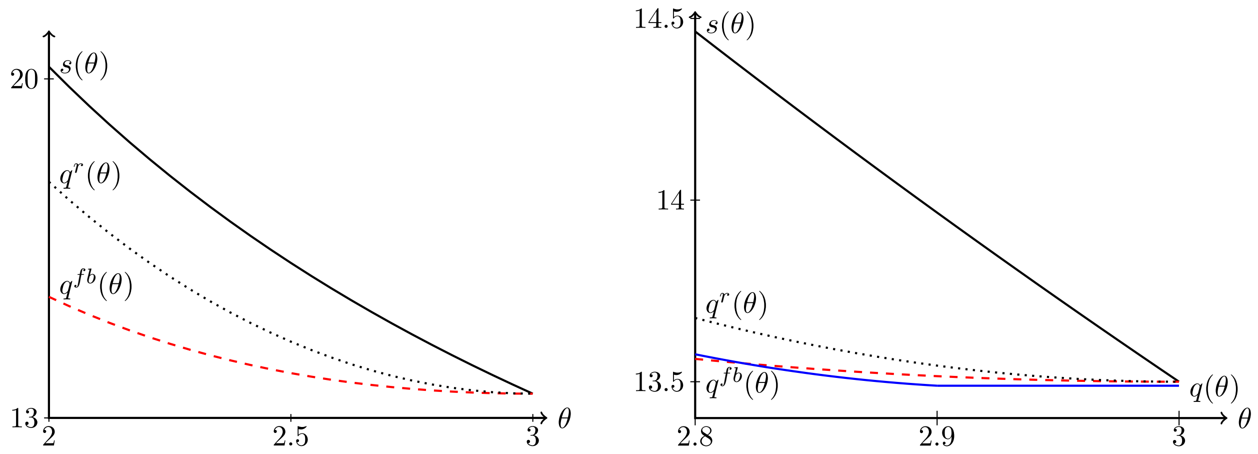


Figure 7: numerical example 2

Here the parameter values $\sigma = 27$ and $\gamma = 12$ are used. In this case, third derivatives have the following signs in the relevant range of the decision: $c_{qq\theta} < 0$ and $c_{q\theta\theta} < 0$. Consequently, the sign switching decision $s(\theta)$ is downward sloping. As depicted in figure 7a, first best decision and relaxed decision are also downward sloping.

Although the example looks different on first sight, it is equivalent to the model of the main text and all results apply accordingly. It turns out that also in this example $(\bar{\theta}, \underline{\theta})$ minimize $\Phi^{\eta}(\theta, \hat{\theta})$ and therefore only the non-local incentive constraint from the highest to the lowest type is binding. However, the monotonicity constraint is binding for the highest types. For each $q(\theta, \eta)$, the optimal bunching interval $[\theta_s(\eta), \bar{\theta}]$ is determined by the condition

$$\int_{\theta_s(\eta)}^{\bar{\theta}} [u_q(q(\theta,\eta),\theta) - c_q(q(\theta,\eta),\theta)]f(\theta) + (1 - F(\theta) - \eta(\theta))c_{q\theta}(q(\theta,\eta),\theta)\, d\theta = 0.$$

Here, $\bar{\eta}$ turns out to be approximately 0.18 and the solution for the highest types is depicted in figure 7b. The solution exhibits bunching of types in $[2.9, 3]$.

Second, I want to compare the obtained solution with solutions of screening models with single crossing. Such a comparison will pin down those effects which can only be explained by a violation of single crossing. In the standard textbook model with single crossing, see for example Fudenberg and Tirole (1991) or Bolton and Dewatripont (2005), decisions are downward distorted for rent extraction reasons. The solution is continuous and under some regularity conditions, e.g. monotone hazard rate, strictly increasing. This paper shows that a violation of single crossing can lead to a reduction of distortion and even to decision levels above first best. The reason is that binding non-local incentive constraints distort the decision upwards while binding local incentive constraints distort it downwards. The underlying cause is the one time violation of single crossing: A high type misrepresenting as a low type can have higher marginal cost at the low type's decision (this is impossible with single crossing). To make the decision of the low type less attractive for the high type it is then best to increase the low type's decision. By increasing also the decisions of the types in between, the slope of the rent function is increased. Consequently, the high type gets a higher rent at his own contract which also prevents misrepresentation.

Even with the monotone hazard rate assumption bunching can occur if the decision of some types is distorted sufficiently above first best. In contrast to the standard model, a violation of single crossing can lead to distortions at the top. Distortion at the top will occur if non-local incentive constraints bind from the best type to a mass of types or to the lowest type.

Jumps and bunching can also be part of the standard model if one allows for arbitrary type distributions as in Hellwig (2010). However, this will not lead to decisions above first best. Furthermore, a no distortion at the top result remains valid in Hellwigs's model. The reason is that with single crossing only local incentive constraints bind while non-local incentive constraints remain lax.

The reader familiar with the literature on adverse selection models might have noticed the similarity between the "first order condition"

$$\{u_q(q,\theta) - c_q(q(\theta),\theta)\}f(\theta) + (1 - F(\theta) - \eta(\theta))c_{q\theta}(q(\theta),\theta) = 0$$

and the first order condition in Jullien (2000). In Jullien's paper type dependent participation constraints are analyzed in a framework with single crossing. If one writes $\gamma(\theta)$ instead of $1 - \eta(\theta)$ in the condition above, the first order condition of his model results. There $\gamma(\theta)$ is the Lagrange parameter denoting the shadow value of relaxing the participation constraint for all types below θ.

A technical difference is that $\gamma(\theta)$ is monotonically increasing while $\eta(\theta)$ is first in- and later decreasing. Intuitively, one can start thinking from the relaxed decision. If a participation constraint is violated in the interior at type θ', the response is to reduce the distortion for all types below θ'. This will increase the slope of the profit function for all types below θ' and therefore increase the payoffs of θ'. If, on the other hand, the non-local incentive constraint is violated between two types $\hat{\theta}', \theta'$ under the relaxed decision, there is no reason to change the decision of types below $\hat{\theta}'$. The problem is solved by increasing the decision only for types between $\hat{\theta}'$ and θ'.

The overproduction result, i.e. q above first best, can occur with type dependent participation constraints as well. It can even occur at the highest type, so there can be distortion at the top. However, with type dependent participation constraints this peculiarity is caused by upward binding incentive constraints, i.e. low types want to misrepresent as high types. With violations of single crossing, the same results is obtained although incentive constraints are only downward binding.

Although the model is the same, it is not straightforward to compare the optimal solution obtained in this paper with the one in Araujo and Moreira (2010). Both, the monotone and the inversely U-shaped solution, are closer to first best than the relaxed solution (and might even cross first best). In contrast to this paper there is a no distortion at the top result in Araujo and Moreira (2010): The type with the highest first best decision, i.e. the type where $q^{fb}(\theta)$ crosses $s(\theta)$, will be assigned his first best decision in the optimal solution. Another difference is that the monotone solution can be continuous without bunching intervals of types. This difference is partly due to the direction non-local incentive constraints bind: In the monotone solution they bind only downward while they bind in both directions in an inversely U-shaped solution.

10. Conclusion

This paper characterizes monotone solutions in a screening environment where single crossing is violated. Although the model restricts itself to a one time violation of single crossing, the main effects of a violation of single crossing can be illustrated. Non-local incentive constraints can become binding. The distortion caused by non-locally binding incentive constraints can counteract the normal rent extraction distortion. Therefore, the solution can be partly above as well as below the first best decision. There can be distortion at the top if non local incentive constraints are binding from the top type to a mass of types (or the lowest type). Furthermore, sufficient conditions for monotonicity and continuity are provided and an algorithm for determining such a continuous, monotone solution is proposed.

Possible applications can be found in various fields of economics. While the paper uses the notation of a regulation or procurement setting, the same model is applicable, for example, in models of labor, insurance, monopoly pricing or optimal taxation. The characterization of continuous and monotone solutions is relatively simple and reasonable classes of functions satisfy sufficient conditions for falling into this class of solutions.

I conclude with some immediate implications of the qualitative results in this paper. In optimal taxation models where single crossing is violated negative marginal tax rates for top incomes can be rationalized because of the distortion at the top result. Note that distortion at the top always is in an "unusual" direction, i.e. above first best. The rough intuition is that subsidizing productive types to work more increases their rent and therefore relaxes their incentive compatibility constraint.

Overinsurance can be optimal in insurance models where single crossing is violated. This gives an alternative explanation for so called "Cadillac" insurance plans. While the political debate focuses on viewing them as (insufficiently taxed) part of a compensation package, screening by insurers with market power could also explain parts of the phenomenon.

Concerning the regulation example, it was mentioned in example 1 that the estimation results in Beard et al. (1991) provide evidence of a violation of single crossing in the cost functions of savings and loan associations. My results show that optimal regulation might induce a subset of such associations to offer more loans than first best optimal.

In Martimort and Stole (2009) the ordering of first best quantities and the competitive menu under substitutes is no longer clear cut if one considers the cases without single crossing. Put differently, firms using non-linear pricing might optimally offer packages which lead to overconsumption of the good. Telecommunication might be an example for this: Consumers often buy packages where an additional unit of calling (or internet use) is for free. If the marginal costs of the provider are only ε above zero, such a price scheme will lead to consumption above the socially optimal consumption.[22]

[22]Of course, there are alternative explanation based on the theory of two-sided markets. However, the two explanations are not mutually exclusive and, for example, two-sidedness is less obvious in case of internet access.

References

Akerlof, G. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics 84*(3), 488–500.

Araujo, A. and H. Moreira (2010). Adverse selection problems without the spence-mirrlees condition. *Journal of Economic Theory 145*(5), 1113–1141.

