PIER Working Paper 16-007 v.2

"Efficient Mechanisms with Information Acquisition "

BY

Yunan Li

http://ssrn.com/abstract=2757692

Efficient Mechanisms with Information Acquisition*

Yunan Li†

University of Pennsylvania

June 23, 2017

Abstract

This paper studies the design of ex ante efficient mechanisms in situations where a single object is for sale, and agents have positively interdependent values and can covertly acquire information at some cost before participating in a mechanism. We find that when the strength of interdependence is low or the number of agents is large, the ex post efficient mechanism is also ex ante efficient. In cases of high interdependence or a small number of agents, ex ante efficient mechanisms discourage agents from acquiring excessive information by introducing randomization to the ex post efficient allocation rule in areas where the information's accuracy increases most rapidly if an addition piece of information is acquired. In special cases, there exists an ex ante efficient mechanism that has a simple and appealing implementation: standard auctions with discrete bids.

Keywords: Auctions, Mechanism Design, Information Acquisition, Efficiency

JEL Classification Numbers: C70, D44, D82, D86

*This paper is based on a chapter of my dissertation at University of Pennsylvania. I would like to express my gratitude to my advisor Rakesh Vohra for his encouraging support since the very beginning of the project. For their valuable discussions, I am grateful to (in no particular order) Steven Matthews, Mallesh Pai, Aislinn Bohren, Hanming Fang, Jason Hartline, Sangmok Lee, George Mailath, Andrew Postlewaite and the participants in seminars at University of Pennsylvania, 3rd summer school of the Econometric Society and 11th World Congress of the Econometric Society. All remaining errors are my responsibility.

†Author's address: Yunan Li, Economics Department, University of Pennsylvania, Philadelphia, PA, USA, *yunanli@sas.upenn.edu*.

1 Introduction

Most of the mechanism design literature assumes that the amount of information possessed by agents is exogenous. In many important applications, however, this assumption does not apply. For example, in auctions for offshore oil and gas leases in the U.S., companies use seismic surveys to collect information about the tracts offered for sale before participating in the auctions. Another example is the sale of financial or business assets, in which buyers perform due diligence to investigate the quality and compatibility of the assets before submitting offers. In these settings, the information held by agents is endogenous. Moreover, it is costly to acquire information. In the example of U.S. auctions for offshore oil and gas leases (see Haile et al. (2010)), companies can choose to conduct two-dimensional (2-D) or three-dimensional (3-D) seismic surveys. 3-D surveys can produce more accurate information, and thus were used in 80% of wells drilled in the Gulf of Mexico by 1996. However, this number was only 5% in 1989 when 3-D surveys were more expensive than 2-D surveys.[1] Similarly, the legal and accounting costs of performing due diligence often amount to millions of dollars in the sale of a business asset (see Quint and Hendricks (2013) and Bergemann et al. (2009)).

Earlier studies have analyzed agents' private incentives to acquire information and compared that with the social incentives. Maskin (1992) and Bergemann and Välimäki (2002), among others, focus on the ex post efficient mechanism that implements the ex post efficient allocations given acquired private information. They find that, if valuations are private, agents' incentives to acquire information coincide with the social incentives and the ex ante efficient information acquisition is achieved. However, if valuations are interdependent, the ex post efficient mechanism will result in socially sub-optimal information acquisition. The case of interdependent values is pertinent to many economic applications. For example, in the U.S. auctions for offshore oil and gas leases, if one company find that the other have a lower estimate on the amount of extractable oil, they may revise their valuation downward. In a follow-up paper, Bergemann et al. (2009) study the equilibrium level of information acquisition when agents face binary information decisions and their values are positively interdependent. They find that the ex post efficient mechanism leads to excessive information acquisition in equilibrium. In summary, when valuations are interdependent, there is a conflict between the provision of ex ante efficient incentives to acquire information and the ex post efficient use of information. The question regarding how to design an ex ante efficient mechanism to balance the two trade-offs remains open.

[1] For instance, it costs \$1 million to examine a 50 square mile 3-D seismic survey in 1990, while this number was less than \$100, 000 in 2000.

In this paper, we study the design of ex ante efficient mechanisms in the sale of a single object when agents' values are positively interdependent. The true value of the object to each agent is initially unknown. Before participating in a mechanism, agents can simultaneously and independently decide how much information to acquire. An agent acquires information by increasing the accuracy of the signal he receives. Both the accuracy and the realization of his signal are an agent's private information, and the signals are independent across agents. An agent must incur a higher cost to receive a more accurate signal. We assume that the accuracy of the signals is *supermodular ordered*. This notion of information order was first introduced into the literature by Ganuza and Penalva (2010) and later used by Shi (2012) when studying revenue-maximizing mechanisms in the independent private value setting with endogenous information.

In most parts of the paper, we focus on symmetric mechanisms and symmetric equilibria in which all agents acquire the same amount of information before participating in a mechanism. There are three main results.

Consistent with Bergemann and Välimäki (2002) and Bergemann et al. (2009), it is never socially optimal to encourage agents acquire more information than they would when they face the ex post efficient mechanism. To discourage agents from acquiring excessive information, the social planner can either withhold the object with some probability, or introduce pooling or randomization into the allocation rule. The first main result of the paper is to show that the object is never withheld in an ex ante efficient mechanism. Intuitively, whenever the social planner withholds the object, she can also allocate it randomly. By doing so, the allocative efficiency increases while an agent's ex ante incentive to acquire information remains unaffected. In fact, this property holds more generally even if we consider asymmetric mechanisms and asymmetric equilibria in which agents acquire different amounts of information. Though intuitive, the proof of this result is non-trivial because of the presence of the non-standard information acquisition constraint. This result is also important technically as it facilitates the analysis by allowing us to work with the interim allocation rule directly.

Second, for any given information choice, we fully characterize all mechanisms that implement this choice and maximize the expected social surplus. Remember that an ex ante efficient mechanism discourages agents from excessive information acquisition by introducing pooling or randomization into the allocation rule. Specifically, the optimal pooling areas are those in which the accuracy of an agent's posterior mean increases most rapidly if an additional piece of information is acquired. If this marginal effect of an additional piece of information is the same for all possible posterior means, we say that the information structures are *uniformly supermodular ordered*. In this special case, there exists an ex ante efficient mechanism that has a simple and appealing implementation: standard auctions with discrete bids.

In the model's environment, standard auctions such as first-price or second-price auctions are ex post efficient. By restricting bids to discrete levels, we bunch nearby posterior means together, which clearly reduces agents' marginal benefits from acquiring information. Restricting bids to discrete levels is not uncommon in auctions in practice. For example, eBay auctions require that the next bid must exceed the current price plus a bid increment and FCC spectrum auctions adopt a minimum clock price increment. Most existing auction theories predict that discrete bids lead to inefficiency and they are mainly used in practice to simplify communication processes and speed auctions (see Ausubel and Cramton (2004)). Our results suggest an alternative justification for the prevalence of discrete bids. That is, when agents have positively interdependent values and can covertly acquire information at some cost, the use of discrete bids can improve ex ante efficiency.

Despite the simple implementation in special cases, ex ante efficient mechanisms generally consist of complex pooling areas. Therefore, it is worthwhile to understand when the relatively simple ex post efficient mechanism is also ex ante efficient. As the third main result of the paper, we show that this is likely to be the case when the level of interdependence is low or the number of agents is large. Intuitively, when the level of interdependence is low, the discrepancy between individual and social incentives to acquire information is small and therefore the ex post efficient mechanism is likely to be ex ante efficient. When the number of agents is large, an individual's marginal benefit from acquiring information is already small because of the fierce competition. Hence, the social planner need not further discourage them from acquiring information by distorting the ex post efficient allocation rule.

Finally, we briefly discuss general ex ante efficient mechanisms without restricting attention to symmetric mechanisms or symmetric equilibria. First, as mentioned above, a robust property of ex ante efficient mechanisms is that the object is never withhold. To obtain some further results, we restrict attention to the special case in which the information structures are uniformly supermodular ordered. In this special case, we provide conditions under which the socially optimal information choices are the same for all agents and a symmetric ex ante efficient mechanism exists. These conditions are likely to be satisfied when the level of interdependence is high or the number of agents is small. When these conditions are not satisfied, we provide an example in which an asymmetric mechanism generates higher expected social surplus than the optimal symmetric mechanism does. The intuition behind this result is as follows. As in Bergemann et al. (2009), the agents' information acquisition decisions are strategic substitutes. Therefore, it could be socially optimal to discourage other agents from acquiring information by encouraging one agent to do so.

Technically, the problem considered in this paper is challenging for two reasons. First, we want to work directly with interim rather than ex post allocation rules, which has been proved to be a very useful method in the literature.[2] However, when valuations are interdependent, it is hard to write the social planner's objective function in terms of interim allocation rules. We overcome this difficulty by proving that the object is never withhold in an ex ante efficient mechanism. The second challenge arise because of the non-standard information acquisition constraint. To overcome this difficulty, we use an approach first proposed by Reid (1968) and later introduced into the mechanism design literature by Mierendorff (2009). The proof, however, is not a straightforward modification of Mierendorff (2009). In Mierendorff (2009), the interim allocation rule is discontinuous at one known point. In this paper, the interim allocation rule could be discontinuous at most countably many times, at unknown points.

1.1 Related literature

This paper is related to the literature studying agents' incentives to acquire information in some commonly used mechanisms. Earlier papers focus on the comparison between first-price and second-price auctions. For example, Matthews (1984a) considers a first-price auction with pure common values, and examines how an increase in the number of agents affects information acquisition. Stegeman (1996) finds that both auctions lead to identical and, more importantly, efficient incentives for information acquisition when agents' values are private and independent. In contrast, Persico (2000) finds that a first-price auction provides stronger incentives for agents to acquire information than a second-price auction does when their values are affiliated. The two most closely related papers are Bergemann and Välimäki (2002) and Bergemann et al. (2009). Both study the efficiency of information acquisition by agents when ex post efficient mechanisms is used. Instead of focusing on a particular mechanism, this paper studies the design of ex ante efficient mechanisms.

This paper is also related to papers that study the revenue-maximizing mechanisms with endogenous information acquisition. The two most closely related papers are Shi (2012) and Crémer et al. (2009). Shi (2012) considers the sale of a single asset when buyers have independent private values and who, before the auction, can simultaneously and independently decide how much information to acquire. He finds that the optimal reserve price is always below the standard monopoly price to encourage information acquisition. The focus of this paper is on efficiency in environments where valuations are interdependent. In Crémer et al. (2009), the revenue-maximizing mechanism also achieves ex ante efficiency. However, in Crémer et al. (2009), agents face binary information decisions and the seller can control

[2] See, for example, Maskin and Riley (1984), Matthews (1983), Mierendorff (2011), Pai and Vohra (2014) and Li (2017).

the access to information. The latter assumption assumes away the problem considered in this paper, which is how to design a mechanism to discourage agents from acquiring excessive information. Several other papers model the information cost as an entry cost (see, for example, Levin and Smith (1994), Ye (2004) and Lu and Ye (2014)). As a result, agents' information decisions are observable. But in this paper agents' information choices are also his private information.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 contains the main results. Specifically, Section 3.2 characterizes optimal symmetric mechanisms for each given information choice and Section 3.3 studies the socially optimal information choice. Section 4 examines ex ante efficient mechanisms without imposing symmetry restrictions. Section 5 concludes. All omitted proofs are relegated to appendix.

2 Model

There are n agents, indexed by $i \in \{1, \cdots, n\}$, who compete for a single object. Each agent i has a payoff-relevant type θ_i, which is unknown to the agent or to the social planner ex ante. Agent i's valuation depend not only on his own type but on others' as well. Specifically, the true value of the object to agent i is

$$v_i(\boldsymbol{\theta}) := \theta_i + \gamma \sum_{j \neq i} \theta_j,$$

where $\boldsymbol{\theta} := (\theta_1, \ldots, \theta_n)$ is the type profile and $\gamma \geq 0$ is a measure of interdependence. We assume that $\gamma \leq 1$, which asserts that a marginal change in agent i's type affects his valuation at least as much as it does that of any other agent. This single crossing condition is necessary and efficient for the implementablility of the ex post efficient allocation in the current setting. If $\gamma = 0$, the model is one of private values; and if $\gamma = 1$, the model is of pure common values. Each agent has a quasi-linear utility. Specifically, agent i's payoff is $q_i v_i(\boldsymbol{\theta}) - t_i$ if he receives the object with probability $q_i \in [0, 1]$ and pays $t_i \in \mathbb{R}$.

Initially, agents know only that $\{\theta_i\}$ are independently distributed with a common cumulative distribution F with support $[\underline{\theta}, \overline{\theta}] \subset \mathbb{R}_+$. The distribution F has a positive and continuous density function f and a mean valuation

$$\mu := \int_{\underline{\theta}}^{\overline{\theta}} \theta f(\theta) \mathrm{d}\theta.$$

Each agent i can *covertly* acquire a signal $x_i \in \mathbb{R}$ about his type θ_i by selecting a joint distribution of (x_i, θ_i) from a family of joint distributions $\{G(x_i, \theta_i | \alpha_i)\}$, indexed by their *accuracy* $\alpha_i \in [\underline{\alpha}, \overline{\alpha}] \subset \mathbb{R}$. For each $\alpha \in [\underline{\alpha}, \overline{\alpha}]$, we also refer the joint distribution $G(\cdot, \cdot|\alpha)$ as an *information structure*. Let g denote the density function associated with G. For all $\alpha \in [\underline{\alpha}, \overline{\alpha}]$, $G(\cdot, \cdot|\alpha)$ admits the same marginal distribution of θ as the prior (i.e. $\int_{\mathbb{R}} g(x, \theta|\alpha)\mathrm{d}x = f(\theta)$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$), and the posterior mean $\mathbb{E}[\theta|x, \alpha]$ is strictly increasing in x, which asserts that a higher signal leads to a higher conditional expectation.[3] A signal with a higher α is more accurate (in a sense defined below). Let $C(\alpha)$ denote the cost of acquiring a signal with *accuracy* α. As is standard in the literature, we assume that the cost function C is non-negative, strictly increasing, twice continuously differentiable and strictly convex, and satisfies $C(\underline{\alpha}) = C'(\underline{\alpha}) = 0$ and $C''(\underline{\alpha}) > 0$.

2.1 Information order

Let $G(x|\alpha)$ denote the marginal distribution of signal x for a given accuracy α. We can define a new signal by applying the probability integral transformation on the original signal. Specifically, let $s := G(x|\alpha)$. The transformed signal s is uniformly distributed on $[0, 1]$.[4] Clearly, the transformed signal has the same informational content as the original one. Furthermore, because any two transformed signals have the same marginal distribution, their realizations are directly comparable. Therefore, we will henceforth work with the transformed signal directly. Let $w(s, \alpha) := \mathbb{E}[\theta|s, \alpha]$ denote the conditional expectation or posterior mean of θ given signal s and accuracy α. By assumption, $w(s, \alpha)$ is strictly increasing in s. For each $\alpha \in [\underline{\alpha}, \overline{\alpha}]$, let $H(w|\alpha) := \mathbb{P}(w(s, \alpha) \le w)$ denote the cumulative distribution function of $w(s, \alpha)$ and $h(w|\alpha)$ its corresponding density function. We assume that both $H(w|\alpha)$ and $h(w|\alpha)$ are twice continuously differentiable in w and α. Throughout the paper, we assume that the information structures are *supermodular ordered*:

Definition 1 *The information structures are supermodular ordered if for all* $\alpha \in [\underline{\alpha}, \overline{\alpha}]$,

$$-\frac{H_\alpha(w|\alpha)}{h(w|\alpha)} \text{ is strictly increasing in } w \text{ on } [w(0, \alpha), w(1, \alpha)].$$

This notion of "supermodular precision" was first introduced into the literature by Ganuza and Penalva (2010). More recently, Shi (2012) also assumes that the information structures are supermodular ordered for some of his results. To understand the definition,

[3] For each $\alpha \in [\underline{\alpha}, \overline{\alpha}]$, let $G(\theta|x, \alpha)$ denote the conditional distribution of θ given x. Then one sufficient condition for this is to assume that $G(\theta|x, \alpha)$ satisfies the monotone likelihood ratio property.

[4] s is uniform on $[0, 1]$ only if $G(\cdot|\alpha)$ is continuous and strictly increasing. This can be assumed without loss of generality. If $G(\cdot|\alpha)$ has a discontinuity at z, where $\mathbb{P}(\tilde{x} = z|\alpha) = p$, then x can be transformed into x^*, which has a continuous and strictly increasing distribution function using the following construction proposed in Lehmann (1988): $x^* = x$ for $x < z$, $x^* = x + pU$ if $x = z$, where U is uniform on $(0, 1)$, and $x^* = x + p$ for $x > z$.

observe that $w_\alpha(s,\alpha) = -H_\alpha(w(s,\alpha)|\alpha)/h(w(s,\alpha)|\alpha)$, which is strictly increasing in s. We prove in Appendix A that this implies that $w(s,\alpha)$ is strictly supermodular:[5]

Lemma 1 *Suppose that the information structures are supermodular ordered. Then $w(\cdot,\cdot)$ is strictly supermodular: for all $s, s' \in (0,1)$, $s > s'$ and $\alpha > \alpha'$*

$$w(s,\alpha) - w(s',\alpha) > w(s,\alpha') - w(s',\alpha'). \tag{1}$$

Intuitively, if the signal s contains little information about θ, the posterior mean $w(s,\alpha)$ does not vary much as s changes and its distribution concentrates around μ. As s becomes more informative about θ, $w(s,\alpha)$ varies more as s changes (see (1)) and its distribution becomes more dispersed.[6]

For some results of this paper, we further require that the information structures are *uniformly supermodular ordered*. Recall that if an information structure is more informative, $w(s,\alpha)$ changes more dramatically as s changes. Hence, we can interpret the change rate $w_s(s,\alpha)$ as a local measure of the information structure's accuracy around s. Then $w_{s,\alpha}(s,\alpha)/w_s(s,\alpha)$ is the percentage change of the information structure's accuracy around s as α increases. We say that the information structures are uniformly supermodular ordered if

$$\frac{w_{s,\alpha}(s,\alpha^*)}{w_s(s,\alpha^*)} = \frac{\partial}{\partial w}\left[-\frac{H_\alpha(w(s,\alpha^*)|\alpha^*)}{h(w(s,\alpha^*)|\alpha^*)}\right]$$

is independent of s (or equivalently w). In other words, when α increases, the information structure becomes more informative "uniformly" over $[0,1]$. The formal definition is given as follows:

Definition 2 *The information structures are uniformly supermodular ordered if there exists a positive function $b : [\underline{\alpha}, \overline{\alpha}] \to \mathbb{R}_{++}$ such that, for all $\alpha \in [\underline{\alpha}, \overline{\alpha}]$ and $w \in [w(0,\alpha), w(1,\alpha)]$,*

$$-\frac{H_\alpha(w|\alpha)}{h(w|\alpha)} = \frac{w-\mu}{b(\alpha)}.$$

[5] Lemma 1 is not an equivalent definition of supermoduler ordered information structures. If $w(\cdot,\cdot)$ is strictly supermodular (i.e. satisfies inequality (1)), $-H_\alpha(\cdot|\alpha)/h(\cdot|\alpha)$ is non-decreasing but not necessarily strictly increasing. I conjecture that all the results still hold if we require only that $-H_\alpha(\cdot|\alpha)/h(\cdot|\alpha)$ is non-decreasing, but the stronger assumption simplifies the analysis.

[6] (See Ganuza and Penalva (2010)) Formally, if $\alpha > \alpha'$, $w(s,\alpha)$ is strictly larger than $w(s,\alpha')$ in the *dispersive order*. Let Y and Z be two real-valued random variables with distributions F and G, respectively. We say Y is greater than Z in the dispersive order if for all $q, p \in (0,1)$ and $q > p$,

$$F^{-1}(q) - F^{-1}(p) \geq G^{-1}(q) - G^{-1}(p).$$

When the information structures are uniformly supermodular ordered, we can obtain a sharper and simpler characterization of the ex ante efficient mechanisms. The following two commonly used information technologies in the literature are uniformly supermodular ordered:[7]

Example 1 (Linear experiments) *Consider the following information structures, which are called "truth-or-noise" in Lewis and Sappington (1994), Johnson and Myatt (2006) and Shi (2012). x_i is equal to agent i's true type θ_i with probability $\alpha_i \in [0,1]$ and is an independent draw from F with probability $1-\alpha_i$. Because the marginal distribution of x_i is F, the transformed signal is $s_i = F(x_i)$. The posterior mean of an agent who chooses α_i and receives s_i is $w(s_i,\alpha_i) = \alpha_i F^{-1}(s_i) + (1-\alpha_i)\mu$. It is easy to verify that*

$$-\frac{H_{\alpha_i}(w_i|\alpha_i)}{h(w_i|\alpha_i)} = \frac{w_i - \mu}{\alpha_i}.$$

Hence, the information structures are uniformly supermodular ordered.

Example 2 (Normal experiments) *Let $\{\theta_i\}$ be independently distributed with a normal distribution: $\theta_i \overset{iid}{\sim} \mathcal{N}(\mu, 1/\beta)$ and $\beta > 0$. Agent i can obtain a costly signal $x_i = \theta_i + \varepsilon_i$, where $\varepsilon_i \overset{iid}{\sim} \mathcal{N}(0, 1/\alpha_i)$ and $\alpha_i > 0$. Because the marginal distribution of x_i is also normal with $x_i \sim \mathcal{N}(\mu, (\beta+\alpha_i)/\beta\alpha_i)$, the transformed signal is $s_i = \Phi\left(\sqrt{\beta\alpha_i}(x_i-\mu)/\sqrt{\beta+\alpha_i}\right)$, where Φ is the distribution function of the standard normal distribution. The posterior mean of an agent who chooses α_i and receives s_i is*

$$w(s_i,\alpha_i) = \mu + \frac{\sqrt{\alpha_i}\Phi^{-1}(s_i)}{\sqrt{\beta(\alpha_i+\beta)}}.$$

It is easy to verify that

$$-\frac{H_{\alpha_i}(w_i|\alpha_i)}{h(w_i|\alpha_i)} = \frac{\beta(w_i-\mu)}{2\alpha_i(\alpha_i+\beta)}.$$

Hence, the information structures are uniformly supermodular ordered.

2.2 Timing

The game proceeds in the following way: The social planner announces a mechanism. After observing the mechanism, the agents simultaneously choose their information structures $\{\alpha_i\}$ and observe the realized signals $\{s_i\}$. Both α_i and s_i are agent i's private information.

[7] See, for example, Ganuza and Penalva (2010) and Shi (2012).

The agents then simultaneously decide whether to participate in the mechanism. All participating agent report their private information to the mechanism. Based on their reports, an allocation and payments are implemented according to the announced mechanism.

We assume that the payoff structure, the timing of the game, the prior distribution and the family of information structures are common knowledge. As is standard in the literature, the solution concept is Bayesian Nash equilibrium.

2.3 Mechanisms

The private information of agent i is two-dimensional including his choice of information structure α_i and the realized signal s_i. However, similar to Biais et al. (2000), Szalay (2009) and Shi (2012), the usual difficulties inherent in multi-dimensional mechanism design problem do not arise here. This is because the posterior mean, $w(s_i, \alpha_i)$, summarizes all the private information needed to compute agent i's expected valuation of the object:

$$\mathbb{E}_{\boldsymbol{\theta}}[v_i(\boldsymbol{\theta})|\alpha_i, s_i] = w(s_i, \alpha_i) + \gamma \sum_{j \neq i} \mathbb{E}[\theta_j].$$

Furthermore, the social planner cannot screen agents with the same posterior mean but different choices of information structures. Hence, we can appeal to the revelation principle and focus on direct mechanisms in which agents report their posterior means directly. For ease of notation, we use w_i to denote $w_i(s_i, \alpha_i)$ and $\boldsymbol{w} := (w_1, \ldots, w_n)$ to denote a vector of posterior means. A direct mechanism is a pair $(\boldsymbol{q}, \boldsymbol{t})$, where $\boldsymbol{q} := (q_1, \ldots, q_n)$ and $\boldsymbol{t} := (t_1, \ldots, t_n)$. For each $i = 1, \ldots, n$, $q_i : [\underline{\theta}, \overline{\theta}]^n \to [0, 1]$ maps a vector of reported posterior means $\boldsymbol{w}$ to agent i's probability of receiving the object and $t_i : [\underline{\theta}, \overline{\theta}]^n \to \mathbb{R}$ maps $\boldsymbol{w}$ to his payment. We note here that the message space for each agent in a direct mechanism is the support of the prior distribution $[\underline{\theta}, \overline{\theta}]$ because any $w_i \in [\underline{\theta}, \overline{\theta}]$ can arise in the game without further knowledge on agents' choices of information structures.

Given a mechanism $(\boldsymbol{q}, \boldsymbol{t})$, let $\boldsymbol{\alpha}^* := (\alpha_1^*, \ldots, \alpha_n^*)$ denote the equilibrium vector of information structures. Define agent i's interim probability of receiving the object as

$$Q_i(w_i) := \mathbb{E}_{w_{-i}}[q_i(w_i, w_{-i})|\alpha_{-i}^*], \; \forall w_i \in [\underline{\theta}, \overline{\theta}], \tag{2}$$

where α_{-i}^* are his opponents' information structures. Then the interim utility of agent i who has a posterior mean w_i and reports w_i' is

$$U_i(w_i, w_i') := w_i Q_i(w_i') + \mathbb{E}_{w_{-i}}\left[\gamma \left(\sum_{j \neq i} w_j\right) q_i(w_i', w_{-i}) - t_i(w_i', w_{-i}) \middle| \alpha_{-i}^*\right].$$

Note that $Q_i(w_i)$ and $U_i(w_i, w_i')$ also depend on α^*_{-i}. Here we suppress the dependence for ease of notation.

We require the mechanism chosen by the social planner to satisfy the following constraints. First, the mechanism must be (interim) individually rational (IR):

$$U_i(w_i) := U_i(w_i, w_i) \geq 0,\ \forall w_i \in [\underline{\theta}, \overline{\theta}], \tag{IR}$$

so that the agents are willing to participate in the mechanism. Because the social planner's goal is to maximize the social surplus, and transfers between the agents and the social planner do not affect the social surplus, we can guarantee that (IR) is satisfied by making sufficiently large lump sum transfers to the agents. Furthermore, as demonstrated later, under some regularity condition the lump sum transfers can be chosen so that (IR) is satisfied and the social planner's revenue is non-negative. Hence, unless noted otherwise, we will ignore (IR) in the remainder of the paper. Second, the mechanism must be Bayesian incentive compatible (IC):

$$U_i(w_i) \geq U_i(w_i, w_i'), \forall w_i, w_i' \in [\underline{\theta}, \overline{\theta}], \tag{IC}$$

so that truth-telling is a Bayesian Nash equilibrium. Lastly, because the information structure chosen by an agent is unobservable, the mechanism must also satisfy the information acquisition constraint (IA). That is, no agent stands to gain by deviating from his equilibrium choice: for each agent i,

$$\alpha_i^* \in \underset{\alpha_i}{\operatorname{argmax}}\, \mathbb{E}_{\boldsymbol{w}} \left[q_i(\boldsymbol{w}) \left(w_i + \gamma \sum_{j\neq i} \mathbb{E}[\theta_j] \right) - t_i(\boldsymbol{w}) \middle| \alpha_i, \alpha_j = \alpha_j^*\ \forall j \neq i \right] - C(\alpha_i). \tag{IA}$$

The social planner's problem, denoted by ($\mathcal{P}$), is to choose a mechanism $(\boldsymbol{q}, \boldsymbol{t})$ and a vector of recommendations of information structures $\boldsymbol{\alpha}^*$ to maximize the expected social surplus:

$$\max_{\boldsymbol{\alpha}^*, (\boldsymbol{q}, \boldsymbol{t})} \mathbb{E}_{\boldsymbol{w}} \left[\sum_i (w_i + \gamma \sum_{j \neq i} w_j) q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha_i^*\ \forall i \right] - \sum_i C(\alpha_i^*), \tag{$\mathcal{P}$}$$

subject to (IC), (IA) and the feasibility constraint (F):

$$0 \leq q_i(\boldsymbol{w}) \leq 1,\ \sum_i q_i(\boldsymbol{w}) \leq 1, \forall \boldsymbol{w}. \tag{F}$$

We say that a mechanism $(\boldsymbol{q}, \boldsymbol{t})$ is *ex ante efficient* or *optimal* if there exists $\boldsymbol{\alpha}^*$ such that $\boldsymbol{\alpha}^*$ and $(\boldsymbol{q}, \boldsymbol{t})$ solve the social planner's problem. We say that a mechanism is *ex post efficient* or *optimal* if the allocation is efficient given acquired information: for all i, $q_i(\boldsymbol{w}) = 1$

if $w_i > \max_j w_j$ and $q_i(\boldsymbol{w}) = 0$ if $w_i < \max_j w_j$.

3 Efficient mechanisms

In this section, we restrict attention to mechanisms that treat all agents symmetrically[8] as well as symmetric equilibria in which all agents choose the same information structure (i.e. $\alpha_i^* = \alpha^*$ for all i). This restriction significantly simplifies the analysis but it may result in a loss of generality. We will relax the symmetric restriction in Section 4.

When the ex post allocation rule $\boldsymbol{q}$ is symmetric and all agents choose the same information structure, the corresponding interim allocation rule Q_i is independent of i. From here on, we drop the subscript i from Q, w and α whenever the meaning is clear. By the standard argument,[9] (IC) holds if and only if

$$Q(w) \text{ is non-decreasing in } w, \tag{MON}$$

and $U_i(w)$ is absolutely continuous and satisfies the following envelope condition

$$U_i(w) = U_i(\underline{\theta}) + \int_{\underline{\theta}}^{w} Q(\tilde{w})\mathrm{d}\tilde{w}, \ \forall w \in [\underline{\theta}, \overline{\theta}]. \tag{3}$$

Consider next the agent's incentives to acquire information. Supposing that agent i chooses α_i, his (ex ante) expected payoff is given by

$$\begin{aligned}&\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} U_i(w)\mathrm{d}H(w|\alpha_i) - C(\alpha_i)\\ =&U_i(w(0,\alpha_i)) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} [1 - H(w_i|\alpha_i)]\, Q(w_i)\mathrm{d}w_i - C(\alpha_i).\end{aligned} \tag{4}$$

(The derivation of equality (4) can be found in Appendix A.) Hence, (IA) becomes

$$\alpha^* \in \underset{\alpha_i}{\operatorname{argmax}}\, U_i(w(0,\alpha_i)) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} [1 - H(w_i|\alpha_i)]\, Q(w_i)\mathrm{d}w_i - C(\alpha_i).$$

This condition is hard to work with directly. We follow the standard first-order approach, and replace (IA) by a one-sided first-order necessary condition. In an earlier paper, Bergemann and Välimäki (2002) show that if the social planner adopts the ex post efficient mechanism, the agents tend to acquire more information than the socially desired level. This result

[8] The formal definition of symmetric mechanisms can be found at the beginning of Appendix A.

[9] See, for example, Myerson (1981).

suggests that an ex ante efficient mechanism would distort the allocation of the object to discourage agents from acquiring excessive information. Hence, we hypothesize that to ensure that (IA) holds in an ex ante efficient mechanism, it suffices to ensure that no agent has incentives to acquire more accurate signalsq than recommended: for all $\alpha_i > \alpha^*$,

$$\begin{aligned} &U_i(w(0,\alpha^*)) + \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} [1 - H(w|\alpha^*)]\, Q(w)\mathrm{d}w - C(\alpha^*) \\ &\geq U_i(w(0,\alpha_i)) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} [1 - H(w|\alpha_i)]\, Q(w)\mathrm{d}w - C(\alpha_i). \end{aligned}$$

This implies the following one-sided first-order condition:[10]

$$\mathbb{E}_w\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)} Q(w) \middle| \alpha_i = \alpha^* \right] \leq C'(\alpha^*). \tag{IA$'$}$$

The left-hand side of the above inequality is agent i's marginal benefit from choosing α^*, and the right-hand side is the marginal information cost. We show that, for any non-decreasing interim allocation rule Q, an agent's marginal benefit from acquiring information is non-negative:

Lemma 2 *Suppose that $Q : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$ is non-decreasing on $[w(0,\alpha_i), w(1,\alpha_i)]$, then*

$$\mathbb{E}_{w_i}\left[-\frac{H_{\alpha_i}(w_i|\alpha_i)}{h(w_i|\alpha_i)} Q(w_i) \middle| \alpha_i \right] \geq 0, \tag{5}$$

where the equality holds if Q is constant.

An important implication of Lemma 2 is that if an interim allocation rule Q is "steeper" than another one Q' in the sense that their difference $Q - Q'$ is non-decreasing, agent i's marginal benefit from acquiring information is higher under Q. Intuitively, a steeper interim allocation rule implies that the mechanism's outcome is more sensitive to an agent's private information and gives him stronger incentives to acquire information. (In fact, an agent's marginal benefit from acquiring information is higher under Q if there exists a partition of $[\underline{\theta}, \overline{\theta}]$ such that $Q - Q'$ is non-decreasing in each interval; we omit the details here.) This property is important for understanding the structure of the ex ante efficient mechanisms later.

[10] The main technical reason why we consider a one-sided first-order condition here is to sign the Lagrangian multiplier associated with (IA$'$). Admittedly, this relaxation also simplifies the proof of Theorem 1. But my conjecture is that Theorem 1 can be proved even if we require the first-order condition holds with equality.

Subsequently, we first consider the social planner's relaxed problem by replacing (IA) by (IA′), and then show that (IA′) holds with equality when α^* is chosen optimally. Finally, the first-order approach is valid if the second-order condition of the agents' optimization problem is satisfied. Appendix A.3 provides sufficient conditions that ensure the first-order approach to be valid.

Although (IA′) is easier to work with than (IA), it is still nonstandard and prevents us from solving the social planner's problem directly as in Myerson (1981). To overcome this difficulty, we focus on reduced form auctions. Formally, a reduced form allocation rule is defined as follows:

Definition 3 *An allocation rule* $\boldsymbol{q}$ implements $Q : [\underline{\theta}, \overline{\theta}] \to [0, 1]$ *and* Q *is the* reduced form *of* $\boldsymbol{q}$ *if* $\boldsymbol{q}$ *satisfies* (2) *and* (F) *for all* $w \in [\underline{\theta}, \overline{\theta}]$. Q *is* implementable *if* $\boldsymbol{q}$ *exists implementing* Q.

One important prior result we use in this paper is the necessary and sufficient condition of Maskin and Riley (1984), Matthews (1984b) and Border (1991), which characterizes the interim allocation rules implementable by symmetric mechanisms. By Theorem 1 in Matthews (1984b), any non-decreasing function $Q : [\underline{\theta}, \overline{\theta}] \to [0, 1]$ is implementable if and only if it satisfies

$$Y(w) := \int_w^{\overline{\theta}} \left[H(z|\alpha^*)^{n-1} - Q(z)\right] h(z|\alpha^*)\mathrm{d}z \geq 0, \ \forall w \in [\underline{\theta}, \overline{\theta}]. \tag{F′}$$

The above condition says that the probability of assigning the object to an agent whose posterior mean is above w, $n \int_w^{\overline{\theta}} Q(z)h(z|\alpha^*)\mathrm{d}z$, must not exceed the probability with which an agent whose posterior mean is above w exists, $1 - H(w|\alpha^*)^n = n \int_w^{\overline{\theta}} H(z|\alpha^*)^{n-1}h(z|\alpha^*)\mathrm{d}z$. Clearly, this is a necessary condition for Q to be implementable. If Q is non-decreasing, Theorem 1 in Matthews (1984b) proves that this condition is also sufficient. Hence, given (MON), we can replace (F) by (F′). Note that in equilibrium the support of the posterior means is give by $W := [w(0, \alpha^*), w(1, \alpha^*)] \subset [\underline{\theta}, \overline{\theta}]$. Therefore, (F′) imposes no restriction on Q outside W.

