Optimality of sequential screening with multiple units and ex post participation constraints

Daniel Krähmer*and Roland Strausz†

First version: September 10, 2015
This version: March 1, 2016

Abstract

We show that in sequential screening problems with ex post participation constraints, optimal contracts elicit the agent's pay-off irrelevant ex ante information when the principal and agent can trade multiple units, in contrast to when they can trade a single unit only. The difference arises because with multiple units, the principal can price each unit differently, giving rise to a larger number of screening instruments. Optimal contracts implement output schedules that are not monotone in the initial information. We identify regularity conditions which ensure that non-monotone schedules are incentive compatible.

Keywords: sequential screening, static screening, ex post participation constraints
JEL codes: D82

*University Bonn, Department of Economics, Institute for Microeconomics, Adenauer Allee 24-42, D-53113 Bonn (Germany), kraehmer@hcm.uni-bonn.de.

†Humboldt-Universität zu Berlin, Institute for Economic Theory 1, Spandauer Str. 1, D-10178 Berlin (Germany), strauszr@wiwi.hu-berlin.de.

We would like to thank an anonymous referee for very useful comments.

1 Introduction

Krähmer and Strausz (2015) shows that in a unit good sequential screening problem with ex post participation constraints, the optimal contract is, under appropriate regularity conditions, static, that is, it does not elicit the agent's information sequentially.[1] In this note, we show that in sequential screening problems with multiple units, optimal contracts do screen sequentially, despite the presence of ex post participation constraints.[2]

Intuitively, the difference arises because in the unit good case, the principal's instruments to screen sequentially are much more limited. With only one unit to trade, there is, by definition, only one marginal unit, i.e. the number of units traded can change only from zero to one, whereas in the case of trading multiple units there are multiple marginal units. Because the principal can, effectively, set different prices for different marginal units, her screening instruments are much richer in a setup with multiple units. As we show and explain in more detail below, the principal can exploit these richer screening opportunities to profitably screen the agent sequentially.

2 The Setup

We consider a seller (she, principal) who can sell multiple units $x \geq 0$ of a good or service to a buyer (he, agent). The seller's cost of producing x units of the good are $c(x)$ with $c'(x) \geq 0$ and $c''(x) \geq 0$. The buyer's marginal valuation is $\theta \in \Theta \equiv [\underline{\theta}, \overline{\theta}]$, with $\underline{\theta} < \overline{\theta}$.

The terms of trade are the number of traded units, $x \geq 0$, and a payment $t \in \mathbb{R}$ from the buyer to the seller. Parties have quasi–linear utility functions. That is, under the terms of trade x and t, the seller receives utility $t - c(x)$, and the buyer receives utility $\theta x - t$.

There are three periods. In the first period, where contracting takes place, the buyer has a private ex ante signal $i \in \{1, 2\}$ that is informative about his marginal valuation which realizes in period 2. The probability of the signal i is p_i. In period 2, the buyer privately observes his

[1] The sequential screening problem without ex post participation constraints has been introduced by Courty and Li (2000). In this model, the agent has private, yet imperfect information ex ante and privately learns more information ex post when trading takes place. Without ex post participation constraints, that is, when the agent cannot quit the contract and can sustain losses ex post, optimal contracts typically elicit the agent's information sequentially. While Courty and Li (2000) study a problem with a single unit, Dai et al. (2006) study a problem with multiple units.

[2] From an applied side, ex post participation constraints arise in contexts in which the agent has a legal right to withdraw from the contract ex post, such as in online sales contracts within the EU or in employment relationships, or when his ability to post bonds is limited due to cash constraints.

marginal valuation θ which is drawn from the distribution $G_i(\theta)$ with common support $[\underline{\theta}, \overline{\theta}]$.[3] While the buyer's ex ante and ex post types are his private information, the distributions G_1 and G_2 are common knowledge as well as the ex ante distribution of i.

Importantly, we assume that, at the end of period 2, after having learned his true valuation θ, the agent can always quit and receive his ex post outside option of zero. This assumption leads to the consideration of ex post rather than ex ante participation constraints. If the buyer does not withdraw from the contract, then, in period 3, the terms of trade are enforced, that is the seller delivers the contractually specified quantity, and the buyer makes payments.

