The Benefits of Sequential Screening

Daniel Krähmer*and Roland Strausz†

First version: October 12, 2011
This version: October 12, 2011

Abstract

This paper considers the canonical sequential screening model and shows that when the agent has an ex post outside option, the principal does not benefit from eliciting the agent's information sequentially. Unlike in the standard model without ex post outside options, the optimal contract is static and conditions only on the agent's aggregate final information. The benefits of sequential screening in the standard model are therefore due to relaxed participation rather than relaxed incentive compatibility constraints. We argue that in the presence of ex post participation constraints, the classical, local approach fails to identify binding incentive constraints and develop a novel, inductive procedure to do so instead. The result extends to the multi–agent version of the problem.

Keywords: Sequential screening, dynamic mechanism design, participation constraints, Mirrlees approach
JEL codes: D82, H57

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

†Humboldt-Universität zu Berlin, Institute for Microeconomic Theory, Spandauer Str. 1, D-10178 Berlin (Germany), strauszr@wiwi.hu-berlin.de. We thank Marco Battaglini, Hao Li, and seminar participants at UBC. Both authors gratefully acknowledge financial support by the DFG (German Science Foundation) under SFB/TR-15. Roland Strausz further acknowledges support under DFG–grant STR991/2-2.

1 Introduction

This paper considers the canonical sequential screening model and shows that when the agent has an ex post outside option, the principal does not benefit from eliciting the agent's information sequentially. Unlike in the standard model with only ex ante outside options, the optimal contract is, instead, static and conditions only on the agent's aggregate final information.[1]

Introducing ex post participation constraints contributes to the understanding of dynamic adverse selection problems from both a conceptual and practical perspective. Conceptually, our approach allows us to identify the reason why, in the absence of ex post participation constraints, sequential screening is strictly better than static screening. Compared to a static screening model where all of the agent's information arrives ex ante, the constraints in a sequential screening problem are weaker for two reasons: First, sequential screening relaxes incentive compatibility constraints because it is easier to prevent the agent from lying about his ex ante information when he does not yet know his ex post information. Second, in the sequential model with only an ex ante outside option, the contract needs to give the agent his outside option only in expectation rather than for all possible contingencies of ex ante and ex post information as in the static model. Our result makes clear that the value of sequential screening in the standard model without ex post participation constraints arises solely from the second reason—relaxed participation constraints—rather than from the first reason—relaxed incentive constraints. This conclusion follows because ex post participation constraints affect only the participation constraints while leaving the incentive constraints unaffected.

Comparing sequential screening models with ex ante and ex post participation constraints in terms of information rents reveals striking qualitative differences. The results of Esö and Szentes (2007a,b) imply that when there are only ex ante participation constraints then the principal can extract at no cost the entire value of the agent's ex post information. Hence, the agent does not obtain any rents from his ex post, but only from his ex ante private information. In contrast, our result implies that with ex post participation constraints the agent receives information rents from both his ex ante and ex post private information.

[1] The (strict) optimality of sequential screening in the absence of ex post participation constraints has been most cleanly established in Courty and Li (2000) and features also in Baron and Besanko (1984), Battaglini (2005), Esö and Szentes (2007a, b), Dai et al. (2006), Krähmer and Strausz (2008, 2011), Inderst and Hoffmann (2009), Pavan et al. (2008).

This insight is reminiscent of Sappington (1983), who observes that in static adverse selection problems, when the agent's private information arrives after contracting and the agent cannot sustain ex post losses, the optimal contract is the same as when the agent's information arrives before contracting. Analogously, in our dynamic setup, the optimal contract with ex post participation constraints is the same as when *all* of the agent's information arrives before contracting. This dynamic extension of Sappington's result is, however, not obvious. In a static problem, ex post participation constraints reduce the set of implementable contracts to the same set as when the agent's information arrives ex ante. The result therefore follows directly from implementability considerations. For dynamic setups, we show in contrast that ex post participation constraints do not render sequential contracts infeasible.[2] The principal, however, does not benefit from offering a sequential contract. Hence, our result follows from optimality rather than implementability considerations.

We obtain our result in the canonical, unit good sequential screening framework with an arbitrary finite number of ex ante and a continuum of ex post agent types and non–shifting support. In particular, we consider a procurement context where the principal seeks to acquire a good from the agent who, while observing a private signal ex ante, learns his true costs only as the relation proceeds. Without ex post participation constraints, the optimal contract can be implemented by a menu of option contracts. An option contract consists of a (possibly negative) *up–front payment* from the principal to the agent, and gives the agent the option to deliver the good at a pre–specified *exercise price* after having observed his true costs. Because, ex ante, agent types have different priors about the likelihood of exercising the option, the principal can screen the agent's prior by offering different combinations of up–front payments and exercise prices.

Our result implies that offering a menu with a variety of different option contracts is no longer optimal in the presence of ex post participation constraints. To see the reason for this, assume to the contrary that, at the optimum, different ex ante types select different option contracts. Observe first that when the agent's true ex post costs happen to equal the exercise price, the agent is indifferent between production and not and, thus, obtains no additional payoff from production. Therefore, with ex post participation constraints, the principal cannot demand an up–front fee, because it would imply an ex post loss for the agent if his true costs equal the exercise price. Clearly, not all contracts in the optimal menu can have positive up–

[2]Indeed, Courty and Li (2000)'s footnote 8 can be understood in this way.

front payments to the agent, because by lowering all of them slightly the principal could do better. Now consider an ex ante type who, at the optimum, selects a contract with a zero up–front payment. This type's contract must display the largest exercise price (otherwise he would have an incentive to pick a contract with a larger exercise price and take advantage of both the higher up–front payment and the higher exercise price). Call this contract with the zero up–front payment and the largest exercise price the "high price" contract.

Now, the contracts selected by the other ex ante types, by definition, have a smaller exercise price, therefore these types produce less frequently ex post than under the "high price" contract, and therefore the "high price" contract is more efficient. Moreover, by incentive compatibility they must get at least the same rent which they get if they choose the "high price" contract. But this implies that the principal is better off by offering the "high price" contract to all(!) agent types. Because then she has to pay at most the same rent as under the sequential contract, and production is more efficient. Therefore, with ex post participation constraints, it is not optimal to screen ex ante types, but instead offer only a single (i.e. static) contract.[3]

The previous reasoning only applies to option contracts. The core analytical challenge of our paper is to show that option contracts are optimal. In the absence of ex post participation constraints, the optimality of option contracts can be established by considering a relaxed problem which only considers the "local" ex ante incentive constraints in the spirit of Mirrlees. Under appropriate regularity conditions, the solution to the relaxed problem is automatically monotone in the ex post type and thus ex post incentive compatible. In the unit good framework, monotonicity in the ex post type implies that the good, depending on type, is produced with probability of either zero or one, which, in turn, implies that the contract can be implemented as an option contract. We argue that in our case, such a Mirrleesian–type, "local" approach does not work, since the solution to the corresponding relaxed problem is not automatically monotone in the ex post type. Instead, as the main methodological contribution of the paper, we develop an inductive procedure to identify the binding "global" constraints. The procedure reduces a model with $n+1$ ex ante types to a model with n ex ante types by merging the two

[3] This argument fails if there are only ex ante participation constraints. At the optimum, the principal then charges an up–front fee for the contract with the largest exercise price. Only the agent who is most optimistic about his future costs chooses this contract. The agent type with the most pessimistic prior would make an *expected* loss from this contract. Thus, only offering the "high price" contract would violate ex ante participation constraints.

most extreme ex ante types of the $n+1$ types model. The set of binding incentive constraints for the $n+1$ types model is then obtained by adding an appropriate constraint to those constraints of the n types model which are, by induction, known to be binding.

Investigating the consequences of ex post participation constraints also contributes to a better understanding of real–world contracts. Ex post participation constraints are empirically relevant, since the principal's ability to inflict ex post losses on the agent is often greatly limited in practice. In employment relations, employees typically have the legal right to leave their employer at will. Such non–slavery conditions imply that the employer cannot inflict losses on her employee and must respect the employee's ex post participation constraints. Alternatively, ex post participation constraints arise because workers are credit or wealth constrained. The relevance of ex post participation constraints is even more compelling in procurement relationships, where the agent as a corporation is legally protected by limited liability and, therefore, cannot make losses. Indeed, procurement contracts that inflict losses on the agent simply drive him out of business, leaving the contract unfulfilled. Similarly, legally granted money–back guarantees give consumers the right to return the good and being fully refunded. In the mail order business in Germany, for example, sellers are required by law to grant consumers a full refund (including all postal charges) up to 14 days after purchase.

Our analysis predicts that in the presence of ex post participation constraints, "simple" contracts are optimal, thus providing a rationale for "incomplete" contracts which depend only on the agent's final information instead of on the entire contingent information flow which the agent observes.[4] Likewise, our results imply for multi–agent versions of our setup that standard, static auctions are optimal even when agents obtain their private information sequentially. Indeed, Esö and Szentes (2007b) show that without ex post participation constraints, the optimal contract with multiple agents is a "handicap auction" where in the first round, bidders pick a premium from a menu offered by the auctioneer, and in the second round, bidders play a second price auction where the winner pays the second highest bid plus his premium from round 1. We argue that with ex post participation constraints, the optimal mechanism is static and thus a second price auction with an optimal reserve price.

The rest of this paper is organized as follows. The next section introduces the setup. In section 3, we derive the principal's problem. In section 4, we discuss three benchmark cases. In

[4] For a similar result in a dynamic adverse selection model with contractible ex post information, see Chiu and Sappington (2010).

section 5, we illustrate the main argument and the main intuition behind our result in the case when the agent's ex ante information is binary. In section 6, we solve the principal's problem for the general case and derive our main result. Section 7 discusses extensions, and section 8 concludes.

2 The Setup

Consider a principal (she) who seeks to buy a good or service from an agent (he).[5] The value of the good for the principal is commonly known to be $v > 0$. The agent's costs of production are $\theta \in [0, 1]$. The terms of trade are the probability with which production takes place, $x \in [0, 1]$, and a payment $t \in \mathbb{R}$ from the principal to the agent.

Parties are risk–neutral and have quasi–linear utility functions. That is, under the terms of trade x and t, the principal receives utility $vx - t$, and the agent receives utility $t - \theta x$. Consequently, the *aggregate surplus* is $(v - \theta)x$.

At the time of contracting about the terms of trade, no party knows the true costs, θ, but the agent has private information about the distribution of costs. After the principal offers the contract but before production takes place, the agent privately learns the true costs θ. Formally, there are two periods. In period 1, the agent knows that costs are distributed according to distribution function G_i with non–shifting support [0,1], where i is drawn from the set $\omega \equiv \{1, \dots, n\}$ with probability $p_i > 0$. We refer to i as the agent's *ex ante type*. In period 2, the agent observes his *ex post type* θ which is drawn according to G_i. While the agent's ex ante and ex post types are his private information, the distributions of ex ante and ex post types are common knowledge.

We depart from the existing sequential screening literature and consider the case in which agent can always quit after learning the true costs θ and receive an ex post outside option. We assume that the outside option is type–independent and normalize it to zero.

Next, we state our distributional assumptions and introduce notation. The probability density $g_i(\theta) = G_i'(\theta)$ exists, is differentiable, and is strictly positive for all $\theta \in [0, 1]$. Moreover,

[5] The setup is isomorphic to a buyer seller relationship where the principal acts as a seller with commonly known marginal costs and the agent as a buyer with private information about his willingness to pay.

we define by

$$h_{i,j}(\theta) \equiv \frac{G_i(\theta)}{g_j(\theta)}, \qquad \text{and} \qquad h_i(\theta) \equiv h_{i,i}(\theta),$$

the *cross hazard rate* between the types i and j and the *hazard rate* of type i. We assume that:

$$h_{i,j} \text{ and } h_i \text{ are non–decreasing in } \theta \text{ for all } i, j.$$

We define conditional distributions associated with a subset of types. For $\gamma \subseteq \omega$, let $p_\gamma = \sum_{i\in\gamma} p_i$ be the probability of γ, and define by

$$G_\gamma(\theta) \equiv \frac{1}{p_\gamma}\sum_{i\in\gamma} p_i G_i(\theta), \quad g_\gamma(\theta) \equiv \frac{1}{p_\gamma}\sum_{i\in\gamma} p_i g_i(\theta), \quad h_\gamma(\theta) \equiv \frac{G_\gamma(\theta)}{g_\gamma(\theta)} \tag{1}$$

the *conditional distribution*, *conditional density*, and *conditional hazard rate conditional* on the event that the ex ante type is in γ. Moreover, define for two subsets $\gamma, \delta \subseteq \omega$ the *conditional cross hazard rate*

$$h_{\gamma,\delta}(\theta) \equiv \frac{G_\gamma(\theta)}{g_\delta(\theta)}.$$

Monotonicity of the (cross) hazard rates carries over to the conditional (cross) hazard rates:

Lemma 1 *$h_{\gamma,\delta}$ is non–decreasing in θ for all $\gamma, \delta \subseteq \omega$.*

For each type i, We define the *ex post cutoff type* θ_i implicitly by

$$v = \theta_i + h_i(\theta_i). \tag{2}$$

Because the hazard rate is non–decreasing, there is at most one solution to (2). Without loss of generality, we label the ex ante types according to the order of the ex post cutoff types:

$$\theta_1 \leq \ldots \leq \theta_i \leq \ldots \leq \theta_n.$$

We extend the definition of ex post cutoff type to subsets γ of types by defining θ_γ as the solution to

$$v = \theta_\gamma + h_\gamma(\theta_\gamma) \tag{3}$$

which, by Lemma 1, is unique. The cutoff θ_γ displays an averaging feature in the sense that it lies in between the lowest and highest cutoffs associated to the types in γ:

Lemma 2 *Let $\gamma, \delta \subset \omega$ be disjoint. Then* $\theta_{\gamma \cup \delta} \in [\min\{\theta_\gamma, \theta_\delta\}, \max\{\theta_\gamma, \theta_\delta\}]$.