Armstrong, M. (1996). Multiproduct nonlinear pricing. *Econometrica 64*(1), 51–75.

Baron, D. and R. Myerson (1982). Regulating a monopolist with unknown costs. *Econometrica 50*(4), 911–930.

Baumol, W. J., J. C. Panzar, and R. D. Willig (1982). *Constestable markets and the theory of industry structure.* New York: Harcourt Brace Jovanovich, Inc.

Beard, T., S. Caudill, and D. Gropper (1991). Finite mixture estimation of multiproduct cost functions. *The Review of Economics and Statistics*, 654–664.

Billingsley, P. (1986). *Probability and measure.* New York: Wiley, J.

Bolton, P. and M. Dewatripont (2005). *Contract theory.* Cambridge: The MIT Press.

Boone, J. and C. Schottmüller (2011). Health insurance without single crossing: Why healthy people have high coverage. *CEPR Discussion Paper no. 8501*.

Fudenberg, D. and J. Tirole (1991). *Game Theory.* Cambridge: MIT Press.

Guasch, J. L. and A. Weiss (1981). Self-selection in the labor market. *American Economic Review 71*(3), 275–284.

Guesnerie, R. and J.-J. Laffont (1984). A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm. *Journal of Public Economics 25*(3), 329 – 369.

Harsanyi, J. (1967). Games with incomplete information played by "bayesian" players, i-iii. part i. the basic model. *Management Science 14*(3), 159–182.

Hellwig, M. (2010). Incentive problems with unidimensional hidden characteristics: A unified approach. *Econometrica 78*(4), 1201–1237.

Jullien, B. (2000). Participation constraints in adverse selection models. *Journal of Economic Theory 93*(1), 1–47.

Laffont, J.-J. and J. Tirole (1987). Auctioning incentive contracts. *Journal of Political Economy 95*(5), 921–937.

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

Martimort, D. and L. Stole (2009). Market participation in delegated and intrinsic common-agency games. *RAND Journal of Economics 40*(1), 78–102.

Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies 38*(2), 175–208.

Mussa, M. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic Theory 18*(2), 301–317.

Myerson, R. (1981). Optimal auction design. *Mathematics of Operations Research 6*, 58–73.

Netzer, N. and F. Scheuer (2010). Competitive screening in insurance markets with endogenous wealth heterogeneity. *Economic Theory 44*(2), 187–211.

Rochet, J. and P. Chone (1998). Ironing, sweeping, and multidimensional screening. *Econometrica 66*(4), 783–826.

Rochet, J. and L. Stole (2003). The economics of multidimensional screening. In *Advances in economics and econometrics: Theory and applications, Eighth World Congress*, Volume 1, pp. 150–197. Cambridge University Press.

Smart, M. (2000). Competitive insurance markets with two unobservables. *International Economic Review 41*(1), 153–170.

Stiglitz, J. (1977). Monopoly, non-linear pricing and imperfect information: the insurance market. *Review of Economic Studies 44*(3), 407–430.

Wambach, A. (2000). Introducing heterogeneity in the rothschild-stiglitz model. *Journal of Risk and Insurance 67*(4), 579–591.

Appendix

A. Variational condition

In Araujo and Moreira (2010), it always holds that $q(\theta) = q(\hat{\theta})$ whenever $\Phi(\theta, \hat{\theta}) = 0$. Consequently, (C1) does not play a role. Starting from (C2), they derive the following condition (with $q = q(\theta) = q(\hat{\theta})$):

$$\frac{u_q(q,\theta) - c_q(q,\theta) + \frac{1-F(\theta)}{f(\theta)} c_{q\theta}(q,\theta)}{c_{q\theta}(q,\theta)} f(\theta) = \frac{u_q(q,\hat{\theta}) - c_q(q,\hat{\theta}) + \frac{1-F(\hat{\theta})}{f(\hat{\theta})} c_{q\theta}(q,\hat{\theta})}{c_{q\theta}(q,\hat{\theta})} f(\hat{\theta}) \tag{7}$$

To derive a similar condition for $q(\theta) \neq q(\hat{\theta})$ take θ and $\hat{\theta}$ such that $c_q(q(\hat{\theta}), \hat{\theta}) = c_q(q(\hat{\theta}), \theta)$, $c_\theta(q(\theta), \theta) = c_\theta(q(\hat{\theta}), \theta)$, $\Phi(\theta, \hat{\theta}) = 0$ and assume that $q(\cdot)$ is strictly monotone and continuous at θ and $\hat{\theta}$.

Given θ and $q(\theta)$, the equation $c_\theta(q(\theta), \theta) = c_\theta(q(\hat{\theta}), \theta)$ pins down a decision $q(\hat{\theta})$ where incentive compatibility could be binding. Given this $q(\hat{\theta})$ as well as θ and $q(\theta)$, the equation $c_q(q(\hat{\theta}), \hat{\theta}) = c_q(q(\hat{\theta}), \theta)$ determines $\hat{\theta}$. Therefore, the critical $\hat{\theta}$ can be written as a function of θ and $q(\theta)$, i.e. $\hat{\theta} = \phi(\theta, q(\theta))$.

Differentiating the two conditions, the partial derivatives ϕ_θ and ϕ_q can be obtained as

$$\begin{aligned}
\phi_\theta(\theta, q) &= \frac{c_{q\theta}(\hat{q}, \theta)}{c_{q\theta}(\hat{q}, \hat{\theta})} + \frac{(c_{qq}(\hat{q}, \theta) - c_{qq}(\hat{q}, \hat{\theta}))(c_{\theta\theta}(q, \theta) - c_{\theta\theta}(\hat{q}, \theta))}{c_{q\theta}(\hat{q}, \hat{\theta}) c_{q\theta}(\hat{q}, \theta)} \\
\phi_q(\theta, q) &= \frac{c_{q\theta}(q, \theta)[c_{qq}(\hat{q}, \theta) - c_{qq}(\hat{q}, \hat{\theta})]}{c_{q\theta}(\hat{q}, \hat{\theta}) c_{q\theta}(\hat{q}, \theta)}
\end{aligned}$$

where $\hat{q} = q(\hat{\theta})$ and $q = q(\theta)$.

Denote by h an admissible perturbation of the optimal solution q^* on some interval $[\theta_1, \theta_2]$, i.e. $h(\theta_1) = h(\theta_2) = 0$. Admissibility implies that if the incentive constraint binds from θ to $\hat{\theta}$, then $\hat{\theta} = \phi(\theta, q(\theta))$.[23]

The idea of the variational argument is the following: I want to derive a necessary condition for a type θ such that $\Phi(\theta, \hat{\theta}) = 0$ for some $\hat{\theta}$. To do so, it is assumed that also under the perturbed decision the incentive constraint is binding for θ and some (other) $\hat{\theta}$. The type $\hat{\theta}$ to which the non-local incentive constraint binds depends on the perturbation and is given by $\phi(\theta, q(\theta))$. The way one should think about it is that incentive compatibility is binding from each $\theta \in [\theta_1, \theta_2]$ to some $\hat{\theta}$ in some interval $[\hat{\theta}_1, \hat{\theta}_2]$.[24] The specific type $\hat{\theta}$ to which a non-local incentive constraint binds from a given θ depends on the perturbation h.

[23] Furthermore, admissibility requires monotonicity.

[24] As it turns out, this is indeed the typical structure of a continuous solution, see lemma 4.

For brevity, I denote in the remainder of this section the optimal solution by $q^*(\theta)$ and the perturbed solution by $q(\theta) = q^*(\theta) + \varepsilon h(\theta)$. Hence the part of the principal's objective function affected by the perturbation can be written as[25]

$$\begin{aligned} G(\varepsilon) &= \int_{\theta_1}^{\theta_2} g(q(\theta),\theta)\,d\theta + \int_{\phi(\theta_2,q(\theta_2))}^{\phi(\theta_1,q(\theta_1))} g(q(\theta),\theta)\,d\theta \\ &= \int_{\theta_1}^{\theta_2} \{g(q(\theta),\theta) - g(\hat{q}(\theta,q(\theta)),\phi(\theta,q(\theta)))\,[\phi_q(q(\theta),\theta)q_\theta(\theta) + \phi_\theta(q(\theta),\theta)]\}\,d\theta \end{aligned} \tag{8}$$

where $g(q(\theta),\theta) = \left[u(q(\theta),\theta) - c(q(\theta),\theta) + \frac{1-F(\theta)}{f(\theta)}c_\theta(q(\theta),\theta)\right] f(\theta)$ is the virtual valuation weighted by the density. The second line is a normal change of variables where $\hat{q}(\theta,q)$ denotes the $\hat{q}$ solving $c_\theta(q,\theta) = c_\theta(\hat{q},\theta)$ with $q \neq \hat{q}$. Note that $\partial\hat{q}/\partial q = c_{q\theta}(q,\theta)/c_{q\theta}(\hat{q},\theta)$.