To summarize, the social planner's relaxed problem, denoted by ($\mathcal{P}'$), becomes:

$$\max_{\alpha^*, (\boldsymbol{q}, \boldsymbol{t})} \mathbb{E}\left[\sum_i (w_i + \gamma \sum_{j \neq i} w_j) q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha^* \ \forall i\right] - nC(\alpha^*),$$

subject to (MON), (IA′) and (F′). Note that all three constraints ((IC), (IA) and (F)) in the original problem ($\mathcal{P}$) are replaced by constraints ((MON), (IA′) and (F′)) that are expressed as functionals of the interim allocation rule Q. In order to work with reduced form auctions, we must also express the social planner's objective function or the expected social surplus as a functional of Q. This exercise is trivial when agents have private values ($\gamma = 0$), because in this case an agent's expected valuation of the object and his winning probability are independent conditional on his private information. In general, this is impossible when agents' values are interdependent ($\gamma > 0$), because in this case both an agent's expected valuation of the object and his winning probability depend on other agents' private information. Nonetheless, we can still write the expected social surplus as a functional of Q if Q is the reduced form of an ex ante efficient allocation rule, which never withholds the object. This is the result of Theorem 1.

Theorem 1 *Suppose that the information structures are supermodular ordered, and α^* and $(\boldsymbol{q}, \boldsymbol{t})$ solve the social planner's relaxed problem $(\mathcal{P}')$. Then*

$$\sum_i q_i(\boldsymbol{w}) = 1 \text{ for almost all } \boldsymbol{w} \in W^n. \tag{6}$$

The proof of Theorem 1 can be found in Section 3.1. Using Theorem 1 and the law of iterated expectations, the social planner's objective function can be rewritten as a functional of Q:

$$\mathbb{E}_{\boldsymbol{w}}\left[\sum_i (w_i + \gamma \sum_{j \neq i} w_j) q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha^* \ \forall i\right] = \sum_i \mathbb{E}_{w_i}\left[(1-\gamma) w_i Q(w_i) \middle| \alpha_i = \alpha^*\right] + n\gamma\mu.$$

Because the second term, $n\gamma\mu$, is a constant, we ignore it from here on. Hence, the social planner's relaxed problem $(\mathcal{P}')$ can be rewritten as follows:

$$\max_{\alpha^*, Q} \mathbb{E}_w\left[(1-\gamma) w Q(w) \middle| \alpha^*\right] - C(\alpha^*), \tag{$\mathcal{P}'$}$$

subject to

$$Y(w) = \int_w^{\overline{\theta}} [H(z|\alpha^*)^{n-1} - Q(z)] h(z|\alpha^*) \mathrm{d}z \geq 0, \ \forall w \in [\underline{\theta}, \overline{\theta}]. \tag{F$'$}$$

$$Q(w) \text{ is non-decreasing in } w, \tag{MON}$$

$$\mathbb{E}_w\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)} Q(w) \middle| \alpha^*\right] \leq C'(\alpha^*). \tag{IA$'$}$$

In addition to being instrumental in solving the social planner's problem, Theorem 1 also has some inherent economic interest. Obviously, when information is exogenous, the efficient mechanism never withholds the object. This is not obvious when information is endogenous, because, by withholding the object occasionally, the social planner can discourage agents from acquiring excessive information, which may improve efficiency ex ante. However, intuitively, whenever the social planner withholds the object, she can also allocate it randomly. By doing so, the ex post allocative efficiency improves while the agents' ex ante incentives to acquire information are unaffected.

Though intuitive, the proof of Theorem 1 is non-trivial. This is because the resulting mechanism, by simply randomizing the object whenever it is withheld, is likely to violate (MON) or (IA′). To illustrate this difficulty, let A be a set of types such that $\sum_i q_i(\boldsymbol{w}) < 1$ whenever $\boldsymbol{w} \in A^n$. Suppose that A has a "hole" in the sense that there exists an interval $(\underline{w}, \overline{w})$ such that $(\underline{w}, \overline{w}) \cap A = \emptyset$ and $\inf A < \underline{w} < \overline{w} < \sup A$.

If we simply redefine $\boldsymbol{q}$ such that it remains unchanged outside A^n and $\sum_i q_i(\boldsymbol{w}) = 1$ for all $\boldsymbol{w} \in A^n$, the resulting Q remains unchanged for all $w \in (\underline{w}, \overline{w})$ but increases for all $w \in A$. If we allocate the object too often to agents whose types are in $[\underline{\theta}, \underline{w}] \cap A$, the resulting Q will no longer be non-decreasing and violate (MON). If we allocate the object too often to agents whose types are in $[\overline{w}, \overline{\theta}] \cap A$, the resulting Q is becomes steeper. Because a steeper interim allocation rule gives an agent a higher marginal benefit from acquiring information, this change will lead to a violation of (IA′). Hence, to ensure that the new $\boldsymbol{q}$ generates a higher social surplus while respects all the constraints, we must adjust $\boldsymbol{q}$ not only inside A^n, but also outside A^n.

Finally, Theorem 1 implies that under some regularity condition the lump sum transfers can be chosen so that (IR) is satisfied and the social planner's revenue is non-negative. This result is summarized in the following corollary.

Corollary 1 *Suppose that the information structures are supermodular ordered, and α^* and $(\boldsymbol{q}, \boldsymbol{t})$ solve the social planner's relaxed problem ($\mathcal{P}'$). Suppose, in addition, that $(1-\gamma)w - [1 - H(w|\alpha^*)]/h(w|\alpha^*)$ is non-decreasing in w. Then there exists $\tilde{\boldsymbol{t}}$ such that α^* and $(\boldsymbol{q}, \tilde{\boldsymbol{t}})$ solve ($\mathcal{P}'$), and (IR) is satisfied and the social planner's revenue is non-negative under this new mechanism.*

3.1 Proof of Theorem 1

This section contains the proof of Theorem 1. The readers who are not interested in the proof may skip this section and proceed directly to Section 3.2 without loss of continuity.

We prove Theorem 1 by proving lemmas 3 and 4. Observe first that if $\alpha_i = \alpha^*$ for all i, $Y(w(0, \alpha^*))$ is equal to 1 minus the probability of assigning the object to some agent. Clearly, (6) is violated if and only if $Y(w(0, \alpha^*)) > 0$. Then we have the following lemma:

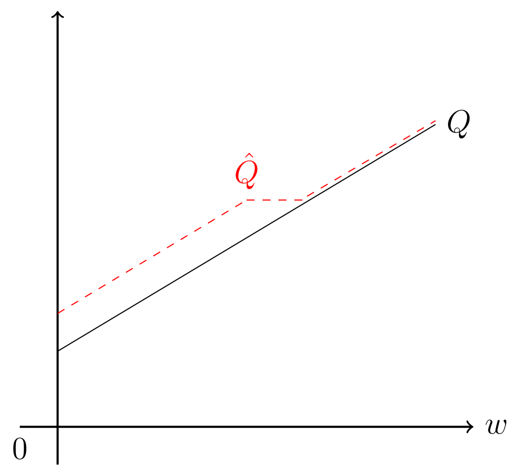


Figure 1: Proof idea of Lemma 3

Lemma 3 *Suppose that the information structures are supermodular ordered and $\alpha_i = \alpha^*$ for all i. Let Q be any interim allocation rule satisfying (F′), (MON), (IA′) and $Y(w(0, \alpha^*)) > 0$. Then there exists $\hat{Q}$ satisfying (F′), (MON) and (IA′) such that*

$$\hat{Q}(w) \geq Q(w), \; \forall w \in W, \tag{7}$$

and the strict inequality holds for a set of w with positive measure.

The intuition behind the proof of Lemma 3 can be illustrated by Figure 1. Supposing that Q satisfies the assumptions in Lemma 3, one can construct another interim allocation rule $\hat{Q}$ by increasing Q at the lower end of its domain as in Figure 1. Clearly, the resulting $\hat{Q}$ is non-decreasing and implementable if the change is sufficiently small. It remains to verify that $\hat{Q}$ also satisfies (IA′). Intuitively, agents have weaker incentives to acquire information if outcomes are less sensitive to changes in their private information. Formally, recall that Lemma 2 implies that if $\hat{Q}$ is less steep than Q in the sense that it differs from Q by a non-increasing function (as in Figure 1), for any amount of information acquired (or any α), $\hat{Q}$ gives agents a smaller marginal benefit from acquiring information that Q does. Hence, $\hat{Q}$ satisfies (IA′) as Q does.

The gap between Lemma 3 and Theorem 1 is that, when $\gamma > 0$, the expected social surplus

$$\mathbb{E}_{\boldsymbol{w}}\left[\sum_i (w_i + \gamma \sum_{j \neq i} w_j) q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha^* \; \forall i\right]$$

does not directly depend on Q. To prove Theorem 1, we need to show that, for any ex-post allocation rule $\boldsymbol{q}$ implementing Q, we can find an ex post allocation rule $\hat{\boldsymbol{q}}$ that implements $\hat{Q}$ and yields higher expected social surplus. This is the result of Lemma 4.

Lemma 4 *Suppose that the information structures are supermodular ordered and $\alpha_i = \alpha^*$ for all i. Let Q and $\hat{Q}$ be two implementable allocation rules satisfying (7). Let $\boldsymbol{q}$ be an ex-post allocation rule that implements Q. Then there exists an ex-post allocation rule $\hat{\boldsymbol{q}}$ that implements $\hat{Q}$ and satisfies*

$$\mathbb{E}_{\boldsymbol{w}}\left[\sum_i (w_i + \gamma \sum_{j\neq i} w_j)\hat{q}_i(\boldsymbol{w}) \middle| \alpha_i = \alpha^* \ \forall i\right] > \mathbb{E}_{\boldsymbol{w}}\left[\sum_i (w_i + \gamma \sum_{j\neq i} w_j)q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha^* \ \forall i\right].$$

The proof of Lemma 4 relies on the following technical lemma. We slightly abuse notation a bit and let h denote the probability measure on W corresponding to $H(w_i|\alpha^*)$.

Lemma 5 *Let $Q : W \to [0,1]$ be an interim allocation rule and $\rho : W^n \to [0,1]$ be a symmetric measurable function. Then there exists a symmetric ex post allocation rule $\boldsymbol{q}$ that implements Q and satisfies $\sum_i q_i(\boldsymbol{w}) \geq \rho(\boldsymbol{w})$ for almost all $\boldsymbol{w} \in W^n$ if and only if, for all measurable sets $A \subset W$, the following inequality holds*

$$\int_{A^n} \rho(\boldsymbol{w})\mathrm{d}h^n(\boldsymbol{w}) \leq n \int_A Q(w_i)\mathrm{d}h(w_i) \leq \int_{A^n} \mathrm{d}h^n(\boldsymbol{w}). \tag{8}$$

To see that inequality (8) is necessary, suppose that there exists an ex post allocation rule $\boldsymbol{q}$ that implements Q and satisfies $\sum_i q_i(\boldsymbol{w}) \geq \rho(\boldsymbol{w})$ for almost all $\boldsymbol{w} \in W^n$. For any measurable set $A \subset W$, the probability with which some agent whose type is in A receives the object is given by $n \int_A Q(w_i)\mathrm{d}h(w_i)$. On the one hand, this probability must exceed the probability with which some agent receives the object when all agents' types are in A, $\int_{A^n} \sum_i q_i(\boldsymbol{w})\mathrm{d}h^n(\boldsymbol{w})$, which is bounded below by $\int_{A^n} \rho(\boldsymbol{w})\mathrm{d}h^n(\boldsymbol{w})$ by assumption. This gives rise to the first inequality. On the other hand, it cannot exceed the probability with which an agent whose type is in A exists, $\int_{A^n} \mathrm{d}h^n(\boldsymbol{w})$. This gives rise to the second inequality. In Appendix A.1, we show that (8) is also sufficient.[11]

With Lemma 5 in hand, it is easy to prove Lemma 4.

Proof of Lemma 4. Consider two implementable allocation rules Q and $\hat{Q}$ satisfying (7). Let $\boldsymbol{q}$ be a symmetric ex-post allocation rule that implements Q. Define $\rho : W^n \to [0,1]$ by $\rho(\boldsymbol{w}) := \sum_i q_i(\boldsymbol{w})$ for all $\boldsymbol{w} \in W^n$. Then ρ is symmetric. By Lemma 5,

$$\int_{A^n} \mathrm{d}h^n(\boldsymbol{w}) \geq n \int_A \hat{Q}(w_i)\mathrm{d}h(w_i) \geq n \int_A Q(w_i)\mathrm{d}h(w_i) \geq \int_{A^n} \rho(\boldsymbol{w})\mathrm{d}h^n(\boldsymbol{w}).$$

By Lemma 5, there exists an allocation rule $\hat{\boldsymbol{q}}$ that implements $\hat{Q}$ and satisfies $\sum \hat{q}_i(\boldsymbol{w}) \geq$

[11] Note also that if $A = [w, \overline{\theta}]$, then the second inequality in (8) becomes (F′).

$\rho(\boldsymbol{w}) = \sum_i q_i(\boldsymbol{w})$ for almost all $\boldsymbol{w} \in W^n$. Hence,

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{w}}\left[\sum_i \left(w_i + \gamma \sum_{j\neq i} w_j\right) \hat{q}_i(\boldsymbol{w}) \middle| \alpha_i = \alpha^* \ \forall i\right] \\
&= \sum_i \mathbb{E}_{w_i}\left[(1-\gamma) w_i \hat{Q}(w_i) \middle| \alpha_i = \alpha^*\right] + \mathbb{E}_{\boldsymbol{w}}\left[\left(\gamma \sum_i w_i\right)\left(\sum_i \hat{q}_i(\boldsymbol{w})\right) \middle| \alpha_i = \alpha^* \ \forall i\right] \\
&> \sum_i \mathbb{E}_{w_i}\left[(1-\gamma) w_i Q(w_i) \middle| \alpha_i = \alpha^*\right] + \mathbb{E}_{\boldsymbol{w}}\left[\left(\gamma \sum_i w_i\right)\left(\sum_i q_i(\boldsymbol{w})\right) \middle| \alpha_i = \alpha^* \ \forall i\right] \\
&= \mathbb{E}_{\boldsymbol{w}}\left[\sum_i \left(w_i + \gamma \sum_{j\neq i} w_j\right) q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha^* \ \forall i\right],
\end{aligned}$$

where the strict inequality holds because Q and $\hat{Q}$ satisfies (7) and $\sum_i \hat{q}_i(\boldsymbol{w}) \geq \sum_i q_i(\boldsymbol{w})$ for almost all $\boldsymbol{w} \in W^n$. This completes the proof. ■

3.2 Optimal mechanisms for fixed α^*

We solve the principal's relaxed problem ($\mathcal{P}'$) in two steps. In this subsection, we solve the following sub-problem for each $\alpha^* \in [\underline{\alpha}, \overline{\alpha}]$, denoted by ($\mathcal{P}'$-$\alpha^*$):

$$V(\alpha^*) := \max_Q \mathbb{E}_w \left[wQ(w) \middle| \alpha^* \right], \tag{$\mathcal{P}'$-α^*}$$

subject to (F′), (MON) and (IA′). In Section 3.3, we solve $\max_{\alpha \in [\underline{\alpha}, \overline{\alpha}]} (1-\gamma)V(\alpha) - C(\alpha)$.

Fix α^*. If the principal adopts the ex post efficient mechanism, the interim allocation rule is given by $Q(w) = H(w|\alpha^*)^{n-1}$ for all w. Clearly, if α^* is such that

$$\mathbb{E}_w \left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)} H(w|\alpha^*)^{n-1} \middle| \alpha^* \right] \leq C'(\alpha^*), \tag{9}$$

the ex post efficient mechanism solves ($\mathcal{P}'$-α^*). In the rest of this subsection, we assume that α^* is such that (9) is violated. In what follows, we consider two cases in turn. In Section 3.2.1, we consider the special case in which the information structures are uniformly supermodular ordered. In this case, we first solve a relaxed problem of ($\mathcal{P}'$-α^*) by ignoring the monotonicity constraint (MON), and then show that if the information structures are uniformly supermodular ordered, the solutions of this relaxed problem automatically satisfy (MON). In Section 3.2.2, we consider the general case when the information structures are supermodular ordered. In this case, the solutions of the relaxed problem violates (MON) in general. In the main text, we present an informal arugument to derive the optimal solutions of ($\mathcal{P}'$-α^*) using Myerson (1981)'s ironing procedure. The formal analysis can be found in Appendix A.2.

3.2.1 Optimal mechanisms in the regular case

If we ignore the monotonicity constraint (MON), the following Lagrangian relaxation can be used to get an intuition of the optimal solution:

$$\mathscr{L} := \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \varphi^{\lambda_X}\left(H(w|\alpha^*), \alpha^*\right) Q(w) h(w|\alpha^*) \mathrm{d}w + \lambda_X C'(\alpha^*), \tag{10}$$

where $\lambda_X > 0$ is the Lagrangian multiplier associated with (IA′) and $\varphi^{\lambda_X}(\cdot, \alpha^*)$ is defined by[12]

$$\varphi^{\lambda_X}(t, \alpha^*) := H^{-1}(t|\alpha^*) + \lambda_X \frac{H_\alpha(H^{-1}(t|\alpha^*)|\alpha^*)}{h(H^{-1}(t|\alpha^*)|\alpha^*)}, \ \forall t \in [0,1].$$

$\varphi^{\lambda_X}(H(w|\alpha^*), \alpha^*)$ can be viewed as the *"virtual value"* associated with posterior mean w:

$$\varphi^{\lambda_X}(H(w|\alpha^*), \alpha^*) = w + \lambda_X \frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}. \ \forall w \in W.$$

The first term in the virtual value, w, is the posterior mean of an agent's type. In the case of private values ($\gamma = 0$), this is equal to his expected valuation of the object. In the ex post efficient mechanism, an agent is rewarded based on his posterior mean. When the agents can choose how much information to acquire, we must subtract, from an agent's posterior mean, λ_X multiplied by the marginal change of his posterior mean if he acquires more precise information:

$$-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)} = w_\alpha(s, \alpha^*),$$

where s is such that $w(s, \alpha^*) = w$. In an ex ante efficient mechanism, an agent is rewarded based on his virtual value. When the information structures are supermodular ordered, $-H_\alpha(w|\alpha^*)/h(w|\alpha^*)$ is strictly increasing in w. Hence, the virtual value (as a function of the posterior mean) is less steeper than the posterior mean. In other words, an agent's virtual value is less sensitive to his private information than his posterior mean. This difference discourages agents from acquiring excessive information as they do under the ex post efficient mechanism.[13]

[12] We define $\varphi^{\lambda_X}(\cdot, \alpha^*)$ as a function of percentiles rather than posterior means simply to make it easier to define "ironed virtual values" later when the pointwise virtual surplus maximizer violates (MON).

[13] The standard virtual value in a revenue maximization problem is defined as the difference between a type's true value and the information rents necessary to induce truthtelling. Here, because the social planner's goal is to maximize the social surplus rather than her revenue, the inverse hazard rate associated with the information rents does not appear.

If the virtual values are non-decreasing in w when λ_X is chosen optimally, the optimal solution to ($\mathcal{P}'$-α^*) can be obtained by maximizing the virtual surplus pointwise because there exists a pointwise virtual surplus maximizer that is non-decreasing and satisfies (MON).

This method works in the simple case in which the information structures are uniformly supermodular ordered. Recall that in this case we have

$$-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)} = \frac{w-\mu}{b(\alpha^*)}, \ \forall w.$$

Hence, the virtual values are given by

$$\varphi^{\lambda_X}(H(w|\alpha^*), \alpha^*) = w - \lambda_X \frac{w-\mu}{b(\alpha^*)}, \ \forall w.$$

We argue that the optimal λ_X is equal to $b(\alpha^*)$. Suppose that $\lambda_X < b(\alpha^*)$. In this case, the virtual value is strictly increasing and the pointwise virtual surplus maximizer is the ex post efficient allocation rule: $Q(w) = H(w|\alpha^*)^{n-1}$ for all w. However, by assumption, α^* is such that (9) is violated. This implies that (IA$'$) is violated, which is a contradiction. Hence, $\lambda_X \geq b(\alpha^*)$. Suppose that $\lambda_X > b(\alpha^*)$. In this case, the virtual value is strictly decreasing and the interim allocation rule Q that maximizes the expected virtual surplus and satisfies (MON) is constant. However, this implies that (IA$'$) holds with strict inequality and therefore $\lambda_X = 0$, which contradicts to the hypothesis that $\lambda_X > b(\alpha^*) > 0$. Hence, the optimal λ_X is equal to $b(\alpha^*)$.

When $\lambda_X = b(\alpha^*)$, the virtual value is constant: $\varphi^{\lambda_X}(w, \alpha^*) = \mu$ for all w. Hence, any feasible non-decreasing allocation rule Q satisfying condition (6) maximizes the expected virtual surplus. If Q also satisfies (IA$'$) with equality, it solves ($\mathcal{P}'$-α^*). These arguments prove the following proposition.

Proposition 1 *Suppose that the first-order approach is valid and the information structures are uniformly supermodular ordered. Suppose, in addition, that α^* is such that (9) is violated. Then Q solves ($\mathcal{P}'$-α^*) if and only if Q is non-decreasing and Q satisfies (6) and (IA$'$) with equality.*

Note that there are typically multiple interim allocation rules that solve ($\mathcal{P}'$-α^*). Corollary 2 below describes one of them whose corresponding direct mechanism has a simple and appealing implementation: standard auctions with discrete bids.

Corollary 2 *Suppose that the first-order approach is valid and the information structures are uniformly supermodular ordered. Suppose, in addition, that α^* is such that (9) is violated. There exists a sequence of $\{w^k\}_{k=0}^{m+1}$ such that $w(0, \alpha^*) = w^0 < w^1 < \cdots < w^m <$*

$w^{m+1} = w(1, \alpha^*)$ *such that the following interim allocation rule solves* ($\mathcal{P}'$-α^*)*: for each* $k = 0, 1, \ldots, m$

$$Q(w) := \frac{\frac{1}{n}\left[H(w^{k+1}|\alpha^*)^n - H(w^k|\alpha^*)^n\right]}{H(w^{k+1}|\alpha^*) - H(w^k|\alpha^*)} \text{ if } w^k < w < w^{k+1}.$$

The corresponding optimal direct mechanism can be implemented by a standard auction with m *discrete bids.*

In a standard auction with m discrete bids, an agent can choose a bid from a set predetermined bids: $b^1 < \cdots < b^m$. The agent with the highest bid wins and pays his or the second highest bid depending on the auction rule. Ties are broken uniformly at random. By restricting bids to discrete levels, we bunch nearby posterior means together, which clearly reduces agents' marginal benefits from gathering information. We further illustrate this idea in the following example in which agents can perform linear experiments.

Example 3 (Linear experiments) *Consider the information structures in Example 1. Assume that* $F(\theta) = \theta$ *with support* $[0, 1]$ *and the cost function is given by*

$$C(\alpha) = \frac{3}{8}\left(\alpha - \frac{1}{2}\right)^2, \ \forall \alpha \in \left[\frac{1}{2}, 1\right],$$

Then, as we demonstrate in Appendix A.3, the first-order approach is valid. Assume that there are $n = 2$ *agents. Finally, let* $\gamma = 7/8$.

As a benchmark, consider first the case in which the social planner runs a second-price auction with no restriction on bids. Suppose that both agents choose the same information structure before partipating in the auction. By Krishna (2009), in this auction there exists a symmetric equilibrium in which an agent's bid is strictly increasing in his posterior mean. Therefore, the allocation is ex post efficient. We argue that in this case it is an equilibrium for both agents to choose $\hat{\alpha} = 11/18$*. Supposing that agent 2 chooses* $\hat{\alpha} = 11/18$*, agent 1's marginal benefit from acquiring information is* $1/12$*. Clearly, it is optimal for agent 1 to choose* $\hat{\alpha} = 11/18$ *as well.*

However, this second-price auction is not socially optimal. As we demonstrate later in Example 4, the socially optimal information choice in this example is $\alpha^* = 7/12$*, which is strictly less than* $\hat{\alpha} = 11/18$*. It can be verified that the following interim allocation rule solves* ($\mathcal{P}'$-α^*)*:*

$$Q(w) = \begin{cases} \frac{1}{4} & \text{if } w < \frac{1}{2} \\ \frac{3}{4} & \text{if } w > \frac{1}{2} \end{cases}.$$

Furthermore, the corresponding ex ante efficient mechanism can be implemented by the following second-price auction with two allowable bids:

$$\underline{b} = -\frac{1}{2} + \frac{\alpha^*}{2} - \frac{\gamma\alpha^*}{2} + \frac{\gamma\alpha^{*2}}{4},$$
$$\overline{b} = \frac{5}{2} - \frac{3\alpha^*}{2} + \frac{5\gamma\alpha^*}{2} - \frac{3\gamma\alpha^{*2}}{4}.$$

Ties are broken uniformly at random. We argue that, if both agents chooses $\alpha^* = 7/12$ ex ante, it is an equilibrium in which an agent bids $\underline{b}$ if $w < 1/2$ and bids $\overline{b}$ otherwise. Supposing that agent 2 follows this strategy, the payoff of agent 1 whose posterior mean is w and bids b is given by

$$U(b; w) = \begin{cases} \frac{1}{4}w + \frac{\alpha^*}{8} - \frac{1}{8} & \text{if } b = \underline{b} \\ \frac{3}{4}w + \frac{\alpha^*}{8} - \frac{3}{8} & \text{if } b = \overline{b} \end{cases}.$$

It is easy to see that it is optimal for agent 1 to bid $\underline{b}$ if $w < 1/2$ and bid $\overline{b}$ otherwise. (In fact, this is true regardless of agent 1's information choice.) Next, consider agent 1's incentives to acquire information. Supposing that agent 2 chooses $\alpha^* = 7/12$, agent 1's marginal benefit from acquiring information is $1/16$. Clearly, it is optimal for agent 1 to choose $\alpha^* = 7/12$ as well.

Restricting bids to discrete levels is not uncommon in auctions in practice. For example, eBay auctions require that the next bid must exceed the current price plus a bid increment and FCC spectrum auctions adopt a minimum clock price increment. Most existing auction theories predict that discrete bids lead to inefficiency and they are mainly used in practice to simplify communication processes and speed auctions (see Ausubel and Cramton (2004)). Our results suggest an alternative justification for the prevalence of discrete bids. That is, when agents have interdependent values and can covertly acquire information at some cost, the use of discrete bids can improve ex ante efficiency.

3.2.2 Optimal mechanisms in the general case

If the information structures are not uniformly supermodular ordered, typically a pointwise virtual surplus maximizer is not non-decreasing (or violates (MON)) and ironing is necessary. In particular, an optimal solution can be obtained by ironing $\varphi^{\lambda_X}(\cdot, \alpha^*)$ in the following procedure first introduced by Myerson (1981). For each $t \in [0, 1]$, define

$$J^{\lambda_X}(t, \alpha^*) := \int_0^t \varphi^{\lambda_X}(\tau, \alpha^*) \mathrm{d}\tau.$$

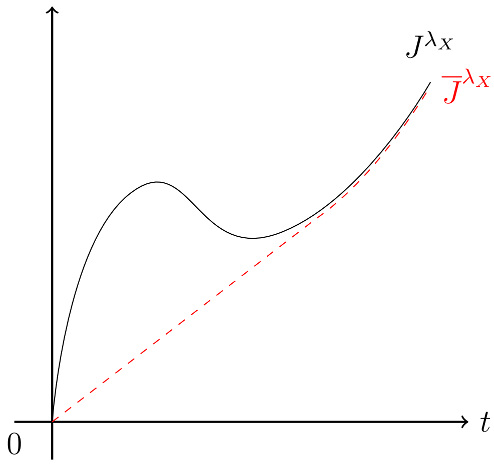


Figure 2: Ironing

Let $\overline{J}^{\lambda_X}$ denote the convex hull of J^{λ_X}, defined by

$$\overline{J}^{\lambda_X}(t, \alpha^*) := \min \left\{ \beta J(t_1, \alpha^*) + (1-\beta)J(t_2, \alpha^*) | t_1, t_2 \in [0,1], \beta t_1 + (1-\beta)t_2 = t \right\}, \ \forall t \in [0,1].$$

This is illustrated by Figure 2. Because $\overline{J}^{\lambda_X}(\cdot, \alpha^*)$ is convex, it is continuously differentiable virtually everywhere. Define the *ironed virtual value* $\overline{\varphi}^{\lambda_X}(\cdot, \alpha^*)$ as follows. First, for each $t \in (0,1)$ such that $\partial \overline{J}^{\lambda_X}(t, \alpha^*)/\partial t$ exists, let $\overline{\varphi}^{\lambda_X}(t, \alpha^*) := \partial \overline{J}^{\lambda_X}(t, \alpha^*)/\partial t$. Second, extend $\overline{\varphi}^{\lambda_X}(\cdot, \alpha^*)$ to $[0,1]$ by right continuity. Because $\overline{J}^{\lambda_X}(\cdot, \alpha^*)$ is convex, $\overline{\varphi}^{\lambda_X}(\cdot, \alpha^*)$ is non-decreasing.

Let λ_X be chosen optimally. By the standard argument, we can show that the expected social surplus is bounded above by the maximum ironed virtual surplus plus a constant:

$$\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} zQ(z)h(z|\alpha^*)\mathrm{d}z \leq \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*), \alpha^*)H(z|\alpha^*)^{n-1}\mathrm{d}H(z|\alpha^*) + \lambda_X C'(\alpha^*). \tag{11}$$

Furthermore, an interim allocation rule Q achieves this upper-bound (or solves ($\mathcal{P}'$-α^*)) if and only if (i) Q satisfies (IA$'$) with equality; (ii) $Y(w(0, \alpha^*)) = 0$ (i.e. Q allocates the object with probability one); and (iii) Q satisfies the following two *pooling properties*:

1. If $J^{\lambda_X}(H(w|\alpha^*), \alpha^*) > \overline{J}^{\lambda_X}(H(w|\alpha^*), \alpha^*)$ for all $w \in (\underline{w}, \overline{w})$ and let $(\underline{w}, \overline{w})$ be chosen maximally, Q is constant on $(\underline{w}, \overline{w})$.

2. If $Y(w) > 0$ for all $w \in (\underline{w}, \overline{w})$ and let $(\underline{w}, \overline{w})$ be chosen maximally, $\overline{\varphi}^{\lambda_X}(H(\cdot|\alpha^*), \alpha^*)$ is constant on $(\underline{w}, \overline{w})$.

(The derivation of inequality (11) can be found in Appendix A.) The first pooling property says that if $J^{\lambda_X}(H(w|\alpha^*), \alpha^*) > \overline{J}^{\lambda_X}(H(w|\alpha^*), \alpha^*)$ for all $w \in (\underline{w}, \overline{w})$ and let $(\underline{w}, \overline{w})$ be chosen maximally, all posterior means in $(\underline{w}, \overline{w})$ are pooled together. The second pooling property says that if posterior means in $(\underline{w}, \overline{w})$ are pooled together, $\overline{\varphi}^{\lambda_X}(H(\cdot|\alpha^*), \alpha^*)$ is constant on $(\underline{w}, \overline{w})$. By construction, $J^{\lambda_X}(\cdot, \alpha^*) \geq \overline{J}^{\lambda_X}(\cdot, \alpha^*)$. If $J^{\lambda_X}(t, \alpha^*) > \overline{J}^{\lambda_X}(t, \alpha^*)$ for some t, $\overline{J}^{\lambda_X}(\cdot, \alpha^*)$ is linear and therefore $\overline{\varphi}^{\lambda_X}(\cdot, \alpha^*)$ is constant in a neighborhood of t (see Figure 2). Therefore, the first pooling property characterizes the minimum pooling area in an optimal allocation rule and the second pooling property characterizes the maximum pooling area in an optimal allocation rule. Intuitively, more pooling leads to less steep allocation rules and weaker incentives to acquire information. The optimal amount of pooling is chosen so that Q satisfies (IA′) with equality.

The main difficulty is to determine the optimal multiplier λ_X. In order to do so, we first define the "steepest" allocation rule Q^+ and the "least steep" allocation rule Q^- satisfying conditions (ii) and (iii) given above. Define $Q^+(\cdot, \lambda_X)$ as follows. If $J^{\lambda_X}(H(w|\alpha^*), \alpha^*) > \overline{J}^{\lambda_X}(H(w|\alpha^*), \alpha^*)$ for $w \in (\underline{w}, \overline{w})$ and let $(\underline{w}, \overline{w})$ be chosen maximally, then let

$$Q^+(w, \lambda_X) := \frac{\frac{1}{n}[H(\overline{w}|\alpha^*)^n - H(\underline{w}|\alpha^*)^n]}{H(\overline{w}|\alpha^*) - H(\underline{w}|\alpha^*)}, \ \forall w \in (\overline{w}, \underline{w}).$$

Otherwise, let $Q^+(w, \lambda^X) := H(w|\alpha^*)^{n-1}$. Define $Q^-(\cdot, \lambda_X)$ as follows. If $\overline{\varphi}^{\lambda_X}(H(\cdot|\alpha^*), \alpha^*)$ is constant on $(\underline{w}, \overline{w})$ with $\underline{w} < \overline{w}$ and let $(\underline{w}, \overline{w})$ be chosen maximally, then let

$$Q^-(w, \lambda_X) := \frac{\frac{1}{n}[H(\overline{w}|\alpha^*)^n - H(\underline{w}|\alpha^*)^n]}{H(\overline{w}|\alpha^*) - H(\underline{w}|\alpha^*)}, \ \forall w \in (\overline{w}, \underline{w}).$$

Otherwise, let $Q^-(w, \lambda_X) := H(w|\alpha^*)^{n-1}$. Clearly, both Q^+ and Q^- are non-decreasing and implementable and satisfy conditions (ii) and (iii). As we have argued above, among all non-decreasing implementable Q's satisfying conditions (ii) and (iii), Q^+ contains the minimum pooling area and Q^- contains the maximum pooling area. We prove in Lemma 23 in Appendix A.2.3 that, for all non-decreasing implementable Q's satisfying conditions (ii) and (iii), Q^+ gives the agents' highest marginal benefit from acquiring information and Q^- gives them the lowest:

$$\mathbb{E}\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}Q^+(w, \lambda_X)\middle|\alpha^*\right] \geq \mathbb{E}\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}Q(w)\middle|\alpha^*\right] \geq \mathbb{E}\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}Q^-(w, \lambda_X)\middle|\alpha^*\right].$$

Hence, there exists a non-decreasing implementable Q satisfying conditions (i)-(iii) if and only if λ_X is such that

$$\mathbb{E}\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}Q^+(w, \lambda_X)\middle|\alpha^*\right] \geq C'(\alpha^*) \geq \mathbb{E}\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)}Q^-(w, \lambda_X)\middle|\alpha^*\right]. \quad (12)$$

The following Lemma 6 proves that such a λ_X exists and is unique, and its proof can be found in Appendix A.2.4.

