Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 188 (2020) 105055

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

The scope of sequential screening with ex post participation constraints ☆

Dirk Bergemann [a,*], Francisco Castro [b], Gabriel Y. Weintraub [c]

[a] *Department of Economics, Yale University, New Haven, USA*
[b] *Anderson School of Management, UCLA, Los Angeles, USA*
[c] *Graduate School of Business, Stanford University, Stanford, USA*

Received 18 July 2019; final version received 1 April 2020; accepted 5 April 2020
Available online 16 April 2020

Abstract

We study the classic sequential screening problem in the presence of ex post participation constraints. We establish necessary and sufficient conditions that determine when the optimal selling mechanism is *either* static *or* sequential. In the *static contract*, the buyers are not screened with respect to their interim type and the object is sold at a posted price. In the *sequential contract*, the buyers are screened with respect to their interim type and a menu of quantities is offered.

We completely characterize the optimal sequential contract with binary interim types and a continuum of ex post values. Importantly, the optimal sequential contract randomizes the allocation of the low-type buyer and awards a deterministic allocation to the high type buyer. Finally, we provide additional results for the case of multiple interim types.
© 2020 Elsevier Inc. All rights reserved.

JEL classification: C72; D82; D83

Keywords: Sequential screening; Ex post participation constraints; Static contract; Sequential contract

☆ The first author acknowledges financial support from NSF Grant SES 1459899. We are grateful to the Editor, Alessandro Pavan, and three anonymous referees for productive suggestions and comments. We thank Ian Ball, Laurens Debo, Alessandro Bonatti, Yang Cai, Tibor Heumann, Daniel Krähmer, Andy Skrzypacz, Roland Strausz, and Matt Weinberg as well as seminar participants at Columbia, Chicago Booth, Kellogg, Michigan Ross, PUC Chile, Rice, Stanford, U. of Chile, and various conferences for helpful conversations.

* Corresponding author.

E-mail addresses: dirk.bergemann@yale.edu (D. Bergemann), francisco.castro@anderson.ucla.edu (F. Castro), gweintra@stanford.edu (G.Y. Weintraub).

https://doi.org/10.1016/j.jet.2020.105055
0022-0531/© 2020 Elsevier Inc. All rights reserved.

1. Introduction

1.1. Motivation

Sequential screening models have been used extensively in economics and revenue management to study optimal contract design when buyers learn their values over time. In the classic formulation of sequential screening pioneered by Courty and Li (2000), a profit-maximizing seller (he) faces a single buyer (she) or, alternatively, a continuum of buyers. The buyer initially has partial and private information about her value, for example the mean, and privately learns her true value at some later time. In the classic setting, each buyer is required to participate *ex interim*: her *expected gains* at the time of contracting have to exceed their outside option. A salient example discussed by Courty and Li (2000) is the airline industry, in which travelers purchase tickets in advance but may only realize their true value as the date of the trip approaches.

Although the optimal contracts that arise may offer partial refunds, the initial advanced price is high enough such that some travelers experience negative *ex post* utility while still being willing to participate ex interim. This situation also arises in other industries, such as hotels, theaters or even railroads where advanced pricing and partial refunds contracts are also offered.

In many online markets, however, the seller is constrained to sell products such that the buyer obtains a nonnegative net utility once she has realized her value, thus *ex post*. For example, in online shopping, buyers may have the option to return a purchased item after delivery, usually at zero or low cost (Krähmer and Strausz (2015)). In the online display advertising market, typical business constraints prohibit publishers from using upfront fees (Balseiro et al. (2018)). Instead, the publishers run auctions, typically some version of first- or second-price auctions that satisfy the ex post participation constraints. Thus, the seller needs to guarantee participation not only initially – at the interim level – but also after the buyers have completely learned their value – at the ex post level.

Motivated by these new markets, we study the sequential screening problem as described by Courty and Li (2000) and incorporate *ex post participation constraints*. Ex post participation constraints rule out the optimal contracts derived by Courty and Li (2000) with upfront fees. As pointed out by Krähmer and Strausz (2015), because different upfront fees cannot be used to price discriminate the different buyers, it may be that a *static* contract, one that does not screen the buyers ex interim, becomes optimal under ex post participation constraints. Building on the work by Krähmer and Strausz (2015), our objective is to understand when the optimal selling mechanism is *static* (buyers are not screened ex interim) or *sequential* (buyers are screened interim), and to obtain a full characterization of such contracts. Our work highlights the significant revenue improvements that can be attained by using a sequential relative to a static contract, even in the presence of ex post participation constraints.

Our model considers a seller who is selling at most one unit of an object to a buyer. The sequence of events unfolds in two periods. In the first period, the buyer privately learns her *interim* type, for example the mean of her value distribution, and the parties contract. We begin the analysis assuming binary interim types of the buyer, thus high and low. The high type has a distribution of *ex post* values that dominates the distribution of the low-type in some stochastic order. The contract specifies allocation and payments as a function of reported interim type and ex post value. In the second period, the buyer privately learns her value, and allocations and transfers are realized. At this point, the buyer accepts the contracting terms only if her realized net utility is weakly larger than her outside option. This model aligns with our aforementioned examples. In online shopping, the first period corresponds to the purchasing time. At this time the buyer possesses private information about her expected value but she only learns her true realized value in the subsequent period. In the second period, the buyer is delivered the item and has the option to return it, at low or no cost. In the case of display advertising, some publishers use a sequence of auctions known as "waterfall auctions" that implicitly impose different priorities over participants.[1] Commonly, higher-priority auctions have higher reserve prices. The first period can be regarded as the time at which the buyer decides in which auction (priority/reserve) to participate. The second period is when the auctions are actually run.

1.2. Results

The first main result characterizes when a static contract—that is, a contract that does not sequentially screen buyers—is optimal. In Theorem 1, we provide a necessary and sufficient condition for the optimality of the static contract, termed the *profit-to-rent condition*. In the optimal static contract the seller offers a single and uniform price to all types.

In Theorem 2, we characterize the optimal mechanism when this *profit-to-rent* condition fails and a static contract is no longer optimal. The scope for revenue improvement through a sequential contract is perhaps easiest to grasp by assuming for a moment that the seller were to know the interim type. From this, admittedly hypothetical, perspective, the uniform static price is too high for the low-type and too low for the high type. As each type has a different ex post distribution of values, the seller would ideally prefer to better tailor the price to the distribution of ex post values. To increase his revenue relative to the static contract, the seller could try to increase the price for the high-type buyer or decrease the price for the low-type buyer. However, either change would lead the high type to mimic the low type. A more promising option is to lower the allocation for (some) low-type buyers while simultaneously reducing the price charged to them. This allows the seller to serve more ex post values of the low type while deterring the high types from taking the low types' contract. Now, the *profit-to-rent* condition establishes exactly when this pricing deviation is not profitable for the seller. The profit-to-rent condition is hence necessary for the optimality of the static contract. Notably, we also show that it is sufficient. The profit-to-rent condition is a weighted monotonicity condition for the virtual value around the optimal static threshold. In the case of exponentially distributed values, we can show that the static contract is optimal if and only if the means of the distributions of the low and high types are sufficiently close.

In line with the above intuition, we find in Theorem 2 that the optimal *sequential* contract provides a lower quantity to the low type, or equivalently randomizes the allocation of the object between 0 and 1, and assigns a deterministic allocation of 1 to the high type. Randomization is needed to deter the high-type buyer from taking the low type's contract. Specifically, the optimal contract is characterized by an allocation probability $x \in (0, 1)$, and three thresholds θ_1, θ_2, and θ_H with $\theta_1 \leq \theta_H \leq \theta_2$. In this contract, the seller allocates the object to a low-type buyer with probability x whenever her value is between θ_1 and θ_2 and asks for a payment of $\theta_1 \cdot x$. When the true value of the low type is above θ_2, then the object is always allocated to her, and the seller demands a payment of $\theta_2 - (\theta_2 - \theta_1) \cdot x$. The high-type buyer obtains the object with certainty and only when her value is above θ_H, at which point the payment she has to make to the seller is θ_H. These parameters are set such that the interim incentive compatibility constraints are satisfied.

[1] See, for example, https://adexchanger.com/the-sell-sider/the-programmatic-waterfall-mystery. A similar dynamic occurs when sellers offer "preferred deals" to advertisers (see, for example, Mirrokni and Nazerzadeh (2017)).

A salient feature of this type of contract is that it discriminates the low type in *two dimensions*. First, we establish that θ_1 is above the threshold a seller would set if she were selling exclusively to low- type buyers. That is, the low-type buyer is allocated the object less often in the presence of high-type buyers. The opposite holds for high-type buyers: they are allocated the object more often than if they were alone. Second, there is a range of values for which the object is sold to the low type with some probability strictly below one, which further reduces the likelihood that a low type will receive the object compared to a case in which there are no high-type buyers. We illustrate these results with the example of the exponential distribution, for which we have explicit solutions. We find that for exponential values, the sequential contract can exhibit revenue improvements exceeding 40% over the static contract.

Towards the end of the paper, we consider several extensions of our base model. Notably, we study the case of many interim types. Theorem 3 generalizes the *profit-to-rent* condition to a setting with an arbitrary number of interim types. We also explore the structure of the optimal sequential contract and the challenges that arise in this setting.

1.3. Related work

Our model builds on the sequential screening literature, as pioneered by Courty and Li (2000), with an *interim participation constraint*.[2] In contrast, in this paper, we impose an *ex post participation constraint*. The most closely related paper to ours that studies sequential screening with ex post participation constraints is Krähmer and Strausz (2015). They establish that the static contract is optimal under a monotonicity condition regarding the cross-hazard rate functions. This condition rules out some common distributions for values such as the exponential distribution. Furthermore, the condition is only sufficient and, therefore, does not provide a complete characterization of when the static contract is optimal. We close this gap by providing a necessary and sufficient condition under which the static contract is optimal. Our condition leverages the economic intuition that lies behind a potential profitable deviation from the optimal static contract. Furthermore and importantly, when the condition fails, we characterize the optimal *sequential* mechanism and show that randomization of one of the interim types is required for optimality.[3]

In terms of approaches, Krähmer and Strausz (2015) relax both the local incentive constraint of the low-type and the monotonicity constraint. Then, they show that under these conditions, the contract that maximizes the Lagrangian is deterministic and that, as a result, the static contract is optimal. In contrast, we also relax the local incentive constraint but maintain the monotonicity constraint. For the relaxed problem, we perform a first-principle analysis, in the style of Samuelson (1984) and Fuchs and Skrzypacz (2015), that leads us to identify the structure of the optimal contract. In turn, this permits us to characterize the optimal sequential contract when the static condition fails. In a recent work, Heumann (2020) considers a setting in which a seller can design the screening mechanism and the information disclosure mechanism with ex post participation constraints.

[2] See Akan et al. (2015) for a recent adaptation of the Courty and Li (2000) formulation to study advanced purchase contracts in revenue management settings.

[3] See also Manelli and Vincent (2007) and Daskalakis et al. (2015) for examples of multi-good environments in which stochastic allocations can improve over deterministic allocations. In a separate contribution, Krähmer and Strausz (2016) establish that with multiple units, as opposed to a single unit, generically, the static contract is not optimal for the sequential screening problem with ex post participation constraints.

The sequential nature of our model and the presence of ex post participation constraints is related to the work of Ashlagi et al. (2016) and Balseiro et al. (2018). These authors consider a model (also motivated by the display advertising market) in which a seller, constrained by ex post participation, repeatedly sells objects to a buyer whose values are independent across periods. Both papers provide characterizations for a nearly optimal mechanism. They are different from ours because we consider a single sale and construct the exact optimal mechanism in a sequential screening model. Krähmer and Kovac (2016) share our concern with static vs. sequential mechanisms in a delegation environment. While the delegation environment in Krähmer and Kovac (2016) is substantially different from the quasi-linear environment that we investigate here, some of our arguments are similar to theirs. In particular, in Theorem 1, we establish that a simple necessary condition for optimality can be extended to a necessary and sufficient condition. The necessary condition involves a comparison of cost and benefits in terms of virtual values, in a manner similar to Proposition 3 in Krähmer and Kovac (2016).[4]

Our optimal mechanism is related to the BIN-TAC auction derived in the context of online display advertising by Celis et al. (2014). This is a *static* auction that offers two options to advertisers: a buy-it-now (BIN) option in which buyers can purchase the impression at a posted high price, and a take-a-chance (TAC) option in which the highest bidders are randomly allocated the impression (if no bidder went for the BIN). This auction is tailored to approximate ironing in the classic static Myerson setting for nonregular distributions that commonly arise in display advertising settings. This mechanism is similar in spirit to ours because it randomizes low-value buyers to separate them from high-values buyers. However, with one bidder, the BIN-TAC auction reduces to a posted price which corresponds to the static contract in our setting. In contrast to their static setting, we study a two-period model in which the buyer is sequentially screened, and randomization occurs even with a single bidder.

2. Model

2.1. Payoffs

We consider a seller (he) who is selling one unit of an object at zero cost to a buyer (she) with an outside option of zero value. Both parties are risk-neutral and have quasilinear utility functions. The sequence of events unfolds in two periods.

In the first period, the buyer privately learns her *interim type* (or simply *type*) and then the parties contract. The type provides information about the distribution of the *ex post values* (or simply *value*) of the buyer— her true willingness-to-pay for the object. The contract specifies allocation and payment as a function of reported interim type and ex post value. In the second period, the buyer privately learns her value, and allocations and transfers are realized.

There are finitely many types, denoted $k \in \{1, \dots, K\}$, and the prior probability of type k is given by α_k with $\alpha_k > 0$ and $\sum_{k=1}^{K} \alpha_k = 1$. In the second period, a buyer of type k privately learns her value θ which we assume to have a continuously differentiable distribution function $F_k(\theta)$ and associated density function $f_k(\theta)$, with full support in $\Theta \subseteq [0, \infty]$. We assume that Θ is a connected interval of the form $[0, \overline{\theta}]$. It will be convenient to denote the upper cumulative distribution function by:

$$\overline{F}_k(\theta) \triangleq 1 - F_k(\theta).$$

[4] We thank the editor and an anonymous referee for drawing our attention to the result in Krähmer and Kovac (2016).

All the distributions are common knowledge. The virtual value of interim type k is given by:

$$\mu_k(\theta) \triangleq \theta - \frac{1-F_k(\theta)}{f_k(\theta)}, \quad \forall k \in \{1, \ldots, K\}, \quad \forall \theta \in \Theta.$$

For the remainder of the paper, we make the standard assumption that the hazard rate

$$\frac{f_k(\theta)}{1-F_k(\theta)}, \quad \text{is increasing in } \theta, \forall k \in \{1, \ldots, K\}. \tag{IHR}$$

This assumption facilitates our discussion. However, our formal results will require a weaker assumption that we introduce later.

The terms of trade are specified by the seller in the first period. For a payment $t \in \mathbb{R}$ and a probability of receiving the object $x \in [0, 1]$, a buyer with value θ receives a utility of $\theta \cdot x - t$, while the seller is paid t.

We assume that the buyer agrees to purchase the object only if she is guaranteed a nonnegative net utility for any possible value of the object she might have. That is, we require $\theta \cdot x - t$ to be nonnegative for all θ. The seller's problem is to design a contract that maximizes his expected payment, satisfying the ex post participation and incentive compatibility constraints.

2.2. *Direct mechanism*

By means of the revelation principle (see, e.g., Myerson (1979)) we can focus on incentive compatible direct revelation mechanisms, with allocations $x_k : \Theta \to [0, 1]$ and transfers $t_k : \Theta \to \mathbb{R}$ that depend on reported interim type k' and ex post value θ'. Then, for a buyer reporting an interim type k' and an ex post type θ', the mechanism allocates the object with probability $x_{k'}(\theta')$ and charges the buyer $t_{k'}(\theta')$.

We define the ex post utility of a buyer who truthfully reported k in the first period and θ' in the second period while her true value is θ as

$$u_k(\theta; \theta') \triangleq \theta \cdot x_k(\theta') - t_k(\theta'),$$

with the understanding that $u_k(\theta) \triangleq u_k(\theta; \theta)$. Similarly, we define the interim expected utility of a buyer whose true interim type is k but reported to the mechanism k' and is truthful in the second period as

$$U_{kk'} \triangleq \int_{\Theta} u_{k'}(z) \cdot f_k(z) dz.$$

We note that with distributions with common support Θ, we can restrict attention to single deviations.

There are two kinds of incentive compatibility constraints that must be satisfied by our mechanism. The first is the ex post incentive compatibility constraint (IC^{xp}), which requires that for any report in the first period, truth-telling is optimal in the second period:

$$u_k(\theta) \geq u_k(\theta; \theta') \quad \forall k \in \{1, \ldots, K\}, \forall \theta \in \Theta. \tag{IC^{xp}}$$

The second is the interim incentive compatibility constraint (IC^i) which requires that truth-telling is optimal in the first period:

$$U_{kk} \geq U_{kk'} \quad \forall k, k' \in \{1, \ldots, K\}. \tag{IC^i}$$

Finally, we require the mechanism to satisfy the ex post individual rationality constraint (IR^{xp}):

$$u_k(\theta) \geq 0, \quad \forall k \in \{1, \dots, K\}, \quad \forall \theta \in \Theta. \tag{IR^{xp}}$$

Then, the seller's problem is

$$\begin{aligned} \max \quad & \sum_{k=1}^{K} \alpha_k \cdot \int_{\Theta} t_k(z) \cdot f_k(z) dz \qquad (\mathcal{P}) \\ & \text{s.t } (IC^i), (IC^{xp}), (IR^{xp}) \\ & 0 \leq \mathbf{x} \leq 1 , \end{aligned}$$

where we use boldfaces to denote the vector $\mathbf{x} = (x_1, ..., x_K)$. Observe that ($IR^{xp}$) implies interim individual rationality (IR). In fact, if we were to relax ($\mathcal{P}$) by considering only interim IR we would be in the setting of Courty and Li (2000) for discrete interim types.

In general, one of two types of contracts can arise as an optimal solution to the seller's problem ($\mathcal{P}$): *static* or *sequential*. A static solution to problem ($\mathcal{P}$) corresponds to the case in which the allocations and transfers (x_k, t_k) do not depend on the interim type k. In this case, we have a single menu (x, t) that is offered to the buyer, and the contract does not screen among interim types. We use ($\mathcal{P}^s$) to denote the version of ($\mathcal{P}$) constrained to static contracts, which we refer to as *the static program*. In contrast, a sequential solution allows for different menus that depend on the interim type k, and each type of buyer self-selects into one of the menus. The problem ($\mathcal{P}$), referred to as *the sequential program*, allows for such solutions.

The main focus of this paper is twofold. The first is to study when the optimal solutions to the static and sequential programs, ($\mathcal{P}^s$) and ($\mathcal{P}$), coincide. Second, when they do not coincide, we aim to characterize the optimal solution to ($\mathcal{P}$).

3. A classic example of sequential screening

We use the opening example of Courty and Li (2000) to illustrate the power of sequential screening in the presence of an ex post participation constraint. We show that a sequential contract outperforms the static contract.

In the opening example, there are two types of potential buyers, low-type and high type. One-third of potential buyers are low-type with value uniformly distributed in $[1, 2]$; two-thirds are high-type buyers with value uniformly distributed in $[0, 1] \cup [2, 3]$.[5] Courty and Li (2000) regard of the low type as a leisure traveler and the high type as a business traveler with the same mean but larger variance in her value. The seller has a production cost equal to 1.

The optimal static contract sets the optimal monopoly price, $\widehat{p}$, equal to 2, which yields a profit of $1/3$. The static contract only serves high types who have high realized values. Courty and Li (2000) show that the seller can significantly increase his profits with sequential screening by offering a menu of advanced payments/partial refund contracts subject to the weaker *interim participation constraints*. The optimal contract offers an advanced payment of 1.5 and no refund to the leisure traveler and an advanced payment of 1.75 and a partial refund of 1 to the business traveler. Note that in this contract some buyers will experience a realized negative net utility. For example, the leisure traveler initially pays 1.5, but her actual value can be any value

[5] We note that the opening example of Courty and Li (2000) violates the common support assumption made above in Section 2. However, the failure of the common support does not affect our argument.

within $[1, 2]$, and therefore, half of the time, she will obtain negative net utility after learning her value. Because of the advanced payment, the contract does not satisfy the *ex post participation constraint*.

By contrast, the following version of a sequential contract does satisfy the ex post participation constraints. The seller offers a menu of two quantities and prices, (x_L, p_L) and (x_H, p_H). The high item is set equal to the optimal static contract, that is, $(x_H, p_H) = (1, 2)$. Thus, the selling price for the high type is 2, and high types that buy receive the full quantity. Next, we determine the optimal quantity and price for the low-type buyer. Given the contract for the high type, the seller's profit is given by:

$$\frac{1}{3} \times x_L \times (p_L - 1) \times (2 - p_L) + \frac{2}{3} \times \frac{1}{2} \times (2 - 1),$$

where $x_L \in [0, 1]$ and $p_L \in [1, 2]$. We need to ensure that the menu is interim incentive compatible. The incentive constraint of the low type is always satisfied (p_H equals 2), and the incentive constraint of the high type is given by:

$$\frac{1}{2} \times \left(\frac{5}{2} - 2\right) \geq \frac{1}{2} \times x_L \times \left(\frac{5}{2} - p_L\right).$$

Profit maximization implies that this constraint must be binding, and therefore, the seller's profit becomes:

$$\frac{1}{3} \times \frac{(p_L - 1) \times (2 - p_L)}{5 - 2p_L} + \frac{1}{3}.$$

The first-order condition yields an optimal price equal to $\left(5 - \sqrt{3}\right)/2$ that, in turn, delivers a profit of $2/3 - 1/(2\sqrt{3})$. The improvement of the sequential contract versus the optimal static contract is then $1 - \sqrt{3}/2 \approx 13\%$.

From this basic exercise, we learn an important lesson: even in this simple setting, a sequential contract can have substantial benefits over a static contract. In this paper, we study more generally when a sequential contract outperforms a static contract and what drives this revenue improvement.

4. Optimality of static contract

In the main result of this section, Theorem 1, we provide a necessary and sufficient condition for the static contract to be optimal. We begin with a reformulation of the problem based on standard techniques that use the envelope theorem, and enable us to solve for the allocation and utilities of the lowest ex post types instead of both allocations and transfers. Using the reformulation we characterize the optimal static contract. In Section 4.2, we use the optimal static contract and a simple deviation analysis to obtain an intuitive necessary condition for its optimality. In Section 4.3, we show that this condition is both necessary and sufficient.

4.1. Problem reformulation and static solution

We obtain a more amenable characterization of the constraints by eliminating the transfers as in the classical Myersonian analysis.

Lemma 1 *(Necessary and Sufficient Conditions for Implementation). The mechanism* $(\mathbf{x}, \mathbf{t})$ *satisfies* (IC^i),(IC^{xp}) *and* (IR^{xp}) *if and only if*

1. $x_k(\cdot)$ *is a nondecreasing function for all* k *in* $\{1, \ldots, K\}$ *and*

$$u_k(\theta) = u_k(0) + \int_0^\theta x_k(z)dz, \quad \forall k \in \{1, \ldots, K\}, \forall \theta \in \Theta. \tag{1}$$

2. $u_k(0) \geq 0$ *for all* k *in* $\{1, \ldots, K\}$.
3. $u_k(0) + \int_\Theta x_k(z)\overline{F}_k(z)dz \geq u_{k'}(0) + \int_\Theta x_{k'}(z)\overline{F}_k(z)dz$ *for all* k, k' *in* $\{1, \ldots, K\}$.

All proofs are provided in the Appendix. The first condition in the lemma is the standard envelope condition and comes from the ex post incentive compatibility constraint. The second condition is derived from the ex post IR constraint and the fact that $u_k(\theta)$ is nondecreasing. The third condition is the envelope formula inserted into the interim incentive compatibility constraint.

Lemma 1 enables us to obtain a more compact formulation of the seller's problem. Specifically, we can use equation (1) and integration by parts to write the objective of $(\mathcal{P})$ in terms of the allocation rule $\mathbf{x}$ and the indirect utilities $\{u_k(0)\}_{k=1}^K$ of the lowest ex post types. To this end, we denote each $u_k(0)$ as a new variable by u_k. The new formulation is then:

$$\max_{0 \leq \mathbf{x} \leq 1, \mathbf{u}} \quad -\sum_{k=1}^K \alpha_k u_k + \sum_{k=1}^K \alpha_k \int_\Theta x_k(z)\mu_k(z) f_k(z)dz \tag{$\mathcal{P}$}$$

$$\text{s.t} \quad x_k(\theta) \quad \text{nondecreasing}, \quad \forall k \in \{1, \ldots, K\}$$

$$u_k \geq 0, \quad \forall k \in \{1, \ldots, K\}$$

$$u_k + \int_\Theta x_k(z)\overline{F}_k(z)dz \geq u_{k'} + \int_\Theta x_{k'}(z)\overline{F}_k(z)dz, \quad \forall k, k' \in \{1, \ldots, K\}.$$

Note that in $(\mathcal{P})$, the variables are the allocation rule $\mathbf{x}$ and the vector of the indirect utilities of the lowest ex post types $\mathbf{u}$. Once we solve for these variables the transfers are determined by equation (1).

As noted above, a solution to $(\mathcal{P})$ that screens the interim types is a sequential contract. In contrast, a static solution to $(\mathcal{P})$ pools the interim types. Formally, we say that a solution to $(\mathcal{P})$ or contract is *static* when $x_k(\cdot) \triangleq x(\cdot)$ and $u_k \triangleq u$ for all k in $\{1, \ldots, K\}$.

We previously defined the virtual value $\mu_k(\cdot)$ of interim type k. Given (IHR), the virtual value for each type k has exactly one zero, which we denote by $\widehat{\theta}_k$. Without loss of generality, we assume *for the remainder of the paper* that we have ordered the interim types such that

$$\widehat{\theta}_1 \leq \cdots \leq \widehat{\theta}_K.$$

It turns out that solving $(\mathcal{P})$ over the space of static contracts is a simpler problem. The (IC^{xi}) constraints disappear from the problem because in this case there is effectively only one interim type. Additionally, it is clear that any optimal solution sets $u_k = 0$ for all k in $\{1, \ldots, K\}$. Therefore, the static version of the seller's problem is given by

$$\max_{0\leq x\leq 1} \int_{\Theta} x(z)\cdot\Big(\sum_{k=1}^{K}\alpha_k\mu_k(z)f_k(z)\Big)dz \tag{$\mathcal{P}^s$}$$
$$\text{s.t}\quad x(\theta)\quad \text{nondecreasing},$$

where a simple calculation shows that the term in parentheses is equal to the virtual value function of the mixture distribution times the density function of the mixture. Hence, this problem corresponds to the classic optimal monopoly price problem applied to the mixture distribution over types. The relevant quantity that shapes the optimal allocation $x(\cdot)$ is:

$$\bar{\mu}(\theta) \triangleq \sum_{k=1}^{K}\alpha_k\mu_k(\theta)f_k(\theta).$$

As shown by Riley and Zeckhauser (1983) in the case of a single buyer, an optimal solution that maximizes

$$\int_{\Theta} x(z)\bar{\mu}(z)dz, \tag{2}$$

is always given by a threshold value $\widehat{\theta}$, which can be implemented by a single posted price $\widehat{p}=\widehat{\theta}$.

Lemma 2 *(Threshold Allocation). A solution to* ($\mathcal{P}^s$) *is a threshold value characterized by* $\widehat{\theta}\in[\widehat{\theta}_1,\widehat{\theta}_K]$ *that maximizes* (2).

4.2. A necessary condition

In the remainder of this and the next section, we state the results for the setting with binary interim types. We denote the low-type by L and the high type by H. In Section 6.1, we return to the general setting with finitely many interim types.

The static optimal solution is characterized by a threshold value $\widehat{\theta}$. In this section, we leverage this characterization and perform an analysis in the style of Bulow and Roberts (1989), to deduce an intuitive necessary condition for the optimality of the static contract. As we will show later in Section 4.3 this condition turns out to be not only necessary but also sufficient.

For ease of exposition, we assume that the high type dominates the low-type in the hazard rate order sense:

$$\frac{1-F_H(\theta)}{f_H(\theta)} \geq \frac{1-F_L(\theta)}{f_L(\theta)}, \qquad \forall\theta\in\Theta. \tag{3}$$

We note that we do not need this assumption for the formal arguments.

Suppose now that a static contract is optimal, that is, setting a single posted price equal to $\widehat{\theta}$ for both types solves ($\mathcal{P}$). Consider Fig. 1, where we have plotted the virtual value weighted by the density function for each type.[6] If the types were public information, the seller would optimally set posted prices equal to $\widehat{\theta}_L$ and $\widehat{\theta}_H$ for types L and H, respectively. In this way, the seller would serve buyers if and only if they have positive virtual values. In contrast, when selecting a single posted price $\widehat{\theta}$, there is surplus that the seller is not extracting; the shaded area shows the regions of the virtual values for each type that the static contract is not capturing. For the high type, the

[6] Representing the virtual value weighted by the density $f_k(\cdot)$ allows for a convenient geometric argument in which the seller's revenue from each type k is the area under the corresponding curve representing $\mu_k(\cdot)f_k(\cdot)$.