Differentiating (8) gives

$$G'(0) = \int_{\theta_1}^{\theta_2} \{g_q h - \hat{g}((\phi_{qq}q_\theta^* + \phi_{q\theta})h + \phi_q h_\theta) - (\hat{g}_q\hat{q}_q + \hat{g}_\theta\phi_q)(\phi_q q_\theta^* + \phi_\theta)h\}\,d\theta = 0$$

where arguments are omitted and a hat denotes evaluation at $(\hat{\theta}, q^*(\hat{\theta}))$. Integrating $\int_{\theta_1}^{\theta_2}(\hat{g}\phi_q)h_\theta\,d\theta$ by parts and substituting yields for the previous equation

$$\int_{\theta_1}^{\theta_2} \{g_q - \hat{g}_q\hat{q}_q\phi_\theta + \hat{g}_q\hat{q}_\theta\phi_q\}\,h\,d\theta = \int_{\theta_1}^{\theta_2} \left\{g_q - \hat{g}_q\frac{c_{q\theta}(q(\theta),\theta)}{c_{q\theta}(q(\hat{\theta}),\hat{\theta})}\right\} h\,d\theta = 0.$$

As h was arbitrary, the following condition has to hold at optimum:

$$g_q(q(\theta),\theta) = g_q(q(\hat{\theta}),\hat{\theta})\frac{c_{q\theta}(q(\theta),\theta)}{c_{q\theta}(q(\hat{\theta}),\hat{\theta})} \tag{C3'}$$

This is condition (C3). For $q(\theta) = q(\hat{\theta})$, (C3') boils down to (7).

B. Proofs

Proof of proposition 1: First, it is shown that the principal's payoff is higher under $q^c(\theta)$ than under $q(\theta)$: The principal maximizes expectation of $u(q,\theta) - c(q,\theta) + (1 - F(\theta))/f(\theta)c_\theta(q,\theta)$. If $q^s(q,\theta) \leq q^r(\theta)$, the principal's objective increases due to the change because of the concavity of (RP) and $q^r(\theta) > s(\theta)$. If $q^s(q(\theta),\theta) > q^{fb}(\theta)$, then the same conclusion follows from $q^v(q(\theta),\theta) \geq q^s(q(\theta),\theta) > q^r(\theta)$ and the concavity of (RP).

Second, the changed decision $q^c(\theta)$ is monotonically increasing: From local incentive compatibility $q(\theta)$ was already increasing wherever it was above $s(\theta)$. At types with $q(\theta) < s(\theta)$ the decision $q(\theta)$ had to be decreasing because of local incentive compatibility. But then $q^s(q(\theta),\theta)$ is clearly increasing

[25]It follows from lemma 2 that $\phi(\theta_1, q(\theta_1)) > \phi(\theta_2, q(\theta_2))$.

in θ for these types because of $c_{q\theta\theta} > 0$. This leaves types at which $q(\theta)$ jumped discontinuously over $s(\theta)$. But at these jump types local incentive compatibility required $c_\theta(q^-(\theta), \theta) - c_\theta(q^+(\theta), \theta) \geq 0$ at downwards jumps (and the converse inequality at upwards jumps) across $s(\theta)$. This implies that also at jump points of $q(\theta)$ monotonicity of $q^c(\theta)$ is guaranteed.

Third, the changed decision $q^c(\theta)$ is incentive compatible: Since $q^c(\theta)$ is monotonically increasing, only downward misrepresentation has to be considered (see lemma 1). Note that the profit function $\pi(\theta)$ was not affected by the change from $q(\theta)$ to $q^c(\theta)$ because of the definition of $q^s(\theta)$ and $\pi_\theta(\theta) = -c_\theta(q(\theta), \theta)$ by local incentive compatibility. Therefore, one has only to check whether any type wants to misrepresent as a lower type $\hat{\theta}$ at which $q(\hat{\theta}) < s(\hat{\theta})$. Since $\pi(\theta)$ is unchanged, one can write incentive compatibility under the changed decision as

$$\begin{aligned}\Phi^c(\theta, \hat{\theta}) &= -\int_{\hat{\theta}}^{\theta} \int_{q^c(\hat{\theta})}^{q(t)} c_{q\theta}(q,t)\, dq\, dt = -\int_{\hat{\theta}}^{\theta} \int_{q^c(\hat{\theta})}^{q(\hat{\theta})} c_{q\theta}(q,t)\, dq\, dt - \int_{\hat{\theta}}^{\theta} \int_{q(\hat{\theta})}^{q(t)} c_{q\theta}(q,t)\, dq\, dt \\ &= \int_{\hat{\theta}}^{\theta} \int_{q(\hat{\theta})}^{q^c(\hat{\theta})} c_{q\theta}(q,t)\, dq\, dt + \Phi(\theta, \hat{\theta}) > 0\end{aligned}$$

where the inequality follows from $\int_{q(\hat{\theta})}^{q^c(\hat{\theta})} c_{q\theta}(q, \hat{\theta})\, dq = 0$ by the definition of $q^s(\cdot)$ and $c_{q\theta\theta} > 0$. □

Proof of lemma 3: First, it is shown that there cannot be a discontinuity at $\hat{\theta}$. Take a type $\hat{\theta}$ to which non-local incentive constraint is binding from some type θ. Suppose that $q(\cdot)$ is discontinuous at $\hat{\theta}$, i.e. $q^-(\hat{\theta}) < q^+(\hat{\theta})$ by local incentive compatibility (monotonicity). Binding incentive constraint means that either (i) $\int_{\hat{\theta}}^{\theta} \int_{q^-(\hat{\theta})}^{q(t)} c_{q\theta}(q,t)\, dq\, dt = 0$ or (ii) $\int_{\hat{\theta}}^{\theta} \int_{q^+(\hat{\theta})}^{q(t)} c_{q\theta}(q,t)\, dq\, dt = 0$ or (iii) $q^-(\hat{\theta}) < q(\hat{\theta}) < q^+(\hat{\theta})$ and $\int_{\hat{\theta}}^{\theta} \int_{q(\hat{\theta})}^{q(t)} c_{q\theta}(q,t)\, dq\, dt = 0$.

In case (i) it must hold that $\int_{\hat{\theta}}^{\theta} c_{q\theta}(q^-(\hat{\theta}), t)\, dt \leq 0$ which is just (C2) adapted to apply for a right hand side discontinuity, i.e. if this did not hold incentive compatibility would be violated for θ and $\hat{\theta} - \varepsilon$. But then $\int_{\hat{\theta}}^{\theta} \int_{q^-(\hat{\theta})}^{q^+(\hat{\theta})} c_{q\theta}(q,t)\, dq\, dt < 0$ from $c_{qq\theta} < 0$. Hence, $\Phi(\theta, \hat{\theta}^+) = \Phi(\theta, \hat{\theta}^-) + \int_{\hat{\theta}}^{\theta} \int_{q^-(\hat{\theta})}^{q^+(\hat{\theta})} c_{q\theta}(q,t)\, dq\, dt < 0$ as $\Phi(\theta, \hat{\theta}^-) = 0$ by assumption. Hence, incentive compatibility is violated from θ to types slightly above $\hat{\theta}$. This is the desired contradiction.

In case (ii) it must hold that $\int_{\hat{\theta}}^{\theta} c_{q\theta}(q^+(\hat{\theta}), t)\, dt \geq 0$. But then $\int_{\hat{\theta}}^{\theta} \int_{q^-(\hat{\theta})}^{q^+(\hat{\theta})} c_{q\theta}(q,t)\, dq\, dt > 0$ from $c_{qq\theta} < 0$. Consequently, $\Phi(\theta, \hat{\theta}^-) = \Phi(\theta, \hat{\theta}^+) - \int_{\hat{\theta}}^{\theta} \int_{q^-(\hat{\theta})}^{q^+(\hat{\theta})} c_{q\theta}(q,t)\, dq\, dt < 0$ and therefore incentive compatibility is violated from θ to types slightly below $\hat{\theta}$.

In case (iii) the same arguments as in case (i) apply if $\int_{\hat{\theta}}^{\theta} c_{q\theta}(q(\hat{\theta}), t)\, dt \leq 0$ while the same arguments as in case (ii) apply if $\int_{\hat{\theta}}^{\theta} c_{q\theta}(q(\hat{\theta}), t)\, dt > 0$.

Second, it is shown that $\theta < \bar{\theta}$ cannot be bunched with some type θ' if $q(\cdot)$ is continuous at θ. Suppose θ and θ' were bunched on q^b (and by monotonicity all types in between them are as well) and suppose for now $\theta < \theta'$. But then $\Phi(\theta', \hat{\theta}) < 0$ and ic is violated as

$$\begin{aligned}\Phi(\theta', \hat{\theta}) &= -\int_{\hat{\theta}}^{\theta'} \int_{q(\hat{\theta})}^{q(t)} c_{s,t}\, ds\, dt = -\int_{\hat{\theta}}^{\theta} \int_{q(\hat{\theta})}^{q(t)} c_{s,t}\, ds\, dt - \int_{\theta}^{\theta} \int_{q(\hat{\theta})}^{q(t)} c_{s,t}\, ds\, dt \\ &= \Phi(\theta, \hat{\theta}) - \int_{\theta}^{\theta'} \int_{q(\hat{\theta})}^{q^b} c_{s,t}\, ds\, dt < 0\end{aligned}$$

where the last inequality follows from (C1) and $c_{q\theta\theta} > 0$.

Now suppose $\theta > \theta'$ and both types are bunched. From condition (C1) for $\theta < \bar{\theta}$ and $c_{q\theta\theta} > 0$ it follows that $\int_{q(\hat{\theta})}^{q(t)} c_{q\theta}(q,t)\, dq < 0$ for every $t \in (\theta - \varepsilon, \theta)$. But then $\Phi(\theta - \varepsilon, \hat{\theta}) = \Phi(\theta, \hat{\theta}) + \int_{\theta-\varepsilon}^{\theta} \int_{q(\hat{\theta})}^{q(t)} c_{q\theta}(q,t)\, dq\, dt < 0$, so incentive compatibility would be violated. □

Proof of proposition 2: Suppose $q(\theta) < q^r(\theta)$ for some types. Since local incentive compatibility does not allow downward jumps, $q(\theta)$ has to be strictly below $q^r(\theta)$ for a mass of types. Consider changing this 'optimal' decision to $q^*(\theta)$ where $q^*(\theta) = max\{q(\theta), q^r(\theta)\}$. Transfers $t^*(\theta)$ are determined such that $\pi(\underline{\theta}) = 0$ and $\pi_\theta(\theta) = -c_\theta(q^*(\theta), \theta)$.