We close this section with the following remarks about our modeling setup:

Remark 1: As is standard in the literature on sequential screening, the agent's ex ante private information does not shift the support of his final ex post type. This non–shifting support assumption facilitates the characterization of incentive compatibility off the equilibrium path.[6] Moreover, the agent's ex ante type i is payoff–irrelevant in the sense that it does not directly affect the final cost type θ. This assumption is, however, without loss of generality, because if final costs are given as a function $\theta(i, s)$ of both the agent's ex ante information i and some ex post information s that he receives in period 2, then we can redefine the agent's ex post type as the value of the random variable $\theta(i, s)$.

Remark 2: Non–decreasing hazard rates h_i are a standard assumption in static screening models, because they ensure that solutions "automatically" exhibit a monotonicity property. To obtain an analogous property in our setting, we also require non–decreasing cross hazard rates. This is satisfied for large and natural families of distributions. It essentially requires that the cumulative distributions increase faster than the densities. Hence, a sufficient condition is that densities are non–increasing.

Remark 3: Our ranking of ex ante types by their ex post cutoff type θ_i is simply a labeling convention. It does not imply any restrictions on the stochastic order ranking of the distributions G_i. In particular, our result does not require that the distributions G_i be ranked in terms of first or second order stochastic dominance, as is the case in standard sequential screening models such as Courty and Li (2000). However, in the special case that the hazard rates h_i are decreasing in i, it is well–known that G_j first order stochastically dominates G_i for $\theta_i > \theta_j$.

3 Principal's problem

The principal's problem is to design a contract that maximizes her expected utility. In this section, we describe the principal's problem formally. Because the agent has private information, the terms of trade optimally depend on communication by the agent to the principal. By the revelation principle for sequential games (e.g., Myerson 1986), the optimal contract can be found in the class of direct and incentive compatible contracts which induces the agent to

[6] See Krähmer and Strausz (2008) for an elaboration of this point.

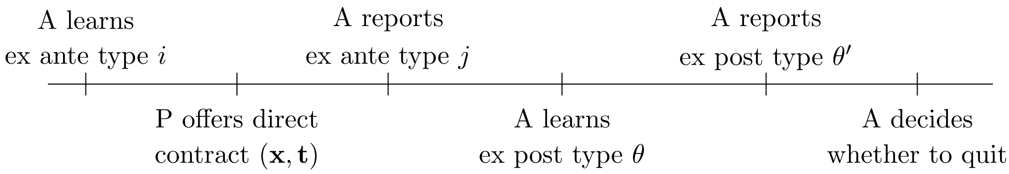

Figure 1: Time line

report his type truthfully at the ex ante as well as at the ex post stage. Formally, a *direct contract*

$$(\mathbf{x}, \mathbf{t}) = (x_j(\theta'), t_j(\theta'))_{j\in\omega,\theta'\in[0,1]}$$

requires the agent to report an ex ante type j in period 1, and an ex post type θ' in period 2. A contract commits the principal to a *production schedule* $x_j(\theta')$ and a *transfer schedule* $t_j(\theta')$. A direct contract induces a game with a timing structure as illustrated in Figure 1.

If the agent's true ex post type is θ and his period 1 report was j, then his utility from reporting θ' in period 2 is

$$u_j(\theta'; \theta) \equiv t_j(\theta') - \theta x_j(\theta').$$

With slight abuse of notation, we denote the agent's period 2 utility from truth–telling by

$$u_j(\theta) \equiv u_j(\theta; \theta).$$

The contract is *incentive compatible in period 2* if it gives the agent an incentive to announce his ex post type truthfully. That is, if for all $j \in \omega$,

$$u_j(\theta) \geq u_j(\theta'; \theta) \quad \text{for all } \theta, \theta' \in [0, 1]. \tag{4}$$

If the contract is incentive compatible in period 2, the agent announces his ex post type truthfully no matter what his report in the first period.[7] Hence, if the agent's true ex ante type is i, then his period 1 utility from reporting j is

$$U_{ji} \equiv \int_0^1 u_j(\theta)\, dG_i(\theta).$$

[7] Observe that the fact that agent's period 2 utility is independent of his ex ante type implies that 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 agent has misreported his ex ante type in period 1. Observe also that, in general, optimality does not require truth–telling off the path. See Krähmer and Strausz (2008) for an elaboration of this point.

We denote, again with a slight abuse of notation, the agent's period 1 utility from truth–telling by

$$U_i = U_{ii}.$$

The contract is *incentive compatible in period 1* if it gives the agent an incentive to announce his ex ante type truthfully:

$$U_i \geq U_{ji} \quad \text{for all } i, j \in \omega. \tag{5}$$

To ensure the agent's participation for all cost realizations, the contract needs to satisfy the *ex post individual rationality constraint*:

$$u_i(\theta) \geq 0 \quad \text{for all } i \in \omega, \theta \in [0,1]. \tag{6}$$

In contrast, an incentive compatible contract is *ex ante individually rational* if

$$U_i \geq 0 \quad \text{for all } i \in \omega. \tag{7}$$

Clearly, ex post individual rationality implies ex ante individual rationality. We say a contract is *feasible* if it is incentive compatible (in both periods) and ex post individually rational.

By definition, the principal's payoff from a feasible contract is the difference between aggregate surplus and the agent's utility. That is, if the agent's ex ante type is i, the principal's conditional expected payoff is

$$W_i \equiv \int_0^1 \{[v - \theta]x_i(\theta) - u_i(\theta)\} \; dG_i(\theta),$$

so that the principal's expected payoff is

$$W \equiv \sum_{i \in \omega} p_i W_i.$$

The principal's problem is therefore to find a direct contract $(\mathbf{x}^*, \mathbf{t}^*)$ that solves the following maximization problem:

$$\max_{(\mathbf{x},\mathbf{t})} W \quad s.t. \qquad (4), (5), (6).$$

3.1 Eliminating transfers from the principal's problem

Our approach to solving the principal's problem is to follow standard procedures of static screening problems as closely as possible. Because for a given first period report, the second period incentive compatibility constraints are the same as in a static screening problem, our first step is to exploit the fact that second period incentive compatibility pins down the agent's utility as a function of the allocation x alone. This yields the familiar result that incentive compatibility is equivalent to monotonicity of the production schedule and to "revenue equivalence", which means that the agent's utility is determined by the production schedule up to a constant. We state this standard result without proof.

Lemma 3 *For all $i \in \omega$, there are transfers $t_i(\theta)$ so that second period incentive compatibility (4) is equivalent to*

$$x_i(\theta) \text{ is non–increasing in } \theta, \tag{MON}$$

$$u_i(\theta) = \int_\theta^1 x_i(z)\, dz + u_i(1). \tag{RE}$$

Lemma 3 has three useful implications. First, we can replace the second period incentive constraints (4) in the principal's problem by the constraints (MON) and (RE). We can then eliminate the constraint (RE) by inserting $u_i(\theta)$ directly in the principal's objective. After an integration by parts, the principal's objective transforms into the familiar expected virtual surplus minus the agent's utility of the least efficient ex post type $\theta = 1$:

$$W_i = \int_0^1 [v - \theta - h_i(\theta)] x_i(\theta)\, dG_i(\theta) - u_i(1). \tag{8}$$

The second implication of Lemma 3 is that (RE) also pins down the agent's period 1 utility U_{ji}, which is simply the expectation over period 2 utility. Applying integration by parts, we arrive at the following characterization of the first period incentive constraints:

Lemma 4 *Consider a contract which satisfies (RE). Then first period incentive compatibility (5) is equivalent to*

$$\int_0^1 [x_i(\theta) - x_j(\theta)] G_i(\theta)\, d\theta + u_i(1) - u_j(1) \geq 0 \quad \text{for all } i, j \in \omega. \tag{IC_{ij}}$$

The third useful implication of Lemma 3 is that because x_i is non–decreasing, the agent's ex post utility $u_i(\theta)$ is non–increasing in his ex post type θ. Thus, ex post individual rationality is satisfied for all types if it holds for the highest type $\theta = 1$:

Lemma 5 *Consider a contract which satisfies (MON) and (RE). Then ex post individual rationality (6) is equivalent to*

$$u_i(1) \geq 0 \text{ for all } i \in \omega. \tag{IR_i}$$

By the previous three lemmas, the following equivalent representation of the principal's problem obtains when we replace the payment $\mathbf{t}$ by the vector $\mathbf{u} = \{u_i(1)\}_{i\in\omega}$ of utilities of the highest ex post type:

$$\begin{aligned} \mathcal{P}: \qquad & \max_{\mathbf{x},\mathbf{u}} \sum_{i\in\omega} p_i \int_0^1 [v - \theta - h_i(\theta)] x_i(\theta)\, dG_i(\theta) - p_i u_i(1) \\ & s.t. \quad (MON),\ (IC_{ij}),\ (IR_i). \end{aligned}$$

Before solving $\mathcal{P}$, it is helpful to introduce some more notation. With slight abuse of notation, we also refer to a pair $(\mathbf{x}, \mathbf{u})$ as a contract. We define a *cutoff schedule with cutoff* $\hat{\theta} \in [0, 1]$ as

$$\bar{x}(\theta|\hat{\theta}) \equiv \begin{cases} 1 & \text{if } \theta \leq \hat{\theta}, \\ 0 & \text{otherwise.} \end{cases}$$

We say that a contract $(\mathbf{x}, \mathbf{u})$ is a *cutoff contract* if each production schedule $x_i(\theta)$ coincides with some cutoff schedule with cutoff $\hat{\theta}_i$. Note that a cutoff contract can be indirectly implemented by a menu of *option contracts*, which consists of an *up–front payment* that the agent receives in period 1 and an *exercise price* which the agent only receives when he decides to produce the good in period 2. To see this note that, under a cutoff contract, the agent is required to produce if he reports an ex post type below $\hat{\theta}_i$ after having announced an ex ante type i. In this case, the ex post type θ obtains utility $\hat{\theta}_i - \theta + u_i(1)$. If, instead, he reports a type above $\hat{\theta}_i$, the agent does not produce and obtains utility $u_i(1)$. Hence, a cutoff contract $(\mathbf{x}, \mathbf{u})$ can be implemented by the menu of $i = 1, \ldots, n$ option contracts with the up–front payment $u_i(1)$ and the exercise price $\hat{\theta}_i$. In what follows, we use the notions of cutoff and option contracts synonymously, whichever interpretation is more convenient.

4 Benchmarks

In this section we discuss three benchmark cases that will play a crucial role in the subsequent analysis. First, we consider the principal's problem when the agent's ex ante type is publicly known. Second, we consider the optimal "static" contract whose terms of trade do not depend on the agent's ex ante information. This latter contract describes the optimal contract when the principal does not engage in sequential screening, but offers the contract only after the agent has obtained all his private information. It is clear that the principal's payoff from an optimal contract lies in between these two benchmarks. Finally, we review the optimal sequential screening contract when the principal has to respect *ex ante* rather than *ex post* participation constraints.

4.1 Publicly known ex ante types

When the agent's ex ante type is publicly known, the incentive constraints (IC_{ij}) are redundant. Absent these constraints, the ex post individual rationality constraints (IR_i) are binding at the optimum. If we now disregard the monotonicity constraint, pointwise maximization of the principal's objective yields that the optimal production schedule is the cutoff schedule with cutoffs θ_i as defined in (2). In particular, it satisfies the monotonicity constraint and must, therefore, be optimal. The next lemma summarizes.

Lemma 6 *If the agent's ex ante type is public information, the optimal contract is a cutoff contract characterized by* $u_i^p(1) = 0$ *and*

$$x_i^p(\theta) = \bar{x}(\theta|\theta_i) \quad \forall i \in \omega.$$

In other words, if the agent's ex ante type i is public information, the principal's problem is that of a unit good monopsonist facing the supply function G_i. At the optimal contract, the transfer is equal to the ex post cutoff type θ_i, and the good is produced whenever costs are smaller than θ_i.

4.2 Optimal static contract

We refer to a contract as static if the contract does not condition on the agent's ex ante type: $x_i = x_j \equiv x^s$ and $u_i(1) = u_j(1) \equiv u^s(1)$ for all $i, j \in \omega$. The principal's objective under a static contract is

$$W^s = \int_0^1 [v - \theta - h_\omega(\theta)]x^s(\theta)\, dG_\omega(\theta) - u^s(1),$$

where h_ω and G_ω are defined in (1) for $\gamma = \omega$.

Under a static contract, the incentive constraints (IC_{ij}) hold trivially, and it follows from inspection of $\mathcal{P}$, that at the optimum, the ex post individual rationality constraints are binding. Observe that the solution to the unconstrained problem which simply maximizes the principal's objective is given by the cutoff schedule with cutoff θ_ω. In particular, it satisfies the monotonicity constraint and is thus a solution to the constrained problem. The next lemma summarizes.

Lemma 7 *The optimal static contract is a cutoff contract characterized by* $u^s(1) = 0$ *and a cutoff schedule*

$$x^s(\theta) = \bar{x}(\theta|\theta_\omega).$$

In other words, if the principal can only offer a static contract, her problem is that of a unit good monopsonist facing the average supply function G_ω. At the optimal contract, the transfer is equal to the critical type θ_ω, and the good is produced whenever costs are smaller than θ_ω.

4.3 Ex ante participation constraints

The main benchmark for our analysis is the standard sequential screening model where the principal has to respect only the ex ante participation constraints (7) rather than the ex post participation constraints (IR_i). In contrast to our main result, the principal does benefit from sequential screening in this case, as shown by Courty and Li (2000). We now review this important benchmark.