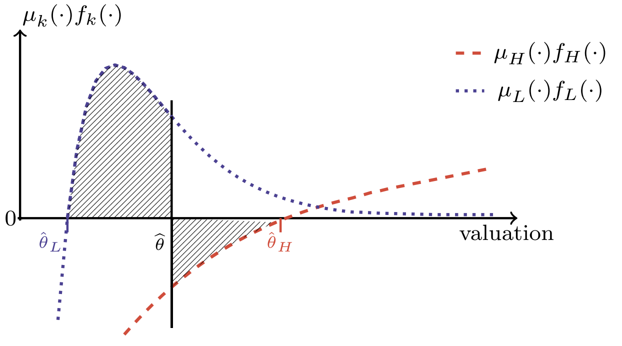

Fig. 1. Weighted virtual valuations for low-type (dotted line) and high type (dashed line) buyer around $\widehat{\theta}$. The shaded areas correspond to the virtual revenue that the seller misses when using a static contract with respect to the case in which the interim types are public information.

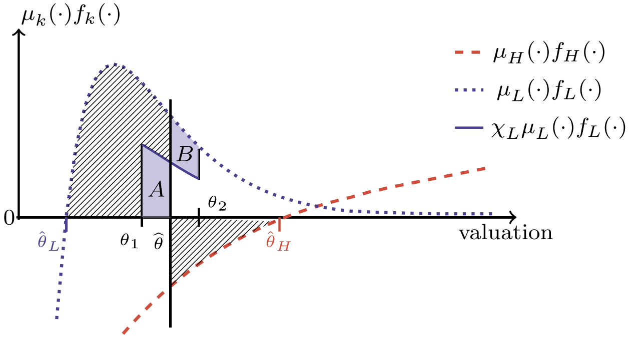

Fig. 2. Weighted virtual valuations for low-type (dotted line) and high-type (dashed line) buyers around $\widehat{\theta}$. The shaded areas correspond to the virtual revenue that the seller leaves on the table when using a static contract with respect to the case in which the interim types are public information. We display the deviation from the static contract for the low-type (solid line). If the solid areas A and B are such that $A - B \geq 0$, the deviation is profitable.

static contract serves too many buyers, some of them with negative virtual values; hence, the seller would be better off by offering a higher price. For the low-type, the static contract serves too few buyers, leaving positive virtual value buyers unserved; hence, the seller would prefer to choose a lower price. A challenge, however, is that the seller faces incentive compatibility constraints that restrict such possible deviations/improvements:

1. Selling to fewer high types implies increasing the price for high types; however, the high types then have an incentive to accept the low-type contract, and such a deviation is not feasible.
2. Selling to more low types amounts to reducing the price from $\widehat{\theta}$ to some value θ_1. However, to prevent the high types from taking the low-type contract the seller must decrease the quantity offered to the low types (or equivalently, randomize their allocation).

This second improvement is feasible by choosing a quantity (probability) $0 < x_L < 1$ for all low types within an interval $[\theta_1, \theta_2]$ with $\theta_1 \leq \widehat{\theta} \leq \theta_2$; see Fig. 2.

Formally, these allocations correspond to the following menu:

$$x_L(\theta) \triangleq \begin{cases} 0 & \text{if } \theta < \theta_1, \\ x_L & \text{if } \theta_1 \leq \theta \leq \theta_2, \\ 1 & \text{if } \theta_2 < \theta; \end{cases} \quad x_H(\theta) \triangleq \begin{cases} 0 & \text{if } \theta < \widehat{\theta}, \\ 1 & \text{if } \widehat{\theta} \leq \theta; \end{cases} \tag{4}$$

with $u_L = u_H = 0$. We refer to this deviation as *an interior variation or improvement*.

The interior improvement is feasible only if it satisfies both incentive compatibility constraints. Inserting the menu (4) into the incentive constraints in ($\mathcal{P}$), we obtain the following for the low-type:

$$x_L \int_{\theta_1}^{\theta_2} (1 - F_L(\theta))d\theta + \int_{\theta_2}^{\bar{\theta}} (1 - F_L(\theta))d\theta \geq \int_{\widehat{\theta}}^{\bar{\theta}} (1 - F_L(\theta))d\theta,$$

and for the high type:

$$\int_{\widehat{\theta}}^{\bar{\theta}} (1 - F_H(\theta))d\theta \geq x_L \int_{\theta_1}^{\theta_2} (1 - F_H(\theta))d\theta + \int_{\theta_2}^{\bar{\theta}} (1 - F_H(\theta))d\theta,$$

and/or in a more compact form as a bracketing inequality:

$$\frac{\int_{\widehat{\theta}}^{\theta_2} (1 - F_L(\theta))d\theta}{\int_{\theta_1}^{\theta_2} (1 - F_L(\theta))d\theta} \leq x_L \leq \frac{\int_{\widehat{\theta}}^{\theta_2} (1 - F_H(\theta))d\theta}{\int_{\theta_1}^{\theta_2} (1 - F_H(\theta))d\theta}, \tag{5}$$

which contains both incentive compatibility constraints. The monotone hazard rate condition (3) guarantees that x_L as given by (5) always exists.[7] The interior variation is thus feasible, and we can select x_L to maximize the seller's revenue.

Indeed, evaluating the interior variation in the seller's objective yields:

$$x_L \cdot \int_{\theta_1}^{\theta_2} \mu_L(\theta) f_L(\theta)d\theta + \int_{\theta_2}^{\bar{\theta}} \mu_L(\theta) f_L(\theta)d\theta,$$

and since $\mu_L(\theta) \geq 0$ in $[\theta_1, \theta_2]$ (see Fig. 2) the right-hand side inequality in (5) must be tight.

With the interior variation, the seller serves more low-value buyers in $[\theta_1, \widehat{\theta}]$ at the level of x_L. This comes at the expense of offering a lower quantity, a loss of $1 - x_L$ to buyers with values in $[\widehat{\theta}, \theta_2]$. In Fig. 2, the area A corresponds to the additional revenue the seller can make due to the variation because he is serving more low-type buyers, and region B is the efficiency loss due to the incentive constraints.

If the static contract is optimal, then this variation cannot be profitable. In terms of Fig. 2 this means the areas must satisfy $A \leq B$. Hence, if the static contract is optimal, then

[7] Indeed, condition (3) is equivalent to $(1 - F_L(\theta))/(1 - F_H(\theta))$ being decreasing. Then $(1 - F_L(\theta))(1 - F_H(\theta')) \leq (1 - F_L(\theta'))(1 - F_H(\theta))$ for $\theta' \leq \widehat{\theta} \leq \theta$. Integrating both sides over $\theta' \in [\theta_1, \widehat{\theta}]$ and $\theta \in [\widehat{\theta}, \theta_2]$ implies $\left(\int_{\widehat{\theta}}^{\theta_2} (1 - F_L(\theta))d\theta\right)\left(\int_{\theta_1}^{\widehat{\theta}} (1 - F_H(\theta))d\theta\right) \leq \left(\int_{\theta_1}^{\widehat{\theta}} (1 - F_L(\theta))d\theta\right)\left(\int_{\widehat{\theta}}^{\theta_2} (1 - F_H(\theta))d\theta\right)$ from which (5) follows.

$$A = x_L \cdot \int_{\theta_1}^{\widehat{\theta}} \mu_L(\theta) f_L(\theta) d\theta \leq (1 - x_L) \cdot \int_{\widehat{\theta}}^{\theta_2} \mu_L(\theta) f_L(\theta) d\theta = B.$$

In turn, since the optimal choice of x_L always equals the right-hand side of (5), we can insert x_L in terms of the ratio, and after some rearranging, we obtain

$$\frac{\int_{\theta_1}^{\widehat{\theta}} \mu_L(\theta) f_L(\theta) d\theta}{\int_{\theta_1}^{\widehat{\theta}} (1 - F_H(\theta)) d\theta} \leq \frac{\int_{\widehat{\theta}}^{\theta_2} \mu_L(\theta) f_L(\theta) d\theta}{\int_{\widehat{\theta}}^{\theta_2} (1 - F_H(\theta)) d\theta}. \tag{6}$$

To better understand this inequality, consider a seller who faces a buyer with values distributed according to $F_k(\cdot)$. Observe that at any given price θ_b the expected profit $\Pi_k(\theta_b)$ of the seller and the expected *informational rent* $I_k(\theta_b)$ of the buyer are given by:

$$\Pi_k(\theta_b) \triangleq \theta_b \cdot (1 - F_k(\theta_b)) = \int_{\theta_b}^{\bar{\theta}} \mu_k(\theta) f_k(\theta) d\theta \quad \text{and} \quad I_k(\theta_b) \triangleq \int_{\theta_b}^{\bar{\theta}} (1 - F_k(\theta))\, d\theta.$$

If the monopolist considers lowering the price from θ_b to θ_a then the change in profit is $\Pi_k(\theta_a) - \Pi_k(\theta_b)$. The lower price positively impacts the information rents which increase by $I_k(\theta_a) - I_k(\theta_b)$. The ratio

$$\frac{\Pi_k(\theta_a) - \Pi_k(\theta_b)}{I_k(\theta_a) - I_k(\theta_b)}$$

is then a measure of the average impact on profits per unit of consumer rents that seller experiences due to the price variation.

Now, condition (6) can be rewritten to obtain a version of this ratio across different interim types. To this end, we set $k = L$ in the numerator and $k = H$ in the denominator. This suggests the following:

Definition 1 *(Average Profit-to-Rent Ratio).* The average profit-to-rent ratio is defined by:

$$R^{jk}(\theta_a, \theta_b) \triangleq \frac{\Pi_j(\theta_a) - \Pi_j(\theta_b)}{I_k(\theta_a) - I_k(\theta_b)} = \frac{\int_{\theta_a}^{\theta_b} \mu_j(\theta) f_j(\theta) d\theta}{\int_{\theta_a}^{\theta_b} (1 - F_k(\theta)) d\theta},$$
$$\forall j, k \in \{L, H\}, \quad 0 \leq \theta_a \leq \theta_b \leq \bar{\theta}.$$

The average profit-to-rent ratio measures the changes in the seller's profit in terms of the information rents he gives away to the consumer due to a change in price. The ratio R^{jk} compares the impact on profit from type j with the increase in the information rent to type k. This cross ratio arises because the incentive compatibility constraint for type k implies that a modification in the contract for type j also affects type k. This was clear from our discussion regarding the interior variation above. There, a price θ_1 (smaller than $\widehat{\theta}$) for type L creates a profit improvement for the seller measured by the numerator of R. Since the seller has to ensure that type H does not take the type-L contract (by reducing quantity), this price decrease generates a loss for the seller quantified by the denominator of R.

Returning to (6), we note that the numerator in either ratio refers to the revenue that the seller makes from the low type over some interval, and the denominator refers to the information rent of the high type over the same interval. Now, since the choice of θ_1, θ_2 was arbitrary, we obtain the following necessary condition by taking the minimum and maximum on both sides of the inequality in (6). If the static contract is optimal, then

$$\max_{\theta_1 \le \widehat{\theta}} R^{LH}(\theta_1, \widehat{\theta}) \le \min_{\widehat{\theta} \le \theta_2} R^{LH}(\widehat{\theta}, \theta_2). \tag{7}$$

The above condition establishes that if the static contract is optimal, then any extra revenue the seller can garner from low-type buyers is offset by the efficiency loss due to the incentive compatibility constraints: $A - B \le 0$ for any possible choice of θ_1 and θ_2.

To prove sufficiency in Theorem 1, we rely on a dualization-type of argument. For the necessity, we assume that condition (7) is not satisfied and then show that there is a profitable deviation as established by the following proposition.

Proposition 1 *(Revenue Improvement). Suppose that $\mu_L(\theta) f_L(\theta)/(1 - F_L(\theta))$ is nondecreasing. Assume that condition (7) does not hold. Then, there exists θ_1, θ_2 such that $\theta_1 < \widehat{\theta} < \theta_2$ and $R^{LH}(\theta_1, \widehat{\theta}) > R^{LH}(\widehat{\theta}, \theta_2)$, for which the allocation in (4) with*

$$x_L = \frac{\int_{\widehat{\theta}}^{\theta_2} \overline{F}_H(\theta) d\theta}{\int_{\theta_1}^{\theta_2} \overline{F}_H(\theta) d\theta},$$

yields a strict improvement in ($\mathcal{P}$) over the static contract.

In the proof of Proposition 1, we see that once condition (7) fails, two things happen. First, a non-static contract becomes feasible, which does not violate the incentive constraints. The mere fact that (7) fails implies the feasibility of the new allocation. Second, the sequential contract guarantees an expected revenue greater than the static revenue.

4.3. A necessary and sufficient condition

We now establish that condition (7) is in fact a sufficient condition for the optimal static solution to coincide with the optimal solution to ($\mathcal{P}$). Before we provide the main theorem, we introduce some notation for the quantities of interest that will help us to further refine our intuition. While we maintain the binary type framework here, we note that all definitions naturally extend to finitely many types as we will see in Section 6.1.

The local version of the average profit-to-rent ratio, when $\theta_a < \widehat{\theta} < \theta_b$ are close to $\widehat{\theta}$, gives rise to the *profit-to-rent ratio*.

Definition 2 *(Profit-to-Rent Ratio).* The profit-to-rent ratio between type j and k is defined by:

$$r^{jk}(\theta) \triangleq \frac{\mu_j(\theta) f_j(\theta)}{1 - F_k(\theta)}, \quad \forall j, k \in \{L, H\}, \forall \theta \in \Theta.$$

The ratio $r^{jk}(\theta_b)$ is obtained as $\lim_{\theta_a \uparrow \theta_b} R^{jk}(\theta_a, \theta_b)$. Observe that condition (IHR) implies that $r^{kk}(\theta)$ is nondecreasing for each $k \in \{L, H\}$. The latter is the condition we use for our formal results.

Now, we are ready to state and discuss the main result of this section.

Theorem 1 *(Optimality of Static Contract). Suppose that $r^{kk}(\theta)$ is nondecreasing for each $k \in \{L, H\}$. The static contract is optimal if and only if*

$$\max_{\theta \leq \widehat{\theta}} R^{LH}(\theta, \widehat{\theta}) \leq \min_{\widehat{\theta} \leq \theta} R^{LH}(\widehat{\theta}, \theta). \tag{APR}$$

This result complements the necessary condition given in Section 4.2 by showing that it is also sufficient. We showed in Section 4.2 that condition (APR) established that the specific deviation that increases the sales to the lower type with a lower quantity is not profitable relative to the static contract.

Theorem 1 now establishes that this in fact is not only a necessary but also a sufficient condition. The sufficient condition is noteworthy because it arises from "simple" deviations, namely, those that assign the low-type an interior allocation in a small interval around the static optimal price. In particular, we do not need to be concerned with either more elaborate deviations that offer the low type several options in her menu, nor do we need to trace simultaneous changes to the offers to the high type. The core of the sufficiency argument is that the nonprofitability of simple deviations from the static optimal contract is enough to establish optimality of the static contract. The present theorem confirms that this type of interior improvement for the low-type is sufficient to study changes in the seller's revenue. In Section 5, we establish that the family of allocations suggested by the interior variation also completely describes the optimal sequential mechanism.

In the Introduction, we noted that Krähmer and Kovac (2016) provided necessary and sufficient conditions for the optimality of a static contract (versus a sequential contract) in a delegation environment similar to Amador and Bagwell (2013). Their Proposition 3 established necessary and (almost) sufficient conditions by considering a ratio of virtual utilities similar to the ratio given by (6). While the exact shape of the virtual utility differs in the quasi-linear and the delegation environment, the logic of the argument is related.

4.4. The exponential example

Before we establish the optimal sequential contract, it might be helpful to build some intuition for the above results. We will consider the case of exponentially distributed values. The main result of this section establishes that the static contract is optimal if and only if the means of the interim types are sufficiently close.

We consider the exponential density functions

$$f_k(\theta) = \lambda_k e^{-\lambda_k \theta}, \quad k = \{L, H\} \quad \theta \geq 0.$$

We assume that $\lambda_L > \lambda_H$, where L and H stand for the low and high type, respectively. Note that H has a higher mean $(1/\lambda_H)$ than L $(1/\lambda_L)$ and that H dominates L in the sense of the hazard rate stochastic order and in first-order stochastic dominance. In addition, for the interim probabilities, we have $\alpha_L + \alpha_H = 1$ with $\alpha_L, \alpha_H > 0$.

We begin by studying the optimal solution to the static formulation. The optimal static contract is given by a threshold allocation. Thus, in the exponential case, the seller's expected revenue for any given threshold θ is

$$\Pi^{\text{static}}(\theta) \triangleq \int_{\theta}^{\infty} (\alpha_L \mu_L(z) f_L(z) + \alpha_H \mu_H(z) f_H(z)) dz = \alpha_L \theta e^{-\lambda_L \theta} + \alpha_H \theta e^{-\lambda_H \theta}.$$

To find the optimal threshold, we simply need to maximize the expression above. The first-order condition yields

$$\alpha_L(\theta - \frac{1}{\lambda_L})\lambda_L e^{-\lambda_L\theta} + \alpha_H(\theta - \frac{1}{\lambda_H})\lambda_H e^{-\lambda_H\theta} = 0. \tag{8}$$

That is, the optimal threshold is a zero of the mixture virtual value. Note that equation (8) cannot be explicitly solved; however, we can (as we do in the forthcoming results) provide comparative statics. Interestingly, in Proposition 4 below, we show that we can obtain explicit expressions for the thresholds characterizing the optimal sequential contract. The following lemma provides some initial properties of the optimal static contract.

Lemma 3. *The optimal solution to* $(\mathcal{P}^s)$ *is a threshold allocation characterized by* $\widehat{\theta}$ *in* $[\frac{1}{\lambda_L}, \frac{1}{\lambda_H}]$, *solving* (8). *Moreover,* $\widehat{\theta}$ *is a nonincreasing function of* α_L *with* $\widehat{\theta}(0) = \frac{1}{\lambda_H}$ *and* $\widehat{\theta}(1) = \frac{1}{\lambda_L}$.

Next, we state a necessary and sufficient condition for the static contract to be optimal.

Proposition 2 *(Necessity and Sufficiency for the Exponential Model). The static contract is optimal if and only if*

$$\lambda_L - \lambda_H \leq \frac{1}{\widehat{\theta}} \tag{9}$$

The result follows from Theorem 1. We note that the threshold value $\widehat{\theta}$ in the inequality is a solution to equation (8) and, therefore, depends on the parameters λ_L and λ_H. Subsequent corollaries provide sharper characterizations that depend solely on the model primitives. We highlight that (9) corresponds to a particular case of condition (APR).

Proposition 2 provides an intuitive characterization for when the seller is better off screening the interim types than not. In terms of equation (9), when λ_L and λ_H are sufficiently close, equation (9) should hold, in which case the static contract is optimal. Conversely, when λ_L and λ_H are sufficiently distant, the static contract will not be optimal.

Intuitively, when the interim types are similar, any contract that screens the types would be close in terms of expected revenue to the static contract because for each type it could obtain at most what it would obtain by setting thresholds $1/\lambda_L$ and $1/\lambda_H$, respectively, but $\widehat{\theta}$ belongs to $[\frac{1}{\lambda_L}, \frac{1}{\lambda_H}]$. However, when screening, the seller has to pay an extra cost to prevent the types from mimicking each other, and since the contracts' revenues will be similar, it is likely that this cost offsets the earnings from screening. On the other hand, when interim types are sufficiently distant in their mean value, the seller can tailor the contract to each type and in this way extract more from them than in the static contract.

Corollary 1 *(Optimality of Static Contract). If* $\lambda_L \in (\lambda_H, 2\lambda_H]$, *then for any* $\alpha_L \in [0, 1]$, *the static contract is optimal.*

This result establishes that when the distributions of the low- and high-type buyers are sufficiently close to each other, the static contract is always optimal, regardless of the proportion between types.

Corollary 2 *(Comparative Statics in α_L). If $\lambda_L > 2\lambda_H$, then there exists $\bar{\alpha} \in (0, 1)$ such that for all $\alpha_L \in (0, \bar{\alpha})$ the sequential contract is strictly optimal, and for all $\alpha_L \in [\bar{\alpha}, 1]$ the static contract is optimal.*

Corollary 2 asserts that when the means of low and high types are sufficiently distinct, the optimality of the static vs. the sequential contract is determined by the frequency of each type. If the proportion of low types is sufficiently low (but not zero), then the seller is better off screening the types. On the other hand, if there is a large proportion of low types, then the static contract is optimal. This follows because the threshold value $\widehat{\theta}$ decreases as α_L increases.

Corollary 3 *(Comparative Statics in λ_L). For fixed λ_H and α_H, there exists $\bar{\lambda}_L$ such that for all $\lambda_L \in (\bar{\lambda}_L, \infty)$, the sequential contract is strictly optimal.*

4.5. Discussion

We previously introduced the increasing hazard rate condition (IHR):

$$h^{kk}(\theta) = \frac{f_k(\theta)}{1 - F_k(\theta)} \quad \text{is increasing.}$$

Krähmer and Strausz (2015) introduced an expanded monotonicity condition that relates any pair of interim types to the hazard rate:

$$h^{jk}(\theta) = \frac{f_j(\theta)}{1 - F_k(\theta)} \quad \text{are increasing in } \theta, \quad \forall j, k \in \{L, H\}. \tag{R}$$

They show that under condition (R), the optimal solutions to ($\mathcal{P}$) and ($\mathcal{P}^s$) coincide; thus, the static contract is optimal. In fact, they show this result for multiple interim types. We discuss our generalization of condition (APR) to multiple types in Section 6.1. However, condition (R) is rather restrictive and not satisfied by some common distributions. For example, the condition is not satisfied by any pair of exponential distributions, because in this case, the cross-hazard rate is given by:

$$h^{jk}(\theta) = \lambda_j e^{-(\lambda_j - \lambda_k)\theta}, \quad j, k = L, H.$$

If, without loss of generality, we consider $\lambda_L > \lambda_H$, then $h^{LH}(\theta)$ is a decreasing function, and therefore, it violates condition (R). However, note that (IHR) is satisfied because the simple hazard rate functions are constant and equal to $1/\lambda_k$.

We can also compare Theorem 1 with Lemma 12 in Krähmer and Strausz (2014). In that Lemma, they assume that $h^{HH}(\theta) < h^{LL}(\theta)$, which implies $\widehat{\theta}_L < \widehat{\theta}_H$, and establish that a necessary condition for the static contract to be optimal is to have the profit-to-rent ratio $r^{LH}(\theta)$ being increasing at $\widehat{\theta}$. Our result contains this lemma because if $r^{LH}(\cdot)$ were decreasing at $\widehat{\theta}$, then we could always find $\theta_1 < \widehat{\theta}$ and $\theta_2 > \widehat{\theta}$ such that

$$R^{LH}(\theta_1, \widehat{\theta}) > R^{LH}(\widehat{\theta}, \theta_2).$$

Thus, (APR) does not hold, and therefore, the static contract would not be optimal. Fig. 3 illustrates how our condition (APR) closes the gap between those offered by Krähmer and Strausz (2015).

We can compare conditions (R) and (APR). Note that condition (R) implies the monotonicity of the profit-to-rent ratios, and therefore condition (APR) holds as

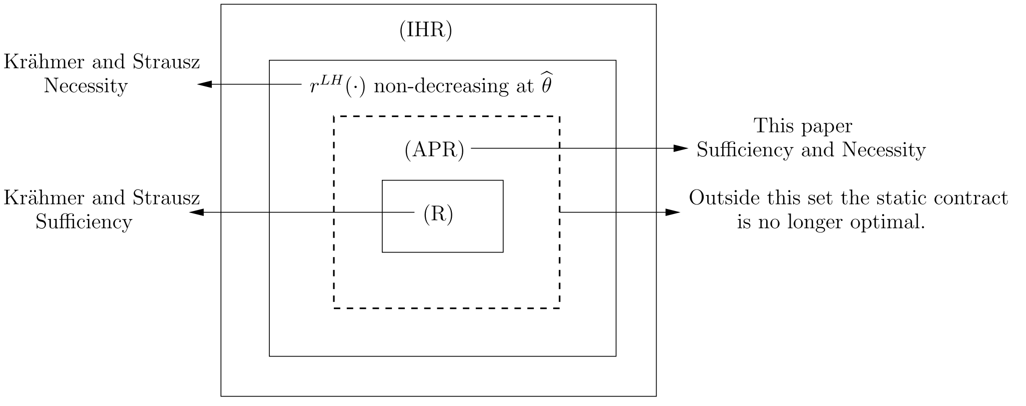

Fig. 3. Optimality of the static contract for (IHR) distributions, with $K = 2$ and a single buyer.

$$R^{LH}(\theta,\widehat{\theta}) = \frac{\int_{\theta}^{\widehat{\theta}} \overline{F}_H(z) r^{LH}(z)dz}{\int_{\theta}^{\widehat{\theta}} \overline{F}_H(z)dz} \leq r^{LH}(\widehat{\theta}), \quad \forall \theta \leq \widehat{\theta},$$

and

$$R^{LH}(\widehat{\theta},\theta) = \frac{\int_{\widehat{\theta}}^{\theta} \overline{F}_H(z) r^{LH}(z)dz}{\int_{\widehat{\theta}}^{\theta} \overline{F}_H(z)dz} \geq r^{LH}(\widehat{\theta}), \quad \forall \theta \geq \widehat{\theta}.$$

Hence, the result obtained by Krähmer and Strausz (2015) that if condition (R) holds then the static contract is optimal follows as a corollary of Theorem 1. We highlight that while condition (R) implies that the profit-to-rent ratios are increasing, our condition (APR) only implies the monotonicity of an appropriately weighted average of the profit-to-rent ratios. This is sensible because we are dealing with interim expected seller's revenues and interim incentive compatibility constraints.

In terms of methodology, our approach differs from that of Krähmer and Strausz (2015). Their approach consists of relaxing the low to high interim incentive constraint and then – by using their condition (R) – they relax the monotonicity constraint and prove that the solution must be a threshold schedule for each type. From there, they show that the threshold for the two types must be equal and, therefore, that the static contract is optimal.

In our approach, we do not use a relaxation of the general formulation or impose conditions on the primitives other than that the ratios $r^{kk}(\theta)$ are nondecreasing. For the sufficiency, we construct a Lagrangian relaxation with multipliers for the incentive compatibility constraints, but we do not relax the monotonicity constraints. The multipliers relate to the profit-to-rent ratios at the static threshold $\widehat{\theta}$; they measure the change in the objective per unit of change in the constraints. Then, by leveraging the result of Riley and Zeckhauser (1983) that an optimal contract is a threshold allocation, we prove that under (APR), the solution to the relaxation is the static contract. The multipliers have a natural structure: the low to high incentive constraint is slack, and for the high to low constraint, the change in the objective is given by the ratio of the seller's profit to the information rent of the high type. Once the multipliers are set, however, the key to the proof is to establish that condition (APR) delivers the optimality of the static contract.

5. Sequential contracts

We now proceed to provide the complete characterization of the optimal sequential contract when the necessary and sufficient condition associated with the static contract fails. As suggested in Section 4.2 and by Proposition 1, the optimal sequential contract provides a deterministic allocation to the high type and, for mid-range values, it randomizes the low-type buyer (or, equivalently, reduces the quantity allocated).

5.1. The structure of the sequential contract

We analyze the following relaxation of ($\mathcal{P}$)

$$\max_{0\leq \mathbf{x}\leq 1} \quad -\sum_{k\in\{L,H\}} \alpha_k u_k + \sum_{k\in\{L,H\}} \alpha_k \int_{\Theta} x_k(z)\mu_k(z) f_k(z)dz \tag{$\mathcal{P}_R$}$$

$$\begin{aligned} \text{s.t} \quad & x_k(\theta) \quad \text{nondecreasing}, \quad \forall k \in \{L, H\} \\ & u_k \geq 0, \forall k \in \{L, H\} \\ & u_H + \int_{\Theta} x_H(z)\overline{F}_H(z)dz \geq u_L + \int_{\Theta} x_L(z)\overline{F}_H(z)dz. \end{aligned}$$

The difference between ($\mathcal{P}_R$) and the original problem ($\mathcal{P}$) is the omission of the incentive constraint for the low-type to report truthfully. Importantly, we do not relax the monotonicity constraint. We obtain a characterization of the optimal solution to ($\mathcal{P}_R$) as stated by the following theorem.