By the definition of $q^r(\theta)$, this change will increase the principal's expected payoff.

It remains to check incentive compatibility, i.e

$$\Phi^*(\theta, \hat{\theta}) = -\int_{\hat{\theta}}^{\theta} \int_{q^*(\hat{\theta})}^{q^*(t)} c_{qt}(q,t)\; dq\, dt \geq 0$$

for arbitrary types θ and $\hat{\theta} < \theta$. If $q^*(\hat{\theta}) = q(\hat{\theta})$, incentive compatibility follows from $q^*(t) \geq q(t)$ and as $q(t) \geq s(t)$ the corresponding 'additional' $c(q,t)$ are negative.

If $q^*(\hat{\theta}) > q(\hat{\theta})$ (and therefore $q^*(\hat{\theta}) = q^r(\hat{\theta})$), there are three possibilities: (i) There exists a type $\theta' \in (\hat{\theta}, \theta)$ with $q(\theta') = q^*(\hat{\theta})$, (ii) all types $\theta' \in (\hat{\theta}, \theta)$ have $q(\theta') < q^*(\hat{\theta})$ and (iii) there are types $\theta' \in (\hat{\theta}, \theta)$ with $q(\theta') > q^*(\hat{\theta})$ but no type θ' with $q(\theta') = q^*(\hat{\theta})$, hence $q(\cdot)$ is discontinuous[26].

If (i), then $\Phi(\theta, \theta') \geq 0$ implies incentive compatibility as $\Phi^*(\theta, \hat{\theta}) > \Phi(\theta, \theta')$. In case (ii) $q^*(\hat{\theta})$ has to be above $q(\theta')$ for all $\theta' \in (\hat{\theta}, \theta)$. But since $q(\theta') > s(\theta')$ for all these types it follows that $q^*(\hat{\theta}) > s(\theta)$ and therefore incentive compatibility is trivially satisfied.

In case (iii) define $\theta' = sup\{t \in (\hat{\theta}, \theta) :\; q(t) < q^*(\hat{\theta})\}$ that is θ' is the jump point. Incentive compatibility between θ and θ' implies $\int_{\theta'}^{\theta} \int_{q^-(\theta')}^{q(t)} c_{qt}(q,t)\; dq\, dt \leq 0$ as well as $\int_{\theta'}^{\theta} \int_{q^+(\theta')}^{q(t)} c_{qt}(q,t)\; dq\, dt \leq 0$ where $q^-(\theta')$ denotes the limit of $q(t)$ as $t \to \theta'$ from below. From $c_{qq\theta} < 0$ and $q^-(\theta') < q^*(\hat{\theta}) < q^+(\theta')$, it follows that $\int_{\theta'}^{\theta} \int_{q^*(\hat{\theta})}^{q(t)} c_{qt}(q,t)\; dq\, dt \leq 0$. But as $\Phi^*(\theta, \hat{\theta}) > -\int_{\theta'}^{\theta} \int_{q^*(\hat{\theta})}^{q(t)} c_{qt}(q,t)\; dq\, dt \geq 0$ incentive compatibility is satisfied. □

Proof of theorem 1: Note that even if the theorem was not true one could define a function $\eta(\theta)$ by rearranging (5). What one has to show are the properties of this function. $\eta(\theta) \geq 0$ follows immediately from proposition 2 and the fact that the left hand side of (5) is decreasing in q.

[26] Given that solutions in Araujo and Moreira (2010) display sometimes discontinuities, one cannot totally exclude this possibility.

Next turn the property that $\eta(\theta)$ is constant on an interval of types on which non local incentive constraints are lax. Suppose to the contrary that $\eta(\theta)$ is not constant. In particular suppose $\eta(\theta)$ was increasing on some interval $[\theta_1, \theta_3]$ where non-local ic is lax for all $\theta \in [\theta_1, \theta_3]$. Denote by θ_2 some interior type of the interval. For each $\theta \in [\theta_2, \theta_2 + \varepsilon]$ define a corresponding type $\theta' \in [\theta_2 - \varepsilon, \theta_2]$ by $\theta' = \theta_2 - (\theta - \theta_2)$ for some small $\varepsilon > 0$. I will show that one can change such a decision on $[\theta_2 - \varepsilon, \theta_2 + \varepsilon]$ in a way which increases the principal's payoff (while keeping incentive compatibility). This contradicts the optimality of $q(\theta)$.

Consider a changed decision $q^c(\cdot)$ such that (i) $q^c(\theta) > q(\theta)$ on $[\theta_2 - \varepsilon, \theta_2)$, (ii) $q^c(\theta) \leq q(\theta)$ on $[\theta_2, \theta_2 + \varepsilon]$, (iii) for corresponding types θ and θ' it holds that $\int_{q(\theta')}^{q^c(\theta')} c_{q\theta}(q, \theta')\, dq = - \int_{q(\theta)}^{q^c(\theta)} c_{q\theta}(q, \theta)\, dq$ and (iv) $q^c_\theta(\theta) \geq 0$ on $[\theta_2 - \varepsilon, \theta_2 + \varepsilon]$. The changed decision will therefore display upwards jumps at $\theta_2 - \varepsilon$ and $\theta_2 + \varepsilon$. For small changes in q (iii) can be written as $\delta(\theta')c_{q\theta}(q(\theta'), \theta') = -\delta(\theta)c_{q\theta}(q(\theta), \theta)$ where $\delta(\theta) = q^c(\theta) - q(\theta)$. This in turn can be written as $\delta(\theta') = -\delta(\theta)k(\theta)$ where $k(\theta)$ is defined as $\frac{c_{q\theta}(q(\theta),\theta)}{c_{q\theta}(q(\theta'(\theta)),\theta'(\theta))}$.

Before proceeding, let me show that a function $q^c(\theta)$ satisfying (i)-(iv) exists. Note that $k(\theta_2) = 1$ and that–due to the differentiability and continuity assumptions on $c(\cdot)$ and the monotonicity of $q(\theta)$– the function $k(\theta)$ is continuously differentiable almost everywhere.[27] First, consider the case where $k^+_\theta(\theta_2) < 0$. Then it is feasible to set $q^c(\theta) = q(\theta_2)$ for types $\theta \in [\theta_2, \theta_2 + \varepsilon]$ if $\varepsilon > 0$ is chosen small enough. Feasibility means that determining $q(\theta')$ by $\delta(\theta') = -\delta(\theta)k(\theta)$ will satisfy all conditions especially (iv). Feasibility of $q^c(\theta) = q(\theta_2)$ for $\theta \in [\theta_2, \theta_2 + \varepsilon]$ and monotonicity of $q(\theta)$ imply that $q^{c*} = \alpha q^c(\theta) + (1 - \alpha)q(\theta)$ is also feasible. The effect of a marginal change of q is the effect changing $q(\cdot)$ to $q^{c*}(\cdot)$ as $\alpha \to 0$.

Second, consider $k_\theta(\theta_2)^+ > 0$. By the same argument, it is feasible to bunch types $\theta \in [\theta_2 - \varepsilon, \theta_2]$ on $q(\theta_2)$ and the remaining argument goes through analogously. Obviously, the third case $k^+_\theta(\theta_2) = 0$ is analogous to either the first or the second case (depending on the second derivative).

The effect of a marginal change on the principal's objective is

$$\int_{\theta_2 - \varepsilon}^{\theta_2 + \varepsilon} \{(u_q(q(\theta), \theta) - c_q(q(\theta), \theta))f(\theta) + (1 - F(\theta))c_{q\theta}(q(\theta), \theta)\}\, \delta(\theta)\, d\theta$$

$$= \int_{\theta_2 - \varepsilon}^{\theta_2 + \varepsilon} \eta(\theta)c_{q\theta}(q(\theta), \theta)\delta(\theta)\, d\theta = \int_{\theta_2}^{\theta_2 + \varepsilon} \delta(\theta)c_{q\theta}(q(\theta), \theta)[\eta(\theta) - \eta(\theta'(\theta))] > 0$$

where the last inequality follows from $\delta(\theta) \leq 0$ for $\theta \in [\theta_2, \theta_2 + \varepsilon]$ and $\eta_\theta(\theta) > 0$. Hence, the principal's objective increases. Due to (iii) incentive compatibility is still satisfied. This contradicts the optimality of $q(\theta)$.

[27] Note that a feasible $q^c(\theta)$ exists even around types θ_2 where $q(\theta)$ is discontinuous: Whether bunching types $[\theta_2 - \varepsilon, \theta_2)$ on $q^-(\theta_2)$ or bunching types $(\theta_2, \theta_2 + \varepsilon]$ on $q^+(\theta_2)$ is feasible is then decided by $k^+_\theta(\theta_2)$ just as in the text.

A similar argument can be made when $\eta(\theta)$ is decreasing almost everywhere on some interval $[\theta_1, \theta_3]$ where non-local ic is lax. The only difference is that (i) and (ii) are substituted by (i) $q^c(\theta) < q(\theta)$ on $[\theta_2 - \varepsilon, \theta_2)$, (ii) $q^c(\theta) \geq q(\theta)$ on $[\theta_2, \theta_2 + \varepsilon]$. The argument for existence is then that for $k_\theta(\theta_2) < 0$ one can choose a $\theta_2 + \varepsilon$ such that setting $q^c(\theta) = q(\theta_2 + \varepsilon)$ for all $\theta \in [\theta_2, \theta_2 + \varepsilon]$ is feasible. Everything else goes through accordingly.