Courty and Li (2000) identify conditions so that the principal's problem can be solved by a "Mirrleesian" approach. That is, the optimal contract obtains from solving a relaxed problem with only the participation constraint for the highest type $i = n$, and all "local downward" incentive constraints $IC_{i,i+1}$. One of the identified conditions is that the distributions G_i are ordered in the sense of first order stochastic dominance.[8] Courty and Li (2000) further show that if, in addition, the solution to the relaxed problem exhibits a production schedule that is monotone in *both* the ex ante and ex post type, then it represents also a solution to the original problem. The need for monotonicity puts additional restrictions on the primitives of the model.

[8] An alternative condition is that the distributions display a particular kind of mean preserving spreads. Combinations of this MPS–ordering and first order dominance are also fine, but second order dominance in general does not work.

The Mirrleesian approach implies that the solution to the relaxed problem exhibits deterministic production schedules $\mathbf{x}^{CL}$ which equal 1 whenever the aggregate surplus exceeds the hazard rate multiplied with an informativeness measure[9]:

$$x_i^{CL}(\theta) = 1 \quad \Leftrightarrow \quad v - \theta \geq \hat{h}_i(\theta) \equiv \frac{p_1 + \ldots + p_{i-1}}{p_i} \cdot \frac{G_{i-1}(\theta) - G_i(\theta)}{g_i(\theta)}.$$

Hence, the remaining question is under which conditions the schedules $\mathbf{x}^{CL}$ are monotone in the ex post type θ and the ex ante type i. For the schedules to be monotone in θ, they must be cutoff schedules with a cutoff θ_i^{CL} that is the unique solution to $v - \theta = \hat{h}_i(\theta)$. A sufficient condition for existence and uniqueness of θ_i^{CL} is that $\hat{h}_i(\theta)$ is convex in θ and $v \leq 1$.

For cutoff schedules to be monotone in the ex ante type, the cutoffs are required to be decreasing: $\theta_n^{CL} \leq \ldots \leq \theta_1^{CL}$. A sufficient condition to obtain this ordering is that $\hat{h}_i(\theta)$ is increasing in i. The following lemma summarizes.

Lemma 8 *Suppose G_i dominates G_{i-1} in the sense of first order stochastic dominance for all $i = 2, \ldots, n$, that $\hat{h}_i(\theta)$ is convex in θ and increasing in i, and that $v \leq 1$. Then, with ex ante participation constraints, the optimal contract $(\mathbf{x}^{CL}, \mathbf{u}^{CL})$ exhibits productions schedules that are characterized by the cutoff schedule $x_i^{CL}(\theta) = \bar{x}(\theta|\theta_i^{CL})$ where θ_i^{CL} is the unique solution to*

$$\theta_1^{CL} = v, \qquad v - \theta_i^{CL} = \hat{h}_i(\theta_i^{CL}) \quad \forall i > 1.$$

Interestingly, the contract $(\mathbf{x}^{CL}, \mathbf{u}^{CL})$ violates *all* ex post participation constraints. To see this note that because type n's ex post utility at the least efficient ex post type, $u_n(1)$, is pinned down by (RE) and the binding ex ante participation constraint (7), it follows that $u_n(1) < 0$. Because $u_i(1)$ is pinned down by the binding incentive constraint $IC_{i,i+1}$, the ordering of the cutoffs implies that the lowest ex ante type gets the lowest utility at the least efficient ex post type:

$$0 > u_n(1) > \ldots > u_1(1).$$

This ordering also reveals the intuition why sequential screening is strictly optimal as well as the role of stochastic dominance: Because lower ex ante types are less likely to become high ex

[9] Courty and Li (2000) present a continuous version of this measure, while Dai et al. (2006) present it for the case with two ex ante types. Baron and Besanko (1984) were the first to interpret the second term as an informativeness measure of the ex ante information.

post types, they are more willing to tolerate higher losses for higher ex post types. The optimal screening contract with ex ante participation constraints exploits this feature. It screens ex ante types by imposing higher ex post losses on lower ex ante types.

5 Two ex ante types

The main result of this paper is that, with ex post participation constraints, the optimal sequential screening contract coincides with the static one. To gain intuition for this result, we analyze in this section the case with two ex ante types. To simplify the exposition, we assume in this section that $v = 1$.

Our approach to solving the principal's problem is to consider an appropriate relaxed problem and to show that its solution also solves the original problem. As in standard screening problems, we ignore, first, the monotonicity constraint. Second, we ignore the "upward" incentive constraint (IC_{21}), because the solution to the problem with publicly known ex ante type violates only the "downward" incentive constraint (IC_{12}).[10] Hence, we consider the relaxed problem

$$\mathcal{R}: \qquad \max_{x_1,x_2,u_1(1),u_2(1)} p_1 \int_0^1 [1-\theta-h_1(\theta)]x_1(\theta)\, dG_1(\theta) - p_1 u_1(1)$$
$$+p_2 \int_0^1 [1-\theta-h_2(\theta)]x_2(\theta)\, dG_2(\theta) - p_2 u_2(1) \qquad s.t.$$

$$\int_0^1 [x_1(\theta)-x_2(\theta)]G_1(\theta)\, d\theta + u_1(1) - u_2(1) \geq 0, \qquad (IC_{12})$$

$$u_1(1) \geq 0, \quad u_2(1) \geq 0. \qquad (IR_i)$$

We now argue that the solution to $\mathcal{R}$ is given by the optimal static contract. It will then also be a solution to the original problem $\mathcal{P}$, because the static contract trivially satisfies all neglected constraints. The argument has two steps. First, we argue that, for any fixed levels $u_1(1)$ and $u_2(1)$, the optimal production schedule must be a cutoff schedule. Then we optimize over $u_1(1)$ and $u_2(1)$ and all possible cutoffs to show that the optimal contract is the optimal static one.

Keeping $u_1(1)$ and $u_2(1)$ fixed, IC_{12} is the only remaining constraint. By the Kuhn–Tucker

[10] In particular, $\int_0^1 [x_1^p(\theta) - x_2^p(\theta)]G_1(\theta)\, d\theta + u_1^p(1) - u_2^p(1) = -\int_{\theta_1}^{\theta_2} G_1(\theta) < 0$, since $\theta_1 < \theta_2$.

theorem[11], a solution to $\mathcal{R}$ maximizes the Lagrange function

$$\begin{aligned}\mathcal{L} \;=\; & p_1 \int_0^1 [1-\theta-h_1(\theta)]x_1(\theta)dG_1(\theta) - p_1u_1(1) + p_2 \int_0^1 [1-\theta-h_2(\theta)]x_2(\theta)dG_2(\theta) - p_2u_2(1) \\ & -\lambda\{\int_0^1 [x_1(\theta)-x_2(\theta)]G_1(\theta)\, d\theta + u_1(1) - u_2(1)\},\end{aligned}$$

where $\lambda \leq 0$ is the multiplier associated to the constraint IC_{12}. Re–arranging delivers

$$\begin{aligned}\mathcal{L} \;=\; & \int_0^1 \{p_1[1-\theta-h_1(\theta)] - \lambda h_1(\theta)\}x_1(\theta)g_1(\theta)\, d\theta - (p_1+\lambda)u_1(1) \\ & + \int_0^1 \{p_2[1-\theta-h_2(\theta)] + \lambda h_{12}(\theta)\}x_2(\theta)g_2(\theta)\, d\theta - (p_2-\lambda)u_2(1).\end{aligned}$$

Observe that we can maximize $\mathcal{L}$ point–wisely. In particular, the production schedules $x_1(\theta)$ and $x_2(\theta)$ are optimally set to 1 whenever the respective expressions in the curly brackets under the integrals,

$$p_1[1-\theta-h_1(\theta)] - \lambda h_1(\theta), \tag{9}$$

$$p_2[1-\theta-h_2(\theta)] + \lambda h_{12}(\theta), \tag{10}$$

are positive, and $x_1(\theta)$ and $x_2(\theta)$ are set to 0 otherwise. This implies that the production schedules are cutoff schedules if (9) and (10) are decreasing in θ. To see that this is indeed the case, recall that $\lambda \leq 0$. Together with h_2 and h_{12} non–decreasing, it then follows that (10) is decreasing in θ. Next consider (9). It is decreasing in θ if $p_1 + \lambda$ is non–negative, because h_1 is non–decreasing in θ. Now let $\hat{\theta}_1 \in [0,1]$ be such that (9) is zero. If $\hat{\theta}_1$ does not exist, then, because (9) is continuous in θ, x_1 is either 0 or 1 everywhere and, hence, a cutoff schedule with cutoff 0 or 1. Otherwise, we have

$$p_1[1-\hat{\theta}_1-h_1(\hat{\theta}_1)] - \lambda h_1(\hat{\theta}_1) = 0 \quad \Leftrightarrow \quad p_1 + \lambda = \frac{p_1(1-\hat{\theta}_1)}{h_1(\hat{\theta}_1)} \geq 0.$$

From this it follows that (9) is decreasing in θ, and $\hat{\theta}_1$ is therefore unique. Hence, also when $\hat{\theta}_1$ exists the optimal production schedules $x_1(\theta)$ and $x_2(\theta)$ are characterized by a cutoff schedule with respective cutoffs $\hat{\theta}_1$ and $\hat{\theta}_2$.

We now turn to the second step and look for the optimal cutoff schedules and utility levels. As argued above, the incentive constraint IC_{12} in problem $\mathcal{R}$ must be binding at the optimum,

[11] See Theorem 1 and 2 in Luenberger (1969, p.187–189).

because disregarding it would yield a solution that violates it (see footnote 10). Therefore, given cutoff schedules, the principal's problem $\mathcal{R}$ can be written as:

$$\mathcal{R}' : \quad \max_{\hat{\theta}_1, \hat{\theta}_1, u_1(1) \geq 0, u_2(1) \geq 0} p_1 \int_0^{\hat{\theta}_1} [1 - \theta - h_1(\theta)] \, dG_1(\theta) - p_1 u_1(1) \tag{11}$$

$$+ p_2 \int_0^{\hat{\theta}_2} [1 - \theta - h_2(\theta)] \, dG_2(\theta) - p_2 u_2(1)$$

$$s.t. \quad \int_{\hat{\theta}_1}^{\hat{\theta}_2} G_1(\theta) \, d\theta = u_1(1) - u_2(1). \tag{12}$$

This representation identifies the principal's fundamental trade-off. The principal may *screen* ex ante types by imposing a different cutoff for each type: $\hat{\theta}_1 \neq \hat{\theta}_2$. This allows her to fine–tune production to the types' different cost distributions. However, by (12), this is feasible only if at least one ex post participation constraint is not binding. In other words, screening ex ante type comes at the cost of giving at least one type a positive ex post utility level $u_i(1)$. We now show that this trade–off is unambiguously resolved in disfavor of screening.

In fact, inspecting $\mathcal{R}$ yields $u_2(1) = 0$ at any optimum, because otherwise lowering $u_2(1)$ would relax IC_{12} and raise the objective. But if $u_2(1) = 0$, then (12) together with the constraint that $u_1(1) \geq 0$ implies that only cutoffs with $\hat{\theta}_1 \leq \hat{\theta}_2$ are feasible. Substituting the constraint (12) with $u_2(1) = 0$ in the objective (11) yields

$$\begin{aligned} & p_1 \int_0^{\hat{\theta}_1} [1 - \theta - h_1(\theta)] \, dG_1(\theta) - p_1 \int_{\hat{\theta}_1}^{\hat{\theta}_2} G_1(\theta) \, d\theta + p_2 \int_0^{\hat{\theta}_2} [1 - \theta - h_2(\theta)] \, dG_2(\theta) \\ = \; & p_1 \int_0^{\hat{\theta}_1} [1 - \theta] \, dG_1(\theta) - p_1 \int_0^{\hat{\theta}_2} G_1(\theta) \, d\theta + p_2 \int_0^{\hat{\theta}_2} [1 - \theta - h_2(\theta)] \, dG_2(\theta). \end{aligned} \tag{13}$$

Notice that, in the second line, the second and the third term do not depend on $\hat{\theta}_1$, and the first term is expected aggregate surplus, conditional on facing type 1. Since aggregate surplus is maximized at $\hat{\theta}_1 = 1$, it is optimal to choose $\hat{\theta}_1$ as large as possible. Because of the restriction $\hat{\theta}_1 \leq \hat{\theta}_2$, it then follows that $\hat{\theta}_1 = \hat{\theta}_2$, or, in other words, that a static contract is optimal. Clearly, among all static contracts the optimal static contract solves the principal's problem. This illustrates our main result for the special case of two types: With ex post participation constraints it is feasible but not optimal for the principal to screen sequentially.

To shed more light on the role of ex post participation constraints, recall that we can interpret a contract $c_i = (\hat{\theta}_i, u_i(1))$ as an option contract with exercise price $\hat{\theta}_i$ and up–front payment $u_i(1)$. Screening ex ante types then corresponds to offering a menu with two different option contracts $c_1 \neq c_2$. To understand intuitively why the principal does not gain from screening ex ante types, suppose that c_2 is the optimal static contract with $\hat{\theta}_2 = \theta_\omega$ and $u_2(1) = 0$. Now observe that when the principal targets type 1 with an additional but different contract c_1, incentive compatibility requires that type 1 gets at least the same rent from c_1 as from c_2. Hence, the principal loses unambiguously from offering a contract c_1 with a smaller exercise price, because the smaller exercise price implies that c_1 is less efficient than c_2 so that on top of paying (at least) the same rent to the agent, c_1 also generates a smaller aggregate surplus.