Proposition 3 *(Relaxed Solution). Suppose that $r^{kk}(\theta)$ is nondecreasing for each $k \in \{L, H\}$. The optimal solution of ($\mathcal{P}_R$) has allocations*

$$x_L^\star(\theta) \triangleq \begin{cases} 0 & \text{if } \theta < \theta_1, \\ x_L & \text{if } \theta_1 \leq \theta \leq \theta_2, \\ 1 & \text{if } \theta_2 < \theta; \end{cases} \qquad x_H^\star(\theta) \triangleq \begin{cases} 0 & \text{if } \theta < \theta_H, \\ 1 & \text{if } \theta_H \leq \theta, \end{cases}$$

for some threshold values $\theta_1, \theta_H, \theta_2$ satisfying $\widehat{\theta}_L \leq \theta_1 \leq \theta_H \leq \theta_2$, $\theta_H \leq \widehat{\theta}_H$ and

$$x_L \triangleq \frac{\int_{\theta_H}^{\theta_2} \overline{F}_H(z)dz}{\int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz}.$$

Note that if $\theta_1 = \theta_H$, we would recover the static contract. Importantly, the optimal contract of ($\mathcal{P}_R$) has the same structure as the profitable deviation to the static contract presented in Proposition 1. The only difference is that in the former, the threshold for the high type may not necessarily be equal to $\widehat{\theta}$ as in the latter. With this generalization, one can show that the proposed profitable deviation is indeed optimal for ($\mathcal{P}_R$). The associated transfers are given by:

$$t_L^\star(\theta) = \begin{cases} 0 & \text{if } \theta < \theta_1, \\ \theta_1 \cdot x_L & \text{if } \theta_1 \leq \theta \leq \theta_2, \\ \theta_2 - (\theta_2 - \theta_1) \cdot x_L & \text{if } \theta_2 < \theta; \end{cases} \qquad t_H^\star(\theta) = \begin{cases} 0 & \text{if } \theta < \theta_H, \\ \theta_H & \text{if } \theta_H \leq \theta. \end{cases}$$

We use an argument based on infinite dimensional linear programming (which may be of more general interest by itself) to show that the extreme points of $(\mathcal{P}_R)$ are step functions with at most one randomization step. We then use an improvement argument to show that the optimal contract of $(\mathcal{P}_R)$ only requires a simple threshold allocation without randomization for the high type.[8]

Further, consider a low-type allocation that randomizes within an interval $[\theta_a, \theta_b]$. Recall the argument in Section 4.3, where we found a revenue improvement while maintaining feasibility, in particular, while maintaining the incentive constraint of the high type. Using a similar reasoning, we can show that feasibly improving upon the random allocation requires the following condition to hold for some $\tilde{\theta}$:

$$R^{LH}(\theta_a, \tilde{\theta}) = \frac{\int_{\theta_a}^{\tilde{\theta}} \overline{F}_H(z) r^{LH}(z)dz}{\int_{\theta_a}^{\tilde{\theta}} \overline{F}_H(z)dz} \leq \frac{\int_{\tilde{\theta}}^{\theta_b} \overline{F}_H(z) r^{LH}(z)dz}{\int_{\tilde{\theta}}^{\theta_b} \overline{F}_H(z)dz} = R^{LH}(\tilde{\theta}, \theta_b). \tag{10}$$

In general, this condition is not satisfied because the profit-to-rent ratio $r^{LH}(\cdot)$ does not need to be a nondecreasing function. Therefore, we cannot find a feasible improvement over the random allocation contract, and hence, we cannot restrict attention to deterministic contracts for the low-type. In contrast, a similar argument for the high type yields the expression $R^{HH}(\theta_a, \tilde{\theta}) \leq R^{HH}(\tilde{\theta}, \theta_b)$, which always holds when $r^{HH}(\cdot)$ is nondecreasing. Hence, we can restrict attention to a deterministic threshold contract for the high type.

The discussion above again highlights the importance of the average profit-to-rent ratios in our analysis, as they quantify revenue improvements while maintaining incentive compatibility. We can now characterize the optimal sequential contract.

Theorem 2 *(Optimal Sequential Contract). Suppose that $r^{kk}(\theta)$ is nondecreasing for each $k \in \{L, H\}$. The optimal sequential contract coincides with the optimal solution of $(\mathcal{P}_R)$ as given by Proposition 3.*

In Proposition 3, we provided the characterization of the optimal solution to $(\mathcal{P}_R)$. In the proof of Theorem 2, we argue that the optimal solution to $(\mathcal{P}_R)$ is feasible for $(\mathcal{P})$ and thus optimal. In turn, we obtain a full characterization of the optimal sequential contract in terms of three parameters ($(\theta_1, \theta_2, \theta_H)$ that we characterize in Lemma B.1 in the Appendix). We note that the proof of the theorem relies on the structure of x and the thresholds derived in Proposition 3. In turn, we do not exploit any single crossing-like property (e.g., stochastic order) but solely the monotonicity of $r^{kk}(\theta)$.

The sequential contract makes the low-type worse off and the high type better off with respect to the contract the seller would offer if he could perfectly screen each type. For the low-type, that contract would set a threshold equal to $\widehat{\theta}_L$ and would always allocate the object when her value is above the threshold. However, the sequential contract allocates the object to the low-type whenever her value is above $\theta_1 \geq \widehat{\theta}_L$ with positive probability. Therefore, the low type is worse off in two dimensions: she is allocated the object less often and with less probability. On the other hand, the high type receives the object more often and with certainty since $\theta_H \leq \widehat{\theta}_H$. A comparison of the thresholds of the optimal static contract with those of the optimal screening contract is more subtle because the optimal static contract may display nonmonotone behavior in the primitives. In the next section, we elaborate more on this issue (cf. Fig. 4).

[8] We thank an anonymous referee for a valuable suggestion regarding the proof technique.

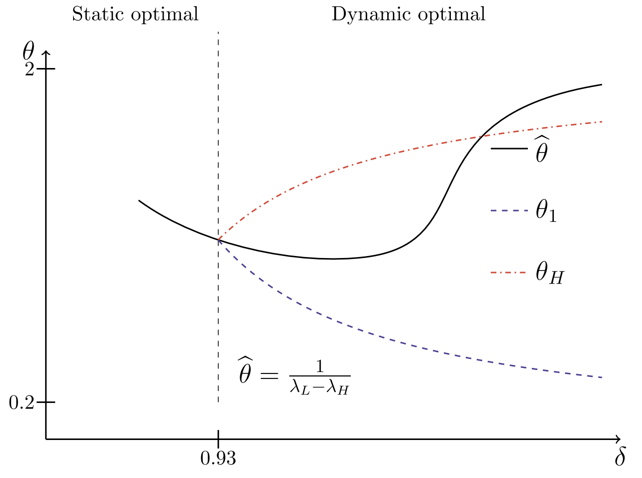

Fig. 4. Optimal thresholds for static and sequential contracts when setting $\lambda_L = \lambda_H + \delta$, with $\alpha_L = 0.7$ and $\lambda_H = 0.5$.

5.2. *The exponential example continued*

In Section 4.4, we studied the properties and structure of the optimal static contract for exponential values. We now derive the optimal sequential contract for this environment.

Proposition 4 *(Optimal Sequential Contract for Exponential Distributions). If condition (9) fails, then the optimal allocation is:*

$$x_L^\star(\theta) = \begin{cases} 0 & \text{if } \theta < \theta_1, \\ x & \text{if } \theta_1 \leq \theta; \end{cases} \quad \text{and} \quad x_H^\star(\theta) = \begin{cases} 0 & \text{if } \theta < \theta_H, \\ 1 & \text{if } \theta_H \leq \theta. \end{cases}$$

The thresholds are given by:

$$\theta_1 = \frac{1}{\lambda_L - \lambda_H} \quad \text{and} \quad \theta_H = \frac{1}{\lambda_H} - \frac{\alpha_L}{\alpha_H} \frac{e^{-1}}{\lambda_L - \lambda_H},$$

with $\theta_1 \leq \theta_H$. The probability of receiving the object for the low-type is:

$$x = \exp\Big(-\lambda_H \Big[\frac{1}{\lambda_H} - \frac{\alpha_L}{\alpha_H} \frac{e^{-1}}{\lambda_L - \lambda_H} - \frac{1}{\lambda_L - \lambda_H} \Big] \Big). \tag{11}$$

This result follows from Theorem 2. We note that in the exponential case, we only have two intervals for the low type's allocation, and thus $\theta_2 = \infty$. That is, the low-type is uniformly restricted to a quantity below one for all realized values $\theta \geq \theta_1$.

We now illustrate our findings below and vary the difference in the mean between the low and high type. Specifically, we fix α_L to be 0.7 and λ_H to be 0.5, that is, the high type has mean 2. Since we are assuming $\lambda_L > \lambda_H$, we consider $\lambda_L = \lambda_H + \delta$ with $\delta > 0$. Fig. 4 shows how the different thresholds vary as δ increases or, equivalently, as the mean of the low-type decreases to zero. As we can see, there is a value of δ (δ =0.93) to the left of which the static contract is optimal, and to its right, the sequential contract is optimal. As suggested by Proposition 2, as δ increases, $(\lambda_L - \lambda_H)$ increases, and therefore, we expect it to be larger than $1/\widehat{\theta}$ (see Corollary 2

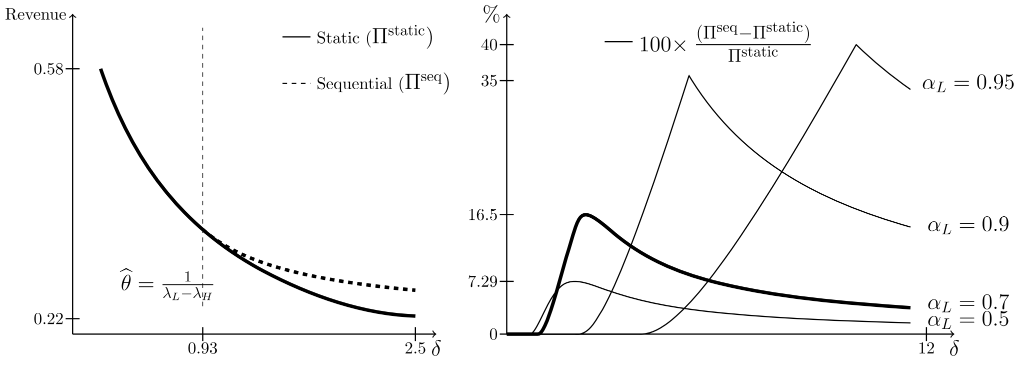

Fig. 5. *Left:* Optimal expected revenue for static and sequential. *Right:* Percentage improvement of the sequential over the static contract. In both figures we set $\lambda_L = \lambda_H + \delta$ with $\lambda_H = 0.5$. In the *left* figure, we set $\alpha_L = 0.7$, while in the *right* figure, α_L takes values in $\{0.5, 0.7, 0.9, 0.95\}$.

and Corollary 3). As δ increases, the two distributions become more distant from each other, and there is a gain from screening the types.

In terms of thresholds, we observe that for the static contract, $\widehat{\theta}$ is initially decreasing and then it increases getting closer to $1/\lambda_H = 2$. This happens because as we increase δ, we are making $1/\lambda_L$ smaller. However, at some point, this value becomes too small, and therefore, the probability of allocating the object to a low type, $P(\text{value low-type} > \widehat{\theta}) = e^{-\lambda_L \widehat{\theta}}$, will be so low that the seller would be better off by choosing a threshold tailored for the high type, that is, close to $1/\lambda_H = 2$. For the sequential thresholds, the threshold for the low type is decreasing while that for the high type is increasing in δ. As δ increases, the distributions become more different, and therefore, it is optimal to set thresholds closer and closer to the threshold that a seller would set if he knew the types in advance, that is, $1/\lambda_L$ and $1/\lambda_H$.

We can also compare the different mechanisms in terms of the resulting revenue. The optimal revenue for the sequential contract Π^{seq} is given by:

$$\Pi^{\text{seq}} = \alpha_L \cdot x \cdot \theta_1 \cdot e^{-\lambda_L \theta_1} + \alpha_H \cdot \theta_H \cdot e^{-\lambda_H \theta_H}.$$

Then, we can plot the different revenues as we vary δ. Fig. 5 (left panel and thick line in right panel) depicts the results. When α_L is large, the static threshold $\widehat{\theta}$ is tailored to the low types, so (9) holds for more values of λ_L. As screening occurs when the mean of the low type is sufficiently small, and thus δ is large, the revenue improvement due to sequential contracts becomes more significant and is above 40% when $\alpha_L = 0.95$. In recent work, Bergemann et al. (2020) compare the revenue of the optimal third-degree price discrimination policy against a uniform pricing policy. The optimal sequential screening policy is upper bounded by the third-degree pricing policy. As a corollary, they establish that the sequential screening policy can yield at most twice the revenue of the uniform pricing policy under some regularity conditions (Corollary 3.3). By means of an example, one can show that the bound can be attained.

5.3. Menu implementation

Next, we discuss how the optimal sequential contract can be implemented in practice. By means of the taxation principle, we can verify that the following menu of contracts is an indirect implementation of our optimal mechanism:

- contract H: there is a single posted price of $p_H = \theta_H$;
- contract L: the buyer can choose between two items:
  (a) buy at a price of $p_L = \theta_1 \cdot x_L$ and be allocated with probability x_L.
  (b) buy at a price of $p_L = \theta_2 - (\theta_2 - \theta_1) \cdot x_L$ and be allocated with probability 1.

The prices in the above menu of contracts are set using the values in Proposition 3. This implementation offers a posted price to the high-type buyer and offers the low-type buyer two options. In option (a), the low-type buyer can pay a low price, but this carries the possibility of not acquiring the item or, equivalently, obtaining a reduced quantity; in (b), the low-type buyer pays a high price and always obtains the object.

An appealing feature of the implementation is that if we regard allocations as quantities, then we can order the per unit prices. In contract L, the per unit prices are θ_1 and $\theta_1 \cdot x_L + \theta_2 \cdot (1 - x_L)$ for (a) and (b), respectively. Hence, the per unit price in (a) is less than or equal to that in (b). That is, the low type in (a) receives less of the good but at a discounted price compared to the low type in (b). For contract H, the per unit price is θ_H, and since θ_1 is less than or equal to θ_H, the low type in (a) also receives less of the good at a discounted price compared to the high-type buyer.

6. Extensions

In this section, we consider three extensions to our base model. First, we consider the case of multiple interim types. Then, for two interim types, we study both a setting with weaker ex post IR constraints and a three-stage setting.

6.1. Multiple types

Thus far, we have studied the optimality of the static and sequential contract for two interim types. In this section, we extend the analysis to an arbitrary number of interim types $\{1, \ldots, K\}$ and investigate some properties of the solution to ($\mathcal{P}$). In particular, we provide a generalized version of condition (APR). Then, we provide numerical evidence and highlight the challenges associated with the characterization of the optimal sequential mechanism when $K > 2$.

6.1.1. A necessary and sufficient condition

Our generalized necessary and sufficient condition continues to rely on small variations in the objective around the static solution. To this end, we consider the following set:

$$\mathcal{A} \triangleq \Big\{ (\lambda_{ij})_{i,j \in \{1,\cdots,K\}^2} \geq 0 : \sum_{j \neq k} \lambda_{jk} \cdot \overline{F}_j(\widehat{\theta}) = \alpha_k \cdot \mu_k(\widehat{\theta}) \cdot f_k(\widehat{\theta}) + \overline{F}_k(\widehat{\theta}) \cdot \sum_{j \neq k} \lambda_{kj},$$
$$\alpha_k \geq \sum_{j \neq k} \lambda_{kj} - \sum_{j \neq k} \lambda_{jk}, \quad \forall k \in \{1, \ldots, K\} \Big\}.$$

The set $\mathcal{A}$ contains the multipliers associated with the incentive constraints that encode the change in the objective as we deviate from the optimal static allocation. Roughly speaking, when the static contract is optimal, allocation perturbations in the contract of each type should equal the dualized costs associated with such perturbations in the incentive constraints. In other words, the derivative of the Lagrangian with respect to the posted price around the static solution equals zero. This is captured by the set of equalities in the definition of $\mathcal{A}$. In addition, the set of inequalities ensures that the optimal ex post utilities of the lowest value buyers are zero. Note that multipliers being in the set $\mathcal{A}$ is a necessary condition for optimality. The next result provides a necessary and sufficient condition.

Theorem 3 *(Necessary and Sufficient Conditions for Finitely Many Types). The set $\mathcal{A}$ is nonempty. If there exists a feasible solution to ($\mathcal{P}$) that strictly satisfies all the incentive constraints, then the static contract is optimal if and only if there exist $(\lambda_{ij})_{i,j\in\{1,\cdots,K\}^2} \in \mathcal{A}$ such that:*

$$\max_{\theta\le\widehat{\theta}} \left\{ \alpha_k \cdot R^{kk}(\theta,\widehat{\theta}) - \sum_{j\neq k} \lambda_{jk} \cdot \frac{\int_\theta^{\widehat{\theta}} \overline{F}_j(z)dz}{\int_\theta^{\widehat{\theta}} \overline{F}_k(z)dz} \right\}$$
$$\le \min_{\widehat{\theta}\le\theta} \left\{ \alpha_k \cdot R^{kk}(\widehat{\theta},\theta) - \sum_{j\neq k} \lambda_{jk} \cdot \frac{\int_{\widehat{\theta}}^{\theta} \overline{F}_j(z)dz}{\int_{\widehat{\theta}}^{\theta} \overline{F}_k(z)dz} \right\}, \tag{APR^M}$$

for all $k \in \{1,\ldots,K\}$.

The strict feasibility for ($\mathcal{P}$) corresponds to the standard Slater condition. Condition (APR^M) is obtained by analyzing the Lagrangian when the static contract is optimal and disentangling the key conditions it must satisfy. To do, so we consider simple threshold deviations from the static contract and study their impact on the Lagrangian. We note that this condition is easy to verify – it amounts to minimizing a convex program. Indeed, both sides of the inequality in (APR^M) correspond to convex (left) and concave (right) functions of λ. Their difference, left side minus right side, is thus a convex function. Moreover, because we can always choose θ equal to $\widehat{\theta}$, this difference is always bounded below by zero. Condition (APR^M) establishes that we can find λ such that this convex function equals zero; that is, its minimum value equals zero. This can be readily verified by using, for example, a subgradient-type method.

To obtain a better understanding of this condition, it is helpful to see how it generalizes the necessary and sufficient condition provided in Theorem 1 for two types. The general condition of Theorem 3 turns in the binary case for the low type (type 1):

$$\max_{\theta\le\widehat{\theta}} \left\{ \alpha_1 \cdot R^{11}(\theta,\widehat{\theta}) - \lambda_{21} \cdot \frac{\int_\theta^{\widehat{\theta}} \overline{F}_2(z)dz}{\int_\theta^{\widehat{\theta}} \overline{F}_1(z)dz} \right\} \le \min_{\widehat{\theta}\le\theta} \left\{ \alpha_1 \cdot R^{11}(\widehat{\theta},\theta) - \lambda_{21} \cdot \frac{\int_{\widehat{\theta}}^{\theta} \overline{F}_2(z)dz}{\int_{\widehat{\theta}}^{\theta} \overline{F}_1(z)dz} \right\}, \tag{12}$$

and for the high type (type 2):

$$\max_{\theta\le\widehat{\theta}} \left\{ \alpha_2 \cdot R^{22}(\theta,\widehat{\theta}) - \lambda_{12} \cdot \frac{\int_\theta^{\widehat{\theta}} \overline{F}_1(z)dz}{\int_\theta^{\widehat{\theta}} \overline{F}_2(z)dz} \right\} \le \min_{\widehat{\theta}\le\theta} \left\{ \alpha_2 \cdot R^{22}(\widehat{\theta},\theta) - \lambda_{12} \cdot \frac{\int_{\widehat{\theta}}^{\theta} \overline{F}_1(z)dz}{\int_{\widehat{\theta}}^{\theta} \overline{F}_2(z)dz} \right\}, \tag{13}$$

where λ_{12} and λ_{21} belong to $\mathcal{A}$. We next argue that condition (APR) holds if and only if there exists $\lambda_{12}, \lambda_{21} \in \mathcal{A}$ such that conditions (12) and (13) hold. Suppose that (APR) holds. Since we expect the incentive constraint of the low type not to be binding, we set λ_{12} equal to zero. Because λ must belong to $\mathcal{A}$, this necessarily implies that λ_{21} is equal to $\alpha_1 r^{12}(\widehat{\theta})$. For this choice of multipliers, the inequality (13) follows directly from the fact that r^{kk} is increasing. Moreover, the choice of multipliers, together with (APR), implies that both the maximum and the minimum in (12) are equal to zero. To see this consider the maximum in (12) and take $\theta = \widehat{\theta}$; since λ_{21} is equal to $\alpha_1 r^{12}(\widehat{\theta})$, the expression inside the brackets is zero. Hence, the maximum in (12) is bounded below by zero. It is also bounded above by zero:

$$\alpha_1 \cdot R^{11}(\theta,\widehat{\theta}) - \lambda_{21} \cdot \frac{\int_\theta^{\widehat{\theta}} \overline{F}_2(z)dz}{\int_\theta^{\widehat{\theta}} \overline{F}_1(z)dz} \le 0 \Leftrightarrow R^{12}(\theta,\widehat{\theta}) \le r^{12}(\widehat{\theta}), \quad \forall \theta \le \widehat{\theta}.$$

When (APR) holds, the right-hand side inequality always holds. A similar argument applies to the minimum. Therefore, the condition provided in Theorem 1 implies (APR^M) for the binary case. The converse implication follows from a contradiction argument, which we omit for the sake of brevity.

The two-type case is amenable to this simplification because one can readily solve for the multipliers: λ_{12} equal to zero is a natural choice, and $\lambda_{21} = \alpha_1 r^{12}(\widehat{\theta})$ then follows from the definition of $\mathcal{A}$. Unfortunately, when $K > 2$, the space of deviations is richer, and so is the possible selection of multipliers. In turn, this precludes a transparent characterization as in the two-type case.

An appealing feature of (APR^M) is that it provides a practical and simple way to verify that for a range of distributions, the static contract is optimal, as shown in the following result.

Proposition 5 *(Alternative Sufficient Conditions). Under the Slater condition of Theorem 3 and when either*

(i) condition (R) *holds or*
(ii) $z \cdot f_k(z)$ is nondecreasing for all k,

the static contract is optimal.

In the proposition above, we show that either *(i)* or *(ii)* implies condition (APR^M) and, consequently, the optimality of the static contract (cf. Theorem 3). Roughly speaking, in the proof of the proposition, we show that under *(i)* or *(ii)*, for all types, an appropriate function is nondecreasing. This function relates to the integrand in the numerator of the expression inside the maximum and minimum in (APR^M). In turn, by leveraging this monotonicity property, we establish that the maximum equals the minimum in (APR^M).

The conditions in Proposition 5 are very different in nature. Condition *(i)* is the same property under which Krähmer and Strausz (2015) prove the optimality of the static contract (here, we provide an alternative proof). This is a "cross" condition, in the sense that it links the distribution of different interim types. It is satisfied when the density of each type is increasing, for example, for natural families of distributions such as $f_k(z) = z^{\beta_k}$ for some $\beta_k > 1$ and $z \in [0,1]$. Condition *(ii)* does not associate the distributions of different types—it is not a cross condition. This property is satisfied by some truncated heavy-tailed distribution, for example, the log-normal distribution truncated between zero and the exponential of the mean of its logarithmic value.

Theorem 3 provides a simple, easy-to-verify set of inequalities for the optimality of the static contract with multiple types. By contrast, a complete characterization of the sequential contract seems substantially more complex with finitely many types. Next, in the context of exponentially distributed ex post types, we briefly describe partial results and highlight the challenges associated with multiple types that already appear in the numerical analysis.

6.1.2. The exponential example continued

Despite the challenges that we discussed above, we are able to provide the following result for the exponential environment.

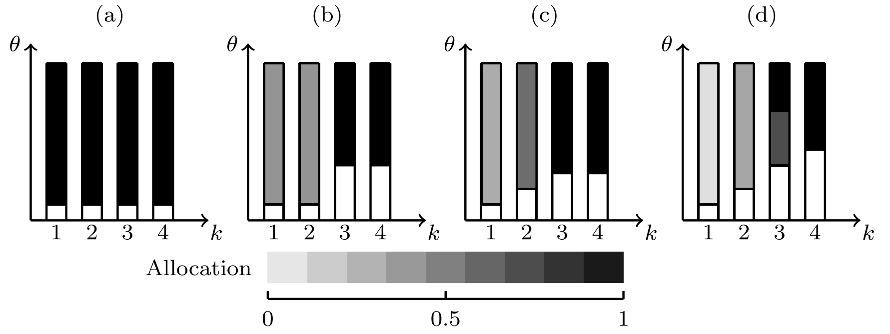

Fig. 6. Optimal allocations for $K = 4$; types have exponential distributions with means $(2.2, 5.0, 12, 50)$ (for numerical simplicity, we use truncated versions of these distributions in the interval [0,60]). In each panel the vertical axis corresponds to buyers' valuations, and the horizontal axis corresponds to the interim type. Each bar represents the allocation for each type; lighter gray indicates lower probability of allocation, while darker gray indicates a higher probability of allocation. White represents no allocation and black full allocation. From panels (a) to (d), the fractions, α_k, for each type are $(0.7, 0.2, 0.05, 0.05)$, $(0.4, 0.1, 0.4, 0.1)$, $(0.3, 0.2, 0.4, 0.1)$ and $(0.25, 0.25, 0.1, 0.4)$, respectively.

Proposition 6 *(Structure of Sequential Contract with Exponential Distributions). For exponential values, the optimal allocations have at most one randomized interval.*

Proposition 6 establishes that for exponentially distributed values, the optimal contract is simple in the sense that each interim type allocation is randomized in at most one interval. The proof proceeds by establishing that the monotonicity constraints form a cone, using duality and complementary slackness. It is worth mentioning that the proof method applies more generally, but the structure of the contract in general depends on the values of the dual variables corresponding to the incentive constraints. In the exponential case, the argument can be simplified to show that the simple structure in the result arises independent of these variables' values.

The characterization in Proposition 6 only establishes the structure of the optimal allocations; it does not provide information on the number of contracts that the optimal solutions will ultimately feature. For example, if $K = 4$, Proposition 6 does not say whether the optimal solution will pool the interim types to create either one, two, three or four different contracts. In general, the full range of contracts from static to fully sequential (K different contracts) is possible.

To further explore the structure of optimal contracts, we provide numerical results. In Fig. 6, we depict the optimal allocations when $K = 4$ and all interim types have exponentially distributed values. A first observation is that for different proportions α_k of interim types, the optimal contract can feature different levels of separation. Panel (a) of the figure corresponds to an optimal static contract (no separation), and panel (d) corresponds to an optimal sequential contract that features a different contract for each interim type (full separation). As a second observation, note that of the four instances depicted in Fig. 6, only one, (d), has four contracts in the optimal solution. Finding the minimal number of contracts that provides a good approximation of the optimal multiple-type sequential contract is a question beyond the scope of this paper but may be of interest to study in the future.

Observe that across the instances in Fig. 6, each optimal contract has at most one interval of value for which randomization occurs (see Proposition 6). This simple structure of the optimal contract does not appear robust to other specifications of the value distributions. When we consider the case of normally distributed values (using truncated normal random variables), the optimal contract might exhibit several different intervals of randomization for a given type. In general, richer contract features may arise when we combine exponential, normal, uniform or other distributions. As a consequence, generally speaking, it is challenging to analytically characterize the optimal solution. The challenge here is that classic relaxation approaches, used in the mechanism design literature, do not apply in our setting. For example, relaxing all the upward incentive constraints and leaving only the local downward incentive constraints does not work because, in general, global downward incentive constraints bind. Moreover, binding constraints are highly sensitive to model primitives. Improving our understanding of this setting may be an interesting avenue for future research.

6.2. Weaker ex post participation constraints

In this section, we generalize our base model and allow for less rigid participation constraints. Consider a scenario in which the seller can ask the buyer to pay a nonrefundable amount upon signing the contract. In this case, the contract must guarantee that the interim utility of the buyer is nonnegative, but the ex post utility can be negative. Effectively, we are relaxing the ex post participation constraints. Krähmer and Strausz (2015) refer to this type of contract as bonds because it is as if the buyer pays a costly bond just before signing the contract. In this setting we can prove, using a similar argument to Theorem 1, that if the nonrefundable payment is not too large, then our necessary and sufficient condition remains valid.