Lemma 6 *Suppose that the first-order approach is valid and the information structures are supermodular ordered. Suppose, in addition, that α^* is such that (9) is violated. There exists a unique $\lambda_X > 0$ such that inequality (12) holds.*

The main result of this section is the following Theorem 2, which demonstrates that the unique λ_X given in Lemma 6 is indeed optimal and the allocation rules we have derived above solve ($\mathcal{P}'$-α^*):

Theorem 2 *Suppose that the first-order approach is valid and the information structures are supermodular ordered. Suppose, in addition, that α^* is such that (9) is violated. Let $\lambda_X > 0$ be such that inequality (12) holds and Q be a non-decreasing implementable allocation rule. Then Q solves ($\mathcal{P}'$-α^*) if and only if $Y(w(0,\alpha^*)) = 0$, and Q satisfies (IA$'$) with equality and the two pooling properties.*

Theorem 2 describes optimal mechanisms in terms of the reduced form but not how to implement them. Similar to the regular case, there exists an ex ante efficient interim allocation rule such that the corresponding direct mechanism can be implemented by restricting the set of allowable bids in standard auctions. But the optimal set of allowable bids can be very complex due to the complex pooling areas. We omit the details here.

We interpret the optimal pooling areas as follows: Optimally, pooling occurs where $\varphi^{\lambda_X}(H(\cdot|\alpha^*), \alpha^*)$ is not strictly increasing, i.e.,

$$\frac{w_{s,\alpha}(s,\alpha^*)}{w_s(s,\alpha^*)} = \frac{\partial}{\partial w}\left[-\frac{H_\alpha(w(s,\alpha^*)|\alpha^*)}{h(w(s,\alpha^*)|\alpha^*)}\right] \geq \frac{1}{\lambda_X}$$

Recall that if an information structure is more informative, then $w(s,\alpha)$ changes more dramatically as s changes, i.e., $w_s(s,\alpha)$ is larger. Hence, one can interpret $w_s(s,\alpha)$ as a local measure of the information structures' accuracy around s. Then, $w_{s,\alpha}(s,\alpha)/w_s(s,\alpha)$ is the percentage change of the information structures' accuracy around s as α increases. Intuitively, the most effective way to discourage agents from acquiring too much information is to introduce randomization to areas where the information structures' accuracy increases most rapidly. If the information structures are uniformly supermodular ordered, $w_{s,\alpha}(s,\alpha)/w_s(s,\alpha)$ is a constant. In other words, when α increases, the information structure becomes more informative *uniformly* over $[0,1]$. As a result, the choice of pooling areas is not important (as it is the case in Proposition 1).[14]

[14]When the information structures are uniformly supermodular ordered, the minimum pooling area is empty and the maximum pooling area is the support of the posterior means W in equilibrium.

Though the result is intuitive, the proof of Theorem 2 is difficult because of the presence of both the non-standard constraint (IA′) and the monotonicity constraint (MON). In this paper, we use the following approach first proposed by Reid (1968) and later introduced into the mechanism design literature by Mierendorff (2009). We first solve ($\mathcal{P}'$-α^*) under an additional restriction, that Q is Lipschitz continuous with global Lipschitz constant K:

$$|Q(w) - Q(w')| \leq K|w - w'|, \ \forall w, w' \in W.$$

Denote the modified maximization problem by ($\mathcal{P}^K$-α^*). We show that the optimal solutions of ($\mathcal{P}^K$-α^*) converge to that of ($\mathcal{P}'$-α^*) as $K \to \infty$. Then we can obtain a characterization of the optimal solutions of ($\mathcal{P}'$-α^*) in the limit. The formal analysis can be found in Appendix A.2.

The proof is not a straightforward modification of Mierendorff (2009). Let Q and Q^K denote the optimal solutions to ($\mathcal{P}^K$-α^*) and ($\mathcal{P}'$-α^*), respectively. In Mierendorff (2009), Q is discontinuous at exactly one known point, and, for K sufficiently large, the slope of Q^K is equal to K only in a neighborhood around the discontinuity point. In this paper, however, Q could be discontinuous at most countably many times, at unknown points. If Q is discontinuous at w, it is possible that every neighborhood of w contains another discontinuity point. Hence, it is non-trivial to characterize Q as the limit of Q^K.

We conclude this subsection by briefly discussing why we cannot apply control theoy directly. In the published version of Mierendorff (2009), Mierendorff (2016) does not use the approach described above, and directly appeals to Theorems 7 and 8 in Seierstad and Sydsæter (1987). However, the problem considered here, ($\mathcal{P}'$-α^*), is more complex for the following two reasons. First, as discussed above, in Mierendorff (2016), state variable Q is discontinuous at exactly one point, while in ($\mathcal{P}'$-α^*), Q could be discontinuous at most countably many points. Second, the problem in Mierendorff (2016) can be written as a control problem without restrictions on the state variables, while ($\mathcal{P}'$-α^*) contains *pure state constraints* (constraints in which control variables do not appear). To the best of my knowledge, no existing theorem can be applied to provide necessary and sufficient conditions for the optimal solutions of ($\mathcal{P}'$-α^*).

3.3 Optimal α^*

Given the optimal solutions of ($\mathcal{P}'$-α^*), we can now study the socially optimal information choice. For each α, let $\pi^s(\alpha) := (1-\gamma)V(\alpha) - C(\alpha)$ denote the maximum average expected social surplus. Let $\alpha^* \in \text{argmax}_{\alpha \in [\underline{\alpha}, \overline{\alpha}]} \pi^s(\alpha)$ denote a socially optimal information structure. We begin the analysis by showing that it is not optimal for the social planner to encourage the agents to acquire more accurate information than they would when facing the ex post efficient mechanism. Then we provide a sharper characterization of the socially optimal α^* in the special case when the information structures are uniformly supermodular ordered.

Recall that the ex post efficient mechanism solves ($\mathcal{P}'$-α^*) if α^* is such that the following inequality holds:

$$\mathbb{E}_w \left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)} H(w|\alpha^*)^{n-1} \middle| \alpha^* \right] \leq C'(\alpha^*), \tag{9}$$

Let

$$\hat{\alpha} := \inf \left\{ \alpha \in [\underline{\alpha}, \overline{\alpha}] | (9) \text{ holds for } \alpha \right\}, \tag{13}$$

which is independent of γ, and satisfies $\hat{\alpha} > \underline{\alpha}$. First, we argue that the socially optimal α^* is bounded above by $\hat{\alpha}$. Lemma 27 in Appendix A.3 proves that if the second-order condition of the agents' optimization problem is satisfied,

$$\int_{w(0,\alpha)}^{w(1,\alpha)} -H_\alpha(w|\alpha) H(w|\alpha)^{n-1} \mathrm{d}w - C'(\alpha) \text{ is strictly decreasing in } \alpha. \tag{14}$$

This implies that inequality (9) holds strictly for all $\alpha > \hat{\alpha}$. Hence, for all $\alpha > \hat{\alpha}$, the optimal solution to ($\mathcal{P}'$-α^*) is the ex post efficient allocation rule: $Q(w) = H(w|\alpha)^{n-1}$ for all w. In this case, the average expected social surplus is

$$\pi^s(\alpha) = (1-\gamma) \int_{w(0,\alpha)}^{w(1,\alpha)} w H(w|\alpha)^{n-1} h(w|\alpha) \mathrm{d}w - C(\alpha).$$

Taking derivative with respect to α gives

$$\pi^{s\prime}(\alpha) = (1-\gamma) \int_{w(0,\alpha)}^{w(1,\alpha)} -H_\alpha(w|\alpha) H(w|\alpha)^{n-1} \mathrm{d}w - C'(\alpha).$$

Because of (14) and the fact that $C'(\alpha)$ is strictly increasing, $\pi^{s\prime}(\alpha)$ is strictly decreasing. By construction, $\pi^{s\prime}(\hat{\alpha}) = -\gamma C'(\hat{\alpha}) \leq 0$. Hence, $\pi^{s\prime}(\alpha) < 0$ for all $\alpha > \hat{\alpha}$ and a socially optimal information choice must satisfy $\alpha^* \leq \hat{\alpha}$. Note that, by construction, if the ex post efficient mechanism is used, all agents choose $\hat{\alpha}$ is the unique symmetric equilibrium. In other words, it is not optimal for the social planner to encourage the agents to acquire more information than they would under the ex post efficient mechanism. This result is consistent with Bergemann and Välimäki (2002) and Bergemann et al. (2009). The fact that $\alpha^* \leq \hat{\alpha}$ also implies that (IA$'$) always holds with equality when α^* is chosen optimally.[15] Hence, it

[15]If $\alpha^* = \hat{\alpha}$, the ex post efficient mechanism is optimal and (IA$'$) holds with equality by the definition of $\hat{\alpha}$. Suppose that $\alpha^* < \hat{\alpha}$. Suppose, to the contrary, that (IA$'$) holds with strict inequality. Then the ex post efficient mechanism is optimal. However, because inequality (9) is violated when $\alpha^* < \hat{\alpha}$, (IA$'$) is violated,

is sufficient to consider the one-sided first-order condition.

Second, the upper-bound $\hat{\alpha}$ converges to $\underline{\alpha}$ as the number of agents n increases because the agents' incentives to acquire information diminish as n increases. To see this, consider the ex post efficient allocation rule: $Q(w) = H(w|\alpha)^{n-1}$ for all w. In this case, the agents' marginal benefit from acquiring information (measured by the first term in (14)) is of order $1/n$.[16] It is nature to conjecture that $\hat{\alpha}$ is decreasing in n. For example, this is the case when the agents can perform linear experiments (see Example 4). But we remark that in general the agents' marginal benefits from acquiring information may not be monotonic in n. This is because when n increases, the interim allocation rule becomes less steep for low posterior means but steeper for high posterior means. Therefore, the aggregate effect is ambiguous.

Finally, we argue that when the agents have private values ($\gamma = 0$), the ex post efficient mechanism is also ex ante efficient. To see this, note that the average expected social surplus is bounded above:

$$\pi^s(\alpha) \leq \int_{w(0,\alpha)}^{w(1,\alpha)} wH(w|\alpha)^{n-1}h(w|\alpha)\mathrm{d}w - C(\alpha),$$

where the right-hand side is maximized at $\hat{\alpha}$. If $\gamma = 0$, this upper-bound is also achieved by the left-hand side at $\hat{\alpha}$. In the special case when agents can perform linear experiments, we show a even stronger result holds: the ex post efficient mechanism is also ex ante efficient when the positive dependence is weak (i.e. γ is small) (see Example 4). This result is also consistent with Bergemann and Välimäki (2002) and Bergemann et al. (2009). Specifically, the latter shows that, in the ex post efficient mechanism, the difference between the equilibrium level of information and socially optimal level of information diminishes as the positive dependence weakens.

These results are summarized in the following proposition:

Proposition 2 *Suppose that the second-order condition of the agents' optimization problem is satisfied and the information structures are supermodular ordered. The socially optimal information choice α^* is bounded above by $\hat{\alpha}$, where $\hat{\alpha}$ is such that (9) holds with equality and $\lim_{n\to\infty}\hat{\alpha}(n) = \underline{\alpha}$. Furthermore, if $\gamma = 0$, $\alpha^* = \hat{\alpha}$.*

For $\alpha \leq \hat{\alpha}$, the analysis in Section 3.2 shows that the maximum average expected social

which is a contradiction. Hence, (IA′) holds with equality.

[16] The first term in (14) is of order $1/n$:

$$\int_{w(0,\alpha)}^{w(1,\alpha)} -\frac{H_\alpha(w|\alpha)}{h(w|\alpha)}H(w|\alpha)^{n-1}h(w|\alpha)\mathrm{d}w < \int_{w(0,\alpha)}^{w(1,\alpha)} -\frac{H_\alpha(\overline{\theta}|\alpha)}{h(\overline{\theta}|\alpha)}H(w|\alpha)^{n-1}h(w|\alpha)\mathrm{d}w = \frac{1}{n}\left(-\frac{H_\alpha(\overline{\theta}|\alpha)}{h(\overline{\theta}|\alpha)}\right).$$

surplus is given by

$$\pi^s(\alpha) = (1-\gamma)\int_{w(0,\alpha)}^{w(1,\alpha)} \overline{\varphi}^{\lambda_X}(H(w|\alpha),\alpha)H(w|\alpha)^{n-1}h(w|\alpha)\mathrm{d}w + (1-\gamma)\lambda_X C'(\alpha) - C(\alpha),$$

where the optimal λ_X depends on α in a complex way. Therefore, it is hard to characterize the optimal α^* in the general case. To obtain a sharper characterization of the socially optimal information structure α^*, we assume in the rest of this section that the information structures are uniformly supermodular ordered. In this case, the average expected social surplus can be written as

$$\pi^s(\alpha) = (1-\gamma)\left[\frac{\mu}{n} + b(\alpha)C'(\alpha)\right] - C(\alpha),\ \forall \alpha \in [\underline{\alpha}, \hat{\alpha}]. \tag{15}$$

Assume that $\max_{\alpha\in[\underline{\alpha},\overline{\alpha}]}\left[(1-\gamma)b(\alpha)C'(\alpha) - C(\alpha)\right]$ has a unique solution and denote it by α°. Clearly, α° is independent of n. The socially optimal information choice is given by $\alpha^*(n,\gamma) = \min\{\alpha^\circ(\gamma), \hat{\alpha}(n)\}$. Taking derivatives of $\pi^s(\alpha)$ with respect to α gives

$$\pi^{s\prime}(\alpha) = \left[(1-\gamma)b'(\alpha) - 1\right]C'(\alpha) + (1-\gamma)b(\alpha)C''(\alpha),\ \forall \alpha \in [\underline{\alpha}, \hat{\alpha}].$$

If agents have pure common values ($\gamma = 1$), $\pi^{s\prime}(\alpha) = -C'(\alpha) \leq 0$ and the inequality holds strictly for all $\alpha > \underline{\alpha}$. Hence, it is socially optimal for the agents not to acquire information: $\alpha^* = \alpha^\circ = \underline{\alpha}$. This is intuitive because information has no social value in the case of pure common values. If $\gamma < 1$, $\pi^{s\prime}(\underline{\alpha}) = (1-\gamma)b(\underline{\alpha})C''(\underline{\alpha}) > 0$. Therefore, $\alpha^\circ > \underline{\alpha}$. Recall that $\hat{\alpha} > \underline{\alpha}$ and $\lim_{n\to\infty}\hat{\alpha} = \underline{\alpha}$. Hence, when n is sufficiently large, $\alpha^*(\gamma, n) = \hat{\alpha}(n)$ and the ex post efficient mechanism is also ex ante efficient. Intuitively, when there is a large number of agents, an individual agent's incentive to acquire information is already small because of the fierce competition, and the social planner need not further discourage them from acquiring information by distorting the allocation rule. These results are summarized by the following proposition:

Proposition 3 *Suppose that the second-order condition of the agents' optimization problem is satisfied and the information structures are uniformly supermodular ordered. The socially optimal information structure is given by* $\alpha^*(n,\gamma) = \min\{\alpha^\circ(\gamma), \hat{\alpha}(n)\}$*, where* $\lim_{\gamma\to 1}\alpha^\circ = \underline{\alpha}$ *and* $\lim_{n\to\infty}\hat{\alpha} = \underline{\alpha}$.

We conclude this section by illustrating the results in Propositions 2 and 3 in the following example in which agents can perform linear experiments.

Example 4 (Linear experiments) *Consider the information structures in Example 1. Assume that* $F(\theta) = \theta$ *with support* $[0,1]$ *and the cost function (used in Persico (2000)) is of the form*

$$C(\alpha) = K\left(\alpha - \underline{\alpha}\right)^2, \ \forall \alpha \in [\underline{\alpha}, 1],$$

where $0 < \underline{\alpha} < 1$ *and* $K \geq 1/8\underline{\alpha}$. *Then, as we demonstrate in Appendix A.3, the first-order approach is valid. In this case* $\hat{\alpha}$ *is such that*

$$2K(\hat{\alpha} - \underline{\alpha}) = \frac{n-1}{2n(n+1)}.$$

The left-hand side of the above equation is the marginal cost of information, which is strictly increasing in $\hat{\alpha}$. *The right-hand side is the agents' marginal benefit from acquiring information, which is strictly decreasing in* n *for* $n \geq 2$ *and converges to* 0 *as* n *goes to infinity. Hence,* $\hat{\alpha}$ *is strictly decreasing in* n *and goes to* $\underline{\alpha}$ *as* n *goes to infinity. Finally,*

$$\pi^{s\prime}(\alpha) = 2K\left[\gamma\underline{\alpha} - (2\gamma - 1)\alpha\right].$$

If $\gamma \leq \frac{1}{2}$, $\pi^{s\prime}(\alpha) \geq 0$ *for all* α *and therefore* $\alpha^* = \hat{\alpha}$. *If* $\gamma > \frac{1}{2}$, $\pi^{s\prime}(\alpha)$ *is strictly decreasing in* α *and therefore*

$$\alpha^* = \min\left\{\frac{\gamma\underline{\alpha}}{2\gamma - 1}, \hat{\alpha}\right\}.$$

Thus, if γ *is sufficiently small or* n *is sufficiently large,* $\alpha^* = \hat{\alpha}$, *and the ex post efficient mechanism is also ex ante efficient. If* γ *is sufficiently large or* n *is sufficiently small, the optimal* α^* *is strictly decreasing in* γ, *and goes to* $\underline{\alpha}$ *as* γ *increases to* 1.

4 Efficient asymmetric mechanisms

In this section, we relax the restriction that mechanisms treat all agents symmetrically and all agents choose the same information structure in equilibrium. First, we show that the result in Theorem 1 can be generalized (i.e. an ex ante efficient mechanism never withholds the object). Second, when the information structures are uniformly supermodular ordered, we provide conditions under which the socially optimal information choices are the same for all agents and therefore a symmetric ex ante efficient mechanism exists. When these conditions are not satisfied, we provide an example in which an asymmetric mechanism generates higher expected social surplus than the optimal symmetric mechanism does.

We follow the same method used in Section 3 to solve the social planner's problem. First, by the standard argument, (IC) holds if and only if

$$Q_i(w) \text{ is non-decreasing in } w, \tag{MON}$$

and $U_i(w)$ is absolutely continuous and satisfies the envelope condition. Next we consider the relaxed problem of the social planner by replacing (IA) by the one-sided first-order conditions

$$\mathbb{E}_w\left[-\frac{H_{\alpha_i}(w|\alpha_i^*)}{h(w|\alpha_i^*)}Q_i(w)\middle|\,\alpha_i=\alpha_i^*\right]\le C'(\alpha_i^*),\ \forall i. \tag{IA$'$}$$

and focus on reduced form auctions. Later on, we show in the appendix that if $\boldsymbol{\alpha}^*$ is chosen optimally, (IA$'$) holds with equality for all i. Let $\boldsymbol{Q} := (Q_1, \ldots, Q_n)$, where $Q_i : [\underline{\theta}, \overline{\theta}] \to [0, 1]$ is non-decreasing for all i. Let $W_i := [w(0, \alpha_i^*), w(1, \alpha_i^*)]$ denote the support of w_i for all i. By Theorem 3 in Mierendorff (2011), $\boldsymbol{Q}$ is implementable if and only if it satisfies

$$\sum_{i=1}^{n}\int_{w_i}^{w(1,\alpha_i^*)} Q_i(z_i)\mathrm{d}H(z_i|\alpha_i^*) \le 1-\prod_{i=1}^{n}H(w_i|\alpha_i^*),\ \forall \boldsymbol{w}\in\prod_{i=1}^{n}W_i. \tag{F$'$}$$

Thus, given (MON), we can replace (F) by (F$'$). Finally, as in the symmetric case, an ex ante efficient mechanism never withholds the object:

Theorem 3 *Suppose that the information structures are supermodular ordered and $\boldsymbol{\alpha}^*$ and $(\boldsymbol{q}, \boldsymbol{t})$ solve the relaxed problem of the social planner. Then*

$$\sum_i q_i(\boldsymbol{w}) = 1 \text{ for almost all } \boldsymbol{w}\in\prod_{i=1}^{n}[w(0,\alpha_i^*), w(1,\alpha_i^*)]. \tag{16}$$

Using Theorem 3 and the law of iterated expectations, the social planner's objective function can be rewritten as a functional of $\boldsymbol{Q}$:

$$\mathbb{E}_{\boldsymbol{w}}\left[\sum_i(w_i+\gamma\sum_{j\ne i}w_j)q_i(\boldsymbol{w})\middle|\,\alpha_i=\alpha_i^*\ \forall i\right]=\sum_i\mathbb{E}_{w_i}\left[(1-\gamma)w_iQ_i(w_i)|\,\alpha_i=\alpha_i^*\right]+n\gamma\mu.$$

Because the second term, $n\gamma\mu$, is a constant, we ignore it from here on. To summarize, the social planner's relaxed problem, denoted by ($\mathcal{P}'$), becomes:

$$\max_{\boldsymbol{\alpha}^*,\boldsymbol{Q}}\mathbb{E}_{\boldsymbol{w}}\left[\sum_i(1-\gamma)w_iQ_i(w_i)\middle|\,\alpha_i=\alpha_i^*\ \forall i\right]-\sum_i C(\alpha_i^*), \tag{$\mathcal{P}'$}$$

subject to

$$\sum_{i=1}^{n}\int_{w_i}^{w(1,\alpha_i^*)} Q_i(z_i)\mathrm{d}H(z_i|\alpha_i^*) \le 1-\prod_{i=1}^{n} H(w_i|\alpha_i^*),\ \forall \boldsymbol{w}\in\prod_{i=1}^{n} W_i, \tag{F$'$}$$

$$Q_i(w_i) \text{ is non-decreasing in } w_i,\ \forall i, \tag{MON}$$

$$\mathbb{E}_{w_i}\left[-\frac{H_{\alpha_i}(w_i|\alpha_i^*)}{h(w_i|\alpha_i^*)}Q_i(w_i)\middle|\, \alpha_i=\alpha_i^*\right] \le C'(\alpha_i^*),\ \forall i. \tag{IA$'$}$$

In general, it is hard to solve ($\mathcal{P}'$) because both (F$'$) and (IA$'$) are much more complex here than they are under the symmetric restriction. In the rest of this section, we focus on the special case in which the information structures are uniformly supermodular ordered.

Analogous to Proposition 2, Proposition 4 below proves that it is not optimal for the principal to encourage all the agents to acquire more information than they would under the ex post efficient mechanism. In particular, if all agents choose the same information structure (i.e. $\alpha_1^*=\cdots=\alpha_n^*$), (17) implies that $\alpha_i^*\le\hat{\alpha}$, where $\hat{\alpha}$ is defined by (13).

Proposition 4 *Suppose that the second-order condition of the agents' optimization problem is satisfied and the information structures are uniformly supermodular ordered. Let $\boldsymbol{\alpha}^*$ be a socially optimal information choices. Then $\boldsymbol{\alpha}^*$ satisfies the following condition: there exists agent i such that*

$$\mathbb{E}_{w_i}\left[-\frac{H_{\alpha_i}(w_i|\alpha_i^*)}{h(w_i|\alpha_i^*)}\prod_{j\neq i}H(w_i|\alpha_j^*)\middle|\, \boldsymbol{\alpha}^*\right] \ge C'(\alpha_i^*). \tag{17}$$

Furthermore, for any $\boldsymbol{\alpha}^*$ that satisfies the above condition, the expected social surplus is given by

$$\pi^s(\boldsymbol{\alpha}^*) = (1-\gamma)\left[\mu+\sum_{i=1}^{n}b(\alpha_i^*)C'(\alpha_i^*)\right]-\sum_i C(\alpha_i^*). \tag{18}$$

Proposition 4 also gives an expression of the expected social surplus. Compare (18) with (15). It follows immediately that if $\alpha^\circ\le\hat{\alpha}$, where α° is the unique solution to $\max_{\alpha\in[\underline{\alpha},\overline{\alpha}]}(1-\gamma)b(\alpha)C'(\alpha)-C(\alpha)$ and $\hat{\alpha}$ is defined by (13), the socially optimal information choices are the same for all agents and a symmetric ex ante efficient mechanism exists:

Proposition 5 *Suppose that the second-order condition of the agents' optimization problem is satisfied and the information structures are uniformly supermodular ordered. Suppose, in addition, that $\alpha^\circ\le\hat{\alpha}$. Then the socially optimal information choices are the same for all agents and a symmetric ex ante efficient mechanism exists.*

The analysis in Section 3 suggests that the condition $\alpha^\circ\le\hat{\alpha}$ is more likely to be satisfied when the number of agents is small or the level of interdependence is high. If $\alpha^\circ>\hat{\alpha}$, we give an example in which an asymmetric mechanism generates higher expected social surplus that the optimal symmetric mechanism does.

Example 5 (Linear experiments) *Consider Example 4. Let* $n = 2$, $\underline{\alpha} = 1/2$ *and* $K = 3/8 \geq 1/8\underline{\alpha}$. *Then, as we demonstrate in Appendix A.3, the first-order approach is valid. By Proposition 4, a socially optimal information choice* $\boldsymbol{\alpha}$ *must be such that*

$$\alpha_1 \leq \frac{9\alpha_2}{18\alpha_2 - 2} \text{ or } \alpha_2 \leq \frac{9\alpha_1}{18\alpha_1 - 2}. \tag{19}$$

When $\boldsymbol{\alpha}$ *satisfies (19), the expected social surplus is given by*

$$\pi^s(\boldsymbol{\alpha}) = (1-\gamma)\left[\frac{1}{2} + \frac{3}{4}\alpha_1\left(\alpha_1 - \frac{1}{2}\right) + \frac{3}{4}\alpha_2\left(\alpha_2 - \frac{1}{2}\right)\right] - \frac{3}{8}\left(\alpha_1 - \frac{1}{2}\right)^2 - \frac{3}{8}\left(\alpha_2 - \frac{1}{2}\right)^2.$$

In this case, $\hat{\alpha}$ *is such that*

$$\frac{3}{4}\left(\hat{\alpha} - \frac{1}{2}\right) = \frac{1}{12} \text{ or } \hat{\alpha} = \frac{11}{18}.$$

Furthermore, $(1-\alpha)b(\alpha)C'(\alpha) - C(\alpha) = (1-\gamma)\frac{3}{4}\alpha(\alpha - \frac{1}{2}) - \frac{3}{8}(\alpha - \frac{1}{2})^2$ *has a unique maximizer on* $[\frac{1}{2}, 1]$. *If* $\gamma \leq \frac{1}{2}$, $\alpha^\circ = 1$; *and if* $\gamma > \frac{1}{2}$, $\alpha^\circ = \gamma/(4\gamma - 2)$. *By Proposition 5, if* $\alpha^\circ \leq \hat{\alpha}$ *or* $\gamma \geq 11/13$, *the socially optimal information choices are the same for all agents.*

Assume that $\gamma < 11/13$. *In this case, the optimal symmetric mechanism is ex ante efficient, and induces the following symmetric equilibrium:* $\alpha_1 = \alpha_2 = \hat{\alpha}$. *In this case, the expected social surplus is given by*

$$\pi^s(\hat{\alpha}, \hat{\alpha}) = \frac{192 - 195\gamma}{324} \approx 0.59 - 0.60\gamma.$$

Consider the following asymmetric mechanism

$$q_1(w_1, w_2) = \begin{cases} 0 & \text{if } \max\left\{\frac{7}{32}, w_1\right\} < w_2 \\ 1 & \text{if } \min\left\{w_1, \frac{25}{32}\right\} > w_2 \end{cases},$$

and $q_2(w_1, w_2) = 1 - q_1(w_1, w_2)$ *for all* $(w_1, w_2) \in [0,1]^2$. *Given this mechanism, the following information choices constitute an equilibrium:* $\alpha_1^* = 9/16$ *and* $\alpha_2^* = 1$. *Given* $\alpha_2^* = 1$, *the interim allocation rule of agent 1 is given by*

$$Q_1(w_1) = \begin{cases} \frac{7}{32} & \text{if } w_1 \in \left[0, \frac{7}{32}\right] \\ w_1 & \text{if } w_1 \in \left[\frac{7}{32}, \frac{25}{32}\right] \\ \frac{25}{32} & \text{if } w_1 \in \left[\frac{25}{32}, 1\right] \end{cases}.$$

It is easy to verify that

$$\int_{w(0,\alpha_1)}^{w(1,\alpha_1)} -H_\alpha(w_1,\alpha_1)Q_1(w_1)\mathrm{d}w_1 - C'(\alpha_1) = -\frac{8\alpha_1}{12} + \frac{3}{8}.$$

Hence, it is optimal for agent 1 *to choose* $\alpha_1^* = 9/16$*. Similarly, given* $\alpha_1^* = 9/16$*, the interim allocation rule of agent* 2 *is given by*

$$Q_2(w_2) = \frac{16}{9}w_2 - \frac{7}{18}, \forall w_2 \in [0,1].$$

It is easy to verify that

$$\int_{w(0,\alpha_2)}^{w(1,\alpha_2)} -H_\alpha(w_2,\alpha_1)Q_2(w_2)\mathrm{d}w_2 - C'(\alpha_2) = -\frac{65\alpha_2}{108} + \frac{3}{8} > 0,\ \forall \alpha_2 \in \left[\frac{1}{2},1\right].$$

Hence, it is optimal for agent 2 *to choose* $\alpha_2^* = 1$*. In this case, the expected social surplus is given by*

$$\pi^s(\alpha_1^*,\alpha_2^*) = \frac{728-923\gamma}{1024} \approx 0.71 - 0.90\gamma.$$

Clearly, if $\gamma < 0.4$*, this asymmetric mechanism generates strictly higher expected social surplus that the optimal symmetric mechanism does.*

To obtain some intuition behind this result, fix $\boldsymbol{\alpha}^*$ and consider the ex post efficient mechanism. If agent j acquires more information, the change of agent i's marginal benefit from acquiring information is negative:

$$\begin{aligned}&\mathbb{E}_{w_i}\left[-\frac{H_{\alpha_i}(w_i|\alpha_i^*)}{h(w_i|\alpha_i^*)}H_{\alpha_j}(w_i|\alpha_j^*)\prod_{k\neq i,j}H(w_i|\alpha_k^*)\middle|\boldsymbol{\alpha}^*\right]\\ =&\mathbb{E}_{w_i}\left[-(w_i-\mu)^2 h(w_i|\alpha_j^*)\prod_{k\neq i,j}H(w_i|\alpha_k^*)\middle|\boldsymbol{\alpha}^*\right] < 0.\end{aligned}$$

This result suggests that the agents' information acquisition decisions are strategic substitutes. Hence, an asymmetric mechanism could possibly outperform the optimal symmetric mechanism because, by encouraging one agent to acquire information, it reduces other agents' incentives to do so. This is exactly the case in Example 5. This finding is also consistent with Bergemann et al. (2009), who model information acquisition as a binary decision and show that an agent's value of being informed is decreasing in the number of informed agents.

5 Conclusion

Consider the sale of a single object when agents have positively interdependent values and can covertly acquire information at some cost before participating in a mechanism. Maskin (1992), Bergemann and Välimäki (2002) and Bergemann et al. (2009), among others, have shown that using the ex post efficient mechanism will give agents too strong incentives to acquire information. In other words, there is a conflict between the provision of ex ante efficient incentives to acquire information and the ex post efficient use of information. In this paper, we show how to design ex ante efficient mechanisms to balance the two trade-offs. We conclude by discussing potential directions for future work.

In this paper, we assume that all agents simultaneously acquire information before participating in the mechanism. One important direction for future research is to allow for the possibility of sequential information acquisitions. It is likely that the efficiency can be improved if agents are asked to acquire information in turn, and one's information acquisition decision can depend on the signals received by those who take actions earlier. It is also interesting to consider the impact of initial private information. In this paper we have only considered static mechanisms in which agents only report their private information once. In general, one can consider a dynamic mechanism in which agents report their private information both before and after acquiring information.

One standard assumption in the auction literature we maintain here is that the payment by an agent depends only on his report and not on his realized valuation. Although it is true that in some cases the value of the auctioned asset is subjective and cannot be verified, in many important applications such as the oil and gas lease auctions and the sales of a financial or business asset, the benefits to the agents are contractible. For example, in U.S. auctions for offshore oil and gas leases, the winner's payment to the government is a bonus plus a fraction of revenues from any oil and/or gas extracted. When selling a company to an acquirer or soliciting venture capital, equity and other securities are commonly used. Another important direction for future research is to consider mechanisms with contingent payments.

A Omitted proofs in sections 2 and 3

Before proceeding to the proofs, we first define *symmetric mechanisms* formally. Let $\sigma_{i,j} : W^n \to W^n$ denote the function that interchanges the ith and the jth coordinates, i.e.,

$$\sigma_{i,j}(w_1, \ldots, w_n) = (w_1, \ldots, w_{i-1}, w_j, w_{i+1}, \ldots, w_{j-1}, w_i, w_{j+1}, \ldots, w_n),\ \forall (w_1, \ldots, w_n).$$

We say that an allocation rule $\boldsymbol{q}$ is *symmetric* if q_1 is such that $q_1(\boldsymbol{w}) = q_1(\sigma_{i,j}(\boldsymbol{w}))$ for all $i, j \neq 1$ and all $\boldsymbol{w}$, $q_i(\boldsymbol{w}) = q_1(\sigma_{1,i}(\boldsymbol{w}))$ for all i all $\boldsymbol{w}$ and $\sum_i q_i(\boldsymbol{w}) \leq 1$ for all $\boldsymbol{w}$. We say that a mechanism $(\boldsymbol{q}, \boldsymbol{t})$ is *symmetric* if its allocation rule $\boldsymbol{q}$ is symmetric.

Proof of Lemma 1. By construction, $H(w(s,\alpha)|\alpha) = s$ for all $s \in [0,1]$ and $\alpha \in [\underline{\alpha}, \overline{\alpha}]$. Taking derivative of both sides of the equation with respect to α yields

$$h(w(s,\alpha)|\alpha)w_\alpha(s,\alpha) + H_\alpha(w(s,\alpha)|\alpha) = 0,$$

or equivalently,

$$-\frac{H_\alpha(w(s,\alpha)|\alpha)}{h(w(s,\alpha)|\alpha)} = w_\alpha(s,\alpha). \tag{20}$$

Because the information structures are supermodular ordered, $-H_\alpha(w|\alpha)/h(w|\alpha)$ is strictly increasing in w. Furthermore, $w(s,\alpha)$ is strictly increasing in s. Hence, $w_\alpha(s,\alpha)$ is strictly increasing in s. Thus, for all $s, s' \in (0,1)$, $s' > s$ and $\alpha' > \alpha''$ we have

$$\begin{aligned} w(s',\alpha') - w(s',\alpha'') &= \int_{\alpha''}^{\alpha'} w_\alpha(s',\alpha)\mathrm{d}\alpha \\ &> \int_{\alpha''}^{\alpha'} w_\alpha(s,\alpha)\mathrm{d}\alpha \\ &= w(s,\alpha') - w(s,\alpha''). \end{aligned}$$

That is, $w(\cdot,\cdot)$ is strictly supermodular. ■

Derivation of equality (4). Supposing that agent i chooses α_i, his expected payoff is

$$\begin{aligned} &\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} U_i(w)\mathrm{d}H(w|\alpha_i) - C(\alpha_i) \\ =&\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} \left[U_i(\underline{\theta}) + \int_{\underline{\theta}}^{w} Q(\tilde{w})\mathrm{d}\tilde{w}\right]\mathrm{d}H(w|\alpha_i) - C(\alpha_i) \\ =&U_i(\underline{\theta}) + \int_{\underline{\theta}}^{w(1,\alpha_i)} Q(w)\mathrm{d}w - \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} H(w|\alpha_i)Q(w)\mathrm{d}w - C(\alpha_i) \\ =&U_i(w(0,\alpha_i)) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} [1 - H(w_i|\alpha_i)]\, Q(w_i)\mathrm{d}w_i - C(\alpha_i), \end{aligned}$$

where the first and the last equalities hold by the envelope condition (3) and the second equality holds by integration by parts. ■

We say a function $J(w)$ is *quasi-monotone* if $w' > w$ and $J(w) > 0$ imply $J(w') \geq 0$.[17] In other words, a quasi-monotone function $J(w)$ crosses the line $y \equiv 0$ at most once and from

[17]See Karamardian and Schaible (1990).

below, as w increases. The following result of Persico (2000) is useful for later proofs.