On the contrary, it is not directly obvious that the principal loses from offering a contract c_1 with a larger, more efficient exercise price $\hat{\theta}_1 > \hat{\theta}_2$. The key observation which helps to understand this is that for $\hat{\theta}_1 > \hat{\theta}_2$ the incentive compatibility constraint (IC_{12}) is necessarily slack, because the up–front payment to type 1 cannot be negative. Hence, when the principal increases the exercise price $\hat{\theta}_1$ beyond $\hat{\theta}_2$, she faces exactly the standard monopsony trade–off between extending supply and paying a higher price, which, by definition, θ_1 solves optimally. But since $\theta_1 < \theta_\omega = \hat{\theta}_2$, raising the exercise price $\hat{\theta}_1$ beyond $\hat{\theta}_2$ is also suboptimal.

It is instructive to see where the previous argument fails when there are only ex ante participation constraints. Clearly, the same reasoning as above implies that it is suboptimal to offer a contract c_1 with a smaller exercise price $\hat{\theta}_1 < \theta_\omega$. But, with ex ante participation constraints, the argument is different for a contract with a higher exercise price $\hat{\theta}_1 > \theta_\omega$. In contrast to the case with ex post participation constraints, the principal can now impose a *negative* up–front payment $u_1(1) < 0$ on type 1. Therefore, she can use $u_1(1)$ to extract exactly that part of type 1's information rent that goes beyond what is needed to guarantee incentive compatibility.[12] In fact, for fixed c_2, it is then optimal to set $u_1(1)$ so that (IC_{12}) is binding. Unlike in the case with ex post participation constraints, increasing the exercise price $\hat{\theta}_1$ does therefore no longer go along with increasing type 1's rent. Consequently, it is optimal to set the exercise price to maximize aggregate surplus, thus $\hat{\theta}_1 = 1$, in accord with the optimal exercise price $\theta_1^{CL} = 1$ from the benchmark in section 4.3.

[12] We may interpret that part of type 1's information rent that goes beyond what is needed to guarantee incentive compatibility as the agent's *ex post information rent*, because it results from the fact that *all* ex post types who produce the good obtain the higher exercise price. That the principal can use the up–front payment to fully extract this ex post information rent is equivalent to Esö and Szentes' (2007a,b) observation that the principal wants to disclose the maximal amount of ex post information available.

6 Arbitrary number of ex ante types

In this section, we extend the result of the previous section to the environment with an arbitrary number of ex ante types. The extension is not straightforward, because in contrast to the two type case, where there are only local incentive constraints, we now have to deal with both local and global incentive constraints. It turns out that, in contrast to sequential screening models with ex ante participation constraints, we cannot use a "Mirrleesian" approach of focusing on local constraints. Indeed, the major challenge in extending our result lies in identifying the relevant incentive constraints.

6.1 Auxiliary problem: $u_i(1) = 0$

We begin by considering the problem when the utilities of the least efficient ex post types, $u_i(1)$, are exogenously set to 0. In the next subsection, we argue that this is indeed optimal. Thus, we first consider the problem

$$\mathcal{P}^0: \qquad \max_{\mathbf{x}} \sum_{i \in \omega} p_i \int_0^1 [v - \theta - h_i(\theta)] x_i(\theta)\, dG_i(\theta) \qquad s.t.$$

$$x_i(\theta) \text{ is non--increasing in } \theta, \tag{MON}$$

$$\int_0^1 [x_i(\theta) - x_j(\theta)] G_i(\theta)\, d\theta \geq 0 \quad \text{for all } i, j \in \omega. \tag{IC^0_{ij}}$$

Our approach to solving $\mathcal{P}^0$ is to solve a relaxed problem where we ignore the monotonicity constraints and consider only a subset of incentive constraints. The main challenge is to identify the relevant incentive constraints such that the solution to the relaxed problem will also be a solution to the original problem, that is, satisfy monotonicity and all ignored constraints.

We identify a subset of constraints IC^0_{ij} with the subset of respective indices (i, j). Let $C_0 \equiv \{(i, j) \in \omega^2 \mid i \neq j\}$. For a subset $C \subseteq C_0$, we denote by $\mathcal{R}^0(C)$ the relaxed problem where only the constraints in C are considered:

$$\mathcal{R}^0(C): \qquad \max_{\mathbf{x}} \sum_{i \in \omega} p_i \int_0^1 [v - \theta - h_i(\theta)] x_i(\theta)\, dG_i(\theta) \quad s.t. \quad IC^0_{ij} \text{ for all } (i, j) \in C.$$

To solve problem $\mathcal{R}^0(C)$, we will work with the Kuhn–Tucker theorem for function spaces. By Theorem 1 and 2 in Luenberger (1969, p.187–189), $\{x_k(\cdot)\}_{k \in \omega}$ solves $\mathcal{R}^0(C)$ if and only if there are multipliers $\lambda_{ij} \leq 0$ associated to constraint IC^0_{ij} such that $\{x_k(\cdot)\}_{k \in \omega}$ maximizes the

Lagrangian

$$\begin{aligned}\mathcal{L}^0(C) &= \sum_{k\in\omega}\int_0^1 p_k[v-\theta-h_k(\theta)]x_k(\theta)g_k(\theta)\,d\theta - \sum_{(i,j)\in C}\lambda_{ij}\int_0^1[x_i(\theta)-x_j(\theta)]G_i(\theta)\,d\theta \\ &= \sum_{k\in\omega}\int_0^1\left\{p_k[v-\theta-h_k(\theta)] - \sum_{j:(k,j)\in C}\lambda_{kj}h_k(\theta) + \sum_{i:(i,k)\in C}\lambda_{ik}h_{i,k}(\theta)\right\}x_k(\theta)g_k(\theta)\,d\theta,\end{aligned}$$

and, moreover, $\lambda_{ij}=0$ only if the inequality in IC^0_{ij} is strict.

By point–wise maximization, the Lagrangian $\mathcal{L}^0(C)$ is maximized if $x_k(\theta)$ is set to 1 whenever the expression in curly brackets under the integral,

$$\Psi_k(\theta,C)\equiv p_k[v-\theta-h_k(\theta)] - \sum_{j:(k,j)\in C}\lambda_{kj}h_k(\theta) + \sum_{i:(i,k)\in C}\lambda_{ik}h_{i,k}(\theta),$$

is positive, and $x_k(\theta)$ is set to 0 otherwise.[13] We summarize this observation in the following lemma.

Lemma 9 *The schedule $\{x_k(\cdot)\}_{k\in\omega}$ is a solution to $\mathcal{R}^0(C)$ if and only if for all $(i,j)\in C$ there is a λ_{ij} so that*

$$\lambda_{ij}\le 0, \tag{KT_1}$$

$$x_k(\theta)=\begin{cases}0 & if \quad \Psi_k(\theta,C)<0\\ 1 & if \quad \Psi_k(\theta,C)>0\end{cases}\qquad \forall k\in\omega, \tag{KT_2}$$

$$\lambda_{ij}\int_0^1[x_i(\theta)-x_j(\theta)]G_i(\theta)d\theta=0 \qquad \forall(i,j)\in C. \tag{KT_3}$$

The main result of this subsection is that the static contract solves problem $\mathcal{P}^0$. We organize the argument in three steps. In step 1, we look for conditions on the constraints C so that a solution to (KT_1)-(KT_3) exhibits a monotone and deterministic production schedule. This will imply that the schedule for a type k is a cutoff schedule with some type specific cutoff. In step 2, we identify conditions so that the resulting cutoffs are the same for all types and equal to the static cutoff. Clearly, this implies that all neglected constraints are satisfied. Finally, in step 3, we construct a set of constraints that satisfies the conditions both from step 1 and step 2.

[13] More precisely, it is sufficient for obtaining a maximum that the previous statement is true for *almost all* θ. For simplicity, we ignore issues of zero measure sets in what follows.

6.1.1 Cutoff contracts

In line with the analysis of the two type case, it seems intuitive to follow a Mirrleesian approach and to relax the original problem $\mathcal{P}^0$ by considering only the "local downward" constraints $IC^0_{i,i+1}$. To see why this does *not* work with more than two types, consider the three types case. When we only consider the local downward constraints IC^0_{12} and IC^0_{23}, then we have with respect to type $k = 2$:

$$\Psi_2(\theta, C) = p_2[v - \theta] - [p_2 + \lambda_{23}]h_2(\theta) + \lambda_{12}h_{12}(\theta). \tag{14}$$

If $\Psi_2(\theta, C)$ were decreasing in θ, then the solution x_2 to (KT_2) would automatically be monotone and deterministic. Observe that since the (cross) hazard rates are non–increasing, $\Psi_2(\theta, C)$ is decreasing provided that $p_2 + \lambda_{23} > 0$. The problem is to show that this is true. In the two types case, we were able to sign the analogous sum of the ex ante probability and the multiplier. Mimicking this argument, suppose there is a solution $\hat{\theta}$ so that $\Psi_2(\hat{\theta}) = 0$, and thus

$$p_2 + \lambda_{23} = \frac{p_2[v - \hat{\theta}]}{h_2(\hat{\theta})} + \lambda_{12}h_{12}(\hat{\theta}).$$

From here, we cannot deduce that $p_2 + \lambda_{23} > 0$ because of the presence of the negative term $\lambda_{12}h_{12}(\hat{\theta})$. For the general case, this suggests that the solution to the relaxed problem does not automatically display monotonicity, if, for some type k, the relaxed problem involves constraints IC^0_{kj} and IC^0_{ik} at the same time.

In the three types case we may however consider the relaxed problem with the constraints IC^0_{13} and IC^0_{23}. Then,

$$\begin{aligned}
\Psi_1(\theta) &= p_1[v - \theta - h_1(\theta)] - \lambda_{13}h_1(\theta), \\
\Psi_2(\theta) &= p_2[v - \theta - h_2(\theta)] - \lambda_{23}h_2(\theta), \\
\Psi_3(\theta) &= p_3[v - \theta - h_3(\theta)] + \lambda_{13}h_{13}(\theta) + \lambda_{23}h_{23}(\theta).
\end{aligned}$$

A similar argument to show monotonicity in (9) and (10) can now be used to show that for all types k, $\Psi_k(\theta)$ is decreasing in θ. This implies that the solution x_k to (KT_2) automatically displays monotonicity. In the general case, this argument extends to any relaxed problem where the set of constraints is what we call *directed*:

Definition 1 *A set $C \subseteq C_0$ is called* directed *if for all i:*

$$(i,j) \in C \text{ for some } j \quad \Rightarrow \quad (k,i) \notin C \text{ for all } k, \text{ and} \tag{15}$$

$$(j,i) \in C \text{ for some } j \quad \Rightarrow \quad (i,k) \notin C \text{ for all } k. \tag{16}$$

For a directed set $C \subseteq C_0$ of constraints, we define by

$$\omega_C \equiv \{i, j \mid (i, j) \in C\}$$

the set of ex ante types that are part of some constraint in C, and

$$\omega_C^+ = \{i \mid (i, j) \in C\}, \quad \omega_C^- = \{j \mid (i, j) \in C\}.$$

Observe that $\omega_C^+ \cap \omega_C^- = \emptyset$, because C is directed. If C is directed, Ψ_k boils down to

$$\Psi_k(\theta, C) = \begin{cases} p_k[v - \theta - h_k(\theta)] & if \quad k \notin \omega_C \\ p_k[v - \theta - h_k(\theta)] + \sum_{i:(i,k)\in C} \lambda_{ik} h_{ik}(\theta) & if \quad k \in \omega_C^- \\ p_k[v - \theta - h_k(\theta)] - \sum_{j:(k,j)\in C} \lambda_{kj} h_k(\theta) & if \quad k \in \omega_C^+. \end{cases} \tag{17}$$

The next lemma shows that for a directed set, the functions Ψ_k are strictly decreasing provided they have a root in the interval $[0, v]$.

Lemma 10 *Let C be directed and $\lambda_{ij} \leq 0$ for all $(i, j) \in C$. If there is a solution $\hat{\theta} \in [0, v]$ to $\Psi_k(\hat{\theta}, C) = 0$, then $\Psi_k(\theta, C)$ is strictly decreasing in θ.*

Next we show that the Kuhn–Tucker conditions (KT_1)-(KT_3) imply that for all k and all $(i, j) \in C$ there is indeed a solution $\hat{\theta}_k \in [0, v]$ to $\Psi_k(\hat{\theta}_k, C) = 0$ with $\lambda_{ij} \leq 0$. Thus, the previous lemma implies that the solution to the problem $\mathcal{R}^0(C)$ automatically satisfies monotonicity.

Lemma 11 *Let C be directed. Then any solution $\{x_k(\cdot)\}_{k\in\omega}$ to $\mathcal{R}^0(C)$ is characterized by a cutoff schedule $x_k(\theta) = \bar{x}(\theta, \hat{\theta}_k)$ with cutoff $\hat{\theta}_k \in [0, v]$ given by $\Psi_k(\hat{\theta}_k, C) = 0$. In particular, the solution satisfies the monotonicity constraint (MON).*

6.1.2 Static solutions

We now identify sufficient conditions on the set of constraints C so that the solution to $\mathcal{R}^0(C)$ is the optimal static contract. By Lemma 11, this amounts to identifying conditions so that the cutoffs $\hat{\theta}_k$ are all equal to the static cutoff θ^s.

Observe first that whenever a constraint $(i, j) \in C$ is *binding*, i.e., $\int_0^1 [x_i - x_j] \; dG_i = 0$, then because x_i and x_j are cutoff schedules by Lemma 11, the respective cutoffs must be the same: $\hat{\theta}_i = \hat{\theta}_j$. Similarly, if C contains the constraints (i, j) and (j, k) and both are binding, then all three cutoffs are the same: $\hat{\theta}_i = \hat{\theta}_j = \hat{\theta}_k$. This argument extends to any set of binding constraints which is connected in the following sense.