Proposition 7. *Let $B > 0$. Suppose that the buyer's ex post utility must be greater or equal than $-B$ and that her interim utility is nonnegative. If $\min_{k\in\{L,H\}} \int_{\hat{\theta}}^{\bar{\theta}} \overline{F}_k(z)dz \geq B$, then the static contract is optimal if and only if condition (APR) is satisfied.*

In the proposition, we consider the following participation constraints

$$u_k \geq -B \quad \text{and} \quad u_k + \int_0^{\bar{\theta}} x_k(z)\overline{F}_k(z)dz \geq 0, \quad \forall k \in \{L, H\}. \tag{14}$$

The proposition establishes that in this setting, when B is not too large, (APR) is still a necessary and sufficient condition for the optimality of the static contract. Krähmer and Strausz (2015) prove a related result that establishes the optimality of the static contract when B is small enough and condition (R) is satisfied.[9]

6.3. A three-stage model

As an extension of our base model, we also study a simple multi-stage setting in which buyers learn progressive information about their valuations over time. In particular, we show that from an initial condition in which the seller offers a static contract, as more information becomes available to the buyers over time and the types become more separated, the seller may wait for this to sequentially screen buyers.

[9] Interestingly, one can also show that in the case $\int_{\hat{\theta}}^{\bar{\theta}} \overline{F}_L(z)dz < B < \int_0^{\bar{\theta}} \overline{F}_L(z)dz$, the optimal static contract may exhibit randomization despite the absence of the interim IC constraints, but due to the presence of the interim IR constraint. We omit the details for brevity.

Consider the following three-stage model. In the first stage, the buyer possesses imperfect information about her type. In the second stage, the buyer learns precisely whether her type is low or high. Later, in the third stage, the buyer learns her valuation. More precisely, in the first stage, the buyer knows that her distribution is the mixture

$$\beta F_L(\cdot) + (1-\beta) F_H(\cdot), \tag{15}$$

where $\beta \in [0, 1]$ is also known by the seller. In the second stage, the buyer learns her type, and from the seller's perspective, there is a probability α_L or $(1-\alpha_L)$ that the buyer is of the low or high type, respectively. That is, from the second stage on, the situation is exactly the same as in our original model.

The seller can either decide to sell the item in the first stage or wait until the second stage. Any contract the seller designs must respect ex post participation constraints. In the first stage, neither the buyer nor the seller possesses private information about the buyer's valuation of the item—both know that it will be drawn from the mixture distribution in equation (15). In turn, the only contract that the seller can offer in the first stage is a static contract without screening. The optimal ex post IR static contract is a posted price against the mixture distribution. Now, the seller could also choose to wait and offer a contract in the second stage. In this case, the buyer gains information because she effectively knows her type while the seller only knows that the buyer is of the low type with probability α_L (and of high type with probability $(1-\alpha_L)$). The optimal contract in this case can be static or sequential depending on the parameters as characterized by condition (APR).

At this point, it is possible for us to assess whether the seller would prefer to offer a static contract in the first stage or to wait and screen in the second stage. Interestingly, it might be optimal for the seller to wait until the second stage despite that the buyer becomes more informed. Suppose that $\beta = \alpha_L$. In this case, the static contracts in the first and second stages coincide. As a result, if (APR) is not satisfied, waiting for the second stage to screen the buyer becomes optimal. In contrast, if (APR) holds, then there is no point in waiting, and offering the static contract in the first stage is optimal. From this, it follows that if α_L and β are different but close to each other, it might indeed be strictly optimal for the seller to wait until the second stage to screen the buyer.

7. Conclusion

We considered the scope of sequential screening in the presence of ex post participation constraints. The ex post participation constraints limit the ability of the seller to extract surplus from the buyer. As the buyer has to be willing to participate in the contractual arrangement following every realization of her value, the surplus has to be extracted ex post rather than at the interim level.

Despite these ex post restrictions, sequential screening frequently allows the seller to increase his revenue beyond the statically optimal revenue. The gains from sequential screening become more pronounced to the extent that the interim types differ in their willingness to pay. A natural implementation of the optimal mechanism simply offers the buyer the choice among different menus in the first stage. The choice of menu in the first period merely restricts the possible choices in the second period. In particular, it is not necessary to ask the buyer for any transfer before the final transaction occurs. Moreover, the buyer only has to make a transfer if she receives the object.

In contrast to the static solution where an optimal policy is always to sell the maximum quantity of 1, the sequential screening policy offers intermediate quantities. This departure from the bang-bang policy in a linear utility setting arises due to the presence of the ex post participation constraint, in conjunction with the incentive compatibility constraints.

There are several natural directions to extend the present work. Our stronger results were for the case of binary interim types while allowing for a continuum of values for each type. We also presented an extension of Theorem 1 to multiple types, as well as a characterization and numerical results for exponential values. We would like to further explore the characterization of the optimal sequential contract to multiple types and general value distributions. An interesting question here concerns the number of randomization intervals per type and whether the number of intermediate allocations increases with the number of interim types. Additionally, is there a fixed number of intermediate allocations that yield a good approximation to the optimal solution for an arbitrary number of interim types? Similarly, is there a fixed number of contracts that yield a good approximation to the optimal solution for an arbitrary number of interim types?

We might also be interested in analyzing how the number of competing buyers may affect the nature of the optimal mechanism. This has important practical consequences, particularly in industries that use market mechanisms such as auctions, for example, in the case of display advertising alluded to at the beginning of the paper. We note that this extension is not immediate because with multiple buyers, we may lose the threshold structure of the optimal static allocation. However, we conjecture that in this case, an approximately optimal market design would consist of running a series of “waterfall auctions” with different priorities across participants.

Appendix A. Proofs of main results

The appendix contains the proof of all results except for those related to the exponential distributions that are contained in the Appendix B.

Proof of Lemma 1. The proof of this result is standard and thus omitted. □

Proof of Lemma 2. The fact that the optimal solution is a threshold allocation is explained in the main text. Thus, we only need to provide a proof of $\widehat{\theta}$ being in the interval $[\widehat{\theta}_1, \widehat{\theta}_K]$; however, this is exactly Lemma 1 in Krähmer and Strausz (2014). □

Proof of Theorem 1. We first show the sufficiency of our condition and then its necessity. We denote by Ω the space of nondecreasing allocations, that is,

$$\Omega \triangleq \{x : [0, \bar{\theta}] \to [0, 1] : x(\cdot) \text{ is nondecreasing}\}.$$

Sufficiency. We assume that condition (APR) holds. We want to verify that the static contract is optimal. In order to do so we dualize the incentive constraints. The Lagrangian is

$$\begin{aligned}\mathcal{L}(\boldsymbol{u}, \mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{w}) = {} & u_L(w_L - \lambda_{HL} - \alpha_L) + u_H(\lambda_{HL} - \alpha_H + w_H) \\ & + \int_0^{\bar{\theta}} x_L(z) \cdot \Big[\alpha_L \mu_L(z) f_L(z) - \lambda_{HL} \overline{F}_H(z)\Big] dz \\ & + \int_0^{\bar{\theta}} x_H(z) \cdot \Big[\alpha_H \mu_H(z) f_H(z) + \lambda_{HL} \overline{F}_H(z)\Big] dz,\end{aligned}$$

where w_L, w_H correspond to the multipliers for the ex post IR constraints, and $\boldsymbol{\lambda} \in \{\lambda_{HL}, \lambda_{LH}\}$ to the multipliers for the incentive constraints. In the Lagrangian above we have chosen the multipliers as follows

$$w_L = \alpha_L - \alpha_H r^{HH}(\widehat{\theta}), w_H = \alpha_H + \alpha_H r^{HH}(\widehat{\theta}), \lambda_{HL} = \alpha_L r^{LH}(\widehat{\theta}), \lambda_{LH} = 0, \tag{A.1}$$

these multipliers are nonnegative because $r^{HH}(\widehat{\theta}) \leq 0$, $r^{LH}(\widehat{\theta}) \geq 0$ and

$$w_H = \alpha_H + \alpha_H r^{HH}(\widehat{\theta}) \geq 0 \Leftrightarrow r^{HH}(\widehat{\theta}) \geq -1 \Leftrightarrow [\widehat{\theta} - \frac{\overline{F}_H}{f_H}(\widehat{\theta})] \geq -\frac{\overline{F}_H}{f_H}(\widehat{\theta}) \Leftrightarrow \widehat{\theta} \geq 0.$$

Hence, maximizing the Lagrangian over nondecreasing allocation x_L and x_H yields an upper bound for the relaxed problem. Note that this choice of multipliers (together with equation (A.4) below) eliminates the u_L and u_H terms in the Lagrangian. We next show that under (APR) the solution to the Lagrangian relaxation is the static solution. We first claim that

$$\max_{x_L \in \Omega} \int_0^{\bar{\theta}} x_L(z) \cdot \Big[\alpha_L \mu_L(z) f_L(z) - \lambda_{HL} \overline{F}_H(z)\Big] dz = \int_{\widehat{\theta}}^{\bar{\theta}} \Big[\alpha_L \mu_L(z) f_L(z) - \lambda_{HL} \overline{F}_H(z)\Big] dz. \tag{A.2}$$

To prove this, first note that the optimal solution x_L on the left-hand side of (A.2) must be of the threshold type, that is, $x_L(\theta) = \mathbf{1}_{\{\theta \geq \theta^\star\}}$, because $x_L(\cdot)$ is nondecreasing (see, e.g., Myerson (1981) or Riley and Zeckhauser (1983)). Hence (A.2) is equivalent to

$$\int_{\theta^\star}^{\bar{\theta}} \Big[\alpha_L \mu_L(z) f_L(z) - \lambda_{HL} \overline{F}_H(z)\Big] dz$$

$$\leq \int_{\widehat{\theta}}^{\bar{\theta}} \Big[\alpha_L \mu_L(z) f_L(z) - \lambda_{HL} \overline{F}_H(z)\Big] dz, \quad \forall \theta^\star \in [0, 1].$$

Replacing the value of λ_{HL}, this equation can be cast over values $\theta_1^\star \leq \widehat{\theta}$ and $\theta_2^\star \geq \widehat{\theta}$ as

$$\frac{\int_{\theta_1^\star}^{\widehat{\theta}} \alpha_L \mu_L(z) f_L(z) dz}{\int_{\theta_1^\star}^{\widehat{\theta}} \overline{F}_H(z) dz} \leq \alpha_L r^{LH}(\widehat{\theta}) \leq \frac{\int_{\widehat{\theta}}^{\theta_2^\star} \alpha_L \mu_L(z) f_L(z) dz}{\int_{\widehat{\theta}}^{\theta_2^\star} \overline{F}_H(z) dz}, \quad \forall \theta_1^\star \leq \widehat{\theta} \leq \theta_2^\star \tag{A.3}$$

Condition (APR) ensures that the equation above always holds. Indeed, condition (APR) implies that for any $\theta_1^\star \leq \widehat{\theta}$ and $\epsilon > 0$

$$\frac{\int_{\theta_1^\star}^{\widehat{\theta}} \alpha_L \mu_L(z) f_L(z) dz}{\int_{\theta_1^\star}^{\widehat{\theta}} \overline{F}_H(z) dz} \leq \frac{\int_{\widehat{\theta}}^{\widehat{\theta}+\epsilon} \alpha_L \mu_L(z) f_L(z) dz}{\int_{\widehat{\theta}}^{\widehat{\theta}+\epsilon} \overline{F}_H(z) dz}.$$

Taking $\epsilon \downarrow 0$ yields the left-hand side inequality in (A.3). The right-hand side inequality in (A.3) can be verified using an analogous argument. This shows (A.2), that is, the static contract maximizes the part of the Lagrangian that corresponds to interim type L. We now prove the same for type H. Note first that the optimality of the static contract implies

$$\lambda_{HL} = \alpha_L r^{LH}(\widehat{\theta}) = -\alpha_H r^{HH}(\widehat{\theta}). \tag{A.4}$$

Then

$$\max_{x_H \in \Omega} \int_0^{\bar{\theta}} x_H(z) \cdot \left[\alpha_H \mu_H(z) f_H(z) + \lambda_{HL} \overline{F}_H(z)\right] dz$$

$$= \max_{x_H \in \Omega} \int_0^{\bar{\theta}} x_H(z) \cdot \alpha_H \cdot \left[\mu_H(z) f_H(z) - r^{HH}(\widehat{\theta}) \overline{F}_H(z)\right] dz$$

$$\overset{(a)}{=} \max_{x_H \in \Omega} \int_0^{\bar{\theta}} x_H(z) \cdot \alpha_H \cdot \left[r^{HH}(z) - r^{HH}(\widehat{\theta})\right] \overline{F}_H(z) dz$$

$$\overset{(b)}{=} \int_{\widehat{\theta}}^{\bar{\theta}} \alpha_H \cdot \left[r^{HH}(z) - r^{HH}(\widehat{\theta})\right] \overline{F}_H(z) dz$$

where in (a) we have used the definition of $r^{HH}(\cdot)$ and in (b) our assumption that $r^{HH}(\cdot)$ is increasing.

Since the value of the Lagrangian coincides with the primal objective at the static solution, and this solution is always primal feasible, we conclude that the static contract is optimal.

Necessity. We defer this proof to the proof of Proposition 1. In it, we show that whenever condition (APR) is not satisfied, there is a contract different from the static one with a strictly larger revenue. □

Proof of Proposition 1. Assume that (APR) does not hold; then, by Lemma A.1 (which we state and prove after the current proof) there exist $\theta_1 < \widehat{\theta} < \theta_2$ such that

$$\frac{\int_{\theta_1}^{\widehat{\theta}} \overline{F}_H(z) r^{LH}(z) dz}{\int_{\theta_1}^{\widehat{\theta}} \overline{F}_H(z) dz} > \frac{\int_{\widehat{\theta}}^{\theta_2} \overline{F}_H(z) r^{LH}(z) dz}{\int_{\widehat{\theta}}^{\theta_2} \overline{F}_H(z) dz}. \tag{A.5}$$

Consider a contract in which we set $u_L = u_H = 0$, and

$$x_L(\theta) = \begin{cases} 0 & \text{if } \theta < \theta_1 \\ x & \text{if } \theta_1 \leq \theta \leq \theta_2 \\ 1 & \text{if } \theta_2 < \theta, \end{cases} \qquad x_H(\theta) = \begin{cases} 0 & \text{if } \theta < \widehat{\theta} \\ 1 & \text{if } \widehat{\theta} \leq \theta, \end{cases}$$

where $x = \int_{\widehat{\theta}}^{\theta_2} \overline{F}_H(z) dz / \int_{\theta_1}^{\theta_2} \overline{F}_H(z) dz$. We next show that this solution is feasible and yields a strict revenue improvement over the static contract.

Feasibility. The ex post participation constraints are clearly satisfied. Additionally, since $\theta_1 < \widehat{\theta} < \theta_2$ we have $x_L \in (0, 1)$, and both $x_L(\cdot)$ and $x_H(\cdot)$ are nondecreasing allocations. We verify the incentive constraints

$$u_L + \int_0^{\bar{\theta}} x_L(\theta) \overline{F}_L(\theta) d\theta \geq u_H + \int_0^{\bar{\theta}} x_H(\theta) \overline{F}_L(\theta) d\theta,$$

$$u_H + \int_0^{\bar\theta} x_H(\theta)\overline{F}_H(\theta)d\theta \geq u_L + \int_0^{\bar\theta} x_L(\theta)\overline{F}_H(\theta)d\theta.$$

By replacing the allocations and ex post utilities we obtain that the incentive constraints are equivalent to

$$\frac{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_H(z)dz}{\int_{\theta_1}^{\theta_2}\overline{F}_H(z)dz} \geq \frac{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_L(z)dz}{\int_{\theta_1}^{\theta_2}\overline{F}_L(z)dz}. \tag{A.6}$$

To see why this is true, rewrite equation (A.5) as

$$\frac{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_H(z)dz}{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_H(z)dz} > \frac{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_H(z)r^{LH}(z)dz}{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_H(z)r^{LH}(z)dz}, \tag{A.7}$$

note that we are using here that by Lemma A.1 the denominator on the right-hand side is strictly positive. Moreover, note that

$$\begin{aligned}
\frac{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_H(z)r^{LH}(z)dz}{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_L(z)dz} &= \frac{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_L(z)r^{LL}(z)dz}{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_L(z)dz}\\
&\geq r^{LL}(\widehat{\theta})\frac{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_L(z)dz}{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_L(z)dz} = r^{LL}(\widehat{\theta})\frac{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_L(z)dz}{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_L(z)dz}\\
&\geq \frac{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_L(z)r^{LL}(z)dz}{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_L(z)dz}\\
&= \frac{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_H(z)r^{LH}(z)dz}{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_L(z)dz},
\end{aligned}$$

where the inequalities come from the fact that $r^{LL}(\cdot)$ is an increasing function and $r^{LL}(\widehat{\theta}) \geq 0$. This gives

$$\frac{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_H(z)r^{LH}(z)dz}{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_H(z)r^{LH}(z)dz} \geq \frac{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_L(z)dz}{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_L(z)dz},$$

note that we are using here that by Lemma A.1 the denominator on the left-hand side is strictly positive. This inequality together with (A.7) yields (A.6), and therefore, the proposed solution is feasible.

Revenue improvement. We need to prove that

$$\int_{\widehat{\theta}}^{\bar\theta}[\alpha_L f_L(z)\mu_L(z) + \alpha_H f_H(z)\mu_H(z)]dz < x\cdot\int_{\theta_1}^{\theta_2}\alpha_L f_L(z)\mu_L(z)dz + \int_{\theta_2}^{\bar\theta}\alpha_L f_L(z)\mu_L(z)dz$$
$$+\int_{\widehat{\theta}}^{\bar\theta}\alpha_H f_H(z)\mu_H(z)dz,$$

this is equivalent to

$$\int_{\widehat{\theta}}^{\theta_2} \alpha_L f_L(z)\mu_L(z)dz < \frac{\int_{\widehat{\theta}}^{\theta_2} \overline{F}_H(z)dz}{\int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz} \cdot \int_{\theta_1}^{\theta_2} \alpha_L f_L(z)\mu_L(z)dz$$

which is the same as

$$\frac{\int_{\widehat{\theta}}^{\theta_2} \overline{F}_H(z)r^{LH}(z)dz}{\int_{\widehat{\theta}}^{\theta_2} \overline{F}_H(z)dz} < \frac{\int_{\theta_1}^{\widehat{\theta}} \overline{F}_H(z)r^{LH}(z)dz}{\int_{\theta_1}^{\widehat{\theta}} \overline{F}_H(z)dz}$$

which is exactly the property satisfied by θ_1, θ_2 in (A.5). □

Lemma A.1. *Suppose that*

$$\max_{0 \le \theta \le \widehat{\theta}} R^{LH}(\theta, \widehat{\theta}) > \min_{\widehat{\theta} \le \theta \le \bar{\theta}} R^{LH}(\widehat{\theta}, \theta).$$

Then, there exist $\theta_a, \theta_b \in [0, \bar{\theta}]$ *with* $\theta_a < \widehat{\theta} < \theta_b$ *such that* $R^{LH}(\theta_a, \widehat{\theta}) > R^{LH}(\widehat{\theta}, \theta_b)$. *Moreover,* $0 < \int_{\theta_a}^{\widehat{\theta}} \overline{F}_H(z)r^{LH}(z)dz = \int_{\theta_a}^{\widehat{\theta}} \overline{F}_L(z)r^{LL}(z)dz$, *and* $0 < \int_{\widehat{\theta}}^{\theta_b} \overline{F}_H(z)r^{LH}(z)dz = \int_{\widehat{\theta}}^{\theta_b} \overline{F}_L(z)r^{LL}(z)dz$.

Proof of Lemma A.1. Note that both $R^{LH}(\cdot, \widehat{\theta})$ and $R^{LH}(\widehat{\theta}, \cdot)$ are continuous functions. Thus the maximum and the minimum in the statement are achieved by some $\tilde{\theta}_a \in [0, \widehat{\theta}]$ and $\tilde{\theta}_b \in [\widehat{\theta}, \bar{\theta}]$, respectively. Therefore, by assumption, we have that

$$R^{LH}(\tilde{\theta}_a, \widehat{\theta}) > R^{LH}(\widehat{\theta}, \tilde{\theta}_b).$$

Using the continuity of both functions, we can find $\theta_a < \widehat{\theta}$ and $\theta_b > \widehat{\theta}$ such that the inequality above is satisfied.

To finalize, we argue why $0 < \int_{\theta_a}^{\widehat{\theta}} \overline{F}_H(z)r^{LH}(z)dz$. Note that since $\theta_b > \widehat{\theta} \ge \widehat{\theta}_L$ (see Lemma 2) we have $R^{LH}(\widehat{\theta}, \theta_b) > 0$. Therefore, $R^{LH}(\theta_a, \widehat{\theta}) > 0$, which implies the desired inequalities. □

Extreme points. We next show that the extreme points in the feasible set of $(\mathcal{P}_R)$ are step functions with at most one intermediate step for the low and high type. We follow notation and definitions from Anderson and Nash (1987).

Let us define the convex cone

$$P \triangleq \{x(\theta) : [0, \bar{\theta}] \to \mathbb{R}_+ : x(\theta) \text{ is a nondecreasing function}\}.$$

We consider P to be a subset of X—the set of Lebesgue-measurable functions defined in $[0, \bar{\theta}]$ taking values in $\mathbb{R}_+$. Let the relation $\ge_P$ be defined by $y \ge_P x$ if and only if $y - x \in P$, for $x, y \in X$. We use 0_X to denote the null vector in X. Furthermore, define the linear functionals

$$A_1 : X \to \mathbb{R}, \quad x \mapsto \int_0^{\bar{\theta}} x(z)\overline{F}_H(z)dz,$$

$$A_2 : X \to \mathbb{R}, \quad x \mapsto x(\bar{\theta}).$$

Under this notation the feasible set in $(\mathcal{P}_R)$ is

$$x_L, x_H \in X, \quad x_L, x_H \geq_P 0_X, \ u_L, u_H \geq 0, \quad u_H + A_1 x_H \geq u_L + A_1 x_L,$$
$$A_2 x_k \leq 1, \ k \in \{L, H\}. \tag{A.8}$$

Note that to study the extreme points of the set above, we can simply focus on either x_L or x_H. For example, we can analyze the set $\{x \in X : x \geq_P 0_X, \ \ A_1 x \geq C, \ \ A_2 x \leq 1\}$ for some constant C. Indeed, note that

$$\exists x_L, x_H \in X : u_H + A_1 x_H \geq u_L + A_1 x_L$$
$$\iff \exists (t, x_L, x_H) \in \mathbb{R} \times X \times X : A_1 x_L + u_L \leq t, -A_1 x_H - u_H \leq -t.$$

In turn, we can fix t and u_L, u_H and obtain two decoupled problems for x_L and x_H for which the feasible sets are

$$\mathcal{F}_L \triangleq \{x \in X : x \geq_P 0_X, \ \ A_1 x \leq t - u_L, \ \ A_2 x \leq 1\} \quad \text{and}$$
$$\mathcal{F}_H \triangleq \{x \in X : x \geq_P 0_X, \ \ A_1 x \geq t - u_H, \ \ A_2 x \leq 1\},$$

respectively. From this, it follows that the extreme points in the feasible set of $(\mathcal{P}_R)$ correspond to the extreme points of $\mathcal{F}_L$ and $\mathcal{F}_H$. We have the following result.

Lemma A.2. *Fix t and u_L, u_H, if x is an extreme point of $\mathcal{F}_L$ or $\mathcal{F}_H$ then*

$$x(\theta) \triangleq \begin{cases} 0 & \text{if } \theta < \theta_1, \\ \chi & \text{if } \theta_1 \leq \theta \leq \theta_2, \\ 1 & \text{if } \theta_2 < \theta, \end{cases}$$

for $\chi \in [0, 1]$ and $0 \leq \theta_1 \leq \theta_2 \leq \bar{\theta}$.

Proof. We next prove the above result from first principles. We only provide a proof for $\mathcal{F}_H$; the proof for $\mathcal{F}_L$ is analogous and thus omitted. Let $C = t - u_H$. We argue that the extreme points of the expanded set

$$\tilde{\mathcal{F}}_H = \{(x, s, r) \in X \times \mathbb{R} \times \mathbb{R} : x \geq_P 0_X, \ s, r \geq 0, \ \ A_2 x + s = 1, \ \ A_1 x - r = C\},$$

correspond to step functions with at most one intermediate step, $s = 0$ and $r \geq 0$.

Since we added slack variables, s and r, we need to consider an expanded cone: $\tilde{P} = P \times \mathbb{R}_+ \times \mathbb{R}_+$. We also define the expanded linear functional $\tilde{A}$ by $(x, s, r) \mapsto (A_2 x + s, A_1 x - r)$. For any $(x, s, r) \in \tilde{P}$ define

$$B((x, s, r)) \triangleq \{(\xi, \eta, \rho) \in X \times \mathbb{R}^2 : (x, s, r) + \lambda(\xi, \eta, \rho) \in \tilde{P},$$
$$(x, s, r) - \lambda(\xi, \eta, \rho) \in \tilde{P} \ \text{ for some scalar } \lambda > 0\},$$
$$N(\tilde{A}) \triangleq \{(\xi, \eta, \rho) \in X \times \mathbb{R}^2 : A_2 \xi + \eta = 0, A_1 \xi - \rho = 0\}.$$

By Theorem 2.2 in Anderson and Nash (1987), we have that (x, s, r) is an extreme of point $\tilde{\mathcal{F}}_H$ if and only if $B((x, s, r)) \cap N(\tilde{A}) = \{(0_X, 0, 0)\}$. Therefore, to characterize the extreme points, it suffices to characterize the points $(x, s, r) \in \tilde{\mathcal{F}}_H$ that make the latter property true. Fix $(x, s, r) \in \tilde{\mathcal{F}}_H$; then, $(\xi, \eta, \rho) \in B((x, s, r)) \cap N(\tilde{A})$ if and only if there exists $\lambda > 0$ such that

$$x + \lambda \xi, x - \lambda \xi \in P, \ \ s + \lambda \eta, s - \lambda \eta \geq 0, \ \ \xi(\bar{\theta}) + \eta = 0, \tag{A.9}$$

and

$$r + \lambda\rho, r - \lambda\rho \geq 0, \quad \int_0^{\bar{\theta}} \xi(\theta)\bar{F}_H(\theta)d\theta = \rho. \tag{A.10}$$

First, note that because $(x, s, r) \in \tilde{\mathcal{F}}_H$, we have

$$r = C - \int_0^{\bar{\theta}} x(\theta)\bar{F}_H(\theta)d\theta.$$

There are two cases, $r > 0$ and $r = 0$. Consider first the case $r > 0$. If x is not a step function, we analyze two subcases: (1) x is strictly increasing and continuous in some interval $[\theta_1, \theta_2]$, or (2) x has two consecutive intermediate steps.

Suppose that we are in (1); by the mean value theorem, there exists $\theta_m \in (\theta_1, \theta_2)$ such that $x(\theta_m) = (x(\theta_1^+) + x(\theta_2^-))/2$. Consider $\eta = 0$ and $\xi(\theta)$ equal to zero outside (θ_1, θ_2) and

$$\xi(\theta) = \begin{cases} \frac{1}{\lambda}(x(\theta) - x(\theta_1^+)) & \text{if } \theta \in (\theta_1, \theta_m] \\ \frac{1}{\lambda}(x(\theta_2^-) - x(\theta)) & \text{if } [\theta_m, \theta_2), \end{cases} \tag{A.11}$$

and we set $\rho = \int_0^{\bar{\theta}} \xi(\theta)\bar{F}_H(\theta)d\theta$ and λ small enough such that $r + \lambda\rho, r - \lambda\rho \geq 0$ (this is possible because $r > 0$). Note that $\xi \neq 0_X$ but (ξ, η, ρ) satisfies conditions (A.9) and (A.10). In turn, no extreme point can be such that is strictly increasing in an interval. Now consider (2), that is, x is such that there are two consecutive intervals in which it takes different and strictly positive values. That is, $x(\theta)$ equals χ_1 in (θ_1, θ_2) and χ_2 in (θ_2, θ_3) with $\chi_1 < \chi_2$ and $x(\theta_1^-) < \chi_1$. We can set $\eta = 0$ and $\xi(\theta) = \mathbf{1}_{\{\theta \in [\theta_1, \theta_2)\}}$ $\rho = \int_0^{\bar{\theta}} \xi(\theta)\bar{F}_H(\theta)d\theta$; and let λ_1 be small enough such that $r + \lambda\rho, r - \lambda\rho \geq 0$. We consider λ equal to $\min\{\lambda_1, \chi_1 - x(\theta_1^-), \chi_2 - \chi_1\}/2$ (here we are assuming, without loss of generality, that x is right continuous). Again, note that $\xi \neq 0_X$ but (ξ, η, ρ) satisfies conditions (A.9) and (A.10). Now, suppose that $x(\theta)$ has a single step, that is, $x(\theta) = \chi \mathbf{1}_{\{\theta \geq \theta_1\}}$. Any ξ that satisfies condition (A.9) must equal zero for $\theta \leq \theta_1$ and it must be constant in $[\theta_1, \bar{\theta}]$. Note that $(x, s, r) \in \tilde{\mathcal{F}}_H$ then $\chi + s = 1$, in turn, this means that if η satisfies condition (A.9) then $\eta \in [-\frac{1-\chi}{\lambda}, \frac{1-\chi}{\lambda}]$. Therefore, if $\chi < 1$ it is possible to find $(\xi, \eta, \rho) \neq (0_X, 0, 0)$ that verify conditions (A.9) and (A.10). In turn, the only possible extreme points of $\tilde{\mathcal{F}}_H$ are such that $\chi = 1$. We have thus proved that the extreme points of $\mathcal{F}_H$ correspond to step functions for the first case $r > 0$.

