THE SCOPE OF SEQUENTIAL SCREENING WITH EX-POST PARTICIPATION

By

Dirk Bergemann, Francisco Castro, and Gabriel Weintraub

February 2017
Revised June 2017

COWLES FOUNDATION DISCUSSION PAPER NO. 2078R

The Scope of Sequential Screening with Ex-Post Participation Constraints*

Dirk Bergemann† Francisco Castro‡ Gabriel Weintraub§

June 11, 2017

Abstract

We study the classic sequential screening problem under ex-post participation constraints. Thus the seller is required to satisfy buyers' ex-post participation constraints. A leading example is the online display advertising market, in which publishers frequently cannot use up-front fees and instead use transaction-contingent fees.

We establish when the optimal selling mechanism is static (buyers are not screened) or dynamic (buyers are screened), and obtain a full characterization of such contracts. We begin by analyzing our model within the leading case of exponential distributions with two types. We provide a necessary and sufficient condition for the optimality of the static contract. If the means of the two types are sufficiently close, then no screening is optimal. If they are sufficiently apart, then a dynamic contract becomes optimal. Importantly, the latter contract randomizes the low type buyer while giving a deterministic allocation to the high type. It also makes the low type worse-off and the high type better-off compared to the contract the seller would offer if he knew the buyer's type. Our main result establishes a necessary and sufficient condition under which the static contract is optimal for general distributions. We show that when this condition fails, a dynamic contract that randomizes the low type buyer is optimal.

Keywords: Sequential screening, ex-post participation constraints, static contract, dynamic contract.

JEL Classification: C72, D82, D83.

---

*The first author acknowledges financial support from NSF Grant SES 1459899. We thank Alessandro Bonatti, Daniel Krähmer and Roland Strausz for helpful conversations.

†Department of Economics, Yale University, New Haven, U.S.A., dirk.bergemann@yale.edu.

‡Graduate School of Business, Columbia University, New York, U.S.A., fcastro19@gsb.columbia.edu.

§Graduate School of Business, Stanford University, Stanford, U.S.A., gweintra@stanford.edu

1 Introduction

1.1 Motivation

In many markets, sellers are constrained to sell products in such a way that buyers obtain a non-negative net utility once they have realized their valuation. A leading example is the online display advertising market. In this setting, typical business constraints impose that publishers cannot use up-front fees and thus instead run a series of "waterfall auctions" that implicitly impose different priorities over participants. Commonly, higher-priority auctions have higher reserves. [1] Another example, is online shopping, in which shoppers have the chance to return the purchased item after delivery, usually at no or low cost.

Motivated by this, we study the sequential screening problem as described by Courty and Li (2000) and in order to match our previous narrative we incorporate ex-post participation constraints. The goal of this work is to understand when the optimal selling mechanism is *static* (buyers are not screened ex-ante) or *dynamic* (buyers are screened ex-ante) and obtain a full characterization of such contracts.

Our model considers a seller who is selling one unit of an object, at no cost to the seller, to a buyer who has an outside option of zero. The sequence of events occurs in two periods. In the first, the buyer privately learns her type and the parties contract—we restrict our analysis to two types of buyers (low and high). The high type has a valuation distribution that dominates the low type one. The contract specifies allocation and payment functions. In the second period, the buyer privately learns her valuation, and allocations and transfers are realized. At this point, the buyer only accepts the contracting terms if her realized net utility is weakly larger than her outside option. This model aligns with our aforementioned examples. In the case of display advertising, the first period can be thought of as the time at which the buyer decide in which auction (priority/reserve) to participate in. The second period is when the auction is actually run. We begin by analyzing our model in the case of exponential valuations, which allow us to obtain clean and intuitive closed-form expressions. We then provide a general version of the results.

1.2 Our Results

One of our main contributions is to characterize when a static contract—that is, a contract that does not sequentially screen buyers—is optimal. We provide a necessary and sufficient condition for the optimality of the aforementioned contract. For further reference, we call this condition (NR). The characterization

[1] See, for example, https://adexchanger.com/the-sell-sider/the-programmatic-waterfall-mystery.

we provide is an average monotonicity condition around the optimal static threshold that encodes information about the similarity of the ex-ante types. For example, in the case of exponential valuations, the static contract is optimal if and only if the means of the distributions of the low and high type are appropriately close.

Our second main contribution characterizes the optimal mechanism when the condition mentioned above does not hold and a static contract is no longer optimal. Specifically, we prove that the optimal *dynamic* contract randomizes the low type and gives a deterministic allocation to the high type. Basically, randomization occurs to prevent the high type buyer from taking the low type's contract. To prove this, we first show that when (NR) is not satisfied, such a sequential screening contract with random allocations becomes feasible and yields an improvement in the seller's revenue compared to the static contract. Even though this contract yields an improvement over the static one, it does not need to be optimal. However, we are able to identify some regularity conditions that imply optimality.

More specifically, the optimal contract is characterized by an allocation probability $\chi \in (0, 1)$, and three thresholds θ_1, θ_2, and θ_3 with $\theta_1 \leq \theta_2 \leq \theta_3$. In this contract, the seller allocates the object to a low-type buyer with probability χ whenever her valuation is between θ_1 and θ_3, and asks for a payment of $\theta_1 \cdot \chi$. When the valuation of this type is above θ_3, the object is always allocated to her and the seller demands a payment of $\theta_3 - (\theta_3 - \theta_1) \cdot \chi$. The high-type buyer gets the object with certainty and only when her valuation is above θ_2, at which point the payment she has to make to the seller is θ_2. These parameters are set in such a way that the ex-ante incentive compatibility constraints are satisfied.

A salient feature of this type of contract is that it discriminates the low type in two dimensions. First, it can be proven that θ_1 is above the optimal threshold a seller would set if she was selling exclusively to low-type buyers. That is, the low type buyer is being allocated the object less often in the presence of high type buyers. An opposite result is true for high-type buyers, they are being 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, which further reduces the chances of a low type to receive the object compared to a case in which there are no high-type buyers. All these values and properties can be clearly expressed for the exponential distribution case. At the end of the paper, we discuss directions on how to expand our analysis and results.

1.3 Related Work

Our model builds on the sequential screening literature as pioneered by Courty and Li (2000), in which there is a buyer who sequentially and privately learns her true valuation. In this classic paper, the buyer only has partial information about her valuation ex-ante when signing a contract to which she can fully commit. This feature is represented by an *ex-ante participation constraint*. In contrast, in this paper we give the option to the buyer to quit the relationship ex-post after acquiring complete information about her valuation. We represent this by an *ex-post participation constraint*.

The closest 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 imposes strong restrictions on the primitives as it rules out common valuation distributions such as the exponential distribution. Furthermore, the condition is only sufficient and thus gives an incomplete characterization of the space of primitives for which the static contract is optimal. However, Krähmer and Strausz (2014) also acknowledge the existence of a necessary condition for the optimality of the static contract. In our paper, for the setting of a single buyer and two ex-ante types, we close this gap by providing a necessary and sufficient condition under which the static contract is optimal. Further and importantly, when our condition breaks we characterize the optimal *dynamic* mechanism and show that randomization of one of the ex-ante types is required for optimality.[2] In terms of approaches, Krähmer and Strausz (2015) relax both the low to high IC and monotonicity constraints and then show that, under their condition, the contract that maximizes the Lagrangian is deterministic and that as a result the static contract is optimal. In contrast, we also relax the same IC constraint but we keep monotonicity. 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 not only the right structure of the optimal contract but also the main objects of analysis. In turn, this permits us not only to determine our necessary and sufficient condition but also to characterize the optimal dynamic contract when our condition breaks. In related recent work, Heumann (2016) considers a setting in which a seller can design the screening mechanism as well as the information disclosure mechanism with ex-post participation constraints.

The sequential nature of our model and the incorporation of ex-post (IR) is related to the work of Ashlagi, Daskalakis, and Haghpanah (2016) and Balseiro, Mirrokni, and Paes Leme (2016). These authors consider a model in which a seller, constrained by ex-post IR (also motivated by the display advertising market), repeatedly sells objects to a buyer whose valuations are independent across periods. Both papers provide characterizations for a nearly optimal mechanism. They are different from ours

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

because we consider a single sale and construct the exactly optimal mechanism in a sequential screening model.

Related to the display advertising setting, our paper is related to the BIN-TAC auction introduced by Celis, Lewis, Mobius, and Nazerzadeh (2014). In this auction, buyers can opt to buy an impression at a relatively high buy-it-now (BIN) price. If no buyers purchase at the BIN price, there is a take-a-chance (TAC) stage in which the winner is randomly selected. The BIN-TAC auction is a relatively simple and intuitive design that aims to approximate the optimal static Myerson auction with ironing in the non-regular case. The paper does not provide a mechanism design analysis of BIN-TAC; however, it numerically compares its outcome with the Myerson optimal auction showing that it achieves close-to-optimal outcomes in many contexts of interest. Even though our setting is different because of sequential screening, it is interesting to note that our optimal dynamic mechanism has a similar flavor to the TAC stage in that it randomizes the allocation for low types.

2 The Model

We consider a seller (he) who is selling one unit of an object at zero cost to a buyer (she) who has an outside option of zero. 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 type and then the parties contract. The type provides information about the buyer's valuation distribution. The contract specifies allocation and payment functions. In the second period, the buyer privately learns her valuation, and allocations and transfers are realized. At this point the buyer only accepts the contracting terms if her net realized utility is larger than her outside option.

In the first period both parties do not possess information about the buyer's valuation θ (or *ex-post type*) but the buyer privately knows her type k (or *ex-ante type*). We assume that the buyer has probability α_k of being of type $k \in \{1, \dots, K\}$, with $\sum_{k=1}^{K} \alpha_k = 1$ and $\alpha_k > 0$. In the second period, a buyer of type k privately learns her valuation θ which we assume to have cdf $F_k(\cdot)$ and pdf $f_k(\cdot)$, with full support in $[0, \bar{\theta}]$ (possibly infinite). It will be convenient to denote the upper cdf by $\bar{F}_k(\cdot) \triangleq 1 - F_k(\cdot)$. All the distributions are common knowledge.

The terms of trade are specified in the first period by the seller. For a payment $t \in \mathbb{R}$ and a probability of receiving the object $x \in [0, 1]$, a buyer with valuation θ receives a utility of $\theta \cdot x - t$, while the seller gets paid t. We assume that the buyer agrees to purchase the object only if she is guaranteed a non-negative net utility for any possible valuation of the object she might have. That is, we require $\theta \cdot x - t$ to be non-negative for all θ. The seller's problem is to design a contract that maximizes his expected payment, constrained to guaranteeing the buyer a non-negative realized ex-post utility.

In general, two types of contract can arise as a solution to the seller's problem: *static* and *dynamic*. A static contract does not screen among ex-ante types and, therefore, offers to all of them the same terms of trade. A dynamic contract offers different contracting conditions for different ex-ante types. For example, if we had only low or high valuation buyers, a static contract would offer a unique menu of transfers and allocations, while a dynamic contract would offer two menus and each type of buyer would self-select into one of the menus.

2.1 Mechanism Design Formulation

By means of the revelation principle (see, e.g., Myerson (1979)) we can focus on incentive compatible direct revelation mechanisms, with allocations $x_k : [0, \bar{\theta}] \to [0,1]$ and transfers $t_k : [0, \bar{\theta}] \to \mathbb{R}$, that depend on the types (k, θ) reported to the mechanism. Then, for a buyer reporting an ex-ante 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 reported k in the first period and θ' in the second period while her true valuation is θ as

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

with the understanding that $u_k(\theta)$ equals $u_k(\theta; \theta)$. Similarly, we define the ex-ante expected utility of a buyer whose true ex-ante type is k but reported to the mechanism k' as

$$U_{kk'} \triangleq \int_0^{\bar{\theta}} \max_{\theta' \in [0, \bar{\theta}]} \{u_{k'}(z; \theta')\} \cdot f_k(z) dz,$$

where the maximum is included because double deviations are allowed.

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

$$u_k(\theta) \geq u_k(\theta; \theta') \quad \forall k \in \{1, \ldots, K\}, \forall \theta \in [0, \bar{\theta}]. \tag{IC^{xp}}$$

The second one is ex-ante incentive compatibility or (IC^{xa}) constraint which requires that truth-telling is optimal in the first period, that is,

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

Also, we require the mechanism to satisfy an ex-post individual rationality constraint or (IR^{xp})

$$u_k(\theta) \geq 0, \quad \forall k \in \{1, \ldots, K\}, \quad \forall \theta \in [0, \bar{\theta}]. \tag{IR^{xp}}$$

Then, the seller's problem is

$$
\begin{aligned}
(\mathcal{P}) \quad \max \quad & \sum_{k=1}^{K} \alpha_k \cdot \int_0^{\bar{\theta}} t_k(z) \cdot f_k(z) dz \\
\text{s.t} \quad & (IC^{xa}), (IC^{xp}), (IR^{xp}) \\
& 0 \leq \mathbf{x} \leq 1 \ .
\end{aligned}
$$

Observe that (IR^{xp}) implies ex-ante individual rationality. In fact, if we were to relax $(\mathcal{P})$ by considering only ex-ante individual rationality we would be in the setting of Courty and Li (2000) for discrete ex-ante types.

3 Elementary Characterizations

We can obtain a more amenable characterization of the constraints by eliminating the transfer from the constraints.

Lemma 1 *The mechanism $(\mathbf{x}, \mathbf{t})$ satisfies (IC^{xa}),(IC^{xp}) and (IR^{xp}) if and only if*

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

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

2. *$u_k(0) \geq 0$ for all k in $\{1, \ldots, K\}$.*

3. *$u_k(0) + \int_0^{\bar{\theta}} x_k(z) \bar{F}_k(z) dz \geq u_{k'}(0) + \int_0^{\bar{\theta}} x_{k'}(z) \bar{F}_k(z) dz$ for all k, k' in $\{1, \ldots, K\}$.*

Proof. The proof of this result is standard and, thus, omitted. ■

The first condition in the lemma is the standard envelope condition and it comes from the ex-post incentive compatibility constraint. The second condition is derived from the ex-post individual rationality constraint and the fact that $u_k(\theta)$ is non-decreasing. The third condition is simply the envelope formula plugged into the ex-ante incentive compatibility constraint.

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

$$
\begin{aligned}
(\mathcal{P}^d) \quad \max_{0\le \mathbf{x}\le 1} \quad & -\sum_{k=1}^{K}\alpha_k u_k + \sum_{k=1}^{K}\alpha_k \int_0^{\bar\theta} x_k(z)\mu_k(z)f_k(z)dz \\
\text{s.t} \quad & x_k(\theta) \quad \text{non-decreasing}, \quad \forall k\in\{1,\dots,K\} \\
& u_k \ge 0, \quad \forall k \in \{1,\dots,K\} \\
& u_k + \int_0^{\bar\theta} x_k(z)\bar F_k(z)dz \ge u_{k'} + \int_0^{\bar\theta} x_{k'}(z)\bar F_k(z)dz, \quad \forall k,k' \in \{1,\dots,K\},
\end{aligned}
$$

where $\mu_k(\cdot)$ is the virtual valuation of the ex-ante type k defined as:

$$
\mu_k(\theta) \triangleq \theta - \frac{\bar F_k(\theta)}{f_k(\theta)}, \quad \forall k \in \{1,\dots,K\}, \forall \theta \in [0,\bar\theta].
$$

Note that in $(\mathcal{P}^d)$ the variables are the allocation rule $\mathbf{x}$ and the vector of lowest ex-post type utilities $\mathbf{u}$. Further, once we solve for this variables the transfers are determined by equation (1).

Note that a solution to $(\mathcal{P}^d)$ that screens the ex-ante types is a dynamic contract. However, note that a solution to $(\mathcal{P}^d)$ can also be static as it might pool the ex-ante types into a single type. Formally, we say that a solution to $(\mathcal{P}^d)$ or contract is *static* when $x_k(\cdot) \equiv x(\cdot)$ and $u_k \equiv u$ for all k in $\{1,\dots,K\}$.

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

$$
\begin{aligned}
(\mathcal{P}^s) \quad \max_{0\le x\le 1} \quad & \int_0^{\bar\theta} x(z)\cdot\Big(\sum_{k=1}^{K}\alpha_k\mu_k(z)f_k(z)\Big)dz \\
\text{s.t} \quad & x(\theta) \quad \text{non-decreasing},
\end{aligned}
$$

which corresponds to the classic optimal mechanism design problem, where the term in parenthesis corresponds to the virtual values of the mixture distribution times the density function of the mixture. The main focus of this paper is two-fold. First, to study when the optimal solutions to the static and dynamic programs, $(\mathcal{P}^s)$ and $(\mathcal{P}^d)$, coincide. Second, when they are not the same, we aim to characterize the optimal solution to $(\mathcal{P}^d)$.

We say that an allocation rule $x(\cdot) : [0,\bar\theta] \to [0,1]$ is a *threshold allocation* characterized by $\tilde\theta \in [0,\bar\theta]$ if

$$
x(\theta) = \begin{cases} 1 & \text{if } \theta \ge \tilde\theta \\ 0 & \text{if } \theta < \tilde\theta. \end{cases}
$$

4 Leading Example: Exponential Distribution

Before we begin developing our general theory and in order to build intuition we provide our results for exponentially distributed valuations. In particular, we provide a necessary and sufficient condition for the static contract to be optimal and, we also give a full characterization of the optimal static contract and the dynamic contract when optimal.

We consider $K = 2$ and the density functions

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

We assume $\lambda_L > \lambda_H$, so L and H stand for low and high type respectively. Note that H dominates L in the sense of the hazard rate stochastic order and the first order stochastic dominance. In addition, for the ex-ante 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.[3] Thus, in the exponential case the seller's expected revenue for any given threshold θ is

$$R^s(\theta) \triangleq \int_{\theta}^{\bar{\theta}} (\alpha_1 \mu_1(\theta) f_1(\theta) + \alpha_2 \mu_2(\theta) f_2(\theta)) d\theta = \alpha_L \theta e^{-\lambda_L \theta} + \alpha_H \theta e^{-\lambda_H \theta}.$$

In order to find the optimal threshold we just need to maximize the expression above. The first order condition yields

$$\alpha_L(\theta - \frac{1}{\lambda_L})\lambda_L e^{-\lambda_L \theta} + \alpha_H(\theta - \frac{1}{\lambda_H})\lambda_H e^{-\lambda_H \theta} = 0, \tag{2}$$

that is, the optimal threshold is a zero of the mixture virtual valuation. Notice that equation (2) cannot be explicitly solve; however, we can (as we do in the forthcoming results) provide comparative statics. Interestingly, in Proposition 2 below, we show that we can obtain explicit expressions for the thresholds characterizing the optimal dynamic contract.

The following lemma provides some initial properties of the optimal static contract.

Lemma 2 *The optimal solution to $(\mathcal{P}^s)$ is a threshold allocation characterized by θ^s in $[\frac{1}{\lambda_L}, \frac{1}{\lambda_H}]$, solving (2). Also, θ^s is a non-increasing function of α_L with $\theta^s(0) = \frac{1}{\lambda_H}$ and $\theta^s(1) = \frac{1}{\lambda_L}$.*

We thus establish that the optimal allocation is given by a threshold allocation between $1/\lambda_L$ and $1/\lambda_H$. Note that the optimal static contract allocates using the mixture of the valuation distributions for the low and high types which cross zero at $1/\lambda_L$ and $1/\lambda_H$, respectively. Finally, the monotonicity

[3] See, e.g., Riley and Zeckhauser (1983).

property in Lemma 2 implies that as the proportion of low types increases, the optimal threshold should be closer to the one the seller would set if the only ex-ante type was the low type.