Throughout, we denote by

$$h_i(\theta) = \frac{1 - G_i(\theta)}{g_i(\theta)}, \quad h_{ij}(\theta) = \frac{1 - G_i(\theta)}{g_j(\theta)} \tag{1}$$

the (inverse) hazard rates and the (inverse) "cross" hazard rates of the distributions of buyer valuations.

3 The maximization problem

To find an optimal contract, we apply the revelation principle for sequential games (e.g., Myerson 1986), which states that the optimal contract can be found in the class of direct and incentive compatible contracts which, on the equilibrium path, induce the buyer to report his type truthfully. Formally, a *direct contract* $\{(x_1(\theta), t_1(\theta)), (x_2(\theta), t_2(\theta))\}_{\theta \in [\underline{\theta}, \overline{\theta}]}$ requires the buyer to report an ex ante type j in period 1, and an ex post type θ' in period 2.

If the buyer's true ex post type is θ and his period 1 report was j, then his utility from reporting θ' in period 2 is $v_j(\theta'; \theta) \equiv \theta x_j(\theta') - t_j(\theta')$. With slight abuse of notation, we denote the buyer's period 2 utility from truth–telling by $v_j(\theta) \equiv v_j(\theta; \theta)$. The contract is *incentive compatible in period 2* if it gives the buyer an incentive to announce his ex post type truthfully: $v_j(\theta) \geq v_j(\theta'; \theta)$ for all $j \in \{1, 2\}, \theta, \theta' \in [\underline{\theta}, \overline{\theta}]$. From this it directly follows that if the contract is incentive compatible in period 2, the buyer announces his ex post type truthfully no matter what his report in the first period.[4] Hence, the contract induces the buyer to announce his ex

[3] We impose assumptions on the ordering of G_2 and G_1 below.

[4] Because the buyer's period 2 utility is independent of his ex ante type, a contract which is incentive compatible in period 2 automatically induces truth–telling in period 2 also off the equilibrium path, that is, if the buyer has misreported his ex ante type in period 1.

ante type truthfully, and is thus *incentive compatible in period 1* if

$$\int_{\underline{\theta}}^{\overline{\theta}} v_i(\theta)\, dG_i(\theta) \geq \int_{\underline{\theta}}^{\overline{\theta}} v_j(\theta)\, dG_i(\theta) \quad \text{for } i, j \in \{1,2\}. \tag{2}$$

In the presence of ex post participation constraints, the direct mechanism must further provide the agent with non–negative utility ex post, that is,[5]

$$v_i(\theta) \geq 0 \quad \text{for all } i \in \{1,2\}, \theta \in [\underline{\theta}, \overline{\theta}]. \tag{3}$$

To state the seller's problem, we proceed in a standard fashion and first eliminate transfers from the problem. By an envelope argument (see Theorem 2 in Milgrom and Segal, 2002), incentive compatibility in the second period is equivalent to the properties (i) that

$$x_i(\theta) \text{ is increasing in } \theta \text{ for } i \in \{1,2\}; \tag{MON_i}$$

and (ii) that $v_i'(\theta) = x_i(\theta)$ for almost all $\theta \in \Theta$ and for all $i \in \{1,2\}$.

Because (ii) implies that v_i is increasing, it follows that (3) is equivalent to

$$v_i(\underline{\theta}) \geq 0 \quad \text{for all } i \in \{1,2\}. \tag{IR_i}$$

Moreover, we can use (ii) to eliminate transfers and obtain the seller's problem as a choice problem over the selling schedules $\mathbf{x} = (x_1(\cdot), x_2(\cdot))$ and the utilities, $\mathbf{v} = (v_1(\underline{\theta}), v_2(\underline{\theta}))$. After integration by parts and re-arranging terms, ex ante type i's incentive constraint (2) rewrites as

$$\int_{\underline{\theta}}^{\overline{\theta}} [x_i(\theta) - x_j(\theta)](1 - G_i(\theta))\, d\theta + v_i(\underline{\theta}) - v_j(\underline{\theta}) \geq 0, \tag{IC_i}$$

and the seller's objective becomes

$$W(\mathbf{x}, \mathbf{v}) = \sum_{i=1,2} p_i \left[\int_{\underline{\theta}}^{\overline{\theta}} (\theta - h_i(\theta))\, x_i(\theta) - c(x_1(\theta))\, dG_i(\theta) - v_i(\underline{\theta}) \right].$$

The following lemma summarizes.