For the second case, suppose that $r = 0$. In turn, condition (A.10) becomes $\rho = 0$ and $\int_0^{\bar{\theta}} \xi(\theta)\bar{F}_H(\theta)d\theta = 0$. Suppose that x is strictly increasing and continuous in some interval (θ_1, θ_2). Consider some $\theta_m \in (\theta_1, \theta_2)$ (to be defined precisely later), and consider θ_a, θ_b such that $\theta_a \leq \theta_m \leq \theta_b$ and

$$x(\theta_a) = \frac{x(\theta_1^+) + x(\theta_m)}{2} \quad \text{and} \quad x(\theta_b) = \frac{x(\theta_2^-) + x(\theta_m)}{2}. \tag{A.12}$$

Given this we can define ξ to be equal to zero outside (θ_1, θ_2) and

$$\xi(\theta) = \begin{cases} x(\theta) - x(\theta_1^+) & \text{if } \theta \in (\theta_1, \theta_a]; \\ x(\theta_m) - x(\theta) & \text{if } \theta \in [\theta_a, \theta_b]; \\ x(\theta) - x(\theta_2^-) & \text{if } \theta \in [\theta_b, \theta_2). \end{cases}$$

Note that for $\lambda = 1$ we have

$$x(\theta) + \lambda\xi(\theta) = \begin{cases} 2x(\theta) - x(\theta_1^+) & \text{if } \theta \in (\theta_1, \theta_a]; \\ x(\theta_m) & \text{if } \theta \in [\theta_a, \theta_b]; \\ 2x(\theta) - x(\theta_2^-) & \text{if } \theta \in [\theta_b, \theta_2), \end{cases}$$

and

$$x(\theta) - \lambda\xi(\theta) = \begin{cases} x(\theta_1^+) & \text{if } \theta \in (\theta_1, \theta_a]; \\ 2x(\theta) - x(\theta_m) & \text{if } \theta \in [\theta_a, \theta_b]; \\ x(\theta_2^-) & \text{if } \theta \in [\theta_b, \theta_2). \end{cases}$$

Note that $\xi \neq 0_X$ but (ξ, η) satisfies condition (A.9), with $\eta = 0$. Therefore, we only need to verify condition (A.10), that is, $\int_0^{\bar{\theta}} \xi(\theta)\bar{F}_H(\theta)d\theta = 0$. We show that this condition can be satisfied by judiciously choosing θ_m as follows. Given our current definition of ξ, the second part of condition (A.10) is equivalent to

$$\underbrace{\int_{\theta_1}^{\theta_a} (x(\theta) - x(\theta_1^+))\overline{F}_H(\theta)d\theta}_{L_1(\theta_m)} + \underbrace{\int_{\theta_a}^{\theta_b} (x(\theta_m) - x(\theta))\overline{F}_H(\theta)d\theta}_{L_2(\theta_m)}$$
$$+ \underbrace{\int_{\theta_b}^{\theta_2} (x(\theta) - x(\theta_2^-))\overline{F}_H(\theta)d\theta}_{L_3(\theta_m)} = 0, \tag{A.13}$$

where each term above is a function of θ_m (because θ_a and θ_b are functions of θ_m) and continuous. Let $\theta_r \in (\theta_1, \theta_2)$ be such that $x(\theta_r) = \frac{x(\theta_1^+)+x(\theta_2^-)}{2}$. Note that $L_1(\theta_1^+) = 0$, and

$$L_2(\theta_1^+) + L_3(\theta_1^+) = \int_{\theta_1}^{\theta_r} (x(\theta_1^+) - x(\theta))\overline{F}_H(\theta)d\theta + \int_{\theta_r}^{\theta_2} (x(\theta) - x(\theta_2^-))\overline{F}_H(\theta)d\theta < 0. \tag{A.14}$$

We also have that $L_3(\theta_2^-) = 0$, and

$$L_1(\theta_2^-) + L_2(\theta_2^-) = \int_{\theta_1}^{\theta_r} (x(\theta) - x(\theta_1^+))\overline{F}_H(\theta)d\theta + \int_{\theta_r}^{\theta_2} (x(\theta_2^-) - x(\theta))\overline{F}_H(\theta)d\theta > 0. \tag{A.15}$$

In turn, there must exist θ_m for which Eq. (A.13) holds. In conclusion, this rules out allocations x that are strictly increasing in some interval as possible extreme points. We next consider the case in which there are two consecutive intermediate steps.

Consider $x(\theta)$ equal to χ_1 in (θ_1, θ_2) and χ_2 in (θ_2, θ_3) with $\chi_1 < \chi_2$, $x(\theta_1^-) < \chi_1$ and $\chi_2 < x(\theta_3^+)$. Without loss of generality, we can assume that $\theta_1 > 0$ and $\theta_3 < \bar{\theta}$ (if this is not satisfied, then we can apply a similar argument to the one we present next). We can consider

$$\xi(\theta)=\begin{cases}\xi_1\cdot\min\{\chi_1-x(\theta_1^+),\frac{(\chi_2-\chi_1)}{2}\} & \text{if } \theta\in(\theta_1,\theta_2);\\ \xi_2\cdot\min\{x(\theta_3^-)-\chi_2,\frac{(\chi_2-\chi_1)}{2}\} & \text{if } \theta\in[\theta_2,\theta_3),\end{cases} \tag{A.16}$$

where the constants $\xi_1\in(0,1)$ and $\xi_2\in(-1,0)$ are defined in such a way that

$$\int_0^{\bar{\theta}}\xi(\theta)\bar{F}_H(\theta)d\theta=0.$$

In turn, $\xi\neq 0_X$ but (ξ,η,ρ) satisfy conditions (A.9) and (A.10), with $\eta=\rho=0$. This shows that there are no extreme points such that there are two intermediate steps $\chi_1,\chi_2\in(0,1)$.

Next, assume that there is only one such intermediate step as in the statement of the lemma. In turn, $s=0$ which implies that $\eta=0$ and that we must have $\xi(\bar{\theta})=0$. Moreover, any ξ that satisfies condition (A.9) must be constant in $[0,\theta_1)$, $[\theta_1,\theta_2)$ and $[\theta_2,\bar{\theta}])$. In turn, $\xi(\theta)$ equals zero in both $[0,\theta_1)$ and $(\theta_2,\bar{\theta}]$, and equals some constant ξ_1 in (θ_1,θ_2). But condition (A.10) requires $\int_0^{\bar{\theta}}\xi(\theta)\bar{F}_H(\theta)d\theta=0$ which in turn implies that $\xi_1=0$. In conclusion, $\xi(\theta)=0$ for all $\theta\in[0,\bar{\theta}]$. We have thus proved that the extreme points of $\tilde{\mathcal{F}}_H$ correspond to step functions with at most one intermediate step. This concludes the proof of the lemma that characterizes the extreme points of $\mathcal{F}_H$. □

We note that in the case of $K>2$ interim types, we can show using a similar argument based on extreme points that in a model with finitely many ex post valuations, one can restrict attention to contracts that have at most $2(K-1)$ intermediate (randomized) step.[10] We believe that by using arguments similar to Winkler (1988) one may be able to extend this argument for $K>2$ types to the setting of continuous valuation distributions, but this may require additional technical arguments that may be worth exploring in future work. More broadly, we believe that the results based on infinite dimensional linear programming presented here and their possible extensions may be of separate interest in mechanism design.

Proof of Proposition 3. We separate this proof into two parts. In part 1 we show that the optimal solution has the structure in the statement of the theorem. Note that it is enough to provide a proof for the structure of the allocation, the transfers can be readily derived from Lemma 1. In part 2 we derive the properties about the thresholds, x_L and u_H and u_L.

Part 1. First we argue that we can restrict attention to allocations that randomize each type in at most one connected interval. Then we show that for the high type there is no need for a randomized allocation.

According to Theorem 2.5 in Anderson and Nash (1987), the optimal solution to $(\mathcal{P}_R)$, which is an infinite dimensional linear program, is achieved at an extreme point. In turn, we must argue that the extreme points in the feasible set of $(\mathcal{P}_R)$ are step functions with at most one intermediate step for the low and high types. However, this follows immediately from Lemma A.2, which we state and prove immediately before the present proof on pages 33 and 34.

To conclude Part 1 of the proof, we show that for the high type, the intermediate step can be eliminated. Suppose $x_H^\star(\cdot)$ is an optimal solution to $(\mathcal{P}_R)$ for which there exists $\theta_1<\theta_2$ and $0<x<1$ such that $x_H^\star(\theta)=x$ in (θ_1,θ_2). Similar to the proof of type L, assume that

[10] Note that in the case of $K=2$, we only have one constraint because we can show that we can relax the low type's IC constraint.

$x^\star_H(\theta_1^-) < x < x^\star_H(\theta_2^+)$. Consider the following deviation x_H^{dev} which coincides with $x^\star_H$ outside (θ_1, θ_2), and on (θ_1, θ_2) is given as

$$x_H^{dev}(\theta) = \begin{cases} x^\star_H(\theta_1^-) & \text{if } \theta \in (\theta_1, \theta_1 + \epsilon_1) \\ x^\star_H(\theta) & \text{if } \theta \in [\theta_1 + \epsilon_1, \theta_2 - \epsilon_2] \\ x^\star_H(\theta_2^+) & \text{if } \theta \in (\theta_2 - \epsilon_2, \theta_2) \end{cases}$$

for some $\epsilon_1, \epsilon_2 > 0$. We can set ϵ_1 and $\epsilon_2(\epsilon_1)$ such that x_H^{dev} is feasible. It is enough to impose that

$$x_H^- \int_{\theta_1}^{\theta_1+\epsilon_1} \overline{F}_H(z)dz + x \int_{\theta_1+\epsilon_1}^{\theta_2-\epsilon_2} \overline{F}_H(z)dz + x_H^+ \int_{\theta_2-\epsilon_2}^{\theta_2} \overline{F}_H(z)dz = x \int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz, \tag{A.17}$$

with $x_H^- = x^\star_H(\theta_1^-)$ and $x_H^+ = x^\star_H(\theta^+)$. Note that this equation defines $\epsilon_2(\epsilon_1)$. Also, $\epsilon_2(\epsilon_1)$ defined in this way is strictly increasing. Moreover, the values that $\epsilon_2(\epsilon_1)$ can take are limited by $\theta_2 - \epsilon_2(\epsilon_1) \geq \theta_1 + \epsilon_1$, that is, the integration interval in the middle term on the left-hand side of (A.17) must be well defined (such that the integral is nonnegative). Therefore, the function $\epsilon_2(\epsilon_1)$ is always bounded above by $\theta_2 - \theta_1 - \epsilon_1$. The unique $\epsilon_1^\star$ such that these two functions are equal, $\epsilon_2(\epsilon_1^\star) = \theta_2 - \theta_1 - \epsilon_1^\star$, represents the upper limit in the domain of $\epsilon_2(\epsilon_1)$. Note that at this point the middle term on the left-hand side of (A.17) vanishes.

Taking the derivative in (A.17) with respect to ϵ_1 yields the following:

$$\epsilon_2'(\epsilon_1) = \frac{(x - x_H^-)\overline{F}_H(\theta_1 + \epsilon_1)}{(x_H^+ - x)\overline{F}_H(\theta_2 - \epsilon_2)}. \tag{A.18}$$

The change in profit for the seller is (proportional to)

$$\Delta = (x_H^- - x) \int_{\theta_1}^{\theta_1+\epsilon_1} \mu_H(z) f_H(z)dz + (x_H^+ - x) \int_{\theta_2-\epsilon_2}^{\theta_2} \mu_H(z) f_H(z)dz,$$

so that

$$\begin{aligned} \frac{d\Delta}{d\epsilon_1} &= (x_H^- - x)\mu_H(\theta_1+\epsilon_1) f_H(\theta_1+\epsilon_1) + (x_H^+ - x)\mu_H(\theta_2-\epsilon_2) f_H(\theta_2-\epsilon_2)\epsilon_2' \\ &\overset{(A.18)}{=} (x_H^- - x)\overline{F}_H(\theta_1+\epsilon_1)\left[\frac{\mu_H(\theta_1+\epsilon_1) f_H(\theta_1+\epsilon_1)}{\overline{F}_H(\theta_1+\epsilon_1)} - \frac{\mu_H(\theta_2-\epsilon_2) f_H(\theta_2-\epsilon_2)}{\overline{F}_H(\theta_2-\epsilon_2)}\right] \\ &= (x_H^- - x)\overline{F}_H(\theta_1+\epsilon_1)\left[r^{HH}(\theta_1+\epsilon_1) - r^{HH}(\theta_2-\epsilon_2)\right] \end{aligned}$$

Because r^{HH} is nondecreasing and $(x_H^- - x) < 0$, this expression is (weakly) positive. In turn, we can conclude that by moving from $\epsilon_1 = 0$ to $\epsilon_1 = \epsilon_1^\star$ we obtain a weak revenue improvement. Since at $\epsilon_1^\star$ the intermediate step, x, vanishes, we obtain the desired result. This completes the proof for interim type 2 and case (2).

In conclusion, we can always consider $x^\star_H$ to be a threshold allocation as in the statement of the proposition.

Part 2. From what we have just proved, we can write $(\mathcal{P}_R)$ as follows

$$\max \quad - \sum_{k \in \{L,H\}} \alpha_k u_k + \alpha_L x \int_{\theta_1}^{\theta_2} \mu_L(z) f_L(z)dz + \alpha_L \int_{\theta_2}^{\bar{\theta}} \mu_L(z) f_L(z)dz$$

$$+\alpha_2 \int_{\theta_H}^{\bar{\theta}} \mu_H(z) f_H(z) dz$$

$$\text{s.t} \quad x \in [0,1], \quad \theta_1 \leq \theta_2$$

$$u_k \geq 0, \;\; k \in \{L, H\}$$

$$u_H + \int_{\theta_H}^{\bar{\theta}} \overline{F}_H(z)dz \geq u_L + x \int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz + \int_{\theta_2}^{\bar{\theta}} \overline{F}_H(z)dz.$$

We prove the properties satisfied by u_L, θ_1, θ_H and θ_2. From the formulation above it is clear that is always optimal to set $u_L = 0$. To see that $\widehat{\theta}_L \leq \theta_1$ suppose the opposite, that is, $\widehat{\theta}_L > \theta_1$. This implies that between θ_1 and $\widehat{\theta}_1$, $\mu_L(\cdot)$ is negative. Then, we can increase θ_1 while maintaining feasibility and, simultaneously, increasing the objective function. Note that this argument is also valid when $\theta_1 = \theta_2$. Additionally, note that we can obtain a strict improvement only when $x > 0$; however, when $x = 0$ we can only obtain a weak improvement. In either case, we can always consider $\widehat{\theta}_L \leq \theta_1$. To see that $\theta_H \leq \widehat{\theta}_H$, suppose the opposite, $\theta_H > \widehat{\theta}_H$. Since $\mu_H(\theta) > 0$ for all $\theta \geq \widehat{\theta}_H$, we can decrease θ_H and obtain an objective improvement while maintaining feasibility.

Next, we argue that $u_H = 0$. Suppose that $u_H > 0$; then, we must have

$$u_H + \int_{\theta_H}^{\bar{\theta}} \overline{F}_H(z)dz = x \int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz + \int_{\theta_2}^{\bar{\theta}} \overline{F}_H(z)dz, \tag{A.19}$$

otherwise, we could decrease u_H and, by doing so, improve the objective. Since $u_H > 0$, equation (A.19) yields

$$0 < u_H = x \int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz + \int_{\theta_2}^{\bar{\theta}} \overline{F}_H(z)dz - \int_{\theta_H}^{\bar{\theta}} \overline{F}_H(z)dz, \tag{A.20}$$

then it must be true that $\theta_1 < \theta_H$; otherwise, from equation (A.20) we would have ($\theta_1 \leq \theta_2$)

$$\int_{\theta_H}^{\theta_1} \overline{F}_H(z)dz + \int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz + \int_{\theta_2}^{\bar{\theta}} \overline{F}_H(z)dz < x \int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz + \int_{\theta_2}^{\bar{\theta}} \overline{F}_H(z)dz,$$

which implies

$$\int_{\theta_H}^{\theta_1} \overline{F}_H(z)dz < 0,$$

a contradiction. Thus, $\theta_1 < \theta_H$.

Now consider a new contract for type H that consists of decreasing the cutoff θ_H by $\epsilon > 0$ sufficiently small, but at the same time maintaining the equality in equation (A.19). Specifically, let $\theta_H(\epsilon) = \theta_H - \epsilon > 0$ (which we can do because as we just saw $\theta_H > \theta_1 \geq 0$) and let $u_H(\epsilon)$ be

$$u_H(\epsilon) = x \int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz + \int_{\theta_2}^{\bar{\theta}} \overline{F}_H(z)dz - \int_{\theta_H(\epsilon)}^{\bar{\theta}} \overline{F}_H(z)dz.$$

Note that by taking ϵ small we still have $u_H(\epsilon) > 0$. We claim that this new contract, characterized by $\theta_1, \theta_2, x, \theta_H(\epsilon)$ and $u_H(\epsilon)$, yields a larger value than the old contract, characterized by $\theta_1, \theta_2, x, \theta_H$ and u_H. The old contract's objective is

$$-\alpha_H u_H + \alpha_L x \int_{\theta_1}^{\theta_2} \mu_L(z) f_L(z) dz + \alpha_L \int_{\theta_2}^{\bar{\theta}} \mu_L(z) f_L(z) dz + \alpha_H \int_{\theta_H}^{\bar{\theta}} \mu_H(z) f_H(z) dz,$$

and using equation (A.19) it becomes

$$x \int_{\theta_1}^{\theta_2} (\alpha_L \mu_L(z) f_L(z) - \alpha_H \overline{F}_H(z)) dz + \int_{\theta_2}^{\bar{\theta}} (\alpha_L \mu_L(z) f_L(z) - \alpha_H \overline{F}_H(z)) dz$$
$$+ \alpha_H \int_{\theta_H}^{\bar{\theta}} z f_H(z) dz.$$

We obtain a similar expression for the new contract's objective. Specifically, the first two terms in the expression above are the same and the third term differs in θ_H. Hence, the new contract yields an improvement over the old one if and only if

$$\int_{\theta_H}^{\bar{\theta}} z f_H(z) dz < \int_{\theta_H(\epsilon)}^{\bar{\theta}} z f_H(z) dz.$$

Since $\theta_H(\epsilon) < \theta_H$ this last inequality is true. Thus, if $u_H > 0$ we can always construct a new contract yielding a larger objective value and, therefore, at any optimal contract we must have $u_H = 0$.

To show that $\theta_H \leq \theta_2$, note that since at any optimal solution $u_H = 0$, the incentive constraint is

$$\int_{\theta_H}^{\bar{\theta}} \overline{F}_H(z) dz \geq x \int_{\theta_1}^{\theta_2} \overline{F}_H(z) dz + \int_{\theta_2}^{\bar{\theta}} \overline{F}_H(z) dz.$$

Hence, if $\theta_H > \theta_2$ from the expression above we would have

$$\int_{\theta_H}^{\bar{\theta}} \overline{F}_H(z) dz \geq x \int_{\theta_1}^{\theta_2} \overline{F}_H(z) dz + \int_{\theta_2}^{\theta_H} \overline{F}_H(z) dz + \int_{\theta_H}^{\bar{\theta}} \overline{F}_H(z) dz,$$

which implies that $\theta_H = \theta_2$, a contradiction.

Next, we argue that $\theta_1 \leq \theta_H$. First, we show that $\theta_1 \leq \widehat{\theta}_H$. Suppose the opposite, that is, $\theta_1 > \widehat{\theta}_H$. Then, since $\widehat{\theta}_H \geq \theta_H$ we must have $\theta_1 > \theta_H$, and therefore,

$$\int_{\theta_H}^{\bar{\theta}} \overline{F}_H(z) dz = \int_{\theta_H}^{\theta_1} \overline{F}_H(z) dz + \int_{\theta_1}^{\bar{\theta}} \overline{F}_H(z) dz$$

$$
\begin{aligned}
&> \int_{\theta_1}^{\bar{\theta}} \overline{F}_H(z)dz \\
&= \int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz + \int_{\theta_2}^{\bar{\theta}} \overline{F}_H(z)dz \\
&\geq x \int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz + \int_{\theta_2}^{\bar{\theta}} \overline{F}_H(z)dz.
\end{aligned}
$$

That is, the incentive constraint is not binding. Therefore, since $\theta_1 > \widehat{\theta}_H \geq \widehat{\theta}_L$, we can slightly decrease θ_1 and, in this way, obtain an objective improvement whenever $x > 0$. When $x = 0$, because $\theta_2 \geq \theta_1$, we can decrease θ_2 and obtain an objective improvement as well. Hence, at any optimal solution we must have $\theta_1 \leq \widehat{\theta}_H$.

To complete the proof, suppose that $\theta_1 > \theta_H$; then, as before, we have

$$
\int_{\theta_H}^{\bar{\theta}} \overline{F}_H(z)dz > x \int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz + \int_{\theta_2}^{\bar{\theta}} \overline{F}_H(z)dz.
$$

Using that $\theta_1 \leq \widehat{\theta}_H$ implies $\theta_H < \widehat{\theta}_H$, we can slightly increase θ_H (maintaining feasibility) and thus obtain an objective improvement. In conclusion, at any optimal solution, we must have $\theta_1 \leq \theta_H$.

Finally we must have that $x = \int_{\theta_H}^{\theta_2} \overline{F}_H(z)dz / \int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz$. Indeed, since $\widehat{\theta}_L \leq \theta$, the part of the objective that involves x is always nonnegative and, therefore, it is optimal to make x as large as possible. The incentive constraints provide an upper bound for x, which is precisely $\int_{\theta_H}^{\theta_2} \overline{F}_H(z)dz / \int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz$, thus the result. □

Proof of Theorem 2. We next show that the solutions to the relaxed problem and the original problem coincide. It is enough to show that the solution of $(\mathcal{P}_R)$ is feasible in $(\mathcal{P})$. From Proposition 3 we know that we can formulate $(\mathcal{P}_R)$ as

$$
\begin{aligned}
(\mathcal{P}_R^d) \quad \max \quad & \alpha_L x \int_{\theta_1}^{\theta_2} \mu_L(z) f_L(z)dz + \alpha_L \int_{\theta_2}^{\bar{\theta}} \mu_L(z) f_L(z)dz + \alpha_H \int_{\theta_H}^{\bar{\theta}} \mu_H(z) f_H(z)dz \\
\text{s.t} \quad & x = \frac{\int_{\theta_H}^{\theta_2} \overline{F}_H(z)dz}{\int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz} \\
& \hat{\theta}_L \leq \theta_1 \leq \theta_H \leq \theta_2, \theta_H \leq \hat{\theta}_H \\
& \int_{\theta_H}^{\bar{\theta}} \overline{F}_H(z)dz \geq x \int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz + \int_{\theta_2}^{\bar{\theta}} \overline{F}_H(z)dz.
\end{aligned}
$$

Let $\theta_1, \theta_H, \theta_2$ and x be the optimal solution to $(\mathcal{P}_R^d)$. If this solution corresponds to the optimal static contract or yields the same objective as it, we are done because this contract is always feasible in ($\mathcal{P}$). If this solution is different from the optimal static contract and yields a strictly larger objective, it must be the case that

$$\int_{\theta_H}^{\bar{\theta}} \bar{\mu}(z)dz < \alpha_L x \int_{\theta_1}^{\theta_2} \mu_L(z) f_L(z)dz + \alpha_L \int_{\theta_2}^{\bar{\theta}} \mu_L(z) f_L(z)dz + \alpha_H \int_{\theta_H}^{\bar{\theta}} \mu_H(z) f_H(z)dz. \tag{A.21}$$

This is true because the contract $(u_1, u_2, x_1, x_2) = (0, 0, \mathbf{1}_{\{\theta \geq \theta_H\}}, \mathbf{1}_{\{\theta \geq \theta_H\}})$ is a feasible static contract, and therefore, its associated revenue is bounded by that of the optimal static contract. From the formulation of ($\mathcal{P}_R^d$) we know that $\widehat{\theta}_L \leq \theta_1 \leq \theta_H \leq \theta_2$, this and equation (A.21) deliver

$$0 \leq \int_{\theta_H}^{\theta_2} \mu_L(z) f_L(z)dz < x \int_{\theta_1}^{\theta_2} \mu_L(z) f_L(z)dz.$$

Hence, $\theta_1 < \theta_2$, $\theta_H < \theta_2$ (otherwise $x = 0$) and

$$\frac{\int_{\theta_H}^{\theta_2} \mu_L(z) f_L(z)dz}{\int_{\theta_1}^{\theta_2} \mu_L(z) f_L(z)dz} < x. \tag{A.22}$$

Also, since $x \leq 1$ we must have $\theta_1 < \theta_H$. Note that since $\widehat{\theta}_L \leq \theta_1 < \theta_2$ the denominator above is strictly positive.