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

Proposition 1 *The static contract is optimal if and only if*

$$\theta^s \leq \frac{1}{\lambda_L - \lambda_H}. \tag{3}$$

We note that the left hand side, θ^s, is a solution to equation (2) and, therefore, it also depends on the parameters λ_L and λ_H. Subsequent corollaries provide sharper characterizations that only depend on model primitives. It is also worth mentioning that equation (3) possesses a general analog. As a matter of fact, in Section 5.2 we state the general version of this necessary and sufficient condition.

Proposition 1 provides an intuitive characterization for when the seller is better-off screening the ex-ante types than not. In terms of equation (3), when λ_L and λ_H are sufficiently close then the ratio $1/(\lambda_L - \lambda_H)$ is large and, therefore, equation (3) should hold, in which case the static contract is optimal. Conversely, when λ_L and λ_H are sufficiently apart from each other the ratio $1/(\lambda_L - \lambda_H)$ is small and potentially smaller than θ^s, so the static contract might not longer be optimal.

At a more intuitive level, when the ex-ante 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 get at most what it would get by setting thresholds $1/\lambda_L$ and $1/\lambda_H$ for each type, but $\theta^s \in [\frac{1}{\lambda_L}, \frac{1}{\lambda_H}]$. However, when screening, the seller has to pay an extra cost to prevent the types from mimicking each other and, since the contracts' revenue will be similar, it is likely that this cost offsets the earnings from screening. On the other hand, when ex-ante types are sufficiently apart in their mean valuation then the seller can tailor the contract to each type and in this way extract more from them than in the static contract.

Corollary 1 *Assume $\lambda_L \in (\lambda_H, 2\lambda_H]$, 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 then no matter in which proportion the types are, the static contract is always optimal.

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

Corollary 2 asserts that when the mean of the low and high type buyers are sufficiently different then both contracts can be optimal. If the proportion of low type is low enough (but not zero) then the seller is better-off screening the types. On the other hand, if there is a very large proportion of low type buyers then the static contract is optimal. This follows because as α_L increases, one can show that θ^s decreases, and at some point condition (3) holds. This discussion suggests our final corollary.

Corollary 3 *For λ_H and α_H fixed, there exists $\bar{\lambda}_L$ larger than $2\lambda_H$ such that for all $\lambda_L \in [\bar{\lambda}_L, \infty)$ the dynamic contract is strictly optimal.*

Now we provide a characterization of the optimal dynamic contract.

Proposition 2 *Assume equation* (3) *does not hold, then the optimal allocations are*

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

with optimal transfers $t_L^\star(\theta) = \theta_L \cdot \chi \cdot \mathbf{1}_{\{\theta \geq \theta_L\}}$ and $t_H^\star(\theta) = \theta_H \cdot \mathbf{1}_{\{\theta \geq \theta_H\}}$. The thresholds are

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

with $\theta_L \leq \theta_H$. And low type's probability of receiving the object is

$$\chi = \exp\Big(- \lambda_H \Big[\frac{1}{\lambda_H} - \frac{\alpha_L}{\alpha_H} \frac{e^{-1}}{\lambda_L - \lambda_H} - \frac{1}{\lambda_L - \lambda_H} \Big] \Big). \tag{4}$$

This results aligns with the discussion succeeding Proposition 1. When the types are different enough from each other or, equivalently, when equation (3) does not hold then it is optimal to screen the types. That's why in Proposition 2 we have different allocations for low and high type buyers. The low type buyers are allocated the object more frequently ($\theta_L \leq \theta_H$) but they are randomized. This is done as a way to prevent the buyers from mimicking each other. Specifically, we must have $\theta_L \leq \theta_H$; otherwise, the low type buyers would have an incentive to pretend being the high type since that would get them allocated the object more often and at a lower price. In general, for exponential valuations the ex-ante IC constraint for the high type is binding.

It is worth noting that the dynamic 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 $1/\lambda_L$ and would always allocate the object when her value is above the threshold. However, the dynamic contract allocates the object to the low type whenever her valuation is above $\theta_L > 1/\lambda_L$ and with some probability. So the low type is worse-off in two dimensions, it is allocated the object less often and with less probability. On the other hand, the high type buyer gets allocated the object more often and with certainty because $\theta_H < 1/\lambda_H$.

In order to better understand the role of the ex-post IR constraints it is useful to compare our solution to the one we would obtain if we only had ex-ante IR constraints. In the latter case, the solution can be derived following Courty and Li (2000).[4] This solution always allocates the object to the high type and it allocates the object to the low type whenever her valuation is above some threshold θ^*(possible infinite). Furthermore, the utilities for the lowest ex-post types satisfy $u_H < u_L \leq 0$. Clearly this solution, that uses up-front fees, is not feasible in our context as the high type buyer has negative utility for her lowest valuation. Moreover, the allocations differ for both the static and dynamic contracts.

We now illustrate our findings with numerical results where we vary the difference in the mean between the low and the 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 λ_L to be $\lambda_H + \delta$ with $\delta > 0$. Figure 1 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 dynamic contract is optimal. This aligns with Proposition 1, because as δ increases $1/(\lambda_L - \lambda_H)$ decreases converging to zero and, therefore, we expect it to be below θ^s (see Corollary 2 and Corollary 3). At a more intuitive level as δ increases both distribution become more and more different from each other with one of them having a larger average value than the other. Thus, there is a gain in screening the types.

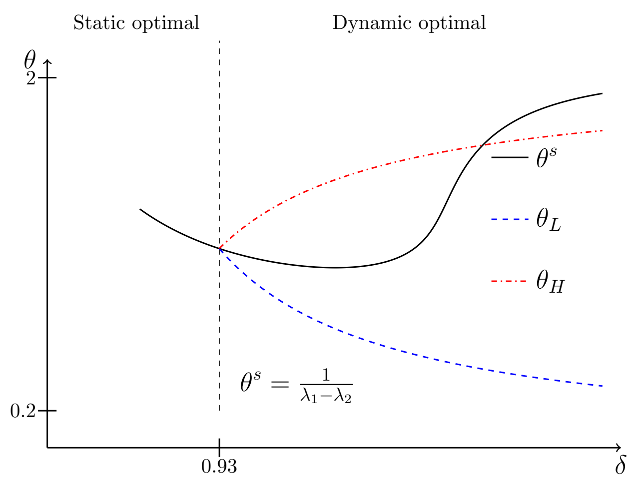

Figure 1: Optimal thresholds for static and dynamic contracts when setting $\lambda_L = \lambda_H + \delta$, with $\alpha_L = 0.7$ and $\lambda_H = 0.5$.

[4] Note that the exponential distribution satisfies first order stochastic dominance; however, the virtual valuations as constructed in the Courty and Li (2000) setting, $\mu_L(\theta) = \theta - \frac{\alpha_H}{\alpha_L} \frac{e^{-\lambda_H \theta} - e^{-\lambda_L \theta}}{\lambda_L e^{-\lambda_L \theta}}$ and $\mu_H(\theta) = \theta$, do not satisfy the regularity condition because $\mu_L(\theta)$ is not non-decreasing. Nonetheless, it is still possible to obtain the optimal mechanism.

In terms of thresholds, for the static contract we observe that θ^s is decreasing at the beginning and then it increases and goes close 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 is too small and, therefore, the probability of allocating the object to a low type, $P(\text{value low type} > \theta^s) = e^{-\lambda_L \theta^s}$, is going to be so low that the seller will be better off by choosing a threshold tailored for the high type, that is, close to $1/\lambda_H = 2$. For the dynamic thresholds, the one for the low type is decreasing while the one for the high type is increasing. This makes sense because in the dynamic case the seller can adjust the threshold for each type; hence, as δ increases the distributions become more and more different and, therefore, is optimal to set thresholds closer and closer to the threshold a seller would set if he knew the types in advance, that is, $1/\lambda_L$ and $1/\lambda_H$. Also, note that from equation (4) we see that χ is a decreasing function of δ because as the mean of the low type goes to zero we are less and less constrained to offer a high probability of allocation; however, in the limit $\chi(\delta) \approx e^{-1}$, hence even though the low type buyers will have values concentrated at zero we still need to offer them a positive probability of allocation so that we prevent them from mimicking the high type buyers.

We can also compare the different mechanism in terms of revenue. Note that from Proposition 2, we can derive the optimal revenue for the dynamic contract:

$$R^d = \alpha_L \cdot \chi \cdot \theta_L \cdot e^{-\lambda_L \theta_L} + \alpha_H \cdot \theta_H \cdot e^{-\lambda_H \theta_H}.$$

Then, we can plot the different revenues as we vary δ. Figure 2 depicts the results. For values of δ above 0.93 the dynamic contract dominates the static one reaching an improvement of 16.5%. Note that when δ grows large the improvement of the dynamic over the static decrease because both contracts set the thresholds to maximize what they can extract from the high type buyer. Actually, we have that

$$\lim_{\delta \to \infty} R^d(\delta) = \lim_{\delta \to \infty} R^s(\delta) = \alpha_H \frac{e^{-1}}{\lambda_H},$$

which equals the optimal revenue a seller could make if he was only selling to the high type buyer.

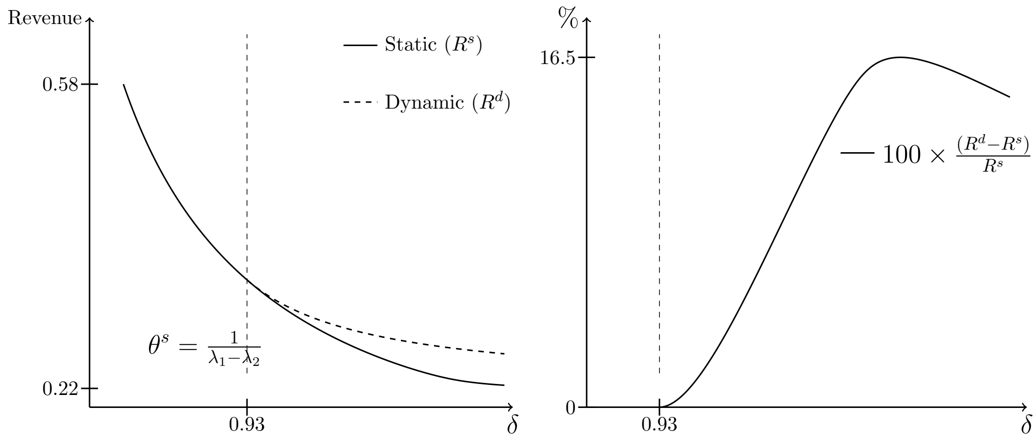

Figure 2: *Left:* Optimal expected revenue for static, dynamic and ex-ante (IR) contracts. *Right:* Percentage improvement of the dynamic over the static contract. In both figures we set set $\lambda_L = \lambda_H + \delta$, with $\alpha_L = 0.7$ and $\lambda_H = 0.5$

5 General Results: $K = 2$

In this section we present our main results for the case when we have two ex-ante types and general valuation distributions. We begin with notation and stating the main objects of our analysis. Then we provide the main results for both the static and dynamic contracts. In particular, this section generalizes Proposition 1 by providing a sharp necessary and sufficient condition for the optimality of the static contract. Then we provide a general statement, similar to Proposition 2, for the characterization of the optimal dynamic contract.

First, we give some definitions that are standard in the mechanism design literature.

Definition 1 (threshold) *We define the smallest threshold to be:*

$$\hat{\theta}_k \triangleq \min\{\theta \in [0, \bar{\theta}] : \mu_k(\theta) \geq 0\}, \quad \forall k \in \{1, \ldots, K\},$$

where $\mu_k(\cdot)$ is the virtual valuation function of type k.

Also we define the hazard rate and cross-hazard rate functions as follows.

Definition 2 (hazard rate and cross-hazard rate functions)

$$h^{\ell k}(\theta) \triangleq \frac{\bar{F}_k(\theta)}{f_\ell(\theta)}, \quad \forall \ell, k \in \{1, \ldots, K\}, \forall \theta \in [0, \bar{\theta}],$$

where when ℓ equals k we refer to this function as hazard rate, and when ℓ does not equal k we refer to it as cross-hazard rate.

With out loss of generality we assume

$$\hat{\theta}_1 \leq \cdots \leq \hat{\theta}_K, \tag{5}$$

one way of thinking about this is in terms of hazard rate stochastic order. For example consider $K = 2$, then:

$$h^{22}(\theta) \geq h^{11}(\theta), \forall \theta \in [0, \bar{\theta}] \Leftrightarrow \theta - h^{22}(\theta) \leq \theta - h^{11}(\theta) \Leftrightarrow \mu_2(\theta) \leq \mu_1(\theta) \Rightarrow \hat{\theta}_1 \leq \hat{\theta}_2.$$

That is, hazard rate stochastic order implies the order of the thresholds $\{\hat{\theta}_k\}_{k=1}^K$. Hence, for $K = 2$ we can think of type 2 as being the high valuation type and type 1 as being the low valuation type.

As it is standard in the mechanism design literature we make the following assumption, which *we keep for the rest of paper*, about the hazard rate function

$$h^{kk}(\theta) \quad \text{are non-increasing in } \theta. \tag{DHR}$$

This assumption is satisfied by a large class of distributions, for example, all log-concave distributions satisfy condition (DHR). An important consequence of this condition is that it implies the virtual valuation functions are increasing and, therefore, the thresholds $\{\hat{\theta}\}_{k=1}^K$ are uniquely determined. Another related condition is about the cross-hazard rate functions,

$$h^{\ell k}(\theta) \quad \text{are non-increasing in } \theta, \quad \forall \ell, k \in \{1, \ldots, n\}. \tag{R}$$

To the best of our knowledge condition (R) was first introduced in the context of sequential screening by Krähmer and Strausz (2015). In that paper the authors show that under condition (R) the optimal solution to $(\mathcal{P}^d)$ and to $(\mathcal{P}^s)$ coincide, that is, the static contract is optimal. However, condition (R) is very demanding and there are many common distribution that do not satisfy it. In fact, our leading example in Section 4 does not satisfy this condition because in this case the hazard rate is

$$h^{\ell k}(\theta) = \frac{e^{-(\lambda_k - \lambda_\ell)\theta}}{\lambda_\ell}, \quad \ell, k = 1, 2.$$

If we consider $\lambda_1 > \lambda_2$ then $h^{12}(\theta)$ is an increasing function and, therefore, it violates conditions (R). However, notice (DHR) is satisfied because the simple hazard rate functions are constant and equal to $1/\lambda_k$.

Now we define the key elements in our analysis, the c-ratios and the k -averaged c-ratios.

Definition 3 (c-ratios)

$$c^{\ell k}(\theta) \triangleq \frac{f_\ell(\theta)\mu_\ell(\theta)}{\bar{F}_k(\theta)} = \frac{\mu_\ell(\theta)}{h^{\ell k}(\theta)}, \quad \forall \ell, k \in \{1, \ldots, K\}, \forall \theta \in [0, \bar{\theta}].$$

In words, the $c^{\ell k}$-ratios correspond to the quotient between type ℓ's virtual valuation and the cross-hazard rate function between types ℓ and k. These quantities are the same as introduced in Krähmer and Strausz (2014) Section 8.4, to study the role of the cross-hazard rate functions. They provide a sufficient condition using these ratios and show that if the static contract is optimal then the c-ratios must be weakly increasing around θ^s. Also, note that condition (R) implies that the c-ratios are increasing. In contrast, condition (DHR) implies that only the c^{kk}-ratios are increasing.

The next definition introduces weighted averages of the c-ratios.

Definition 4 (k-averaged c-ratios)

$$\overline{C}_{\ell k}(\theta_a, \theta_b) \triangleq \frac{\int_{\theta_a}^{\theta_b} \bar{F}_k(z) c^{\ell k}(z) dz}{\int_{\theta_a}^{\theta_b} \bar{F}_k(z) dz}, \quad \forall \ell, k \in \{1, \dots, K\}, \quad 0 \leq \theta_a \leq \theta_b \leq \bar{\theta}.$$

The quantity $\overline{C}_{\ell k}(\theta_a, \theta_b)$ represents the average value of $c^{\ell k}$ weighted by the tail CDF of type k for values of θ between θ_a and θ_b. As it will become clear in Section 5.2 these ratios are relevant for the characterization of the optimal mechanism.

5.1 Static Contract

In Section 2 we defined a static contract as one that does not screen the ex-ante types and, therefore, it sets $u_k \equiv u$ and $x_k \equiv x$ for all k in $\{1, \dots, K\}$. Thus, a static contract can be cast as the optimal solution to $(\mathcal{P}^s)$. From this formulation we see that the relevant quantity that shapes the allocation $x(\cdot)$ is $\bar{\mu}(\theta) \triangleq \sum_{k=1}^{K} \alpha_k \mu_k(\theta) f_k(\theta)$. In general, independent of any regularity assumptions imposed over $\bar{\mu}(\theta)$, the optimal way to choose a non-decreasing allocation $x(\cdot)$ that maximizes

$$\int_0^{\bar{\theta}} x(z) \bar{\mu}(z) dz, \tag{6}$$

is a threshold allocation (see, e.g., Myerson (1981) or Riley and Zeckhauser (1983)). We summarize this and some other properties of the optimal solution in the following lemma.

Lemma 3 *The solution to $(\mathcal{P}^s)$ is a threshold allocation characterized by θ^s in $[\hat{\theta}_1, \hat{\theta}_K]$, maximizing (6). In addition, θ^s satisfies the following properties for $K = 2$:*

1. *Is a solution to $\alpha_1 c^{12}(\theta^s) + \alpha_2 c^{22}(\theta^s) = 0$.*

2. *Is a non-increasing function of α_1 with $\theta^s(0) = \hat{\theta}_2$ and $\theta^s(1) = \hat{\theta}_1$.*

The fact that θ^s is in the interval $[\hat{\theta}_1, \hat{\theta}_K]$ and property (2), follow the same intuition as in the exponential case presented in Section 4. Property (1) is the optimality condition found by Riley and Zeckhauser (1983) written in terms of the c-ratios, it establishes that the optimal threshold must be a zero of $\bar{\mu}(\cdot)$.

5.2 Dynamic Contract

The purpose of this section it to characterize the conditions under which it is optimal to screen the ex-ante types. In particular, we provide a necessary and sufficient condition for the static contract to be optimal. For the cases in which the static contract is not optimal we characterize the optimal dynamic contract.

Our analysis consists in studying the following relaxation to $(\mathcal{P}^d)$

$$
\begin{aligned}
(\mathcal{P}^d_R) \quad \max_{0 \leq \mathbf{x} \leq 1} \quad & -\sum_{k=1}^{2} \alpha_k u_k + \sum_{i=1}^{2} \alpha_k \int_0^{\bar{\theta}} x_k(z)\mu_k(z) f_k(z)dz \\
\text{s.t} \quad & x_k(\theta) \quad \text{non-decreasing}, \quad \forall k = 1, 2 \\
& u_k \geq 0, \forall k = 1, 2 \\
& u_2 + \int_0^{\bar{\theta}} x_2(z)\bar{F}_2(z)dz \geq u_1 + \int_0^{\bar{\theta}} x_1(z)\bar{F}_2(z)dz.
\end{aligned}
$$

The difference between $(\mathcal{P}^d_R)$ and the original dynamic formulation lies in relaxing the type 1 to type 2 IC constraint (or low to high IC constraint). Importantly, we are not relaxing the monotonicity constraint and we can obtain a characterization of the optimal solution to $(\mathcal{P}^d_R)$ as stated by the following result.