Lemma 7 (Lemma 1 in Persico (2000)) *Let (c,d) be an interval of the real line, J be a quasi-monotone function and Q be a non-decreasing function. Assume that for some measure h on $\mathbb{R}$ we have $\int_c^d J(w)\mathrm{d}h(w)=0$. Then $\int_c^d J(w)Q(w)\mathrm{d}h(w)\geq 0$.*

Proof of Lemma 2. Observe that a non-decreasing function is quasi-monotone. Because both Q and $-\frac{H_{\alpha_i}(w_i|\alpha_i)}{h(w_i|\alpha_i)}$ are non-decreasing on $[w(0,\alpha_i),w(1,\alpha_i)]$, it suffices to show that

$$\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} H_{\alpha_i}(w_i|\alpha_i)\mathrm{d}w_i = 0. \tag{21}$$

On the one hand, by integration by parts,

$$\begin{aligned}\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} H(w_i|\alpha_i)\mathrm{d}w_i &= w_i H(w_i|\alpha_i)\,\Big|_{w(0,\alpha_i)}^{w(1,\alpha_i)} - \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} w_i \mathrm{d}H(w_i|\alpha_i),\\ &= w(1,\alpha_i)-\mu.\end{aligned}$$

Taking derivative with respect to α_i yields

$$\frac{\partial}{\partial\alpha_i}\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} H(w_i|\alpha_i)\mathrm{d}w_i = w_{\alpha_i}(1,\alpha_i). \tag{22}$$

On the other hand, by the chain rule, we have

$$\frac{\partial}{\partial\alpha_i}\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} H(w_i|\alpha_i)\mathrm{d}w_i = w_{\alpha_i}(1,\alpha_i) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} H_{\alpha_i}(w_i|\alpha_i)\mathrm{d}w_i. \tag{23}$$

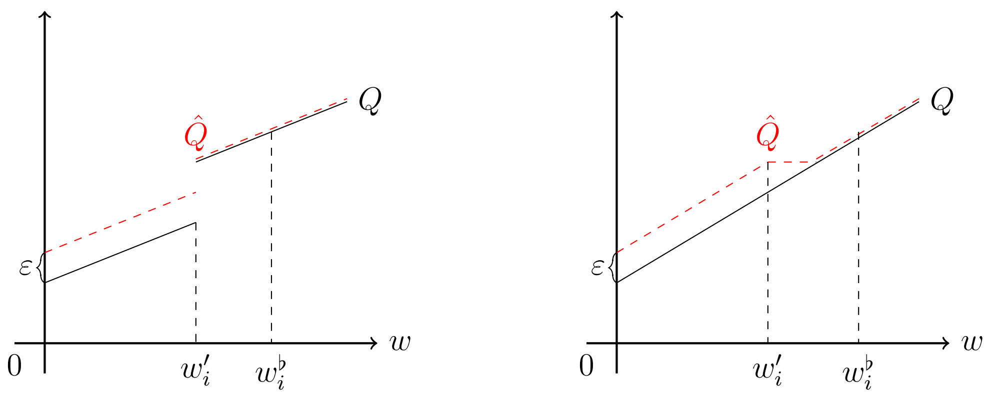
Comparing (22) and (23) proves (21). By Lemma 7, inequality (5) holds. If Q is constant, the equality holds by (21). ■

Proof of Lemma 3. Define $w^\flat := \sup\{w_i\,|Y(w_i')>0,\ \forall w(0,\alpha^*)\leq w_i'\leq w_i\}$. By the continuity of Y, we have $Y(w^\flat)=0$ and $w^\flat > w(0,\alpha^*)$. The proof is by construction. There are four cases to consider.

Case I: Suppose that there exists $w_i'\in(w(0,\alpha^*),w^\flat)$ such that Q is discontinuous at w_i'.

Let $Q(w_i'^{+})$ denote the right-hand limit of Q at w_i' and $Q(w_i'^{-})$ the corresponding left-hand limit. Let $0<\varepsilon\leq\min\left\{\min_{w(0,\alpha^*)\leq w_i\leq w_i'}\frac{Y(w_i)}{H(w_i'|\alpha^*)}, Q(w_i'^{+})-Q(w_i'^{-})\right\}$. Define $\hat{Q}$ as follows. If $w_i\leq w(0,\alpha^*)$, let $\hat{Q}(w_i):=Q(w_i)$; and if $w_i>w(0,\alpha^*)$, let

$$\hat{Q}(w_i) := Q(w_i)+\varepsilon\chi_{\{w_i\leq w_i'\}},$$



(a) Case I   (b) Case II

Figure 3: Proof of Lemma 3

where $\chi_{\{w_i \le w_i'\}}$ is an indicator function. (See Figure 3a for an illustration.) By construction, $\hat{Q}(w) \ge Q(w)$ for all $w \in W$ and the inequality holds strictly on a positive measure set. It is also clear that $\hat{Q}$ satisfies (MON). We now verify that $\hat{Q}$ satisfies (IA′) and (F′). Because $\chi_{\{w_i \le w_i'\}}$ is non-increasing on $[w(0,\alpha^*), w(1,\alpha^*)]$, by Lemma 2, we have

$$
\begin{aligned}
&\mathbb{E}\left[-\frac{H_{\alpha_i}(w_i|\alpha^*)}{h(w_i|\alpha^*)}\hat{Q}(w_i)\middle|\,\alpha_i=\alpha^*\right],\\
=&\mathbb{E}\left[-\frac{H_{\alpha_i}(w_i|\alpha^*)}{h(w_i|\alpha^*)}Q(w_i)\middle|\,\alpha_i=\alpha^*\right]+\varepsilon\mathbb{E}\left[-\frac{H_{\alpha_i}(w_i|\alpha^*)}{h(w_i|\alpha^*)}\chi_{\{w_i\le w_i'\}}\middle|\,\alpha_i=\alpha^*\right],\\
\le& C'(\alpha^*)+0=C'(\alpha^*).
\end{aligned}
$$

Hence, $\hat{Q}$ satisfies (IA′). Finally, let

$$
\hat{Y}(w_i) := \int_{w_i}^{\overline{\theta}} \left[H(z|\alpha^*)^{n-1} - \hat{Q}(z)\right] h(z|\alpha^*)\mathrm{d}z.
$$

If $w_i \le w_i'$, $\hat{Y}(w_i) = Y(w_i) - \varepsilon[H(w_i'|\alpha^*) - H(w_i|\alpha^*)] \ge Y(w_i) - \varepsilon H(w_i'|\alpha^*) \ge 0$. If $w_i > w_i'$, $\hat{Y}(w_i) = Y(w_i) \ge 0$. Hence, $\hat{Q}$ satisfies (F′).

Case II: Suppose that Q is continuous on $[w(0,\alpha^*), w^\flat]$.

We first show that there exists $w_i' \in (w(0,\alpha^*), w^\flat)$ such that $Q(w_i') < Q(w^\flat)$. Suppose, to the contrary, that $Q(w_i) = Q(w^\flat)$ for all $w_i \in (w(0,\alpha^*), w^\flat)$. If $Q(w^\flat) \ge H(w^\flat|\alpha^*)^{n-1}$, $Y(w(0,\alpha^*)) = \int_{w(0,\alpha^*)}^{w^\flat}[H(z|\alpha^*)^{n-1} - Q(z)]h(z|\alpha^*)\mathrm{d}z < 0$, a contradiction. If $Q(w^\flat) < H(w^\flat|\alpha^*)^{n-1}$, by the continuity of Q and H, there exists $\delta > 0$ such that $Q(w_i) < H(w_i|\alpha^*)^{n-1}$

for all $w_i \in [w^\flat, w^\flat + \delta]$. Hence,

$$0 = Y(w^\flat) = \int_{w^\flat}^{w^\flat+\delta} [H(z|\alpha^*)^{n-1} - Q(z)]h(z|\alpha^*)\mathrm{d}z + Y(w^\flat + \delta) > Y(w^\flat + \delta),$$

a contradiction. Thus, there exists $w_i' \in (w(0, \alpha^*), w^\flat)$ such that $Q(w_i') < Q(w^\flat)$.

By the continuity of Q, there exists $w_i'' \in (w_i', w^\flat)$ such that $Q(w_i'') = \frac{1}{2}\left(Q(w_i') + Q(w^\flat)\right)$. Let $0 < \varepsilon \leq \min\left\{\min_{w(0,\alpha^*)\leq w_i \leq w_i''} \frac{Y(w_i)}{H(w_i''|\alpha^*)}, Q(w_i'') - Q(w_i')\right\}$. Let

$$\hat{Q}(w_i) := \begin{cases} \max\{Q(w_i') + \varepsilon, Q(w_i)\} & \text{if } w_i > w_i', \\ Q(w_i) + \varepsilon & \text{if } w(0, \alpha^*) < w_i \leq w_i', \\ Q(w_i) & \text{if } w_i \leq w(0, \alpha^*). \end{cases}$$

(See Figure 3b for an illustration.) Note that if $w_i \geq w_i''$, $Q(w_i) \geq Q(w_i'') \geq Q(w_i') + \varepsilon$. Hence, $\hat{Q}(w_i) = Q(w_i)$ for $w_i \geq w_i''$. By construction, $\hat{Q}(w) \geq Q(w)$ for all $w \in W$ and the inequality holds strictly on a positive measure set. Clearly, $\hat{Q}$ satisfies (MON). We now verify that $\hat{Q}$ satisfies (IA′) and (F′). It is easy to verify that $\hat{Q} - Q$ is non-increasing on $[w(0, \alpha^*), w(1, \alpha^*)]$ and therefore $\hat{Q}$ satisfies (IA′) by Lemma 2. Finally, if $w_i \geq w_i''$, $\hat{Y}(w_i) = Y(w_i)$. If $w_i < w_i''$,

$$\begin{aligned}\hat{Y}(w_i) = &\int_{w_i}^{w^\flat} \left[H(z|\alpha^*)^{n-1} - \hat{Q}(z)\right] h(z|\alpha^*)\mathrm{d}z, \\ =&Y(w_i) - \int_{w_i}^{w_i''} \left[\hat{Q}(z) - Q(z)\right] h(z|\alpha^*)\mathrm{d}z, \\ \geq& Y(w_i) - \varepsilon\left[H(w_i''|\alpha^*) - H(w_i|\alpha^*)\right], \\ \geq& Y(w_i) - \varepsilon H(w_i''|\alpha^*) \geq 0.\end{aligned}$$

Hence, $\hat{Q}$ satisfies (F′).

Case III: Suppose that Q is continuous on $[w(0, \alpha^*), w^\flat)$ and $Q(w^{\flat-}) < H(w^\flat|\alpha^*)^{n-1}$.

Define $R(w_i) := Y(w_i)/(H(w^\flat|\alpha^*) - H(w_i|\alpha^*))$ for $w_i < w^\flat$. Then by L'Hopital's rule,

$$\lim_{w_i \to w^{\flat-}} R(w_i) = H(w^\flat|\alpha^*)^{n-1} - Q(w^{\flat-}) > 0.$$

Let $0 < \varepsilon \leq \min\left\{\inf_{w(0,\alpha^*)\leq w_i < w^\flat} R(w_i), Q(w^{\flat+}) - Q(w^{\flat-})\right\}$. Define $\hat{Q}$ as follows. If $w_i \leq w(0, \alpha^*)$, let $\hat{Q}(w_i) := Q(w_i)$; and if $w_i > w(0, \alpha^*)$, let $\hat{Q}(w_i) := Q(w_i) + \varepsilon\chi_{\{w_i < w^\flat\}}$. By construction, $\hat{Q}(w) \geq Q(w)$ for all $w \in W$ and the inequality holds strictly on a positive measure set. Clearly, $\hat{Q}$ satisfies (MON). We can verify that $\hat{Q}$ satisfies (IA′) following the arguments in Case I. Finally, if $w_i < w^\flat$, $\hat{Y}(w_i) = Y(w_i) - \varepsilon[H(w^\flat|\alpha^*) - H(w_i|\alpha^*)] \geq Y(w_i) - R(w_i)[H(w^\flat|\alpha^*) - H(w_i|\alpha^*)] = 0$. If $w_i \geq w^\flat$, $\hat{Y}(w_i) = Y(w_i) \geq 0$. Hence, $\hat{Q}$ satisfies (F′).

Case IV: Suppose that Q is continuous on $[w(0,\alpha^*), w^\flat)$ and $Q(w^{\flat -}) \geq H^{n-1}(w^\flat|\alpha^*)$.

We first show that $Q(w^{\flat -}) = H^{n-1}(w^\flat|\alpha^*)$. Suppose to the contrary that $Q(w^{\flat -}) > H^{n-1}(w^\flat|\alpha^*)$. Then by the continuity of Q and H on $[w(0,\alpha^*), w^\flat)$, there exists $\delta > 0$ such that $Q(w_i) > H^{n-1}(w_i|\alpha^*)$ for all $w_i \in (w^\flat - \delta, w^\flat)$. Then

$$Y(w^\flat - \delta) = \int_{w^\flat - \delta}^{w^\flat} [H(z|\alpha^*)^{n-1} - Q(z)]h(z|\alpha^*)\mathrm{d}z + Y(w^\flat) < 0,$$

a contradiction. Hence, $Q(w^{\flat -}) = H^{n-1}(w^\flat|\alpha^*)$. Second, we show that there exists $w_i' \in (w(0,\alpha^*), w^\flat)$ such that $Q(w_i') < Q(w^{\flat -})$. Suppose to the contrary that $Q(w_i) = Q(w^{\flat -})$ for all $w_i \in (w(0,\alpha^*), w^\flat)$, then $Y(w(0,\alpha^*)) = \int_{w(0,\alpha^*)}^{w^\flat} [H^{n-1}(z|\alpha^*) - Q(z)]h(z|\alpha^*)\mathrm{d}z + Y(w^\flat) < 0$, a contradiction. Hence, there exists $w_i' \in (w(0,\alpha^*), w^\flat)$ such that $Q(w_i') < Q(w^{\flat -})$. The rest of the proof follows from that of Case II. ■

Proof of Corollary 1. The social planner's revenue is

$$\begin{aligned}
\sum_{i=1}^n \mathbb{E}_{\boldsymbol{w}}\left[t_i(\boldsymbol{w})\right] =& \sum_{i=1}^n \mathbb{E}_{w_i}\left[(1-\gamma)w_i Q_i(w_i) + \mathbb{E}_{w_{-i}}\left[\gamma\left(\sum_{j=1}^n w_j\right) q_i(w_i, w_{-i})\right] - U_i(w_i)\right] \\
=& \sum_{i=1}^n \mathbb{E}_{w_i}\left[\left((1-\gamma)w_i - \frac{1-H(w_i|\alpha^*)}{h(w_i|\alpha^*)}\right) Q_i(w_i)\right] + n\gamma\mu - \sum_{i=1}^n U_i(\underline{\theta}),
\end{aligned}$$

where the last equality follows from Theorem 1 and the envelope condition (3). Let $\tilde{t}_i(\boldsymbol{w}) := t_i(\boldsymbol{w}) + U_i(\underline{\theta})$ for all i and $\boldsymbol{w}$. Clearly, (IR) is satisfied. The social planner's revenue becomes

$$\begin{aligned}
\sum_{i=1}^n \mathbb{E}_{\boldsymbol{w}}\left[\tilde{t}_i(\boldsymbol{w})\right] =& \sum_{i=1}^n \mathbb{E}_{w_i}\left[\left((1-\gamma)w_i - \frac{1-H(w_i|\alpha^*)}{h(w_i|\alpha^*)}\right) Q_i(w_i)\right] + n\gamma\mu \\
=& \sum_{i=1}^n \mathbb{E}_{w_i}\left[\left((1-\gamma)w_i + \gamma\mu - \frac{1-H(w_i|\alpha^*)}{h(w_i|\alpha^*)}\right) Q_i(w_i)\right] \\
\geq& 0,
\end{aligned}$$

where the last inequality follows from Lemma 7 because $(1-\gamma)w_i + \gamma\mu - \frac{1-H(w_i|\alpha^*)}{h(w_i|\alpha^*)}$ is non-decreasing in w_i, $\mathbb{E}_{w_i}\left[(1-\gamma)w_i + \gamma\mu - \frac{1-H(w_i|\alpha^*)}{h(w_i|\alpha^*)}\right] = 0$ and Q_i is non-decreasing. ■

Proof of Corollary 2. We prove the corollary by construction. For $m = 0, 1, \ldots,$ define

Q^m as follows

$$Q^m(w) := \frac{\frac{1}{n}\left[H(w^k|\alpha^*)^n - H(w^{k-1}|\alpha^*)^n\right]}{H(w^k|\alpha^*) - H(w^{k-1}|\alpha^*)} \text{ if } w^{k-1} \leq w < w^k,$$

where $w^k := w(0,\alpha^*) + k[w(1,\alpha^*) - w(0,\alpha^*)]/2^m$ for $k = 0, 1, \ldots, 2^m$. For $k = 0, 1, \ldots, 2^m$, define $Q^{m,k}$ as follows

$$Q^{m,k}(w) := \begin{cases} Q^{m+1}(w) & \text{if } w \leq w^k \\ Q^m(w) & \text{otherwise} \end{cases}.$$

Then $Q^{m,0} = Q^m$ and $Q^{m,2^m} = Q^{m+1}$. It is easy to verify that, for all $k = 0, 1, \ldots, 2^m$, $Q^{m,k+1} - Q^{m,k}$ is quasi-monotone and

$$\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \left[Q^{m,k+1}(w) - Q^{m,k}(w)\right] h(w|\alpha^*)\mathrm{d}w = \frac{1}{n} - \frac{1}{n} = 0.$$

Furthermore, $-H_\alpha(w|\alpha^*)/h(w|\alpha^*)$ is non-decreasing. By Lemma 7,

$$\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)} \left[Q^{m,k+1}(w) - Q^{m,k}(w)\right] h(w|\alpha^*)\mathrm{d}w \geq 0.$$

By induction,

$$\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)} \left[Q^{m+1}(w) - Q^{m}(w)\right] h(w|\alpha^*)\mathrm{d}w \geq 0.$$

If there exists m such that Q^m satisfies (IA′) with equality, we are done. Otherwise, because $\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -H_\alpha(w|\alpha^*)Q^0(w)\mathrm{d}w = 0 < C'(\alpha^*)$ and

$$\lim_{m\to\infty} \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -H_\alpha(w|\alpha^*)Q^m(w)\mathrm{d}w = \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -H_\alpha(w|\alpha^*)H(w|\alpha^*)^{n-1}\mathrm{d}w > C'(\alpha^*),$$

there exists m such that

$$\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -H_\alpha(w|\alpha^*)Q^m(w)\mathrm{d}w < C'(\alpha^*) < \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -H_\alpha(w|\alpha^*)Q^{m+1}(w)\mathrm{d}w.$$

If there exists k such that $Q^{m,k}$ satisfies (IA′) with equality, we are done. Otherwise, there exists k such that

$$\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -H_\alpha(w|\alpha^*)Q^{m,k}(w)\mathrm{d}w < C'(\alpha^*) < \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -H_\alpha(w|\alpha^*)Q^{m,k+1}(w)\mathrm{d}w.$$

For each $r \in [w^k, w^{k+1}]$, define $Q^{m,k,r}$ as follows

$$Q^{m,k,r}(w) := \begin{cases} Q^{m+1}(w) & \text{if } w < w^k \\ \frac{\frac{1}{n}\left[H(r|\alpha^*)^n - H(w^k|\alpha^*)^n\right]}{H(r|\alpha^*) - H(w^k|\alpha^*)} & \text{if } w^k \leq w < r \\ \frac{\frac{1}{n}\left[H(w^{k+1}|\alpha^*)^n - H(r|\alpha^*)^n\right]}{H(w^{k+1}|\alpha^*) - H(r|\alpha^*)} & \text{if } r \leq w < w^{k+1} \\ Q^m(w) & \text{if } w \geq w^{k+1} \end{cases}.$$

By construction, $Q^{m,k,w^{k+1}} = Q^{m,k+1}$ and $Q^{m,k,w^k} = Q^{m,k}$. By continuity, there exists $r \in [w^k, w^{k+1}]$ such that $Q^{m,k,r}$ satisfies (IA′) with equality. By Proposition 1, $Q^{m,k,r}$ solves ($\mathcal{P}'$-α^*). ■

Derivation of Inequality (11)**.** The proof is based on Toikka (2011). For the ease of notation, we suppress the dependence of φ^{λ_X}, $\overline{\varphi}^{\lambda_X}$, J^{λ^X} and $\overline{J}^{\lambda_X}$ on λ_X. Then we have

$$\begin{aligned}
&\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} zQ(z)h(z|\alpha^*)\mathrm{d}z \\
\leq &\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \left[z + \lambda_X \frac{H_\alpha(z|\alpha^*)}{h(z|\alpha^*)}\right] Q(z)h(z|\alpha^*)\mathrm{d}z + \lambda_X C'(\alpha^*) && (24) \\
= &\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \varphi^{\lambda_X}(H(z|\alpha^*))Q(z)\mathrm{d}H(z|\alpha^*) + \lambda_X C'(\alpha^*) \\
= &\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \left[\varphi^{\lambda_X}(H(z|\alpha^*)) - \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))\right] Q(z)\mathrm{d}H(z|\alpha^*) \\
&+ \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))Q(z)\mathrm{d}H(z|\alpha^*) + \lambda_X C'(\alpha^*) \\
= &- \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \left[J^{\lambda_X}(H(z|\alpha^*)) - \overline{J}^{\lambda_X}(H(z|\alpha^*))\right] \mathrm{d}Q(z) \\
&+ \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))Q(z)\mathrm{d}H(z|\alpha^*) + \lambda_X C'(\alpha^*) && (25) \\
\leq &\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))Q(z)\mathrm{d}H(z|\alpha^*) + \lambda_X C'(\alpha^*) && (26) \\
= &\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))H(z|\alpha^*)^{n-1}\mathrm{d}H(z|\alpha^*) + \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))Y'(z)\mathrm{d}z + \lambda_X C'(\alpha^*) && (27) \\
\leq &\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))H(z|\alpha^*)^{n-1}\mathrm{d}H(z|\alpha^*) - \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} Y(z)\mathrm{d}\overline{\varphi}^{\lambda_X}(H(z|\alpha^*)) + \lambda_X C'(\alpha^*) && (28) \\
\leq &\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^{\lambda_X}(H(z|\alpha^*))H(z|\alpha^*)^{n-1}\mathrm{d}H(z|\alpha^*) + \lambda_X C'(\alpha^*). && (29)
\end{aligned}$$

Here, inequality (24) holds because $\lambda_X \geq 0$ and Q satisfies (IA′) and the equality holds if and only if Q satisfies (IA′) with equality. Equality (25) follows from integration by parts. Inequality (26) holds because $J^{\lambda_X} \geq \overline{J}^{\lambda_X}$ and the equality holds if and only if Q satisfies the first pooling property. Equality (27) follows from the definition of Y. Inequality (28) follows from integration by parts and the fact that $Y((w(0,\alpha^*))) \geq 0$ and the equality holds if and only if $Y(w(0,\alpha^*)) = 0$. Finally, inequality (29) holds because $Y \geq 0$ and the equality holds if and only if Q satisfies the second pooling property. ■

A.1 Proof of Lemma 5

The proof uses a network-flow approach (see Che et al. (2013) for detailed discussions of this approach). For simplicity, we prove here the result for the case of finite W. By a similar argument to the proof of Theorem 5 in Che et al. (2013), the result generalizes to the case of continuum W. A direct and more involved proof for the continuum case based on Border (1991) is available upon request.

We abuse notation a bit and let f denote the probability mass function in the case of finite W. In this case, (8) becomes:

$$\sum_{\boldsymbol{w}\in A^n}\prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w}) \le n\sum_{w\in A} f(w)Q(w) \le \sum_{\boldsymbol{w}\in A^n}\prod_{i=1}^{n} f(w_i),\ \forall A\subset W. \tag{30}$$

The proof is similar to that of Theorem 3 in Che et al. (2013). Before proceeding to the proof, we first introduce some notations and definitions. Let $D_i := \{(w_i, i)|w_i \in W\}$ and $D := \cup_{i=1}^{n} D_i$, where the latter is known as the *disjoint union* of the individual posterior mean spaces. To simplify notation, we write typical elements of D as w_i instead of (w_i, i). Given an interim allocation rule Q, define a circulation network (N, E, k, d) as follows. The node set is $N := D \cup W^n \cup \{\circ\}$ consisting of demand nodes D, supply nodes W^n, and a circulation node $\circ$. Directed edges $E \subset N \times N$ specify the pairs of nodes that can carry flows. There are three different kinds of edges:

- Edges from supply nodes to demand nodes: $(\tilde{\boldsymbol{w}}, w_i) \in E$ if $\tilde{w}_i = w_i$.
- Edges from demand nodes to the circulation node $\circ$: $(w_i, \circ) \in E$ for all $w_i \in D$.
- Edges from the circulation node $\circ$ to supply nodes: $(\circ, \boldsymbol{w}) \in E$ for all $\boldsymbol{w} \in W^n$.

Let $d(\nu, N')$ and $k(\nu, N')$ denote a lower and upper bound for the total flow from node ν to subset $N' \subset N\backslash\{\nu\}$. There are three different kinds of flow capacities:

- Flow capacities from supply nodes: For each supply node $\boldsymbol{w} \in W^n$, let $d(\boldsymbol{w}, N') = \prod_{i=1}^{n} f(w_i)\rho(\boldsymbol{w})$ if $N' \supset \{w_1, \ldots, w_n\}$ or else $d(\boldsymbol{w}, N') = 0$; and let $k(\boldsymbol{w}, N') = \prod_{i=1}^{n} f(w_i)$ if $N' \cap \{w_1, \ldots, w_n\} \neq \emptyset$ or else $k(\boldsymbol{w}, N') = 0$.
- Flow capacities from demand nodes: For each demand node $w_i \in D$, let $k(w_i, N') = d(w_i, N') = f(w_i)Q(w_i)$ if $\circ \in N'$ or else $k(w_i, N') = d(w_i, N') = 0$.
- Flow capacities from $\circ$: Let $d(\circ, N') = 0$ and $k(\circ, N') = K$ for some $K > 0$ sufficiently large.

A *feasible circulation flow* on (N, E, k, d) is a function $\zeta : E \to \mathbb{R}_+$ that satisfies the *capacity constraints*

$$d(\nu, N') \leq \sum_{\nu' \in N' : (\nu, \nu') \in E} \zeta(\nu, \nu') \leq k(\nu, N'), \ \forall \nu \in N, \forall N' \subset N \backslash \{\nu\},$$

and the *flow conservation law*

$$\sum_{\nu' : (\nu, \nu') \in E} \zeta(\nu, \nu') = \sum_{\nu' : (\nu', \nu) \in E} \zeta(\nu', \nu), \ \forall \nu \in N.$$

By Theorem 1 in Che et al. (2013), an interim allocation rule Q is implementable by an ex post allocation rule $\boldsymbol{q}$ satisfying $\sum_{i=1}^n q_i(\boldsymbol{w}) \geq \rho(\boldsymbol{w})$ for all $\boldsymbol{w}$ if and only if there exists a feasible circulation flow for the network (N, E, k, d) defined above. It is easy to verify that for every $\nu \in N$, $k(\nu, \cdot)$ and $d(\nu, \cdot)$ are *paramodular*:

1. $k(\nu, \cdot)$ is *submodular*: For any $N', N'' \subset N$, $k(\nu, N') + k(\nu, N'') \geq k(\nu, N' \cup N'') + k(\nu, N' \cap N'')$.

2. $d(\nu, \cdot)$ is *supermodular*: For any $N', N'' \subset N$, $d(\nu, N') + d(\nu, N'') \leq d(\nu, N' \cup N'') + d(\nu, N' \cap N'')$.

3. $k(\nu, \cdot)$ and $d(\nu, \cdot)$ are *compliant*: For any $N', N'' \subset N$, $k(\nu, N') - d(\nu, N'') \geq k(\nu, N' \backslash N'') - d(\nu, N'' \backslash N')$.

Hence, by Theorem 1 in Hassin (1982), a feasible circulation flow $\zeta : E \to \mathbb{R}_+$ exists if and only if

$$\sum_{\nu \in N \backslash N'} d(\nu, N') \leq \sum_{\nu \in N'} k(\nu, N \backslash N'), \ \forall N' \subset N, \tag{31}$$

which requires that the sum of lower bounds on the flows entering N' does not exceed the sum of upper bounds on the flows exiting N'.

Necessity: Suppose that the interim allocation rule Q is the reduced form of an ex post allocation rule $\boldsymbol{q}$ satisfying $\sum_{i=1}^n q_i(\boldsymbol{w}) \geq \rho(\boldsymbol{w})$ for all $\boldsymbol{w}$. Then, by Theorem 1 in Che et al. (2013) and Theorem 1 in Hassin (1982), (31) holds. Let $N' = \cup_{i=1}^n \{(w_i, i) | w_i \in A\} \subset D$, where $A \subset W$ is a measurable set. The right-hand side of (31) becomes

$$\begin{aligned}\sum_{\nu \in N'} k(\nu, N \backslash N') &= \sum_{w_i \in N'} k(w_i, \circ) \\ &= \sum_{w_i \in N'} f(w_i) Q(w_i) \\ &= n \sum_{w \in A} f(w) Q(w)\end{aligned}$$

and the left-hand side of (31) becomes

$$\begin{aligned}\sum_{\nu \in N \backslash N'} d(\nu, N') &= \sum_{\boldsymbol{w}:\{w_i\} \subset N'} d(\boldsymbol{w}, N') \\ &= \sum_{\boldsymbol{w} \in A^n} \prod_{i=1}^{n} f(w_i) \rho(\boldsymbol{w}),\end{aligned}$$

which proves the first inequality in (30). Let $N' = N \backslash \cup_{i=1}^{n} \{(w_i, i) | w_i \in A\}$. The right-hand side of (31) becomes

$$\sum_{\nu \in N'} k(\nu, N \backslash N') = \sum_{\boldsymbol{w} \in A^n} \prod_{i=1}^{n} f(w_i)$$

and the left-hand side of (31) becomes

$$\begin{aligned}\sum_{\nu \in N \backslash N'} d(\nu, N') &= \sum_{i=1}^{n} \sum_{w_i \in A} d(w_i, \circ) \\ &= n \sum_{w \in A} f(w) Q(w),\end{aligned}$$

which proves the second inequality in (30).

Sufficiency: Because ρ is symmetric, by a similar argument to that in the proof of Theorem 7 in Che et al. (2013), (30) holds if and only if

$$\sum_{\boldsymbol{w} \in \prod_i A_i} \prod_{i=1}^{n} f(w_i) \rho(\boldsymbol{w}) \leq \sum_{i=1}^{n} \sum_{w_i \in A_i} f(w_i) Q(w_i) \leq \sum_{\boldsymbol{w} \in \cup_i (A_i \times W^{n-1})} \prod_{i=1}^{n} f(w_i), \ \forall \prod_{i=1}^{n} A_i \subset W^n. \tag{32}$$

For completeness, we include a proof of this claim in Lemma 8 below. In what follows, we show that (32) implies (31). Hence, (30) implies (31). Then sufficiency follows from Theorem 1 in Hassin (1982) and Theorem 1 in Che et al. (2013).

Fix $N' \subset N$. Suppose first that $\circ \notin N'$. Let $A_i = N' \cap D_i$ for all i. In this case,

$$\begin{aligned}
\sum_{\nu \in N'} d(\nu, N') &= d(\circ, N' \cap W^n) + \sum_{\boldsymbol{w} \in W^n \setminus N'} d(\boldsymbol{w}, N' \cap D) \\
&= \sum_{\boldsymbol{w} \in W^n \setminus N'} d(\boldsymbol{w}, N' \cap D) \\
&\leq \sum_{\boldsymbol{w} \in W^n} d(\boldsymbol{w}, N' \cap D) \\
&= \sum_{\boldsymbol{w} \in \prod_{i=1}^n A_i} d(\boldsymbol{w}, N' \cap D) \\
&= \sum_{\boldsymbol{w} \in \prod_{i=1}^n A_i} \prod_{i=1}^n f(w_i) \rho(\boldsymbol{w}) \\
&\leq \sum_{i=1}^n \sum_{w_i \in A_i} f(w_i) Q(w_i) \\
&= \sum_{w_i \in D \cap N'} k(w_i, \circ) \\
&= \sum_{\nu \in N'} k(\nu, N \setminus N').
\end{aligned}$$

Suppose next $\circ \in N'$. Then if $W^n \not\subset N'$, we have $\sum_{\nu \in N'} k(\nu, N \setminus N') \geq k(\circ, N \setminus N') = K > \sum_{\nu \in N \setminus N'} d(\nu, N')$ for K sufficiently large. Otherwise, if $W^n \subset N'$, then let $A_i = D_i \setminus N'$ for all i and

$$\begin{aligned}
\sum_{\nu \in N'} k(\nu, N \setminus N') &= \sum_{\boldsymbol{w} \in \cup_i (A_i \times W^{n-1})} k(\boldsymbol{w}, D \setminus N') \\
&= \sum_{\boldsymbol{w} \in A_i \times W^{n-1}} \prod_{i=1}^n f(w_i) \\
&\geq \sum_{i=1}^n \sum_{w_i \in A_i} f(w_i) Q(w_i) \\
&= \sum_{\nu \in D \setminus N'} d(\nu, N').
\end{aligned}$$

Lemma 8 *Suppose that all the condition in Lemma 5 are satisfied.* (30) *holds if and only if* (32) *holds.*

Proof. Clearly, if (32) holds, (30) holds. Suppose that (30) holds. We prove here only the first inequality in (30). Virtually the same argument can be applied to prove the second inequality in (32).

Suppose, to the contrary, that there exists $\prod_{i=1}^n A_i \subset W^n$ such that the first inequality in (32) is violated. Suppose that $\prod_{i=1}^n A_i$ is minimal in the sense that for all proper subsets $\prod_{i=1}^n A'_i \subsetneq \prod_{i=1}^n A_i$, the first inequality in (32) holds. Let $\overline{A} := \cup_i A_i$. We want to show that the first inequality in (30) is violated for $\overline{A}$, which is a contradiction.