Now we argue that the contract optimizing ($\mathcal{P}_R^d$) characterized by $\theta_1, \theta_H, \theta_2$ and x is feasible for ($\mathcal{P}$). Since the high to low incentive constraint is satisfied, we only need to verify the low to high incentive constraint. That is, we need to verify the following inequality

construct a new contract that is feasible for ($\mathcal{P}^d$) and yields a strictly larger objective value than the optimal static contract. In fact, this new contract is the one that optimizes ($\mathcal{P}_R^d$). Therefore, we only need to check feasibility. Since the high to low IC constraint is satisfied we need to verify the low to high IC constraint, that is, we need to verify the following inequality

$$x \int_{\theta_1}^{\theta_2} \overline{F}_L(z)dz + \int_{\theta_2}^{\bar{\theta}} \overline{F}_L(z)dz \geq \int_{\theta_H}^{\bar{\theta}} \overline{F}_L(z)dz, \tag{A.23}$$

or, equivalently, $x \geq \int_{\theta_H}^{\theta_2} \overline{F}_L(z)dz / \int_{\theta_1}^{\theta_2} \overline{F}_L(z)dz$. In order to see why (A.23) holds, observe that from Lemma A.3 (which we state and prove after the present proof) we have

$$\frac{\int_{\theta_1}^{\theta_2} \mu_L(z) f_L(z)dz}{\int_{\theta_1}^{\theta_2} \overline{F}_L(z)dz} \leq \frac{\int_{\theta_H}^{\theta_2} \mu_L(z) f_L(z)dz}{\int_{\theta_H}^{\theta_2} \overline{F}_L(z)dz} \Leftrightarrow \frac{\int_{\theta_1}^{\theta_H} \mu_L(z) f_L(z)dz}{\int_{\theta_1}^{\theta_H} \overline{F}_L(z)dz} \leq \frac{\int_{\theta_H}^{\theta_2} \mu_L(z) f_L(z)dz}{\int_{\theta_H}^{\theta_2} \overline{F}_L(z)dz}. \tag{A.24}$$

The right-hand side in (A.24) always holds thanks to (IHR), indeed,

$$\begin{aligned}\frac{\int_{\theta_1}^{\theta_H} \mu_L(z) f_L(z)dz}{\int_{\theta_1}^{\theta_H} \overline{F}_L(z)dz} &= \frac{\int_{\theta_1}^{\theta_H} \overline{F}_L r^{LL}(z)dz}{\int_{\theta_1}^{\theta_H} \overline{F}_L(z)dz} \leq r^{LL}(\theta_H) \leq \frac{\int_{\theta_H}^{\theta_2} \overline{F}_L r^{LL}(z)dz}{\int_{\theta_H}^{\theta_2} \overline{F}_L(z)dz} \\ &= \frac{\int_{\theta_H}^{\theta_2} \mu_L(z) f_L(z)dz}{\int_{\theta_H}^{\theta_2} \overline{F}_L(z)dz}.\end{aligned}$$

Thus the left-hand side in (A.24) holds. Equivalently,

$$\frac{\int_{\theta_H}^{\theta_2} \overline{F}_L(z)dz}{\int_{\theta_1}^{\theta_2} \overline{F}_L(z)dz} \leq \frac{\int_{\theta_H}^{\theta_2} \mu_L(z) f_L(z)dz}{\int_{\theta_1}^{\theta_2} \mu_L(z) f_L(z)dz}$$

Using this, together with equation (A.22), delivers equation (A.23). This concludes the proof. □

Lemma A.3. *Let $\theta_i \in [0, \bar{\theta}]$ for $i = 1, 2, 3$ be such that $\theta_1 < \theta_2 < \theta_3$. Additionally, consider functions $f, g : [\theta_1, \theta_3] \to \mathbb{R}$, with $\int_{\theta_1}^{\theta_2} g(z)dz, \int_{\theta_2}^{\theta_3} g(z)dz > 0$. Then,*

$$\frac{\int_{\theta_1}^{\theta_3} f(z)dz}{\int_{\theta_1}^{\theta_3} g(z)dz} \leq \frac{\int_{\theta_2}^{\theta_3} f(z)dz}{\int_{\theta_2}^{\theta_3} g(z)dz} \quad \textit{if and only if} \quad \frac{\int_{\theta_1}^{\theta_2} f(z)dz}{\int_{\theta_1}^{\theta_2} g(z)dz} \leq \frac{\int_{\theta_2}^{\theta_3} f(z)dz}{\int_{\theta_2}^{\theta_3} g(z)dz}.$$

Proof of Lemma A.3.

$$\begin{aligned}
\frac{\int_{\theta_1}^{\theta_3} f(z)dz}{\int_{\theta_1}^{\theta_3} g(z)dz} \leq \frac{\int_{\theta_2}^{\theta_3} f(z)dz}{\int_{\theta_2}^{\theta_3} g(z)dz} &\Leftrightarrow \Big(\int_{\theta_2}^{\theta_3} g(z)dz\Big)\Big(\int_{\theta_1}^{\theta_3} f(\theta)dz\Big) \leq \Big(\int_{\theta_1}^{\theta_3} g(z)dz\Big)\Big(\int_{\theta_2}^{\theta_3} f(z)dz\Big) \\
&\Leftrightarrow \Big(\int_{\theta_2}^{\theta_3} g(z)dz\Big)\Big(\int_{\theta_1}^{\theta_2} f(z)dz\Big) \leq \Big(\int_{\theta_1}^{\theta_2} g(z)dz\Big)\Big(\int_{\theta_2}^{\theta_3} f(z)dz\Big) \\
&\Leftrightarrow \frac{\int_{\theta_1}^{\theta_2} f(z)dz}{\int_{\theta_1}^{\theta_2} g(z)dz} \leq \frac{\int_{\theta_2}^{\theta_3} f(z)dz}{\int_{\theta_2}^{\theta_3} g(z)dz} \quad \square
\end{aligned}$$

Proof of Proposition 7. The problem we analyze in this proposition is:

$$\max_{0 \leq \mathbf{x} \leq 1, \mathbf{u}} \quad -\sum_{k \in \{L,H\}} \alpha_k u_k + \sum_{k \in \{L,H\}} \alpha_k \int_0^{\bar{\theta}} x_k(z)\mu_k(z) f_k(z)dz \tag{$\mathcal{P}_B$}$$

$$\begin{aligned}
\text{s.t} \quad & x_k(\theta) \quad \text{nondecreasing}, \quad \forall k \in \{L, H\} \\
& u_k \geq -B, \quad \forall k \in \{L, H\} \\
& u_k + \int_0^{\bar{\theta}} x_k(z)\overline{F}_k(z)dz \geq u_{k'} + \int_0^{\bar{\theta}} x_{k'}(z)\overline{F}_k(z)dz, \quad \forall k, k' \in \{L, H\} \\
& u_k + \int_0^{\bar{\theta}} x_k(z)\overline{F}_k(z)dz \geq 0, \quad \forall k \in \{L, H\}.
\end{aligned}$$

To prove that (APR) implies the optimality of the static contract we consider $(\mathcal{P}_B)$ and relax the interim IR constraint. The resulting problem is the same as the original screening problem $(\mathcal{P})$ except for the change that $u_k \geq -B$. Then by following the same exact steps in the sufficiency part of the proof of Theorem 1 the implication follows.

Now for the reverse implication, if (APR) does not hold when the static contract is optimal then it is possible to construct a dynamic contract as in the proof of Proposition 1 that gives a strict revenue improvement. The only subtlety is that now we must verify that the constraint $u_k + \int_0^{\bar{\theta}} x(z)\overline{F}_k(z)dz \geq 0$ is satisfied for $k \in \{L, H\}$. This can be readily verified, following the notation in the proof of Proposition 1 we have (for $u_L = u_H = -B$) for the H type

$$u_H + \int_0^{\bar{\theta}} x(z)\overline{F}_H(z)dz \geq 0 \Leftrightarrow \int_{\hat{\theta}}^{\bar{\theta}} \overline{F}_L(z)dz \geq B \tag{A.25}$$

which is true because we are assuming $\min_{k\in\{L,H\}} \int_{\hat{\theta}}^{\bar{\theta}} \overline{F}_k(z)dz \geq B$. For the L type, we have

$$u_L + \int_0^{\bar{\theta}} x(z)\overline{F}_L(z)dz \geq 0 \Leftrightarrow x \geq \frac{B - \int_{\theta_2}^{\bar{\theta}} \overline{F}_L(z)dz}{\int_{\theta_1}^{\theta_2} \overline{F}_L(z)dz}. \tag{A.26}$$

Moreover, in that proof we established that

$$x = \frac{\int_{\hat{\theta}}^{\theta_2} \overline{F}_H(z)dz}{\int_{\theta_1}^{\theta_2} \overline{F}_H(z)dz} \geq \frac{\int_{\hat{\theta}}^{\theta_2} \overline{F}_L(z)dz}{\int_{\theta_1}^{\theta_2} \overline{F}_L(z)dz}, \tag{A.27}$$

but note that since $\int_{\hat{\theta}}^{\bar{\theta}} \overline{F}_L(z)dz \geq \min_{k\in\{L,H\}} \int_{\hat{\theta}}^{\bar{\theta}} \overline{F}_k(z)dz \geq B$ we have that

$$x \geq \frac{\int_{\hat{\theta}}^{\theta_2} \overline{F}_L(z)dz}{\int_{\theta_1}^{\theta_2} \overline{F}_L(z)dz} \geq \frac{B - \int_{\theta_2}^{\bar{\theta}} \overline{F}_L(z)dz}{\int_{\theta_1}^{\theta_2} \overline{F}_L(z)dz}. \tag{A.28}$$

In conclusion, if $\int_{\hat{\theta}}^{\bar{\theta}} \overline{F}_L(z)dz \geq B$, the static contract is optimal if and only if (APR) holds, that is, Theorem 1 still holds. □

Proof of Theorem 3. In Lemma A.4 (which we state and prove after this proof) we show that $\mathcal{A}$ is nonempty. Next, we prove the necessary and sufficient condition.

We prove both directions separately. First we show that if there exists $\boldsymbol{\lambda} \in \mathcal{A}$ satisfying the properties then the static contract is optimal. Then we show that if the static contract is optimal then we can always solve for $\boldsymbol{\lambda}$ satisfying the properties.

Define

$$\Omega \triangleq \{x : [0, \bar{\theta}] \longrightarrow [0, 1] : x(\cdot) \text{ is nondecreasing}\}, \quad \text{and} \quad \Omega^K \triangleq \underbrace{\Omega \times \cdots \times \Omega}_{K \text{ times}}.$$

For the first part we use a Lagrangian relaxation approach. That is, we dualize the incentive constraints for a specific set of multipliers. This gives an upper bound to the seller's problem. Then we show that for our choice of multipliers the relaxation is maximized at the static allocation. The Lagrangian is

$$\mathcal{L}(x, u, \boldsymbol{\lambda}, \mathbf{w}) = \sum_{k=1}^{K} u_k \Big(-\alpha_k + w_k + \sum_{j:j\neq k} \lambda_{kj} - \sum_{j:j\neq k} \lambda_{jk} \Big)$$

$$+\sum_{k=1}^{K}\int_{0}^{\bar{\theta}} x_k(z)\Big(\alpha_k\mu_k(z)f_k(z)+\overline{F}_k(z)\cdot\sum_{j:j\neq k}\lambda_{kj}-\sum_{j:j\neq k}\lambda_{jk}\overline{F}_j(z)\Big)dz,$$

where $\boldsymbol{\lambda}$ correspond to the multipliers associated with the incentives, and $\mathbf{w}$ to the multipliers associated with the ex post IR constraints. Let us define $\boldsymbol{\lambda}$ to be equal to the $(\lambda_{ij})_{i,j\in\{1,\cdots,K\}^2}$ we are assuming to exist, that is $\boldsymbol{\lambda}\in\mathcal{A}$, and let

$$w_k=\alpha_k+\sum_{j:j\neq k}\lambda_{jk}-\sum_{j:j\neq k}\lambda_{kj},\forall k\in\{1,\ldots,K\}. \tag{A.29}$$

Note that by our choice of $\boldsymbol{\lambda}$ ($\boldsymbol{\lambda}\in\mathcal{A}$), w_k is nonnegative for all k. With this choice of $\mathbf{w}$ the first summation in the Lagrangian becomes zero. Now, we need to show that for this choice of multipliers the Lagrangian is maximized at the static contract. In order to show this observe that

$$\begin{aligned}&\max_{x\in\Omega^K,u\geq 0}\ \mathcal{L}(x,u,\boldsymbol{\lambda},\mathbf{w})\\&=\sum_{k=1}^{K}\max_{x_k\in\Omega}\int_{0}^{\bar{\theta}} x_k(z)\Big(\alpha_k\mu_k(z)f_k(z)+\overline{F}_k(z)\cdot\sum_{j:j\neq k}\lambda_{kj}-\sum_{j:j\neq k}\lambda_{jk}\overline{F}_j(z)\Big)dz.\end{aligned} \tag{A.30}$$

Thus we only need to verify that the RHS of (A.30) is bounded above by

$$\sum_{k=1}^{K}\int_{\widehat{\theta}}^{\bar{\theta}}\Big(\alpha_k\mu_k(z)f_k(z)+\overline{F}_k(z)\cdot\sum_{j:j\neq k}\lambda_{kj}-\sum_{j:j\neq k}\lambda_{jk}\overline{F}_j(z)\Big)dz. \tag{A.31}$$

Note that the RHS of (A.30), for each k, is maximized at some threshold contract $\theta_k\in[0,1]$. To prove that (A.31) is an upper bound of (A.30) is enough to show that for all k and for any $\theta_k\in[0,1]$

$$\begin{aligned}&\int_{\theta_k}^{\bar{\theta}}\Big(\alpha_k\mu_k(z)f_k(z)+\overline{F}_k(z)\cdot\sum_{j:j\neq k}\lambda_{kj}-\sum_{j:j\neq k}\lambda_{jk}\overline{F}_j(z)\Big)dz\\&\leq\int_{\widehat{\theta}}^{\bar{\theta}}\Big(\alpha_k\mu_k(z)f_k(z)+\overline{F}_k(z)\cdot\sum_{j:j\neq k}\lambda_{kj}-\sum_{j:j\neq k}\lambda_{jk}\overline{F}_j(z)\Big)dz.\end{aligned} \tag{A.32}$$

Consider $\theta_k\geq\widehat{\theta}$ in (A.32), then (A.32) becomes

$$0\leq\int_{\widehat{\theta}}^{\theta_k}\Big(\alpha_k\mu_k(z)f_k(z)+\overline{F}_k(z)\cdot\sum_{j:j\neq k}\lambda_{kj}-\sum_{j:j\neq k}\lambda_{jk}\overline{F}_j(z)\Big)dz,$$

this is equivalent to

$$-\Big(\sum_{j:j\neq k}\lambda_{kj}\Big)\cdot\int_{\widehat{\theta}}^{\theta_k}\overline{F}_k(z)dz\leq\int_{\widehat{\theta}}^{\theta_k}\Big(\alpha_k\mu_k(z)f_k(z)-\sum_{j:j\neq k}\lambda_{jk}\overline{F}_j(z)\Big)dz,\quad\forall\theta_k\geq\widehat{\theta},$$

which can be rewritten as

$$-\Big(\sum_{j:j\neq k}\lambda_{kj}\Big)\leq\min_{\widehat{\theta}\leq\theta}\left\{\alpha_k\frac{\int_{\widehat{\theta}}^{\theta}\mu_k(z)f_k(z)dz}{\int_{\widehat{\theta}}^{\theta}\overline{F}_k(z)dz}-\sum_{j:j\neq k}\lambda_{jk}\cdot\frac{\int_{\widehat{\theta}}^{\theta}\overline{F}_j(z)dz}{\int_{\widehat{\theta}}^{\theta}\overline{F}_k(z)dz}\right\}. \tag{A.33}$$

Similarly, if $\theta_k\leq\widehat{\theta}$ then (A.32) is equivalent to

$$0\geq\int_{\theta_k}^{\widehat{\theta}}\Big(\alpha_k\mu_k(z)f_k(z)+\overline{F}_k(z)\cdot\sum_{j:j\neq k}\lambda_{kj}-\sum_{j:j\neq k}\lambda_{jk}\overline{F}_j(z)\Big)dz,\quad\forall\theta_k\leq\widehat{\theta},$$

which is equivalent to

$$\max_{\theta\leq\widehat{\theta}}\left\{\alpha_k\frac{\int_{\theta}^{\widehat{\theta}}\mu_k(z)f_k(z)dz}{\int_{\theta}^{\widehat{\theta}}\overline{F}_k(z)dz}-\sum_{j:j\neq k}\lambda_{jk}\cdot\frac{\int_{\theta}^{\widehat{\theta}}\overline{F}_j(z)dz}{\int_{\theta}^{\widehat{\theta}}\overline{F}_k(z)dz}\right\}\leq-\Big(\sum_{j:j\neq k}\lambda_{kj}\Big). \tag{A.34}$$

In summary, proving that (A.32) holds is equivalent to showing that both (A.33) and (A.34) hold. To see why this is true, note that

$$\begin{aligned}&\lim_{\theta\to\widehat{\theta}+}\alpha_k\frac{\int_{\widehat{\theta}}^{\theta}\mu_k(z)f_k(z)dz}{\int_{\widehat{\theta}}^{\theta}\overline{F}_k(z)dz}-\sum_{j:j\neq k}\lambda_{jk}\cdot\frac{\int_{\widehat{\theta}}^{\theta}\overline{F}_j(z)dz}{\int_{\widehat{\theta}}^{\theta}\overline{F}_k(z)dz}\\&\quad=\frac{\alpha_k\cdot\mu_k(\widehat{\theta})\cdot f_k(\widehat{\theta})-\sum_{j:j\neq k}\lambda_{jk}\cdot\overline{F}_j(\widehat{\theta})}{\overline{F}_k(\widehat{\theta})}=-\Big(\sum_{j:j\neq k}\lambda_{kj}\Big),\end{aligned} \tag{A.35}$$

where the last equality comes from the choice of the multipliers. Since the limit is taken for values above $\widehat{\theta}$, this implies that

$$\begin{aligned}&\min_{\widehat{\theta}\leq\theta}\left\{\alpha_k\frac{\int_{\widehat{\theta}}^{\theta}\mu_k(z)f_k(z)dz}{\int_{\widehat{\theta}}^{\theta}\overline{F}_k(z)dz}-\sum_{j:j\neq k}\lambda_{jk}\cdot\frac{\int_{\widehat{\theta}}^{\theta}\overline{F}_j(z)dz}{\int_{\widehat{\theta}}^{\theta}\overline{F}_k(z)dz}\right\}\\&\quad\leq\lim_{\theta\to\widehat{\theta}+}\alpha_k\frac{\int_{\widehat{\theta}}^{\theta}\mu_k(z)f_k(z)dz}{\int_{\widehat{\theta}}^{\theta}\overline{F}_k(z)dz}-\sum_{j:j\neq k}\lambda_{jk}\cdot\frac{\int_{\widehat{\theta}}^{\theta}\overline{F}_j(z)dz}{\int_{\widehat{\theta}}^{\theta}\overline{F}_k(z)dz}\\&\quad=-\Big(\sum_{j:j\neq k}\lambda_{kj}\Big).\end{aligned}$$

A similar argument (taking the limit for values below $\widehat{\theta}$ this time) can be used to show that

$$-\Big(\sum_{j:j\neq k}\lambda_{kj}\Big)\leq\max_{\theta\leq\widehat{\theta}}\left\{\alpha_k\frac{\int_{\theta}^{\widehat{\theta}}\mu_k(z)f_k(z)dz}{\int_{\theta}^{\widehat{\theta}}\overline{F}_k(z)dz}-\sum_{j:j\neq k}\lambda_{jk}\cdot\frac{\int_{\theta}^{\widehat{\theta}}\overline{F}_j(z)dz}{\int_{\theta}^{\widehat{\theta}}\overline{F}_k(z)dz}\right\}.$$

Since we are assuming that the minimum is an upper bound to the maximum above, we can conclude that both (A.33) and (A.34) hold (with equality). This concludes the proof for the first direction.

For the second direction we need to show that if the static contract is optimal then we can find $\boldsymbol{\lambda}$ satisfying condition (APR^M). Theorem 1 in Luenberger (1969, p. 217) gives then the existence of Lagrange multipliers such that the static contract maximizes the Lagrangian (here we use the interior point condition in the assumptions). In other words, $\exists\boldsymbol{\lambda},\boldsymbol{w}\geq 0$ such that

$$\mathcal{L}(\mathbf{x}^{\mathbf{s}},\mathbf{0},\boldsymbol{\lambda},\boldsymbol{w})\geq\mathcal{L}(\mathbf{x},\boldsymbol{u},\boldsymbol{\lambda},\boldsymbol{w}),\quad\forall\boldsymbol{u},\mathbf{x}\in\mathbb{R}_+^K\times\Omega^K. \tag{A.36}$$

Note that (A.36) holds for any $\boldsymbol{u}, \mathbf{x} \in \mathbb{R}_+^K \times \Omega^K$. Thus we can first consider $\mathbf{x}$ equal to $\mathbf{x^s}$ in (A.36), this yields

$$0 \geq \sum_{k=1}^{K} u_k \Big(-\alpha_k + w_k + \sum_{j:j\neq k} \lambda_{kj} - \sum_{j:j\neq k} \lambda_{jk} \Big), \quad \forall \boldsymbol{u} \in \mathbb{R}_+^K.$$

Which implies that

$$-\alpha_k + w_k + \sum_{j:j\neq k} \lambda_{kj} - \sum_{j:j\neq k} \lambda_{jk} = 0, \quad \forall k,$$

and since $w_k \geq 0$ we can conclude that

$$\alpha_k \geq \sum_{j:j\neq k} \lambda_{kj} - \sum_{j:j\neq k} \lambda_{jk}, \quad \forall k,$$

as required. Now, fix k and consider a solution $\mathbf{x} \in \Omega^K$ such that $x_j \triangleq x^s$ for all $j \neq k$ and x_k is $\mathbf{1}_{\{\theta \geq \theta_k\}}$ for some $\theta_k \in [0, 1]$. Then equation (A.36) delivers equation (A.32). And we already saw that (A.32) is equivalent to both equations (A.33) and (A.34). Combining these two equations yields

$$\max_{\theta \leq \widehat{\theta}} \left\{ \alpha_k \frac{\int_\theta^{\widehat{\theta}} \mu_k(z) f_k(z) dz}{\int_\theta^{\widehat{\theta}} \overline{F}_k(z) dz} - \sum_{j:j\neq k} \lambda_{jk} \cdot \frac{\int_\theta^{\widehat{\theta}} \overline{F}_j(z) dz}{\int_\theta^{\widehat{\theta}} \overline{F}_k(z) dz} \right\} \leq -\Big(\sum_{j:j\neq k} \lambda_{kj} \Big)$$

$$\leq \min_{\widehat{\theta} \leq \theta} \left\{ \alpha_k \frac{\int_{\widehat{\theta}}^{\theta} \mu_k(z) f_k(z) dz}{\int_{\widehat{\theta}}^{\theta} \overline{F}_k(z) dz} - \sum_{j:j\neq k} \lambda_{jk} \cdot \frac{\int_{\widehat{\theta}}^{\theta} \overline{F}_j(z) dz}{\int_{\widehat{\theta}}^{\theta} \overline{F}_k(z) dz} \right\},$$

that is, condition (APR^M) holds for any k. We only need to check that $\boldsymbol{\lambda} \in \mathcal{A}$. Observe that both the maximum and the minimum are bounded from below and above (respectively) by

$$\frac{\alpha_k \cdot \mu_k(\widehat{\theta}) \cdot f_k(\widehat{\theta}) - \sum_{j:j\neq k} \lambda_{jk} \cdot \overline{F}_j(\widehat{\theta})}{\overline{F}_k(\widehat{\theta})}. \tag{A.37}$$

To see this, we can take the limit as before. For the maximum we take the limit of θ approaching $\widehat{\theta}$ from below. This limit converges to the expression in (A.37) and is bounded above by the maximum. The same argument applies to the minimum but this time taking the limit from above $\widehat{\theta}$. This in turn implies that

$$\frac{\alpha_k \cdot \mu_k(\widehat{\theta}) \cdot f_k(\widehat{\theta}) - \sum_{j:j\neq k} \lambda_{jk} \cdot \overline{F}_j(\widehat{\theta})}{\overline{F}_k(\widehat{\theta})} = -\Big(\sum_{j:j\neq k} \lambda_{kj} \Big),$$

and we can conclude that $\boldsymbol{\lambda} \in \mathcal{A}$. □

Lemma A.4. *The set $\mathcal{B} \subset \mathcal{A}$ defined by*

$$\mathcal{B} \triangleq \Big\{ (\lambda_{ij})_{i,j \in \{1,\cdots,K\}^2} \geq 0 : \sum_{j\neq k} \lambda_{jk} \cdot \overline{F}_j(\widehat{\theta}) = \alpha_k \cdot \mu_k(\widehat{\theta}) \cdot f_k(\widehat{\theta}) + \overline{F}_k(\widehat{\theta}) \cdot \sum_{j\neq k} \lambda_{kj},$$
$$\alpha_k \geq \sum_{j\neq k} \lambda_{kj}, \quad \forall k \in \{1, \ldots, K\} \Big\},$$

is non-empty. Hence, the set $\mathcal{A}$ is non-empty.

Proof of Lemma A.4. We want to show that $\mathcal{B} \neq \emptyset$, which amount to proving that the linear system

$$\sum_{j=1, j\neq k}^{K} \lambda_{jk} \cdot \overline{F}_j(\widehat{\theta}) = \alpha_k \cdot \mu_k(\widehat{\theta}) \cdot f_k(\widehat{\theta}) + \overline{F}_k(\widehat{\theta}) \cdot \sum_{j=1, j\neq k}^{K} \lambda_{kj}, \quad \forall k \in \{1, \dots, K\},$$

$$\alpha_k = w_k + \sum_{j=1, j\neq k}^{K} \lambda_{kj} \quad \forall k \in \{1, \dots, K\},$$

with $(\boldsymbol{\lambda}, \mathbf{w}) \geq 0$ has a solution. We begin by writing down the system with matrices and then we apply Farkas' lemma.

First, the vector $\boldsymbol{\lambda}$ is given by

$$(\underbrace{\lambda_{12}, \lambda_{13}, \cdots, \lambda_{1K}}_{Type1}, \underbrace{\lambda_{21}, \lambda_{23}, \cdots, \lambda_{2K}}_{Type2}, \cdots, \underbrace{\lambda_{K1}, \lambda_{K2}, \cdots, \lambda_{KK-1}}_{TypeK}).$$

Note that the terms λ_{kk} for any $k \in \{1, \dots, K\}$ do not form part of the vector. Now, consider matrix A with $K(K-1)+K$ columns and $2K$ rows given by

$$A = \begin{bmatrix} \mathbf{F}^1 & \mathbf{F}^2 & \cdots & \mathbf{F}^K & 0_{K\times K} \\ B^1 & B^2 & \cdots & B^K & I_{K\times K} \end{bmatrix},$$

where $0_{K\times K}$ is the zero matrix of dimension $K \times K$ and $I_{K\times K}$ is the identity matrix of dimension $K \times K$. Furthermore, $\mathbf{F}^k$ and B^k are matrices of dimension $K \times (K-1)$ defined by

$$\mathbf{F}^k_{ij} = \begin{cases} -\overline{F}_k(\widehat{\theta}) & \text{if } i = k \\ \overline{F}_k(\widehat{\theta}) & \text{if } i < k, j = i \\ \overline{F}_k(\widehat{\theta}) & \text{if } i > k, j = i-1 \\ 0 & \text{if } o.w, \end{cases} \qquad B^k_{ij} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } o.w. \end{cases}$$

Finally, let b be a vector defined by $b = (\alpha_L \mu_1(\widehat{\theta}) f_1(\widehat{\theta}), \alpha_2 \mu_2(\widehat{\theta}) f_2(\widehat{\theta}), \cdots, \alpha_K \mu_K(\widehat{\theta}) f_K(\widehat{\theta}), \alpha_L, \cdots, \alpha_K)$. Then, the linear system can be rewritten as

$$A \cdot \begin{bmatrix} \boldsymbol{\lambda} \\ \mathbf{w} \end{bmatrix} = b, \quad \boldsymbol{\lambda}, \mathbf{w} \geq 0.$$

Now we use Farkas' lemma, if this system does not have a solution then it must be the case that the following system has a solution

$$A^\intercal \cdot \begin{bmatrix} y^F \\ y^B \end{bmatrix} \geq 0, \quad b^\intercal \cdot \begin{bmatrix} y^F \\ y^B \end{bmatrix} < 0. \tag{A.38}$$

Explicitly, we have (y^F, y^B) solve

$$\overline{F}_k(\widehat{\theta}) \cdot (y^F_j - y^F_k) + y^B_k \geq 0, \quad \forall k, \forall j \neq k$$

$$y^B_k \geq 0, \quad \forall k$$

$$\sum_{k=1}^{K} \alpha_k \mu_k(\widehat{\theta}) f_k(\widehat{\theta}) \cdot y^F_k + \sum_{k=1}^{K} \alpha_k \cdot y^B_k < 0.$$

Let y^F_m be equal to $\min_k\{y^F_k\}$ (m is the index that achieves the minimum) then

$$
\begin{aligned}
&\sum_{k=1}^{K}\alpha_k\mu_k(\widehat{\theta})f_k(\widehat{\theta})\cdot y_k^F+\sum_{k=1}^{K}\alpha_k\cdot y_k^B\\
&\overset{(a)}{=}\sum_{k=1}^{K}\alpha_k\mu_k(\widehat{\theta})f_k(\widehat{\theta})\cdot(y_k^F-y_m^F)+\sum_{k=1}^{K}\alpha_k\cdot y_k^B\\
&=\sum_{k=1}^{K}\alpha_k\left(\widehat{\theta}-\frac{\overline{F}_k(\widehat{\theta})}{f_k(\widehat{\theta})}\right)f_k(\widehat{\theta})\cdot(y_k^F-y_m^F)+\sum_{k=1}^{K}\alpha_k\cdot y_k^B\\
&=\sum_{k=1}^{K}\alpha_k\left(\widehat{\theta}f_k(\widehat{\theta})-\overline{F}_k(\widehat{\theta})\right)\cdot(y_k^F-y_m^F)+\sum_{k=1}^{K}\alpha_k\cdot y_k^B\\
&\overset{(b)}{\geq}-\sum_{k=1}^{K}\alpha_k\overline{F}_k(\widehat{\theta})\cdot(y_k^F-y_m^F)+\sum_{k=1}^{K}\alpha_k\cdot y_k^B\\
&=\sum_{k=1}^{K}\alpha_k\overline{F}_k(\widehat{\theta})\cdot(y_m^F-y_k^F)+\sum_{k=1}^{K}\alpha_k\cdot y_k^B\\
&\overset{(c)}{\geq}-\sum_{k=1}^{K}\alpha_k\cdot y_k^B+\sum_{k=1}^{K}\alpha_k\cdot y_k^B\\
&=0,
\end{aligned}
$$

a contradiction. Where in (a) we use the fact that $\sum_{k=1}^{K}\alpha_k\mu_k(\widehat{\theta})f_k(\widehat{\theta})=0$, in (b) we use the definition of y_m^F, and in (c) we use the first set of equations in (A.38). □

Proof of Proposition 5. We apply Theorem 3. For any k, consider the function

$$
L_k(z)\triangleq\frac{\alpha_k\mu_k(z)f_k(z)-\sum_{j\neq k}\lambda_{jk}\overline{F}_j(z)}{\overline{F}_k(z)}. \tag{A.39}
$$

We next show that under any of the two conditions in the statement of the proposition we can always find $\lambda\in\mathcal{A}$ such that (APR^M) holds. To prove this, it is enough to verify that (a) $L_k(z)\leq L_k(\widehat{\theta})$ for all $z\leq\widehat{\theta}$, and (b) $L_k(z)\geq L_k(\widehat{\theta})$ for all $z\geq\widehat{\theta}$, for some suitable $\lambda\in\mathcal{A}$, for all k. Indeed, if such λ exists then for any k, any $\theta_1\leq\widehat{\theta}$ and $\theta_2\geq\widehat{\theta}$ we have

$$
\begin{aligned}
\alpha_k\cdot R^{kk}(\theta_1,\widehat{\theta})-\sum_{j\neq k}\lambda_{jk}\cdot\frac{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_j(z)dz}{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_k(z)dz}
&=\frac{\int_{\theta_1}^{\widehat{\theta}}L_k(z)\overline{F}_k(z)dz}{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_k(z)dz}\\
&\leq\frac{\int_{\theta_1}^{\widehat{\theta}}L_k(\widehat{\theta})\overline{F}_k(z)dz}{\int_{\theta_1}^{\widehat{\theta}}\overline{F}_k(z)dz}\\
&=L_k(\widehat{\theta})\\
&\leq\frac{\int_{\widehat{\theta}}^{\theta_2}L_k(z)\overline{F}_k(z)dz}{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_k(z)dz}\\
&=\alpha_k\cdot R^{kk}(\widehat{\theta},\theta_2)-\sum_{j\neq k}\lambda_{jk}\cdot\frac{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_j(z)dz}{\int_{\widehat{\theta}}^{\theta_2}\overline{F}_k(z)dz},
\end{aligned}
$$

which is precisely (APR^M). The first inequality above comes from (a) and the second from (b).