Theorem 1 *Consider problem $(\mathcal{P}^d_R)$, the optimal solution has allocations*

$$
x_1^\star(\theta) = \begin{cases} 0 & \textit{if } \theta < \theta_1 \\ \chi & \textit{if } \theta_1 \leq \theta \leq \theta_3 \\ 1 & \textit{if } \theta_3 < \theta, \end{cases} \quad x_2^\star(\theta) = \begin{cases} 0 & \textit{if } \theta < \theta_2 \\ 1 & \textit{if } \theta_2 \leq \theta, \end{cases}
$$

and transfers

$$
t_1^\star(\theta) = \begin{cases} 0 & \textit{if } \theta < \theta_1 \\ \theta_1 \cdot \chi & \textit{if } \theta_1 \leq \theta \leq \theta_3 \\ \theta_3 - (\theta_3 - \theta_1) \cdot \chi & \textit{if } \theta_3 < \theta, \end{cases} \quad t_2^\star(\theta) = \begin{cases} 0 & \textit{if } \theta < \theta_2 \\ \theta_2 & \textit{if } \theta_2 \leq \theta. \end{cases}
$$

for some values $\theta_1, \theta_2, \theta_3$ with $\hat{\theta}_1 \leq \theta_1 \leq \theta_2 \leq \theta_3$, $\theta_2 \leq \hat{\theta}_2$. And $u_2 = u_1 = 0$.

Next we provide an informal and intuitive description of what leads us to Theorem 1. This description adapts techniques from Fuchs and Skrzypacz (2015) that enable us to derive the allocations' structure. We begin with ex-ante type 2, consider an allocation $x_2^\star(\theta)$ equal to some χ in some interval (θ_a, θ_b), with $\chi \in (0,1)$. Then the part of the objective for the ex-ante type 2 in this interval is

$$\alpha_2 \cdot \chi \cdot \int_{\theta_a}^{\theta_b} \mu_2(z) f_2(z) dz.$$

If $\mu_2(\hat{\theta}) \geq 0$ for some $\hat{\theta} \in (\theta_a, \theta_b)$ then because of (DHR), $\mu_2(\theta) \geq 0$ for all $\theta \geq \hat{\theta}$ and, therefore, we can always find a better solution by setting $x_2^\star(\theta) = 1$ for all $\theta \geq \hat{\theta}$ (note that this does not affect feasibility in $(\mathcal{P}_R^d)$). On the other hand, if $\mu_2(\theta) < 0$ for all $\theta \in (\theta_a, \theta_b)$, then it must be the case that

$$u_2 + \int_0^{\bar{\theta}} x_2(z) \bar{F}_2(z) dz = u_1 + \int_0^{\bar{\theta}} x_1(z) \bar{F}_2(z) dz,$$

otherwise we could decrease χ and obtain a strict improvement in the objective. Now, consider splitting the interval in half, that is, take $\hat{\theta} = (\theta_a + \theta_b)/2$. Doing some manipulations and using (DHR) we get:

$$\frac{\int_{\theta_a}^{\hat{\theta}} \mu_2(z) f_2(z) dz}{\int_{\theta_a}^{\hat{\theta}} \bar{F}_2(z) dz} = \frac{\int_{\theta_a}^{\hat{\theta}} \bar{F}_2(z) c^{22}(z) dz}{\int_{\theta_a}^{\hat{\theta}} \bar{F}_2(z) dz} \leq c^{22}(\hat{\theta}) \leq \frac{\int_{\hat{\theta}}^{\theta_b} \bar{F}_2(z) c^{22}(z) dz}{\int_{\hat{\theta}}^{\theta_b} \bar{F}_2(z) dz} = \frac{\int_{\hat{\theta}}^{\theta_b} \mu_2(z) f_2(z) dz}{\int_{\hat{\theta}}^{\theta_b} \bar{F}_2(z) dz}, \tag{7}$$

or equivalently $\overline{C}_{22}(\theta_a, \hat{\theta}) \leq \overline{C}_{22}(\hat{\theta}, \theta_b)$. We can use this to find an improvement to the objective function. To do so, modify $x_2^\star(\theta)$ to be $\chi - \epsilon_1$ for $\theta \in (\theta_a, \hat{\theta})$ and$\chi + \epsilon_2$ for $\theta \in (\hat{\theta}, \theta_b)$, with $\epsilon_1, \epsilon_2 > 0$ and such that

$$-\epsilon_1 \cdot \int_{\theta_a}^{\hat{\theta}} \bar{F}_2(z) dz + \epsilon_2 \cdot \int_{\hat{\theta}}^{\theta_b} \bar{F}_2(z) dz = 0,$$

so the IC constraint still binds. Now, replacing in the objective:

$$-\alpha_2 \cdot \frac{\epsilon_2 \cdot \int_{\hat{\theta}}^{\theta_b} \bar{F}_2(z) dz}{\int_{\theta_a}^{\hat{\theta}} \bar{F}_2(z) dz} \cdot \int_{\theta_a}^{\hat{\theta}} \mu_2(z) f_2(z) dz + \alpha_2 \cdot \epsilon_2 \cdot \int_{\hat{\theta}}^{\theta_b} \mu_2(z) f_2(z) dz,$$

which by (7) is non-negative. Then we can keep increasing ϵ_2 until either $\chi - \epsilon_1 = x_2^\star(\theta_a)$ or $\chi + \epsilon_2 = x_2^\star(\theta_b)$. Hence, we can weakly improve the objective function by modifying the solution in a way that for one of the two halves of the interval the step, χ, reaches the boundary given by either $x_2^\star(\theta_a)$ or $x_2^\star(\theta_b)$. For the half that did not reach the boundary we can do the same procedure, which we can then repeat until we eliminate the intermediate step χ completely. This argument shows why the high type allocation is deterministic.

For the ex-ante type 1 (the low type) we can try to follow a similar argument; but, in order to obtain an improvement to the objective function we would need the analogue of condition (7) to hold, which considering the high to low IC constraint is:

$$\overline{C}_{12}(\theta_a,\hat{\theta}) \leq \overline{C}_{12}(\hat{\theta},\theta_b). \tag{8}$$

However, in general this condition is not satisfied, as the c-ratio $c^{12}(\cdot)$ does not need to be a non-decreasing function. Therefore, we cannot apply a similar argument to show we can restrict attention to deterministic contracts for the low type. Nonetheless, the optimal contract can be shown to have the structure given in Theorem 1. To see this, suppose for example that $x_1^{\star}(\theta)$ equals χ_a in $(\theta_a,\hat{\theta})$ and χ_b in $(\hat{\theta},\theta_b)$ with $0<\chi_a<\chi_b<1$, and also assume (8) does not hold. Then, we can increase χ_a and decrease χ_b (maintaining feasibility) and obtain an improvement to the objective function. We can do this until χ_a and χ_b collapse in a single value.

This discussion not only provides intuition about the structure of the optimal dynamic contract but also highlights the importance of the k -averaged c-ratios. Roughly speaking, when (7) and (8) hold we can find an improvement over a stochastic allocation, because we can modify the allocation to put more weight where the average virtual valuation is higher.

Now, we state an important corollary for our analysis that follows from the proof of Theorem 1 that allows us to focus on $(\mathcal{P}^d_R)$ when characterizing the static contract's optimality.

Corollary 4 *The static contract is an optimal solution to $(\mathcal{P}^d)$ if and only if it is an optimal solution to $(\mathcal{P}^d_R)$.*

Thus, to find a necessary and sufficient condition for the optimality of the static contract in $(\mathcal{P}^d)$ it suffices to find such a condition for $(\mathcal{P}^d_R)$. The next theorem, which is one of our main results, provides the condition. Note that the condition implies not only that the allocation must be deterministic, as discussed in equation (8), but also that the static contract is optimal.

Theorem 2 *The static contract is optimal if and only if*

$$\max_{0\leq\theta\leq\theta^s} \overline{C}_{12}(\theta,\theta^s) \leq \min_{\theta^s\leq\theta\leq\bar{\theta}} \overline{C}_{12}(\theta^s,\theta). \tag{NR}$$

Condition (NR) is the general version of the condition in Proposition 1 for exponential valuations regarding the similarity of the distributions of both types. Note that condition (R) implies the monotonicity of the c-ratios, and therefore condition (NR) holds, because:

$$\overline{C}_{12}(\theta,\theta^s) = \frac{\int_{\theta}^{\theta^s} \bar{F}_2(z)c^{12}(z)dz}{\int_{\theta}^{\theta^s} \bar{F}_2(z)dz} \leq c^{12}(\theta^s), \quad \forall\theta\leq\theta^s,$$

and

$$\overline{C}_{12}(\theta^s, \theta) = \frac{\int_{\theta^s}^{\theta} \bar{F}_2(z) c^{12}(z) dz}{\int_{\theta^s}^{\theta} \bar{F}_2(z) dz} \geq c^{12}(\theta^s), \quad \forall \theta \geq \theta^s.$$

Hence, the result by Krähmer and Strausz (2015) that if condition (R) holds then the static contract is optimal follows as corollary of Theorem 2. We highlight that while condition (R) implies the c-ratios are increasing, our condition (NR) only implies a type of monotonicity over an appropriate weighted average of the c-ratios.

We can also compare Theorem 2 with Lemma 12 in Krähmer and Strausz (2014). In that lemma they assume $h^{22}(\theta) > h^{11}(\theta)$, which we already saw it implies $\hat{\theta}_1 < \hat{\theta}_2$, and they establish that a necessary condition for the static contract to be optimal is to have the c-ratio $c^{12}(\theta)$ being increasing at θ^s. Our results also contains this lemma, because if $c^{12}(\cdot)$ was decreasing at θ^s we can always find $\underline{\theta} < \theta^s$ and $\overline{\theta} > \theta^s$ such that

$$\overline{C}_{12}(\underline{\theta}, \theta^s) > \overline{C}_{12}(\theta^s, \overline{\theta}),$$

so (NR) does not hold and, therefore, the static contract would not be optimal. Figure 3 illustrates how our condition closes the gap between the ones by Krähmer and Strausz.

In terms of methodology, our approach differs from that of Krähmer and Strausz (2015). Their approach consists of relaxing the low to high ex-ante IC 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 both types must be equal and, therefore, the static contract is optimal. In our approach we also use a relaxation of the general formulation, but we do not impose any condition on the primitives and we do not relax the monotonicity constraint. We perform a first principle analysis which allows us to understand what are the conditions under which is possible to find objective improvements. From this analysis we not only determine the structure of the optimal contract but we also identify the main objects of analysis, the k-averaged c-ratios. This leads us to Theorem 1, once we have that result we realize that the static contract is optimal for the relaxed problem if and only if it optimal for the original problem. Hence, to obtain a complete characterization of the optimality of the static contract we can study ($\mathcal{P}_R^d$). Then using this formulation, we use the KKT conditions to show that condition (NR) yields a sharp characterization as stated in Theorem 2.

An important contribution of our work is that, to the best of our knowledge, we provide the first characterization of the optimal dynamic contract when the necessary and sufficient condition associated to the static contract being optimal fails. A first step towards this is given by the following proposition.

Proposition 3 *Assume condition* (NR) *does not hold. Then there exists* θ_a, θ_b *such that* $\theta_a < \theta^s < \theta_b$ *and* $\overline{C}_{12}(\theta_a, \theta^s) > \overline{C}_{12}(\theta^s, \theta_b)$, *for which the following allocation yields a strict improvement over the*

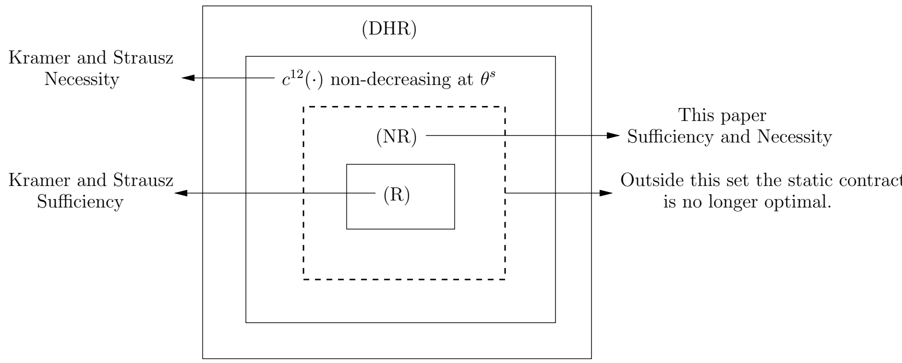

Figure 3: Optimality of the static contract for (DHR) distributions, with $K = 2$ and a single buyer.

static contract:

$$x_1(\theta) = \begin{cases} 0 & \text{if } \theta < \theta_a \\ \chi & \text{if } \theta_a \leq \theta \leq \theta_b \\ 1 & \text{if } \theta_b < \theta, \end{cases} \quad x_2(\theta) = \begin{cases} 0 & \text{if } \theta < \theta^s \\ 1 & \text{if } \theta^s \leq \theta, \end{cases}$$

where $\chi = \int_{\theta^s}^{\theta_b} \bar{F}_2(z)dz / \int_{\theta_a}^{\theta_b} \bar{F}_2(z)dz$. *And we set* $u_1 = u_2 = 0$.

In the proof of Proposition 3 we can see that as soon as condition (NR) breaks two things happen. First, a non-static contract becomes feasible as it does not violate the incentive compatibility constraints. Second, the same contract obtains a larger expected revenue than the static one. So, from this we see that (NR) is preventing both the feasibility and optimality of a dynamic contract. In terms of the k-averaged c-ratios, when (NR) fails to hold type 1's average virtual value in (θ_a, θ^s) is larger than that in (θ^s, θ_b). Hence, by increasing the static allocation in (θ_a, θ^s) and reducing it in (θ^s, θ_b) (in a feasible manner) we can find a better dynamic allocation.

The next result characterizes the optimal dynamic contract and it also provides conditions that allow to compute the optimal thresholds.

Theorem 3 *Assume condition* (NR) *does not hold. Suppose there exist* $\underline{\theta}_1 \leq \theta_2 \leq \overline{\theta}_1$ *such that*

1. $\overline{C}_{12}(\underline{\theta}_1, \overline{\theta}_1) \leq \min_{\overline{\theta}_1 \leq \theta \leq \bar{\theta}} \overline{C}_{12}(\overline{\theta}_1, \theta)$.

2. $\max_{0 \leq \theta \leq \overline{\theta}_1} \overline{C}_{12}(\theta, \overline{\theta}_1) \leq \overline{C}_{12}(\underline{\theta}_1, \overline{\theta}_1)$

3. $\alpha_1 \cdot \overline{C}_{12}(\underline{\theta}_1, \overline{\theta}_1) + \alpha_2 c^{22}(\theta_2) = 0$.

Then $(\underline{\theta}_1, \theta_2, \overline{\theta}_1)$ *characterize the optimal contract of Theorem 1, with* $\theta_1 = \underline{\theta}_1, \theta_2 = \theta_2, \theta_3 = \overline{\theta}_1$ *and* $\chi = \int_{\theta_2}^{\overline{\theta}_1} \bar{F}_2(z)dz / \int_{\underline{\theta}_1}^{\overline{\theta}_1} \bar{F}_2(z)dz$.

Conditions (1), (2), and (3) of the theorem ensure that the IC constraint from the low ex-ante type is satisfied, and can be thought as optimality conditions that characterize the thresholds. Intuitively, condition (1) asserts that for any θ larger that $\overline{\theta}_1$, type 1's average virtual valuation is always above its average value in the interval $(\underline{\theta}_1, \overline{\theta}_1)$ and , therefore, always allocating the object for these values is optimal. Condition (2) identifies an intermediate interval for which type 1's average virtual valuation is largest, thus maximizing what the seller can make from randomizing the low type in the interval $(\underline{\theta}_1, \overline{\theta}_1)$. Finally, condition (3) is simple a first order optimality condition on θ_2.

This result generalizes Proposition 2. For the exponential distribution the conditions of the theorem are always met. We note that in the exponential case we only have two intervals for the low type's allocation as we can show that $\overline{\theta}_1 = \infty$.

We would like to stress that conditions (1), (2), and (3) in the theorem are only sufficient conditions for optimality. However, there is a more general approach to compute the optimal solution. The key is that the optimal solution to $(\mathcal{P}^d_R)$ can be shown to be feasible for $(\mathcal{P}^d)$. Hence, it is enough to solve $(\mathcal{P}^d_R)$. From Theorem 1 we already know that the optimal contract depends solely on the variables $\theta_1, \theta_2, \theta_3$ and χ. It can be shown then that at optimality the IC constraint must bind and, therefore, χ is a function of the thresholds. This implies that we can optimize the objective in $(\mathcal{P}^d_R)$ as a function of the thresholds only and with no IC constraint.

5.3 Indirect Implementation

Next, we discuss how the optimal dynamic 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 $p_2 = \theta_2$.
- Contract L: in this case, the buyer can choose between two items:

  (a) Buy at a price of $p_1 = \underline{\theta}_1 \cdot \chi$ and be allocated with probability χ.

  (b) Buy at a price of $p_1 = \overline{\theta}_1 - (\overline{\theta}_1 - \underline{\theta}_1) \cdot \chi$ and be allocated with probability 1.

The prices in the above menu of contracts are set using the values in Theorem 3. This implementation offers a posted price to the high type buyer, and gives to the low type buyer two options. In option (a) the low type buyer can pay a low price but it can potentially not acquire the item; in (b), the low type buyer pays a high price and always gets the object.

An appealing feature of the implementation is that if we think of allocations as quantities, then we can order the per unit prices. In contract L, the per unit prices are $\underline{\theta}_1$ and $\underline{\theta}_1 \cdot \chi + \overline{\theta}_1 \cdot (1 - \chi)$ for (a) and (b), respectively. Hence, the per unit price in (a) is less than or equal to the one in (b). That is, the low type in (a) receives less of the good but at a discounted price compare to the low type in (b). For contract H, the per unit price is θ_2 and, since $\underline{\theta}_1$ is less than or equal to θ_2, the low type in (a) receives less of the good at a discounted price compared to the high type buyer. To contrast the per unit prices of the low type in (b) and the high type is less straightforward. Even-though θ_2 is between $\underline{\theta}_1$ and $\overline{\theta}_1$ we are not able to compare it to $\underline{\theta}_1 \cdot \chi + \overline{\theta}_1 \cdot (1 - \chi)$. However, intuitively, if the high type puts a large mass in values larger than $\overline{\theta}_1$ then we expect the per unit price of the high type to be below the one of the low type in (b) because, otherwise, the high type buyer would have an incentive to take contract L. Equivalently, the high type or the low type in (b) have to pay a premium for the additional quantity. We can also refer back to the exponential case of Section 4. From Proposition 2, the premium the high type has to pay is given by $\theta_H - \theta_L = \log(1/\chi)/\lambda_H$ and, therefore, the larger the quantity the lower is the premium. Finally, note that this implementation accommodates the case in which the static contract is optimal. In that case, we have $\chi = 1$ and $\underline{\theta}_1 = \theta_2 = \overline{\theta}_1$ thus both contracts are the same.

6 Future Work and Extensions

There are several directions in which we would like to extend the work presented in this paper.

The theoretical and numerical results we presented in this paper are for two ex-ante types. Even though this setting is already rich and, as it was shown in Section 4 provides good economic insights, a more general setting with more than just two ex-ante types is an important venue for future work. We would like to extend Theorems 2 and 3 for multiple ex-ante types. In this sense, an interesting question concerns the number of 'intermediate classes'. Theorem 3 establishes that the low type buyer has an interval in which she is allocated the object with some probability $\chi \in (0, 1)$; hence, there is one intermediate class. An interesting question is whether the number of intermediate classes increases with the number of ex-ante types. Also, is there a fixed number of intermediate classes that yield a good approximation to the optimal solution for an arbitrary number of ex-ante types?

Related to the possible number of intermediate classes is the (DHR) assumption. As we saw in the discussion after Theorem 1, the order of the k-averaged c-ratios plays an important role in determining whether an allocation has an intermediate class or not. This order is guaranteed (or partially guaranteed) by conditions over the cross-hazard rate or hazard rate functions such as (R) or (DHR). In fact, as we weaken condition (R) one intermediate class for the low type buyer emerges. Therefore we postulate that if we weaken condition (DHR), then more intermediate class might appear, in particular, the high type buyer now could have an intermediate class. However, in order to make this last point one has to be careful because in this case relaxing the high type IC constraint might no longer be a valid relaxation to the original formulation.