Hence, $\eta(\theta)$ is constant on all intervals on which non-local incentive constraints do not bind.[28]

To see that $\eta(\theta)$ is non-decreasing at types $\hat{\theta}$ to which a non-local incentive constraint is binding one can use the same steps as above for types where non-local incentive constraints were lax. The key insight is that such a change is feasible due to the structure given by lemma 1 and lemma 2 (see also figure 4): Increasing q for slightly higher types than $\hat{\theta}$ (and reducing for slightly lower types than $\hat{\theta}$) will relax (or not affect) binding non-local incentive constraints because these constraints are downward binding and not overlapping.

The argument why $\eta(\theta)$ is non-increasing at types θ from which non-local incentive constraints bind is also equivalent to the one above. The key with respect to feasibility is now that reducing q for types slightly below θ (and increasing for types slightly above θ) will again relax (or not affect) binding non-local incentive constraints because these constraints are downward binding.

Now turn to $\eta(\bar{\theta}) = 0$ (and therefore $q(\bar{\theta}) = q^{fb}(\bar{\theta})$) whenever no non-local incentive constraint is binding from $\bar{\theta}$. Clearly, $q(\bar{\theta})$ does not affect non-local incentive constraints of other types, see figure 1b for an illustration. Consequently, the principal's payoff is maximized by setting $q(\bar{\theta}) = q^r(\bar{\theta})$. The only thing to show is that the monotonicity constraint is not binding at $\bar{\theta}$. Suppose to the contrary that types $[\theta', \bar{\theta}]$ were bunched on $q^b > q^{fb}(\bar{\theta})$. By lemma 3, non-local incentive constraints cannot be binding for types in $(\theta', \bar{\theta}]$. First, note that $q(\theta)$ has to be continuous at θ' as otherwise the principal's payoff could be increased by reducing q^b. Therefore–by the same argument as in the proof of lemma 3–non-local incentive constraints cannot bind from types $[\theta' - \varepsilon, \theta']$ for some small $\varepsilon > 0$. Given that $q(\theta) > q^{fb}(\bar{\theta}) > q^r(\theta)$ for all $\theta \in [\theta' - \varepsilon, \bar{\theta})$, the principal's payoff could be increased by changing $q(\theta)$ to $q(\theta' - \varepsilon)$ for all $\theta \in [\theta' - \varepsilon, \bar{\theta}]$. This contradicts the optimality of $q(\theta)$.

The part that $\eta(\underline{\theta}) = 0$ if no non-local incentive constraint is binding to $\underline{\theta}$ is even simpler: Reducing $q(\theta)$ to $q^r(\theta)$ cannot violate the monotonicity constraint as $q(\theta) \geq q^r(\theta) \geq q^r(\underline{\theta})$ by proposition 2. □

Proof of lemma 4: I proof the stronger statement, i.e. non-local incentive constraints do not only bind at isolated interior types. The proof is by contradiction.

Suppose, non-local incentive constraints bound only from isolated interior types. Denote by θ' the

[28]Note that $\eta(\theta)$ cannot be different for isolated types in such an interval: This would, by (5) and the continuity of the derivatives of $c(\cdot)$, lead to $q(\theta)$ being discontinuous at isolated points. Such a discontinuity, however, violates local incentive compatibility.

supremum of all types with $\eta(\theta) > 0$, i.e. $\theta' = \sup\{\theta : \eta(\theta) > 0\}$. By theorem 1, a non-local incentive constraint is binding from θ' and $\eta(\theta) = 0$ for all $\theta > \theta'$.[29] As the set of types from which non-local incentive constraints bind consists only of isolated types, there exists an $\varepsilon > 0$ such that non-local incentive constraints are lax for all $\theta \in (\theta' - \varepsilon, \theta')$. By theorem 1, $\eta(\theta)$ is constant on $(\theta' - \varepsilon, \theta')$ and by the definition of θ' there has to be a discontinuity in $\eta(\theta)$ at θ', i.e. $\eta^-(\theta') > \eta^+(\theta') = 0$. The definition of $\eta(\theta)$ in (5) implies then that $q^-(\theta') > q^+(\theta')$. But this violates the monotonicity constraint. Hence, θ' cannot be isolated in the set of types from which non-local incentive constraints bind.

Similarly, take $\hat{\theta}' = \inf(\hat{\theta} : \eta(\hat{\theta}) > 0)$. It holds that $\eta(\theta) = 0$ for all $\theta < \hat{\theta}'$. Therefore, by proposition 2, $\hat{\theta}'$ cannot be bunched. Consequently, a non-local incentive constraint has to bind to $\hat{\theta}'$. If $\hat{\theta}'$ is isolated in the set of types to which non-local incentive constraints are binding, $\eta(\theta)$ has to be discontinuous at $\hat{\theta}'$ by the definition of $\hat{\theta}'$. Then also $q(\theta)$ is discontinuous at $\hat{\theta}'$. But this is impossible by lemma 3. Hence, $\hat{\theta}'$ cannot be isolated in the set of types to which non-local incentive constraints bind.

It remains to show the closedness part of the lemma. Note first that a monotone solution is continuous almost everywhere. Consequently, the principal's payoff is not changed if $q(\cdot)$ is changed at its discontinuity points. I want to resolve this ambiguity using the following convention: Say $q(\theta)$ is discontinuous at θ'. Then $q(\theta') = q^-(\theta')$ if there exists an increasing sequence of types θ_i $i = 1, 2, \ldots$ such that (i) $\lim_{i\to\infty} \theta_i = \theta'$ and (ii) a non-local incentive constraint is binding from or to each θ_i. If such a sequence does not exist, $q(\theta') = q^+(\theta')$.

With this convention in mind, consider a sequence of types θ_n with $n = 1, 2, \ldots$ such that a non-local incentive constraint is binding from each θ_n to some $\hat{\theta}_n$. Assume that $\lim_{n\to\infty} \theta_n = \theta'$. Then it has to be shown that $\Phi(\theta', \hat{\theta}') = 0$ for some $\hat{\theta}'$. Since all $\hat{\theta}_n$ belong to the closed and bounded interval $[\underline{\theta}, \bar{\theta}]$, there is a convergent subsequence of $\hat{\theta}_n$. I will denote the elements of this subsequence by $\hat{\theta}_k$ with $k = 1, 2, \ldots$. The corresponding type from which a non-local incentive constraint is binding to $\hat{\theta}_k$ is denoted by θ_k. Now, take $\hat{\theta}' = \lim_{k\to\infty} \hat{\theta}_k$. Note that there always exists a *monotone* subsequence of θ_k. It is therefore without loss of generality to assume θ_k to be monotone. For concreteness, assume $\theta_{k+1} \geq \theta_k$ for all $k = 1, 2, \ldots$. As $\Phi(\theta_k, \hat{\theta}_k) = 0$ for all $k = 1, 2, \ldots$, continuity of $\Phi(\cdot)$ at $(\theta', \hat{\theta}')$ is sufficient for $\Phi(\theta', \hat{\theta}') = 0$. As $\pi(\cdot)$ is continuous by local incentive compatibility and $c(\cdot)$ is continuous by assumption, continuity of $\Phi(\cdot)$ at $(\theta', \hat{\theta}')$ follows if $q(\cdot)$ is continuous at $\hat{\theta}'$. Since θ_k is monotonically increasing, continuity from below is actually sufficient. But this is ensured by the convention above.

If $\theta_{k+1} \leq \theta_k$ for all $k = 1, 2, \ldots$, the convention establishes $q(\hat{\theta}') = q^+(\hat{\theta}')$ which is needed in this case.

[29] Note that θ' cannot be bunched because of proposition 4 and $q^-(\theta') = q^r(\theta')$.

The proof for the closedness of the set of types to which non-local incentive constraints bind works in the same way. □

Proof of lemma 5: From lemma 3, non-local incentive constraints cannot bind from any $\theta \in [\theta_s^b, \theta_e^b)$. To satisfy similar properties as in theorem 1, $\eta(\theta)$ has therefore to be non-decreasing on (θ_s^b, θ_e^b).

Let $\eta(\theta)$ be defined by (5) for all types that are not bunched. Define $\eta(\theta)$ on the bunching interval using the following two step procedure: First, all $\hat{\theta} \in (\theta_s^b, \theta_e^b)$ such that $\Phi(\theta, \hat{\theta}) = 0$ and (C1') as well as (C2') are satisfied are assigned $\eta(\hat{\theta}) = \eta(\theta)$. Second, types in $\theta \in (\theta_s^b, \theta_e^b)$ who are not assigned a value for $\eta(\theta)$ in step 1 are assigned the same η as the highest type $\theta' < \theta$ that was already assigned a value $\eta(\theta')$.

Now it is shown that the constructed $\eta(\theta)$ is non-decreasing on (θ_s^b, θ_e^b): Say, there are two types $\hat{\theta}_1, \hat{\theta}_2 \in (\theta_s^b, \theta_e^b)$ with $\hat{\theta}_2 > \hat{\theta}_1$ which are assigned an η in the first step. Then (C2') implies that $\theta_1 > \theta_2$. From theorem 1 and the structure of the solution as depicted in figure 4, it follows that $\eta(\theta_2) \geq \eta(\theta_1)$. Therefore, $\eta(\hat{\theta}_2) \geq \eta(\hat{\theta}_1)$. The second step does not change the monotonicity of $\eta(\theta)$ which proves that $\eta(\theta)$ is non-decreasing on (θ_s^b, θ_e^b).