Lemma 1 *The optimal, possibly dynamic, screening contract solves*

$$\mathscr{P}: \qquad \max_{\mathbf{x},\mathbf{v}} W(\mathbf{x}, \mathbf{v}) \quad s.t. \quad (MON_i),\ (IR_i),\ (IC_i) \quad \textit{for } i \in \{1,2\}.$$

[5] Put differently, if the seller offered a contract for which the buyer would make an ex post loss for some θ, then the buyer would withdraw from the contract for such a θ, and effectively enforce the terms of trade $x_i(\theta) = t_i(\theta) = 0$.

4 Optimality of sequential screening

In a linear environment—where the seller has constant marginal costs, $c'' = 0$, and there is an upper bound on the quantity, $x \leq \bar{x}$—Krähmer and Strausz (2015) show that there is a solution to $\mathscr{P}$ which is "static", that is, the contract conditions only on the ex post type θ and not the ex ante signal i:

$$x_1(\theta) = x_2(\theta) \text{ for all } \theta \in [\underline{\theta}, \overline{\theta}], \text{ and } v_1(\underline{\theta}) = v_2(\underline{\theta}). \tag{$STAT$}$$

The point of this note is to show that with non-constant marginal cost, the optimal contract is, in contrast, dynamic and does screen sequentially. To show this claim, we first illustrate that a solution to the first-order conditions to problem $\mathscr{P}$ is generically not static. We then state conditions under which such a solution indeed solves the original problem.

By the Kuhn-Tucker theorem (Luenberger, 1969, p.220), $(\mathbf{x}, \mathbf{v})$ is a solution to $\mathscr{P}$ if and only if there are multipliers $\lambda_i \geq 0, \mu_i \geq 0$ so that $(\mathbf{x}, \mathbf{v})$ maximizes the Lagrangian

$$\begin{aligned} L \;=\; & p_1 \left[\int_{\underline{\theta}}^{\overline{\theta}} \left(\theta - h_1(\theta) + \frac{\lambda_1}{p_1} h_1(\theta) - \frac{\lambda_2}{p_1} h_{21}(\theta) \right) x_1(\theta) - c(x_1(\theta))\, dG_1(\theta) \right] \\ & + p_2 \left[\int_{\underline{\theta}}^{\overline{\theta}} \left(\theta - h_2(\theta) + \frac{\lambda_2}{p_2} h_2(\theta) - \frac{\lambda_1}{p_1} h_{12}(\theta) \right) x_2(\theta) - c(x_2(\theta))\, dG_2(\theta) \right] \\ & - \big(p_1 - \lambda_1 + \lambda_2 - \mu_1) v_1(\underline{\theta})\big) - p_2 \big(1 - \lambda_2 + \lambda_1 - \mu_2) v_2(\underline{\theta})\big) \end{aligned} \tag{4}$$

subject to the monotonicity constraints (MON_i). If we consider the relaxed problem where we ignore (MON_i), then a solution to this relaxed problem is a maximizer of L and thus satisfies the first-order conditions (for point-wise optimality)

$$c'(x_1(\theta)) = \theta - h_1(\theta) + \frac{\lambda_1}{p_1} h_1(\theta) - \frac{\lambda_2}{p_1} h_{21}(\theta), \text{ and } c'(x_2(\theta)) = \theta - h_2(\theta) + \frac{\lambda_2}{p_2} h_2(\theta) - \frac{\lambda_1}{p_1} h_{12}(\theta). \tag{5}$$