Another direction for future work is increasing the number of buyers. This has important practical consequences particularly in industries that use market mechanisms like auctions, such as display advertising alluded at the beginning of the paper. We believe that our extension to multiple buyers will allow us to study whether the market design of running a series of "waterfall auctions" with different priorities over participants and reserves is effective in screening buyers and how close it is to the optimal dynamic mechanism.

APPENDIX

A Proofs for section 5

We will need the following auxiliary lemmata.

Lemma 4 *Assume $\hat{\theta}_1 \leq \hat{\theta}_2$, then*

1. $\alpha_1 \min_{\theta^s \leq \theta \leq \bar{\theta}} \overline{C}_{12}(\theta^s, \theta) \leq -\alpha_2 c^{22}(\theta_s) \leq \alpha_1 \max_{0 \leq \theta \leq \theta^s} \overline{C}_{12}(\theta, \theta^s)$.
2. $\max_{0 \leq \theta \leq \theta^s} \overline{C}_{12}(\theta, \theta^s) \leq \min_{\theta^s \leq \theta \leq \bar{\theta}} \overline{C}_{12}(\theta^s, \theta)$ *if and only if* $\alpha_1 \max_{0 \leq \theta \leq \theta^s} \overline{C}_{12}(\theta, \theta^s) = -\alpha_2 c^{22}(\theta_s) = \alpha_1 \min_{\theta^s \leq \theta \leq \bar{\theta}} \overline{C}_{12}(\theta^s, \theta)$.

Proof of Lemma 4. We prove (1) first. For the inequality involving the min consider $\epsilon > 0$. Then, from the definition of $\overline{C}_{12}$ we have

$$\alpha_1 \min_{\theta^s \leq \theta \leq \bar{\theta}} \overline{C}_{12}(\theta^s, \theta) \leq \alpha_1 \cdot \frac{\int_{\theta^s}^{\theta^s+\epsilon} \bar{F}_2(z) c^{12}(z) dz}{\int_{\theta^s}^{\theta^s+\epsilon} \bar{F}_2(z) dz} \xrightarrow{\epsilon \to 0} \alpha_1 c^{12}(\theta_s) = -\alpha_2 c^{22}(\theta_s),$$

where in the equality we used Lemma 3. A similar argument applies to $\max_{0 \leq \theta \leq \theta^s} \overline{C}_{12}(\theta, \theta^s)$. Property (2) is a direct consequence of what we have just proved for (1). ■

Lemma 5 *Let $\theta_i \in [0, \bar{\theta}]$ for $i = 1, 2, 3$ be such that $\theta_1 < \theta_2 < \theta_3$. Also, consider functions $f, g : [\theta_1, \theta_3] \to \mathbb{R}_+$, with $f, g > 0$ almost everywhere in $[\theta_1, \theta_3]$. 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 5.

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

■

Lemma 6 *Suppose*

$$\max_{0\le\theta\le\theta^s}\overline{C}_{12}(\theta,\theta^s) > \min_{\theta^s\le\theta\le\bar{\theta}}\overline{C}_{12}(\theta^s,\theta).$$

Then, there exist $\theta_a,\theta_b\in[0,\bar{\theta}]$ with $\theta_a<\theta^s<\theta_b$ such that $\overline{C}_{12}(\theta_a,\theta^s)>\overline{C}_{12}(\theta^s,\theta_b)$. Note that this implies that: $0<\int_{\theta_a}^{\theta^s}\bar{F}_2(z)c^{12}(z)dz=\int_{\theta_a}^{\theta^s}\bar{F}_1(z)c^{11}(z)dz$.

Proof of Lemma 6. Suppose the result is not true. That is, for all $\theta_a,\theta_b\in[0,\bar{\theta}]$ with $\theta_a<\theta^s<\theta_b$ we have

$$\overline{C}_{12}(\theta_a,\theta^s)\le\overline{C}_{12}(\theta^s,\theta_b). \tag{9}$$

Take $\epsilon>0$ and consider $\theta_b(\epsilon)=\theta^s+\epsilon$, then from equation (9) we have

$$\overline{C}_{12}(\theta_a,\theta^s)\le\overline{C}_{12}(\theta^s,\theta^s+\epsilon),\quad\forall\epsilon>0,$$

taking the limit as ϵ approaches to 0 yields $\overline{C}_{12}(\theta_a,\theta^s)\le c^{12}(\theta^s)$ for all $\theta_a<\theta^s$. This implies that

$$\max_{0\le\theta\le\theta^s}\overline{C}_{12}(\theta,\theta^s)=c^{12}(\theta^s).$$

Using equation (9) again, we can do the same for the minimum and, therefore, we obtain a contradiction.

To finalize, we argue why $0<\int_{\theta_a}^{\theta^s}\bar{F}_2(z)c^{12}(z)dz$. Note that since $\theta_b>\theta^s\ge\hat{\theta}_1$ we have $\overline{C}_{12}(\theta^s,\theta_b)>0$. Therefore, $\overline{C}_{12}(\theta_a,\theta^s)>0$ which implies the desired inequality. ■

Proof of Lemma 3. The fact that the optimal solution is a threshold allocation and property (1) are explained in the main text. Thus, we only need to provide a proof for θ^s being in the interval $[\hat{\theta}_1,\hat{\theta}_K]$ and property (2).

We begin showing that θ^s belongs to the interval $[\hat{\theta}_1,\hat{\theta}_K]$. Note that for all θ below $\hat{\theta}_1$, $\mu_k(\theta)$ is negative for all $k\in\{1,\dots,K\}$. Therefore, $\bar{\mu}(\theta)$ is negative for all θ below $\hat{\theta}_1$. Similarly, for all θ above $\hat{\theta}_K$, $\bar{\mu}(\theta)$ is positive. Since the allocation is of the threshold type, it is optimal to set $x(\theta)$ equal to 0 for θ below $\hat{\theta}_1$ and to set $x(\theta)$ equal to 1 for θ above $\hat{\theta}_K$. This necessarily implies that θ^s is in $[\hat{\theta}_1,\hat{\theta}_K]$.

As for Property (2), note first that θ^s can be seen as a function of α_1 and α_2 but since α_2 equals $1-\alpha_1$, we can effectively consider θ^s just a function of α_1. Then, when α_1 equals 0 is as we only had type 2 buyers and, therefore, the optimal threshold is $\hat{\theta}_2$. While when α_1 equals 1 is as we only had type 1 buyers so the optimal threshold is $\hat{\theta}_1$. Hence, $\theta^s(0)$ equals $\hat{\theta}_2$ and $\theta^s(1)$ equals $\hat{\theta}_1$.

Now we prove that $\theta^s(\alpha_1)$ is non-increasing. Consider $\alpha_1^a<\alpha_1^b$ and suppose that $\theta^s(\alpha_1^a)<\theta^s(\alpha_1^b)$. Define

$$\ell(\theta,\alpha_1)\triangleq\int_\theta^{\bar{\theta}}\alpha_1 f_1(z)\mu_1(z)+(1-\alpha_1)f_2(z)\mu_2(z)dz,$$

note that this is a linear function of α_1 and, for fixed α_1, it is maximized at $\theta^s(\alpha_1)$. Hence,

$$\begin{aligned}\ell(\theta^s(\alpha_1^a),\alpha_1^b) &\leq \ell(\theta^s(\alpha_1^b),\alpha_1^b)\\ &= \ell(\theta^s(\alpha_1^b),\alpha_1^b-\alpha_1^a)+\ell(\theta^s(\alpha_1^b),\alpha_1^a)\\ &\leq \ell(\theta^s(\alpha_1^b),\alpha_1^b-\alpha_1^a)+\ell(\theta^s(\alpha_1^a),\alpha_1^a)\end{aligned}$$

therefore

$$\int_{\theta^s(\alpha_1^a)}^{\theta^s(\alpha_1^b)} \alpha_1^b f_1(z)\mu_1(z)+(1-\alpha_1^b)f_2(z)\mu_2(z)dz \leq \int_{\theta^s(\alpha_1^a)}^{\theta^s(\alpha_1^b)} \alpha_1^a f_1(z)\mu_1(z)+(1-\alpha_1^a)f_2(z)\mu_2(z)dz. \tag{10}$$

Recall that θ^s is in $[\hat{\theta}_1,\hat{\theta}_2]$ and, therefore, $\hat{\theta}_1 \leq \theta^s(\alpha_1^a) < \theta^s(\alpha_1^b) \leq \hat{\theta}_2$. This in turn implies that

$$\mu_1(z) > 0 \quad \text{and} \quad \mu_2(z) < 0, \quad \forall z \in (\theta^s(\alpha_1^a),\theta^s(\alpha_1^b)),$$

so for z in $(\theta^s(\alpha_1^a),\theta^s(\alpha_1^b))$ we have

$$\alpha_1^a f_1(z)\mu_1(z)+(1-\alpha_1^a)f_2(z)\mu_2(z) < \alpha_1^b f_1(z)\mu_1(z)+(1-\alpha_1^b)f_2(z)\mu_2(z),$$

which contradicts (10). ■

Proof of Theorem 1. For easy of exposition we restate the problem's formulation.

$$\begin{aligned}(\mathcal{P}_R^d) \quad \max_{0\leq \mathbf{x}\leq 1} \quad & -\sum_{k=1}^{2}\alpha_k u_k + \sum_{i=1}^{2}\alpha_k \int_0^{\bar{\theta}} x_k(z)\mu_k(z)f_k(\theta)d\theta\\ \text{s.t} \quad & x_k(\theta) \quad \text{non-decreasing}, \quad \forall k=1,2\\ & u_k \geq 0, \forall k=1,2\\ & u_2+\int_0^{\bar{\theta}} x_2(z)\bar{F}_2(z)dz \geq u_1+\int_0^{\bar{\theta}} x_1(z)\bar{F}_2(z)dz.\end{aligned}$$

For any optimal solution to $(\mathcal{P}_R^d)$ two possible situations may arise:

1. The allocation has an interval in which is continuously strictly increasing.

2. The allocation does not have an interval in which is continuously strictly increasing, but is a piecewise constant non-decreasing function.

The proof idea is as follows. For each ex-ante type, we prove that if we are in case (2), we can modify the allocation in that interval to be constant and obtain at least a weak improvement in the objective. This implies that for any optimal allocation, we can construct another optimal allocation that is a piecewise constant non-decreasing function. Therefore, we can always assume we are in case

(3). In this case, we show that for ex-ante type 1 there is only one intermediate step, and for ex-ante type 2 there is no intermediate step. We split the proof in ex-ante type 1 and 2.

Let $x_1^\star(\theta)$ and $x_2^\star(\theta)$ denote the optimal allocations. We begin with ex-ante type 1.

-

- **Ex-ante type 1 case (1):** Suppose there is an interval (θ_1, θ_2) in which $x_1^\star(\theta)$ is continuously strictly increasing. Before we start with the main argument, note that if $\hat{\theta}_1 > \theta_1$ then we can set $x_1^\star(\theta)$ to be equal to $x_1^\star(\theta_1)$ for all θ in $(\theta_1, \hat{\theta}_1)$. This strictly increases the objective function while maintaining feasibility. So we can assume $\hat{\theta}_1 \leq \theta_1$, which in turn implies that $\mu_1(\cdot)$ is non-negative in the interval (θ_1, θ_2).

  Now we give the main argument. Note that by Theorem 1 in Luenberger (1969, p. 217), $x_1^\star(\theta)$ must maximize the Lagrangean

$$\mathcal{L}(\boldsymbol{u}, \mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{w}) = u_1(w_1 - \lambda - \alpha_1) + u_2(\lambda - \alpha_2 + w_2) \\ + \int_0^{\bar{\theta}} x_1(z) \cdot \Big[\alpha_1 \mu_1(z) f_1(z) - \lambda \bar{F}_2(z)\Big] dz + \int_0^{\bar{\theta}} x_2(z) \cdot \Big[\alpha_2 \mu_2(z) f_2(z) + \lambda \bar{F}_2(z)\Big] dz,$$

  with $\lambda, w_1, w_2 \geq 0$. Define $L_1(\cdot)$ by

$$L_1(\theta) \triangleq \alpha_1 \mu_1(\theta) f_1(\theta) - \lambda \bar{F}_2(\theta),$$

  then it must be the case that $L_1(\theta) = 0$ for all $\theta \in (\theta_1, \theta_2)$. Suppose this is not true, then we could have $\hat{\theta} \in (\theta_1, \theta_2)$ such that $L_1(\hat{\theta}) > 0$, since $L_1(\cdot)$ is a continuous function this must also be true for all $\theta \in (\hat{\theta} - \epsilon, \hat{\theta} + \epsilon)$ for $\epsilon > 0$ small enough. But then we can obtain an strict improvement by setting $x_1(\theta) = x_1^\star(\hat{\theta} + \epsilon)$ for all $\theta \in (\hat{\theta} - \epsilon, \hat{\theta} + \epsilon)$. A similar argument holds when $L_1(\hat{\theta}) < 0$. Therefore, we have just proved that $L_1(\theta) = 0$ for all $\theta \in (\theta_1, \theta_2)$. In other words,

$$\alpha_1 \frac{\mu_1(\theta) f_1(z)}{\bar{F}_2(\theta)} = \lambda, \quad \forall \theta \in (\theta_1, \theta_2). \tag{11}$$

  Also, by the second mean value theorem for integrals there exists $\hat{\theta} \in (\theta_1, \theta_2)$ such that

$$x_1^\star(\hat{\theta}) = \frac{\int_{\theta_1}^{\theta_2} x_1^\star(z) \bar{F}_2(z) dz}{\int_{\theta_1}^{\theta_2} \bar{F}_2(z) dz}. \tag{12}$$

  Going back to $(\mathcal{P}_R^d)$, we have that the part of objective associated to $x_1^\star$ in (θ_1, θ_2) is

$$\int_{\theta_1}^{\theta_2} \alpha_1 x_1^\star(z) \mu_1(z) f_1(z) dz = \lambda \cdot \int_{\theta_1}^{\theta_2} x_1^\star(z) \bar{F}_2(z) dz, \tag{13}$$

where in the equality we have used (11). Now, consider modifying $x_1^\star$ to be $\tilde{x}_1^\star$ equal to $x_1^\star(\hat{\theta})$ in (θ_1, θ_2). Then from (11), (12) and (13) we get

$$\begin{aligned}
\int_{\theta_1}^{\theta_2} x_1^\star(z)\alpha_1\mu_1(z)f_1(z)dz &= \lambda \cdot x_1^\star(\hat{\theta}) \cdot \int_{\theta_1}^{\theta_2} \bar{F}_2(z)dz \\
&= x_1^\star(\hat{\theta}) \cdot \int_{\theta_1}^{\theta_2} \alpha_1\mu_1(z)f_1(z)dz \\
&= \int_{\theta_1}^{\theta_2} \tilde{x}_1^\star(z)\alpha_1\mu_1(z)f_1(z)dz,
\end{aligned}$$

therefore, the modified $\tilde{x}_1^\star$ has the same objective value than the old one. Also, note that we have preserved feasibility because

$$\begin{aligned}
u_1 + \int_0^{\bar{\theta}} \tilde{x}_1^\star(z)\bar{F}_2(z)dz &= u_1 + \int_{\theta_1}^{\theta_2} \tilde{x}_1^\star(z)\bar{F}_2(z)dz + \int_{(\theta_1,\theta_2)^c} \tilde{x}_1^\star(z)\bar{F}_2(z)dz \\
&= u_1 + x_1^\star(\hat{\theta}) \cdot \int_{\theta_1}^{\theta_2} \bar{F}_2(z)dz + \int_{(\theta_1,\theta_2)^c} x_1^\star(z)\bar{F}_2(z)dz \\
&\overset{(a)}{=} u_1 + \int_{\theta_1}^{\theta_2} x_1^\star(z)\bar{F}_2(z)dz + \int_{(\theta_1,\theta_2)^c} x_1^\star(z)\bar{F}_2(z)dz \\
&= u_1 + \int_0^{\bar{\theta}} x_1^\star(z)\bar{F}_2(z)dz,
\end{aligned}$$

where in (a) we used equation (12).

- **Ex-ante type 1 case (2):** Suppose for $x_1^\star(\cdot)$ there exists $\theta_1 < \theta_2 < \theta_3$ and $0 < \chi_1 < \chi_2 < 1$ such that $x_1^\star(\theta) = \chi_1$ in (θ_1, θ_2) and $x_1^\star(\theta) = \chi_2$ in (θ_2, θ_3). Since type's 1 allocation is piecewise constant we must have $x_1^\star(\theta_1^-) < \chi_1$ and $\chi_2 < x_1^\star(\theta_3^+)$.

  Then, the part of objective associated to ex-ante type 1 in these intervals is

$$\alpha_1 \cdot \chi_1 \cdot \int_{\theta_1}^{\theta_2} \mu_1(z)f_1(z)dz + \alpha_1 \cdot \chi_2 \cdot \int_{\theta_2}^{\theta_3} \mu_1(z)f_1(z)dz. \tag{14}$$

  If $\mu_1(\hat{\theta}) \leq 0$ for some $\hat{\theta} \in (\theta_1, \theta_3)$ then because of (DHR), $\mu_1(\theta) \leq 0$ for all $\theta \leq \hat{\theta}$ and, therefore, we can always find a better solution by setting $x_1^\star(\theta) = 0$ for all $\theta \leq \hat{\theta}$ (note that this does not affect feasibility in $(\mathcal{P}_R^d)$). So assume $\mu_1(\theta) > 0$ for all $\theta \in (\theta_1, \theta_3)$, then it must be the case that

$$u_2 + \int_0^{\bar{\theta}} x_2(z)\bar{F}_2(z)dz = u_1 + \int_0^{\bar{\theta}} x_1(z)\bar{F}_2(z)dz, \tag{15}$$

  otherwise we could increase χ_1 and obtain an strict improvement in the objective. There are two cases:

a) $\frac{\int_{\theta_1}^{\theta_2}\mu_1(z)f_1(z)dz}{\int_{\theta_1}^{\theta_2}\bar{F}_2(z)dz} \geq \frac{\int_{\theta_2}^{\theta_3}\mu_1(z)f_1(z)dz}{\int_{\theta_2}^{\theta_3}\bar{F}_2(z)dz}$: In this case consider decreasing χ_2 by $\epsilon_2 > 0$ and increasing χ_1 by $\epsilon_1 > 0$, in such a way that equation (15) remains with equality, that is,

$$\epsilon_1 \cdot \int_{\theta_1}^{\theta_2} \bar{F}_2(z)dz - \epsilon_2 \cdot \int_{\theta_2}^{\theta_3} \bar{F}_2(z)dz = 0. \tag{16}$$

The change in equation (14) is

$$\alpha_1 \cdot \frac{\epsilon_2 \cdot \int_{\theta_2}^{\theta_3}\bar{F}_2(z)dz}{\int_{\theta_1}^{\theta_2}\bar{F}_2(z)dz} \cdot \int_{\theta_1}^{\theta_2} \mu_1(z)f_1(z)dz - \alpha_1 \cdot \epsilon_2 \cdot \int_{\theta_2}^{\theta_3} \mu_1(z)f_1(z)dz, \tag{17}$$

which under our current assumption is non-negative. So we can weakly improve our objective, indeed we can do it so until $\chi_1 + \epsilon_1$ and $\chi_2 - \epsilon_2$ are equal,

$$\chi_1 + \epsilon_1 = \chi_2 - \epsilon_2 \Leftrightarrow \chi_1 + \epsilon_2 \cdot \frac{\int_{\theta_2}^{\theta_3}\bar{F}_2(z)dz}{\int_{\theta_1}^{\theta_2}\bar{F}_2(z)dz} = \chi_2 - \epsilon_2 \Leftrightarrow \epsilon_2 = \frac{(\chi_2 - \chi_1)}{1 + \frac{\int_{\theta_2}^{\theta_3}\bar{F}_2(z)dz}{\int_{\theta_1}^{\theta_2}\bar{F}_2(z)dz}},$$

since $\chi_2 > \chi_1$ we have $\epsilon_2 > 0$ and, therefore, we have shown that it is possible to increase χ_1 and to decrease χ_2 in such a way the objective is weakly improved and the solution is constant in (θ_1, θ_3).

b) $\frac{\int_{\theta_1}^{\theta_2}\mu_1(z)f_1(z)dz}{\int_{\theta_1}^{\theta_2}\bar{F}_2(z)dz} < \frac{\int_{\theta_2}^{\theta_3}\mu_1(z)f_1(z)d\theta}{\int_{\theta_2}^{\theta_3}\bar{F}_2(z)dz}$: In this case consider increasing χ_2 by $\epsilon_2 > 0$ and decreasing χ_1 by $\epsilon_1 > 0$ in such a way that equation (15) remains with equality. By doing this the change in the objective is strictly positive, and we do it until either $\chi_1 = x^\star(\theta_1^-)$ or $\chi_2 = x^\star(\theta_3^+)$.