To conclude we next verify conditions (a) and (b). We start by choosing $\lambda \in \mathcal{A}$ such that $\alpha_k \geq \sum_{j \neq k} \lambda_{jk}$, for all k. Lemma A.4 guarantees the existence of such λ. Next note that because $\lambda \in \mathcal{A}$ we have that $L_k(\widehat{\theta}) = -\sum_{j \neq k} \lambda_{kj}$ for all k. Hence, (a) is equivalent to

$$\alpha_k z f_k(z) - (\alpha_k - \sum_{j \neq k} \lambda_{kj}) \overline{F}_k(z) - \sum_{j \neq k} \lambda_{jk} \overline{F}_j(z) \leq 0, \quad \forall z \leq \widehat{\theta}, \ \forall k.$$

Note that $(\alpha_k - \sum_{j \neq k} \lambda_{kj}) \geq 0$ for all k. If condition *(i)* holds, we can divide the inequality above by $f_k(z)$ and use that $\overline{F}_j(z)/f_k(z)$ is non-increasing for any j (this is true under *(i)*) to conclude that the resulting function on the left-hand side is nondecreasing. If condition *(ii)* holds then because all $\overline{F}_j(z)$ are non-increasing functions and $z f_k(z)$ is nondecreasing then the resulting function on the left-hand side is nondecreasing. In conclusion the left-hand side in the equation above is bounded above by its value at $\widehat{\theta}$; however, since $\lambda \in \mathcal{A}$, this value equals zero. This establishes (a). Condition (b) can be verified in an analogous manner. □

Appendix B. Proofs for leading example: exponential distribution

This appendix contains the proofs for all the results related to the exponential distribution.

Proof of Lemma 3. From Lemma 2 we have that $\widehat{\theta}_L \leq \widehat{\theta} \leq \widehat{\theta_H}$. For exponential distributions, $\widehat{\theta}_L = 1/\lambda_L$ and $\widehat{\theta}_H = 1/\lambda_H$. Therefore, $\widehat{\theta} \in [1/\lambda_L, 1/\lambda_L]$. Moreover, $\widehat{\theta}$ must satisfy (8); if not, we could increase it or decrease it and obtain a strict revenue improvement.

We provide a proof for the rest of the properties for general distributions satisfying (IHR). Note first that $\widehat{\theta}$ can be seen as a function of α_L and α_H but since α_H equals $1 - \alpha_L$, we can effectively consider $\widehat{\theta}$ just a function of α_L. Then, when α_L equals 0 is as we only had type H buyers and, therefore, the optimal threshold is $\widehat{\theta}_H$. While when α_L equals 1 is as we only had type L buyers so the optimal threshold is $\widehat{\theta}_L$. Hence, $\widehat{\theta}(0)$ equals $\widehat{\theta}_H$ and $\widehat{\theta}(1)$ equals $\widehat{\theta}_L$.

Now we prove that $\widehat{\theta}(\alpha_L)$ is non-increasing. Consider $\alpha_L^a < \alpha_L^b$ and suppose that $\widehat{\theta}(\alpha_L^a) < \widehat{\theta}(\alpha_L^b)$. Define

$$\ell(\theta, \alpha_L) \triangleq \int_{\theta}^{\bar{\theta}} \alpha_L f_L(z) \mu_L(z) + (1 - \alpha_L) f_H(z) \mu_H(z) dz,$$

note that this is a linear function of α_L and, for fixed α_L, it is maximized at $\widehat{\theta}(\alpha_L)$. Hence,

$$\begin{aligned}
\ell(\widehat{\theta}(\alpha_L^a), \alpha_L^b) &\leq \ell(\widehat{\theta}(\alpha_L^b), \alpha_L^b) \\
&= \ell(\widehat{\theta}(\alpha_L^b), \alpha_L^b - \alpha_L^a) + \ell(\widehat{\theta}(\alpha_L^b), \alpha_L^a) \\
&\leq \ell(\widehat{\theta}(\alpha_L^b), \alpha_L^b - \alpha_L^a) + \ell(\widehat{\theta}(\alpha_L^a), \alpha_L^a)
\end{aligned}$$

therefore

$$\int_{\widehat{\theta}(\alpha_L^a)}^{\widehat{\theta}(\alpha_L^b)} \alpha_L^b f_L(z) \mu_L(z) + (1 - \alpha_L^b) f_H(z) \mu_H(z) dz$$

$$\leq \int\limits_{\widehat{\theta}(\alpha_L^a)}^{\widehat{\theta}(\alpha_L^b)} \alpha_L^a f_L(z)\mu_L(z) + (1-\alpha_L^a) f_H(z)\mu_H(z) dz. \tag{B.1}$$

Recall that $\widehat{\theta}$ is in $[\widehat{\theta}_L, \widehat{\theta}_H]$, and therefore, $\widehat{\theta}_L \leq \widehat{\theta}(\alpha_L^a) < \widehat{\theta}(\alpha_L^b) \leq \widehat{\theta}_H$. This in turn implies that

$$\mu_L(z) > 0 \quad \text{and} \quad \mu_H(z) < 0, \quad \forall z \in (\widehat{\theta}(\alpha_L^a), \widehat{\theta}(\alpha_L^b)),$$

hence for z in $(\widehat{\theta}(\alpha_L^a), \widehat{\theta}(\alpha_L^b))$ we have

$$\alpha_L^a f_L(z)\mu_L(z) + (1-\alpha_L^a) f_H(z)\mu_H(z) < \alpha_L^b f_L(z)\mu_L(z) + (1-\alpha_L^b) f_H(z)\mu_H(z),$$

which contradicts (B.1). □

Proof of Proposition 2. We make use of Theorem 1. Condition (APR) for the exponential distribution is

$$\max_{\theta \leq \widehat{\theta}} \left\{ \frac{\widehat{\theta} e^{-\lambda_L \widehat{\theta}} - \theta e^{-\lambda_L \theta}}{e^{-\lambda_H \widehat{\theta}} - e^{-\lambda_H \theta}} \right\} \leq \min_{\widehat{\theta} \leq \theta} \left\{ \frac{\theta e^{-\lambda_L \theta} - \widehat{\theta} e^{-\lambda_L \widehat{\theta}}}{e^{-\lambda_H \theta} - e^{-\lambda_H \widehat{\theta}}} \right\}. \tag{B.2}$$

Before we begin the proof, we need some definitions and observations. Define the following functions

$$\underline{g}(\theta) \triangleq \frac{\widehat{\theta} e^{-\lambda_L \widehat{\theta}} - \theta e^{-\lambda_L \theta}}{e^{-\lambda_H \widehat{\theta}} - e^{-\lambda_H \theta}} \quad \text{and} \quad \overline{g}(\theta) \triangleq \frac{\theta e^{-\lambda_L \theta} - \widehat{\theta} e^{-\lambda_L \widehat{\theta}}}{e^{-\lambda_H \theta} - e^{-\lambda_H \widehat{\theta}}}.$$

Note the following

$$\lim_{\theta \to \widehat{\theta}+} \overline{g}(\theta) = \lim_{\theta \to \widehat{\theta}-} \underline{g}(\theta) = \frac{(\lambda_L \widehat{\theta} - 1)}{\lambda_H} \cdot e^{-\widehat{\theta}(\lambda_L - \lambda_H)}, \tag{B.3}$$

and

$$\lim_{\theta \to \infty} \overline{g}(\theta) = \widehat{\theta} \cdot e^{-\widehat{\theta}(\lambda_L - \lambda_H)}. \tag{B.4}$$

Finally note that

$$\frac{(\lambda_L \widehat{\theta} - 1)}{\lambda_H} \cdot e^{-\widehat{\theta}(\lambda_L - \lambda_H)} \leq \widehat{\theta} \cdot e^{-\widehat{\theta}(\lambda_L - \lambda_H)} \iff \widehat{\theta} \leq \frac{1}{\lambda_L - \lambda_H}. \tag{B.5}$$

Now, suppose that condition (APR) holds and

$$\widehat{\theta} > \frac{1}{\lambda_L - \lambda_H} \tag{B.6}$$

From equations (B.3),(B.4) and (B.5) we see that

$$\overline{g}(\widehat{\theta}) = \underline{g}(\widehat{\theta}) > \lim_{\theta \to \infty} \underline{g}(\theta),$$

which implies

$$\max_{\theta \leq \widehat{\theta}} \left\{ \frac{\widehat{\theta} e^{-\lambda_L \widehat{\theta}} - \theta e^{-\lambda_L \theta}}{e^{-\lambda_H \widehat{\theta}} - e^{-\lambda_H \theta}} \right\} > \min_{\widehat{\theta} \leq \theta} \left\{ \frac{\theta e^{-\lambda_L \theta} - \widehat{\theta} e^{-\lambda_L \widehat{\theta}}}{e^{-\lambda_H \theta} - e^{-\lambda_H \widehat{\theta}}} \right\} \tag{B.7}$$

contradicting the fact that condition (APR) holds.

For the other direction, assume that equation (9) holds. We first prove that for $\theta \le \widehat{\theta}$ we have $\underline{g}(\theta) \le \underline{g}(\widehat{\theta})$; indeed,

$$\begin{aligned}\underline{g}(\theta) \le \underline{g}(\widehat{\theta}) &\iff \frac{\widehat{\theta}e^{-\lambda_L\widehat{\theta}} - \theta e^{-\lambda_L\theta}}{e^{-\lambda_H\widehat{\theta}} - e^{-\lambda_H\theta}} \le \frac{(\lambda_L\widehat{\theta}-1)}{\lambda_H} \cdot e^{-\widehat{\theta}(\lambda_L-\lambda_H)} \\ &\iff \lambda_H \cdot (\widehat{\theta}e^{-\lambda_L\widehat{\theta}} - \theta e^{-\lambda_L\theta}) \ge (e^{-\lambda_H\widehat{\theta}} - e^{-\lambda_H\theta}) \cdot (\lambda_L\widehat{\theta}-1) \cdot e^{-\widehat{\theta}(\lambda_L-\lambda_H)} \\ &\iff \lambda_H\widehat{\theta} \cdot (1 - \frac{\theta}{\widehat{\theta}}e^{-\lambda_L(\theta-\widehat{\theta})}) - (1 - e^{-\lambda_H(\theta-\widehat{\theta})}) \cdot (\lambda_L\widehat{\theta}-1) \ge 0,\end{aligned}$$

and hence we simply need to verify that this last inequality holds for $\theta \le \widehat{\theta}$. For doing so define

$$H(\theta) \triangleq \lambda_H\widehat{\theta} \cdot (1 - \frac{\theta}{\widehat{\theta}}e^{-\lambda_L(\theta-\widehat{\theta})}) - (1 - e^{-\lambda_H(\theta-\widehat{\theta})}) \cdot (\lambda_L\widehat{\theta}-1),$$

and note that $H(\widehat{\theta}) = 0$ and

$$H(0) = \lambda_H\widehat{\theta} + (e^{\lambda_H\widehat{\theta}} - 1) \cdot (\lambda_L\widehat{\theta}-1) \ge \lambda_H\widehat{\theta} + \lambda_H\widehat{\theta}(\lambda_L\widehat{\theta}-1) = \lambda_H\widehat{\theta} \cdot \lambda_L\widehat{\theta} > 0,$$

where the inequality comes from convexity of the exponential function and the fact that $\widehat{\theta} \ge 1/\lambda_L$. Furthermore the derivative of H is given by

$$\frac{dH}{d\theta} = \lambda_H(\lambda_L\theta - 1)e^{-\lambda_L(\theta-\widehat{\theta})} - \lambda_H(\lambda_L\widehat{\theta}-1)e^{-\lambda_H(\theta-\widehat{\theta})},$$

and it can be easily verified that for $\theta \le \widehat{\theta}$ we have $dH/d\theta \le 0$. This together with the facts that $H(0) > 0$ and $H(\widehat{\theta}) = 0$ imply that $\underline{g}(\theta) \le \underline{g}(\widehat{\theta})$ for all $\theta \le \widehat{\theta}$. This in turn implies

$$\max_{\theta \le \widehat{\theta}} \left\{ \frac{\widehat{\theta}e^{-\lambda_L\widehat{\theta}} - \theta e^{-\lambda_L\theta}}{e^{-\lambda_H\widehat{\theta}} - e^{-\lambda_H\theta}} \right\} = \frac{(\lambda_L\widehat{\theta}-1)}{\lambda_H} \cdot e^{-\widehat{\theta}(\lambda_L-\lambda_H)}.$$

Now we prove that for $\theta \ge \widehat{\theta}$ we have $\overline{g}(\theta) \ge \overline{g}(\widehat{\theta})$. Note that if we prove this we are done because this and what we have just proven imply condition (APR). As before we do

$$\begin{aligned}\overline{g}(\theta) \ge \overline{g}(\widehat{\theta}) &\iff \frac{\theta e^{-\lambda_L\theta} - \widehat{\theta}e^{-\lambda_L\widehat{\theta}}}{e^{-\lambda_H\theta} - e^{-\lambda_H\widehat{\theta}}} \ge \frac{(\lambda_L\widehat{\theta}-1)}{\lambda_H} \cdot e^{-\widehat{\theta}(\lambda_L-\lambda_H)} \\ &\iff \lambda_H(\widehat{\theta}e^{-\lambda_L\widehat{\theta}} - \theta e^{-\lambda_L\theta}) \ge (\lambda_L\widehat{\theta}-1) \cdot (e^{-\lambda_H\widehat{\theta}} - e^{-\lambda_H\theta}) \cdot e^{-\widehat{\theta}(\lambda_L-\lambda_H)} \\ &\iff \lambda_H(\widehat{\theta} - \theta e^{-\lambda_L(\theta-\widehat{\theta})}) - (\lambda_L\widehat{\theta}-1) \cdot (1 - e^{-\lambda_H(\theta-\widehat{\theta})}) \ge 0,\end{aligned}$$

note that the LHS of this last inequality is again the function $H(\cdot)$ but this time defined for $\theta \ge \widehat{\theta}$. We have $H(\widehat{\theta}) = 0$. It is easy to prove that for $\widehat{\theta} \le \theta \le \tilde{\theta}$ the function $H(\theta)$ is increasing, and then for $\theta > \tilde{\theta}$ is decreasing, where $\tilde{\theta} > \widehat{\theta}$ and $dH(\tilde{\theta})/d\theta = 0$. Additionally,

$$\lim_{\theta\to\infty} H(\theta) = \lambda_H\widehat{\theta} - (\lambda_L\widehat{\theta}-1) \ge 0.$$

Hence, for $\theta \ge \widehat{\theta}$, we have $H(\theta) \ge 0$, and therefore, $\overline{g}(\theta) \ge \overline{g}(\widehat{\theta})$ for all $\theta \ge \widehat{\theta}$, as desired. □

Proof of Corollary 1. Recall that for any $\lambda_L > \lambda_H$, from Lemma 3, we have

$$\frac{1}{\lambda_L} \le \widehat{\theta}(\alpha_L) \le \frac{1}{\lambda_H},$$

and

$$\lambda_L \le 2\lambda_H \iff \frac{1}{\lambda_H} \le \frac{1}{\lambda_L - \lambda_H},$$

therefore, for any $\alpha_L \in [0, 1]$ equation (9) is satisfied. Then by Proposition 2 we conclude that the static contract is optimal for any $\alpha_L \in [0, 1]$. □

Proof of Corollary 2. First, we show that $\widehat{\theta}(\cdot)$ is continuous from the right at zero. Let $\{\alpha_L^n\} \in [0, 1]$ be any sequence such that

$$\lim_{n\to\infty} \alpha_L^n = 0,$$

and suppose that $\widehat{\theta}(\alpha_L^n)$ does not converge to $\widehat{\theta}(0) = 1/\lambda_H$. That is,

$$\exists \epsilon > 0, \forall n_0, \exists n \ge n_0, \quad |\frac{1}{\lambda_H} - \widehat{\theta}(\alpha_L^n)| > \epsilon,$$

since $\widehat{\theta}(\alpha_L^n) \le \frac{1}{\lambda_H}$ we have

$$|\frac{1}{\lambda_H} - \widehat{\theta}(\alpha_L^n)| > \epsilon \iff \frac{1}{\lambda_H} - \widehat{\theta}(\alpha_L^n) > \epsilon.$$

This in turn means that we can create a subsequence $\{\alpha_L^{\ell_n}\} \subset \{\alpha_L^n\}$ such that

$$\forall n, \quad \frac{1}{\lambda_H} - \epsilon > \widehat{\theta}(\alpha_L^{\ell_n}). \tag{B.8}$$

However, since $\widehat{\theta}(\alpha_L^{\ell_n})$ is a maximizer of $\Pi^{\text{static}}(\cdot)$, we must have

$$\alpha_L^{\ell_n}\widehat{\theta}(\alpha_L^{\ell_n})e^{-\lambda_L\widehat{\theta}(\alpha_L^{\ell_n})} + (1-\alpha_L^{\ell_n})\widehat{\theta}(\alpha_L^{\ell_n})e^{-\lambda_H\widehat{\theta}(\alpha_L^{\ell_n})}$$
$$\ge \alpha_L^{\ell_n}\frac{1}{\lambda_H}e^{-\lambda_L\frac{1}{\lambda_H}} + (1-\alpha_L^{\ell_n})\frac{1}{\lambda_H}e^{-\lambda_H\frac{1}{\lambda_H}},$$

because $\lambda_L > \lambda_H$ we can bound the LHS above to obtain

$$\widehat{\theta}(\alpha_L^{\ell_n})e^{-\lambda_H\widehat{\theta}(\alpha_L^{\ell_n})} \ge \alpha_L^{\ell_n}\frac{1}{\lambda_H}e^{-\lambda_L\frac{1}{\lambda_H}} + (1-\alpha_L^{\ell_n})\frac{1}{\lambda_H}e^{-\lambda_H\frac{1}{\lambda_H}}. \tag{B.9}$$

Note that the function $\theta e^{-\lambda_H\theta}$ has a unique maximum at $\theta = 1/\lambda_H$ and since $\widehat{\theta}(\alpha_L^{\ell_n})$ satisfies equation (B.8), we can always find $\delta(\epsilon) > 0$ such that

$$\Big(\frac{1}{\lambda_H} + \delta(\epsilon)\Big)e^{-\lambda_H(\frac{1}{\lambda_H}+\delta(\epsilon))} > \widehat{\theta}(\alpha_L^{\ell_n})e^{-\lambda_H\widehat{\theta}(\alpha_L^{\ell_n})}, \quad \forall n,$$

plugging this into equation (B.9) yields

$$\Big(\frac{1}{\lambda_H} + \delta(\epsilon)\Big)e^{-\lambda_H(\frac{1}{\lambda_H}+\delta(\epsilon))} > \alpha_L^{\ell_n}\frac{1}{\lambda_H}e^{-\lambda_L\frac{1}{\lambda_H}} + (1-\alpha_L^{\ell_n})\frac{1}{\lambda_H}e^{-\lambda_H\frac{1}{\lambda_H}}, \quad \forall n,$$

therefore, taking the limit over n gives a contradiction. In conclusion, we have proved that $\widehat{\theta}(\cdot)$ is continuous from the right at zero. Now, to finalize the proof, recall that we are assuming that $\lambda_L > 2\lambda_H$ or equivalently $\frac{1}{\lambda_H} > \frac{1}{\lambda_L-\lambda_H}$. However, since $\widehat{\theta}(0) = 1/\lambda_H$ and $\widehat{\theta}(\cdot)$ is continuous from the right, we can always find $\bar{\alpha}_L \in (0, 1]$ such that

$$\frac{1}{\lambda_H} \ge \widehat{\theta}(\bar{\alpha}_L) \ge \frac{1}{\lambda_L - \lambda_H},$$

so thanks to Proposition 2, the sequential contract is optimal when we set $\alpha_L > \bar{\alpha}_L$. Note that the same argument is valid for $1/\lambda_L$. That is, we can show that $\widehat{\theta}(\alpha_L)$ is continuous from the left at 1 and then using the fact that

$$\frac{1}{\lambda_L - \lambda_H} > \frac{1}{\lambda_L},$$

we can find $\bar{\alpha}_H \in [\bar{\alpha}_L, 1)$ such that

$$\frac{1}{\lambda_L - \lambda_H} > \widehat{\theta}(\bar{\alpha}_H) \geq \frac{1}{\lambda_L}.$$

Hence, in $[\bar{\alpha}_H, 1]$, the static contract is optimal. All of this implies that since $\widehat{\theta}(\cdot)$ is a nonincreasing function, we can always find $\bar{\alpha} \in (0, 1)$ with the desired property. □

Proof of Corollary 3. Fix λ_H and α_L. Suppose the result is not true, that is,

$$\forall \bar{\lambda}_L \geq 2\lambda_H, \exists \lambda_L \geq \bar{\lambda}_L, \quad \widehat{\theta}(\lambda_L) \leq \frac{1}{\lambda_L - \lambda_H}.$$

From this we can construct a sequence $\lambda_L^n \geq 2\lambda_H$ such that

$$\lim_{n\to\infty} \lambda_L^n = \infty \quad \text{and} \quad \widehat{\theta}(\lambda_L^n) \leq \frac{1}{\lambda_L^n - \lambda_H}, \qquad \forall n \in \mathbb{N},$$

therefore $\widehat{\theta}(\lambda_L^n)$ converges to 0, and we have

$$\Pi^{\text{static}}(\widehat{\theta}(\lambda_L^n)) = \widehat{\theta}(\lambda_L^n) e^{-\lambda_H \widehat{\theta}(\lambda_L^n)} \left(\alpha_L e^{-(\lambda_L^n - \lambda_H)\widehat{\theta}(\lambda_L^n)} + \alpha_H\right) \leq \widehat{\theta}(\lambda_L^n) e^{-\lambda_H \widehat{\theta}(\lambda_L^n)} \overset{n\to\infty}{\to} 0.$$

However, since $\widehat{\theta}(\lambda_L^n)$ maximizes $\Pi^{\text{static}}(\cdot)$ it must be the case that

$$\Pi^{\text{static}}(1/\lambda_H) \leq \Pi^{\text{static}}(\widehat{\theta}(\lambda_L^n)),$$

that is,

$$\alpha_L \frac{1}{\lambda_H} e^{-\lambda_L^n \frac{1}{\lambda_H}} + \alpha_H \frac{1}{\lambda_H} e^{-\lambda_H \frac{1}{\lambda_H}} \leq \Pi^{\text{static}}(\widehat{\theta}(\lambda_L^n)).$$

Taking the limit over n on both sides of the previous equation yields

$$\alpha_H \frac{1}{\lambda_H} e^{-\lambda_H \frac{1}{\lambda_H}} \leq 0,$$

a contradiction. □

Proof of Proposition 4. We use the sufficient conditions in Lemma B.1 (which we state and proof after the present prove). First note that since the support of the exponential distribution is unbounded from above, we can take $\theta_2 = \infty$ which eliminates condition (1). Conditions (2) and (3) can be cast as

$$\begin{aligned} &\theta_1 e^{-\theta_1(\lambda_L - \lambda_H)} \geq \theta e^{-\theta(\lambda_L - \lambda_H)} \quad \forall \theta \geq 0 \quad \text{and} \\ &\alpha_L \cdot \lambda_H \theta_1 e^{-\theta_1(\lambda_L - \lambda_H)} = -\alpha_H \cdot (\lambda_H \theta_H - 1). \end{aligned} \tag{B.10}$$

By optimizing the first term in (B.10), we obtain

$$\theta_1 = \frac{1}{\lambda_L - \lambda_H},$$

and then solving for θ_H yields

$$\theta_H = \frac{1}{\lambda_H} - \frac{\alpha_L}{\alpha_H}\frac{e^{-1}}{\lambda_L - \lambda_H}.$$

What we need to check is that $\theta_1 \le \theta_H$. First, we show that

$$Q \triangleq \alpha_L(\theta_1 - \frac{1}{\lambda_L})\lambda_L e^{-\lambda_L\theta_1} + \alpha_H(\theta_1 - \frac{1}{\lambda_H})\lambda_H e^{-\lambda_H\theta_1} < 0. \tag{B.11}$$

To prove this inequality, note that since $\widehat{\theta}$ is the optimal static cutoff, we have

$$\alpha_L\widehat{\theta}e^{-\lambda_L\widehat{\theta}} + \alpha_H\widehat{\theta}e^{-\lambda_H\widehat{\theta}} \ge \alpha_L\theta_1 e^{-\lambda_L\theta_1} + \alpha_H\theta_1 e^{-\lambda_H\theta_1}. \tag{B.12}$$

Then, we have

$$\begin{aligned}
Q &= \alpha_L\theta_1(\lambda_L - \lambda_H)e^{-\lambda_L\theta_1} + \alpha_L\theta_1\lambda_H e^{-\lambda_L\theta_1} + \alpha_H\theta_1\lambda_H e^{-\lambda_H\theta_1} - \alpha_L e^{-\lambda_L\theta_1} - \alpha_H e^{-\lambda_H\theta_1} \\
&= \alpha_L e^{-\lambda_L\theta_1} + \lambda_H(\alpha_L\theta_1 e^{-\lambda_L\theta_1} + \alpha_H\theta_1 e^{-\lambda_H\theta_1}) - \alpha_L e^{-\lambda_L\theta_1} - \alpha_H e^{-\lambda_H\theta_1} \\
&\overset{(a)}{\le} \lambda_H(\alpha_L\widehat{\theta}e^{-\lambda_L\widehat{\theta}} + \alpha_H\widehat{\theta}e^{-\lambda_H\widehat{\theta}}) - \alpha_H e^{-\lambda_H\theta_1} \\
&\overset{(b)}{<} \lambda_H(\alpha_L\widehat{\theta}e^{-\lambda_L\widehat{\theta}} + \alpha_H\widehat{\theta}e^{-\lambda_H\widehat{\theta}}) - \alpha_H e^{-\lambda_H\widehat{\theta}} \\
&= \lambda_H\alpha_L\widehat{\theta}e^{-\lambda_L\widehat{\theta}} + \lambda_H\alpha_H e^{-\lambda_H\widehat{\theta}}(\widehat{\theta} - \frac{1}{\lambda_H}) \\
&\overset{(c)}{=} \lambda_H\alpha_L\widehat{\theta}e^{-\lambda_L\widehat{\theta}} - \lambda_L\alpha_L e^{-\lambda_L\widehat{\theta}}(\widehat{\theta} - \frac{1}{\lambda_L}) \\
&= \alpha_L e^{-\lambda_L\widehat{\theta}}\Big(-\widehat{\theta}(\lambda_L - \lambda_H) + 1\Big) \\
&\overset{(d)}{<} 0,
\end{aligned}$$

where (a) comes from equation (B.12), (b) is true because the function $-e^{-\lambda_H\theta}$ increasing and $\theta_1 < \widehat{\theta}$, and (c) comes from equation (8). Moreover, (d) comes from $\theta_1 < \widehat{\theta}$. With this, we have proven (B.11), and thus

$$\begin{aligned}
\lambda_L\alpha_H \cdot (\theta_H - \frac{1}{\lambda_H}) &\overset{(a)}{=} -\lambda_L\alpha_L \cdot \theta_1 e^{-\theta_1(\lambda_L - \lambda_H)} \\
&= -\lambda_L\alpha_L \cdot \Big(\theta_1 - \frac{1}{\lambda_L}\Big)e^{-\theta_1(\lambda_L - \lambda_H)} - \lambda_L\alpha_L \cdot \frac{1}{\lambda_L}e^{-\theta_1(\lambda_L - \lambda_H)} \\
&\overset{(b)}{>} \alpha_H(\theta_1 - \frac{1}{\lambda_H})\lambda_H - \alpha_L \cdot e^{-\theta_1(\lambda_L - \lambda_H)} \\
&\overset{(c)}{=} \alpha_H(\theta_1 - \frac{1}{\lambda_H})\lambda_H + \frac{\alpha_H}{\theta_1} \cdot (\theta_H - \frac{1}{\lambda_H}),
\end{aligned}$$

where in (a) and (c) we used the definition of θ_H, and in (b) we used equation (B.11). From this we have that

$$(\theta_H - \frac{1}{\lambda_H}) \cdot \Big(\lambda_L\alpha_H - \frac{\alpha_H}{\theta_1}\Big) > \alpha_H(\theta_1 - \frac{1}{\lambda_H})\lambda_H,$$

but replacing θ_1 with $1/(\lambda_L - \lambda_H)$ in this last expression we get $\theta_H > \theta_1$.