To show this, we show that starting from $\prod_{i=1}^n A_i$, we can construct a finite sequence of sets $\prod_{i=1}^n A_i = \mathscr{S}^1 \subsetneq \mathscr{S}^2 \subsetneq \cdots \subsetneq \mathscr{S}^M = \overline{A}^n$ such that the first inequality in (32) is violated for all $\mathscr{S}^m$. The sequence is constructed inductively:

Step 1. Let $\mathscr{S}^1 := \prod_{i=1}^n A_i$.

Step m. If $\mathscr{S}^{m-1} = \overline{A}^n$, we are done. Otherwise, there exist $j, k \in \{1, \ldots, n\}$ such that $B_j := A_j \backslash \mathscr{S}_k^{m-1} \neq \emptyset$ or $B_k := A_k \backslash \mathscr{S}_j^{m-1} \neq \emptyset$. Let $\mathscr{S}^m := (\mathscr{S}_j^{m-1} \cup B_k) \times (\mathscr{S}_k^{m-1} \cup B_j) \times \prod_{i \neq j,k} \mathscr{S}_i^{m-1}$.

Because there are a finite number of agents, the construction stops after a finite number of steps. Next we show that if the first inequality in (32) is violated for $\mathscr{S}^m$, it is also violated for $\mathscr{S}^{m+1}$. Recall that the first inequality in (32) is violated for $\prod_{i=1}^n A_i$:

$$\sum_{i=1}^n \sum_{w_i \in A_i} f(w_i) Q(w_i) < \sum_{\boldsymbol{w} \in \prod_{i=1}^n A_i} \prod_{i=1}^n f(w_i) \rho(\boldsymbol{w}) \tag{33}$$

Because $\prod_{i=1}^n A_i$ is chosen minimally, we have

$$\begin{aligned} &\sum_{i=1}^n \sum_{w_i \in A_i} f(w_i) Q(w_i) - \sum_{w_j \in B_j} f(w_j) Q(w_j) - \sum_{w_k \in B_k} f(w_k) Q(w_k) \\ \geq &\sum_{\boldsymbol{w} \in (A_j \backslash B_j) \times (A_k \backslash B_k) \times \prod_{i \neq j,k} A_i} \prod_{i=1}^n f(w_i) \rho(\boldsymbol{w}). \end{aligned}$$

Hence,

$$\begin{aligned} &\sum_{w_j \in B_j} f(w_j) Q(w_j) + \sum_{w_k \in B_k} f(w_k) Q(w_k) \\ < &\sum_{\boldsymbol{w} \in \prod_{i=1}^n A_i} \prod_{i=1}^n f(w_i) \rho(\boldsymbol{w}) - \sum_{\boldsymbol{w} \in (A_j \backslash B_j) \times (A_k \backslash B_k) \times \prod_{i \neq j,k} A_i} \prod_{i=1}^n f(w_i) \rho(\boldsymbol{w}). \end{aligned}$$

For $\mathscr{S}^{m+1} = (\mathscr{S}_j^m \cup B_k) \times (\mathscr{S}_k^m \cup B_j) \times \prod_{i \neq j,k} \mathscr{S}_i^m$, we have

$$
\begin{aligned}
&\sum_{i=1}^{n} \sum_{w_i \in \mathscr{S}_i^m} f(w_i) Q(w_i) + \sum_{w_j \in B_k} f(w_j) Q(w_j) + \sum_{w_k \in B_j} f(w_k) Q(w_k) \\
=&\sum_{i=1}^{n} \sum_{w_i \in \mathscr{S}_i^m} f(w_i) Q(w_i) + \sum_{w_j \in B_j} f(w_j) Q(w_j) + \sum_{w_k \in B_k} f(w_k) Q(w_k) \\
<& \sum_{\boldsymbol{w} \in \mathscr{S}^m} \prod_{i=1}^{n} f(w_i) \rho(\boldsymbol{w}) + \sum_{\boldsymbol{w} \in A_j \times A_k \times \prod_{i \neq j,k} A_i} \prod_{i=1}^{n} f(w_i) \rho(\boldsymbol{w}) - \sum_{\boldsymbol{w} \in (A_j \backslash B_j) \times (A_k \backslash B_k) \times \prod_{i \neq j,k} A_i} \prod_{i=1}^{n} f(w_i) \rho(\boldsymbol{w}) \\
=& \sum_{\boldsymbol{w} \in \mathscr{S}^m} \prod_{i=1}^{n} f(w_i) \rho(\boldsymbol{w}) + \sum_{\boldsymbol{w} \in A_k \times A_j \times \prod_{i \neq j,k} A_i} \prod_{i=1}^{n} f(w_i) \rho(\boldsymbol{w}) - \sum_{\boldsymbol{w} \in (A_k \backslash B_k) \times (A_j \backslash B_j) \times \prod_{i \neq j,k} A_i} \prod_{i=1}^{n} f(w_i) \rho(\boldsymbol{w}) \\
\leq& \sum_{\boldsymbol{w} \in \mathscr{S}^m} \prod_{i=1}^{n} f(w_i) \rho(\boldsymbol{w}) + \sum_{\boldsymbol{w} \in (\mathscr{S}_j^m \cup B_k) \times (\mathscr{S}_k^m \cup B_j) \times \prod_{i \neq j,k} \mathscr{S}_i^m} \prod_{i=1}^{n} f(w_i) \rho(\boldsymbol{w}) \\
&- \sum_{\boldsymbol{w} \in \mathscr{S}_j^m \times \mathscr{S}_k^m \times \prod_{i \neq j,k} \mathscr{S}_i^m} \prod_{i=1}^{n} f(w_i) \rho(\boldsymbol{w}) \\
=& \sum_{\boldsymbol{w} \in \mathscr{S}^{m+1}} \prod_{i=1}^{n} f(w_i) \rho(\boldsymbol{w}),
\end{aligned}
$$

where the fourth line holds because ρ is symmetric, and the fifth line holds because

$$\sum_{\boldsymbol{w} \in \prod_{i=1}^n A_i} \prod_{i=1}^{n} f(w_i) \rho(\boldsymbol{w})$$

is supermodular over $\prod_{i=1}^n A_i$. ■

A.2 Solving ($\mathcal{P}'$-α^*)

Throughout Appendix A.2 we assume that all the conditions in Theorem 2 are satisfied. Recall that the sub-problem ($\mathcal{P}'$-α^*) is

$$V(\alpha^*) := \max_{Q} \mathbb{E}\left[wQ(w) \middle| \, \alpha^* \right],$$

subject to

$$Y(w) := \int_w^{\bar\theta} [H(z|\alpha^*)^{n-1} - Q(z)]h(z|\alpha^*)\mathrm{d}z \geq 0,\ \forall w \in [\underline{\theta}, \bar\theta]. \tag{F'}$$

$$Q(w) \text{ is non-decreasing in } w, \tag{MON}$$

$$\mathbb{E}\left[-\frac{H_\alpha(w|\alpha^*)}{h(w|\alpha^*)} Q(w) \middle| \alpha^* \right] \leq C'(\alpha^*). \tag{IA'}$$

For brevity, denote $w(0,\alpha^*)$ by $\underline{w}$, $w(1,\alpha^*)$ by $\overline{w}$, $h(w|\alpha^*)$ by $h(w)$, $H(w|\alpha^*)$ by $H(w)$ and $H_\alpha(w|\alpha^*)$ by $H_\alpha(w)$. Let $X(w) := \int_0^w H_\alpha(z)Q(z)\mathrm{d}z$ for all $w \in [\underline{w}, \overline{w}]$. Then this is a control problem with state variables X, Y and Q, and a control variable $a \geq 0$. The evolution of the state variables is governed by

$$X'(w) = H_\alpha(w)Q(w), \tag{34}$$

$$Y'(w) = -[H(w)^{n-1} - Q(w)]h(w), \tag{35}$$

$$Q'(w) = a(w), \tag{36}$$

where the last equality holds if $Q(w)$ is differentiable at w. The non-negativity constraint for a guarantees that Q is non-decreasing. This implies some regularity on Q, but still leaves some problems to apply control theory directly. First, we have to allow for (upward) jumps in the state variable Q. Second, Q is not guaranteed to be piecewise continuous and piecewise continuously differentiable.

These problems can be circumvented by solving the maximization problem under the additional restriction that Q is Lipschitz continuous with global Lipschitz constant K:

$$Q \in \mathcal{L}^K := \{Q : W \to [0,1] | |Q(z) - Q(z')| \leq K|z - z'|\ \forall z, z' \in [0,1]\}.$$

We define the maximization problem ($\mathcal{P}^K$-α^*) as ($\mathcal{P}'$-α^*) subject to the additional constraint $Q \in \mathcal{L}^K$.

We say that Q is a *feasible solution* of ($\mathcal{P}'$-α^*) if it satisfies (MON), (F′) and (IA′), and an *optimal solution* of ($\mathcal{P}'$-α^*) if it maximizes $\mathbb{E}[wQ(w)|\alpha^*]$ subject to (MON), (F′) and (IA′). Similarly, we say that $Q \in \mathcal{L}^K$ is a *feasible solution* of ($\mathcal{P}^K$-α^*) if it satisfies (MON), (F′) and (IA′), and an *optimal solution* of ($\mathcal{P}^K$-α^*) if it maximizes $\mathbb{E}[wQ(w)|\alpha^*]$ subject to (MON), (F′) and (IA′).

Lemma 9 in Appendix A.2.1 shows that an optimal solution of ($\mathcal{P}'$-α^*) exists, and for every $K > 0$, an optimal solution of ($\mathcal{P}^K$-α^*) exists. Lemma 10 in Appendix A.2.1 shows that there exists an optimal solution of ($\mathcal{P}'$-α^*), which is the pointwise limit of the optimal solutions of ($\mathcal{P}^K$-α^*) as K goes to infinity.

The rest of Appendix A.2 is organized as follows. Appendix A.2.1 introduces and proves Lemmas 9 and 10. Appendix A.2.2 gives the necessary conditions that an optimal solution of ($\mathcal{P}^K$-α^*) must satisfy. Appendix A.2.3 proves Theorem 2. Appendix A.2.4 proves Lemma 6.

A.2.1 Existence of optimal solutions

Before introducing and proving Lemmas 9 and 10, we first introduce some notations. We abuse notation a bit and let h denote the probability measure on W corresponding to $H(w)$. In what follows, let $L_2(h)$ denote the set of measurable functions whose absolute value raised to the 2nd power has finite integral. For brevity, denote $L_2(h)$ by L_2. Because L_2 is the dual of L_2 under the duality $\langle f, g\rangle = \mathbb{E}_w\left[f(w)g(w)|\alpha = \alpha^*\right]$, topologize L_2 with its weak*, or $\sigma(L_2, L_2)$, topology.

Lemma 9 *The following two statements are true.*

1. *An optimal solution of ($\mathcal{P}'$-α^*) exists.*
2. *For every $K > 0$, an optimal solution of ($\mathcal{P}^K$-α^*) exists.*

Proof. The proof is based on Mierendorff (2009).

1. Let $\{Q^\nu\}$ be a sequence of feasible solutions of ($\mathcal{P}'$-α^*) such that

$$\int_{\underline{w}}^{\overline{w}} zQ^\nu(z)h(z)\mathrm{d}z \to V(\alpha^*).$$

By Helly's selection theorem, there exists a subsequence $\{Q^{\nu_\kappa}\}$ and a non-decreasing function Q such that Q^{ν_κ} converges pointwise to Q. Let $\mathscr{D}$ collect all $Q : W \to [0, 1]$ that satisfies (F′), (MON) and (IA′). Consider $\mathscr{D}$ as a subset of L_2. Recall that h is the probability measure on W corresponding to $H(z)$. Then $\mathscr{D}$ is $\sigma(L_2, L_2)$ compact by a proof similar to that of Lemma 5.4 in Border (1991) and Lemma 8 in Mierendorff (2011). Therefore, after taking subsequences again, Q^{ν_κ} converges to Q in $\sigma(L_2, L_2)$ topology and $Q \in \mathscr{D}$. Because $z \in L_2$ and $h \in L_2$, the weak convergence of $\{Q^{\nu_\kappa}\}$ implies that

$$\int_{\underline{w}}^{\overline{w}} zQ(z)h(z)\mathrm{d}z = V(\alpha^*).$$

2. Let $\{Q^\nu\}$ be a sequence of feasible solutions of ($\mathcal{P}^K$-α^*) such that

$$\int_{\underline{w}}^{\overline{w}} zQ^\nu(z)h(z)\mathrm{d}z \to V^K(\alpha^*).$$

After taking subsequences, we can assume that Q^ν converges to Q pointwise and in $\sigma(L_2, L_2)$ topology, and $Q \in \mathscr{D}$ as in part 1. Because $Q^\nu \in \mathcal{L}^K$, for all $z, z' \in W$,

$$|Q(z) - Q(z')| = \lim_{\nu\to\infty} |Q^\nu(z) - Q^\nu(z')| \leq K|z - z'|.$$

Hence, $Q \in \mathcal{L}^K$.

■

Lemma 10 *Let $\{Q^K\}$ be a sequence of optimal solutions of ($\mathcal{P}^K$-α^*) where $K \to \infty$. Then there exists a feasible solution Q of ($\mathcal{P}'$-α^*) and a subsequence Q^{K_ν} such that Q^{K_ν} converges to Q for almost every $w \in W$. Furthermore, Q is optimal, i.e.,*

$$\int_{\underline{w}}^{\overline{w}} zQ(z)h(z)\mathrm{d}z = V(\alpha^*).$$

Proof. The proof is based on Reid (1968) and Mierendorff (2009). After taking a subsequence, we can assume that Q^K converges pointwise to a feasible solution $\hat{Q}$ of ($\mathcal{P}'$-α^*) (see the proof of Lemma 9). To show the optimality of $\hat{Q}$, let Q be an optimal solution of ($\mathcal{P}'$-α^*). We can extend Q to $\mathbb{R}$ by setting $Q(z) := 0$ for $z < \underline{w}$ and $Q(z) := 1$ for $z > \overline{w}$. Define $Q_d : \mathbb{R} \to [0, 1]$ as

$$Q_d(z) := \frac{1}{d}\int_{z-d}^{z} Q(\zeta)\mathrm{d}\zeta, \ \forall z \in \mathbb{R}.$$

By the Lebesgue differentiation theorem (see, e.g., Theorem 3.21 in Folland (1999)), $Q_d(z) \to Q(z)$ for almost every $z \in W$ as $d \to 0$. Because Q is non-decreasing and $Q(z) \in [0, 1]$ for all z, Q_d is non-decreasing, $Q_d \leq Q$, and $Q_d(z) \in [0, 1]$ for all z. Furthermore, $Q_d \in \mathcal{L}^{\frac{1}{d}}$: For all $z > z'$,

$$\begin{aligned}
0 \leq Q_d(z) - Q_d(z') =& \frac{1}{d}\left(\int_{z-d}^{z} Q(\zeta)\mathrm{d}\zeta - \int_{z'-d}^{z'} Q(\zeta)\mathrm{d}\zeta\right) \\
=& \frac{1}{d}\left(\int_{z'}^{z} Q(\zeta)\mathrm{d}\zeta - \int_{z'-d}^{z-d} Q(\zeta)\mathrm{d}\zeta\right) \\
\leq& \frac{1}{d}\int_{z'}^{z} Q(\zeta)\mathrm{d}\zeta \\
\leq& \frac{1}{d}(z - z').
\end{aligned}$$

Finally, Q_d satisfies (F′) because $Q_d \leq Q$ and Q satisfies (F′).

Define $\tilde{Q}_d := Q_d$ if $-\int_{\underline{w}}^{\overline{w}} H_{\alpha_i}(z)Q_d(z)\mathrm{d}z \leq C'(\alpha^*)$ and otherwise $\tilde{Q}_d := \beta_d Q_d + (1-\beta_d)/n$, where

$$\beta_d = \frac{C'(\alpha^*)}{-\int_{\underline{w}}^{\overline{w}} H_{\alpha_i}(z)Q_d(z)\mathrm{d}z}.$$

Then by Lemma 2, $-\int_{\underline{w}}^{\overline{w}} H_{\alpha_i}(z)\tilde{Q}_d(z)\mathrm{d}z \leq C'(\alpha^*)$. Thus, $\tilde{Q}_d$ is a feasible solution of ($\mathcal{P}^K$-α^*), where $K = 1/d$. Because $\beta_d \to 0$, $\tilde{Q}_d \to Q$ almost everywhere as $d \to 0$. By the dominated convergence theorem,

$$\int_{\underline{w}}^{\overline{w}} z\tilde{Q}_d(z)h(z)\mathrm{d}z \to \int_{\underline{w}}^{\overline{w}} zQ(z)h(z)\mathrm{d}z, \text{ as } d \to 0$$

and

$$\int_{\underline{w}}^{\overline{w}} zQ^K(z)h(z)\mathrm{d}z \to \int_{\underline{w}}^{\overline{w}} z\hat{Q}(z)h(z)\mathrm{d}z, \text{ as } K \to \infty.$$

Let $d = 1/K$. Then, for all K, $\tilde{Q}_d$ is a feasible solution of ($\mathcal{P}^K$-α^*) and therefore

$$\int_{\underline{w}}^{\overline{w}} z\tilde{Q}_d(z)h(z)\mathrm{d}z \leq \int_{\underline{w}}^{\overline{w}} zQ^K(z)h(z)\mathrm{d}z.$$

Hence,

$$\int_{\underline{w}}^{\overline{w}} z\hat{Q}(z)h(z)\mathrm{d}z = \int_{\underline{w}}^{\overline{w}} zQ(z)h(z)\mathrm{d}z.$$

This completes the proof. ■

A.2.2 Solving ($\mathcal{P}^K$-α^*)

In this subsection, we derive the necessary conditions that an optimal solution of ($\mathcal{P}^K$-α^*) must satisfy. The problem ($\mathcal{P}^K$-α^*) can be summarized as follows:

$$\max_{X,Y,Q,a} \int_{\underline{w}}^{\overline{w}} zQ(z)h(z)\mathrm{d}z, \tag{$\mathcal{P}^K$-α^*}$$

subject to

$$X'(z) = H_\alpha(z)Q(z), \tag{34}$$

$$Y'(z) = -[H(z)^{n-1} - Q(z)]h(z), \tag{35}$$

$$Q'(z) = a(z), \tag{36}$$

$$X(\underline{w}) = 0,\ X(\overline{w}) \geq -C'(\alpha^*), \tag{37}$$

$$Y(\underline{w}) = 0,\ Y(\overline{w}) = 0, \tag{38}$$

$$Q(\underline{w}) \geq 0,\ Q(\overline{w}) \leq 1, \tag{39}$$

$$0 \leq a(z) \leq K, \tag{40}$$

$$Y(z) \geq 0. \tag{41}$$

We say that some property holds virtually everywhere if the property is fulfilled at all z except for a countable number of z's. We use the following abbreviation for "virtually everywhere": $v.e.$ By Theorem 6.7.15 in Seierstad and Sydsæter (1987), we have

Lemma 11 *Let (X, Y, Q, a) be an admissible pair that solves ($\mathcal{P}^K$-α^*). Then there exist a number λ_0, vector functions $(\lambda_X, \check{\lambda}_Y, \lambda_Q)$ and $(\underline{\eta}_a, \overline{\eta}_a)$, and a non-decreasing function η_Y, all* *having one-sided limits everywhere, such that the following condition holds:*

$$\lambda_0 = 0 \textit{ or } \lambda_0 = 1, \tag{42}$$

$$(\lambda_0, \lambda_X(z), \check{\lambda}_Y(z), \lambda_Q(z), \eta_Y(\overline{w}) - \eta_Y(\underline{w})) \neq 0, \ \forall z, \tag{43}$$

$$\lambda_Q(z)a(z) \geq \lambda_Q(z)a, \ \forall a \in (0, K), v.e. \tag{44}$$

$$\lambda_Q(z) - \overline{\eta}_a(z) + \underline{\eta}_a(z) = 0, v.e. \tag{45}$$

$$\eta_Y \textit{ is constant on any interval where } Y > 0. \tag{46}$$

$$\lambda_X \textit{ and } \lambda_Q \textit{ are continuous.} \tag{47}$$

$$\lambda_X'(z) = 0, \ \ v.e. \tag{48}$$

$$\lambda_Q'(z) = -\left[\lambda_0 z + \lambda_X(z)\frac{H_\alpha(z)}{h(z)} + \check{\lambda}_Y(z)\right] h(z) + \eta_Y(z)h(z), \ \ v.e. \tag{49}$$

$$\check{\lambda}_Y(z) + \eta_Y(z) \textit{ is continuous,} \tag{50}$$

$$\check{\lambda}_Y'(z) + \eta_Y'(z) = 0, \ \ v.e. \tag{51}$$

$$\lambda_X(\overline{w}) \geq 0 (= 0 \textit{ if } X(\overline{w}) > -C'(\alpha^*)), \tag{52}$$

$$\lambda_Q(\overline{w}) \leq 0 (= 0 \textit{ if } Q(\overline{w}) < 1), \tag{53}$$

$$\lambda_Q(\underline{w}) \leq 0 (= 0 \textit{ if } Q(\underline{w}) > 0). \tag{54}$$

$$\underline{\eta}_a(z) \geq 0 (= 0 \textit{ if } a(z) > 0), \tag{55}$$

$$\overline{\eta}_a(z) \geq 0 (= 0 \textit{ if } a(z) < K). \tag{56}$$

In what follows, we assume that (X, Y, Q, a) is an admissible pair that solves ($\mathcal{P}^K$-α^*) and $(X, Y, Q, a, \lambda_0, \lambda_X, \check{\lambda}_Y, \lambda_Q, \underline{\eta}_a, \overline{\eta}_a, \eta_Y)$ satisfy the conditions in Lemma 11. We begin the analysis by simplifying the conditions in Lemma 11.

Because λ_X is continuous and $\lambda_X'(z) = 0$ virtually everywhere, $\lambda_X(z)$ is constant on $[\underline{w}, \overline{w}]$. We abuse notation a bit and denote this constant by λ_X. Then (52) is equivalent to

$$\lambda_X \geq 0 (= 0 \text{ if } X(\overline{w}) > -C'(\alpha^*)).$$

Similarly, because $\check{\lambda}_Y + \eta_Y$ is continuous and $\check{\lambda}_Y'(z) + \eta_Y'(z) = 0$ virtually everywhere, $\check{\lambda}_Y(z) + \eta_Y(z)$ is constant on $[\underline{w}, \overline{w}]$. We can assume without loss of generality that $\check{\lambda}_Y(z) + \eta_Y(z) = 0$ for all $z \in [\underline{w}, \overline{w}]$. Let $\lambda_Y := 2\check{\lambda}_Y$. Then $\eta_Y = -\lambda_Y/2$ and condition (46) is equivalent to

$$\lambda_Y(z) \text{ is constant on any interval where } Y(z) > 0, \tag{57}$$

and (49) is equivalent to

$$\lambda_Q'(z) = -\left[\lambda_0 z + \lambda_X(z)\frac{H_{\alpha_i}(z)}{h(z)} + \lambda_Y(z)\right] h(z),\ \text{v.e.}$$

Furthermore, η_Y is non-decreasing if and only if λ_Y is non-increasing. Because λ_Y has one-sided limits everywhere, we can assume without loss of generality that $\lambda_Y(\underline{w}) = \lim_{z\to\underline{w}} \lambda_Y(z)$ and $\lambda_Y(\overline{w}) = \lim_{z\to\overline{w}} \lambda_Y(z)$. Finally, (44), (45), (55) and (56) can be simplified to for virtually all $z \in (\underline{w}, \overline{w})$: If $0 < a(z) < K$, $\lambda_Q(z) = \overline{\eta}_a(z) = \underline{\eta}_a(z) = 0$. If $a(z) = 0, \overline{\eta}_a(z) = 0$ and $-\underline{\eta}_a(z) = \lambda_Q(z) \leq 0$. If $a(z) = K$, $\underline{\eta}_a(z) = 0$ and $\overline{\eta}_a(z) = \lambda_Q(z) \geq 0$.

Then the conditions in Lemma 11 can be simplified as follows:

Corollary 3 *Let (X, Y, Q, a) be an admissible pair that solves ($\mathcal{P}^K$-α^*). If (X, Y, Q, a) is optimal, there exist a constant λ_X, a continuous and piecewise continuously differentiable function λ_Q, and a non-increasing function λ_Y such that the following holds:*

$$\lambda_X \geq 0 \ (= 0 \text{ if } X(\bar{w}) > -C'(\alpha^*)). \tag{58}$$

$$\lambda_Q'(z) = -\left[z + \lambda_X\frac{H_{\alpha_i}(z)}{h(z)} + \lambda_Y(z)\right] h(z),\ \textit{v.e.} \tag{59}$$

$$\lambda_Y \textit{ is constant on any interval where } Y > 0. \tag{60}$$

$$\lambda_Q(\overline{w}) = 0. \tag{61}$$

$$\lambda_Q(\underline{w}) = 0. \tag{62}$$

$$a(z) = \begin{cases} = 0 & \textit{if } \lambda_Q(z) \leq 0, \\ \in [0, K] & \textit{if } \lambda_Q(z) = 0, \quad \textit{v.e.} \\ = K & \textit{if } \lambda_Q(z) \geq 0. \end{cases} \tag{63}$$

Proof. We prove Corollary 3 by proving the following two lemmas.

Lemma 12 $\lambda_Q(\underline{w}) = \lambda_Q(\overline{w}) = 0$.

Proof. By the transversality condition (54), $\lambda_Q(\underline{w}) \leq 0$ and the equality holds if $Q(\underline{w}) > 0$. Suppose, to the contrary, that $\lambda_Q(\underline{w}) < 0$. Then $Q(\underline{w}) = 0$. By continuity there exists $\delta > 0$ such that $\lambda_Q(z) < 0$ for all $z \in (\underline{w}, \underline{w} + \delta)$. Hence, by (44), $a(z) = 0$ for all $z \in (\underline{w}, \underline{w} + \delta)$. This implies that $Q(z) = 0$ for all $z \in (\underline{w}, \underline{w} + \delta)$. Let $z \in (\underline{w}, \underline{w} + \delta)$, then

$$0 = Y(\underline{w}) = \int_{\underline{w}}^{z} H(\zeta)^{n-1} h(z) \mathrm{d}z + Y(z) > Y(z),$$

a contradiction. Hence, $\lambda_Q(\underline{w}) = 0$. A similar argument proves that $\lambda_Q(\overline{w}) = 0$. ■

Lemma 13 *(Non-triviality)* $\lambda_0 = 1$.

Proof. Suppose, to the contrary, that $\lambda_0 = 0$. Then

$$\lambda_Q'(z) = -\left[\lambda_X \frac{H_\alpha(z)}{h(z)} + \lambda_Y(z)\right] h(z), \text{ v.e.}$$

Hence,

$$\begin{aligned}\lambda_Q(\overline{w}) =& \lambda_Q(\underline{w}) - \int_{\underline{w}}^{\overline{w}} \left[\lambda_X \frac{H_\alpha(z)}{h(z)} + \lambda_Y(z)\right] h(z)\mathrm{d}z, \\ =& \lambda_Q(\underline{w}) - \int_{\underline{w}}^{\overline{w}} \lambda_Y(z)h(z)\mathrm{d}z.\end{aligned}$$

Because $\lambda_Q(\underline{w}) = \lambda_Q(\overline{w}) = 0$, we have

$$\int_{\underline{w}}^{\overline{w}} \lambda_Y(z)h(z)\mathrm{d}z = 0.$$

Because λ_Y is non-increasing, it must be that $\lambda_Y(\underline{w}) \geq 0$ and $\lambda_Y(\overline{w}) \leq 0$.

We argue that $\lambda_X = 0$. Suppose, to the contrary, that $\lambda_X > 0$. Then because $H_{\alpha_i}(z)/h(z)$ is strictly decreasing and $\lambda_Y(z)$ is non-increasing, $\lambda_Q(H^{-1}(\cdot))$ is strictly convex. Hence, $\lambda_Q(z) < 0$ for all $z \in (\underline{w}, \overline{w})$, and therefore $a(z) = 0$ for all $z \in (\underline{w}, \overline{w})$. That is, Q is constant. However, if Q is constant, $X(\overline{w}) = 0 > -C'(\alpha^*)$, a contradiction to that $\lambda_X > 0$. Hence, $\lambda_X = 0$ and therefore

$$\lambda_Q(z) = -\int_{\underline{w}}^{z} \lambda_Y(\zeta)h(\zeta)\mathrm{d}\zeta.$$

Suppose that $\lambda_Y(\underline{w}) = \lambda_Y(\overline{w}) = 0$, then $\lambda_Y(z) = 0$ for all $z \in (\underline{w}, \overline{w})$. Hence, $\lambda_Q(z) = 0$ for all $z \in (\underline{w}, \overline{w})$ and $\eta_Y(\overline{w}) - \eta_Y(\underline{w}) = -\lambda_Y(\overline{w})/2 + \lambda_Y(\underline{w})/2 = 0$. Then

$$(\lambda_0, \lambda_X(z), \lambda_Y(z), \lambda_Q(z), \eta_Y(\overline{w}) - \eta_Y(\underline{w})) = 0, \ \forall z,$$

which is a contradiction to (43). Hence, $\lambda_Y(\underline{w}) > 0$ and $\lambda_Y(\overline{w}) < 0$. Thus, $\lambda_Q(z) < 0$ for all $z \in (\underline{w}, \overline{w})$ and therefore Q is constant. Hence, $Y(z) > 0$ for all $z \in [\underline{w}, \overline{w}]$. This, by (57), implies that λ_Y is constant on $(\underline{w}, \overline{w})$, which is a contradiction to the fact that that $\lambda_Y(\underline{w}) > 0$ and $\lambda_Y(\overline{w}) < 0$. Hence, $\lambda_0 = 1$. ■

This completes the proof of Corollary 3. ■

Before proceeding, we first introduce some notations and proves two technical lemmas (Lemmas 14 and 16) that will be useful for later proof. For the ease of notation, we suppress the dependence of φ^{λ_X}, J^{λ_X}, $\overline{\varphi}^{\lambda_X}$ and $\overline{J}^{\lambda_X}$ on α^*. For each $w \in W$, define

$$m_Y(w) := -\int_{\underline{w}}^{w} \lambda_Y(z)h(z)\mathrm{d}z.$$

It follows from (59) that for any $\underline{z}, \overline{z} \in W$ and $\underline{z} < \overline{z}$, we have

$$\begin{aligned}\lambda_Q(\overline{z}) =& \lambda_Q(\underline{z}) - \int_{\underline{z}}^{\overline{z}} \left[z + \lambda_X \frac{H_{\alpha_i}(z)}{h(z)} + \lambda_Y(z)\right] h(z)\mathrm{d}z,\\ =& \lambda_Q(\underline{z}) - \int_{\underline{z}}^{\overline{z}} \left[\varphi^{\lambda_X}(H(z)) + \lambda_Y(z)\right] h(z)\mathrm{d}z.\end{aligned}$$

If $\underline{z} = \underline{w}$,

$$\lambda_Q(\overline{z}) = \lambda_Q(\underline{w}) - J^{\lambda_X}(H(\overline{z})) + m_Y(\overline{z}). \tag{64}$$

Hence, (63) can be rewritten as, for virtually all $z \in (\underline{w}, \overline{w})$,

$$a(z) = \begin{cases} = 0 & \text{if } \lambda_Q(\underline{w}) + m_Y(z) \leq J^{\lambda_X}(H(z)),\\ \in [0, K] & \text{if } \lambda_Q(\underline{w}) + m_Y(z) = J^{\lambda_X}(H(z)),\\ = K & \text{if } \lambda_Q(\underline{w}) + m_Y(z) \geq J^{\lambda_X}(H(z)).\end{cases}$$

Lemma 14 *For all $t \in [0, 1]$,*

$$\lambda_Q(\underline{w}) + m_Y(H^{-1}(t)) \geq \overline{J}^{\lambda_X}(t).$$

Proof. The proof of Lemma 14 uses the following lemma.

Lemma 15 (Reid) *Suppose that $\lambda_Q(\underline{w}) + m_Y(H^{-1}(t)) = J^{\lambda_X}(t)$ for $t \in \{\underline{t}, \overline{t}\}$. Let $a, b \in \mathbb{R}$ and $l(t) = a + bt$. If $J^{\lambda_X}(t) \geq l(t)$ for all $t \in [\underline{t}, \overline{t}]$,*

$$\lambda_Q(\underline{w}) + m_Y(H^{-1}(t)) \geq l(t),\ \forall t \in [\underline{t}, \overline{t}].$$

Proof. Suppose, to the contrary, that $\lambda_Q(\underline{w}) + m_Y(H^{-1}(t)) < l(t)$ for some $t \in (\underline{t}, \overline{t})$. Then by continuity there exist $\varepsilon > 0$ and $t_1, t_2 \in (\underline{t}, \overline{t})$ such that $\underline{t} < t_1 < t_2 < \overline{t}$, $\lambda_Q(\underline{w}) + m_Y(H^{-1}(\tau)) < l(\tau) - \varepsilon$ for $\tau \in (t_1, t_2)$, and

$$\begin{aligned}\lambda_Q(\underline{w}) + m_Y(H^{-1}(t_1)) = l(t_1) - \varepsilon,\\ \lambda_Q(\underline{w}) + m_Y(H^{-1}(t_2)) = l(t_2) - \varepsilon.\end{aligned}$$

On the one hand, this implies that $\lambda_Y((H^{-1}(\cdot)) = -m_Y'(H^{-1}(\cdot))$ cannot be constant on (t_1, t_2).

On the other hand, $\lambda_Q(\underline{w}) + m_Y(H^{-1}(\tau)) < l(\tau) - \varepsilon < J^{\lambda_X}(\tau)$ for $\tau \in (t_1, t_2)$. Hence, $a(H^{-1}(\tau)) = 0$ for $\tau \in (t_1, t_2)$, which implies that $Y(H^{-1}(\tau)) > 0$ on the interval (t_1, t_2). To see this, note that $Y'(z) = Q(z) - H^{n-1}(z)$ is strictly decreasing if Q is constant. Hence, Y is strictly concave on $(H^{-1}(t_1), H^{-1}(t_2))$. For any $\tau \in (t_1, t_2)$ there exists $\lambda \in (0,1)$ such that $H^{-1}(\tau) = \lambda H^{-1}(t_1) + (1-\lambda)H^{-1}(t_2)$. By strict concavity, $Y(H^{-1}(\tau)) > \lambda Y(H^{-1}(t_1)) + (1-\lambda)Y(H^{-1}(t_2)) \geq 0$. By (60), $Y(H^{-1}(\cdot)) > 0$ on (t_1, t_2) implies that $\lambda_Y(H^{-1}(\cdot))$ is constant on (t_1, t_2), a contradiction. ■

By (64) and Lemma 12, $\lambda_Q(\underline{w}) + m_Y(H^{-1}(0)) = 0 = J^{\lambda_X}(0)$ and $\lambda_Q(\underline{w}) + m_Y(H^{-1}(1)) = J^{\lambda_X}(1)$. Hence, by Lemma 15, Lemma 14 holds. ■

Lemma 16 *If* $K > \overline{K} := \max_{z \in W}(n-1)H(z)^{n-2}h(z)$,

$$\lambda_X \leq \overline{\lambda}_X := \left[\min_{z \in W} \frac{\partial}{\partial z}\left[-\frac{H_\alpha(z)}{h(z)}\right]\right]^{-1}.$$

Proof. The proof of Lemma 16 uses Lemmas 17 and 18.