This proves the result for ex-ante type 1 and case (2).

In conclusion, putting together what we have proved for cases (1) and (2), we can always consider $x_1^\star$ to be a step function with at most one intermediate step.

Now we proceed with ex-ante type 2.

- **Ex-ante type 2 case (1):** Suppose there is an interval (θ_1, θ_2) in which $x_2^\star(\theta)$ is continuously strictly increasing. Before we start with the main argument, note that if $\hat{\theta}_2 < \theta_2$ then we can set $x_2^\star(\theta)$ to be equal to $x_2^\star(\theta_2)$ for all θ in $(\hat{\theta}_2, \theta_2)$. This strictly increases the objective function and maintains feasibility. So we can assume $\hat{\theta}_2 \geq \theta_2$, which in turn implies that $\mu_2(\cdot)$ is non-positive in the interval (θ_1, θ_2).

Now we give the main argument. Note that by Theorem 1 in Luenberger (1969, p. 217), $x_2^\star(\theta)$ must maximize the Lagrangean

$$\begin{aligned}\mathcal{L}(\boldsymbol{u}, \mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{w}) &= u_1(w_1 - \lambda - \alpha_1) + u_2(\lambda - \alpha_2 + w_2) \\ &\quad + \int_0^{\bar{\theta}} x_1(z) \cdot \Big[\alpha_1 \mu_1(z) f_1(z) - \lambda \bar{F}_2(z)\Big] dz + \int_0^{\bar{\theta}} x_2(z) \cdot \Big[\alpha_2 \mu_2(z) f_2(z) + \lambda \bar{F}_2(z)\Big] dz,\end{aligned}$$

with $\lambda, w_1, w_2 \geq 0$. Define $L_2(\cdot)$ by

$$L_2(\theta) \triangleq \alpha_2 \mu_2(\theta) f_2(\theta) + \lambda \bar{F}_2(\theta),$$

then it must be the case that $L_2(\theta) = 0$ for all $\theta \in (\theta_1, \theta_2)$. Suppose this is not true, then we could have $\hat{\theta} \in (\theta_1, \theta_2)$ such that $L_2(\hat{\theta}) > 0$, since $L_2(\cdot)$ is a continuous function this must also be true for all $\theta \in (\hat{\theta} - \epsilon, \hat{\theta} + \epsilon)$ for $\epsilon > 0$ small enough. But then we can obtain an strict improvement by setting $x_2(\theta) = x_2^\star(\hat{\theta} + \epsilon)$ for all $\theta \in (\hat{\theta} - \epsilon, \hat{\theta} + \epsilon)$. A similar argument holds when $L_2(\hat{\theta}) < 0$. Therefore, we have just proved that $L_2(\theta) = 0$ for all $\theta \in (\theta_1, \theta_2)$. In other words,

$$\alpha_2 \frac{\mu_2(\theta) f_2(\theta)}{\bar{F}_2(\theta)} = -\lambda, \quad \forall \theta \in (\theta_1, \theta_2). \tag{18}$$

Also note that by the second mean value theorem for integrals, there exists $\hat{\theta} \in (\theta_1, \theta_2)$ such that

$$x_2^\star(\hat{\theta}) = \frac{\int_{\theta_1}^{\theta_2} x_2^\star(z) \bar{F}_2 dz}{\int_{\theta_1}^{\theta_2} \bar{F}_2(z) dz}. \tag{19}$$

Going back to $(\mathcal{P}_R^d)$, we have that the part of objective associated to $x_2^\star$ in (θ_1, θ_2) is

$$\int_{\theta_1}^{\theta_2} \alpha_2 x_2^\star(z) \mu_2(z) f_2(z) dz = -\lambda \cdot \int_{\theta_1}^{\theta_2} x_2^\star(z) \bar{F}_2(z) dz, \tag{20}$$

where in the equality we have used (18). Now, consider modifying $x_2^\star$ to be $\tilde{x}_2^\star$ equal to $x_2^\star(\hat{\theta})$ in (θ_1, θ_2). Then from (18), (19) and (20) we get

$$\begin{aligned}\int_{\theta_1}^{\theta_2} x_2^\star(z) \alpha_2 \mu_2(z) f_2(z) dz &= -\lambda \cdot x_2^\star(\hat{\theta}) \cdot \int_{\theta_1}^{\theta_2} \bar{F}_2(z) dz \\ &= x_2^\star(\hat{\theta}) \cdot \int_{\theta_1}^{\theta_2} \alpha_2 \mu_2(z) f_2(z) dz \\ &= \int_{\theta_1}^{\theta_2} \tilde{x}_2^\star(z) \alpha_2 \mu_2(z) f_2(z) dz,\end{aligned}$$

therefore, the modified $\tilde{x}_2^\star$ has the same objective value than the old one. Also, note that we have preserved feasibility because

$$\begin{aligned}
u_2 + \int_0^{\bar{\theta}} \tilde{x}_2^\star(z)\bar{F}_2(z)dz &= u_2 + \int_{\theta_1}^{\theta_2} \tilde{x}_2^\star(z)\bar{F}_2(z)dz + \int_{(\theta_1,\theta_2)^c} \tilde{x}_2^\star(z)\bar{F}_2(z)dz \\
&= u_2 + x_2^\star(\hat{\theta}) \cdot \int_{\theta_1}^{\theta_2} \bar{F}_2(z)dz + \int_{(\theta_1,\theta_2)^c} x_2^\star(z)\bar{F}_2(z)dz \\
&\overset{(a)}{=} u_2 + \int_{\theta_1}^{\theta_2} x_2^\star(z)\bar{F}_2(z)dz + \int_{(\theta_1,\theta_2)^c} x_2^\star(z)\bar{F}_2(z)dz \\
&= u_2 + \int_0^{\bar{\theta}} x_2^\star(z)\bar{F}_2(z)dz,
\end{aligned}$$

where in (a) we used equation (19).

- **Ex-ante type 2 case (2):** Suppose $x_2^\star(\cdot)$ is an optimal solution to $(\mathcal{P}_R^d)$ for which there exists $\theta_1 < \theta_2$ and $0 < \chi < 1$ such that $x_2^\star(\theta) = \chi$ in (θ_1, θ_2). Similar to the proof of type 1 assume $x_2^\star(\theta_1^-) < \chi < x_2^\star(\theta_2^+)$.

  Then the part of the objective for the ex-ante type 2 in this interval is

$$\alpha_2 \cdot \chi \cdot \int_{\theta_1}^{\theta_2} \mu_2(z) f_2(z)dz. \tag{21}$$

  If $\mu_2(\hat{\theta}) \geq 0$ for some $\hat{\theta} \in (\theta_1, \theta_2)$ then because of (DHR), $\mu_2(\theta) \geq 0$ for all $\theta \geq \hat{\theta}$ and, therefore, we can always find a better solution by setting $x_2^\star(\theta) = 1$ for all $\theta \geq \hat{\theta}$ (note that this does not affect feasibility in $(\mathcal{P}_R^d)$). So assume $\mu_2(\theta) < 0$ for all $\theta \in (\theta_1, \theta_2)$, then it must be the case that

$$u_2 + \int_0^{\bar{\theta}} x_2(z)\bar{F}_2(z)dz = u_1 + \int_0^{\bar{\theta}} x_1(z)\bar{F}_2(z)dz, \tag{22}$$

  otherwise we could decrease χ and obtain an strict improvement in the objective. Now, consider splitting the interval in half, that is, take $\hat{\theta} = (\theta_1 + \theta_2)/2$ and note that because of (DHR) we always have

$$\frac{\int_{\theta_1}^{\hat{\theta}} \mu_2(z) f_2(z)dz}{\int_{\theta_1}^{\hat{\theta}} \bar{F}_2(z)dz} \leq \frac{\int_{\hat{\theta}}^{\theta_2} \mu_2(z) f_2(z)dz}{\int_{\hat{\theta}}^{\theta_2} \bar{F}_2(z)dz}. \tag{23}$$

  We can modify $x_2^\star(\theta)$ in (θ_1, θ_2) as follows and obtain an, at least weakly, objective improvement. For $\theta \in (\theta_1, \hat{\theta})$ set $x_2^\star(\theta) = \chi - \epsilon_1$ and for $\theta \in (\hat{\theta}, \theta_2)$ set $x_2^\star(\theta) = \chi + \epsilon_2$ with $\epsilon_1, \epsilon_2 > 0$, and such that equation (22) remains with equality. That is,

$$-\epsilon_1 \cdot \int_{\theta_1}^{\hat{\theta}} \bar{F}_2(z)dz + \epsilon_2 \cdot \int_{\hat{\theta}}^{\theta_2} \bar{F}_2(z)dz = 0.$$

With this modification the change in the objective is

$$-\alpha_2 \cdot \frac{\epsilon_2 \cdot \int_{\hat{\theta}}^{\theta_2} \bar{F}_2(z)dz}{\int_{\theta_1}^{\hat{\theta}} \bar{F}_2(z)dz} \cdot \int_{\theta_1}^{\hat{\theta}} \mu_2(z)f_2(z)dz + \alpha_2 \cdot \epsilon_2 \cdot \int_{\hat{\theta}}^{\theta_2} \mu_2(z)f_2(z)dz,$$

which thanks to equation (23) is non-negative. Then we can keep increasing ϵ_2 until either $\chi - \epsilon_1 = x_2^\star(\theta_1^-)$ or $\chi + \epsilon_2 = x_2^\star(\theta_2^+)$. This proofs we can, at least weakly, improve the objective. It also proves that we can modify the solution in such a way that for one of the two halves of the intervals the step reaches the boundary bound given by either $x_2^\star(\theta_1^-)$ or $x_2^\star(\theta_2^+)$. For the half that did not reach the boundary, we can do the same procedure described above and then repeat this procedure until we completely get rid of the intermediate step between $(x_2^\star(\theta_1^-), x_2^\star(\theta_2^+))$. Note that this process can be potentially infinite, in which case a more rigorous argument is required.

Suppose the process described above goes for infinitely many steps. In this case, an allocation sequence $\{x_2^n(\theta)\}_{n\in\mathbb{N}}$ defined in $[\theta_1, \theta_2]$ is generated. To prove that the argument works, we need to show that there exists $\theta_\infty \in [\theta_1, \theta_2]$ such that

$$\lim_{n\to\infty} \int_{\theta_1}^{\theta_2} x_2^n(z)\mu_2(z)f_2(z)dz = x_2^\star(\theta_1) \int_{\theta_1}^{\theta_\infty} \mu_2(z)f_2(z)dz + x_2^\star(\theta_2) \int_{\theta_\infty}^{\theta_2} \mu_2(z)f_2(z)dz. \tag{24}$$

To prove this, let $\{\underline{\theta}_n, \overline{\theta}_n, \hat{\theta}_n\}_{n\in\mathbb{N}}$ be the sequence generated in the infinite process where:

- $\underline{\theta}_n$ and $\overline{\theta}_n$ correspond to the lower and upper bound of the interval. For example, at the beginning $\underline{\theta}_1 = \theta_1$ and $\overline{\theta}_1 = \theta_2$. At the next iteration we will have either $\underline{\theta}_2 = \theta_1$ and $\overline{\theta}_2 = \hat{\theta}$ or $\underline{\theta}_2 = \hat{\theta}$ and $\overline{\theta}_2 = \theta_2$. Note that for all $n \in \mathbb{N}$: $\underline{\theta}_n, \overline{\theta}_n \in [\theta_1, \theta_2]$.
- $\hat{\theta}_n$ is defined to be the half of the interval. So $\hat{\theta}_1 = \hat{\theta}$, and $\hat{\theta}_2 = (\underline{\theta}_2 + \overline{\theta}_2)/2$.

From these definitions we have that $\underline{\theta}_n$ and $\overline{\theta}_n$ are bounded monotonic sequences (the first non-decreasing and the second non-increasing), thus both converge to a limit. Also,

$$\hat{\theta}_n = \frac{\underline{\theta}_n + \overline{\theta}_n}{2},$$

then all three quantities, $\underline{\theta}_n, \overline{\theta}_n$ and $\hat{\theta}_n$, converge to the same limit which we denote by $\theta_\infty \in [\theta_1, \theta_2]$ (if the limit was not the same we could continue iterating the process). Now, from this we have that

$$\lim_{n\to\infty} x_2^n(\theta) = \begin{cases} x_2^\star(\theta_1^-) & \text{if } \theta < \theta_\infty \\ x_2^\star(\theta_2^+) & \text{if } \theta \geq \theta_\infty, \end{cases} \quad \text{a.s in } [\theta_1, \theta_2]. \tag{25}$$

To see why (25) holds, consider $\theta \in [\theta_1, \theta_\infty)$ then from the convergence of $\underline{\theta}_n$ we have

$$\exists n_0 \in \mathbb{N}, \forall n \geq n_0, \quad \theta < \underline{\theta}_n \leq \theta_\infty.$$

Then, from the way x_2^n is constructed, it must be the case that $x_2^n(\theta)$ equals $x_2^\star(\theta_1^-)$. A similar argument holds for $\theta \in (\theta_\infty, \theta_2]$. Thus, $x_2^n(\theta)$ satisfies the almost surely convergence in equation (25). Finally, we can use the almost surely version of the dominated convergence theorem to obtain (24). This completes the proof for ex-ante type 2 and case (2).

Proof for the reminder of the properties:

From the previous discussion we can write down $(\mathcal{P}_R^d)$ as follows

$$\begin{aligned}
\max \quad & -\sum_{k=1}^{2} \alpha_k u_k + \alpha_1 \chi \int_{\theta_1}^{\theta_3} \mu_1(z) f_1(z) dz + \alpha_1 \int_{\theta_3}^{\bar{\theta}} \mu_1(z) f_1(z) dz + \alpha_2 \int_{\theta_2}^{\bar{\theta}} \mu_2(z) f_2(z) dz \\
\text{s.t} \quad & \chi \in [0,1], \quad \theta_1 \leq \theta_3 \\
& u_k \geq 0, k = 1, 2 \\
& u_2 + \int_{\theta_2}^{\bar{\theta}} \bar{F}_2(z) dz \geq u_1 + \chi \int_{\theta_1}^{\theta_3} \bar{F}_2(z) dz + \int_{\theta_3}^{\bar{\theta}} \bar{F}_2(z) dz.
\end{aligned}$$

- $u_1 = 0$: From the formulation above it is clear that is always optimal to set $u_1 = 0$.

- $\hat{\theta}_1 \leq \theta_1$: Suppose the opposite, that is, $\hat{\theta}_1 > \theta_1$. This implies that between θ_1 and $\hat{\theta}_1$, $\mu_1(\cdot)$ is negative. Then, we can increase θ_1 while keeping feasibility and, at the same time, increasing the objective function. Note this argument is also valid when $\theta_1 = \theta_3$. Also, note that we can obtain a strict improvement only when $\chi > 0$; however, when $\chi = 0$ we can only obtain a weak improvement. In either case, we can always consider $\hat{\theta}_1 \leq \theta_1$.

- $\theta_2 \leq \hat{\theta}_2$: Suppose the opposite, $\theta_2 > \hat{\theta}_2$. Since $\mu(\theta) > 0$ for all $\theta \geq \hat{\theta}_2$, we can can decrease θ_2 and obtain an objective improvement while maintaining feasibility.

- $u_2 = 0$: Suppose $u_2 > 0$, then we must have

$$u_2 + \int_{\theta_2}^{\bar{\theta}} \bar{F}_2(z) dz = \chi \int_{\theta_1}^{\theta_3} \bar{F}_2(z) dz + \int_{\theta_3}^{\bar{\theta}} \bar{F}_2(z) dz, \tag{26}$$

  otherwise, we could decrease u_2 and, by doing so, improve the objective.

  Since $u_2 > 0$, equation (26) yields

$$0 < u_2 = \chi \int_{\theta_1}^{\theta_3} \bar{F}_2(z) dz + \int_{\theta_3}^{\bar{\theta}} \bar{F}_2(z) dz - \int_{\theta_2}^{\bar{\theta}} \bar{F}_2(z) dz, \tag{27}$$

then it must be true that $\theta_1 < \theta_2$; otherwise, from equation (27) we would have

$$\int_{\theta_2}^{\theta_1} \bar{F}_2(z)dz + \int_{\theta_1}^{\theta_3} \bar{F}_2(z)dz + \int_{\theta_3}^{\bar{\theta}} \bar{F}_2(z)dz < \chi \int_{\theta_1}^{\theta_3} \bar{F}_2(z)dz + \int_{\theta_3}^{\bar{\theta}} \bar{F}_2(z)dz,$$

which implies

$$\int_{\theta_2}^{\theta_1} \bar{F}_2(z)dz < 0,$$

a contradiction. Thus, $\theta_1 < \theta_2$.

Now consider, a new contract for type 2 which consists on decreasing the cut-off θ_2 by $\epsilon > 0$ sufficiently small, but at the same time maintaining the equality in equation (26). Specifically, let $\theta_2(\epsilon) = \theta_2 - \epsilon > 0$ (this is true because we just saw that $\theta_2 > \theta_1 \geq 0$) and let $u_2(\epsilon)$ be

$$u_2(\epsilon) = \chi \int_{\theta_1}^{\theta_3} \bar{F}_2(z)dz + \int_{\theta_3}^{\bar{\theta}} \bar{F}_2(z)dz - \int_{\theta_2(\epsilon)}^{\bar{\theta}} \bar{F}_2(z)dz,$$

note that by taking ϵ small we still have $u_2(\epsilon) > 0$. We claim that this new contract, characterized by $\theta_1, \theta_3, \chi, \theta_2(\epsilon)$ and $u_2(\epsilon)$, yields a larger objective that the old contract, characterized by $\theta_1, \theta_3, \chi, \theta_2$ and u_2. The old contract objective's is

$$-\alpha_2 u_2 + \alpha_1 \chi \int_{\theta_1}^{\theta_3} \mu_1(z) f_1(z)dz + \alpha_1 \int_{\theta_3}^{\bar{\theta}} \mu_1(z) f_1(z)dz + \alpha_2 \int_{\theta_2}^{\bar{\theta}} \mu_2(z) f_2(z)dz,$$

and using equation (26) it becomes

$$\chi \int_{\theta_1}^{\theta_3} (\alpha_1 \mu_1(z) f_1(z) - \alpha_2 \bar{F}_2(z))dz + \int_{\theta_3}^{\bar{\theta}} (\alpha_1 \mu_1(z) f_1(z) - \alpha_2 \bar{F}_2(z))dz + \alpha_2 \int_{\theta_2}^{\bar{\theta}} z f_2(z)dz.$$

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_2. Hence, the new contract yields an improvement over the old one if and only if

$$\int_{\theta_2}^{\bar{\theta}} z f_2(z)dz < \int_{\theta_2(\epsilon)}^{\bar{\theta}} z f_2(z)dz.$$

Since $\theta_2(\epsilon) < \theta_2$ this last inequality is true. Thus, if $u_2 > 0$ we can always construct a new contract yielding a larger objective value and, therefore, at any optimal contract we must have $u_2 = 0$.