If non-local incentive constraints are not binding for the bunched types, no type is assigned a value for $\eta(\theta)$ in step 1. Consequently, $\eta(\theta)$ is constant on (θ_s^b, θ_e^b).

Next, it is shown that $\eta(\theta)$ is also non-decreasing at the types θ_s^b and θ_e^b. First, note that the proof of theorem 1 can be easily extended to show that $\eta(\theta_s^b) \leq \eta(\theta_e^b)$: If this inequality did not hold, reduce $q(\theta)$ on $(\theta_s^b - \varepsilon, \theta_s^b)$ and increase $q(\theta)$ marginally on $(\theta_e^b, \theta_e^b + \varepsilon)$ such that $\int_{\theta_s^b - \varepsilon}^{\theta_e^b + \varepsilon} \int_{q(\theta_s^b - \varepsilon)}^{q(t)} c_{q\theta}(q, t)\, dq\, dt$ remains the same before and after the change. As in the proof of theorem 1, this change would increase the principal's payoff without impeding incentive compatibility (note that non-local incentive constraints cannot bind from the bunched types because of lemma 3). Consequently, $\eta(\theta_s^b) \leq \eta(\theta_e^b)$.

Second, it is necessary to show that–with the above constructed $\eta(\theta)$ on (θ_s^b, θ_e^b)–there is no upward jump of $\eta(\theta)$ at θ_e^b (no downward jump of $\eta(\theta)$ at θ_s^b). If no type is assigned an η in the first step of the procedure above, this is obvious. Therefore, take the case where some type in the bunching interval is assigned a value $\eta(\theta)$ in the first step of the procedure. Then the claim follows from theorem 1: Say, $\eta^-(\theta_e^b) = \eta(\theta_1)$ for some type θ_1 from which a non-local incentive constraint binds. The structure of the solution (as depicted in figure 4) and theorem 1 imply that $\eta^+(\theta_e^b) = \eta^-(\theta_1)$.[30] Since $\eta(\theta)$ is non-increasing at θ_1 according to theorem 1, it follows that $\eta^-(\theta_1) \geq \eta^+(\theta_1)$ and therefore $\eta^-(\theta_e^b) \geq \eta^+(\theta_e^b)$. A similar argument holds for θ_s^b.

[30] If non-local incentive constraints bind from types $\theta' \in (\theta_e^b, \theta_1)$ to types $\hat{\theta}' \in (\theta_e^b, \theta_1)$, this holds still true because of the necessary condition (C3). Also discontinuities at $\theta'' \in (\theta_e^b, \theta_1)$ do not matter as by lemma 3 and theorem 1 $\eta(\theta)$ is non-increasing at θ''. If there are several bunching intervals, the argument holds for the highest interval and given this, it holds for the second highest etc..

It remains to show $\int_{\theta_s^b}^{\theta_e^b} \nu_\theta(\theta)\, d\theta = 0$. But this follows directly from $\nu(\theta_s^b) = \nu(\theta_e^b) = 0$. □

Proof of proposition 3: By lemma 3, $q(\theta)$ cannot be discontinuous at a type to which a non-local incentive constraint binds (with the exception of boundary types of bunching intervals). Therefore, theorem 1 implies that a solution could only be discontinuous at types where $\eta(\theta)$ is non-increasing or at the boundary types of a bunching interval to which a non-local incentive constraint is binding.

First, it is shown that $\eta(\theta)$ is also non-increasing at such boundary types of a bunching interval. To see this take a bunching interval $[\hat{\theta}_1, \hat{\theta}_2]$ to which non-local incentive constraints bind and suppose the solution was discontinuous at $\hat{\theta}$, i.e. $q^-(\hat{\theta}_2) < q^+(\hat{\theta}_2)$. By the arguments in the proof of lemma 3, $\int_{\hat{\theta}_2}^{\theta} c_{q\theta}(q^-(\hat{\theta}_2), t)\, dt > 0$ for any θ such that $\Phi(\theta, \hat{\theta}_2) = 0$. But then an argument as in the proof of theorem 1 applies: There is an incentive compatible way to increase $q(\hat{\theta})$ for $\hat{\theta} \in [\hat{\theta}_2 - \varepsilon, \hat{\theta}_2]$ and decrease the decision for types in $[\hat{\theta}_2, \hat{\theta}+\varepsilon]$. Incentive compatible means that binding non-local incentive constraints are not violated and the decision remains monotone (details in the proof of theorem 1). If $\eta(\cdot)$ was strictly increasing at $\hat{\theta}_2$, such a change would increase the principal's payoff. Therefore, $\eta(\cdot)$ has to be decreasing at $\hat{\theta}_2$. A similar argument applies at $\hat{\theta}_1$. A discontinuity is only possible at $\hat{\theta}_1$ if $\int_{\hat{\theta}_1}^{\theta} c_{q\theta}(q(\hat{\theta}_1), t)\, dt < 0$ for all θ such that $\Phi(\theta, \hat{\theta}_1) = 0$. Therefore, decreasing the decision on $[\hat{\theta}_1, \hat{\theta}_1 + \varepsilon]$ and increasing the decision on $[\hat{\theta}_1 - \varepsilon, \hat{\theta}_1)$ can be done in an incentive compatible way. If $\eta(\cdot)$ was strictly increasing, such a change would increase the principal's payoff.

Hence, $q(\theta)$ can only be discontinuous at types where $\eta(\theta)$ is non-increasing. Second, it is shown that a discontinuity in $q(\theta)$ would lead to an upward jump of $\eta(\theta)$ at the discontinuity type which implies that there cannot be a discontinuity in $q(\theta)$.

By local incentive compatibility, $q(\theta)$ can only jump upwards, i.e. $q^-(\theta') < q^+(\theta')$ at a hypothetical discontinuity type θ'. Using the definition of $\eta(\theta)$ in (5) one can calculate the change in $\eta(\theta')$ at the discontinuity type

$$\begin{aligned}
\eta^+(\theta') - \eta^-(\theta') &= \int_{q^-(\theta')}^{q^+(\theta')} \frac{d\,\eta(\theta')}{d\,q(\theta')}\, dq \\
&= \int_{q^-(\theta')}^{q^+(\theta')} \frac{(u_{qq} - c_{qq}) f c_{q\theta} + (1-F) c_{qq\theta} c_{q\theta} - (u_q - c_q) f c_{qq\theta} - (1-F) c_{q\theta} c_{qq\theta}}{c_{q\theta}^2}\, dq
\end{aligned}$$

where all functions are evaluated at (q, θ'). Note that the integrand is positive whenever $q \leq q^{fb}(\theta')$. If $q > q^{fb}(\theta')$, the integrand can be written as

$$\frac{f(u_q - c_q)}{c_{q\theta}} \left(\frac{u_{qq} - c_{qq}}{u_q - c_q} - \frac{c_{qq\theta}}{c_{q\theta}} \right)$$

which is also positive due to the condition of the proposition. Hence, $\eta(\theta)$ would jump up at θ' but this contradicts that $q(\theta)$ can only be discontinuous at types where $\eta(\theta)$ is non-increasing. □

Proof of proposition 4: The proof is by contradiction. Suppose the optimal decision $q(\theta)$ was above the first best decision for some types. Since there is no distortion at the top by assumption and since the optimal decision cannot drop discontinuously downward (local incentive compatibility), there has to be a type θ' at which the optimal decision intersects $q^{fb}(\theta)$ from above. The proof works now in two steps. First, I show that a non local incentive constraint must bind from θ' and second that then non local incentive compatibility is violated for some type close to θ'.

Note that $q(\theta) > q^{fb}(\theta)$ if and only if $\eta(\theta) > 1 - F(\theta)$. Since $1 - F(\theta)$ is decreasing and $q(\theta) > (<) q^{fb}(\theta)$ slightly above (below) θ', it follows that $\eta_\theta(\theta')$ is negative. But then, by theorem 1, a non local incentive constraint has to be binding from θ' to some $\hat{\theta}'$. Furthermore, the necessary condition $\int_{q(\hat{\theta}')}^{q(\theta')} c_{q\theta}(q, \theta')\, dq = 0$ has to hold.

Next consider a type $\theta'' = \theta' - \epsilon$ with $\epsilon > 0$ very small. Since $q^m(\theta)$ is increasing and $\int_{q(\hat{\theta})}^{q(\theta)} c_{q\theta}(q, \theta')\, dq = 0$, clearly $\int_{q(\hat{\theta}')}^{q^{fb}(\theta'')} c_{q\theta}(q, \theta'')\, dq < 0$. Since $q(\theta'') > q^{fb}(\theta'')$, it has to hold that $\int_{q(\hat{\theta}')}^{q(\theta'')} c_{q\theta}(q, \theta'')\, dq < 0$ as well. The same inequality holds for all $\theta \in (\theta'', \theta')$. But then $\Phi(\theta'', \hat{\theta}') = \Phi(\theta', \hat{\theta}') + \int_{\theta''}^{\theta'} \int_{q(\hat{\theta}')}^{q(t)} c_{q\theta}(q, t)\, dq\, dt < 0$, i.e. incentive compatibility from θ'' to $\hat{\theta}'$ is violated. Hence, the optimal decision cannot be above the first best decision.

Continuity of the optimal decision is now straightforward: $q(\theta) \leq q^{fb}(\theta)$ implies that $1 - F(\theta) - \eta(\theta) \geq 0$. Therefore, the left hand side of the first order condition $u_q - c_q + (1 - F - \eta)c_{q\theta} = 0$ is strictly decreasing in q. The same arguments as in the proof of proposition 3 show that $q(\theta)$ has to be continuous.