Lemma 17 (interior solution) *Suppose that* $a(z) \in (0, K)$ *for* $z \in (\underline{z}, \overline{z})$, *then* $\lambda_Y(z) = -\varphi^{\lambda_X}(H(z))$ *for virtually every* $z \in (\underline{z}, \overline{z})$.

Proof. If $a(z) \in (0, K)$ for $z \in (\underline{z}, \overline{z})$, $\lambda_Q(\underline{w}) + m_Y(z) = J^{\lambda_X}(H(z))$ for virtually every $z \in (\underline{z}, \overline{z})$. Differentiating this equality with respect to z yields for virtually every $z \in (\underline{z}, \overline{z})$:

$$-\lambda_Y(z)h(z) = \varphi^{\lambda_X}(H(z))h(z),$$

Because $h > 0$, $-\lambda_Y(z) = \varphi^{\lambda_X}(H(z))$ for virtually every $z \in (\underline{z}, \overline{z})$. ■

Lemma 18 (a(z)=K) *Suppose that* $a(z) = K$ *on* $(\underline{z}, \overline{z})$ *with* $\underline{z} < \overline{z}$ *and let* $(\underline{z}, \overline{z})$ *be chosen maximally. Then*

$$\lambda_Q(z) = 0,$$
$$\lambda_Q(\underline{w}) + m_Y(z) = J^{\lambda_X}(H(z)),$$

for $z = \underline{z}$ *if* $\underline{z} > \underline{w}$, *and* $z = \overline{z}$ *if* $\overline{z} < \overline{w}$. *Furthermore,*

$$\varphi^{\lambda_X}(H(\underline{z})) + \lambda_Y(\underline{z}^-) \leq 0, \text{ if } \underline{z} > \underline{w},$$
$$\varphi^{\lambda_X}(H(\overline{z})) + \lambda_Y(\overline{z}^+) \geq 0, \text{ if } \overline{z} < \overline{w}.$$

Proof. Because $a(z) = K$ on $(\underline{z}, \overline{z})$,

$$\lambda_Q(\underline{w}) + m_Y(z) \geq J^{\lambda_X}(H(z)), \text{ v.e. } z \in (\underline{z}, \overline{z}).$$

Suppose that $\underline{z} > \underline{w}$ and let $S_- := \{z < \underline{z} | a(z) < K\}$. Because $(\underline{z}, \overline{z})$ is chosen maximally and Q is absolutely continuous, there exists a sequence $\{z_k\} \in S_-$ converging to $\underline{z}$ with $\lambda_Q(\underline{w}) + m_Y(z_k) \leq J^{\lambda_X}(H(z_k))$ for all k. By continuity, if $\underline{z} > \underline{w}$, $\lambda_Q(\underline{w}) + m_Y(\underline{z}) = J^{\lambda_X}(H(\underline{z}))$, and therefore $\lambda_Q(\underline{z}) = \lambda_Q(\underline{w}) + m_Y(\underline{z}) - J^{\lambda_X}(H(\underline{z})) = 0$. A similar argument proves that $\lambda_Q(\overline{z}) = 0$ and $\lambda_Q(\overline{w}) + m_Y(\overline{z}) = J^{\lambda_X}(H(\overline{z}))$ if $\overline{z} < \overline{w}$.

If $\underline{z} > \underline{w}$, for virtually all $z \in S_-$,

$$\begin{aligned} 0 &= \lambda_Q(\underline{z}) \\ &= \lambda_Q(z) - \int_z^{\underline{z}} \left[\varphi^{\lambda_X}(H(\zeta)) + \lambda_Y(\zeta)\right] h(\zeta) \mathrm{d}\zeta \\ &\leq - \int_z^{\underline{z}} \left[\varphi^{\lambda_X}(H(\zeta)) + \lambda_Y(\zeta)\right] h(\zeta) \mathrm{d}\zeta. \end{aligned}$$

Thus, there exists a sequence $\{z_k\} \in S_-$ converging to $\underline{z}$ such that

$$\int_{z_k}^{\underline{z}} \left[\varphi^{\lambda_X}(H(\zeta)) + \lambda_Y(\zeta)\right] h(\zeta) \mathrm{d}\zeta \leq 0, \ \forall k.$$

Hence,

$$\varphi^{\lambda_X}(H(\underline{z})) + \lambda_Y(\underline{z}^-) \leq 0, \ \text{if } \underline{z} > \underline{w}.$$

A similar argument proves that

$$\varphi^{\lambda_X}(H(\overline{z})) + \lambda_Y(\overline{z}^+) \geq 0, \ \text{if } \overline{z} < \overline{w}.$$

■

Suppose, to the contrary, that $\lambda_X > \overline{\lambda}_X$. Then $\varphi^{\lambda_X}(H(z))$ is strictly decreasing. Suppose that there exists an interval $(\underline{z}, \overline{z})$ such that $a(z) \in (0, K)$ for $z \in (\underline{z}, \overline{z})$. Then, by Lemma 17, $\lambda_Y(z) = -\varphi^{\lambda_X}(H(z))$ for virtually every $z \in (\underline{z}, \overline{z})$. Thus, λ_Y is strictly increasing on $(\underline{z}, \overline{z})$, which is a contradiction to the fact that λ_Y is non-increasing. Because a is piecewise continuous by assumption, $a(z) \in \{0, K\}$ for almost every $z \in W$.

Suppose that there exists an interval $(\underline{z}, \overline{z})$ such that $a(z) = K$ on $(\underline{z}, \overline{z})$ and let $(\underline{z}, \overline{z})$ be chosen maximally. Then $Y'(z) = Q(z) - H(z)^{n-1}$ which is strictly increasing because $K > \max_{z \in W}(n-1)H(z)^{n-2}h(z)$, and therefore $Y(z)$ is strictly convex on $[\underline{z}, \overline{z}]$. This implies that $Y(z) > 0$ on $[\underline{z}, \overline{z}]$ except at most one point. Suppose that $Y(z) > 0$ for all $z \in (\underline{z}, \overline{z})$. Then λ_Y is constant on $(\underline{z}, \overline{z})$, and $\lambda_Q(H^{-1}(t)) = \lambda_Q(\underline{z}) - \int_{H(\underline{z})}^t \left[\varphi^{\lambda_X}(\tau) + \lambda_Y(H^{-1}(\tau))\right] \mathrm{d}\tau$ is strictly convex on $(H(\underline{z}), H(\overline{z}))$. By Lemma 18 and the fact that $\lambda_Q(\underline{w}) = \lambda_Q(\overline{w}) = 0$, $\lambda_Q(\underline{z}) = \lambda_Q(\overline{z}) = 0$. Then the strict convexity of $\lambda_Q(H^{-1}(t))$ implies that $\lambda_Q(z) < 0$ for all $z \in (\underline{z}, \overline{z})$. However, $a(z) = K$ on $(\underline{z}, \overline{z})$ implies that $\lambda_Q(z) \geq 0$ for virtually every $z \in (\underline{z}, \overline{z})$, a contradiction. Hence, there exists a unique $z_0 \in (\underline{z}, \overline{z})$ such that $Y(z_0) = 0$, and therefore $Y(\underline{z}) > 0$ and $Y(\overline{z}) > 0$. Because $Y(\underline{w}) = Y(\overline{w}) = 0$, we have $\underline{w} < \underline{z} < \overline{z} < \overline{w}$. Note that this also implies that λ_Y is constant in a neighborhood of and therefore continuous at $z \in \{\underline{z}, \overline{z}\}$. By Lemma 18, we have

$$\varphi^{\lambda_X}(H(\underline{z})) + \lambda_Y(\underline{z}) \leq 0,$$
$$\varphi^{\lambda_X}(H(\overline{z})) + \lambda_Y(\overline{z}) \geq 0.$$

Hence,

$$\lambda_Y(\underline{z}) \leq -\varphi^{\lambda_X}(H(\underline{z})) < -\varphi^{\lambda_X}(H(\overline{z})) \leq \lambda_Y(\overline{z}),$$

where the second inequality holds because φ^{λ_X} is strictly decreasing and H is strictly increasing. However, this is a contradiction to that λ_Y is non-increasing. Hence, $a(z) = 0$ for almost all $z \in W$.

Because Q is absolutely continuous, this implies that Q is constant on W. However, by Lemma 2, $X(\overline{w}) = 0 > -C'(\alpha^*)$ when Q is constant on W, which implies that $\lambda_X = 0$, a contradiction to the supposition that $\lambda_X > \bar{\lambda}_X > 0$. Hence, $\lambda_X \leq \bar{\lambda}_X$. ■

A.2.3 Proof of Theorem 2

Let $\{Q^\nu\}$ be a sequence of optimal solutions of ($\mathcal{P}^K$-α^*) where $K = K_\nu > \overline{K}$ for each ν, $\overline{K}$ is defined in Lemma 16, and $K_\nu \to \infty$ as $\nu \to \infty$. After taking a subsequence, we can assume that $\{Q^\nu\}$ converges pointwise. Let Q^∞ denote the pointwise limit of this sequence. By Lemma 10, Q^∞ is an optimal solution of $\mathcal{P}'$-α^*. Denote the corresponding joint variables associated with Q^ν by λ^ν. By Lemma 16, $\{\lambda_X^\nu\}$ is bounded for ν sufficiently large. After taking a subsequence, we can assume that $\{\lambda_X^\nu\}$ converges, and let $\lambda_X^\infty := \lim_{\nu\to\infty} \lambda_X^\nu$.

For brevity, let φ^ν (or φ^∞) denote $\varphi^{\lambda_X^\nu}$ (or $\varphi^{\lambda_X^\infty}$), J^ν (or J^∞) denote $J^{\lambda_X^\nu}$ (or $J^{\lambda_X^\infty}$), $\overline{\varphi}^\nu$ (or $\overline{\varphi}^\infty$) denote $\overline{\varphi}^{\lambda_X^\nu}$ (or $\overline{\varphi}^{\lambda_X^\infty}$), and $\overline{J}^\nu$ (or $\overline{J}^\infty$) denote $\overline{J}^{\lambda_X^\nu}$ (or $\overline{J}^{\lambda_X^\infty}$).

Because Q^ν satisfies (IA$'$) with equality for all ν and Q^∞ is the pointwise limit of $\{Q^\nu\}$, Q^∞ satisfies (IA$'$) with equality. By a similar argument, $Y^\infty(\underline{w}) = 0$. Lemmas 20 and 21 below show that Q^∞ satisfies the two pooling properties when $\lambda_X = \lambda_X^\infty$. Finally, Lemma 22 below proves that that $\lambda_X^\infty = \lambda_X^*$, where $\lambda_X^* > 0$ is such that inequality (12) holds. By the arguments in Section 3.2, this completes the proof Theorem 2.

Before introducing and proving Lemmas 20, 21 and 22, we first prove the following technical lemma, which is used in the proofs of Lemmas 20 and 21.

Lemma 19 *The following four statments are true.*

1. *The sequence $\{\varphi^\nu\}$ is uniformly convergent with limit φ^∞.*

2. *The sequence $\{J^\nu\}$ is uniformly convergent with limit J^∞.*

3. *The sequence $\{\varphi^{\nu\prime}\}$ is uniformly convergent with limit $\varphi^{\infty\prime}$.*

4. *The sequence $\{\overline{J}^\nu\}$ is uniformly convergent with limit $\overline{J}^\infty$.*

Proof. Let $\gamma_1 := \max_{z\in W} |H_{\alpha_i}(z)/h(z)| > 0$,

$$\gamma_2 := \max_{z\in W} \left| \frac{\partial}{\partial z} \left[-\frac{H_{\alpha_i}(z)}{h(z)} \right] \right| > 0,$$

and $\gamma_3 := \max_{z\in W} 1/h(z) > 0$. Here γ_1, γ_2 and γ_3 are well define because H and h are twice continuously differentiable and W is compact.

1. $$|\varphi^\nu(t) - \varphi^\infty(t)| = |\lambda_X^\nu - \lambda_X^\infty| \left| \frac{H_{\alpha_i}(H^{-1}(t))}{h(H^{-1}(t))} \right| \le \gamma_1 |\lambda_X^\nu - \lambda_X^\infty| \to 0,$$

   as $\nu \to \infty$. Hence, the sequence $\{\varphi^\nu\}$ is uniformly convergent with limit φ^∞.

2. $$\begin{aligned} |J^\nu(t) - J^\infty(t)| &= \int_0^t |\varphi^\nu(\tau) - \varphi^\infty(\tau)| \, \mathrm{d}\tau \\ &\le t\gamma_1 |\lambda_X^\nu - \lambda_X^\infty| \\ &\le \gamma_1 |\lambda_X^\nu - \lambda_X^\infty| \to 0, \end{aligned}$$

   as $\nu \to \infty$. Hence, the sequence $\{J^\nu\}$ is uniformly convergent with limit J^∞.

3. $$|\varphi^{\nu\prime}(t) - \varphi^{\infty\prime}(t)| = |\lambda_X^\nu - \lambda_X^\infty| \left| \frac{\partial}{\partial z} \left[\frac{H_{\alpha_i}(H^{-1}(t))}{h(H^{-1}(t))} \right] \frac{1}{h(H^{-1}(t))} \right| \le \gamma_2\gamma_3 |\lambda_X^\nu - \lambda_X^\infty| \to 0,$$

   as $\nu \to \infty$. Hence, the sequence $\{\varphi^{\nu\prime}\}$ is uniformly convergent with limit $\varphi^{\infty\prime}$.

4. Because the sequence $\{J^\nu\}$ is uniformly convergent with limit J^∞, for any $\varepsilon > 0$, there exists $\overline{\nu} > 0$ such that for all $\nu > \overline{\nu}$, $|J^\infty(t) - J^\nu(t)| \le \varepsilon$ for all $t \in [0,1]$. Fix $t \in [0,1]$. Let $t_1, t_2, \beta \in [0,1]$ be such that $\beta t_1 + (1-\beta)t_2 = t$. Then for any $\nu > \overline{\nu}$

   $$\begin{aligned} \overline{J}^\infty(t) \le& \beta J^\infty(t_1) + (1-\beta)J^\infty(t_2) \\ \le& \beta J^\nu(t_1) + (1-\beta)J^\nu(t_2) + \beta|J^\infty(t_1) - J^\nu(t_1)| + (1-\beta)|J^\infty(t_2) - J^\nu(t_2)| \\ \le& \beta J^\nu(t_1) + (1-\beta)J^\nu(t_2) + \varepsilon \end{aligned}$$

Hence,

$$\overline{J}^{\infty}(t) \leq \min\{\beta J^{\nu}(t_1) + (1-\beta)J^{\nu}(t_2) | \beta t_1 + (1-\beta)t_2 = t\} + \varepsilon = \overline{J}^{\nu}(t) + \varepsilon.$$

Similarly, we can show that $\overline{J}^{\nu}(t) \leq \overline{J}^{\infty}(t) + \varepsilon$. Hence, $|\overline{J}^{\infty}(t) - \overline{J}^{\nu}(t)| \leq \varepsilon$. Because this inequality holds for any $t \in [0,1]$, the sequence $\{\overline{J}^{\nu}\}$ is uniformly convergent with limit $\overline{J}^{\infty}$.

■

Lemma 20 (the first pooling property) *Suppose that $J^{\infty}(H(z)) > \overline{J}^{\infty}(H(z))$ for $z \in (\underline{z}, \overline{z})$ with $\underline{z} < \overline{z}$ and let $(\underline{z}, \overline{z})$ be chosen maximally. Then Q^{∞} is constant on $(\underline{z}, \overline{z})$.*

Proof. For each $0 < \delta < (\overline{z} - \underline{z})/2$, let $\varepsilon(\delta) := \min_{z \in [\underline{z}+\delta, \overline{z}-\delta]}\{J^{\infty}(H(z)) - \overline{J}^{\infty}(H(z))\}$. Then $\varepsilon(\delta)$ is non-increasing in δ and converges to zero as δ converges to zero. Fix $0 < \delta_0 < (\overline{z} - \underline{z})/2$. Let $\varepsilon_0 := \frac{1}{4}\varepsilon(\delta_0) > 0$. There exist $0 < \delta_1 < \delta_2 < \delta_0$ such that $\varepsilon(\delta_1) = \varepsilon_0$ and $\varepsilon(\delta_2) = 2\varepsilon_0$. We claim that there exists $\overline{\nu}$ such that for all $\nu > \overline{\nu}$,

$$J^{\nu}(H(z)) - \bar{J}^{\nu}(H(z)) \geq \frac{7\varepsilon_0}{2} \text{ if } z \in [\underline{z} + \delta_0, \overline{z} - \delta_0], \tag{65}$$

$$J^{\nu}(H(z)) - \bar{J}^{\nu}(H(z)) \geq \frac{\varepsilon_0}{2} \text{ if } z \in [\underline{z} + \delta_1, \overline{z} - \delta_1], \tag{66}$$

$$J^{\nu}(H(z)) - \bar{J}^{\nu}(H(z)) \leq \frac{5\varepsilon_0}{2} \text{ if } J^{\infty}(H(z)) - \overline{J}^{\infty}(H(z)) \leq 2\varepsilon_0. \tag{67}$$

We begin by prove (66). Because the sequence $\{J^{\nu}\}$ is uniformly convergent with limit J^{∞}, there exists $\overline{\nu}$ such that for all $\nu > \overline{\nu}$, $|J^{\nu}(t) - J^{\infty}(t)| < \varepsilon_0/8$ for all $t \in [0,1]$. Let $t \in [H(\underline{z} + \delta_1), H(\overline{z} - \delta_1)]$. Then, by construction, $J^{\infty}(t) - \overline{J}^{\infty}(t) \geq \varepsilon_0$. Hence, there exists $\beta, t_1, t_2 \in [0,1]$ such that $\beta t_1 + (1-\beta)t_2 = t$ and $\beta J^{\infty}(t_1) + (1-\beta)J^{\infty}(t_2) < J^{\infty}(t) - 3\varepsilon_0/4$. Then

$$\begin{aligned}
\bar{J}^{\nu}(t) \leq& \beta J^{\nu}(t_1) + (1-\beta)J^{\nu}(t_2) \\
\leq& \beta J^{\infty}(t_1) + \beta|J^{\nu}(t_1) - J^{\infty}(t_1)| + (1-\beta)J^{\infty}(t_2) + (1-\beta)|J^{\nu}(t_2) - J^{\infty}(t_2)| \\
\leq& \beta J^{\infty}(t_1) + (1-\beta)J^{\infty}(t_2) + \frac{\varepsilon_0}{8} \\
\leq& J^{\infty}(t) - \frac{3\varepsilon_0}{4} + \frac{\varepsilon_0}{8} \\
\leq& J^{\nu}(t) + |J^{\nu}(t) - J^{\infty}(t)| - \frac{5\varepsilon_0}{8} \\
\leq& J^{\nu}(t) + \frac{\varepsilon}{8} - \frac{5\varepsilon_0}{8} \\
=& J^{\nu}(t) - \frac{\varepsilon_0}{2}.
\end{aligned}$$

Hence, $J^\nu(H(z)) - \bar{J}^\nu(H(z)) \geq \varepsilon/2$ for all $z \in [\underline{z}+\delta_1, \overline{z}-\delta_1]$. A similar argument proves (65). To prove (67), let z be such that $J^\infty(H(z)) - \overline{J}^\infty(H(z)) \leq 2\varepsilon_0$. For any $\beta, t_1, t_2 \in [0,1]$ such that $\beta t_1 + (1-\beta)t_2 = t$, where $t = H(z)$, we have

$$\begin{aligned}
&\beta J^\nu(t_1) + (1-\beta)J^\nu(t_2) \\
\geq&\beta J^\infty(t_1) - \beta|J^\nu(t_1) - J^\infty(t_1)| + (1-\beta)J^\infty(t_2) - (1-\beta)|J^\nu(t_2) - J^\infty(t_2)| \\
\geq&\beta J^\infty(t_1) + (1-\beta)J^\infty(t_2) - \frac{\varepsilon_0}{8} \\
\geq&\overline{J}^\infty(t) - \frac{\varepsilon_0}{8} \\
=&J^\infty(t) - [J^\infty(t) - \overline{J}^\infty(t)] - \frac{\varepsilon_0}{8} \\
\geq&J^\infty(t) - 2\varepsilon_0 - \frac{\varepsilon_0}{8}, \\
\geq&J^\nu(t) - |J^\nu(t) - J^\infty(t)| - 2\varepsilon_0 - \frac{\varepsilon_0}{8}, \\
\geq&J^\nu(t) - \frac{\varepsilon_0}{8} - 2\varepsilon_0 - \frac{\varepsilon_0}{8}, \\
\geq&J^\nu(t) - \frac{5\varepsilon_0}{2}.
\end{aligned}$$

Hence,

$$\overline{J}^\nu(t) := \min\{\beta J^\nu(t_1) + (1-\beta)J^\nu(t_2) | \beta, t_1, t_2 \in [0,1] \text{ and } \beta t_1 + (1-\beta)t_2 = t\} \geq J^\nu(t) - \frac{5\varepsilon_0}{2}.$$

We abuse notation a bit and let h denote the probability measure on W corresponding to $H(w)$. Because $\varepsilon(\delta_1) = \varepsilon_0$ and $\varepsilon(\delta_2) = 2\varepsilon_0$, by continuity,

$$\begin{aligned}
h_\delta := \min\big\{&h(\{z \in [\underline{z}+\delta_1, \underline{z}+\delta_2] | J^\infty(H(z)) - \overline{J}^\infty(H(z)) \leq 2\varepsilon_0\}), \\
&h(\{z \in [\overline{z}-\delta_2, \overline{z}-\delta_1] | J^\infty(H(z)) - \overline{J}^\infty(H(z)) \leq 2\varepsilon_0\})\big\} > 0.
\end{aligned}$$

Fix $\nu > \overline{\nu}$ such that $K_\nu > 1/h_\delta$. Suppose that there exists $(b_1, b_2) \subset (\underline{z}+\delta_0, \overline{z}-\delta_0)$ such that $a^\nu(z) > 0$ for $z \in (b_1, b_2)$. Then $\lambda_Q(b_1), \lambda_Q(b_2) \geq 0$. Because λ_Y^ν is non-increasing, we have $\lambda_Y^\nu(b_2) \leq \lambda_Y^\nu(b_1)$. By (66), $\overline{J}^\nu$ is linear and therefore $\overline{\varphi}^\nu$ is constant on $[\underline{z}+\delta_1, \overline{z}-\delta_1]$. Hence, we have either $-\lambda_Y^\nu(b_2) \geq \overline{\varphi}^\nu(H(z))$ for all $z \in [\underline{z}+\delta_1, \overline{z}-\delta_1]$, or $-\lambda_Y^\nu(b_1) \leq \overline{\varphi}^\nu(H(z))$ for all $z \in [\underline{z}+\delta_1, \overline{z}-\delta_1]$. Assume without loss of generality that $-\lambda_Y^\nu(b_2) \geq \overline{\varphi}^\nu(H(z))$ for all

$z \in [\underline{z} + \delta_1, \overline{z} - \delta_1]$. For any $z \in [\overline{z} - \delta_2, \overline{z} - \delta_1]$ with $J^\infty(H(z)) - \overline{J}^\infty(H(z)) \le 2\varepsilon_0$, we have

$$\begin{aligned}
\lambda_Q^\nu(z) &= \lambda_Q^\nu(b_2) - \int_{b_2}^z [\lambda_Y^\nu(\zeta) + \varphi^\nu(H(\zeta))]\, h(\zeta)\mathrm{d}\zeta \\
&\ge \int_{b_2}^z \overline{\varphi}^\nu(H(\zeta))h(\zeta)\mathrm{d}\zeta - \int_{b_2}^z \varphi^\nu(H(\zeta))h(\zeta)\mathrm{d}\zeta \\
&= \overline{J}^\nu(H(z)) - \overline{J}^\nu(H(b_2)) - J^\nu(H(z)) + J^\nu(H(b_2)) \\
&= J^\nu(H(b_2)) - \overline{J}^\nu(H(b_2)) - \left[J^\nu(H(z)) - \overline{J}^\nu(H(z))\right] \\
&\ge \frac{7\varepsilon_0}{2} - \frac{5\varepsilon_0}{2} = \varepsilon_0 > 0,
\end{aligned}$$

where the first inequality holds because $\lambda_Q^\nu(b_2) \ge 0$ and $-\lambda_Y^\nu(\zeta) \ge -\lambda_Y^\nu(b_2) \ge \overline{\varphi}^\nu(H(\zeta))$ for all $b_2 \le \zeta \le z$. That is, $a^\nu(z) = K_\nu$ for almost every $z \in [\overline{z} - \delta_2, \overline{z} - \delta_1]$ with $J^\infty(H(z)) - \overline{J}^\infty(H(z)) \le 2\varepsilon_0$. However, this is a contradiction to that $K_\nu > 1/h_\delta$ because $0 \le Q \le 1$.

Hence, $a^\nu(z) = 0$ for almost every $z \in [\underline{z} + \delta_0, \overline{z} - \delta_0]$ for ν sufficiently large. Let ν goes to infinity and we have Q^∞ is constant on $[\underline{z} + \delta_0, \overline{z} - \delta_0]$. Because this is true for any $0 < \delta_0 < (\overline{z} - \underline{z})/2$, we have that Q^∞ is constant on $(\underline{z}, \overline{z})$. ■

Lemma 21 (the second pooling property) *Suppose that $Y^\infty(z) > 0$ for all $z \in (\underline{z}, \overline{z})$ with $\underline{z} < \overline{z}$ and let $(\underline{z}, \overline{z})$ be chosen maximally. Then $\overline{\varphi}^\infty$ is constant on $(H(\overline{z}), H(\underline{z}))$.*

Proof. Suppose, to the contrary, that $\overline{\varphi}^\infty(H(\overline{z})^-) > \overline{\varphi}^\infty(H(\underline{z}))$. Because $\overline{\varphi}^\infty$ is non-decreasing and right-continuous, there exists $\overline{\delta} > 0$ such that $\overline{\varphi}^\infty(H(\overline{z} - \delta)) > \overline{\varphi}^\infty(H(\underline{z} + \delta))$ for all $\delta \in (0, \overline{\delta})$. Fix $\delta \in (0, \min\{\overline{\delta}/2, (\overline{z} - \underline{z})/4\})$. Because $\overline{J}^\infty$ is convex and $\overline{\varphi}^\infty$ is not constant on $(\underline{z} + \delta, \overline{z} - \delta)$, we have

$$\overline{J}^\infty(H(z)) < \overline{J}^\infty(H(\underline{z}+\delta)) + [H(z) - H(\underline{z}+\delta)]\frac{\overline{J}^\infty(H(\overline{z} - \delta)) - \overline{J}^\infty(H(\underline{z} + \delta))}{H(\overline{z} - \delta) - H(\underline{z} + \delta)}, \ \forall z \in (\underline{z}+\delta, \overline{z}-\delta).$$

Let

$$\varepsilon_1 := \min_{z \in [\underline{z}+2\delta, \overline{z}-2\delta]} \left\{ \overline{J}^\infty(H(\underline{z} + \delta)) + [H(z) - H(\underline{z} + \delta)]\frac{\overline{J}^\infty(H(\overline{z} - \delta)) - \overline{J}^\infty(H(\underline{z} + \delta))}{H(\overline{z} - \delta) - H(\underline{z} + \delta)} - \overline{J}^\infty(H(z)) \right\} > 0,$$

$$\varepsilon_2 := \min_{z \in [\underline{z}+\delta, \overline{z}-\delta]} Y^\infty(z) > 0,$$

and

$$M_1 := 2 \max_{z \in [\underline{z}, \overline{z}]} |\varphi^\infty(H(z))| > 0.$$

Because the sequence $\{\overline{J}^\nu\}$ is uniformly convergent with limit $\overline{J}^\infty$, the sequence $\{Y^\nu\}$ is uniformly convergent with limit Y^∞, and the sequence $\{\varphi^\nu\}$ is uniformly convergent with limit φ^∞, there exists $\overline{\nu}$ such that for $\nu > \overline{\nu}$, $|Y^\nu(z) - Y^\infty(z)| < \varepsilon_2/2$ for all $z \in W$, $|\overline{J}^\infty(t) - \overline{J}^\nu(t)| \leq \varepsilon_1/8$ for all $t \in [0,1]$, and $|\varphi^\nu(t) - \varphi^\infty| \leq M_1/2$ for all $t \in [0,1]$. Then for all $\nu > \overline{\nu}$ and $z \in [\underline{z} + 2\delta, \overline{z} - 2\delta]$ we have

$$
\begin{aligned}
&\overline{J}^\nu(H(\underline{z}+\delta)) + [H(z) - H(\underline{z}+\delta)]\frac{\overline{J}^\nu(H(\overline{z}-\delta)) - \overline{J}^\nu(H(\underline{z}+\delta))}{H(\overline{z}-\delta) - H(\underline{z}+\delta)} - \overline{J}^\nu(H(z)) \\
\geq &\overline{J}^\infty(H(\underline{z}+\delta)) - \left|\overline{J}^\infty(H(\underline{z}+\delta)) - \overline{J}^\nu(H(\underline{z}+\delta))\right| - \overline{J}^\infty(H(z)) - \left|\overline{J}^\infty(H(z)) - \overline{J}^\nu(H(z))\right| \\
&+ \frac{H(z) - H(\underline{z}+\delta)}{H(\overline{z}-\delta) - H(\underline{z}+\delta)}\Big[\overline{J}^\infty(H(\overline{z}-\delta)) - \left|\overline{J}^\infty(H(\overline{z}-\delta)) - \overline{J}^\nu(H(\overline{z}-\delta))\right| \\
&-\overline{J}^\infty(H(\underline{z}+\delta)) - \left|\overline{J}^\infty(H(\underline{z}+\delta)) - \overline{J}^\nu(H(\underline{z}+\delta))\right|\Big] \\
\geq &\overline{J}^\infty(H(\underline{z}+\delta)) + [H(z) - H(\underline{z}+\delta)]\frac{\overline{J}^\infty(H(\overline{z}-\delta)) - \overline{J}^\infty(H(\underline{z}+\delta))}{H(\overline{z}-\delta) - H(\underline{z}+\delta)} - \overline{J}^\infty(H(z)) - \frac{\varepsilon_1}{2} \\
\geq &\frac{\varepsilon_1}{2}.
\end{aligned}
\tag{68}
$$

For all $\nu > \overline{\nu}$ and $z \in [\underline{z}+\delta, \overline{z}-\delta]$, we have

$$
Y^\nu(z) \geq Y^\infty(z) - |Y^\nu(z) - Y^\infty(z)| \geq \frac{\varepsilon_2}{2} > 0.
$$

Because $Y^\nu(z) > 0$ on $[\underline{z}+\delta, \overline{z}-\delta]$, λ_Y is constant and therefore $m_Y^\nu(H^{-1}(\cdot))$ is affine on $[H(\underline{z}+\delta), H(\overline{z}-\delta)]$. By Lemmas 12 and 14, we have $m_Y^\nu(z) \geq \overline{J}^\nu(H(z))$ for all $z \in W$. In particular, $m_Y^\nu(\underline{z}+\delta) \geq \overline{J}^\nu(H(\underline{z}+\delta))$ and $m_Y^\nu(\overline{z}-\delta) \geq \overline{J}^\nu(H(\overline{z}-\delta))$. Hence, for all $z \in [\underline{z}+\delta, \overline{z}-\delta]$,

$$
m_Y^\nu(z) \geq \overline{J}^\nu(H(\underline{z}+\delta)) + [H(z) - H(\underline{z}+\delta)]\frac{\overline{J}^\nu(H(\overline{z}-\delta)) - \overline{J}^\nu(\underline{z}+\delta)}{H(\overline{z}-\delta) - H(\underline{z}+\delta)}.
$$

For all $\nu > \overline{\nu}$ and $z \in [\underline{z}+\delta, \overline{z}-\delta]$, we have

$$
\varphi^\nu(H(z)) \leq \varphi^\infty(H(z)) + |\varphi^\nu(H(z)) - \varphi^\infty(H(z))| \leq \frac{M_1}{2} + \frac{M_1}{2} = M_1.
$$

Finally, let $M_3 := \min_{z\in[\underline{z},\overline{z}]} h(z) > 0$ and

$$
M_2 := \left|\frac{\overline{J}^\nu(H(\overline{z}-\delta)) - \overline{J}^\nu(H(\underline{z}+\delta))}{H(\overline{z}-\delta) - H(\underline{z}+\delta)}\right| > 0.
$$

Fix $\nu > \overline{\nu}$ such that $K_\nu > 1/\min\{\underline{z}+2\delta+\varepsilon_1/8M_1M_3, \underline{z}+2\delta+\varepsilon_1/8M_2M_3, \overline{z}-\delta\}$. Suppose that $J^\nu(H(z_0)) = \overline{J}^\nu(H(z_0))$ for some $z_0 \in [\underline{z}+2\delta, \overline{z}-2\delta]$. Then $K_\nu > 1/\min\{z_0+\varepsilon_1/8M_1M_3, z_0+\varepsilon_1/8M_2M_3, \overline{z}-\delta\}$. For all $z \in (z_0, \overline{z}-\delta)$ such that $H(z) - H(z_0) \leq \min\{\varepsilon_1/8M_1, \varepsilon_2/8M_2\}$

we have

$$\begin{aligned}
&J^{\nu}(H(z))\\
=&J^{\nu}(H(z_0))+\int_{H(z_0)}^{H(z)}\varphi^{\nu}(\tau)\mathrm{d}\tau\\
\leq&\overline{J}^{\nu}(H(z_0))+M_1(H(z)-H(z_0))\\
\leq&\overline{J}^{\nu}(H(\underline{z}+\delta))+[H(z)-H(\underline{z}+\delta)-H(z)+H(z_0)]\frac{\overline{J}^{\nu}(H(\overline{z}-\delta))-\overline{J}^{\nu}(H(\underline{z}+\delta))}{H(\overline{z}-\delta)-H(\underline{z}+\delta)}-\frac{\varepsilon_1}{2}+\frac{\varepsilon_1}{8}\\
\leq&m_Y^{\nu}(z)+[H(z)-H(z_0)]M_2-\frac{\varepsilon_1}{2}+\frac{\varepsilon_1}{8}\\
\leq&m_Y^{\nu}(z)+\frac{\varepsilon_1}{8}-\frac{\varepsilon_1}{2}+\frac{\varepsilon_1}{8}=m_Y^{\nu}(z)-\frac{\varepsilon_1}{4},
\end{aligned}$$

where the second inequality holds by (68). Hence, $J^{\nu}(H(z))<m_Y^{\nu}(z)$ and therefore $a^{\nu}(z)=K_{\nu}$ for all $z\in(z_0,\overline{z}-\delta)$. Because $H(z)-H(z_0)\leq M_3(z-z_0)$, we have $a^{\nu}(z)=K_{\nu}$ for all $z\in(z_0,\min\{z_0+\varepsilon_1/8M_1M_3,z_0+\varepsilon_1/8M_2M_3,\overline{z}-\delta\})$, a contradiction to the fact that $K_{\nu}>1/\min\{z_0+\varepsilon_1/8M_1M_3,z_0+\varepsilon_1/8M_2M_3,\overline{z}-\delta\}$ because $0\leq Q\leq 1$.