- $\theta_2 \leq \theta_3$: Since at any optimal solution $u_2 = 0$, the IC constraint is

$$\int_{\theta_2}^{\bar{\theta}} \bar{F}_2(z)dz \geq \chi \int_{\theta_1}^{\theta_3} \bar{F}_2(z)dz + \int_{\theta_3}^{\bar{\theta}} \bar{F}_2(z)dz.$$

Hence, if $\theta_2 > \theta_3$ from the expression above we would have

$$\int_{\theta_2}^{\bar{\theta}} \bar{F}_2(z)dz \geq \chi \int_{\theta_1}^{\theta_3} \bar{F}_2(z)dz + \int_{\theta_3}^{\theta_2} \bar{F}_2(z)dz + \int_{\theta_2}^{\bar{\theta}} \bar{F}_2(z)dz,$$

which implies $\theta_2 = \theta_3$, a contradiction.

- $\theta_1 \leq \theta_2$: First we show that $\theta_1 \leq \hat{\theta}_2$. Suppose the opposite, that is, $\theta_1 > \hat{\theta}_2$. Then, since $\hat{\theta}_2 \geq \theta_2$ we must have $\theta_1 > \theta_2$ and, therefore,

$$\begin{aligned}\int_{\theta_2}^{\bar{\theta}} \bar{F}_2(z)dz &= \int_{\theta_2}^{\theta_1} \bar{F}_2(z)dz + \int_{\theta_1}^{\bar{\theta}} \bar{F}_2(z)dz \\ &> \int_{\theta_1}^{\bar{\theta}} \bar{F}_2(z)dz \\ &= \int_{\theta_1}^{\theta_3} \bar{F}_2(z)dz + \int_{\theta_3}^{\bar{\theta}} \bar{F}_2(z)dz \\ &\geq \chi \int_{\theta_1}^{\theta_3} \bar{F}_2(z)dz + \int_{\theta_3}^{\bar{\theta}} \bar{F}_2(z)dz.\end{aligned}$$

That is, the IC constraint is not binding. Therefore, since $\theta_1 > \hat{\theta}_2 \geq \hat{\theta}_1$ we can slightly decrease θ_1 and, in this way, obtain an objective improvement whenever $\chi > 0$. When $\chi = 0$, because $\theta_3 \geq \theta_1$, we can decrease θ_3 and obtain an objective improvement as well. Hence, at any optimal solution we must have $\theta_1 \leq \hat{\theta}_2$.

In order to complete the proof, suppose $\theta_1 > \theta_2$ then, as before, we have

$$\int_{\theta_2}^{\bar{\theta}} \bar{F}_2(z)dz > \chi \int_{\theta_1}^{\theta_3} \bar{F}_2(z)dz + \int_{\theta_3}^{\bar{\theta}} \bar{F}_2(z)dz.$$

Using that $\theta_1 \leq \hat{\theta}_2$ implies $\theta_2 < \hat{\theta}_2$, we can slightly increase θ_2 (maintaining feasibility) and thus obtain an objective improvement. In conclusion, at any optimal solution we must have $\theta_1 \leq \theta_2$.

■

Proof of Corollary 4. From Theorem 1 we know that we can formulate $(\mathcal{P}_R^d)$ as

$$\begin{aligned}(\mathcal{P}_R^d) \quad \max \quad & \alpha_1\chi \int_{\theta_1}^{\theta_3} \mu_1(z)f_1(z)dz + \alpha_1 \int_{\theta_3}^{\bar{\theta}} \mu_1(z)f_1(z)dz + \alpha_2 \int_{\theta_2}^{\bar{\theta}} \mu_2(z)f_2(z)dz \\ \text{s.t} \quad & \chi \in [0,1] \\ & \hat{\theta}_1 \leq \theta_1 \leq \theta_2 \leq \theta_3, \theta_2 \leq \hat{\theta}_2 \\ & \int_{\theta_2}^{\bar{\theta}} \bar{F}_2(z)dz \geq \chi \int_{\theta_1}^{\theta_3} \bar{F}_2(z)dz + \int_{\theta_3}^{\bar{\theta}} \bar{F}_2(z)dz.\end{aligned}$$

It is easy to see that if the static contract is an optimal solution to $(\mathcal{P}_R^d)$ then it is also an optimal solution to $(\mathcal{P}^d)$. This is true because the optimal value of $(\mathcal{P}_R^d)$ is always an upper bound to the optimal value of $(\mathcal{P}^d)$, and the static contract is always feasible for $(\mathcal{P}^d)$.

For the other direction, suppose that the static contract is an optimal solution to $(\mathcal{P}^d)$ but is not an optimal solution to $(\mathcal{P}_R^d)$. We will find a contract that is feasible for $(\mathcal{P}^d)$ and yields a larger objective than the the static contract.

Let $\theta_1, \theta_2, \theta_3$ and χ be the optimal solution to $(\mathcal{P}_R^d)$. Then, it must be the case that

$$\int_{\theta_2}^{\bar{\theta}} \bar{\mu}(z)dz < \alpha_1 \chi \int_{\theta_1}^{\theta_3} \mu_1(z) f_1(z)dz + \alpha_1 \int_{\theta_3}^{\bar{\theta}} \mu_1(z) f_1(z)dz + \alpha_2 \int_{\theta_2}^{\bar{\theta}} \mu_2(z) f_2(z)dz. \tag{28}$$

This is true because the static contract $(u_1, u_2, x_1, x_2) = (0, 0, \mathbf{1}_{\{\theta \geq \theta_2\}}, \mathbf{1}_{\{\theta \geq \theta_2\}})$ is a feasible contract for $(\mathcal{P}^d)$ and, therefore, it must yield a lower objective than the optimal static contract. Under the current assumption, the optimal static contract yields a strictly lower objective than the solution to $(\mathcal{P}_R^d)$. Therefore, equation (28) holds.

From the formulation of $(\mathcal{P}_R^d)$ we know that $\hat{\theta}_1 \leq \theta_1 \leq \theta_2 \leq \theta_3$. Then, this and equation (28) deliver

$$0 \leq \int_{\theta_2}^{\theta_3} \mu_1(z) f_1(z)dz < \chi \int_{\theta_1}^{\theta_3} \mu_1(\theta) f_1(z)dz.$$

Hence, $\theta_1 < \theta_3$ and

$$\frac{\int_{\theta_2}^{\theta_3} \mu_1(z) f_1(z)dz}{\int_{\theta_1}^{\theta_3} \mu_1(z) f_1(z)dz} < \chi. \tag{29}$$

Also, since $\chi \leq 1$ we must have $\theta_1 < \theta_2$.

Now we argue that the contract optimizing $(\mathcal{P}_R^d)$ is feasible for $(\mathcal{P}^d)$. Since the high to low IC constraint is satisfied, we only need to verify the low to high IC constraint. That is, we need to verify the following inequality

$$\chi \int_{\theta_1}^{\theta_3} \bar{F}_1(z)dz + \int_{\theta_3}^{\bar{\theta}} \bar{F}_1(z)dz \geq \int_{\theta_2}^{\bar{\theta}} \bar{F}_1(z)dz,$$

or equivalently

$$\chi \geq \frac{\int_{\theta_2}^{\theta_3} \bar{F}_1(z)dz}{\int_{\theta_1}^{\theta_3} \bar{F}_1(z)dz}. \tag{30}$$

When $\theta_2 = \theta_3$, equation (30) trivially holds. So, assume $\theta_2 < \theta_3$. Then, to see why (30) continues to hold in this case, observe that from Lemma 5 we have

$$\frac{\int_{\theta_1}^{\theta_3} \mu_1(z) f_1(z)dz}{\int_{\theta_1}^{\theta_3} \bar{F}_1(z)dz} \leq \frac{\int_{\theta_2}^{\theta_3} \mu_1(z) f_1(z)dz}{\int_{\theta_2}^{\theta_3} \bar{F}_1(z)dz} \Leftrightarrow \frac{\int_{\theta_1}^{\theta_2} \mu_1(z) f_1(z)dz}{\int_{\theta_1}^{\theta_2} \bar{F}_1(z)dz} \leq \frac{\int_{\theta_2}^{\theta_3} \mu_1(z) f_1(z)dz}{\int_{\theta_2}^{\theta_3} \bar{F}_1(z)dz}. \tag{31}$$

The right hand side in (31) always hold because thanks to (DHR) we have:

$$\frac{\int_{\theta_1}^{\theta_2} \mu_1(z) f_1(z) dz}{\int_{\theta_1}^{\theta_2} \bar{F}_1(z) dz} = \frac{\int_{\theta_1}^{\theta_2} \bar{F}_1 c^{11}(z) dz}{\int_{\theta_1}^{\theta_2} \bar{F}_1(z) dz} \leq c^{11}(\theta_2) \leq \frac{\int_{\theta_2}^{\theta_3} \bar{F}_1 c^{11}(z) dz}{\int_{\theta_2}^{\theta_3} \bar{F}_1(z) dz} = \frac{\int_{\theta_2}^{\theta_3} \mu_1(z) f_1(z) dz}{\int_{\theta_2}^{\theta_3} \bar{F}_1(z) dz}.$$

Thus the left hand side in (31) holds. Equivalently,

$$\frac{\int_{\theta_2}^{\theta_3} \bar{F}_1(z) dz}{\int_{\theta_1}^{\theta_3} \bar{F}_1(z) dz} \leq \frac{\int_{\theta_2}^{\theta_3} \mu_1(z) f_1(z) dz}{\int_{\theta_1}^{\theta_3} \mu_1(z) f_1(z) dz}$$

Using this, together with equation (29), delivers equation (30). This concludes the proof. ■

Proof of Theorem 2. The proof relies on the global theory of constrained optimization. Specifically, we make use of Theorem 1 in Luenberger (1969, p. 217) and of Theorem 1 in Luenberger (1969, p. 220).

We begin by seeting up the stage for the proof. Define the set of functions

$$\mathcal{F} \triangleq \{x : [0, \bar{\theta}] \longrightarrow [0,1] : x(\cdot) \text{ is non-decreasing}\},$$

that is, $\mathcal{F}$ is the set of all feasible allocations. Then, the domain we are optimizing on is

$$\Omega \triangleq \mathbb{R} \times \mathbb{R} \times \mathcal{F} \times \mathcal{F}.$$

An element of Ω is (u_1, u_2, x_1, x_2), the ex-post utility for the lowest ex-post type for both ex-ante types and, the allocation schedule for each ex-ante type. The constraints of the problem are the ex-post individually rationality constraints and the ex-ante incentive compatibility constraints. The optimization problem is then

$$(\mathcal{P}^d) \qquad \max_{(u_1,u_2,x_1,x_2)\in\Omega} \quad -\alpha_1 u_1 - \alpha_2 u_2 + \int_0^{\bar{\theta}} \Big(\alpha_1 x_1(z)\mu_1(z) f_1(z) + \alpha_2 x_2(z)\mu_2(z) f_2(z)\Big) d\theta$$

$$\text{s.t} \quad u_1 + \int_0^{\bar{\theta}} x_1(z)\bar{F}_1(z) dz \geq u_2 + \int_0^{\bar{\theta}} x_2(z)\bar{F}_1(z) dz$$

$$u_2 + \int_0^{\bar{\theta}} x_2(z)\bar{F}_2(z) dz \geq u_1 + \int_0^{\bar{\theta}} x_1(z)\bar{F}_2(z) dz$$

$$u_1, u_2 \geq 0.$$

Note that from Corollary 4 we can relax the low to high IC constraint. We present here a more general proof which does not relay on any relaxation. Nonetheless, the next argument also applies to the relaxed formulation.

The lagrangian for this problem is

$$\begin{aligned}\mathcal{L}(\boldsymbol{u}, \mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{w}) \triangleq u_1 \cdot (-\alpha_1 + w_1 + \lambda_1 - \lambda_2) + \int_0^{\bar{\theta}} x_1(z) \cdot \Big[\alpha_1 \mu_1(z) f_1(z) + \lambda_1 \bar{F}_1(z) - \lambda_2 \bar{F}_2(z)\Big] dz \\ + u_2 \cdot (-\alpha_2 + w_2 - \lambda_1 + \lambda_2) + \int_0^{\bar{\theta}} x_2(z) \cdot \Big[\alpha_2 \mu_2(z) f_2(z) - \lambda_1 \bar{F}_1(z) + \lambda_2 \bar{F}_2(z)\Big] dz,\end{aligned}$$

where $w_1, w_2 \geq 0$ are the multipliers associated to the ex-post individually rationality constraints and, $\lambda_1, \lambda_2 \geq 0$ are the multipliers associated to the ex-ante incentive compatibility constraints.

Now we are ready to begin the proof. We prove both implications separately. Suppose first that the static contract is optimal, we want to prove that condition (NR) holds. We proceed by contradiction. So assume (NR) does not hold, then by Lemma 6 there exist $\theta_a < \theta^s < \theta_b$ such that

$$\frac{\int_{\theta_a}^{\theta^s} \bar{F}_2(z) c^{12}(z) dz}{\int_{\theta_a}^{\theta^s} \bar{F}_2(z) dz} > \frac{\int_{\theta^s}^{\theta_b} \bar{F}_2(z) c^{12}(z) dz}{\int_{\theta^s}^{\theta_b} \bar{F}_2(z) dz}. \tag{32}$$

We resort to Theorem 1 in Luenberger (1969, p. 217). In order to use the theorem we need to verify the interior, or Slater, condition. So we need to find $(u_1, u_2, x_1, x_2) \in \Omega$ such that

$$\begin{aligned}&u_1 > 0, u_2 > 0, \\ &u_1 + \int_0^{\bar{\theta}} x_1(\theta) \bar{F}_1(\theta) d\theta > u_2 + \int_0^{\bar{\theta}} x_2(\theta) \bar{F}_1(\theta) d\theta \\ &u_2 + \int_0^{\bar{\theta}} x_2(\theta) \bar{F}_2(\theta) d\theta > u_1 + \int_0^{\bar{\theta}} x_1(\theta) \bar{F}_2(\theta) d\theta.\end{aligned}$$

To see why this is true, consider equation (32) and note that together with Lemma 5 it implies

$$\frac{\int_{\theta^s}^{\theta_b} \bar{F}_2(z) dz}{\int_{\theta_a}^{\theta_b} \bar{F}_2(z) dz} > \frac{\int_{\theta^s}^{\theta_b} \bar{F}_2(z) c^{12}(z) dz}{\int_{\theta_a}^{\theta_b} \bar{F}_2(z) c^{12}(z) dz}. \tag{33}$$

Also, note that

$$\begin{aligned}
\frac{\int_{\theta^s}^{\theta_b} \bar{F}_2(z)c^{12}(z)dz}{\int_{\theta^s}^{\theta_b} \bar{F}_1(z)dz} &= \frac{\int_{\theta^s}^{\theta_b} \bar{F}_1(z)c^{11}(z)dz}{\int_{\theta^s}^{\theta_b} \bar{F}_1(z)dz} \\
&\geq c^{11}(\theta^s)\frac{\int_{\theta^s}^{\theta_b} \bar{F}_1(z)dz}{\int_{\theta^s}^{\theta_b} \bar{F}_1(z)dz} \\
&= c^{11}(\theta^s) \\
&\geq c^{11}(\theta^s)\frac{\int_{\theta_a}^{\theta^s} \bar{F}_1(z)dz}{\int_{\theta_a}^{\theta^s} \bar{F}_1(z)dz} \\
&\geq \frac{\int_{\theta_a}^{\theta^s} \bar{F}_1(z)c^{11}(z)dz}{\int_{\theta_a}^{\theta^s} \bar{F}_1(z)dz} \\
&= \frac{\int_{\theta_a}^{\theta^s} \bar{F}_2(z)c^{12}(z)dz}{\int_{\theta_a}^{\theta^s} \bar{F}_1(z)dz},
\end{aligned}$$

where the inequalities come from the fact that $c^{11}(\cdot)$ is an increasing function. This in turn yields

$$\frac{\int_{\theta^s}^{\theta_b} \bar{F}_2(z)c^{12}(z)dz}{\int_{\theta_a}^{\theta_b} \bar{F}_2(z)c^{12}(z)dz} \geq \frac{\int_{\theta^s}^{\theta_b} \bar{F}_1(z)dz}{\int_{\theta_a}^{\theta_b} \bar{F}_1(z)dz}. \tag{34}$$

Putting equations (33) and (34) together implies

$$\exists \chi \in (0,1): \quad \frac{\int_{\theta^s}^{\theta_b} \bar{F}_2(z)dz}{\int_{\theta_a}^{\theta_b} \bar{F}_2(z)dz} > \chi > \frac{\int_{\theta^s}^{\theta_b} \bar{F}_1(z)dz}{\int_{\theta_a}^{\theta_b} \bar{F}_1(z)dz}.$$

Now, take $u_1 = u_2 > 0$ and

$$x_1(\theta) = \begin{cases} 0 & \text{if } \theta < \theta_a \\ \chi & \text{if } \theta_a \leq \theta \leq \theta_b \\ 1 & \text{if } \theta_b < \theta \end{cases}$$

and

$$x_2(\theta) = \begin{cases} 0 & \text{if } \theta < \theta^s \\ 1 & \text{if } \theta^s \leq \theta, \end{cases}$$

then is not hard to check that for this choice of $(u_1, u_2, x_1, x_2) \in \Omega$ the interior condition is satisfied, as required. Theorem 1 in Luenberger (1969, p. 217) gives then the existence of Lagrange multipliers and it also states that the static contract

$$(u_1, u_2, x_1, x_2) = (0, 0, \mathbf{1}_{\{\theta \geq \theta^s\}}, \mathbf{1}_{\{\theta \geq \theta^s\}}),$$

should maximize the lagrangean. In other words, $\exists, \boldsymbol{\lambda}, \boldsymbol{w} \geq 0$ such that

$$\mathcal{L}(0,0,\mathbf{1}_{\{\theta\geq\theta^s\}},\mathbf{1}_{\{\theta\geq\theta^s\}},\boldsymbol{\lambda},\boldsymbol{w}) \geq \mathcal{L}(\boldsymbol{u},\mathbf{x},\boldsymbol{\lambda},\boldsymbol{w}), \quad \forall (u_1,u_2,x_1,x_2)\in\Omega. \tag{35}$$

Since this is for any $(u_1, u_2, x_1, x_2) \in \Omega$ we can take $u_1, u_2 = 0$, x_2 defined as above and two possible x_1: one equal to $\mathbf{1}_{\{\theta\geq\theta_a\}}$ and the other equal to $\mathbf{1}_{\{\theta\geq\theta_b\}}$. Then from (35) we get

$$\alpha_1 \cdot \frac{\int_{\theta_a}^{\theta^s} \bar{F}_2(z)c^{12}(z)dz}{\int_{\theta_a}^{\theta^s} \bar{F}_2(z)dz} + \lambda_1 \cdot \frac{\int_{\theta_a}^{\theta^s} \bar{F}_1(z)dz}{\int_{\theta_a}^{\theta^s} \bar{F}_2(z)dz} \leq \lambda_2 \leq \alpha_1 \cdot \frac{\int_{\theta^s}^{\theta_b} \bar{F}_2(z)c^{12}(z)dz}{\int_{\theta^s}^{\theta_b} \bar{F}_2(z)dz} + \lambda_1 \cdot \frac{\int_{\theta^s}^{\theta_b} \bar{F}_1(z)dz}{\int_{\theta^s}^{\theta_b} \bar{F}_2(z)dz},$$

this and equation (32) imply

$$\lambda_1 \cdot \frac{\int_{\theta_a}^{\theta^s} \bar{F}_1(z)dz}{\int_{\theta_a}^{\theta^s} \bar{F}_2(z)dz} < \lambda_1 \cdot \frac{\int_{\theta^s}^{\theta_b} \bar{F}_1(z)dz}{\int_{\theta^s}^{\theta_b} \bar{F}_2(z)dz},$$

if $\lambda_1 = 0$ we get a contradiction. While if $\lambda_1 > 0$ from equation (32) we deduce

$$\frac{\int_{\theta^s}^{\theta_b} \bar{F}_1(z)c^{11}(z)dz}{\int_{\theta_a}^{\theta^s} \bar{F}_1(z)c^{11}(z)dz} < \frac{\int_{\theta^s}^{\theta_b} \bar{F}_2(z)dz}{\int_{\theta_a}^{\theta^s} \bar{F}_2(z)dz} < \frac{\int_{\theta^s}^{\theta_b} \bar{F}_1(z)dz}{\int_{\theta_a}^{\theta^s} \bar{F}_1(z)dz},$$

which in turn implies

$$c^{11}(\theta^s) \leq \frac{\int_{\theta^s}^{\theta_b} \bar{F}_1(z)c^{11}(z)dz}{\int_{\theta^s}^{\theta_b} \bar{F}_1(z)dz} < \frac{\int_{\theta_a}^{\theta^s} \bar{F}_1(z)c^{11}(z)dz}{\int_{\theta_a}^{\theta^s} \bar{F}_1(z)dz} \leq c^{11}(\theta_s), \tag{36}$$

a contradiction.