Hence, $x_1(\cdot)$ and $x_2(\cdot)$ can coincide while satisfying (5) only if there are λ_1, λ_2 so that for all $\theta \in [\underline{\theta}, \overline{\theta}]$, we have

$$-h_1(\theta) + \frac{\lambda_1}{p_1} h_1(\theta) - \frac{\lambda_2}{p_1} h_{21}(\theta) = -h_2(\theta) + \frac{\lambda_2}{p_2} h_2(\theta) - \frac{\lambda_1}{p_1} h_{12}(\theta). \tag{6}$$

This condition is highly special, since the two variables λ_1 and λ_2 have to satisfy an infinite number of equations. Or stated more formally, condition (6) is "non–generic" in the sense that if the condition is satisfied for two distributions G_1 and G_2, it will be violated if we perturb them only slightly. Therefore, the relaxed problem where we ignore (MON_i) generically has no static solution. Hence, if any solution to the relaxed problem automatically satisfies (MON_i), we can conclude that also the original problem generically has no static solution.

We now provide sufficient conditions for the solution to (5) to be indeed the solution to the original problem. To do so, we impose "regularity" conditions on the distribution functions so that it will be sufficient to solve a reduced problem in which the "low" type's ex ante incentive constraint and the monotonicity constraints are neglected, and both participation constraints are binding.

In particular, we consider the problem

$$\mathscr{R}: \qquad \max_{\mathbf{x}} \quad W(\mathbf{x},0,0) \quad \text{s.t.} \quad \int_{\underline{\theta}}^{\overline{\theta}} [x_2(\theta)-x_1(\theta)](1-G_2(\theta))\,d\theta \geq 0.$$

We say that distributions G_1 and G_2 are *regular* if

(i) G_2 dominates G_1 in the likelihood ratio order, that is, g_2/g_1 is increasing;

(ii) h_1, h_2, and h_{21} are decreasing;

(iii) h_2/h_1 is decreasing.

The next proposition shows that, for regular distributions, any schedule $\mathbf{x}$ which is optimal for $\mathscr{R}$ also solves $\mathscr{P}$. Moreover, any solution $\mathbf{x}$ of $\mathscr{R}$ satisfies the first order condition (5) with $\lambda_1 = 0$ and $\lambda_2 > 0$ so that it is generically not static.

Proposition 1 *Suppose G_1 and G_2 are regular. Then for any solution $\mathbf{x}$ to problem $\mathscr{R}$, the tuple $(\mathbf{x},0,0)$ is a solution to problem $\mathscr{P}$. In particular, the optimal schedules $x_1(\cdot)$ and $x_2(\cdot)$ are given as a solution to (5) where $\lambda_1 = 0$, and $\lambda_2 > 0$ is such that (IC_2) is binding. Moreover, x_1 crosses x_2 once in $(\underline{\theta},\bar{\theta})$ and from above.*

Our regularity conditions, and in particular conditions (i) and (iii), are new to the literature.[6] It is therefore instructive to discuss their role, and how they relate to the regularity conditions the literature usually imposes. Condition (ii) corresponds to monotone hazard rate conditions which, as is familiar in screening models, ensure that the solution to the relaxed problem satisfies the monotonicity condition (MON_i).[7]

[6]Conditions (i) to (iii) are satisfied for large classes of distributions. An example is the class of power distributions: $G_i(\theta) = \theta^{a_i}$ with $a_1 < a_2$.

[7]We already introduced this condition in Krähmer and Strausz (2015), where we also explained its role more carefully.

The role of conditions (i) and (iii) is to ensure that the solution to the relaxed problem automatically satisfies the neglected (IC_1). To see this, it is best to compare them to the corresponding conditions in sequential screening models where there are no ex post participation constraints.[8] Without ex post participation constraints, ex ante types are typically required to be ordered in terms of first order stochastic dominance (which is weaker than likelihood ratio dominance). In this case, the "high" ex ante type's incentive constraint (IC_2) implies the "low" ex ante type's incentive constraint (IC_1) if the schedule $\mathbf{x}$ is strongly monotone, that is, $x_i(\theta)$ is increasing in both θ and i. One then imposes an additional regularity condition that ensures that the solution to the relaxed problem where (IC_1) is neglected indeed displays strong monotonicity.