Hence, $J^{\nu}(H(z))>\overline{J}^{\nu}(H(z))$ for all $z\in[\underline{z}+2\delta,\overline{z}-2\delta]$. This implies that $\overline{\varphi}^{\nu}$ is constant on $[H(\underline{z}-2\delta),H(\overline{z}+2\delta)]$. Clearly, $\{\overline{\varphi}^{\nu}\}$ converges uniformly on $[H(\underline{z}-2\delta),H(\overline{z}+2\delta)]$ and $\lim_{\nu\to\infty}\overline{\varphi}^{\nu}$ is constant on $[H(\underline{z}-2\delta),H(\overline{z}+2\delta)]$. Because, on $[H(\underline{z}-2\delta),H(\overline{z}+2\delta)]$, $\{\overline{J}^{\nu}\}$ converges uniformly to $\overline{J}^{\infty}$, each $\overline{J}^{\nu}$ is differentiable with derivative $\overline{\varphi}^{\nu}$, and $\{\overline{\varphi}^{\nu}\}$ converges uniformly, we have $\overline{J}^{\infty}$ is differentiable on $[H(\underline{z}-2\delta),H(\overline{z}+2\delta)]$ and its derivative $\overline{\varphi}^{\infty}(t)=\lim_{\nu\to\infty}\overline{\varphi}^{\nu}(t)$ for all $t\in[H(\underline{z}-2\delta),H(\overline{z}+2\delta)]$. Thus, $\overline{\varphi}^{\infty}$ is constant on $[H(\underline{z}-2\delta),H(\overline{z}+2\delta)]$ and $2\delta<\overline{\delta}$, a contradiction to the supposition. Hence, $\overline{\varphi}^{\infty}$ is constant on $(H(\underline{z}),H(\overline{z}))$. ■

The following corollary of Lemma 21 is used in the proof of Lemma 22.

Corollary 4 *Suppose that $\overline{\varphi}^{\infty}(H(z))$ is constant on $(\underline{z},\overline{z})$ with $\underline{z}<\overline{z}$ and let $(\underline{z},\overline{z})$ be chosen maximally. Then $Y^{\infty}(\underline{z})=Y^{\infty}(\overline{z})=0$, i.e.,*

$$\int_{\underline{z}}^{\overline{z}}\left[H(\zeta)^{n-1}-Q^{\infty}(\zeta)\right]h(\zeta)\mathrm{d}\zeta=0.$$

Proof. This is an immediate corollary of Lemma 21. Suppose, to the contrary, that $Y^{\infty}(\overline{z})>0$. Then by Lemma 21, $\overline{\varphi}^{\infty}(H(\cdot))$ is constant in a neighborhood of $\overline{z}$, a contradiction to the fact that $(\underline{z},\overline{z})$ is chosen maximally. Hence, $Y^{\infty}(\overline{z})=0$. Similarly, $Y^{\infty}(\underline{z})=0$. ■

Lemma 22 *$\lambda_X^{\infty}=\lambda_X^{*}$, where $\lambda_X^{*}>0$ is such that inequality (12) holds .*

Proof. For any $\lambda_X>0$, recall that $Q^{+}(\cdot,\lambda_X)$ and $Q^{-}(\cdot,\lambda_X)$ are defined as follows: If $J^{\lambda_X}(H(w|\alpha^*),\alpha^*)>\overline{J}^{\lambda_X}(H(w|\alpha^*),\alpha^*)$ for $w\in(\underline{w},\overline{w})$ and let $(\underline{w},\overline{w})$ be chosen maximally, let

$$Q^+(w,\lambda_X) := \frac{\frac{1}{n}[H(\overline{w}|\alpha^*)^n - H(\underline{w}|\alpha^*)^n]}{H(\overline{w}|\alpha^*) - H(\underline{w}|\alpha^*)},\ \forall w \in (\underline{w},\overline{w}).$$

Otherwise, let $Q^+(w,\lambda^X) := H(w|\alpha^*)^{n-1}$. If $\overline{\varphi}^{\lambda_X}(H(\cdot|\alpha^*),\alpha^*)$ is constant on $(\underline{w},\overline{w})$ with $\underline{w} < \overline{w}$ and let $(\underline{w},\overline{w})$ be chosen maximally, let

$$Q^-(w,\lambda_X) := \frac{\frac{1}{n}[H(\overline{w}|\alpha^*)^n - H(\underline{w}|\alpha^*)^n]}{H(\overline{w}|\alpha^*) - H(\underline{w}|\alpha^*)},\ \forall w \in (\underline{w},\overline{w}).$$

Otherwise, let $Q^-(z,\lambda_X) := H(z|\alpha^*)^{n-1}$. For the ease of notation, denote $Q^+(w,\lambda_X^\infty)$ (or $Q^-(w,\lambda_x^\infty)$) by $Q^+(w)$ (or $Q^-(w)$). Note that Q^+, Q^- and Q^∞ are implementable and non-decreasing, allocate the object with probability one, and satisfy the two pooling properties. Hence, by the arguments in the derivation of inequality (11), for $Q \in \{Q^+, Q^-, Q^\infty\}$,

$$\begin{aligned}&\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \left[z + \lambda_X^\infty \frac{H_{\alpha_i}(z)}{h(z)}\right] Q(z)h(z)\mathrm{d}z + \lambda_X^\infty C'(\alpha^*)\\ =&\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \overline{\varphi}^\infty(H(z))H(z)^{n-1}\mathrm{d}H(z) + \lambda_X^\infty C'(\alpha^*).\end{aligned} \tag{69}$$

Next we show that $Y^+ \le Y^\infty \le Y^-$. Let $S^+ := \{z \in W | Y^+(z) > 0\}$, $S^- := \{z \in W | Y^-(z) > 0\}$ and $S := \{z \in W | Y^\infty(z) > 0\}$. By construction,

$$S^+ = \cup\{(\underline{z},\overline{z}) | J^\infty(H(z)) > \overline{J}^\infty(H(z))\ \forall z \in (\underline{z},\overline{z})\},$$

and

$$S^- = \cup\{(\underline{z},\overline{z}) | \overline{\varphi}^\infty(H(\cdot)) \text{ is constant on } (\underline{z},\overline{z})\}.$$

It follows from Lemma 20 and Lemma 21 that $S^+ \subset S \subset S^-$. If $z \notin S^-$, $Y^+(z) = Y^\infty(z) = Y^-(z) = 0$. If $z \in S^-\backslash S$, $Y^+(z) = Y^\infty(z) = 0 < Y^-(z)$. Consider $z \in S \subset S^-$. There exists an interval $(\underline{z},\overline{z})$ with $\underline{z} < z < \overline{z}$ such that $\overline{\varphi}^\infty(H(\cdot))$ is constant on $(\underline{z},\overline{z})$. Let $(\underline{z},\overline{z})$ be chosen maximally, then by construction $Y^-(\overline{z}) = Y^-(\underline{z}) = 0$. By Corollary 4, $Y^\infty(\overline{z}) = Y^\infty(\underline{z}) = 0$. For any $z \in (\underline{z},\overline{z})$,

$$Y^-(z) - Y^\infty(z) = \int_z^{\overline{z}} \left[Q^\infty(\zeta) - Q^-(\zeta)\right] \mathrm{d}H(\zeta).$$

Then for any $t \in (H(\underline{z}), H(\overline{z}))$,

$$\left[Y^-(H^{-1}(t)) - Y^\infty(H^{-1}(t))\right]' = Q^-(H^{-1}(t)) - Q^\infty(H^{-1}(t)),$$

which is non-increasing on $(H(\underline{z}), H(\overline{z}))$ because Q^∞ is non-decreasing and Q^- is constant on $(\underline{z}, \overline{z})$ by construction. Hence, $Y^-(H^{-1}(t)) - Y^\infty(H^{-1}(t))$ is concave on $(H(\underline{z}), H(\overline{z}))$. Because $Y^-(\underline{z}) - Y^\infty(\underline{z}) = 0$ and $Y^-(\overline{z}) - Y^\infty(\overline{z}) = 0$, we have $Y^-(z) - Y^\infty(z) \geq 0$ for all $z \in (\underline{z}, \overline{z})$. Thus, $Y^-(z) - Y^\infty(z) \geq 0$ for all $z \in S$. If $z \in S \backslash S^+$, $Y^-(z) \geq Y^\infty(z) \geq 0 = Y^+(z)$. Finally, consider $z \in S^+ \subset S$. It suffices to show that $Y^+(z) \leq Y^\infty(z)$. By construction, there exists an interval $(\underline{z}, \overline{z})$ with $\underline{z} < z < \overline{z}$ such that $\overline{J}^\infty(H(z)) < J^\infty(H(z))$ for all $z \in (\underline{z}, \overline{z})$. Let $(\underline{z}, \overline{z})$ be chosen maximally, then by construction $Y^+(\overline{z}) = Y^+(\underline{z}) = 0$. For any $z \in (\underline{z}, \overline{z})$

$$Y^+(z) - Y^\infty(z) = \int_z^{\overline{z}} \left[Q^\infty(\zeta) - Q^-(\zeta)\right] \mathrm{d}H(\zeta) - Y^\infty(\overline{z}).$$

Then for any $t \in (H(\underline{z}), H(\overline{z}))$,

$$\left[Y^+(H^{-1}(t)) - Y^\infty(H^{-1}(t))\right]' = Q^+(H^{-1}(t)) - Q^\infty(H^{-1}(t)),$$

which is constant on $(H(\underline{z}), H(\overline{z}))$ because Q^∞ is constant on $(\underline{z}, \overline{z})$ by Lemma 20 and Q^+ is constant on $(\underline{z}, \overline{z})$ by construction. Hence, $Y^+(H^{-1}(t)) - Y^\infty(H^{-1}(t))$ is affine on $(H(\underline{z}), H(\overline{z}))$. Because $Y^+(\underline{z}) = 0 \leq Y^\infty(\underline{z})$ and $Y^+(\overline{z}) = 0 \leq Y^\infty(\overline{z})$, we have $Y^+(z) - Y^\infty(z) \leq 0$ for all $z \in (\underline{z}, \overline{z})$. Thus, $Y^+(z) - Y^\infty(z) \leq 0$ for all $z \in S^+$.

Furthermore, for any implementable allocation rule Q, we have

$$\begin{aligned}
&\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} zQ(z)\mathrm{d}H(z) \\
=&\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} zY'(z)\mathrm{d}z + \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} zH(z)^{n-1}\mathrm{d}H(z) \\
=&\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} zH(z)^{n-1}\mathrm{d}H(z) - \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} Y(z)\mathrm{d}z,
\end{aligned} \tag{70}$$

where the second line holds by the definition of Y and the third line holds by integration by parts. Hence,

$$\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} zQ^+(z)\mathrm{d}H(z) \geq \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} zQ^\infty(z)\mathrm{d}H(z) \geq \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} zQ^-(z)\mathrm{d}H(z).$$

Because $\lambda_X^\infty \geq 0$ and $-\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} H_\alpha(z)Q^\infty(z)\mathrm{d}z = C'(\alpha^*)$, combining this and (69) yields

$$\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} H_\alpha(z)Q^+(z)\mathrm{d}z \leq -C'(\alpha^*) \leq \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} H_\alpha(z)Q^-(z)\mathrm{d}z. \tag{12}$$

By Lemma 6, there exists an unique $\lambda_X > 0$ such that (12) holds. Hence, $\lambda_X^\infty = \lambda_X^*$. ■

The arguments in Lemma 22 also proves the following lemma:

Lemma 23 *Let $\lambda_X \geq 0$. For any non-decreasing implementable Q that allocates the object with probability one and satisfies the two pooling properties, the following inequality holds:*

$$\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} zQ^+(z,\lambda_X)\mathrm{d}H(z) \geq \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} zQ(z)\mathrm{d}H(z) \geq \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} zQ^-(z,\lambda_X)\mathrm{d}H(z).$$

A.2.4 Proof of Lemma 6

I break the proof into several lemmas. For each λ_X, recall that $Q^+(\cdot,\lambda_X)$ is the "steepest" allocation rule associated with λ_X, and $Q^-(\cdot,\lambda_X)$ is the "least steep" allocation rule associated with λ_X. By Lemma 23, $Q^-(\cdot,\lambda_X)$ gives agents' lower marginal benefit from acquiring information than $Q^+(\cdot,\lambda_X)$ does:

$$\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -\frac{H_\alpha(z)}{h(z)}Q^-(z,\lambda_X)h(z)\mathrm{d}z \leq \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -\frac{H_\alpha(z)}{h(z)}Q^+(z,\lambda_X)\mathrm{d}z.$$

Let $\lambda'_X > \lambda_X$. Lemma 24 below proves that if $\overline{\varphi}^{\lambda_X}$ is constant on $(\underline{t},\overline{t})$ with $\underline{t} < \overline{t}$ and let $(\underline{t},\overline{t})$ be chosen maximally, $J^{\lambda'_X}(t) > \overline{J}^{\lambda'_X}(t)$ for all $t \in (\underline{t}-\delta, \overline{t}+\delta)$ for some $\delta > 0$. Let

$$S := \left\{ t \in [0,1] \,\middle|\, J^{\lambda'_X}(t) > \overline{J}^{\lambda'_X}(t) \right\}. \tag{71}$$

It follows from Lemma 24 and the construction of Q^+ and Q^- that, if $t \notin S$, $Q^+(w,\lambda'_X) = Q^-(w,\lambda_X) = H(w)^{n-1}$, where $w = H^{-1}(t)$. Let

$$I := \left\{ (\underline{t},\overline{t}) | J^{\lambda'_X}(t) > \overline{J}^{\lambda'_X}(t) \text{ on } (\underline{t},\overline{t}) \text{ with } \underline{t} < \overline{t} \text{ and let } (\underline{t},\overline{t}) \text{ be chosen maximally} \right\}.$$

Clearly, different intervals in I are disjoint. Then we have

$$S = \bigsqcup_{(\underline{t},\overline{t})\in I} (\underline{t},\overline{t}).$$

For each $(\underline{t},\overline{t}) \in I$, by construction, $Q^-(\cdot,\lambda_X) - Q^+(\cdot,\lambda'_x)$ is non-decreasing on $[H(\underline{t}), H(\overline{t})]$ because $Q^+(\cdot,\lambda'_x)$ is constant on $[H(\underline{t}), H(\overline{t})]$, and $\int_{H(\underline{t})}^{H(\overline{t})} [Q^-(w,\lambda_X) - Q^+(w,\lambda'_x)]\, h(w)\mathrm{d}w = 0$. Furthermore, by Lemma 24, the set $\left\{ w \in [H(\underline{t}), H(\overline{t})]\, | Q^+(w,\lambda'_X) \neq Q^-(w,\lambda_X) \right\}$ has a positive measure. We prove in Lemma 26 below that these condition implies that

$$\int_{H(\underline{t})}^{H(\overline{t})} -\frac{H_\alpha(w)}{h(w)} \left[Q^-(w,\lambda_X) - Q^+(w,\lambda'_x)\right] h(w)\mathrm{d}w > 0.$$ [18]

Hence, $Q^-(\cdot,\lambda_X)$ gives agents' higher marginal benefit from acquiring information than $Q^+(\cdot,\lambda'_X)$ does:

$$\begin{aligned}
&\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -\frac{H_\alpha(z)}{h(z)} \left[Q^-(z,\lambda_X) - Q^+(z,\lambda'_X)\right] h(z)\mathrm{d}z \\
=&\int_{w=H(t),t\in S} -\frac{H_\alpha(z)}{h(z)} \left[Q^-(z,\lambda_X) - Q^+(z,\lambda'_X)\right] h(z)\mathrm{d}z \\
&+\sum_{(\underline{t},\overline{t})\in I}\int_{H(\underline{t})}^{H(\overline{t})} -\frac{H_\alpha(z)}{h(z)} \left[Q^-(z,\lambda_X) - Q^+(z,\lambda'_X)\right] h(z)\mathrm{d}z > 0
\end{aligned}$$

Furthermore, if λ_X is sufficiently large, $Q^+(\cdot,\lambda_X)$ is constant and $\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -\frac{H_\alpha(z)}{h(z)}Q^+(z,\lambda_X)h(z)\mathrm{d}z = 0 < C'(\alpha^*)$; if $\lambda_X = 0$, $Q^-(\cdot,\lambda_X) = H^{n-1}(\cdot)$ and $\int_{w(0,\alpha^*)}^{w(1,\alpha^*)} -\frac{H_\alpha(z)}{h(z)}Q^-(z,\lambda_X)h(z)\mathrm{d}z > C'(\alpha^*)$ by assumption. Hence, there exists a unique $\lambda_X > 0$ such that inequality (12) holds.

Lemma 24 *Let $\lambda'_X > \lambda_X$. Suppose that $\overline{\varphi}^{\lambda_X}$ is constant on $(\underline{t},\overline{t})$ with $\underline{t} < \overline{t}$ and let $(\underline{t},\overline{t})$ be chosen maximally. Then there exists $\delta > 0$ such that $J^{\lambda'_X}(t) > \overline{J}^{\lambda'_X}(t)$ for all $t \in (\underline{t}-\delta,\overline{t}+\delta)$.*

The proof of Lemma 24 uses the following technical lemma.

Lemma 25 *Let $t \in (0,1)$. If $J^{\lambda_X}(t) = \overline{J}^{\lambda_X}(t)$, $\overline{J}^{\lambda_X}$ is continuously differentiable at t with derivative $\overline{\varphi}^{\lambda_X}(t) = \varphi^{\lambda_X}(t)$ and $\varphi^{\lambda_X\prime}(t) \geq 0$. Furthermore, $J^{\lambda_X}(t) = \overline{J}^{\lambda_X}(t)$ if and only if*

$$J^{\lambda_X}(\tau) \geq (\tau - t)\varphi^{\lambda_X}(t) + J^{\lambda_X}(t),\ \forall \tau \in [0,1]. \tag{72}$$

Proof. For ease of notation, we suppress the dependence of $J, \overline{J}, \varphi$ and $\overline{\varphi}$ on λ_X. Let $t \in (0,1)$. Suppose that $J(t) = \overline{J}(t)$. Suppose, to the contrary, that $\overline{J}$ is not continuously differentiable at t, then $\overline{\varphi}(t^-) < \overline{\varphi}(t^+)$. Then either $\varphi(t) > \overline{\varphi}(t^-)$ or $\varphi(t) < \overline{\varphi}(t^+)$. Assume without loss of generality that $\varphi(t) < \overline{\varphi}(t^+)$. Because φ is continuous and $\overline{\varphi}$ is non-decreasing, there exists $\delta > 0$ such that $\varphi(\tau) < \overline{\varphi}(t^+) \leq \overline{\varphi}(\tau)$ for all $\tau \in (t, t+\delta)$. Then

$$J(t+\delta) = J(t) + \int_t^{t+\delta} \varphi(\tau)\mathrm{d}\tau < \overline{J}(t) + \int_t^{t+\delta} \overline{\varphi}(\tau)\mathrm{d}\tau = \overline{J}(t+\delta),$$

[18] The weak inequality holds by Lemma 7, but we need the strict inequality for this proof to work.

a contradiction. Hence, $\overline{J}$ is continuously differentiable at t. It follows from a similar argument that $\overline{\varphi}(t) = \varphi(t)$ with $\varphi'(t) \geq 0$. Furthermore, for all $\tau \in [0,1]$,

$$J(\tau) \geq \overline{J}(\tau) \geq (\tau - t)\overline{\varphi}(t) + \overline{J}(t) = (\tau - t)\varphi(t) + J(t),$$

where the second inequality holds because $\overline{J}$ is convex.

Suppose that (72) holds. Then $\tau :\to (\tau - t)\varphi(t) + J(t)$ is a convex function below J. Because $\overline{J}$ is the greatest convex function below J, we have

$$\overline{J}(\tau) \geq (\tau - t)\varphi(t) + J(t) \ \forall \tau \in [0,1].$$

If $\tau = t$, then $\overline{J}(t) \geq J(t)$. Hence, $\overline{J}(t) = J(t)$. ∎

Proof of Lemma 24. First, we claim that $\overline{J}^{\lambda_X}(\underline{t}) = J^{\lambda_X}(\underline{t})$ and $\overline{J}^{\lambda_X}(\bar{t}) = J^{\lambda_X}(\bar{t})$. To see that $\overline{J}^{\lambda_X}(\underline{t}) = J^{\lambda_X}(\underline{t})$, suppose to the contrary that $\overline{J}^{\lambda_X}(\underline{t}) < J^{\lambda_X}(\underline{t})$. Then $\overline{\varphi}^{\lambda_X}(t)$ is constant in a neighborhood of $\underline{t}$. A contradiction to that $(\underline{t}, \bar{t})$ is chosen maximally. A similar argument proves that $\overline{J}^{\lambda_X}(\bar{t}) = J^{\lambda_X}(\bar{t})$.

Consider $t \in (\underline{t}, \bar{t})$. Let $\beta \in (0,1)$ be such that $\beta\underline{t} + (1-\beta)\bar{t} = t$. To show that $J^{\lambda'_X}(t) > \overline{J}^{\lambda'_X}(t)$, it suffices to show that

$$J^{\lambda'_X}(t) > \beta J^{\lambda'_X}(\underline{t}) + (1-\beta) J^{\lambda'_X}(\bar{t}).$$

Because $\overline{\varphi}^{\lambda_X}$ is constant on $(\underline{t}, \bar{t})$, $\overline{J}^{\lambda_X}(\underline{t}) = J^{\lambda_X}(\underline{t})$ and $\overline{J}^{\lambda_X}(\bar{t}) = J^{\lambda_X}(\bar{t})$, we have

$$\begin{aligned} J^{\lambda_X}(t) &\geq \overline{J}^{\lambda_X}(t) \\ &= \beta\overline{J}^{\lambda_X}(\underline{t}) + (1-\beta)\overline{J}^{\lambda_X}(\bar{t}) \\ &= \beta J^{\lambda_X}(\underline{t}) + (1-\beta) J^{\lambda_X}(\bar{t}). \end{aligned}$$

Hence,

$$\begin{aligned} 0 \leq & J^{\lambda_X}(t) - \beta J^{\lambda_X}(\underline{t}) - (1-\beta) J^{\lambda_X}(\bar{t}) \\ = & \int_0^t H^{-1}(\tau)\mathrm{d}\tau - \beta \int_0^{\underline{t}} H^{-1}(\tau)\mathrm{d}\tau - (1-\beta)\int_0^{\bar{t}} H^{-1}(\tau)\mathrm{d}\tau \\ & + \lambda_X \left[\int_0^t \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\mathrm{d}\tau - \beta \int_0^{\underline{t}} \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\mathrm{d}\tau - (1-\beta)\int_0^{\bar{t}} \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\mathrm{d}\tau \right]. \end{aligned}$$

Because $H^{-1}(\cdot)$ is strictly increasing, $\int_0^t H^{-1}(\tau)\mathrm{d}\tau$ is strictly convex in t and therefore

$$\int_0^t H^{-1}(\tau)\mathrm{d}\tau - \beta \int_0^{\underline{t}} H^{-1}(\tau)\mathrm{d}\tau - (1-\beta)\int_0^{\bar{t}} H^{-1}(\tau)\mathrm{d}\tau < 0.$$

Hence,

$$\int_0^t \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\mathrm{d}\tau - \beta \int_0^{\underline{t}} \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\mathrm{d}\tau - (1-\beta)\int_0^{\bar{t}} \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\mathrm{d}\tau > 0.$$

Then

$$\begin{aligned}
&J^{\lambda'_X}(t) - \beta J^{\lambda'_X}(\underline{t}) - (1-\beta)J^{\lambda'_X}(\bar{t}) \\
=&\int_0^t H^{-1}(\tau)\mathrm{d}\tau - \beta \int_0^{\underline{t}} H^{-1}(\tau)\mathrm{d}\tau - (1-\beta)\int_0^{\bar{t}} H^{-1}(\tau)\mathrm{d}\tau \\
&+ \lambda'_X \left[\int_0^t \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\mathrm{d}\tau - \beta \int_0^{\underline{t}} \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\mathrm{d}\tau - (1-\beta)\int_0^{\bar{t}} \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\mathrm{d}\tau\right] \\
>&\int_0^t H^{-1}(\tau)\mathrm{d}\tau - \beta \int_0^{\underline{t}} H^{-1}(\tau)\mathrm{d}\tau - (1-\beta)\int_0^{\bar{t}} H^{-1}(\tau)\mathrm{d}\tau \\
&+ \lambda_X \left[\int_0^t \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\mathrm{d}\tau - \beta \int_0^{\underline{t}} \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\mathrm{d}\tau - (1-\beta)\int_0^{\bar{t}} \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\mathrm{d}\tau\right] \\
=&J^{\lambda_X}(t) - \beta J^{\lambda_X}(\underline{t}) - (1-\beta)J^{\lambda_X}(\bar{t}) \\
\geq&0.
\end{aligned}$$

Consider $\underline{t}$. Since $\overline{\varphi}^{\lambda_X}$ is constant on $(\underline{t}, \bar{t})$ with $J^{\lambda_X}(\underline{t}) = \overline{J}^{\lambda_X}(\underline{t})$ and $J^{\lambda_X}(\bar{t}) = \overline{J}^{\lambda_X}(\bar{t})$, by Lemma 25, we have

$$\begin{aligned}
\overline{J}^{\lambda_X}(\bar{t}) =&J^{\lambda_X}(\bar{t}) \\
\geq&(\bar{t}-\underline{t})\varphi^{\lambda_X}(\underline{t}) + J^{\lambda_X}(\underline{t}) \\
=&(\bar{t}-\underline{t})\overline{\varphi}^{\lambda_X}(\underline{t}) + \overline{J}^{\lambda_X}(\underline{t}) = \overline{J}^{\lambda_X}(\bar{t}).
\end{aligned}$$

Hence, $J^{\lambda_X}(\bar{t}) = (\bar{t}-\underline{t})\varphi^{\lambda_X}(\underline{t}) + J^{\lambda_X}(\underline{t})$ or equivalently

$$\int_{\underline{t}}^{\bar{t}} H^{-1}(\tau)\mathrm{d}\tau - (\bar{t}-\underline{t})H^{-1}(\underline{t}) = \lambda_X \left[(\bar{t}-\underline{t})\frac{H_\alpha(H^{-1}(\underline{t}))}{h(H^{-1}(\underline{t}))} - \int_{\underline{t}}^{\bar{t}} \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))}\right].$$

Since $H^{-1}(t)$ is strictly increasing, the left-hand side of the above equality is strictly positive.

Hence for $\lambda'_X > \lambda_X > 0$ we have

$$\int_{\underline{t}}^{\bar{t}} H^{-1}(\tau)\mathrm{d}\tau - (\bar{t}-\underline{t})H^{-1}(\underline{t}) < \lambda'_X \left[(\bar{t}-\underline{t}) \frac{H_\alpha(H^{-1}(\underline{t}))}{h(H^{-1}(\underline{t}))} - \int_{\underline{t}}^{\bar{t}} \frac{H_\alpha(H^{-1}(\tau))}{h(H^{-1}(\tau))} \right],$$

i.e.,

$$J^{\lambda'_X}(\bar{t}) < (\bar{t}-\underline{t})\varphi^{\lambda'_X}(\underline{t}) + J^{\lambda'_X}(\underline{t}).$$

By Lemma 25, $J^{\lambda'_X}(\underline{t}) > \overline{J}^{\lambda'_X}(\underline{t})$.

A similar argument proves that $J^{\lambda'_X}(\bar{t}) > \overline{J}^{\lambda'_X}(\bar{t})$. By continuity, there exists $\delta > 0$ such that $J^{\lambda'_X}(t) > \overline{J}^{\lambda'_X}(t)$ for all $t \in (\underline{t}-\delta, \bar{t}+\delta)$. ■

Lemma 26 *Let $[\underline{z}, \overline{z}] \subset W$ with $\underline{z} < \overline{z}$, and $z^0 \in (\underline{z}, \overline{z})$. Suppose that $Q : [\underline{\theta}, \overline{\theta}] \to [0,1]$ and $\hat{Q} : [\underline{\theta}, \overline{\theta}] \to [0,1]$ satisfy the following two conditions:*

$$\int_{\underline{z}}^{\overline{z}} Q(z)h(z)\mathrm{d}z = \int_{\underline{z}}^{\overline{z}} \hat{Q}(z)h(z)\mathrm{d}z, \tag{73}$$

and

$$Q(z) \geq \hat{Q}(z) \text{ if } z > z^0, \text{ and } Q(z) \leq \hat{Q}(z) \text{ if } z < z^0. \tag{74}$$

Then

$$\int_{\underline{z}}^{\overline{z}} -\frac{H_{\alpha_i}(z)}{h(z)}[Q(z) - \hat{Q}(z)]h(z)\mathrm{d}z \geq 0, \tag{75}$$

where the inequality holds strictly if the set $\{z \in [\underline{z}, \overline{z}] | Q(z) \neq \hat{Q}(z)\}$ has a positive measure.

Proof. Because $-\frac{H_\alpha(w)}{h(w)}$ is strictly increasing in w, and Q and $\hat{Q}$ satisfy (74), we have

$$\int_{\underline{z}}^{\overline{z}} \left[-\frac{H_{\alpha_i}(z)}{h(z)} + \frac{H_{\alpha_i}(z^0)}{h(z^0)} \right] [Q(z) - \hat{Q}(z)]h(z)\mathrm{d}z \geq 0,$$

where the inequality holds strictly if the set $\{z \in [\underline{z}, \overline{z}] | Q(z) \neq \hat{Q}(z)\}$ has a positive measure. This implies inequality (75) by (73). ■

A.3 Sufficient conditions for the first-order approach

In this section we provide sufficient conditions for the first-order approach to be valid. Let $\pi(\alpha_i)$ denote an agent i's payoff from choosing α_i given mechanism $(\boldsymbol{q}, \boldsymbol{t})$ and $\alpha_j = \alpha^*$ for all $j \neq i$. Then

$$\pi(\alpha_i) := U(w(0, \alpha_i)) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} [1 - H(w_i|\alpha_i)]\, Q(w_i)\mathrm{d}w_i - C(\alpha_i),$$

where Q is defined by (2) for $\alpha_j = \alpha^*$ for all $j \neq i$. Then

$$\begin{aligned}\pi'(\alpha_i) =& U'(w(0,\alpha_i))w_{\alpha_i}(0,\alpha_i) + [1 - H(w(1,\alpha_i)|\alpha_i)]\, Q(w(1,\alpha_i))w_{\alpha}(1,\alpha_i) \\ & - [1 - H(w(0,\alpha_i)|\alpha_i)]\, Q(w(0,\alpha_i))w_{\alpha_i}(0,\alpha_i) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -H_{\alpha_i}(w_i|\alpha_i)Q(w_i)\mathrm{d}w_i - C'(\alpha_i) \\ =& \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -H_{\alpha_i}(w_i|\alpha_i)Q(w_i)\mathrm{d}w_i - C'(\alpha_i),\end{aligned}$$

where the second line holds because $H(w(1,\alpha_i)|\alpha_i) = 1$, $H(w(0,\alpha_i)|\alpha_i) = 0$, and $U'(w(0,\alpha_i)) = Q(w(0,\alpha_i))$ by the envelope condition. A sufficient condition for the first-order approach to be valid is that $\pi'(\alpha_i)$ is strictly decreasing for all non-decreasing implementable allocation rule Q.

If the support of the conditional expectation $[w(0,\alpha_i), w(1,\alpha_i)]$ is invariant, $\pi'(\alpha_i)$ is strictly decreasing if $-H_{\alpha_i}(w_i|\alpha_i)$ has the *single-crossing property* in $(\alpha_i; w_i)$ and $C'(\alpha_i)$ is strictly decreasing.

In many important applications (e.g. the two leading examples in this paper), the support of the conditional expectation $[w(0,\alpha_i), w(1,\alpha_i)]$ is not invariant. In this case, we have

$$\begin{aligned}\pi''(\alpha_i) =& \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -\frac{\partial^2 H(w_i|\alpha_i)}{\partial \alpha_i^2} Q(w_i)\mathrm{d}w_i - H_{\alpha_i}(w(1,\alpha_i)|\alpha_i)w_{\alpha_i}(1,\alpha_i)Q(w(1,\alpha_i)) \\ & + H_{\alpha_i}(w(0,\alpha_i)|\alpha_i)w_{\alpha_i}(0,\alpha_i)Q(w(0,\alpha_i)) - C''(\alpha_i), \\ \leq & \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -\frac{\partial^2 H(w_i|\alpha_i)}{\partial \alpha_i^2} Q(w_i)\mathrm{d}w_i - H_{\alpha_i}(w(1,\alpha_i)|\alpha_i)w_{\alpha_i}(1,\alpha_i)Q(w(1,\alpha_i)) - C''(\alpha_i),\end{aligned}$$

where the inequality holds because $H_{\alpha_i}(w(0,\alpha_i)|\alpha_i) \geq 0$ and $w_{\alpha_i}(0,\alpha_i) \leq 0$ when the information structures are supermodular ordered. The following proposition from Shi (2012) gives sufficient conditions for $\pi''(\alpha_i) < 0$ for the two leading examples.

Proposition 6 (Shi (2012)) *The following conditions are sufficient for first order approach:*

- *In the linear experiments, if $\alpha_i C''(\alpha_i) \geq f\left(\overline{\theta}\right)(\overline{\theta} - \mu)^2$ for all α_i, then $\pi''(\alpha_i) < 0$ when either $F(\theta)$ is convex, or $F(\theta) = \theta^b$ $(b > 0)$ with support $[0,1]$.*
- *In the normal experiments, $\pi''(\alpha_i) < 0$ if $\sqrt{\beta^3 / \left[\alpha_i^3(\alpha_i + \beta)^5\right]} < 2\sqrt{2\pi}C''(\alpha_i)$ for all α_i.*

We conclude this section by proving the following lemma which is used in the proof of Proposition 2.