Finally, x is given by

$$x = \frac{\int_{\theta_H}^{\theta_3} \overline{F}_H(z)dz}{\int_{\theta_1}^{\theta_3} \overline{F}_H(z)dz} = \frac{e^{-\lambda_H \theta_H}}{e^{-\lambda_H \theta_1}} = \exp\Big(-\lambda_H \Big[\frac{1}{\lambda_H} - \frac{\alpha_L}{\alpha_H} \frac{e^{-1}}{\lambda_L - \lambda_H} - \frac{1}{\lambda_L - \lambda_H} \Big] \Big). \quad \square$$

Lemma B.1. *The following conditions for the thresholds $\theta_1 \le \theta_H \le \theta_2$ (as in Proposition 3) are sufficient for their optimality in $(\mathcal{P}_R)$:*

1. $R^{LH}(\theta_1, \theta_2) \le \min_{\theta_2 \le \theta} R^{LH}(\theta_2, \theta)$;
2. $\max_{\theta \le \theta_2} R^{LH}(\theta, \theta_2) \le R^{LH}(\theta_1, \theta_2)$;
3. $\alpha_L \cdot R^{LH}(\theta_1, \theta_2) + \alpha_H r^{HH}(\theta_H) = 0$.

Proof of Lemma B.1. It is enough to prove that under these conditions, the optimal contract characterized by $(\theta_1, \theta_H, \theta_2)$ is optimal for $(\mathcal{P}_R)$. To prove this we use a Lagrangian relaxation (we do not relax the monotonicity constraints) and show that this relaxation is optimized by the contract characterized by $(\theta_1, \theta_H, \theta_2)$.

First, we establish some properties that can be derived from conditions (1) to (3). Condition (3) implies that $\theta_2 \ge \widehat{\theta}_L$; otherwise, $\theta_1, \theta_2 < \widehat{\theta}_L$ which would imply that $R^{LH}(\theta_1, \theta_2) < 0$. In turn, condition (3) would give $R^{HH}(\theta_H) > 0$ which would imply that $\widehat{\theta}_H < \theta_H$. Since $\theta_H \le \theta_2$ we would have $\widehat{\theta}_H < \theta_H \le \theta_2 < \widehat{\theta}_L$, that is, $\widehat{\theta}_H < \widehat{\theta}_L$ which is not possible. Moreover, condition (2) together with the fact that $\theta_2 \ge \widehat{\theta}_L$ imply that $\theta_1 \ge \widehat{\theta}_L$. This yields $R^{LH}(\theta_1, \theta_2) \ge 0$, and thus we can use condition (3) again to deduce that $\theta_H \le \widehat{\theta}_H$. In summary, $\widehat{\theta}_L \le \theta_1$ and $\theta_H \le \widehat{\theta}_H$.

Now, we provide the main argument. If $\theta_1 = \theta_2$, then we also have $\theta_1 = \theta_2 = \theta_H$. Condition (3) implies that the contract characterized by $(\theta_1, \theta_H, \theta_2)$ is the static contract. Conditions (1) and (2) together yield (APR), and therefore, from Theorem 1, we deduce that the static contract is optimal. Next suppose that $\theta_1 < \theta_2$, and define

$$\Omega \triangleq \{x : [0, \bar{\theta}] \to [0, 1] : x(\cdot) \text{ is nondecreasing}\}.$$

We use $\mathbf{x}^\star$ to denote the solution characterized by $(\theta_1, \theta_H, \theta_2)$. The Lagrangian for $(\mathcal{P}_R)$ is

$$\begin{aligned}\mathcal{L}(\boldsymbol{u}, \mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{w}) &= u_L(w_L - \lambda - \alpha_L) + u_H(\lambda - \alpha_H + w_H) \\ &+ \int_0^{\bar{\theta}} x_L(z) \cdot \Big[\alpha_L \mu_L(z) f_L(z) - \lambda \overline{F}_H(z)(z) \Big] dz \\ &+ \int_0^{\bar{\theta}} x_H(z) \cdot \Big[\alpha_H \mu_H(z) f_H(z) + \lambda \overline{F}_H(z) \Big] dz.\end{aligned}$$

Consider the following multipliers

$$\lambda = \alpha_L \cdot R^{LH}(\theta_1, \theta_2), \quad w_L = \lambda + \alpha_L, w_H = -\lambda + \alpha_H.$$

Note that λ and w_L are nonnegative, and for w_H we have

$$\begin{aligned}w_H \ge 0 &\Leftrightarrow \alpha_H + \alpha_H r^{HH}(\theta_H) \ge 0 \Leftrightarrow r^{HH}(\theta_H) \\ &\ge -1 \Leftrightarrow [\theta_H - h^{HH}(\theta_H)] \ge -h^{HH}(\theta_H) \Leftrightarrow \theta_H \ge 0,\end{aligned}$$

where in the first if and only if we made use of condition (3) above. Thus when we optimize the Lagrangian we obtain:

$$\max_{(\boldsymbol{u},\mathbf{x})\in\Omega} \mathcal{L}(\mathbf{x},\boldsymbol{u},\boldsymbol{\lambda},\boldsymbol{w}) = \max_{0\le\theta\le\bar{\theta}} \int_{\theta}^{\bar{\theta}} \Big[\alpha_1\mu_1(z)f_1(z) - \lambda\overline{F}_2(z)\Big]dz + \max_{0\le\theta\le\bar{\theta}} \int_{\theta}^{\bar{\theta}} \Big[\alpha_H\mu_H(z)f_H(z) + \lambda\overline{F}_H(z)\Big]dz\ , \tag{B.13}$$

where we can reduce attention to threshold strategies because $x_L(\cdot), x_H(\cdot)$ are nondecreasing (see, e.g., Myerson (1981) or Riley and Zeckhauser (1983)). If we are able to show that $\mathcal{L}(\mathbf{x},\boldsymbol{u},\boldsymbol{\lambda},\boldsymbol{w})$ evaluated at our candidate solution is an upper bound for the RHS above we are done. Let us begin with the second term. Take any $0\le\theta\le\bar{\theta}$; then,

$$\begin{aligned}
\int_{\theta}^{\bar{\theta}} \Big[\alpha_H\mu_H(z)f_H(z) + \lambda\overline{F}_H(z)\Big]dz &= \int_{\theta}^{\bar{\theta}} \Big[\alpha_H\mu_H(z)f_H(z) - \alpha_H r^{HH}(\theta_H)\overline{F}_H(z)\Big]dz \\
&= \int_{\theta}^{\bar{\theta}} \alpha_H\overline{F}_H(z)\Big[r^{HH}(z) - r^{HH}(\theta_H)\Big]dz \\
&\le \int_{\theta_H}^{\bar{\theta}} \alpha_H\overline{F}_H(z)\Big[r^{HH}(z) - r^{HH}(\theta_H)\Big]dz \\
&= \int_{0}^{\bar{\theta}} x_H^{\star}(z)\Big[\alpha_H\mu_H(z)f_H(z) + \lambda\overline{F}_H(z)\Big]dz,
\end{aligned}$$

where in the first equality we used condition (3) and the inequality comes from the fact that $r^{HH}(\cdot)$ is nondecreasing. Now we look into the first term in equation (B.13), consider first $\theta\ge\theta_2$

$$\begin{aligned}
\int_{\theta}^{\bar{\theta}} \Big[\alpha_L\mu_L(z)f_L(z) - \lambda\overline{F}_H(z)\Big]dz &= \int_{\theta_L^2}^{\bar{\theta}} \Big[\alpha_L\mu_L(z)f_L(z) - \lambda\overline{F}_H(z)\Big]dz \\
&\quad - \int_{\theta_L^2}^{\theta} \Big[\alpha_L\mu_L(z)f_L(z) - \lambda\overline{F}_H(z)\Big]dz \\
&\le \int_{\theta_L^2}^{\bar{\theta}} \Big[\alpha_L\mu_L(z)f_L(z) - \lambda\overline{F}_H(z)\Big]dz,
\end{aligned}$$

where we have used the following

$$-\int_{\theta_2}^{\theta}\Big[\alpha_L\mu_L(z)f_L(z)-\lambda\overline{F}_2(z)\Big]dz\le 0 \Leftrightarrow \alpha_L\cdot\frac{\int_{\theta_1}^{\theta_2}\overline{F}_H(z)r^{LH}(z)dz}{\int_{\theta_1}^{\theta_2}\overline{F}_H(z)dz}$$
$$=\lambda\le\alpha_L\cdot\frac{\int_{\theta_2}^{\theta}\overline{F}_2(z)r^{LH}(z)dz}{\int_{\theta_2}^{\theta}\overline{F}_H(z)dz},$$

which thanks to condition (1) in our hypothesis is true. A similar argument holds for $\theta\le\theta_2$, but using condition (2). Since $\mathcal{L}(\mathbf{x}^\star,0,\boldsymbol{\lambda},\boldsymbol{w})$ equals

$$x\int_{\theta_1}^{\theta_2}\Big[\alpha_1\mu_1(z)f_1(z)-\lambda\overline{F}_H(z)\Big]dz+\int_{\theta_2}^{\bar\theta}\Big[\alpha_1\mu_1(z)f_1(z)-\lambda\overline{F}_H(z)\Big]dz$$
$$+\int_{\theta_H}^{\bar\theta}\Big[\alpha_H\mu_H(z)f_H(z)+\lambda\overline{F}_H(z)\Big]dz,$$

which by the definition of λ simplifies to

$$\int_{\theta_2}^{\bar\theta}\Big[\alpha_1\mu_1(z)f_1(z)-\lambda\overline{F}_H(z)\Big]dz+\int_{\theta_H}^{\bar\theta}\Big[\alpha_H\mu_H(z)f_H(z)+\lambda\overline{F}_H(z)\Big]dz,$$

we conclude that $\max_{(\boldsymbol{u},\mathbf{x})\in\Omega}\mathcal{L}(\boldsymbol{u},\mathbf{x},\boldsymbol{\lambda},\boldsymbol{w})\le\mathcal{L}(0,\mathbf{x}^\star,\boldsymbol{\lambda},\boldsymbol{w})$, as required. □

Proof of Proposition 6. We make use of Lemma B.2 which we state and prove after the present proof. In that lemma we need to define the function

$$L_k(z|\boldsymbol{\lambda})\triangleq\alpha_k\mu_k(z)+\frac{\bar F_k(z)}{f_k(z)}\cdot\sum_{\ell:\ell\neq k}\lambda_{k\ell}-\sum_{\ell:\ell\neq k}\lambda_{\ell k}\frac{\bar F_\ell(z)}{f_k(z)},$$

for any $\boldsymbol{\lambda}\ge 0$. For exponential distributions $L_k(z|\boldsymbol{\lambda})$ becomes:

$$L_k(z|\boldsymbol{\lambda})=\underbrace{\alpha_k\cdot z+\frac{1}{\lambda_k}\cdot\Big(\sum_{\ell:\ell\neq k}\lambda_{k\ell}-\alpha_k\Big)}_{\text{linear}}-\underbrace{\sum_{\ell:\ell>k}\lambda_{\ell k}\frac{e^{-z(\lambda_l-\lambda_k)}}{\lambda_k}}_{\text{increasing and convex}}-\underbrace{\sum_{\ell:\ell<k}\lambda_{\ell k}\frac{e^{-z(\lambda_l-\lambda_k)}}{\lambda_k}}_{\text{decreasing and convex}}.$$

Hence, $L_k(\cdot|\boldsymbol{\lambda})$ is concave, which means that it crosses zero at most two times. Using Lemma B.2 we conclude that in the exponential case allocations have at most one step in which randomization occurs. □

Lemma B.2. *For any dual-feasible variable $\boldsymbol{\lambda}$ associated with the incentive constraints, define*

$$L_k(z|\boldsymbol{\lambda})\triangleq\alpha_k\mu_k(z)+\frac{\bar F_k(z)}{f_k(z)}\cdot\sum_{\ell:\ell\neq k}\lambda_{k\ell}-\sum_{\ell:\ell\neq k}\lambda_{\ell k}\frac{\bar F_\ell(z)}{f_k(z)}.\tag{F}$$

If $L_k(z|\boldsymbol{\lambda})$ crosses zero at most p, times then the optimal allocation x_k has at most $\lfloor p/2\rfloor$ intervals where randomization occurs.

Proof of Lemma B.2. We divide the proof into two parts. In the first part, we construct a new dual problem and state the complementary slackness conditions. This part of the proof follows the general theory of linear programming in infinite dimensional space developed by Anderson and Nash (1987). In the second part we exploit the complementary slackness conditions to show that the optimal allocation x_k has at most $\lfloor p/2 \rfloor$ intervals where randomization occurs.

Part 1. Define the cone of nonnegative nondecreasing functions

$$\mathcal{K} \triangleq \{x : [0, \bar{\theta}] \to \mathbb{R} | x \text{ is nonnegative and nondecreasing function}\}.$$

The general formulation of the seller's problem is

$$
\begin{aligned}
(\mathcal{P}) \quad \max \quad & -\sum_{k=1}^{K} \alpha_k u_k + \sum_{k=1}^{K} \alpha_k \int_0^{\bar{\theta}} x_k(z) \mu_k(z) f_k(z) dz \\
\text{s.t} \quad & x_k(\cdot) \in \mathcal{K}, \quad \forall k \in \{1, \dots, K\} \\
& x_k(\theta) \leq 1, \quad \forall \theta \in [0, \bar{\theta}] \quad, \forall k \in \{1, \dots, K\} \\
& u_k \geq 0, \quad \forall k \in \{1, \dots, K\} \\
& u_k + \int_0^{\bar{\theta}} x_k(z) \bar{F}_k(z) dz \geq u_{k'} + \int_0^{\bar{\theta}} x_{k'}(z) \bar{F}_k(z) dz, \quad \forall k, k' \in \{1, \dots, K\}.
\end{aligned}
$$

Note that the dual cone of $\mathcal{K}$ is

$$\mathcal{K}^* = \{\beta : \int_{\theta}^{\bar{\theta}} \beta(z) dz \geq 0, \quad \forall \theta \in [0, \bar{\theta}]\}.$$

The Lagrangian is

$$
\begin{aligned}
& \mathcal{L}(x, u, \boldsymbol{\lambda}, \boldsymbol{\beta}, \mathbf{w}) \\
& = \sum_{k=1}^{K} u_k \cdot \Big(-\alpha_k + w_k + \sum_{\ell:\ell \neq k} \lambda_{k\ell} - \sum_{\ell:\ell \neq k} \lambda_{\ell k} \Big) \\
& + \sum_{k=1}^{K} \int_0^{\bar{\theta}} x_k(z) \Big(\alpha_k \mu_k(z) f_k(z) + \bar{F}_k(z) \cdot \sum_{\ell:\ell \neq k} \lambda_{k\ell} - \sum_{\ell:\ell \neq k} \lambda_{\ell k} \bar{F}_\ell(z) + \beta_k(z) - \eta_k(z) \Big) dz \\
& + \sum_{k=1}^{K} \int_0^{\bar{\theta}} \eta_k(z) dz,
\end{aligned}
$$

where β_k are the dual variables associated with the monotonicity constraints, η_k are dual variables associated with the constraints $x_k(\theta) \leq 1$, and $\boldsymbol{\lambda}$, $\boldsymbol{w}$ correspond to the dual variables associated with the incentive an non-negativity constraints respectively. This yields the following dual program (D):

$$(D) \quad \min \quad \sum_{k=1}^{K} \int_0^{\bar{\theta}} \eta_k(z) dz$$

$$\text{s.t} -\alpha_k + w_k + \sum_{\ell:\ell\neq k} \lambda_{k\ell} - \sum_{\ell:\ell\neq k} \lambda_{\ell k} = 0, \quad \forall k$$

$$\alpha_k \mu_k(z) f_k(z) + \bar{F}_k(z) \cdot \sum_{\ell:\ell\neq k} \lambda_{k\ell}$$

$$- \sum_{\ell:\ell\neq k} \lambda_{\ell k} \bar{F}_\ell(z) = \eta_k(z) - \beta_k(z), \quad \forall k, \quad \forall z \in [0, \bar{\theta}]$$

$$\boldsymbol{\lambda}, \mathbf{w}, \eta_k(\cdot) \geq 0, \quad \beta_k \in \mathcal{K}^*, \quad \forall k.$$

We must have complementary slackness. That is, for the monotonicity constraints (the cone constraints) this means that if $x_k(\cdot)$ changes at some θ, then $\int_\theta^{\bar{\theta}} \beta_k(z)dz = 0$. Moreover, $x(0) \cdot \int_0^{\bar{\theta}} \beta(z)dz = 0$. All of this for all k. For the upper bound constraints, we must have $(1 - x_k(\theta)) \cdot \eta_k(\theta) = 0$ for all $\theta \in [0, \bar{\theta}]$ and for all k.

Part 2. Consider an optimal primal-dual pair. Let x_k be the primal solution for interim type k, and β_k, η_k and $\boldsymbol{\lambda}, \mathbf{w}$ be the corresponding dual solutions. Observe that from dual feasibility, we must have

$$f_k(z) \cdot L_k(z|\boldsymbol{\lambda}) = \eta_k(z) - \beta_k(z), \quad \forall z \in [0, \bar{\theta}]. \tag{B.14}$$

Let us denote by $\hat{z}_1 < \cdots < \hat{z}_p$ the points where $L_k(\cdot|\boldsymbol{\lambda})$ crosses zero, and we let $\hat{z}_0 = 0$ and $\hat{z}_{p+1} = \bar{\theta}$. Note that $L_k(\bar{\theta}|\boldsymbol{\lambda}) = \alpha \cdot \bar{\theta} > 0$, and by the feasibility of $\boldsymbol{\lambda}$ we have $L_k(0|\boldsymbol{\lambda}) = -w_k/f_k(0) \leq 0$.

Let $z_1^\star \triangleq \inf\{z \in [0, \bar{\theta}] : x_k(z) = 1\}$ (if $x_k(z)$ never equals 1 we take $z_1^\star = \bar{\theta}$). We can assume that $z_1^\star > 0$; otherwise, $x_k(z)$ would be equal to 1 everywhere in $[0, \bar{\theta}]$ and the result would follow. In turn, there has to be a change in x_k around $z_1^\star$, and therefore, complementary slackness implies that $\int_{z_1^\star}^{\bar{\theta}} \beta_k(z)dz = 0$. Moreover, since $x_k(z) < 1$ for all $z < z_1^\star$, complementary slackness implies that $\eta_k(z) = 0$ for all $z < z_1^\star$. Therefore, Eq. (B.14) becomes

$$f_k(z) \cdot L_k(z|\boldsymbol{\lambda}) = -\beta_k(z), \quad \forall z \in [0, z_1^\star). \tag{B.15}$$

Let q be the largest index in $\{0, 1, \ldots, p\}$ such that $\hat{z}_q \leq z_1^\star$. Note that $z_1^\star \in [\hat{z}_q, \hat{z}_{q+1}]$. We show the following claim:

Claim 1. $L_k(\cdot|\boldsymbol{\lambda})$ is positive in $(\hat{z}_q, \hat{z}_{q+1})$ and $z_1^\star = \hat{z}_q$.

Proof of Claim 1. First, suppose that $L_k(\cdot|\boldsymbol{\lambda})$ is positive in $(\hat{z}_q, \hat{z}_{q+1})$; we show that $z_1^\star = \hat{z}_q$. If not, then for any $z \in (\hat{z}_q, z_1^\star)$, we have $L_k(z|\boldsymbol{\lambda}) > 0$, which thanks to Eq. (B.15) yields $\beta_k(z) < 0$ for any $z \in (\hat{z}_q, z_1^\star)$, and therefore,

$$\int_z^{\bar{\theta}} \beta_k(z)dz = \int_z^{z_1^\star} \beta_k(z)dz + \underbrace{\int_{z_1^\star}^{\bar{\theta}} \beta_k(z)dz}_{=0} = \int_z^{z_1^\star} \beta_k(z)dz < 0, \tag{B.16}$$

but, this contradicts the fact that $\beta_k \in \mathcal{K}^*$. That is, $z_1^\star \leq \hat{z}_q$ but since $\hat{z}_q \leq z_1^\star$ we conclude that $\hat{z}_q = z_1^\star$. To complete the argument, suppose that $L_k(\cdot|\boldsymbol{\lambda})$ is negative in $(\hat{z}_q, \hat{z}_{q+1})$ then, in particular, $L_k(\cdot|\boldsymbol{\lambda})$ is negative in $(z_1^\star, \hat{z}_{q+1})$, and from Eq. (B.14), we deduce that $\beta_k(z') > 0$ for all $z_1'^\star, \hat{z}_{q+1})$. Hence, for any $z_1'^\star, \hat{z}_{q+1})$

$$0 = \int_{z_1^\star}^{\bar{\theta}} \beta_k(z)dz = \underbrace{\int_{z_1^\star}^{z'} \beta_k(z)dz}_{>0} + \underbrace{\int_{z'}^{\bar{\theta}} \beta_k(z)dz}_{\geq 0} > 0, \tag{B.17}$$

a contradiction. In the second bracket, we use the fact that $\beta_k \in \mathcal{K}^*$. This concludes the proof of Claim 1.

This shows that $x_k(\cdot)$ equals 1 in $(\hat{z}_q, \bar{\theta}]$ and that it changes value at $\hat{z}_q$. Now, from Claim 1, we know that $L_k(\cdot|\boldsymbol{\lambda})$ is negative in $(\hat{z}_{q-1}, \hat{z}_q)$, and therefore, from Eq. (B.15) we deduce that $\beta_k(\cdot)$ is positive in $(\hat{z}_{q-1}, \hat{z}_q)$. This together with $\int_{z_1^\star}^{\bar{\theta}} \beta_k(z)dz = 0$ imply that $x_k(\cdot)$ is constant in $(\hat{z}_{q-1}, \hat{z}_q)$ (by means of complementary slackness any change would yield a contradiction). Let us denote the value of $x_k(\cdot)$ in $(\hat{z}_{q-1}, \hat{z}_q)$ by χ_q. Note that if $\chi_q = 0$, we are done. Similar to what we did before, we define $z_2^\star \triangleq \inf\{z \in [0, \hat{z}_{q-1}] : x_k(z) = \chi_q\}$. Note that $z_2^\star < \hat{z}_{q-1}$. If $z_2^\star = 0$, then $x_k(\cdot)$ equals χ_q for all values below z_q, and therefore, there is nothing more to prove. Thus, assume that $z_2^\star > 0$. If $z_2^\star = \hat{z}_{q-1}$ then $x_k(\cdot)$ changes value at $\hat{z}_{q-1}$ and, therefore, by complementary slackness $\int_{\hat{z}_{q-1}}^{\bar{\theta}} \beta_k(z)dz = 0$. However, $L_k(\cdot|\boldsymbol{\lambda})$ is positive in $(\hat{z}_{q-2}, \hat{z}_{q-1})$ which by Eq. (B.15) implies that β_k is negative in $(\hat{z}_{q-2}, \hat{z}_{q-1})$, but this would contradict the dual feasibility of β_k. Hence, we can assume that $z_2^\star < \hat{z}_{q-1}$.

Let q_2 be the largest index in $\{0, 1, \ldots, q-1\}$ such that $\hat{z}_{q_2} \leq z_2^\star$. Note that $z_2^\star \in [\hat{z}_{q_2}, \hat{z}_{q_2+1}]$. As before, we can show that $L_k(\cdot|\boldsymbol{\lambda})$ is positive in $(\hat{z}_{q_2}, \hat{z}_{q_2+1})$ and $z_2^\star = \hat{z}_{q_2}$. Note that this implies that the value χ_q of $x_k(\cdot)$ extends for at least two intervals, namely, $(\hat{z}_{q-2}, \hat{z}_{q-1})$ and $(\hat{z}_{q-1}, \hat{z}_q)$.

The previous argument can be applied iteratively over all intervals defined by $\hat{z}_1 < \cdots < \hat{z}_p$. Since in each step of the argument we cover two intervals, we deduce that there can be at most $\lfloor p/2 \rfloor$ different values of $\chi_{q'}$, where q' is defined in every step as we did before. Moreover, if $L_k(0|\boldsymbol{\lambda}) < 0$, then in the interval $(0, \hat{z}_1)$, the dual variable $\beta_k(\cdot)$ is positive. Because $\int_{\hat{z}_1}^{\bar{\theta}} \beta_k(z)dz = 0$ (this follows from the steps of the argument) and $x(0) \cdot \int_0^{\bar{\theta}} \beta(z)dz = 0$, we must have $x(0) = 0$, and so in the last interval x_k, equals 0. □

References

Akan, M., Ata, B., Dana, J., 2015. Revenue management by sequential screening. J. Econ. Theory 159, 728–774.

Amador, M., Bagwell, K., 2013. The theory of optimal delegation with an application to tariff caps. Econometrica 81 (4), 1541–1599.

Anderson, E.J., Nash, P., 1987. Linear Programming in Infinite-Dimensional Spaces: Theory and Applications. John Wiley & Sons.

Ashlagi, I., Daskalakis, C., Haghpanah, N., 2016. Sequential mechanisms with ex-post participation guarantees. In: ACM 2016 Conference on Economics and Computation.

Balseiro, S., Mirrokni, V., Paes Leme, R., 2018. Dynamic mechanisms with martingale utilities. Manag. Sci. 64, 5062–5082.

Bergemann, D., Castro, F., Weintraub, G., 2020. Uniform Pricing vs Third Degree Price Discrimination. Discussion Paper Discussion Paper 2213R. Cowles Foundation for Economic Research, Yale University.

Bulow, J., Roberts, J., 1989. The simple economics of optimal auctions. J. Polit. Econ. 97, 1060–1090.

Celis, L.E., Lewis, G., Mobius, M., Nazerzadeh, H., 2014. Buy-it-now or take-a-chance: price discrimination through randomized auctions. Manag. Sci. 60, 2927–2948.

Courty, P., Li, H., 2000. Sequential screening. Rev. Econ. Stud. 67, 697–717.

Daskalakis, C., Deckelbaum, A., Tzamos, C., 2015. Strong duality for a multiple-good monopolist. In: Proceedings of the Sixteenth ACM Conference on Economics and Computation. ACM, pp. 449–450.

Fuchs, W., Skrzypacz, A., 2015. Government interventions in a dynamic market with adverse selection. J. Econ. Theory 158, 371–406.

Heumann, T., 2020. Information design and sequential screening with ex-post participation constraints. Theor. Econ. 15 (1), 319–359. https://doi.org/10.3982/TE2818.

Krähmer, D., Kovac, E., 2016. Optimal sequential delegation. J. Econ. Theory 163, 849–888.

Krähmer, D., Strausz, R., 2014. Optimal Sales Contracts with Withdrawal Rights. Mimeo.

Krähmer, D., Strausz, R., 2015. Optimal sales contracts with withdrawal rights. Rev. Econ. Stud. 82, 762–790.

Krähmer, D., Strausz, R., 2016. Optimality of sequential screening with multiple units and ex-post participation constraints. Econ. Lett. 142, 64–68.

Manelli, A.M., Vincent, D.R., 2007. Multidimensional mechanism design: revenue maximization and the multiple-good monopoly. J. Econ. Theory 137, 153–185.

Mirrokni, V., Nazerzadeh, H., 2017. Deals or no deals: contract design for online advertising. In: Proceedings of the World Wide Web Conference 2017, pp. 7–14. International World Wide Web Conferences Steering Committee.

Myerson, R.B., 1979. Incentive compatibility and the bargaining problem. Econometrica, 61–73.

Myerson, R.B., 1981. Optimal auction design. Math. Oper. Res. 6, 58–73.

Riley, J., Zeckhauser, R., 1983. Optimal selling strategies: when to haggle, when to hold firm. Q. J. Econ. 98, 267–289.

Samuelson, W., 1984. Bargaining under asymmetric information. Econometrica, 995–1005.

Winkler, G., 1988. Extreme points of moment sets. Math. Oper. Res. 13 (4), 581–587.