In our reduced problem $\mathscr{R}$, however, *any* non-static solution that satisfies the constraint with equality violates strong monotonicity. Hence, in our setting, the condition that G_2 first order stochastically dominates G_1 is generally too weak to ensure that (IC_2) implies (IC_1). We therefore impose the stronger likelihood ratio order (i) on the ex ante type distributions and show that in this case, the "high" ex ante type's incentive constraint (IC_2) implies the "low" ex ante type's incentive constraint (IC_1) if the schedule $\mathbf{x}$ satisfies a weaker condition than strong monotonicity, namely only that x_1 crosses x_2 at most once and from above. Condition (iii) is then the additional regularity condition that ensures that the solution to the relaxed problem where (IC_1) is neglected indeed displays this weaker condition.

5 Discussion

The result of Proposition 1 that even in the presence of ex post participation constraints, the optimal contract screens sequentially when there are multiple units stands in contrast to the result in Krähmer and Strausz (2015) that sequential screening has no benefits when there is a single unit. An intuition for this discrepancy follows from considering the question, why, with multiple units, the optimal static contract, which is optimal in the class of contracts that do not screen ex ante types, is not optimal within the class of sequential screening contracts.

All else equal, the buyer type 2 has a higher ex ante willingness to pay because he considers it more likely than type 1 to have a high valuation θ ex post. Hence, the seller would like to screen the buyer so as to extract type 2's higher ex ante willingness to pay. Relative to pooling ex ante types and offering only the optimal static contract, she can do so by offering a menu

[8] See Courty and Li (2000). For a textbook treatment, see Krähmer and Strausz (2015b).

which contains an additional second contract with a higher level of consumption for high ex post valuations and a lower level of consumption for low ex post valuations. This second contract is relatively more attractive to type 2, because type 2 considers high ex post valuations more likely than type 1. Hence, by setting the specific terms of this additional contract so that type 2 is indifferent between this contract and the optimal static one, the principal is able to sequentially screen the two ex ante types.[9]

Now observe that when the seller offers such a sequential menu, then by construction, both buyer types get the same utility as when only the static contract is offered. Thus, information rents are the same under the static and the sequential contract. Therefore, the principal obtains a strictly lower payoff from offering only the optimal static contract, if the sequential contract generates a strictly larger surplus. Because type 2 is more likely to have a high valuation θ and because the optimal static contract generically exhibits downward distortions, it is indeed possible for the principal to construct a more efficient contract for type 2 on average by raising output for high ex post valuations and lowering it for small ex post valuations.

To see that this procedure to improve upon the optimal static contract is not applicable in the single unit case, note that when only a single unit is sold, the optimal static contract is a simple threshold contract, so that the good is traded for high ex post valuations and not traded for low ex post valuations. Hence, even though the optimal contract in the single unit case is also distorted downwards (in the sense that the threshold is too high), it is not possible to offer an additional contract which, relative to the optimal static contract, has both higher output for larger ex post types and lower output for smaller ex post types.

We finally point out that, while the optimal contract with ex post participation constraints does not coincide with the optimal static contract, it does also not coincide with the optimal sequential screening without ex post participation constraints. This follows directly from the observation in Courty and Li (2000) that the optimal sequential screening contract with only an ex ante participation constraint violates the ex post participation constraints for low ex post types.

Appendix

Proof of Proposition 1 Observe first that the Lagrangian to problem $\mathscr{R}$ is given by (4) with $\lambda_1 = 0$ (and $v_1(\underline{\theta}) = v_2(\underline{\theta}) = 0$). Therefore, $\mathbf{x}$ is a solution to $\mathscr{R}$ if and only if there is a $\lambda_2 \geq 0$ so that $\mathbf{x}$

[9] Because type 1 is less "optimistic" than type 2, he will, if type 2 is indifferent, prefer the static contract.

satisfies condition (5) for $\lambda_1 = 0$, that is,

$$c'(x_1(\theta)) = \theta - h_1(\theta) - \frac{\lambda_2}{p_1} h_{21}(\theta), \quad \text{and} \quad c'(x_2(\theta)) = \theta - h_2(\theta) + \frac{\lambda_2}{p_2} h_2(\theta), \tag{7}$$

and, moreover, $\lambda_2 = 0$ if the constraint is not binding. We proceed by showing five auxiliary claims which will imply the proposition.