Lemma 27 *If $\pi''(\alpha_i) < 0$ for all α_i and all non-decreasing implementable allocation rule Q,*

$$\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -H_{\alpha_i}(w_i|\alpha_i)H(w_i|\alpha_i)^{n-1}\mathrm{d}w_i - C'(\alpha_i) \text{ is strictly decreasing in } \alpha_i. \qquad (14)$$

Proof. In particular, $\pi''(\alpha_i) < 0$ if $Q(w_i) = H(w_i|\alpha_i)^{n-1}$ for all w_i. Then

$$\begin{aligned}
&\frac{\partial}{\partial \alpha_i}\left[\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -H_{\alpha_i}(w_i|\alpha_i)H(w_i|\alpha_i)^{n-1}\mathrm{d}w_i - C'(\alpha_i)\right] \\
=&\int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -\frac{\partial^2 H(w_i|\alpha_i)}{\partial \alpha_i^2}Q(w_i)\mathrm{d}w_i - H_{\alpha_i}(w(1,\alpha_i)|\alpha_i)w_{\alpha_i}(1,\alpha_i)Q(w(1,\alpha_i)) \\
&+ H_{\alpha_i}(w(0,\alpha_i)|\alpha_i)w_{\alpha_i}(0,\alpha_i)Q(w(0,\alpha_i)) - C''(\alpha_i) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -(n-1)H_{\alpha_i}(w_i|\alpha_i)^2 H(w_i|\alpha_i)^{n-2}\mathrm{d}w_i, \\
=&\pi''(\alpha_i) + \int_{w(0,\alpha_i)}^{w(1,\alpha_i)} -(n-1)H_{\alpha_i}(w_i|\alpha_i)^2 H(w_i|\alpha_i)^{n-2}\mathrm{d}w_i \\
<&0,
\end{aligned}$$

where $Q(w_i) = H(w_i|\alpha_i)^{n-1}$ for all w_i. ■

B Efficient asymmetric mechanisms

B.1 Proof of Theorem 3

As in the proof of Theorem 1, we prove Theorem 3 by proving the following two lemmas. Define

$$Y(\boldsymbol{w}) := 1 - \prod_{i=1}^{n} H(w_i|\alpha_i^*) - \sum_{i=1}^{n}\int_{w_i}^{w(1,\alpha_i^*)} Q_i(z_i)\mathrm{d}H(z_i|\alpha_i^*), \forall \boldsymbol{w} \in \prod_{i=1}^{n}[w(0,\alpha_i^*), w(1,\alpha_i^*)].$$

Recall that $1 - \prod_{i=1}^{n} H(w_i|\alpha_i^*)$ is the probability with which there exists an agent i whose type is above w_i; and $\sum_{i=1}^{n}\int_{w_i}^{w(1,\alpha_i^*)} Q_i(z_i)\mathrm{d}H(z_i|\alpha_i^*)$ is the probability with which an agent whose type is above w_i receives the object. Let $\underline{\boldsymbol{w}} := (w(0,\alpha_1^*), \ldots, w(0,\alpha_n^*))$. Then $Y(\underline{\boldsymbol{w}})$ is the difference between 1 and the probability with which some agent receives the object. Clearly, (16) is violated if and only if $Y(\underline{\boldsymbol{w}}) > 0$.

Lemma 28 *Suppose that the information structures are supermodular ordered and $\alpha_i = \alpha_i^*$ for all i. Let $\boldsymbol{Q}$ be any interim allocation rule satisfying (F′) (MON), (IA′) and $Y(\underline{\boldsymbol{w}}) > 0$. Then, for any i, there exists $\hat{\boldsymbol{Q}}$ satisfying (F′), (MON) and (IA′) such that $\hat{Q}_j = Q_j$ for*

$j \neq i$ and

$$\hat{Q}_i(w_i) \geq Q_i(w_i),\ \forall w_i \in W_i, \tag{76}$$

and the strict inequality holds for a set of w_i with positive measure.

Proof. Fix i. Define $Y_i(w_i) := \inf_{w_{-i}} Y(\boldsymbol{w})$ for all $w_i \in [w(0,\alpha_i^*), w(1,\alpha_i^*)]$. By Theorem 3 in Milgrom and Segal (2002), Y_i is differentiable and $Y_i'(w_i) = -h(w_i|\alpha_i^*)\prod_{j\neq i} H(w_j^*|\alpha_j^*) + Q_i(w_i)h(w_i|\alpha_i^*)$, where $w_{-i}^*(w_i)$ is such that $Y(w_i, w_{-i}^*(w_i)) = Y_i(w_i)$ for all $w_i \in (w(0,\alpha_i^*), w(1,\alpha_i^*))$. Note that

$$Y(w(0,\alpha_i^*), w_{-i}) = 1 - \int_{w(0,\alpha_i^*)}^{w(1,\alpha_i^*)} Q_i(z_i)\mathrm{d}H(z_i|\alpha_i^*) - \sum_{j\neq i}^{n} \int_{w_j}^{w(1,\alpha_j^*)} Q_j(z_j)\mathrm{d}H(z_j|\alpha_j^*),$$

which is strictly increasing in w_j for all $j \neq i$. Hence, $Y_i(w(0,\alpha_i^*)) = Y(\underline{w}) > 0$. Define $w^\flat := \sup\{w_i\,|Y_i(w_i') > 0\ \forall w(0,\alpha_i^*) \leq w_i' \leq w_i\}$. By the continuity of Y_i, we have $Y_i(w^\flat) = 0$ and $w^\flat > w(0,\alpha_i^*)$. There are four cases to consider.

Case I: Suppose that there exists $w_i' \in (w(0,\alpha_i^*), w^\flat)$ such that Q_i is discontinuous at w_i'.

Let $Q_i(w_i'^{+})$ denote the right-hand limit of Q_i at w_i' and $Q_i(w_i'^{-})$ the corresponding left-hand limit. Let $0 < \varepsilon \leq \min\left\{\min_{w(0,\alpha_i^*)\leq w_i \leq w_i'} \frac{Y_i(w_i)}{H(w_i'|\alpha_i^*)}, Q_i(w_i'^{+}) - Q_i(w_i'^{-})\right\}$. Define $\hat{\boldsymbol{Q}}$ as follows. If $w_i \leq w(0,\alpha^*)$, let $\hat{Q}_i(w_i) := Q_i(w_i)$; otherwise let

$$\hat{Q}_i(w_i) := Q_i(w_i) + \varepsilon\chi_{\{w_i \leq w_i'\}},$$

where $\chi_{\{w_i \leq w_i'\}}$ is the indicator function. Let $\hat{Q}_j := Q_j$ for all $j \neq i$. By construction, $\hat{Q}_i(w) \geq Q_i(w)$ for all $w_i \in W_i$ and the inequality holds strictly on a positive measure set. By a similar argument to that in the proof of Lemma 3, $\hat{Q}_i$ satisfies (MON) and (IA′). We now verify that $\hat{\boldsymbol{Q}}$ satisfies (F′). If $w_i \leq w_i'$, $\hat{Y}(w_i, w_{-i}) = Y(w_i, w_{-i}) - \varepsilon[H(w_i'|\alpha^*) - H(w_i|\alpha^*)] \geq Y(w_i, w_{-i}) - \varepsilon H(w_i'|\alpha_{-i}^*) \geq 0$ for all w_{-i}. If $w_i > w_i'$, $\hat{Y}(w_i, w_{-i}) = Y(w_i, w_{-i}) \geq 0$ for all w_{-i}. That is, $\hat{\boldsymbol{Q}}$ satisfies (F′).

Case II: Suppose that Q_i is continuous on $[w(0,\alpha_i^*), w^\flat]$.

We first show that there exists $w_i' \in (w(0,\alpha_i^*), w^\flat)$ such that $Q_i(w_i') < Q_i(w^\flat)$. Suppose, to the contrary, that $Q_i(w_i) = Q_i(w^\flat)$ for all $w_i \in (w(0,\alpha_i^*), w^\flat)$. Let w_{-i}^* be such that $Y(w^\flat, w_{-i}^*) = Y_i(w^\flat) = 0$. If $Q_i(w^\flat) \geq \prod_{j\neq i} H(w_j^*|\alpha_j^*)$,

$$Y(w(0,\alpha_i^*), w_{-i}^*) = Y(w^\flat, w_{-i}^*) + \int_{w(0,\alpha_i^*)}^{w^\flat} \left[\prod_{j\neq i} H(w_j^*|\alpha_j^*) - Q_i(z)\right] h(z|\alpha_i^*)\mathrm{d}z < 0,$$

a contradiction. Hence, $Q_i(w^\flat) < \prod_{j\neq i} H(w_j^*|\alpha_j^*)$. Then, by the continuity of Q_i and H, there exists $\delta > 0$ such that $Q_i(w_i) < \prod_{j\neq i} H(w_j^*|\alpha_j^*)$ for all $w_i \in [w^\flat, w^\flat+\delta]$. Moreover,

$$0 = Y(w^\flat, w_{-i}^*) = \int_{w^\flat}^{w^\flat+\delta} \left[\prod_{j\neq i} H(w_j^*|\alpha_j^*) - Q(z)\right] h(z|\alpha_i^*)\mathrm{d}z + Y(w^\flat+\delta, w_{-i}^*) > Y(w^\flat+\delta, w_{-i}^*),$$

a contradiction. Hence, there exists $w_i' \in (w(0,\alpha_i^*), w^\flat)$ such that $Q_i(w_i') < Q_i(w^\flat)$.

By the continuity of Q_i, there exists $w_i'' \in (w_i', w^\flat)$ such that $Q_i(w_i'') = \frac{1}{2}\left(Q_i(w_i') + Q_i(w^\flat)\right)$. Let $0 < \varepsilon \leq \min\left\{\min_{w(0,\alpha_i^*)\leq w_i \leq w_i''} \frac{Y_i(w_i)}{H(w_i''|\alpha_i^*)}, Q_i(w_i'') - Q_i(w_i')\right\}$. Let $\hat{Q}_j := Q_j$ for $j \neq i$ and

$$\hat{Q}_i(w_i) := \begin{cases} \max\{Q_i(w_i')+\varepsilon, Q_i(w_i)\} & \text{if } w_i > w_i', \\ Q_i(w_i)+\varepsilon & \text{if } w(0,\alpha_i^*) < w_i \leq w_i', \\ Q_i(w_i) & \text{if } w_i \leq w(0,\alpha_i^*). \end{cases}$$

Note that if $w_i \geq w_i''$, $Q_i(w_i) \geq Q_i(w_i'') \geq Q_i(w_i') + \varepsilon$. Hence, $\hat{Q}_i(w_i) = Q_i(w_i)$ for $w_i \geq w_i''$. By construction, $\hat{Q}_i(w_i) \geq Q(w_i)$ for all $w_i \in W_i$ and the inequality holds strictly on a positive measure set. By a similar argument to that in the proof of Lemma 3, $\hat{Q}_i$ satisfies (MON) and (IA′). We now verify that $\hat{\boldsymbol{Q}}$ satisfies (F′). If $w_i \geq w_i''$, $\hat{Y}(w_i, w_{-i}) = Y(w_i, w_{-i}) \geq 0$ for all w_{-i}. If $w_i < w_i''$, for all w_{-i},

$$\begin{aligned}\hat{Y}(w_i, w_{-i}) =& Y(w_i, w_{-i}) - \int_{w_i}^{w_i''} \left[\hat{Q}_i(z) - Q_i(z)\right] h_i(z|\alpha_i^*)\mathrm{d}z, \\ \geq& Y(w_i, w_{-i}) - \varepsilon\left[H(w_i''|\alpha_i^*) - H(w_i|\alpha_i^*)\right], \\ \geq& Y_i(w_i) - \varepsilon H(w_i''|\alpha^*) \geq 0.\end{aligned}$$

Hence, $\hat{\boldsymbol{Q}}$ satisfies (F′).

Case III: Let w_{-i}^* be such that $Y(w^\flat, w_{-i}^*) = Y_i(w^\flat) = 0$. Suppose that Q_i is continuous on $[w(0,\alpha_i^*), w^\flat)$ and $Q_i(w^{\flat-}) < \prod_{j\neq i} H(w_j^*|\alpha_j^*)$.

Define $R(w_i) := Y_i(w_i)/(H(w^\flat|\alpha_i^*) - H(w_i|\alpha_i^*))$ for $w_i < w^\flat$. Then, by Theorem 3 in Milgrom and Segal (2002) and L'Hopital's rule,

$$\lim_{w_i \to w^{\flat-}} R(w_i) = \prod_{j\neq i} H_j(w_j^*|\alpha_j^*) - Q_i(w^{\flat-}) > 0.$$

Let $0 < \varepsilon \leq \min\left\{\inf_{w(0,\alpha_i^*)\leq w_i < w^\flat} R(w_i), Q_i(w^{\flat+}) - Q_i(w^{\flat-})\right\}$. Let $\hat{Q}_j := Q_j$ for all $j \neq i$. If $w_i \leq w(0,\alpha_i^*)$, let $\hat{Q}_i(w_i) := Q_i(w_i)$; otherwise let $\hat{Q}_i(w_i) := Q_i(w_i) + \varepsilon\chi_{\{w_i < w^\flat\}}$.

By construction, $\hat{Q}_i(w_i) \geq Q_i(w_i)$ for all $w_i \in W_i$ and the inequality holds strictly on a positive measure set. One can verify that $\hat{Q}_i$ satisfies (MON) and (IA′) by an argument similar to that in the proof of Lemma 3. Finally, if $w_i < w^\flat$, $\hat{Y}(w_i, w_{-i}) = Y(w_i, w_{-i}) - \varepsilon[H(w^\flat|\alpha_i^*) - H(w_i|\alpha_i^*)] \geq Y_i(w_i) - R(w_i)[H(w^\flat|\alpha^*) - H(w_i|\alpha^*)] = 0$ for all w_{-i}. If $w_i \geq w^\flat$, $\hat{Y}(w_i, w_{-i}) = Y(w_i, w_{-i}) \geq 0$ for all w_{-i}. Hence, $\hat{\boldsymbol{Q}}$ satisfies (F′).

Case IV: Let w_{-i}^* be such that $Y(w^\flat, w_{-i}^*) = Y_i(w^\flat) = 0$. Suppose that Q_i is continuous on $[w(0, \alpha_i^*), w^\flat)$ and $Q_i(w^{\flat-}) \geq \prod_{j \neq i} H(w_j^*|\alpha_j^*)$.

We first show that $Q_i(w^{\flat-}) = \prod_{j \neq i} H(w_j^*|\alpha_j^*)$. Suppose, to the contrary, that $Q_i(w^{\flat-}) > \prod_{j \neq i} H(w_j^*|\alpha_j^*)$. Then, by the continuity of Q_i and H on $[w(0, \alpha^*), w^\flat)$, there exists $\delta > 0$ such that $Q_i(w_i) > \prod_{j \neq i} H(w_j^*|\alpha_j^*)$ for all $w_i \in (w^\flat - \delta, w^\flat)$. Then

$$Y(w^\flat - \delta, w_{-i}) = \int_{w^\flat - \delta}^{w^\flat} \left[\prod_{j \neq i} H(w_j^*|\alpha_j^*) - Q_i(z) \right] h(z|\alpha_i^*) \mathrm{d}z < 0,$$

a contradiction. Hence, $Q_i(w^{\flat-}) = \prod_{j \neq i} H(w_j^*|\alpha_j^*)$. Second, we show that there exists $w_i' \in (w(0, \alpha_i^*), w^\flat)$ such that $Q_i(w_i') < Q_i(w^{\flat-})$. Suppose, to the contrary, that $Q_i(w_i) = Q_i(w^{\flat-})$ for all $w_i \in (w(0, \alpha^*), w^\flat)$. Then

$$Y(w(0, \alpha^*), w_{-i}^*) = \int_{w(0,\alpha^*)}^{w^\flat} \left[\prod_{j \neq i} H(w_j^*|\alpha_j^*) - Q_i(z) \right] h(z|\alpha_i^*) \mathrm{d}z < 0,$$

a contradiction. Hence, there exists $w_i' \in (w(0, \alpha_i^*), w^\flat)$ such that $Q_i(w_i') < Q_i(w^{\flat-})$. The rest of the proof follows from that of Case II. ■

Lemma 29 *Suppose that the information structures are supermodular ordered and $\alpha_i = \alpha_i^*$ for all i. Let $\boldsymbol{Q}$ and $\hat{\boldsymbol{Q}}$ be two implementable allocation rules satisfying (7). Let $\boldsymbol{q}$ be an ex-post allocation rule that implements $\boldsymbol{Q}$. Then there exists an ex-post allocation rule $\hat{\boldsymbol{q}}$ that implements $\hat{\boldsymbol{Q}}$ and satisfies*

$$\mathbb{E}_{\boldsymbol{w}} \left[\sum_i (w_i + \gamma \sum_{j \neq i} w_j) \hat{q}_i(\boldsymbol{w}) \middle| \alpha_i = \alpha_i^* \ \forall i \right] > \mathbb{E}_{\boldsymbol{w}} \left[\sum_i (w_i + \gamma \sum_{j \neq i} w_j) q_i(\boldsymbol{w}) \middle| \alpha_i = \alpha_i^* \ \forall i \right].$$

The proof of Lemma 29 relies on the following technical lemma. For each i, let h_i denote the probability measure on $[w(0, \alpha_i^*), w(1, \alpha_i^*)]$ corresponding to $H(w_i|\alpha_i^*)$, then

Lemma 30 *Let $\boldsymbol{Q} : \prod_i W_i \to [0, 1]^n$ be an interim allocation rule and $\rho : \prod_i W_i \to [0, 1]$ be a measurable function. Then there exists an ex post allocation rule $\boldsymbol{q}$ that implements $\boldsymbol{Q}$ and satisfies $\sum_i q_i(\boldsymbol{w}) \geq \rho(\boldsymbol{w})$ for almost all $\boldsymbol{w} \in \prod_i [w(0, \alpha_i^*), w(1, \alpha_i^*)]$ if and only if for each measurable set $\boldsymbol{A} = (A_1, \ldots, A_n)$, where $A_i \subset [w(0, \alpha_1^*), w(0, \alpha_i^*)]$ for all i, the following* *inequality holds:*

$$\int_{\boldsymbol{A}} \rho(\boldsymbol{w}) \mathrm{d}h_1(w_1)\ldots \mathrm{d}h_n(w_n) \le \sum_i \int_{A_i} Q(w_i)\mathrm{d}h_i(w_i) \le \int_{\boldsymbol{A}} \mathrm{d}h_1(w_1)\ldots \mathrm{d}h_n(w_n). \tag{77}$$

The proof of Lemma 5 can be readily extended to prove Lemma 30 and is neglected here. With Lemma 30 in hand, the proof of Lemma 4 can be readily extended to prove Lemma 29 and is also neglected here. Theorem 3 follows immediately from Lemmas 28 and 29.

B.2 Other omitted proofs

Proof of Proposition 4. As in Section 3, we solve ($\mathcal{P}'$) in two steps. First, for each $\boldsymbol{\alpha}^* \in [\underline{\alpha}, \overline{\alpha}]^n$, we solve the following sub-problem, denoted by ($\mathcal{P}'$-$\boldsymbol{\alpha}^*$):

$$V(\boldsymbol{\alpha}^*) := \max_{\boldsymbol{Q}} \mathbb{E}_{\boldsymbol{w}} \left[\sum_i w_i Q_i(w_i) \middle| \boldsymbol{\alpha}^* \right] \text{ subject to (F}'\text{), (MON) and (IA}'\text{) },$$

Second, we solve $\max_{\boldsymbol{\alpha} \in [\underline{\alpha}, \overline{\alpha}]^n} \pi^s(\boldsymbol{\alpha}) := (1-\gamma)V(\boldsymbol{\alpha}) - \sum_i C(\alpha_i)$.

Fix $\boldsymbol{\alpha}^*$. If the principal adopts the ex post efficient mechanism, the interim allocation rule is given by $Q_i(w_i) = \prod_{j \neq i} H(w_i|\alpha_i^*)$ for all w_i and all i. Clearly, if $\boldsymbol{\alpha}^*$ is such that

$$\mathbb{E}_{w_i} \left[-\frac{H_{\alpha_i}(w_i|\alpha_i^*)}{h(w_i|\alpha_i^*)} \prod_{j \neq i} H(w_i|\alpha_j^*) \middle| \boldsymbol{\alpha}^* \right] \le C'(\alpha_i^*), \forall i, \tag{78}$$

the ex post efficient mechanism solves ($\mathcal{P}'$-$\boldsymbol{\alpha}^*$). Furthermore, Lemma 31 below proves that if $\boldsymbol{\alpha}^*$ is chosen optimally, (78) (or equivalently (IA$'$) in this case) holds with equality for all i and the expected social surplus is given by

$$\pi^s(\boldsymbol{\alpha^*}) = (1-\gamma) \left[\mu + \sum_{i=1}^n b(\alpha_i^*) C'(\alpha_i^*) \right] - \sum_i C(\alpha_i^*). \tag{18}$$

Suppose that $\boldsymbol{\alpha}^*$ is such that

$$\mathbb{E}_{w_i} \left[-\frac{H_{\alpha_i}(w_i|\alpha_i^*)}{h(w_i|\alpha_i^*)} \prod_{j \neq i} H(w_i|\alpha_j^*) \middle| \boldsymbol{\alpha}^* \right] > C'(\alpha_i^*) \text{ for some } i. \tag{79}$$

Assume, without loss of generality, that (IA$'$) binds for the first k ($0 < k \le n$) agents. Then we can ignore (IA$'$) for the last $n-k$ agents. Let λ_i denote the Lagrangian multiplier associated with (IA$'$) for agent i ($i \le k$). By a similar argument to that in Section 3.2, the optimal Lagrangian multipliers satisfy $\lambda_i = b(\alpha_i^*)$ for all $i \le k$. Then the Lagrangian of ($\mathcal{P}'$-$\boldsymbol{\alpha}^*$) can be written as

$$\begin{aligned}
\mathscr{L} =& \sum_{i \le k} \int_{w(0,\alpha_i^*)}^{w(1,\alpha_i^*)} \left[w_i + \lambda_i \frac{H(w_i|\alpha_i^*)}{h(w_i|\alpha_i^*)} \right] Q_i(w_i) h(w_i|\alpha_i^*) \mathrm{d}w_i \\
&+ \sum_{i > k} \int_{w(0,\alpha_i^*)}^{w(1,\alpha_i^*)} w_i Q_i(w_i) h(w_i|\alpha_i^*) \mathrm{d}w_i + \sum_{i \le k} \lambda_i C'(\alpha_i^*) \\
=& \sum_{i \le k} \int_{w(0,\alpha_i^*)}^{w(1,\alpha_i^*)} \mu Q_i(w_i) h(w_i|\alpha_i^*) \mathrm{d}w_i + \sum_{i > k} \int_{w(0,\alpha_i^*)}^{w(1,\alpha_i^*)} w_i Q_i(w_i) h(w_i|\alpha_i^*) \mathrm{d}w_i + \sum_{i \le k} b(\alpha_i^*) C'(\alpha_i^*) \\
=& \int_{w(0,\alpha_1^*)}^{w(1,\alpha_1^*)} \cdots \int_{w(0,\alpha_n^*)}^{w(1,\alpha_n^*)} \left(\sum_{i \le k} \mu q_i(\boldsymbol{w}) + \sum_{i > k} w_i q_i(\boldsymbol{w}) \right) \prod_{i=1}^{n} h(w_i|\alpha_i^*) \mathrm{d}w_1 \ldots \mathrm{d}w_n + \sum_{i \le k} b(\alpha_i^*) C'(\alpha_i^*).
\end{aligned}$$

Suppose that $k < n$, then a pointwise virtual surplus maximizer must satisfy for all $\boldsymbol{w}$,

$$\sum_{i \le k} q_i(\boldsymbol{w}) = \begin{cases} 1 & \text{if } \max_{j>k}\{w_j\} < \mu \\ 0 & \text{if } \max_{j>k}\{w_j\} > \mu \end{cases}$$

and

$$q_i(\boldsymbol{w}) = \begin{cases} 1 & \text{if } w_i > \mu \text{ and } w_i > \max_{j>k}\{w_j\} \\ 0 & \text{if } w_i < \mu \text{ or } w_i < \max_{j>k}\{w_j\} \end{cases}, \ \forall i > k.$$

Therefore, for all $i > k$, the optimal interim allocation rule is given by $Q_i(w_i) = \prod_{j>k, j \ne i} H(w_i|\alpha_j^*)$ if $w_i > \mu$ and $Q_i(w_i) = 0$ if $w_i < \mu$. Hence,

$$V(\boldsymbol{\alpha}^*) = \mu \prod_{i>k} H(\mu|\alpha_i^*) + \sum_{i>k} \int_{\mu}^{w(1,\alpha_i^*)} w_i \prod_{j>k, j \ne i} H(w_i|\alpha_j^*) h(w_i|\alpha_i^*) \mathrm{d}w_i + \sum_{i \le k} b(\alpha_i^*) C'(\alpha_i^*).$$

Finally, (IA$'$) holds for $i > k$ if and only if

$$\int_{\mu}^{w(1,\alpha_i^*)} -H_{\alpha_i}(w_i|\alpha_i^*) \prod_{j>k, j \ne i} H(w_i|\alpha_j^*) \mathrm{d}w_i \le C'(\alpha_i^*). \tag{80}$$

Consider an agent i ($i > k$). We argue that if $\boldsymbol{\alpha}^*$ is chosen optimally, (80) holds with equality. Suppose, to the contrary, that (80) holds with strict inequality, then

$$\frac{\partial \pi^s(\boldsymbol{\alpha}^*)}{\partial \alpha_i} = -(1-\gamma) \int_{\mu}^{w(1,\alpha_i^*)} H_{\alpha_i}(w_i|\alpha_i^*) \prod_{j>k, j \ne i} H(w_i|\alpha_j^*) \mathrm{d}w_i - C'(\alpha_i^*) < -\gamma C'(\alpha_i^*) \le 0,$$

a contradiction to the optimality of α_i^*. Hence, (80) holds with equality for all $i > k$.

Furthermore, because the information structures are uniformly supermodular ordered, we have

$$\int_{\mu}^{w(1,\alpha_i^*)} w_i \prod_{j>k, j\neq i} H(w_i|\alpha_j^*)\mathrm{d}w_i = C'(\alpha_i^*)b(\alpha_i^*) + \int_{\mu}^{w(1,\alpha_i^*)} \mu \prod_{j>k, j\neq i} H(w_i|\alpha_j^*)\mathrm{d}w_i.$$

Substituting this into the expression of $V(\boldsymbol{\alpha}^*)$ yields

$$\begin{aligned} V(\boldsymbol{\alpha}^*) =& \mu \prod_{i>k} H(\mu|\alpha_i^*) + \sum_{i>k} \int_{\mu}^{w(1,\alpha_i^*)} \mu \prod_{j>k, j\neq i} H(w_i|\alpha_j^*)h(w_i|\alpha_i^*)\mathrm{d}w_i + \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*) \\ =& \mu \prod_{i>k} H(\mu|\alpha_i^*) + \mu \int_{\mu}^{\overline{\theta}} \mathrm{d} \prod_{j>k} H(w|\alpha_j^*) + \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*) \\ =& \mu \prod_{i>k} H(\mu|\alpha_i^*) + \mu \left[1 - \prod_{i>k} H(\mu|\alpha_i^*)\right] + \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*) = \mu + \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*). \end{aligned}$$

Hence,

$$\pi^s(\boldsymbol{\alpha^*}) = (1-\gamma)\left[\mu + \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*)\right] - \sum_i C(\alpha_i^*). \tag{18}$$

Suppose that $k = n$, then a pointwise virtual surplus maximizer must satisfy $\sum_i q_i(\boldsymbol{w}) = 1$ for all $\boldsymbol{w}$. Hence, (18) still holds in this case. ■

Lemma 31 *Suppose that the second-order condition of the agents' optimization problem is satisfied and the information structures are uniformly supermodular ordered. Let $\boldsymbol{\alpha}^*$ be a socially optimal information choice. Suppose, in addition, that (78) holds. Then (78) holds with equality for all i. Furthermore,*

$$\pi^s(\boldsymbol{\alpha^*}) = (1-\gamma)\left[\mu + \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*)\right] - \sum_i C(\alpha_i^*). \tag{18}$$

Proof. Let $\boldsymbol{\alpha}^*$ be a socially optimal information choice. Suppose, to the contrary, that $\boldsymbol{\alpha}^*$

is such that (78) holds with strictly inequality for some i. Note that

$$\begin{aligned}
V(\boldsymbol{\alpha}^*) &= \sum_{i=1}^{n} \int_{w(0,\alpha_i^*)}^{w(1,\alpha_i^*)} w \prod_{j\neq i} H(w|\alpha_j^*)h(w|\alpha_i^*)\mathrm{d}w \\
&= \int_{\underline{\theta}}^{\overline{\theta}} w \mathrm{d} \prod_{i} H(w|\alpha_i^*) \\
&= \overline{\theta} - \int_{\underline{\theta}}^{\overline{\theta}} \prod_{i} H(w|\alpha_i^*)\mathrm{d}w,
\end{aligned}$$

where the last line holds by integration by parts. Hence,

$$\begin{aligned}
\frac{\partial}{\partial \alpha_i}\pi^s(\boldsymbol{\alpha}^*) &= (1-\gamma)\frac{\partial}{\partial \alpha_i}V(\boldsymbol{\alpha}^*) - C'(\alpha_i^*) \\
&= (1-\gamma)\int_{\underline{\theta}}^{\overline{\theta}} -H_{\alpha_i}(w|\alpha_i^*)\prod_{j\neq i} H(w|\alpha_j^*)\mathrm{d}w - C'(\alpha_i^*),
\end{aligned}$$

which is strictly decreasing in α_i when the second-order condition of the agents' optimization problem is satisfied. Because $\boldsymbol{\alpha}^*$ is such that (78) holds with strictly inequality for i,

$$\frac{\partial}{\partial \alpha_i}\pi^s(\boldsymbol{\alpha}^*) < -\gamma C'(\alpha_i^*) \leq 0.$$

Hence, $\boldsymbol{\alpha}^*$ is not optimal, a contradiction. Hence, if $\boldsymbol{\alpha}^*$ is chosen optimally, then (IA′) holds with equality for all i.

Let $\boldsymbol{\alpha}^*$ be such that (78) holds. Because the information structures are uniformly supermodular ordered, and (IA′) holds with equality, we have

$$\int_{w(0,\alpha_i^*)}^{w(1,\alpha_i^*)} \frac{w-\mu}{b(\alpha_i)} \prod_{j\neq i} H(w|\alpha_j^*)h(w|\alpha_i^*)\mathrm{d}w = C'(\alpha_i^*).$$

Hence,

$$\begin{aligned}
V(\boldsymbol{\alpha}^*) &= \sum_{i=1}^{n} \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} w \prod_{j\neq i} H(w|\alpha_j^*)h(w|\alpha_i^*)\mathrm{d}w \\
&= \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*) + \mu \sum_{i=1}^{n} \int_{w(0,\alpha^*)}^{w(1,\alpha^*)} \prod_{j\neq i} H(w|\alpha_j^*)h(w|\alpha_i^*)\mathrm{d}w \\
&= \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*) + \mu \int_{\underline{\theta}}^{\overline{\theta}} \mathrm{d}\prod_{i} H(w|\alpha_i) \\
&= \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*) + \mu.
\end{aligned}$$

Hence,

$$\pi^s(\boldsymbol{\alpha^*}) = (1-\gamma)\left[\mu + \sum_{i=1}^{n} b(\alpha_i^*)C'(\alpha_i^*)\right] - \sum_i C(\alpha_i^*). \tag{18}$$

This completes the proof. ■

References

Ausubel, L. M., Cramton, P., 2004. Auctioning many divisible goods. Journal of the European Economic Association 2 (2-3), 480–493.

Bergemann, D., Shi, X., Välimäki, J., 2009. Information acquisition in interdependent value auctions. Journal of the European Economic Association 7 (1), 61–89.

Bergemann, D., Välimäki, J., 2002. Information acquisition and efficient mechanism design. Econometrica 70 (3), 1007–1033.

Biais, B., Martimort, D., Rochet, J.-C., 2000. Competing mechanisms in a common value environment. Econometrica 68 (4), 799–837.

Border, K. C., 1991. Implementation of reduced form auctions: A geometric approach. Econometrica, 1175–1187.

Che, Y.-K., Kim, J., Mierendorff, K., 2013. Generalized reduced-form auctions: A network-flow approach. Econometrica 81 (6), 2487–2520.

Crémer, J., Spiegel, Y., Zheng, C. Z., 2009. Auctions with costly information acquisition. Economic Theory 38 (1), 41–72.

Folland, G., 1999. Real analysis: modern techniques and their applications. Pure and applied mathematics. Wiley.

Ganuza, J.-J., Penalva, J. S., 2010. Signal orderings based on dispersion and the supply of private information in auctions. Econometrica 78 (3), 1007–1030.

Haile, P., Hendricks, K., Porter, R., 2010. Recent us offshore oil and gas lease bidding: A progress report. International Journal of Industrial Organization 28 (4), 390–396.

Hassin, R., 1982. Minimum cost flow with set-constraints. Networks 12 (1), 1–21.

Johnson, J. P., Myatt, D. P., 2006. On the simple economics of advertising, marketing, and product design. The American Economic Review, 756–784.

Karamardian, S., Schaible, S., 1990. Seven kinds of monotone maps. Journal of Optimization Theory and Applications 66 (1), 37–46.

Krishna, V., 2009. Auction theory. Academic press.

Lehmann, E. L., 1988. Comparing location experiments. The Annals of Statistics 16 (2), pp. 521–533.

Levin, D., Smith, J. L., 1994. Equilibrium in auctions with entry. The American Economic Review, 585–599.

Lewis, T. R., Sappington, D. E., 1994. Supplying information to facilitate price discrimination. International Economic Review, 309–327.

Li, Y., 2017. Mechanism design with costly verification and limited punishments.

Lu, J., Ye, L., 2014. Optimal two-stage auctions with costly information acquisition. Ohio State University Discussion paper.

Maskin, E., 1992. Auctions and Privatization. J.C.B. Mohr Publisher, pp. 115–136.

Maskin, E., Riley, J., 1984. Optimal auctions with risk averse buyers. Econometrica: Journal of the Econometric Society, 1473–1518.

Matthews, S., 1984a. Information acquisition in discriminatory auctions. Bayesian models in economic theory, ed. by M. Boyer, and R. Kihlstrom 49, 1477–1500.

Matthews, S. A., 1983. Selling to risk averse buyers with unobservable tastes. Journal of Economic Theory 30 (2), 370–400.

Matthews, S. A., 1984b. On the implementability of reduced form auctions. Econometrica: Journal of the Econometric Society, 1519–1522.

Mierendorff, K., 2009. Optimal dynamic mechanism design with deadlines. University of Bonn.

Mierendorff, K., 2011. Asymmetric reduced form auctions. Economics Letters 110 (1), 41–44.

Mierendorff, K., 2016. Optimal dynamic mechanism design with deadlines. Journal of Economic Theory 161, 190–222.

Milgrom, P., Segal, I., 2002. Envelope theorems for arbitrary choice sets. Econometrica 70 (2), 583–601.

Myerson, R. B., 1981. Optimal auction design. Mathematics of operations research 6 (1), 58–73.

Pai, M. M., Vohra, R., 2014. Optimal auctions with financially constrained buyers. Journal of Economic Theory 150, 383–425.

Persico, N., 2000. Information acquisition in auctions. Econometrica 68 (1), 135–148.

Quint, D., Hendricks, K., 2013. Indicative bidding in auctions with costly entry. Tech. rep., Working paper.

Reid, W. T., 1968. A simple optimal control problem involving approximation by monotone functions. Journal of Optimization Theory and Applications 2 (6), 365–377.

Seierstad, A., Sydsæter, K., 1987. Optimal Control Theory with Economic Applications. Advanced textbooks in economics. North-Holland.

Shi, X., 2012. Optimal auctions with information acquisition. Games and Economic Behavior 74 (2), 666–686.

Stegeman, M., 1996. Participation costs and efficient auctions. Journal of Economic Theory 71 (1), 228–259.

Szalay, D., 2009. Contracts with endogenous information. Games and Economic Behavior 65 (2), 586–625.

Toikka, J., 2011. Ironing without control. Journal of Economic Theory 146 (6), 2510–2526.

Ye, L., 2004. Optimal auctions with endogenous entry. Contributions in Theoretical Economics 4 (1).