Definition 2 *Consider a subset $C \subseteq C_0$.*
(i) C is called connected *if for all $(i,j),(i',j') \in C$, C contains a sequence of pairs $(i_s, j_s)_{s=1}^S$ so that*

$$(i_1, j_1) = (i,j), \quad (i_2, j_2) = (i_2, j), \quad (i_3, j_3) = (i_2, j_3), \quad \ldots, (i_S, j_S) = (i', j').$$

(ii) C is called binding *if for any solution $\{x_k(\cdot)\}_{k\in\omega_C}, \{\lambda_{ij}\}_{(i,j)\in C}$ to (KT_1)-(KT_3), it holds $\lambda_{ij} < 0$ for all $(i,j) \in C$.*

The next lemma expresses the insight that if the set of constraints is directed, connected and binding, then for all types in the set of constraints, the solution to the relaxed problem is given by the same cutoff schedule.

Lemma 12 *Let C be directed, connected, and binding. Then for any solution $\{x_k(\cdot)\}_{k\in\omega}$ to $\mathcal{R}^0(C)$ there is a $\hat{\theta} \in [0, v]$ such that $x_k(\theta) = \bar{x}(\theta; \hat{\theta})$ for all $k \in \omega_C$.*

By Lemma 11, the cutoff $\hat{\theta}$ satisfies the equation $\Psi_k(\hat{\theta}, C) = 0$ for all types $k \in \omega_C$. Thus, solving this system of $|\omega_C|$ equations pins down the optimal cutoff $\hat{\theta}$. It turns out that $\hat{\theta}$ actually coincides with the optimal static monopsony cutoff, θ_{ω_C}, when the principal faces only types in ω_C.

Lemma 13 *Let C be directed, connected, and binding. Then the cutoff in Lemma 12 is given by $\hat{\theta} = \theta_{\omega_C}$.*

An immediate implication of the previous lemma is that if $\omega_C = \omega$ so that any type appears in some constraint, then the cutoff is equal to the optimal static cutoff θ_ω. We call such a set C with $\omega_C = \omega$ *exhausting*. This means that for a directed, connected, binding, and exhausting set of constraints C, the solution to $\mathcal{R}^0(C)$ is the static contract. Since the static contract (trivially) satisfies all original constraints, it is also a solution to the original problem $\mathcal{P}^0$:

Lemma 14 *Let C be directed, connected, binding, and exhausting. Then the solution to $\mathcal{R}^0(C)$ is the static contract. In particular, the optimal static contract solves the problem $\mathcal{P}^0$.*

6.1.3 Identifying directed, connected, binding, and exhausting constraints.

We now develop a constructive algorithm which, for any problem $\mathcal{P}^0$, yields a directed, connected, binding, and exhausting set of constraints. The construction is non–trivial, because it turns out that the relevant set of constraints depends on how the various cutoffs θ_γ, $\gamma \subseteq \omega$, introduced in (3) are ordered. To illustrate this, turn again to the three types case. In the previous subsection, we relaxed the problem by considering only IC^0_{13} and IC^0_{23}. These two constraints form a directed, connected, and exhausting set $C = \{(1,3);(2,3)\}$. To see whether the set is also binding, we check, for instance, IC^0_{23}. Ignoring this constraint yields the problem

$$\mathcal{R}^0(\{(1,3)\}): \max_{x_1,x_2,x_3} \sum_{i=1}^{3} p_i \int_0^1 [v-\theta-h_i(\theta)]x_i(\theta)\, dG_i(\theta) \ s.t. \int_0^1 [x_1(\theta)-x_3(\theta)]G_1(\theta)\, d\theta \geq 0.$$

In $\mathcal{R}(\{(1,3)\})$, the choice variable x_2 is unconstrained, and moreover, with respect to the choice variables x_1 and x_3, the problem is isomorphic to the two types problem with the types 1 and 3. Thus, the solution x_2 is characterized by the public information cutoff θ_2. Moreover, by "induction", the solution x_1 and x_3 is given by the optimal static contract in the case in which the principal faces only the two types 1 and 3. The latter is characterized by the cutoff $\theta_{\{1,3\}}$ defined by (3). Therefore, the constraint IC^0_{23} writes $\int_0^1 [x_2 - x_3]G_1 d\theta = \int_{\theta_{\{1,3\}}}^{\theta_2} G_2 d\theta \geq 0$, and, hence, is violated if and only if

$$\theta_2 < \theta_{\{1,3\}}. \tag{18}$$

Accordingly, if (18) holds, IC^0_{23} must be binding at the optimum. As a consequence, $C = \{(1,3),(2,3)\}$ is directed, connected, exhausting *and* binding so that we have found a solution by Lemma 14.

For $\theta_2 > \theta_{\{1,3\}}$, however, C is not binding and we have to look for a different set of constraints C'. It is straightforward to check that in this case the set $C' = \{(1,2),(1,3)\}$ is directed, connected, exhausting *and* binding. Therefore, whether the appropriate set of constraints is C or C' depends on the ordering of the cutoffs θ_2 and $\theta_{\{1,3\}}$. This insight is key for extending our result to an arbitrary number of types.

Finally note that for the special case $\theta_2 = \theta_{\{1,3\}}$, Lemma 2 implies $\theta_2 = \theta_{\{1,3\}} = \theta_\omega$ so that the solution to $\mathcal{R}^0(\{(1,3)\})$ itself already coincides with the static contract. Hence, it trivially satisfies both IC^0_{23} and IC^0_{13} and solves the overall problem. Consequently, we obtain our result that the static contract is optimal, even though neither C nor C' are binding by our definition. This illustrates that our result also obtains in the special non–generic cases, where cutoffs coincide, but requires a different (and easier) treatment. For expositional clarity, we will concentrate on the case where the cutoffs do not coincide and only note that our result also hold for the special cases where cutoffs do coincide. Formally, we assume

$$\theta_\gamma \neq \theta_\delta \quad \text{for all } \gamma, \delta \subseteq \omega, \gamma \neq \delta. \tag{19}$$

We now construct an explicit algorithm that yields a directed, connected, binding and exhausting subset of constraints for any configuration of cutoffs θ_γ that satisfies (19).

Our construction is inductive. We first reduce the number of types by merging the largest and the smallest types and proceed until we are left with two types. The next definition formalizes this idea. We interpret the set $2^\omega \setminus \emptyset$ of non–empty subsets of ω as an index set that encodes the types of the compressed type space. We abuse notation and identify a type $i \in \omega$ with the singleton $\{i\}$ so that $i \cup j$ denotes $\{i, j\}$ etc.

Definition 3 *Inductively, for $m = n, \ldots, 2$, define*

- *Basis ($m = n$):*
  - ◇ $\omega_n = \omega$, $\alpha_n = 1$, $\beta_n = n$.
- *Step ($m \to m-1$):*
  - ◇ $\omega_{m-1} = (\omega_m \setminus \{\alpha_m, \beta_m\}) \cup \{\alpha_m \cup \beta_m\}$,
  - ◇ $\alpha_{m-1} = \arg\max\{\theta_\gamma \mid \gamma \in \omega_{m-1}\}$ *is the set index with the lowest monopsony cutoff,*
  - ◇ $\beta_{m-1} = \arg\min\{\theta_\gamma \mid \gamma \in \omega_{m-1}\}$ *is the set index with the highest monopsony cutoff.*[14]

The algorithm results in a type space $\omega_2 = \{\alpha_2, \beta_2\}$ that has two types and by construction exhibits $\theta_{\alpha_2} < \theta_{\beta_2}$. For this case, we already know from the analysis of the two types case in Section 5 that the set $C_2 = \{(\alpha_2, \beta_2)\}$ of constraints is directed, connected, binding and exhausting. Starting with this constraint, we now expand the type space again in reverse order and essentially add to C_2 the constraint which requires that in the expanded type space ω_3, the type α_3 does not mimic type β_3. The resulting set C_3 is our candidate for a directed, connected, binding and exhausting set of constraints for ω_3. Proceeding in this fashion, we generate a set C_n of $n-1$ constraints for the original type space ω. The procedure is formally described in the next definition and subsequently illustrated for the three types case.

Definition 4 *Inductively, for $m = 2, \ldots, n-1$, define*

- *Basis ($m = 2$):*
  - ◇ $C_2 = \{(\alpha_2, \beta_2)\}$.

[14]By assumption (19), the types α_{m-1} and β_{m-1} are unique.

- *Step* ($m \to m+1$):

Define the re–labeling function $\rho_{m+1} : C_m \to \omega_{m+1}^2$ *by*

$$\rho_{m+1}((\gamma,\delta)) = \begin{cases} (\alpha_{m+1},\delta) & if \quad \gamma = \alpha_{m+1} \cup \beta_{m+1} \\ (\gamma,\beta_{m+1}) & if \quad \delta = \alpha_{m+1} \cup \beta_{m+1} \\ (\gamma,\delta) & else \end{cases}$$

and define the set

$$C_{m+1} = \rho_{m+1}(C_m) \cup \{(\alpha_{m+1},\beta_{m+1})\}.$$

To highlight how the construction of C_n depends on the primitives, consider explicitly the case $n = 3$. To generate ω_2, the procedure first merges the highest and the lowest type in ω so that we get $\omega_2 = \{\{2\},\{1,3\}\}$. Now if (18) holds, the highest type in ω_2 is $\alpha_2 = \{2\}$, and the lowest type is $\beta_2 = \{1,3\}$. Therefore, $C_2 = \{(\{2\},\{1,3\})\}$. To create C_3, the procedure re-labels first the merged type $\beta_2 = \{1,3\}$ as type $\{3\}$. This yields the set of constraints $\rho_3(C_2) = \{(\{2\},\{3\})\}$ which we identify with $\{(2,3)\}$. We subsequently add to this set the constraint $(\alpha_1,\beta_1) = (1,3)$, resulting in $C_3 = \{(2,3),(1,3)\}$. In contrast, the expansion procedure starts with $C_2' = \{(\{1,3\},\{2\})\}$ for the case $\theta_2 > \theta_{\{1,3\}}$ and finally yields $C_3' = \{(1,2),(1,3)\}$. By construction, the sets C_3 and C_3' are directed, connected, and exhausting. Moreover, as we have argued above, they are binding for the respective cases $\theta_2 < \theta_{\{1,3\}}$ and $\theta_2 > \theta_{\{1,3\}}$. We now prove that this insight extends to any configuration of cutoffs θ_γ that satisfies (19). This is the key step to establish the main result of our paper.

Lemma 15 *The set* C_n *is directed, connected, binding and exhausting.*

Together with Lemma 14, Lemma 15 implies that the static contract solves problem $\mathcal{P}^0$. In problem $\mathcal{P}^0$, we set $u_i(1)$ exogenously to zero. We now consider the original problem $\mathcal{P}$, in which $u_i(1)$ is a choice variable of the principal.

6.2 Original problem: $u_i(1)$ as a choice variable

To solve problem $\mathcal{P}$, we consider the relaxed problem where we ignore the monotonicity constraints and consider only the incentive constraints in the set C_n constructed in Definition

4:

$$\mathcal{R}: \qquad \max_{\mathbf{x},u_1(1),\ldots,u_n(1)} \sum_{i\in\omega} p_i \int_0^1 [v-\theta-h_i(\theta)]x_i(\theta)\;dG_i(\theta) - p_i u_i(1) \qquad s.t.$$

$$\int_0^1 [x_i(\theta)-x_j(\theta)]G_i(\theta)\;d\theta + u_i(1) - u_j(1) \geq 0 \quad \text{for all } (i,j)\in C_n,$$

$$u_i(1)\geq 0 \quad \text{for all } i\in\omega.$$

We now prove that the static contract solves problem $\mathcal{R}$. This establishes the main result of the paper that the static contract solve problem $\mathcal{P}$.

Theorem 1 *The static contract is a solution to problem $\mathcal{R}$. Because the static contract satisfies all neglected constraints, it is also a solution to the original problem $\mathcal{P}$.*

7 Extensions

7.1 Different ex ante and ex post outside options

In the analysis so far, we assumed that the agent's ex ante outside option coincides with his ex post outside option. This is the natural assumption when the ex post outside option is type–independent and does not change over time. Yet from a practical perspective, it is important to know to what extent our results are robust and extend to differences in ex ante and ex post outside options. For example, in a procurement relationship, where the agent as a firm is typically protected by limited liability, the agent can incur some losses without going bankrupt when he has pledgable assets or other sources of income. In this case, the agent's ex post outside option is lower than his ex ante outside option.

To allow for different outside options, we normalize the ex ante outside option to zero and set the ex post outside option to $\bar{u}\leq 0$. Thus, the ex post individual rationality (6) changes to

$$u_i(\theta)\geq \bar{u} \quad \text{for all } i\in\omega, \theta\in[0,1], \tag{20}$$

while the ex ante individual rationality constraint (7) remains the same.

We argue that our result that the static contract is optimal still holds as long as $\bar{u}$ is not too negative. To see this, note first that if we solve for the optimal contract with the adapted ex post individual rationality constraint (20) while disregarding the ex ante individual rationality constraint (7), the only change is that the principal can extract more utility from the agent.

In particular, the optimal production schedule is equal to the cutoff schedule with the optimal static cutoff θ^s, and since the ex post participation constraint (20) is binding at the optimum, we have that $u_i(1) = \bar{u}$. It follows that the expected utility of ex ante type i is

$$U_i = \bar{u} + \int_0^{\theta^s} G_i d\theta.$$

Hence, if we define

$$\bar{u}^p \equiv -\min_i \int_0^{\theta^s} G_i d\theta,$$

then for $\bar{u} \geq \bar{u}^p$, the solution satisfies automatically the ex ante participation constraint (7). Thus, for all ex post outside options $\bar{u} \in [\bar{u}^p, 0]$, the optimal contract is the static one. Because $\bar{u}^p < 0$, our result that sequential screening is not helpful with ex post participation constraints is robust and extends to differences in ex ante and ex post outside options.