Last it has to be shown that the decision is strictly monotone when it is below first best. This will be done in two steps. The first step is to show that $q(\theta)$ is strictly increasing if $\eta_\theta(\theta) \geq 0$. The second step is to show that in a hypothetical bunching interval there are types θ at which $\eta_\theta(\theta) \geq 0$ which by the first step contradicts that these types are bunched.

First, the decision $q(\theta)$ has to satisfy

$$[u_q(q(\theta), \theta) - c_q(q(\theta), \theta)] + \frac{(1 - F(\theta) - \eta(\theta))}{f(\theta)} c_{q\theta}(q(\theta), \theta) = 0 \tag{9}$$

by theorem 1. From the implicit function theorem, the sign of $q_\theta(\theta)$ can be determined. Note that $q(\theta) \leq q^{fb}(\theta)$ implies $1 - F(\theta) - \eta(\theta) \geq 0$. This in turn implies that the derivative of the left hand side of (9) with respect to q is negative. Hence, the sign of $q_\theta(\theta)$ is the sign of the partial derivative of the equation above with respect to θ. Denoting $(1 - F(\theta) - \eta(\theta))$ by $\lambda(\theta)$ this derivative is

$$u_{q\theta}(q(\theta), \theta) - c_{q\theta}(q(\theta), \theta) + \frac{\lambda(\theta)}{f(\theta)} c_{q\theta\theta} + \frac{\partial\, \lambda(\theta)/f(\theta)}{\partial \theta} c_{q\theta}(q(\theta), \theta). \tag{10}$$

Now take a bunching interval $[\theta_1, \theta_2]$ (closed or open). The first three terms are clearly positive as

$q(\theta_1) \leq q^{fb}(\theta_1)$ implies $\lambda(\theta) \geq 0$. The fourth term is positive if $\eta_\theta(\theta) \geq 0$ as then

$$\frac{\partial\, \lambda(\theta)/f(\theta)}{\partial\theta} = \frac{-f^2(\theta) - f_\theta(\theta)(1-F(\theta))}{f^2(\theta)} - \frac{\eta_\theta(\theta)}{f(\theta)} + \frac{f_\theta(\theta)\eta(\theta)}{f^2(\theta)} < 0$$

where the inequality comes from the monotone hazard rate assumption if $f_\theta(\theta) \leq 0$. If $f_\theta(\theta) > 0$, then $q^{fb}(\theta) \geq q(\theta)$ implies $\lambda(\theta) \geq 0$ which ensures the inequality above.

Now turn to the second step. Suppose contrary to the proposition that an interval (θ_1, θ_2) exists in which types are bunched and non-local incentive constraints are either binding to these types or are lax.[31] Using the same argument as in the proof of theorem 1, it becomes evident that $\eta(\theta)$ as defined by (5) cannot be decreasing on the whole interval (θ_1, θ_2): If this was the case, increasing $q(\theta)$ for types $((\theta_2+\theta_1)/2, \theta_2)$ and decreasing $q(\theta)$ slightly for the other bunched types would increase the principal's payoff (and can be done in an incentive compatible way). From the definition of $\eta(\theta)$ and the differentiability of q on the bunching interval, it follows that $\eta(\theta)$ is continuous and differentiable on this interval. Consequently, there has to be some type in the interior of the bunching interval where $\eta_\theta(\theta) \geq 0$. But then the first step shows that this type cannot be bunched. □

Proof of proposition 5: Take two types θ' and $\hat{\theta}'$ such that a non-local incentive constraint is binding from θ to $\hat{\theta}$ under the optimal decision $q(\theta)$. By (C3), $\eta(\theta') = \eta(\hat{\theta}')$ and for this proof η (in $\Phi^\eta()$) simply denotes this common value $\eta(\theta') = \eta(\hat{\theta}')$.

First, suppose that $(\theta', \hat{\theta}')$ does not minimize $\Phi^\eta(\theta, \hat{\theta})$ on $[\hat{\theta}', \theta]$ and call the minimizer $(\theta'', \hat{\theta}'')$. Then incentive compatibility under the optimal decision requires $\Phi(\theta'', \hat{\theta}'') \geq 0$. If $q(\theta)$ was $\tilde{q}(\theta)$ for all types in $[\hat{\theta}', \hat{\theta}''] \cup [\theta'', \theta']$, then $\Phi(\theta', \hat{\theta}') = \Phi^\eta(\theta', \hat{\theta}') + \Phi(\theta'', \hat{\theta}'') - \Phi^\eta(\theta'', \hat{\theta}'') > 0$ where the inequality stems from the definition of $(\theta'', \hat{\theta}'')$ as global minimizer of $\Phi^\eta(\theta, \hat{\theta})$. Therefore ic would not be binding between θ' and $\hat{\theta}'$.

If $q(\theta) \neq \tilde{q}(\theta)$ for some types in $[\hat{\theta}', \hat{\theta}''] \cup [\theta'', \theta']$, then ic must be binding for some of these types.[32] But this will only relax ic, i.e. $q(\theta) > \tilde{q}(\theta)$ in a monotone solution. Therefore $\Phi(\theta', \hat{\theta}')$ will be even higher than when $q(\theta) = \tilde{q}(\theta)$ and therefore ic cannot bind between θ' and $\hat{\theta}'$. This is the desired contradiction. Consequently, $(\theta', \hat{\theta}')$ has to minimize $\Phi^\eta(\theta, \hat{\theta})$ on $[\hat{\theta}', \theta']$.

Second, suppose that $(\theta'', \hat{\theta}'')$ with $\hat{\theta}'' < \hat{\theta}' < \theta' < \theta''$ has $\Phi^\eta(\theta', \hat{\theta}') > \Phi'(\theta'', \hat{\theta}'')$. In fact choose θ'' and $\hat{\theta}''$ such that it is the global minimizer of $\Phi^\eta(\theta, \hat{\theta})$ under the constraint $\hat{\theta} < \hat{\theta}' < \theta' < \theta$.

Now suppose for the moment that all types in $[\hat{\theta}'', \hat{\theta}'] \cup [\theta', \theta'']$ had $q(\theta) = \tilde{q}(\theta)$. Then since $\Phi(\theta', \hat{\theta}') = 0$ but $(\theta'', \hat{\theta}'')$ minimizes $\Phi^\eta(\theta, \hat{\theta})$, ic would be violated for θ'' and $\hat{\theta}''$.

If $q(\theta) \neq \tilde{q}(\theta)$ for some types in $[\hat{\theta}'', \hat{\theta}'] \cup [\theta', \theta'']$, then ic was binding for some types in those intervals. In a monotone solution this implies that $q(\theta) < \tilde{q}(\theta)$ for these types. Put differently, ic is stricter under

[31] By lemma 3, types from which non-local incentive constraints bind cannot be bunched.

[32] Because of lemma 2 ic cannot bind from outside $[\hat{\theta}', \theta]$ into the interval (neither the other way round).

$q(\theta)$ than under $\tilde{q}(\theta)$.[33] But then ic will be even more violated for θ'' and $\hat{\theta}''$ under $q(\theta)$ than under $\tilde{q}(\theta)$. Therefore, there cannot be a global minimizer $(\theta'', \hat{\theta}'')$ with $\hat{\theta}'' < \hat{\theta}' < \theta' < \theta''$. □

Proof of corollary 1: Note first that the highest type θ from which a non-local incentive constraint is binding must have $q(\theta) = q^r(\theta)$ if θ is interior. This follows from the reasoning in the proof of lemma 4. The same holds for the lowest type $\hat{\theta}$ to which a non-local incentive constraint binds. Therefore, there is a type pair such that (i) $q(\theta') = q^r(\theta')$, (ii) $q(\hat{\theta}') = q^r(\hat{\theta}')$ and (iii) $\Phi(\theta', \hat{\theta}') = 0$.

Since $(\theta', \hat{\theta}')$ satisfy (C2) and (C1) with q^r and given the results of proposition 5, $(\theta', \hat{\theta}')$ locally minimize $\Phi^r(\theta, \hat{\theta})$. Proposition 5 rules out that $\hat{\theta}^r < \hat{\theta}' < \theta' < \theta^r$ and also $\hat{\theta}' < \hat{\theta}^r < \theta^r < \theta'$. Hence, it still has to be shown that there cannot be an overlap between the two type pairs, i.e. $\hat{\theta}' < \hat{\theta}^r < \theta' < \theta^r$ or $\hat{\theta}^r < \hat{\theta}' < \theta^r < \theta'$. To get a contradiction suppose $\hat{\theta}' < \hat{\theta}^r < \theta' < \theta^r$. In a similar way as in lemma 2, one can now show that in this case $\Phi^r(\theta^r, \hat{\theta}') < \Phi^r(\theta^r, \hat{\theta}^r)$ thereby contradicting that $(\theta^r, \hat{\theta}^r)$ is the global minimizer of $\Phi^r(\theta, \hat{\theta})$:

$$\begin{aligned}\Phi^r(\theta^r, \hat{\theta}') &= \Phi^r(\theta^r, \hat{\theta}^r) + \Phi^r(\theta', \hat{\theta}') + \int_{\hat{\theta}^r}^{\theta'} \int_{q^r(\hat{\theta}^r)}^{q^r(t)} c_{q\theta}(q,t)\, dq\, dt - \int_{\theta'}^{\theta^r} \int_{q^r(\hat{\theta}')}^{q^r(\hat{\theta}^r)} c_{q\theta}(q,t)\, dq\, dt \\ &= \Phi^r(\theta^r, \hat{\theta}^r) + \Phi^r(\theta', \hat{\theta}') - \Phi^r(\theta', \hat{\theta}^r) - \int_{\theta'}^{\theta^r} \int_{q^r(\hat{\theta}')}^{q^r(\hat{\theta}^r)} c_{q\theta}(q,t)\, dq\, dt\end{aligned}$$

By proposition 5, $\Phi^r(\theta', \hat{\theta}') - \Phi^r(\theta', \hat{\theta}^r) \leq 0$. Furthermore, $\int_{q^r(\hat{\theta}')}^{q^r(\theta')} c_{q\theta}\, dq = 0$ since $(\theta', \hat{\theta}')$ locally minimize $\Phi^r(\theta, \hat{\theta})$. Therefore, $\int_{q^r(\hat{\theta}')}^{q(\hat{\theta}^r)} c_{q\theta}\, dq > 0$ as $q^r(\theta') > q^r(\hat{\theta}^r)$ and $c_{qq\theta} < 0$. From $c_{q\theta\theta} > 0$ it follows that $\int_{\theta'}^{\theta^r} \int_{q^r(\hat{\theta}')}^{q^r(\hat{\theta}^r)} c_{q\theta}(q,t)\, dq\, dt > 0$ which shows that $\Phi^r(\theta^r, \hat{\theta}') < \Phi^r(\theta^r, \hat{\theta}^r)$. This is the desired contradiction.