Claim 1: At any solution to $\mathscr{R}$, (IC_2) is binding, that is, $\int_{\underline{\theta}}^{\overline{\theta}} [x_2(\theta) - x_1(\theta)](1 - G_2(\theta))\, d\theta = 0$.

Indeed, otherwise, $\lambda_2 = 0$ so that (7) implies that $c'(x_i(\theta)) = \theta - h_i(\theta)$. Since likelihood ratio dominance implies (inverse) hazard rate dominance[10], assumption (i) implies that $h_2(\theta) \geq h_1(\theta)$ for all θ, and hence, $x_1(\theta) \geq x_2(\theta)$ for all θ, yielding the contradiction that the solution would violate (IC_2).

Claim 2: At any solution to $\mathscr{R}$, x_1 and x_2 are increasing.

To see this, observe first that since $\lambda_2 > 0$ by Claim 1, (7), $c'' > 0$, and assumption (ii) imply that x_1 is increasing. Moreover, since $\lambda_2 > 0$ and $c'' > 0$, we have that $x_1(\theta) \leq x^{FB}(\theta)$ for all θ where $x^{FB}(\cdot)$ is the first-best schedule that satisfies $c'(x(\theta)) = \theta$ for all θ. This implies that $\lambda_2 \leq p_2$, because otherwise, (7) would imply that $x_2(\theta) \geq x^{FB}(\theta)$ for all θ, and hence $x_2(\theta) \geq x_1(\theta)$ for all θ, thereby yielding a contradiction to (IC_2) being binding by Claim 1. Having thus established that $\lambda_2 \leq p_2$, it is now immediate from (7) and assumption (ii) that also x_2 is increasing.

Claim 3: At any solution to $\mathscr{R}$, x_1 crosses x_2 exactly once in $(\underline{\theta}, \bar{\theta})$ and from above.

Indeed, by (7) and since $c'' > 0$, we have that

$$x_1(\theta) \geq x_2(\theta) \quad \Leftrightarrow \quad \phi(\theta) \geq \lambda_2, \quad \text{with } \phi(\theta) = \frac{h_2(\theta) - h_1(\theta)}{h_{21}(\theta)/p_1 + h_2(\theta)/p_2}. \tag{8}$$

Since (IC_2) is binding by Claim 1, x_1 and x_2 have to cross at least once in $(\underline{\theta}, \bar{\theta})$. To show the claim, it is therefore sufficient to show that ϕ is decreasing. To show this, recall that for non-negative, differentiable functions $\alpha, \beta : \Theta \to \mathbb{R}$, we have that $\frac{\alpha}{\beta}$ is decreasing if and only if $\frac{\alpha'}{\alpha} \leq \frac{\beta'}{\beta}$. Assumption (i) implies that $h_2 - h_1$ is non–negative.[11] Therefore,

$$\phi \text{ is decreasing} \quad \Leftrightarrow \quad \frac{h_2' - h_1'}{h_2 - h_1} \leq \frac{h_{21}'/p_1 + h_2'/p_2}{h_{21}/p_1 + h_2/p_2}. \tag{9}$$

[10] G_2 dominates G_1 in the (inverse) hazard rate order if $h_2(\theta) \geq h_1(\theta)$ for all θ, which is also equivalent to $(1 - G_1)/(1 - G_2)$ to be decreasing. See Shaked and Shanthikumar (2007, p. 43).

[11] See footnote 10.

To prove that conditions (i) and (iii) imply the inequality, we show that

$$\frac{h_2' - h_1'}{h_2 - h_1} \leq \frac{h_2'}{h_2} \quad \text{and} \quad \frac{h_2'}{h_2} \leq \frac{h_{21}'/p_1 + h_2'/p_2}{h_{21}/p_1 + h_2/p_2}. \tag{10}$$

Indeed, the left inequality holds if and only if $(h_2 - h_1)/h_2$ is decreasing, which holds by assumption (iii). The right inequality holds if and only if $h_2/(h_{21}/p_1 + h_2/p_2)$ is decreasing, or, equivalently, if $h_{21}/h_2 = g_2/g_1$ is increasing, which holds by assumption (i). This completes the proof of Claim 3.