Taking the opposite approach and solving the model with the ex ante individual rationality constraint (7) while disregarding the ex post individual rationality constraint (20) yields the solution of Lemma 8. In particular, the individual rationality constraint of the highest type and all local downward incentive constraints are binding. This means that the ex post type $\theta = 1$ of ex ante type 1 obtains the lowest ex post utility of all ex post agent types and, in particular,

$$u_1(1) = \bar{u}^a \equiv -\sum_{i=1}^{n} \int_{\theta_{i+1}^{CL}}^{\theta_i^{CL}} G_i(\theta) d\theta,$$

where $\theta_{n+1}^{CL} \equiv 0$. Hence, the solution satisfies the neglected ex ante individual rationality constraint (20) whenever $\bar{u} < \bar{u}^a$.

It follows that as we vary the ex post outside option $\bar{u}$, we obtain the sequential screening models with ex ante and ex post participation constraints as two extremes: the model with ex ante constraints for $\bar{u} \leq \bar{u}^a$ and the model with ex post constraints for $\bar{u} \geq \bar{u}^p$. A fully fledged analysis of the intermediate case $\bar{u} \in (\bar{u}^a, \bar{u}^p)$ lies, due to intractability issues, outside the scope of this paper. We only mention that simulation exercises for simple models with two ex ante types show that it depends on the exact magnitude of u^p which of the incentive, ex ante, and ex post individual rationality constraints are binding, and that in the interval $(\bar{u}^a, \bar{u}^p)$ different types of static and sequential contracts can be optimal.

7.2 Auctions

Our techniques and results extend readily to settings with multiple agents. Consequently, the optimal mechanism with ex post participation constraints is equivalent to the static Myerson (1981) auction that is optimal for the principal when he faces the agents after they received all their private information. Again, this stands in stark contrast to sequential screening models with ex ante participation constraints only. In particular, Esö and Szentes (2007b) show that with multiple agents sequential screening allows the principal to extract all the additional information embodied in the ex post private information by means of an augmented second price auction. With ex post outside options, the optimal contract is simpler and the principal cannot extract the agent's information rents.

8 Conclusion

This paper shows that introducing ex post participation constraints in a sequential screening problem eliminates the value of eliciting the agent's information sequentially. Instead, a static contract, which conditions only on the agent's final information is optimal. In this sense, the value of dynamic over static contracting in the absence of ex post participation constraints is due to relaxed participation rather than relaxed incentive constraints.

In this paper, we have taken the agent's outside option as exogenous. In practice, the outside option is often endogenously determined, for example, by the presence of a spot market where the agent can trade in any period. Our paper raises the question if dynamic, long–term contracting has some value when spot markets offer the agent an outside option at any point in time.

Appendix

Proof of Lemma of 1 Since h_{ij} is non–decreasing by assumption, we have $h'_{ij} = (g_j^2)^{-1} \cdot (-g_i g_j - G_i g'_j) \geq 0$ for all $i, j \in \omega$. Hence, for $\gamma, \delta \subset \omega$:

$$h'_{\gamma,\delta} = (g_\delta^2)^{-1}[-g_\gamma g_\delta - G_\gamma g'_\delta] = \frac{1}{p_\gamma p_\delta}(g_\delta^2)^{-1} \cdot \sum_{i\in\gamma, j\in\delta} p_i p_j(-g_i g_j - G_i g'_j) \geq 0, \qquad (21)$$

and this proves the claim. Q.E.D.

Proof of Lemma 2 By definition, $\theta_{\gamma\cup\delta}$ satisfies the equation

$$(v - \theta_{\gamma\cup\delta}) \sum_{i\in\gamma\cup\delta} p_i g_i(\theta_{\gamma\cup\delta}) = \sum_{i\in\gamma\cup\delta} p_i G_i(\theta_{\gamma\cup\delta}). \qquad (22)$$

Now suppose that contrary to the claim, we have $\theta_{\gamma\cup\delta} > \max\{\theta_\gamma, \theta_\delta\}$. (Similar arguments apply to the case $\theta_{\gamma\cup\delta} < \min\{\theta_\gamma, \theta_\delta\}$.) Then, by monotonicity of the hazard rate and the definition of $\theta_\gamma, \theta_\delta$:

$$v - \theta_{\gamma\cup\delta} < \frac{G_\gamma(\theta_{\gamma\cup\delta})}{g_\gamma(\theta_{\gamma\cup\delta})} \text{ and } v - \theta_{\gamma\cup\delta} < \frac{G_\delta(\theta_{\gamma\cup\delta})}{g_\delta(\theta_{\gamma\cup\delta})} \qquad (23)$$

$$\Leftrightarrow \quad (v - \theta_{\gamma\cup\delta}) \cdot g_\gamma(\theta_{\gamma\cup\delta}) < G_\gamma(\theta_{\gamma\cup\delta}) \text{ and } (v - \theta_{\gamma\cup\delta}) \cdot g_\delta(\theta_{\gamma\cup\delta}) < G_\delta(\theta_{\gamma\cup\delta}). \qquad (24)$$

Now multiply the left inequality with p_γ and the right with p_δ and add them up to get

$$(v - \theta_{\gamma\cup\delta}) \left[\sum_{i\in\gamma} p_i g_i(\theta_{\gamma\cup\delta}) + \sum_{i\in\delta} p_i g_i(\theta_{\gamma\cup\delta})\right] < \sum_{i\in\gamma} p_i G_i(\theta_{\gamma\cup\delta}) + \sum_{i\in\delta} p_i G_i(\theta_{\gamma\cup\delta}). \qquad (25)$$

Since γ and δ are disjoint, this can be written as

$$(v - \theta_{\gamma\cup\delta}) \sum_{i\in\gamma\cup\delta} p_i g_i(\theta_{\gamma\cup\delta}) < \sum_{i\in\gamma\cup\delta} p_i G_i(\theta_{\gamma\cup\delta}), \qquad (26)$$

a contradiction to (22). Q.E.D.

Derivation of (8) **and Proof of Lemma 4** By (RE),

$$\int_0^1 u_j(\theta)\, dG_i(\theta) = \int_0^1 \int_\theta^1 x_j(z)\, dz\, g_i(\theta)\, d\theta + u_j(1) \qquad (27)$$

$$= \int_\theta^1 x_j(z)\, dz \cdot G_i(\theta)\Big|_0^1 - \int_0^1 -x_j(\theta) G_i(\theta)\, d\theta + u_j(1) \qquad (28)$$

$$= \int_0^1 x_j(\theta) G_i(\theta)\, d\theta + u_j(1), \qquad (29)$$

where we have used integration by parts in the second line.

Thus, for $j = i$, it follows that

$$\int_0^1 u_i(\theta)\, dG_i(\theta) = \int_0^1 x_i(\theta) h_i(\theta)\, dG_i(\theta) + u_i(1). \tag{30}$$

Plugging this in the principal's objective delivers (8).

Moreover, since $U_{ij} = \int_0^1 u_j(\theta)\, dG_i(\theta)$, (29) implies that the first period incentive compatibility condition $U_i - U_{ij} \geq 0$ is equivalent to (IC_{ij}), and this is what we wanted to show. Q.E.D.

Proof of Lemma 10 Because the hazard rate $h_k(\theta)$ is non–decreasing and $p_k[v - \theta]$ is strictly decreasing, it follows that $p_k[v - \theta - h_k(\theta)]$ is strictly decreasing. This establishes that $\Psi_k(\theta, C)$ is strictly decreasing in θ for $k \notin \omega_C$. In addition, $\lambda_{kj} \leq 0$ and non–decreasing cross hazard rates $h_{kj}(\theta)$ imply that $\sum_{i:(i,k)\in C} \lambda_{ik} h_{ik}(\theta)$ is non–increasing in θ. Hence, $\Psi_k(\theta, C)$ is strictly decreasing in θ also for $k \in \omega_C^-$. Finally, to see that $\Psi_k(\theta, C)$ is strictly decreasing in θ also for $k \in \omega_C^+$, first rewrite $\Psi_k(\theta, C)$ for $k \in \omega_C^+$ as

$$\Psi_k(\theta, C) = p_k[v - \theta] - \left(p_k + \sum_{j:(k,j)\in C} \lambda_{kj} \right) h_k(\theta). \tag{31}$$

By assumption, $\Psi_k(\hat{\theta}, C) = 0$ for some $\hat{\theta} \in [0, v]$. For $k \in \omega_C^+$, this implies that

$$p_k + \sum_{j:(k,j)\in C} \lambda_{kj} = \frac{p_k[v - \hat{\theta}]}{h_k(\hat{\theta})} \geq 0. \tag{32}$$

The non–decreasing hazard rate $h_k(\cdot)$ therefore implies that $(p_k + \sum_{j:(k,j)\in C} \lambda_{kj}) h_k(\theta)$ is non–decreasing. Due to the term $p_k[v - \theta]$, it then follows that (31) is strictly decreasing in θ. Q.E.D.

Proof of Lemma 11 By Lemma 9 any solution $\{x_k(\cdot)\}_{k\in\omega}$ to $\mathcal{R}^0(C)$ satisfies (KT_1)–KT_3. We distinguish the three possible cases:

Case $k \notin \omega_C$: In this case, $\Psi(\theta_k, C) = 0$ by definition of θ_k. Since $\theta_k < v$, Lemma 10 implies that $\Psi_k(\theta, C) > 0$ for all $\theta \in [0, \theta_k)$ and $\Psi_k(\theta, C) < 0$ for all $\theta \in (\theta_k, 1]$. By (KT_2), any solution therefore exhibits $x_k(\theta) = \mathbf{1}_{[0,\theta_k]}(\theta)$.[15]

Case $k \in \omega_C^-$: In this case, $\Psi_k(0, C) = p_k v > 0$ and, by (KT_1), $\Psi_k(\theta_k, C) = \sum_{i:(i,k)\in C} \lambda_{ik} h_{ik}(\theta_k) \leq 0$. Continuity of $\Psi_k(\theta, C)$ in θ then implies there exists a $\hat{\theta}_k \in (0, \theta_k]$ such that $\Psi_k(\hat{\theta}_k, C) = 0$. Because $\hat{\theta}_k \leq \theta_k < v$, Lemma 10 applies so that $\Psi_k(\theta, C)$ is strictly decreasing in θ. Hence,

[15] Let $\mathbf{1}_A(a)$ express the indicator function: It takes value 1 if $a \in A$ and 0 otherwise.

$\Psi_k(\theta, C) > 0$ for all $\theta \in [0, \hat{\theta}_k)$ and $\Psi_k(\theta, C) < 0$ for all $\theta \in (\hat{\theta}_k, 1]$. By (KT_2), any solution therefore exhibits $x_k(\theta) = \mathbf{1}_{[0,\theta_k]}(\theta)$.

Case $k \in \omega_C^+$: We prove by contradiction that there exists a $\hat{\theta}_k \in [0, v]$ such that $\Psi_k(\hat{\theta}_k, C) = 0$. For suppose the contrary, then, by $\Psi_k(0, C) = p_k v > 0$ and continuity of $\Psi_k(\theta, C)$ in θ, it must hold $\Psi_k(\theta, C) > 0$ for all $\theta \in [0, v]$. First, this implies, by (KT_2), that for any solution it must hold $x_k(\theta) = 1$ for any $\theta \in [0, v]$. Second, it implies that $\Psi_k(v, C) = -[p_k + \sum_{j:(k,j)\in C} \lambda_{kj}]h_k(v) > 0$ so that necessarily $\sum_{j:(k,j)\in C} \lambda_{kj} < -p_k$. Hence, there must be at least one $j \in \omega_C^-$ such that $\lambda_{kj} < 0$, implying by (KT_3) that IC_{kj} is satisfied in equality. But because $j \in \omega_C^-$, we just established that for any solution there exists a $\hat{\theta}_j < v$ such that $x_j(\theta) = \mathbf{1}_{[0,\hat{\theta}_j]}(\theta)$. It therefore follows

$$\int_0^1 [x_k(\theta) - x_j(\theta)]G_k(\theta)d\theta \geq \int_{\hat{\theta}_j}^v G_k(\theta)d\theta > 0.$$

Using (KT_3), this leads to the contradiction that $\lambda_{kj} = 0$. Consequently, there must exist a $\hat{\theta}_k \in [0, v]$ such that $\Psi_k(\hat{\theta}_k, C) = 0$. By Lemma 10, $\Psi_k(\theta, C) > 0$ for all $\theta \in [0, \hat{\theta}_k)$ and $\Psi_k(\theta, C) < 0$ for all $\theta \in (\hat{\theta}_k, 1]$. By (KT_2), any solution therefore exhibits $x_k(\theta) = \mathbf{1}_{[0,\hat{\theta}_k]}(\theta)$.

We conclude that any solution is characterized by a cutoff schedule $x_k(\theta) = \bar{x}(\theta, \hat{\theta}_k)$, where the cutoff $\hat{\theta}_k \in [0, v]$ solves $\Psi_k(\hat{\theta}_k, C) = 0$. This solution trivially satisfies the monotonicity constraint (MON). Q.E.D.

Proof of Lemma 12 Because C is binding, for any $(i, j) \in C$ it holds $\lambda_{ij} < 0$ such that (KT_3) implies

$$\int_0^1 [x_i(\theta) - x_j(\theta)]G_i(\theta))d\theta = 0. \tag{33}$$

Because by Lemma 11 $x_i(\cdot)$ and $x_j(\cdot)$ are increasing cutoff schedules with respective cutoffs $\hat{\theta}_i$ and $\hat{\theta}_j$, (33) can only hold if $\hat{\theta}_i = \hat{\theta}_j$. Connectedness then implies that for any $i, j \in \omega_C$, the cutoffs are the same. Q.E.D.