For the other direction we assume condition (NR) holds and we want to verify the static contract is optimal. In order to do so we use Theorem 1 in Luenberger (1969, p. 220). This theorem states that if we are able to find lagrange multipliers $\boldsymbol{\lambda}, \boldsymbol{w} \geq 0$ for which equation (35) holds, then the static contract is optimal.

So, set the lagrange multipliers as follows

$$w_1 = \alpha_1 - \alpha_2 c^{22}(\theta_s), w_2 = \alpha_2 + \alpha_2 c^{22}(\theta_s), \lambda_1 = 0, \lambda_2 = -\alpha_2 c^{22}(\theta_s), \tag{37}$$

these multipliers are non-negative because $c^{22}(\theta_s) \leq 0$ and

$$w_2 = \alpha_2 + \alpha_2 c^{22}(\theta_s) \geq 0 \Leftrightarrow c^{22}(\theta_s) \geq -1 \Leftrightarrow [\theta_s - h^{22}(\theta_s)] \geq -h^{22}(\theta_s) \Leftrightarrow \theta_s \geq 0.$$

We first claim that

$$\max_{x_1\in\Omega} \int_0^{\bar{\theta}} x_1(z) \cdot \Big[\alpha_1\mu_1(z)f_1(z) - \lambda_2\bar{F}_2(z)\Big]dz = \int_{\theta_s}^{\bar{\theta}} \Big[\alpha_1\mu_1(z)f_1(z) - \lambda_2\bar{F}_2(z)\Big]dz. \tag{38}$$

To prove this first note that the optimal solution x_1 on the left hand side of (38) is of the threshold type, that is, $x_1 = \mathbf{1}_{\{\theta \geq \theta^\star\}}$. We proceed by contradiction, suppose

$$\int_{\theta^\star}^{\bar{\theta}} \Big[\alpha_1\mu_1(z)f_1(z) - \lambda_2\bar{F}_2(z)\Big]dz > \int_{\theta_s}^{\bar{\theta}} \Big[\alpha_1\mu_1(z)f_1(z) - \lambda_2\bar{F}_2(z)\Big]dz, \tag{39}$$

if $\theta^\star > \theta_s$ then the previous equation is equivalent to

$$0 > \int_{\theta_s}^{\theta^\star} \Big[\alpha_1\mu_1(z)f_1(z) - \lambda_2\bar{F}_2(z)\Big]dz,$$

or put it in another way

$$-\alpha_2 c^{22}(\theta_s) > \frac{\int_{\theta_s}^{\theta^\star} \alpha_1\mu_1(z)f_1(z)dz}{\int_{\theta_s}^{\theta^\star} \bar{F}_2(z)dz} = \alpha_1 \cdot \frac{\int_{\theta_s}^{\theta^\star} \bar{F}_2(z)c^{12}(z)dz}{\int_{\theta_s}^{\theta^\star} \bar{F}_2(z)dz},$$

which thanks to Lemma 4 implies

$$\max_{0\leq\theta\leq\theta^s} \overline{C}_{12}(\theta,\theta^s) > \min_{\theta^s\leq\theta\leq\bar{\theta}} \overline{C}_{12}(\theta^s,\theta),$$

contradicting condition (NR). If $\theta^\star < \theta_s$ then equation (39) is equivalent to

$$\int_{\theta^\star}^{\theta_s} \Big[\alpha_1\mu_1(z)f_1(z) - \lambda_2\bar{F}_2(z)\Big]dz > 0,$$

which is the same that

$$\alpha_1 \cdot \frac{\int_{\theta^\star}^{\theta^s} \bar{F}_2(z)c^{12}(z)dz}{\int_{\theta^\star}^{\theta^s} \bar{F}_2(z)dz} > -\alpha_2 c^{22}(\theta_s),$$

which thanks to Lemma 4 contradicts condition (NR). This proves equation (38).

Now, recall our choice of lagrange multipliers in (37) and consider the lagrangean evaluated at some arbitrary $(u_1, u_2, x_1, x_2) \in \Omega$. Then, we can verify (35)

$$\begin{aligned}
\mathcal{L}(\boldsymbol{u},\mathbf{x},\boldsymbol{\lambda},\boldsymbol{w}) &= \int_0^{\bar{\theta}} x_1(z)\cdot\Big[\alpha_1\mu_1(z)f_1(z) - \lambda_2\bar{F}_2(z)\Big]dz + \int_0^{\bar{\theta}} x_2(z)\cdot\Big[\alpha_2\mu_2(z)f_2(z) + \lambda_2\bar{F}_2(z)\Big]dz \\
&\leq \max_{x_1\in\Omega}\int_0^{\bar{\theta}} x_1(z)\cdot\Big[\alpha_1\mu_1(z)f_1(z) - \lambda_2\bar{F}_2(z)\Big]dz \\
&+ \max_{x_2\in\Omega}\int_0^{\bar{\theta}} x_2(z)\cdot\Big[\alpha_2\mu_2(z)f_2(z) + \lambda_2\bar{F}_2(z)\Big]dz \\
&\stackrel{(a)}{=} \int_{\theta_s}^{\bar{\theta}} \Big[\alpha_1\mu_1(z)f_1(z) - \lambda_2\bar{F}_2(z)\Big]dz + \int_{\theta^s}^{\bar{\theta}} \Big[\alpha_2\mu_2(z)f_2(z) + \lambda_2\bar{F}_2(z)\Big]dz \\
&= \int_{\theta^s}^{\bar{\theta}} \Big[\alpha_1\mu_1(z)f_1(z) + \alpha_2c^{22}(\theta_s)\bar{F}_2(z)\Big]dz + \int_{\theta^s}^{\bar{\theta}} \Big[\alpha_2\mu_2(z)f_2(z) - \alpha_2c^{22}(\theta_s)\bar{F}_2(z)\Big]dz \\
&= \int_{\theta^s}^{\bar{\theta}} \Big[\alpha_1\mu_1(z)f_1(z) + \alpha_2\mu_2(z)f_2(z)\Big]dz \\
&= \mathcal{L}(0,0,\mathbf{1}_{\{\theta\geq\theta^s\}},\mathbf{1}_{\{\theta\geq\theta^s\}},\boldsymbol{\lambda},\boldsymbol{w}),
\end{aligned}$$

where in (a) we have used (38) and the fact that

$$\alpha_2\mu_2(\theta)f_2(\theta)+\lambda_2\bar{F}_2(\theta)\geq 0 \Leftrightarrow c^{22}(\theta)\geq c^{22}(\theta^s),$$

which, since $c^{22}(\cdot)$ is increasing, holds if and only $\theta \geq \theta^s$. Thus, we have proved that for this choice of lagrange multipliers the static contract maximizes the lagrangen and, therefore, thanks to Theorem 1 in Luenberger (1969, p. 220) it is optimal. ■

Proof of Proposition 3. Take θ_a, θ_b from Lemma 6. In the proof of Theorem 2 we already saw that this allocation is feasible. So we only need to verify that it yields a larger payoff than the static contract, that is, we want

$$\int_{\theta^s}^{\bar{\theta}}[\alpha_1 f_1(z)\mu_1(z)+\alpha_2 f_2(z)\mu_2(z)]dz < \chi\cdot\int_{\theta_a}^{\theta_b}\alpha_1 f_1(z)\mu_1(z)dz+\int_{\theta_b}^{\bar{\theta}}\alpha_1 f_1(z)\mu_1(z)dz + \int_{\theta^s}^{\bar{\theta}}\alpha_2 f_2(z)\mu_2(z)dz,$$

this is equivalent to

$$\frac{\int_{\theta^s}^{\theta_b}\bar{F}_2(z)dz}{\int_{\theta_a}^{\theta_b}\bar{F}_2(z)dz}\cdot\int_{\theta_a}^{\theta_b}\alpha_1 f_1(z)\mu_1(z)dz > \int_{\theta^s}^{\theta_b}\alpha_1 f_1(z)\mu_1(z)dz,$$

which is the same as

$$\frac{\int_{\theta_a}^{\theta^s}\bar{F}_2(z)c^{12}(z)dz}{\int_{\theta_a}^{\theta^s}\bar{F}_2(z)dz} > \frac{\int_{\theta^s}^{\theta_b}\bar{F}_2(z)c^{12}(z)dz}{\int_{\theta^s}^{\theta_b}\bar{F}_2(z)dz}.$$

which is exactly the property satisfied by θ_a, θ_b. ■

Proof of Theorem 3. We relax the constraint associated to type 1 from $(\mathcal{P}^d)$, that is, we relax the following inequality

$$u_1+\int_0^{\bar{\theta}}x_1(\theta)\bar{F}_2(\theta)d\theta \geq u_2+\int_0^{\bar{\theta}}x_2(\theta)\bar{F}_1(\theta)d\theta \tag{40}$$

thus we end up with the same optimization problem that in the proof of Theorem 1: $(\mathcal{P}^d_R)$.

We use Theorem 1 in Luenberger (1969, p. 220) to prove that the proposed solution actually optimizes $(\mathcal{P}^d_R)$. Then, we show that this solution is feasible for the original problem and, therefore, optimal.

For easy of notation set $\theta_1=\underline{\theta}_1$, $\theta_2=\theta_2$ and $\theta_3=\overline{\theta}_1$. With this notation our assumptions are: there exists $\theta_1\leq\theta_2\leq\theta_3$ such that

1. $\overline{C}_{12}(\theta_1,\theta_3)\leq \min_{\theta_3\leq\theta\leq\bar{\theta}}\overline{C}_{12}(\theta_3,\theta)$.

2. $\max_{0\leq\theta\leq\theta_3}\overline{C}_{12}(\theta,\theta_3)\leq\overline{C}_{12}(\theta_1,\theta_3)$

3. $\alpha_1 \cdot \overline{C}_{12}(\theta_1, \theta_3) + \alpha_2 c^{22}(\theta_2) = 0$.

Note that from this condition is immediate that $\hat{\theta}_1 \leq \theta_1$ and $\theta_2 \leq \hat{\theta}_2$. For easy of exposition we assume $\theta_1 < \theta_2 < \theta_3$, but the next argument still goes through if we don't assume this.

Now, the Lagrangean for $(\mathcal{P}_R^d)$ is

$$\begin{aligned}\mathcal{L}(\boldsymbol{u}, \mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{w}) &= u_1(w_1 - \lambda - \alpha_1) + u_2(\lambda - \alpha_2 + w_2) \\ &\quad + \int_0^{\bar{\theta}} x_1(z) \cdot \Big[\alpha_1 \mu_1(z) f_1(z) - \lambda \bar{F}_2(z)(z)\Big] dz + \int_0^{\bar{\theta}} x_2(\theta) \cdot \Big[\alpha_2 \mu_2(z) f_2(z) + \lambda \bar{F}_2(z)\Big] dz,\end{aligned}$$

consider the following multipliers

$$\lambda = \alpha_1 \cdot \frac{\int_{\theta_1}^{\theta_3} \bar{F}_2(z) c^{12}(z) dz}{\int_{\theta_1}^{\theta_3} \bar{F}_2(z) dz}, \quad w_1 = \lambda + \alpha_1, w_2 = -\lambda + \alpha_2,$$

note that λ and w_1 are trivially non-negative, and for w_2 we have

$$w_2 \geq 0 \Leftrightarrow \alpha_2 + \alpha_2 c^{22}(\theta_2) \geq 0 \Leftrightarrow c^{22}(\theta_2) \geq -1 \Leftrightarrow [\theta_2 - h^{22}(\theta_2)] \geq -h^{22}(\theta_2) \Leftrightarrow \theta_2 \geq 0,$$

where in the first if and only if we used condition (3). in our hypothesis. Thus when we optimize the lagrangean we obtain

$$\max_{(\boldsymbol{u}, \mathbf{x}) \in \Omega} \mathcal{L}(\mathbf{x}, \boldsymbol{u}, \boldsymbol{\lambda}, \boldsymbol{w}) = \max_{0 \leq \theta \leq \bar{\theta}} \int_\theta^{\bar{\theta}} \Big[\alpha_1 \mu_1(z) f_1(z) - \lambda \bar{F}_2(z)\Big] dz + \max_{0 \leq \theta \leq \bar{\theta}} \int_\theta^{\bar{\theta}} \Big[\alpha_2 \mu_2(z) f_2(z) + \lambda \bar{F}_2(z)\Big] dz. \quad (41)$$

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's begin with the second term, take any $0 \leq \theta \leq \bar{\theta}$ then

$$\begin{aligned}\int_\theta^{\bar{\theta}} \Big[\alpha_2 \mu_2(z) f_2(z) + \lambda \bar{F}_2(z)\Big] dz &= \int_\theta^{\bar{\theta}} \Big[\alpha_2 \mu_2(z) f_2(z) - \alpha_2 c^{22}(\theta_2) \bar{F}_2(z)\Big] dz \\ &= \int_\theta^{\bar{\theta}} \alpha_2 \bar{F}_2(z) \Big[c^{22}(z) - c^{22}(\theta_2)\Big] dz \\ &\leq \int_{\theta_2}^{\bar{\theta}} \alpha_2 \bar{F}_2(z) \Big[c^{22}(z) - c^{22}(\theta_2)\Big] dz \\ &= \int_0^{\bar{\theta}} x_2^\star(z) \Big[\alpha_2 \mu_2(z) f_2(z) + \lambda \bar{F}_2(z)\Big] dz,\end{aligned}$$

where in the first equality we used condition (3) and the inequality comes from the fact that $c^{22}(\cdot)$ is non-decreasing. Now we look into the first term in equation (41), consider first $\theta \geq \theta_3$

$$\begin{aligned}\int_\theta^{\bar{\theta}} \Big[\alpha_1 \mu_1(z) f_1(z) - \lambda \bar{F}_2(z)\Big] dz &= \int_{\theta_3}^{\bar{\theta}} \Big[\alpha_1 \mu_1(z) f_1(z) - \lambda \bar{F}_2(z)\Big] dz \\ &\quad - \int_{\theta_3}^{\theta} \Big[\alpha_1 \mu_1(z) f_1(z) - \lambda \bar{F}_2(z)\Big] dz \\ &\leq \int_{\theta_3}^{\bar{\theta}} \Big[\alpha_1 \mu_1(z) f_1(z) - \lambda \bar{F}_2(z)\Big] dz,\end{aligned}$$

where we have used the following

$$-\int_{\theta_3}^{\theta}\Big[\alpha_1\mu_1(z)f_1(z)-\lambda\bar{F}_2(z)\Big]dz\leq 0\Leftrightarrow\alpha_1\cdot\frac{\int_{\theta_1}^{\theta_3}\bar{F}_2(z)c^{12}(z)dz}{\int_{\theta_1}^{\theta_3}\bar{F}_2(z)dz}=\lambda\leq\alpha_1\cdot\frac{\int_{\theta_3}^{\theta}\bar{F}_2(z)c^{12}(z)dz}{\int_{\theta_3}^{\theta}\bar{F}_2(z)dz},$$

which thanks to condition (1) in our hypothesis is true. A similar argument holds for $\theta\leq\theta_3$, but using condition (2).

Since $\mathcal{L}(0,\mathbf{x}^\star,\boldsymbol{\lambda},\boldsymbol{w})$ equals

$$\chi\int_{\theta_1}^{\theta_3}\Big[\alpha_1\mu_1(z)f_1(z)-\lambda\bar{F}_2(z)\Big]dz+\int_{\theta_3}^{\bar{\theta}}\Big[\alpha_1\mu_1(z)f_1(z)-\lambda\bar{F}_2(z)\Big]dz+\int_{\theta_2}^{\bar{\theta}}\Big[\alpha_2\mu_2(z)f_2(z)+\lambda\bar{F}_2(z)\Big]dz,$$

which by the definition of λ simplifies to

$$\int_{\theta_3}^{\bar{\theta}}\Big[\alpha_1\mu_1(\theta)f_1(\theta)-\lambda\bar{F}_2(\theta)\Big]d\theta+\int_{\theta_2}^{\bar{\theta}}\Big[\alpha_2\mu_2(\theta)f_2(\theta)+\lambda\bar{F}_2(\theta)\Big]d\theta,$$

we conclude that

$$\max_{(\boldsymbol{u},\mathbf{x})\in\Omega}\mathcal{L}(\boldsymbol{u},\mathbf{x},\boldsymbol{\lambda},\boldsymbol{w})\leq\mathcal{L}(0,\mathbf{x}^\star,\boldsymbol{\lambda},\boldsymbol{w}),$$

as required.