Claim 4: At any solution to $\mathscr{R}$, (IC_1) is slack, that is, $\int_{\underline{\theta}}^{\overline{\theta}} [x_2(\theta) - x_1(\theta)](1 - G_1(\theta))\, d\theta \leq 0$.

Indeed, let $\Delta(\theta) = [x_2(\theta) - x_1(\theta)](1 - G_2(\theta))$. By Claim 3, Δ crosses 0 exactly once, say at θ_0, and from below $(\underline{\theta}, \bar{\theta})$. Moreover, by Claim 1, $\int_{\underline{\theta}}^{\overline{\theta}} \Delta(\theta)\, d\theta = 0$. Therefore,

$$\begin{aligned}
\int_{\underline{\theta}}^{\overline{\theta}} [x_2(\theta) - x_1(\theta)](1 - G_1(\theta))\, d\theta &= \int_{\underline{\theta}}^{\overline{\theta}} \Delta(\theta) \frac{1 - G_1(\theta)}{1 - G_2(\theta)}\, d\theta - \int_{\underline{\theta}}^{\overline{\theta}} \Delta(\theta) \frac{1 - G_1(\theta_0)}{1 - G_2(\theta_0)}\, d\theta \\
&= \int_{\underline{\theta}}^{\theta_0} \Delta(\theta) \left[\frac{1 - G_1(\theta)}{1 - G_2(\theta)} - \frac{1 - G_1(\theta_0)}{1 - G_2(\theta_0)} \right] d\theta \qquad (11) \\
&\quad + \int_{\theta_0}^{\overline{\theta}} \Delta(\theta) \left[\frac{1 - G_1(\theta)}{1 - G_2(\theta)} - \frac{1 - G_1(\theta_0)}{1 - G_2(\theta_0)} \right] d\theta. \qquad (12)
\end{aligned}$$

We now argue that each of the two integrals in (11) and (12) is negative. Consider the integral in (11). By definition, Δ is negative for all $\theta \leq \theta_0$. Moreover, $(1 - G_1)/(1 - G_2)$ is decreasing by assumption (i).[12] Thus, the term in the square brackets under the integral is positive for all $\theta \leq \theta_0$. Consequently, the first integral is negative. The argument for the integral in (12) is analogous. Hence, we have established that (IC_1) is slack.

Claim 5. The previous four claims imply that any solution to $\mathscr{R}$ is a solution to $\mathscr{P}$ under the constraint that $v_1(\underline{\theta}) = v_2(\underline{\theta}) = 0$. To complete the proof of Proposition 1, we thus have to show that any solution to $\mathscr{R}$ is also a solution to $\mathscr{P}$ under the weaker constraints that $v_1(\underline{\theta}) \geq 0$ and $v_2(\underline{\theta}) \geq 0$. But this can be shown with identical steps as in Step 3 in the proof of Proposition 5 in Krähmer and Strausz (2015). This completes the proof. Q.E.D.

References

Courty, P. and H. Li (2000). “Sequential Screening." *Review of Economic Studies* 67, 697–717.

[12]See footnote 10.

Dai, C., T. R. Lewis, and G. Lopomo (2006).“Delegating Management to Experts." *RAND Journal of Economics* 37, 503–520.

Krähmer, D. and R. Strausz (2015). “Optimal Sales Contracts with Withdrawal Rights." *Review of Economic Studies* 82, 762–790.

Krähmer, D. and R. Strausz (2015b). “Dynamic Mechanism Design." Chapter 11 in ed. T. Börgers, *An Introduction to the Theory of Mechanism Design*, 204–234.

Luenberger, F. (1969). *Optimization by Vector Space Methods.*, John Wiley & Sons: New York.

Milgrom, P. and I. Segal (2002). “Envelope Theorems for Arbitrary Choice Sets." *Econometrica* 70, 583–601.

Shaked, M. and J.G. Shanthikumar (2007). *Stochastic Orders.* Springer: New York.