Proof of Lemma 13 From Lemma 11 and Lemma 12 it follows that $\hat{\theta}$ satisfies $\Psi_k(\hat{\theta}, C) = 0$ for all $k \in \omega_C$. Because C is directed, we have for all $k \in \omega_C$ that either $k \in \omega_C^+$ or (exclusively)

$k \in \omega_C^-$. Multiplying $\Psi_k(\hat{\theta}, C)$ by $g_k(\hat{\theta})$ and adding up delivers

$$
\begin{aligned}
0 &= \sum_{k\in\omega_C} \Psi_k(\hat{\theta}, C) g_k(\hat{\theta}) && (34)\\
&= \sum_{k\in\omega_C} p_k[v-\hat{\theta}]g_k(\hat{\theta}) - p_k G_k(\hat{\theta}) && (35)\\
&\quad - \sum_{k\in\omega_C^+} \sum_{j:(k,j)\in C} \lambda_{kj} h_k(\hat{\theta}) g_k(\hat{\theta}) + \sum_{k\in\omega_C^-} \sum_{i:(i,k)\in C} \lambda_{ik} h_{i,k}(\hat{\theta}) g_k(\hat{\theta}) && (36)\\
&= \sum_{k\in\omega_C} p_k[v-\hat{\theta}]g_k(\hat{\theta}) - p_k G_k(\hat{\theta}) && (37)\\
&\quad - \sum_{k\in\omega_C^+} \sum_{j:(k,j)\in C} \lambda_{kj} G_k(\hat{\theta}) + \sum_{j\in\omega_C^-} \sum_{k:(k,j)\in C} \lambda_{kj} G_k(\hat{\theta}). && (38)
\end{aligned}
$$

The last inequality follows by re-labeling the summation index of the second double sum and by definition of h_k and $h_{i,k}$. Now observe that in the last line, every pair (k, j) that appears under the first double sum also appears under the second double sum (and vice versa). Therefore, the last line is zero, and we obtain

$$
0 = \sum_{k\in\omega_C} p_k[v-\hat{\theta}]g_k(\hat{\theta}) - p_k G_k(\hat{\theta}) \;\Leftrightarrow\; v = \hat{\theta} + \frac{\sum_{k\in\omega_C} p_k G_k(\hat{\theta})}{\sum_{k\in\omega_C} p_k g_k(\hat{\theta})}, \tag{39}
$$

which by (3) implies $\hat{\theta} = \theta_{\omega_C}$. Q.E.D.

Proof of Lemma 14 The claim is a direct implication of Lemmata 11, 12, and 13. Q.E.D.

Proof of Lemma 15 The set C_n is directed, connected, and exhausting by construction. To show that it is binding, we have to show that any solution to (KT_1)-(KT_3) satisfies

$$
\lambda_{\gamma,\delta} < 0 \quad \forall (\gamma, \delta) \in C_n. \tag{40}
$$

We only consider the case in which $\theta_{\{1,n\}} < \theta_{\omega\setminus\{1,n\}}$. (The argument for the reverse case is analogous.)

Step 1: We begin by showing that

$$
\lambda_{1n} < 0. \tag{41}
$$

Contrary to (41), suppose there is a solution $\{x_\eta(\cdot)\}_{\eta\in\omega}$ to (KT_1)-(KT_3) so that $\lambda_{1n} = 0$. By construction of C_n, the pair $(1, n)$ is the only pair in C_n that involves the index n. Together with $\lambda_{1n} = 0$, this implies that

$$
\Psi_n(\theta, C_n) = p_n[v - \theta - h_n(\theta)]. \tag{42}
$$

Hence, (KT_2) implies that $x_n = 1_{[0,\theta_n]}$. Next, we determine x_1. We distinguish two cases.

(1) If $\lambda_{1\delta} = 0$ for all δ with $(1, \delta) \in C_n$, the same reasoning as in the previous paragraph delivers that $x_1 = 1_{[0,\theta_1]}$. Since $\theta_1 < \theta_n$ by assumption, $\{x_\eta(\cdot)\}_{\eta \in \omega}$ violates the constraint $(1, n) \in C_n$, a contradiction to the assumption that $\{x_\eta(\cdot)\}_{\eta \in \omega}$ is a solution to $\mathcal{R}^0(C_n)$.

(2) Suppose that $\lambda_{1\delta^*} < 0$ for some $(1, \delta^*) \in C_n$. We shall derive a contradiction in a similar fashion as in the previous paragraph by first determining x_1 and then comparing it to x_n. Let

$$A = \{(\gamma, \delta) \in C_n \mid \lambda_{\gamma\delta} < 0\} \subset C_n \tag{43}$$

be the set of binding constraints in C_n. It is non–empty, as it contains $(1, \delta^*)$. Let B be the largest connected set in A which contains $(1, \delta^*)$. By definition of B, we have

$$\eta \in \omega_B \quad \Leftrightarrow \quad (\eta, \delta) \in B \text{ for all } (\eta, \delta) \in C_n \text{ with } \lambda_{\eta\delta} < 0, or \tag{44}$$

$$(\gamma, \eta) \in B \text{ for all } (\gamma, \eta) \in C_n \text{ with } \lambda_{\gamma\eta} < 0. \tag{45}$$

This implies that for all $\eta \in \omega_B$ we can write

$$\begin{aligned} \Psi_\eta(\theta, C_n) &= p_\eta[v - \theta - h_\eta(\theta)] - \sum_{\delta:(\eta,\delta)\in C_n} \lambda_{\eta\delta} h_\eta(\theta) + \sum_{\gamma:(\gamma,\eta)\in C_n} \lambda_{\gamma\eta} h_{\gamma,\eta}(\theta) && (46) \\ &= p_\eta[v - \theta - h_\eta(\theta)] - \sum_{\delta:(\eta,\delta)\in B} \lambda_{\eta\delta} h_\eta(\theta) + \sum_{\gamma:(\gamma,\eta)\in B} \lambda_{\gamma\eta} h_{\gamma,\eta}(\theta) && (47) \\ &= \Psi_\eta(\theta, B). && (48) \end{aligned}$$

Therefore, (KT_2) for $C = C_n$ implies that

$$x_\eta(\theta) = \begin{cases} 0 & if \quad \Psi_\eta(\theta, B) < 0 \\ 1 & if \quad \Psi_\eta(\theta, B) > 0 \end{cases} \qquad \forall \eta \in \omega_B. \tag{49}$$

Moreover, (KT_1) and (KT_3) for $C = C_n$ (trivially) imply

$$\lambda_{\gamma\delta} \leq 0, \quad and \quad \lambda_{\gamma\delta} \int_0^1 [x_\gamma(\theta) - x_\delta(\theta)] G_\gamma(\theta) d\theta = 0 \qquad \forall (\gamma, \delta) \in B. \tag{50}$$

The conditions (49) and (50) are the first order conditions for the problem $\mathcal{R}^0(B)$ (when the type space is ω_B). Therefore, the solution $x_\eta(\cdot)$ for $\eta \in \omega_B$ to $\mathcal{R}^0(C_n)$ is given by the solution $x_\eta(\cdot)$ to $\mathcal{R}^0(B)$. Because B as a subset of the directed set C_n is itself directed and, because it is connected and binding by definition, Lemma 12 and 13 imply that for all $\eta \in \omega_B$ and in particular for $\eta = 1 \in \omega_B$, we have:

$$x_\eta = x_1 = 1_{[0,\theta_{\omega_B}]}. \tag{51}$$

Because $n \notin \omega_B$, it follows, by assumption, $\theta_\eta < \theta_n$ for all $\eta \in \omega_B$. By Lemma 2 it then follows $\theta_{\omega_B} < \theta_n$, and thus $x_1(\cdot) < x_n(\cdot)$. But this violates the constraint $(1,n) \in C_n$, a contradiction to the assumption that $\{x_\eta(\cdot)\}_{\eta\in\omega}$ is a solution to $\mathcal{R}^0(C_n)$.

Step 2: We now show the rest of (40) by induction over the number of types n. More precisely the induction hypothesis is that (40) is true for any model with $n-1$ types.

Base ($n = 2$): By assumption $\theta_1 < \theta_2$ so that $C_2 = \{(1,2)\}$. Thus, Step 1 implies that $\lambda_{12} < 0$, as desired.

Step ($n-1 \to n$): Let $(\mathbf{x}, \lambda) = (\{x_k(\cdot)\}_{k\in\omega}, \{\lambda_{ij}\}_{(i,j)\in C_n})$ be a solution to (KT_1)-(KT_3) (for $C = C_n$). Contrary to the claim, suppose that

$$\lambda_{ij} = 0 \text{ for some } (i,j) \in C_n \text{ with } (i,j) \neq (1,n). \tag{52}$$

We will derive a contradiction to the induction hypothesis by constructing a model with $n-1$ types and a solution to $(KT_1)^{n-1}$-$(KT_3)^{n-1}$ (for $C = C_{n-1}$) which violates $(40)^{n-1}$.[16]

By Lemma 11, $(\mathbf{x}, \lambda)$ displays for each k a cutoff $\hat{\theta}_k \in [0, v]$ so that $x_k = 1_{[0,\hat{\theta}_k]}$, and

$$\Psi_k(\hat{\theta}_k, C_n) = 0. \tag{53}$$

Moreover, the argument in the proof of Lemma 12 implies that for all $(i,j) \in C_n$:

$$\lambda_{ij} < 0 \quad \Rightarrow \quad \hat{\theta}_i = \hat{\theta}_j. \tag{54}$$

Consider now the model with $n-1$ types when types 1 and n are merged so that the new type space is $\omega_{n-1} = \{\{1,n\}, 2, \ldots, n-1\}$. We indicate the variables pertaining to this model by a superindex $n-1$. From the solution $(\mathbf{x}, \lambda)$ to $(KT_1)^n$-$(KT_3)^n$ (for $C = C_n$), we propose the following candidate solution $(\mathbf{x}^{n-1}, \lambda^{n-1})$ to $(KT_1)^{n-1}$-$(KT_3)^{n-1}$ (for $C = C_{n-1}$):

$$x_k^{n-1} = x_k = 1_{[0,\hat{\theta}_k]} \quad \text{for all } k = 2, \ldots, n-1, \tag{55}$$

$$x_{\{1,n\}}^{n-1} = x_1 = 1_{[0,\hat{\theta}_1]}, \tag{56}$$

$$\lambda_{\gamma\delta}^{n-1} = \lambda_{\gamma\delta} \quad \text{for all } (\gamma,\delta) \in C_{n-1} \text{ with } \gamma \neq \{1,n\}, \tag{57}$$

$$\lambda_{\{1,n\}\delta}^{n-1} = \lambda_{1\delta} \cdot \frac{G_1(\hat{\theta}_1)}{G_{\{1,n\}}(\hat{\theta}_1)} \quad \text{for all } (\{1,n\},\delta) \in C_{n-1}. \tag{58}$$

[16] In what follows, the exponent on the equation reference refers to the number of types of the model under consideration.

(Observe that because $\theta_{\{1,n\}} > \theta_{\omega \setminus \{1,n\}}$ by assumption, we have by construction of the set C_{n-1} that $(1, \delta) \in C_n$ if and only if $(\{1, n\}, \delta) \in C_{n-1}$.)

By (52), this definition implies that there is a $(\gamma, \delta) \in C_{n-1}$ so that $\lambda^{n-1}_{\gamma\delta} = 0$. Thus, it is sufficient to show that

$$(\mathbf{x}^{n-1}, \lambda^{n-1}) \text{ satisfies } (KT_1)^{n-1}\text{-}(KT_3)^{n-1} \text{ (for } C = C_{n-1}) \ , \tag{59}$$

because then we would have found a model with $n-1$ types that violates $(40)^{n-1}$, a contradiction to the induction hypothesis.

We now show (59).

- That $(\mathbf{x}^{n-1}, \lambda^{n-1})$ satisfies $(KT_1)^{n-1}$ is trivial.
- To see that $(\mathbf{x}^{n-1}, \lambda^{n-1})$ satisfies $(KT_2)^{n-1}$, observe that since x^{n-1}_k is a cutoff schedule for all $k \in \omega_{n-1}$, we only have to show that $\Psi^{n-1}_k(\cdot, C_{n-1})$ equals zero at the respective cutoff of x^{n-1}_k. By the definition (55) and (56) of x^{n-1}_k, this means we have to show:

$$\Psi^{n-1}_k(\hat{\theta}_k, C_{n-1}) = 0 \text{ for all } k = 2, \ldots, n-1, \quad \textit{and} \quad \Psi^{n-1}_{\{1,n\}}(\hat{\theta}_1, C_{n-1}) = 0. \tag{60}$$

Recall that $\theta_{\{1,n\}} > \theta_{\omega \setminus \{1,n\}}$ by assumption. Thus, the set C_{n-1} is constructed by removing $(1, n)$ from C_n and then re-labeling the index 1 as $\{1, n\}$. Moreover, $(1, n)$ is the only pair in C_n that involves the index n. Thus,

$$\text{For all } k \in \omega^+_{n-1} \setminus \{1, n\} : (k, \delta) \in C_{n-1} \Leftrightarrow (k, \delta) \in C_n, \tag{61}$$

$$\text{For all } k \in \omega^-_{n-1} : (\gamma, k) \in C_{n-1} \Leftrightarrow (\gamma, k) \in C_n, \text{ or } (\gamma = \{1, n\} \text{ and } (1, k) \in C_n). \tag{62}$$

We now establish the left part of (60). For $k \in \omega^+_{n-1} \setminus \{1, n\}$, (61) together with (53) and (57) implies that

$$\Psi^{n-1}_k(\hat{\theta}_k, C_{n-1}) = p_k[v - \hat{\theta}_k - h_k(\hat{\theta}_k)] - \sum_{\delta:(k,\delta) \in C_{n-1}} \lambda^{n-1}_{k\delta} h_k(\hat{\theta}_k) \tag{63}$$

$$= p_k[v - \hat{\theta}_k - h_k(\hat{\theta}_k)] - \sum_{\delta:(k,\delta) \in C_n} \lambda_{k\delta} h_k(\hat{\theta}_k) = \Psi_k(\hat{\theta}_k, C_n) = 0. \tag{64}$$