In order to conclude the proof we need to verify the proposed solution is indeed feasible for the original problem. That is, we need to verify it satisfies equation (40) (note the other (IC) constraint is satisfied with equality), which is equivalent to

$$\chi\cdot\int_{\theta_1}^{\theta_3}\bar{F}_1(z)dz+\int_{\theta_3}^{\bar{\theta}}\bar{F}_1(z)dz\geq\int_{\theta_2}^{\bar{\theta}}\bar{F}_1(z)dz,$$

or in a more compact form

$$\frac{\int_{\theta_2}^{\theta_3}\bar{F}_2(z)dx}{\int_{\theta_1}^{\theta_3}\bar{F}_2(z)dz}\geq\frac{\int_{\theta_2}^{\theta_3}\bar{F}_1(z)dz}{\int_{\theta_1}^{\theta_3}\bar{F}_1(z)dz}. \tag{42}$$

In order to verify equation (42) consider condition (2) with $\theta=\theta_2$, that gives us

$$\frac{\int_{\theta_2}^{\theta_3}\bar{F}_2(z)dz}{\int_{\theta_1}^{\theta_3}\bar{F}_2(z)dz}\geq\frac{\int_{\theta_2}^{\theta_3}\bar{F}_2(z)c^{12}(z)dz}{\int_{\theta_1}^{\theta_3}\bar{F}_2(z)c^{12}(z)dz},$$

but thanks to (DHR) is not hard to check that

$$\frac{\int_{\theta_2}^{\theta_3}\bar{F}_2(z)c^{12}(z)dz}{\int_{\theta_1}^{\theta_3}\bar{F}_2(z)c^{12}(z)dz}\geq\frac{\int_{\theta_2}^{\theta_3}\bar{F}_1(z)dz}{\int_{\theta_1}^{\theta_3}\bar{F}_1(z)dz},$$

and, therefore, under our assumption equation (42) is verified. This concludes the proof. ■

B Proofs for section 4

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

$$\max_{\theta\leq\theta^s}\left\{\frac{\theta^s e^{-\lambda_1\theta^s}-\theta e^{-\lambda_1\theta}}{e^{-\lambda_2\theta^s}-e^{-\lambda_2\theta}}\right\}\leq\min_{\theta^s\leq\theta}\left\{\frac{\theta e^{-\lambda_1\theta}-\theta^s e^{-\lambda_1\theta^s}}{e^{-\lambda_2\theta}-e^{-\lambda_2\theta^s}}\right\}. \tag{43}$$

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

$$g_L(\theta)\triangleq\frac{\theta^s e^{-\lambda_1\theta^s}-\theta e^{-\lambda_1\theta}}{e^{-\lambda_2\theta^s}-e^{-\lambda_2\theta}}\quad\text{and}\quad g_U(\theta)\triangleq\frac{\theta e^{-\lambda_1\theta}-\theta^s e^{-\lambda_1\theta^s}}{e^{-\lambda_2\theta}-e^{-\lambda_2\theta^s}}.$$

Note the following

$$\lim_{\theta\to\theta^s+}g_U(\theta)=\lim_{\theta\to\theta^s-}g_L(\theta)=\frac{(\lambda_1\theta^s-1)}{\lambda_2}\cdot e^{-\theta^s(\lambda_1-\lambda_2)}, \tag{44}$$

and

$$\lim_{\theta\to\infty}g_U(\theta)=\theta^s\cdot e^{-\theta^s(\lambda_1-\lambda_2)}. \tag{45}$$

Finally note that

$$\frac{(\lambda_1\theta^s-1)}{\lambda_2}\cdot e^{-\theta^s(\lambda_1-\lambda_2)}\leq\theta^s\cdot e^{-\theta^s(\lambda_1-\lambda_2)}\iff\theta^s\leq\frac{1}{\lambda_1-\lambda_2}. \tag{46}$$

Now, suppose condition (NR) holds and

$$\theta^s>\frac{1}{\lambda_1-\lambda_2} \tag{47}$$

From equations (44),(45) and (46) we see that

$$g_U(\theta^s)=g_L(\theta^s)>\lim_{\theta\to\infty}g_L(\theta),$$

which implies

$$\max_{\theta\leq\theta^s}\left\{\frac{\theta^s e^{-\lambda_1\theta^s}-\theta e^{-\lambda_1\theta}}{e^{-\lambda_2\theta^s}-e^{-\lambda_2\theta}}\right\}>\min_{\theta^s\leq\theta}\left\{\frac{\theta e^{-\lambda_1\theta}-\theta^s e^{-\lambda_1\theta^s}}{e^{-\lambda_2\theta}-e^{-\lambda_2\theta^s}}\right\} \tag{48}$$

contradicting the fact that condition (NR) holds.

For the other direction, assume equation (3) holds. We first prove that for $\theta\leq\theta^s$ we have $g_L(\theta)\leq g_L(\theta^s)$, indeed

$$\begin{aligned}g_L(\theta)\leq g_L(\theta^s)&\iff\frac{\theta^s e^{-\lambda_1\theta^s}-\theta e^{-\lambda_1\theta}}{e^{-\lambda_2\theta^s}-e^{-\lambda_2\theta}}\leq\frac{(\lambda_1\theta^s-1)}{\lambda_2}\cdot e^{-\theta^s(\lambda_1-\lambda_2)}\\&\iff\lambda_2\cdot(\theta^s e^{-\lambda_1\theta^s}-\theta e^{-\lambda_1\theta})\geq(e^{-\lambda_2\theta^s}-e^{-\lambda_2\theta})\cdot(\lambda_1\theta^s-1)\cdot e^{-\theta^s(\lambda_1-\lambda_2)}\\&\iff\lambda_2\theta^s\cdot(1-\frac{\theta}{\theta^s}e^{-\lambda_1(\theta-\theta^s)})-(1-e^{-\lambda_2(\theta-\theta^s)})\cdot(\lambda_1\theta^s-1)\geq 0,\end{aligned}$$

so we just need to see that this las inequality holds for $\theta\leq\theta^s$. For doing so define

$$H(\theta)\triangleq\lambda_2\theta^s\cdot(1-\frac{\theta}{\theta^s}e^{-\lambda_1(\theta-\theta^s)})-(1-e^{-\lambda_2(\theta-\theta^s)})\cdot(\lambda_1\theta^s-1),$$

and note that $H(\theta^s) = 0$ and

$$H(0) = \lambda_2\theta^s + (e^{\lambda_2\theta^s} - 1)\cdot(\lambda_1\theta^s - 1) \geq \lambda_2\theta^s + \lambda_2\theta^s(\lambda_1\theta^s - 1) = \lambda_2\theta^s \cdot \lambda_1\theta^s > 0,$$

where the inequality comes from convexity of the exponential function and the fact that $\theta^s \geq 1/\lambda_1$. Furthermore the derivative of H is given by

$$\frac{dH}{d\theta} = \lambda_2(\lambda_1\theta - 1)e^{-\lambda_1(\theta-\theta^s)} - \lambda_2(\lambda_1\theta^s - 1)e^{-\lambda_2(\theta-\theta^s)},$$

and it can be easily verified that for $\theta \leq \theta^s$ we have $dH/d\theta \leq 0$. This together to the facts that $H(0) > 0$ and $H(\theta^s) = 0$ imply that $g_L(\theta) \leq g_L(\theta^s)$ for all $\theta \leq \theta^s$. Which in turn implies

$$\max_{\theta\leq\theta^s}\left\{\frac{\theta^s e^{-\lambda_1\theta^s} - \theta e^{-\lambda_1\theta}}{e^{-\lambda_2\theta^s} - e^{-\lambda_2\theta}}\right\} = \frac{(\lambda_1\theta^s - 1)}{\lambda_2}\cdot e^{-\theta^s(\lambda_1-\lambda_2)}.$$

Now we prove that for $\theta \geq \theta^s$ we have $g_U(\theta) \geq g_U(\theta^s)$. Note that if we prove this we are done because this and what we have just proven imply condition (NR). As before we do

$$\begin{aligned} g_U(\theta) \geq g_U(\theta^s) &\iff \frac{\theta e^{-\lambda_1\theta} - \theta^s e^{-\lambda_1\theta^s}}{e^{-\lambda_2\theta} - e^{-\lambda_2\theta^s}} \geq \frac{(\lambda_1\theta^s - 1)}{\lambda_2}\cdot e^{-\theta^s(\lambda_1-\lambda_2)} \\ &\iff \lambda_2(\theta^s e^{-\lambda_1\theta^s} - \theta e^{-\lambda_1\theta}) \geq (\lambda_1\theta^s - 1)\cdot(e^{-\lambda_2\theta^s} - e^{-\lambda_2\theta})\cdot e^{-\theta^s(\lambda_1-\lambda_2)} \\ &\iff \lambda_2(\theta^s - \theta e^{-\lambda_1(\theta-\theta^s)}) - (\lambda_1\theta^s - 1)\cdot(1 - e^{-\lambda_2(\theta-\theta^s)}) \geq 0, \end{aligned}$$

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

$$\lim_{\theta\to\infty} H(\theta) = \lambda_2\theta^s - (\lambda_1\theta^s - 1) \geq 0,$$

hence for $\theta \geq \theta^s$ we have $H(\theta) \geq 0$ and, therefore, $g_U(\theta) \geq g_U(\theta^s)$ for all $\theta \geq \theta^s$, as desired. ■

Proof of Corollary 1. Recall that for any $\lambda_L > \lambda_H$ from Lemma 2 we have

$$\frac{1}{\lambda_L} \leq \theta^s(\alpha_L) \leq \frac{1}{\lambda_H},$$

and

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

therefore, for any $\alpha_L \in [0, 1]$ equation (3) is satisfied. Then by Proposition 1 we conclude that the static contract is optimal for any $\alpha_L \in [0, 1]$. ■

Proof of Corollary 2. First we show $\theta^s(\cdot)$ 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 $\theta^s(\alpha_L^n)$ does not converge to $\theta^s(0) = 1/\lambda_H$. That is,

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

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

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

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

$$\forall n, \quad \frac{1}{\lambda_H} - \epsilon > \theta^s(\alpha_L^{\ell_n}). \tag{49}$$

But since $\theta^s(\alpha_L^{\ell_n})$ is a maximizer of $R^s(\cdot)$ we must have

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

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

$$\theta^s(\alpha_L^{\ell_n})e^{-\lambda_H\theta^s(\alpha_L^{\ell_n})} \geq \alpha_L^{\ell_n}\frac{1}{\lambda_H}e^{-\lambda_L\frac{1}{\lambda_H}} + (1-\alpha_L^{\ell_n})\frac{1}{\lambda_H}e^{-\lambda_H\frac{1}{\lambda_H}}. \tag{50}$$

Note that the function $\theta e^{-\lambda_H\theta}$ has a unique maximum at $\theta = 1/\lambda_H$ and since $\theta^s(\alpha_L^{\ell_n})$ satisfies equation (49), we can always find $\delta(\epsilon) > 0$ such that

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

plugging this in equation (50) yields

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

so taking the limit over n gives a contradiction. In conclusion we have proved that $\theta^s(\cdot)$ is continuous from the right at zero. Now, to finalize the proof recall that we are assuming $\lambda_L > 2\lambda_H$ or in other words

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

but since $\theta^s(0) = 1/\lambda_H$ and $\theta^s(\cdot)$ is continuous from the right we can always find $\bar{\alpha}_1 \in (0,1]$ such that

$$\frac{1}{\lambda_H} \geq \theta^s(\bar{\alpha}_1) \geq \frac{1}{\lambda_L - \lambda_H},$$

so thanks to Proposition 1, the dynamic contract is optimal when we set $\alpha_L > \bar{\alpha}_1$. Note that the same arguments is valid for $1/\lambda_L$. That is, we can show that $\theta^s(\alpha_L)$ is continuos from the left at $1/\lambda_L$ and then using the fact that

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

we can find $\bar{\alpha}_2 \in [\bar{\alpha}_1, 1)$ such that

$$\frac{1}{\lambda_L - \lambda_H} > \theta^s(\bar{\alpha}_2) \geq \frac{1}{\lambda_L},$$

hence in $[\bar{\alpha}_2, 1]$ the static contract is optimal. All of this implies that since $\theta^s(\cdot)$ is a non-increasing 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,

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

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

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

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

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

However, since $\theta^s(\lambda_L^n)$ maximizes $R^s(\cdot)$ it must be the case that $R^s(1/\lambda_H) \leq R^s(\theta^s(\lambda_L^n))$, that is,

$$\alpha_L \frac{1}{\lambda_H} e^{-\lambda_L^n \frac{1}{\lambda_H}} + \alpha_H \frac{1}{\lambda_H} e^{-\lambda_H \frac{1}{\lambda_H}} \leq R^s(\theta^s(\lambda_L^n)).$$

Taking limit over n at 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 2. We need to find $\theta_1 = \theta_L$, $\theta_2 = \theta_H$ and θ_3 such that $\theta_1 \leq \theta_2 \leq \theta_3$ and the following conditions are satisfied

1. $\overline{C}_{12}(\theta_1, \theta_3) \leq \min_{\theta_3 \leq \theta \leq \bar{\theta}} \overline{C}_{12}(\theta_3, \theta)$.

2. $\max_{0 \leq \theta \leq \theta_3} \overline{C}_{12}(\theta, \theta_3) \leq \overline{C}_{12}(\theta_1, \theta_3)$

3. $\alpha_1 \cdot \overline{C}_{12}(\theta_1, \theta_3) + \alpha_2 c^{22}(\theta_2) = 0$.

And then we can apply Theorem 3.

First note that since the support of the exponential distribution is unbounded from above, we can take $\theta_3 = \infty$ which eliminates condition (1). Conditions (2) and (3) can be cast as

$$\theta_1 e^{-\theta_1(\lambda_1-\lambda_2)} \geq \theta e^{-\theta(\lambda_1-\lambda_2)} \quad \forall \theta \geq 0 \quad \text{and} \quad \alpha_1 \cdot \lambda_2 \theta_1 e^{-\theta_1(\lambda_1-\lambda_2)} = -\alpha_2 \cdot (\lambda_2\theta_2 - 1), \tag{51}$$

By optimizing the first term in (51) we obtain

$$\theta_1 = \frac{1}{\lambda_1 - \lambda_2},$$

and then solving for θ_2 yields

$$\theta_2 = \frac{1}{\lambda_2} - \frac{\alpha_1}{\alpha_2}\frac{e^{-1}}{\lambda_1 - \lambda_2}.$$

What we need to check (and it is not obvious at a first glance) is that $\theta_1 \leq \theta_2$. First, we show

$$\alpha_1(\theta_1 - \frac{1}{\lambda_1})\lambda_1 e^{-\lambda_1\theta_1} + \alpha_2(\theta_1 - \frac{1}{\lambda_2})\lambda_2 e^{-\lambda_2\theta_1} < 0. \tag{52}$$

To prove this inequality notice that since θ^s is the optimal static cutoff we have

$$\alpha_1\theta^s e^{-\lambda_1\theta^s} + \alpha_2\theta^s e^{-\lambda_2\theta^s} \geq \alpha_1\theta_1 e^{-\lambda_1\theta_1} + \alpha_2\theta_1 e^{-\lambda_2\theta_1}, \tag{53}$$

then

$$\begin{aligned}
\alpha_1(\theta_1 - \frac{1}{\lambda_1})\lambda_1 e^{-\lambda_1\theta_1} + \alpha_2(\theta_1 - \frac{1}{\lambda_2})\lambda_2 e^{-\lambda_2\theta_1} &= \alpha_1\theta_1(\lambda_1-\lambda_2)e^{-\lambda_1\theta_1} + \alpha_1\theta_1\lambda_2 e^{-\lambda_1\theta_1} + \alpha_2\theta_1\lambda_2 e^{-\lambda_2\theta_1} \\
&- \alpha_1 e^{-\lambda_1\theta_1} - \alpha_2 e^{-\lambda_2\theta_1} \\
&= \alpha_1 e^{-\lambda_1\theta_1} + \lambda_2(\alpha_1\theta_1 e^{-\lambda_1\theta_1} + \alpha_2\theta_1 e^{-\lambda_2\theta_1}) - \alpha_1 e^{-\lambda_1\theta_1} \\
&- \alpha_2 e^{-\lambda_2\theta_1} \\
&\overset{(a)}{\leq} \lambda_2(\alpha_1\theta^s e^{-\lambda_1\theta^s} + \alpha_2\theta^s e^{-\lambda_2\theta^s}) - \alpha_2 e^{-\lambda_2\theta_1} \\
&\overset{(b)}{<} \lambda_2(\alpha_1\theta^s e^{-\lambda_1\theta^s} + \alpha_2\theta^s e^{-\lambda_2\theta^s}) - \alpha_2 e^{-\lambda_2\theta^s} \\
&= \lambda_2\alpha_1\theta^s e^{-\lambda_1\theta^s} + \lambda_2\alpha_2 e^{-\lambda_2\theta^s}(\theta^s - \frac{1}{\lambda_2}) \\
&\overset{(c)}{=} \lambda_2\alpha_1\theta^s e^{-\lambda_1\theta^s} - \lambda_1\alpha_1 e^{-\lambda_1\theta^s}(\theta^s - \frac{1}{\lambda_1}) \\
&= \alpha_1 e^{-\lambda_1\theta^s}\Big(-\theta^s(\lambda_1-\lambda_2) + 1\Big) \\
&\overset{(d)}{<} 0,
\end{aligned}$$

where (a) comes from equation (53), (b) is true because the function $-e^{-\lambda_2\theta}$ increasing and $\theta_1 < \theta^s$, (c) comes from equation (2). And (d) comes from $\theta_1 < \theta^s$. With this we have proven (52) and thus

$$\begin{aligned}
\lambda_1\alpha_2 \cdot (\theta_2 - \frac{1}{\lambda_2}) &\overset{(a)}{=} -\lambda_1\alpha_1 \cdot \theta_1 e^{-\theta_1(\lambda_1-\lambda_2)} \\
&= -\lambda_1\alpha_1 \cdot \Big(\theta_1 - \frac{1}{\lambda_1}\Big)e^{-\theta_1(\lambda_1-\lambda_2)} - \lambda_1\alpha_1 \cdot \frac{1}{\lambda_1}e^{-\theta_1(\lambda_1-\lambda_2)} \\
&\overset{(b)}{>} \alpha_2(\theta_1 - \frac{1}{\lambda_2})\lambda_2 - \alpha_1 \cdot e^{-\theta_1(\lambda_1-\lambda_2)} \\
&\overset{(c)}{=} \alpha_2(\theta_1 - \frac{1}{\lambda_2})\lambda_2 + \frac{\alpha_2}{\theta_1} \cdot (\theta_2 - \frac{1}{\lambda_2}),
\end{aligned}$$

where in (a) and (c) we used the definition of θ_2, and in (b) we used equation (52). From this we have that

$$(\theta_2 - \frac{1}{\lambda_2}) \cdot \Big(\lambda_1\alpha_2 - \frac{\alpha_2}{\theta_1}\Big) > \alpha_2(\theta_1 - \frac{1}{\lambda_2})\lambda_2,$$

but replacing θ_1 with $1/(\lambda_1 - \lambda_2)$ in this last expression we get $\theta_2 > \theta_1$.

Finally, following the result in Theorem 3 for χ we have

$$\chi = \frac{\int_{\theta_2}^{\theta_3} f_2(x)h^{22}(x)dx}{\int_{\theta_1}^{\theta_3} f_2(x)h^{22}(x)dx} = \frac{e^{-\lambda_2\theta_2}}{e^{-\lambda_2\theta_1}} = \exp\Big(- \lambda_2\Big[\frac{1}{\lambda_2} - \frac{\alpha_1}{\alpha_2}\frac{e^{-1}}{\lambda_1 - \lambda_2} - \frac{1}{\lambda_1 - \lambda_2}\Big]\Big).$$

■

References

Ashlagi, I., C. Daskalakis, and N. Haghpanah (2016): "Sequential Mechanisms with Ex-post Participation Guarantees," .

Balseiro, S., V. Mirrokni, and R. Paes Leme (2016): "Dynamic Mechanisms with Martingale Utilities," .

Celis, L. E., G. Lewis, M. Mobius, and H. Nazerzadeh (2014): "Buy-it-Now or Take-a-Chance: Price Discrimination through Randomized Auctions," *Management Science*, 60(12), 2927–2948.

Courty, P., and H. Li (2000): "Sequential Screening," *The Review of Economic Studies*, 67(4), 697–717.

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

Fuchs, W., and A. Skrzypacz (2015): "Government Interventions in a Dynamic Market with Adverse Selection," *Journal of Economic Theory*, 158, 371–406.

Heumann, T. (2016): "Optimal Sequential Screening with Ex-Post Participation Constraint," .

Krähmer, D., and R. Strausz (2014): "Optimal Sales Contracts with Withdrawal Rights," mimeo.

——— (2015): "Optimal Sales Contracts with Withdrawal Rights," *The Review of Economic Studies*, 82(2), 762–790.

——— (2016): "Optimality of sequential screening with multiple units and ex post participation constraints," *Economics Letters*, 142, 64–68.

Luenberger, D. G. (1969): *Optimization by Vector Space Methods*. John Wiley & Sons.

Manelli, A. M., and D. R. Vincent (2007): "Multidimensional Mechanism Design: Revenue Maximization and The Multiple-Good Monopoly," *Journal of Economic theory*, 137(1), 153–185.

Myerson, R. B. (1979): "Incentive Compatibility and The Bargaining Problem," *Econometrica*, pp. 61–73.

——— (1981): "Optimal Auction Design," *Mathematics of operations research*, 6(1), 58–73.

Riley, J., and R. Zeckhauser (1983): "Optimal Selling Strategies: When to Haggle, When to Hold Firm," *The Quarterly Journal of Economics*, 98(2), 267–289.

Samuelson, W. (1984): "Bargaining Under Asymmetric Information," *Econometrica*, pp. 995–1005.