A similar argument can be made for the case $\hat{\theta}^r < \hat{\theta}' < \theta^r < \theta'$. Consequently, the only possibility is that $(\theta', \hat{\theta}') = (\theta^r, \hat{\theta}^r)$ which had to be shown.

If the highest/lowest type from/to which a non-local incentive constraint is binding is a boundary type, this type's decision is not necessarily the relaxed decision. However, the minimization argument does not change which concludes the proof. □

[33] Strictly speaking one also has to show that ic did not bind from outside $[\hat{\theta}'', \theta'']$ into this interval (or the other way round), thereby increasing $q(\theta)$ for some types in say (θ', θ''). If however this was the case and the increase in $q(\theta)$ was such that ic between θ'' and $\hat{\theta}''$ was relaxed by it, then there has to exist a type $\hat{\theta}''' \in (\theta', \theta'')$ and a type $\theta''' > \theta''$ with $\Phi(\theta''', \hat{\theta}''') = 0$ and $q(\hat{\theta}''') = \tilde{q}(\hat{\theta}''')$. But this would contradict that $(\theta'', \hat{\theta}'')$ is a global minimum of $\Phi^\eta(\theta, \hat{\theta})$ (analogously to the proof of lemma 2), i.e. $\Phi'(\theta''', \hat{\theta}'') < \Phi'(\theta'', \hat{\theta}'')$.

C. Existence of an optimal contract

This appendix shows that an optimal contract exists and therefore the characterization done in the paper is meaningful. It is assumed that $q^v(q,\theta) \geq q^s(q,\theta)$ for all $q \in [0, q^f(\theta)]$ and all $\theta \in [\underline{\theta}, \bar{\theta}]$ and therefore proposition 1 applies. Before showing existence, two useful lemmata are derived.

Define $\tilde{q}$ such that $\int_0^{\tilde{q}} c_{q\theta}(q,\bar{\theta})\, dq = 0$. Since $c_{qq\theta} < 0$, $\tilde{q}$ is unique and therefore properly defined.

Lemma 6. *Any incentive compatible contract with a decision $q(\theta)$ above $\bar{q} = max\{q^{fb}(\bar{\theta}), \tilde{q}\}$ for some type is dominated by a contract consisting of decision*

$$q^c(\theta) = min\{q(\theta), \bar{q}\}$$

and transfers such that $\pi(\underline{\theta}) = 0$ and $\pi_\theta(\theta) = \int_{\underline{\theta}}^{\theta} -c_\theta(q(t), t)\, dt$.

Proof. The concavity of the virtual valuation implies that the principal's payoff under $q^c(\theta)$ is higher than under $q(\theta)$. Hence, the lemma holds if the changed contract is incentive compatible.

Note that incentive compatibility of $q^c(\theta)$ is obvious if $q(\theta) > \bar{q}$ for all θ. Now define $\theta^m = inf\{\theta : q(\theta) > \bar{q}\}$. Note that incentive compatibility from θ^m to any lower type is not affected by the change from $q(\cdot)$ to $q^c(\cdot)$ since $\Phi(\theta^m, \hat{\theta})$ does not change.

The next step is to see that $q(\theta) > \bar{q}$ for all $\theta > \theta^m$. The reason is that local incentive compatibility does not allow for any decision in $[s(\theta), \bar{q}]$ as long as $q(\theta)$ stays above $s(\theta)$. Furthermore, downward jumps to a decision below $s(\theta)$ would require that $\int_{q^+(\theta^j)}^{q^-(\theta^j)} c_{q\theta}(q, \theta^j)\, dq \geq 0$ at the jump type θ^j (for local incentive compatibility). But by the definition of $\bar{q}$ and from $c_{q\theta\theta} > 0$, this inequality cannot hold for any type below $\bar{\theta}$ (and a jump at the boundary type $\bar{\theta}$ would not hurt the following argument).

Therefore, all types above θ^m will have $\bar{q}$ as their changed decision. From lemma 1 it follows that only incentive compatibility from types above θ^m to types below θ^m has to be checked. Therefore take an arbitrary $\theta > \theta^m$ and some $\hat{\theta} < \theta^m$. Then $\Phi(\theta, \hat{\theta}) = \Phi(\theta^m, \hat{\theta}) - \int_{\theta^m}^{\theta} \int_{q(\hat{\theta})}^{\bar{q}} c_{q\theta}(q,t)\, dq\, dt > 0$ where the inequality follows from the incentive compatibility between θ^m and $\hat{\theta}$ under $q(\theta)$, the definition of $\tilde{q}$ and $c_{q\theta\theta} > 0$. $\square$

Lemma 7. *Take a sequence of incentive compatible decision functions*[34] $q^n(\theta) \leq \bar{q}$, $n = 1, 2 \ldots,$ *and let this sequence converge to $q(\theta)$. Then $q(\theta)$ is incentive compatible.*

Proof. Define $\tilde{c}_{q\theta} = \max_{q \in [0,\bar{q}],\, \theta \in [\underline{\theta}, \bar{\theta}]} |c_{q\theta}(q,\theta)|$. Since $[0, \bar{q}] \times [\underline{\theta}, \bar{\theta}]$ is compact and $c_{q\theta}(\cdot)$ is continuous by assumption, $\tilde{c}_{q\theta}$ exists.

[34] An incentive compatible decision is a decision such that the menu consisting of this decision and transfers defined by $\pi(\theta) = \int_{\underline{\theta}}^{\theta} -c_\theta(q(t), t)\, dt$ is incentive compatible.

Now suppose contrary to the lemma that $\Phi(\theta, \hat{\theta}) = -\varepsilon$ for some θ, $\hat{\theta} \in \Theta$ and $\varepsilon > 0$ and therefore incentive compatibility is violated under $q(\theta)$. From convergence of $\{q^n(\theta)\}$, for each $\delta > 0$ there exists an N_δ such that $|q^n(\theta) - q(\theta)| \leq \delta$ for all types and all $n > N_\delta$. Therefore,

$$\Phi(\theta, \hat{\theta}) = \int_{\hat{\theta}}^{\theta} \int_{q(\hat{\theta})}^{q(t)} -c_{q\theta}(q,t)\; dq\, dt \geq \int_{\hat{\theta}}^{\theta} \int_{q^n(\hat{\theta})}^{q^n(t)} -c_{q\theta}(q,t)\; dq\, dt - \int_{\hat{\theta}}^{\theta} 2\delta \tilde{c}_{q\theta}\; dt$$

for an arbitrary $n > N_\delta$. But then choosing a $\delta < \frac{\varepsilon}{2\tilde{c}_{q\theta}(\bar{\theta} - \underline{\theta})}$ shows that $\Phi(\theta, \hat{\theta}) > -\varepsilon$ as $\Phi^n(\theta, \hat{\theta}) \geq 0$ where $\Phi^n(\cdot)$ denotes $\Phi(\cdot)$ under $q^n(\cdot)$. This contradicts the definition of ε and therefore $q(\theta)$ is incentive compatible. □

Given proposition 1 and the previous two results, the existence proof in Jullien (2000) applies. For completeness, I replicate the proof briefly. The problem of the principal is the program:

$$\max_{q(\theta)} \int_{\underline{\theta}}^{\bar{\theta}} (u(q(\theta), \theta) - c(q(\theta), \theta)) f(\theta) + (1 - F(\theta)) c_\theta(q(\theta), \theta)\; d\theta$$

subject to

$$\begin{aligned} \Phi(\theta, \hat{\theta}) &\geq 0 \quad \text{for all} \quad \theta, \hat{\theta} \in [\underline{\theta}, \bar{\theta}] \\ 0 &\leq q(\theta) \leq \bar{q} \end{aligned}$$

Let W^* be the maximum value of the program. Take a sequence of decision functions such that $q^n(\theta)$ induce a value larger than $W^* - \frac{1}{n}$ and each $q^n(\theta)$ is incentive compatible . Because of proposition 1, the sequence can be chosen such that each $q^n(\theta)$ is an increasing function. Then Helly's selection theorem, see Billingsley (1986) Thm. 25.9, yields that there exists a non-decreasing function $q(\theta)$ which is the limit of a subsequence $q^{n_k}(\theta)$ at every point of continuity of $q(\theta)$ and therefore almost everywhere on $[\underline{\theta}, \bar{\theta}]$. Lebesgue's dominated convergence theorem, see Billingsley (1986) Thm. 16.4, yields that the principal's payoff under $q(\theta)$ is W^*. By lemma 7, $q(\theta)$ is implementable and therefore an optimal contract exists.