Moreover, for $k \in \omega_{n-1}^-$, (62) together with (57) and (58) implies that

$$\begin{aligned}
\Psi_k^{n-1}(\hat{\theta}_k, C_{n-1}) &= p_k[v - \hat{\theta}_k - h_k(\hat{\theta}_k)] + \sum_{\gamma:(\gamma,k)\in C_{n-1}} \lambda_{\gamma k}^{n-1} h_{\gamma,k}(\hat{\theta}_k) && (65)\\
&= p_k[v - \hat{\theta}_k - h_k(\hat{\theta}_k)] + \sum_{\substack{\gamma:(\gamma,k)\in C_{n-1}\\ \gamma\neq\{1,n\}}} \lambda_{\gamma k}^{n-1} h_{\gamma,k}(\hat{\theta}_k) + \lambda_{\{1,n\}k}^{n-1} h_{\{1,n\},k}(\hat{\theta}_k) && (66)\\
&= p_k[v - \hat{\theta}_k - h_k(\hat{\theta}_k)] + \sum_{\substack{\gamma:(\gamma,k)\in C_{n}\\ \gamma\neq 1}} \lambda_{\gamma k} h_{\gamma,k}(\hat{\theta}_k) && (67)\\
&\quad + \lambda_{1k} \cdot \frac{G_1(\hat{\theta}_1)}{G_{\{1,n\}}(\hat{\theta}_1)} h_{\{1,n\},k}(\hat{\theta}_k),
\end{aligned}$$

where we make use of the convention: $\lambda_{\{1,n\}k}^{n-1} = 0$ if $(\{1,n\},k) \notin C_{n-1}$ and $\lambda_{1k} = 0$ if $(1,k) \notin C_{n-1}$. We now distinguish two cases. If $\lambda_{1k} = 0$, then the last term in the previous expression vanishes, and we can as well write

$$\Psi_k^{n-1}(\hat{\theta}_k, C_{n-1}) = p_k[v - \hat{\theta}_k - h_k(\hat{\theta}_k)] + \sum_{\substack{\gamma:(\gamma,k)\in C_{n}\\ \gamma\neq 1}} \lambda_{\gamma k} h_{\gamma,k}(\hat{\theta}_k) + \lambda_{1k} h_{1,k}(\hat{\theta}_k) = \Psi_k(\hat{\theta}_k, C_n), \quad (68)$$

which is zero by (53), as desired. If $\lambda_{1k} < 0$, then $\hat{\theta}_1 = \hat{\theta}_k$ by (54), so that

$$\frac{G_1(\hat{\theta}_1)}{G_{\{1,n\}}(\hat{\theta}_1)} h_{\{1,n\},k}(\hat{\theta}_k) = \frac{G_1(\hat{\theta}_1)}{G_{\{1,n\}}(\hat{\theta}_1)} \cdot \frac{G_{\{1,n\}}(\hat{\theta}_1)}{g_k(\hat{\theta}_1)} = h_{1,k}(\hat{\theta}_k). \quad (69)$$

Consequently, we also have in this case that $\Psi_k^{n-1}(\hat{\theta}_k, C_{n-1}) = \Psi_k(\hat{\theta}_k, C_n)$, which is zero by (53). This completes the proof of the left part of (60).

We are left to show the right part of (60), i.e. $\Psi_{\{1,n\}}^{n-1}(\hat{\theta}_1, C_{n-1}) = 0$. By assumption, $\theta_{\{1,n\}} > \theta_{\omega\setminus\{1,n\}}$. It is easy to see that this implies $\{1,n\} \in \omega_{n-1}^+$. Hence, TO DO

$$\begin{aligned}
\Psi_{\{1,n\}}^{n-1}(\hat{\theta}_1, C_{n-1}) &= p_{\{1,n\}}[v - \hat{\theta}_1 - h_{\{1,n\}}(\hat{\theta}_1)] - \sum_{\delta:(\{1,n\},\delta)\in C_{n-1}} \lambda_{\{1,n\}\delta}^{n-1} h_{\{1,n\}}(\hat{\theta}_1) && (70)\\
&= \frac{[v - \hat{\theta}_1] p_{\{1,n\}} g_{\{1,n\}}(\hat{\theta}_1) - p_{\{1,n\}} G_{\{1,n\}}(\hat{\theta}_1)}{g_{\{1,n\}}(\hat{\theta}_1)} && (71)\\
&\quad - \sum_{\delta:(\{1,n\},\delta)\in C_{n-1}} \lambda_{\{1,n\}\delta}^{n-1} h_{\{1,n\}}(\hat{\theta}_1),
\end{aligned}$$

where in the second line we have took $1/g_{\{1,n\}}(\hat{\theta}_1)$ out of the square brackets. By (58) and the definition of $G_{\{1,n\}}$ and $g_{\{1,n\}}$, and since $(\{1,n\},\delta) \in C_{n-1}$ if and only if $(1,\delta) \in C_n$ and $\delta \neq n$, this can be re-written as

$$\Psi^{n-1}_{\{1,n\}}(\hat\theta_1, C_{n-1}) = \frac{[v-\hat\theta_1]\cdot\{p_1 g_1(\hat\theta_1)+p_n g_n(\hat\theta_1)\} - p_1 G_1(\hat\theta_1) - p_n G_n(\hat\theta_1)}{g_{\{1,n\}}(\hat\theta_1)} - \sum_{\substack{\delta:(1,\delta)\in C_n \\ \delta\neq n}} \lambda_{1,\delta}\frac{G_1(\hat\theta_1)}{g_{\{1,n\}}(\hat\theta_1)} \tag{72}$$

Now add and subtract $\lambda_{1n}G_1(\hat\theta_1)/g_1(\hat\theta_1)$ to obtain

$$\begin{aligned}\Psi^{n-1}_{\{1,n\}}(\hat\theta_1, C_{n-1}) &= \frac{1}{g_{\{1,n\}}(\hat\theta_1)}\Bigg\{p_1[(v-\hat\theta_1)g_1(\hat\theta_1) - G_1(\hat\theta_1)] - \lambda_{1n}G_1(\hat\theta_1) \\ &\quad - \sum_{\substack{\delta:(1,\delta)\in C_n \\ \delta\neq n}} \lambda_{1,\delta}G_1(\hat\theta_1)\Bigg\} \\ &\quad + \frac{1}{g_{\{1,n\}}(\hat\theta_1)}\Big\{p_n[(v-\hat\theta_1)g_n(\hat\theta_1) - G_n(\hat\theta_1)] + \lambda_{1n}G_1(\hat\theta_1)\Big\}.\end{aligned} \tag{73}$$

The first two lines on the right hand side can be written as

$$\frac{g_1(\hat\theta_1)}{g_{\{1,n\}}(\hat\theta_1)}\cdot\frac{p_1[(v-\hat\theta_1)g_1(\hat\theta_1) - G_1(\hat\theta_1)] - \sum_{\delta:(1,\delta)\in C_n}\lambda_{1\delta}G_1(\hat\theta_1)}{g_1(\hat\theta_1)} = \frac{g_1(\hat\theta_1)}{g_{\{1,n\}}(\hat\theta_1)}\cdot\Psi_1(\hat\theta_1, C_n). \tag{74}$$

Moreover, since $(1,n)$ is the only pair in C_n that involves the index n, the third line becomes

$$\frac{g_1(\hat\theta_1)}{g_{\{1,n\}}(\hat\theta_1)}\cdot\frac{p_n[(v-\hat\theta_1)g_n(\hat\theta_1) - G_n(\hat\theta_1)] + \lambda_{1n}G_1(\hat\theta_1)}{g_1(\hat\theta_1)} = \frac{g_1(\hat\theta_1)}{g_{\{1,n\}}(\hat\theta_1)}\cdot\Psi_n(\hat\theta_1, C_n) \tag{75}$$

$$= \frac{g_1(\hat\theta_1)}{g_{\{1,n\}}(\hat\theta_1)}\cdot\Psi_n(\hat\theta_n, C_n), \tag{76}$$

where in the last line we have used that $\lambda_{1n} < 0$ by Step 1 which implies $\hat\theta_1 = \hat\theta_n$ by (54). Hence, by (53):

$$\Psi^{n-1}_{\{1,n\}}(\hat\theta_1, C_{n-1}) = \frac{g_1(\hat\theta_1)}{g_{\{1,n\}}(\hat\theta_1)}\cdot\Psi_1(\hat\theta_1, C_n) + \frac{g_1(\hat\theta_1)}{g_{\{1,n\}}(\hat\theta_1)}\cdot\Psi_n(\hat\theta_n, C_n) = 0, \tag{77}$$

and this completes the proof of the right part of (60).

• To complete the proof of (59), it remains to be shown that $(\mathbf{x}^{n-1}, \lambda^{n-1})$ satisfies $(KT_3)^{n-1}$. Consider $(\gamma,\delta)\in C_{n-1}$. If $\lambda^{n-1}_{\gamma\delta} = 0$, $(KT_3)^{n-1}$ holds trivially. If $\lambda^{n-1}_{\gamma\delta} < 0$, then (54)-(58) imply that $x^{n-1}_\gamma = x^{n-1}_\delta$. Accordingly, $(KT_3)^{n-1}$ also holds in this case.

This establishes (59) and completes the proof. Q.E.D.

Proof of Proposition 1 For all $k \in \omega$, let $x_k^s(\theta) = 1_{[0,\theta_\omega](\theta)}$ be the production schedule and $u_k^s(1) = 0$ the least efficient ex post type's utility level under the static contract. By the Kuhn–Tucker theorem, we have to show that there are multipliers $\lambda_{ij} \leq 0$, $(i,j) \in C_n$, and $\mu_k \leq 0$, $k \in \omega$, so that $(\{x_k^s(\cdot)\}_{k\in\omega}, \{u_k^s(1)\}_{k\in\omega})$ maximizes the Lagrangian

$$\begin{aligned}\mathcal{L} &= \sum_{k\in\omega}\left\{\int_0^1 p_k[v-\theta-h_k(\theta)]x_k(\theta)g_k(\theta)\,d\theta - p_k u_k(1)\right\} \qquad (78)\\ &\quad -\sum_{(i,j)\in C_n}\lambda_{ij}\left[\int_0^1[x_i(\theta)-x_j(\theta)]G_i(\theta)\,d\theta + u_i(1)-u_j(1)\right] - \sum_{k\in\omega}\mu_k u_k(1)\\ &= \sum_{k\in\omega}\int_0^1\left\{p_k[v-\theta-h_k(\theta)] - \sum_{j:(k,j)\in C_n}\lambda_{kj}h_k(\theta) + \sum_{i:(i,k)\in C_n}\lambda_{ik}h_{i,k}(\theta)\right\}x_k(\theta)g_k(\theta)\,d\theta \qquad (79)\\ &\quad -\sum_{k\in\omega}\left\{p_k + \sum_{j:(k,j)\in C_n}\lambda_{kj} - \sum_{i:(i,k)\in C_n}\lambda_{ik} + \mu_k\right\}u_k(1),\end{aligned}$$

where $\lambda_{ij} = 0$ or $\mu_k = 0$ only if the respective constraints are not binding. Now, let $\lambda_{ij} < 0$ be as in the proof of Lemma 15, and define

$$\mu_k = \begin{cases} -p_k - \sum_{j:(k,j)\in C_n}\lambda_{kj} & if \quad k \in \omega_{C_n}^+ \\ -p_k + \sum_{i:(i,k)\in C_n}\lambda_{ik} & if \quad k\in\omega_{C_n}^- \end{cases}. \qquad (80)$$

Then the curly brackets in the last line are zero, and the Lagrangian $\mathcal{L}$ is identical to the Lagrangian for the problem $\mathcal{R}^0(C)$. Therefore, by Lemma 11 and 15, $(\{x_k^s(\cdot)\}_{k\in\omega}, \{u_k^s(1)\}_{k\in\omega})$ maximizes $\mathcal{L}$. It remains to be shown that $\mu_k \leq 0$. Since $\lambda_{ik} < 0$, the claim is trivial for $k \in \omega_{C_n}^-$. For $k \in \omega_{C_n}^+$, recall from (32) in the proof of Lemma 10 that $-p_k - \sum_{j:(k,j)\in C_n}\lambda_{kj} \geq 0$. This completes the proof. Q.E.D.

References

Baron, D. and D. Besanko (1984). "Regulation and information in a continuing relationship." *Information Economics and Policy* 1, 267-302.

Battaglini, M. (2005). "Long-Term Contracting with Markovian Consumers." American Economic Review 95, 637-658.

Chiu, L. Y. and D. E. M. Sappington (2010). "Contracting with Private Knowledge of Signal Quality." *RAND Journal of Economics* 41, 244-269.

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

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

Esö, P. and B. Szentes (2007a). "The Price of Advise." *RAND Journal of Economics* 38, 863–880.

Esö, P. and B. Szentes (2007b). "Optimal Information Disclosure in Auctions and the Handicap Auction." *Review of Economic Studies* 74, 705-731.

Hoffmann, F. and R. Inderst (2009). "Price Discrimination and the Provision of Information." mimeo.

Krähmer, D. and R. Strausz (2008). "Ex Post Private Information and Monopolistic Screening." *The B.E. Journal of Theoretical Economics, Topics* 8(1), Article 25.

Krähmer, D. and R. Strausz (2011). "Optimal Procurement Contracts with Pre-Project Planning." *Review of Economic Studies* 78, 2011, 1015-1041.

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

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

Pavan, A., I. Segal, and J. Toikka (2008). "Dynamic Mechanism Design: Revenue Equivalence, Profit Maximization and Information Disclosure." mimeo.

Sappington, D. E. M. (1983). "Limited liability contracts between principal and agent." *Journal of Economic Theory* 29, 1–21.