Rent Extraction with Information Acquisition

Krittanai Laohakunakorn*

July 2, 2026

Abstract

This paper provides a framework for modelling information acquisition and revisits the classic mechanism design question of when privately informed buyers in an auction setting can expect to receive economic rents. It is well known that under standard assumptions, the seller can fully extract rent for generic prior distributions over the valuations of the buyers. However, a crucial assumption underlying this result is that the buyers are not able to acquire any additional information about each other. Our framework allows us to model the possibility of information acquisition in a general way. We provide necessary and sufficient conditions on the information acquisition technology for the seller to be able to fully extract rent. Unlike in the standard model, these conditions are not generically satisfied. Indeed, under suitable continuity assumptions, the set of information acquisition technologies that allow full rent extraction is negligible from a topological point of view.

Keywords: Rent extraction, information acquisition, auctions

JEL classification: D44, D82

*School of Economics, University of Surrey, UK. Email: k.laohakunakorn@surrey.ac.uk. I am extremely grateful to Andrew Ellis, Péter Eső, Francesco Nava, Ronny Razin, and Balázs Szentes for helpful comments and suggestions.

1 Introduction

Agents with private information who interact in strategic situations often have incentives to acquire information about each other before making their decisions. This paper explores the implications of this observation for a seller in an auction setting who wishes to implement some allocation rule and fully extract rent from the buyers.

We model information acquisition in a general way by introducing a set of additional signals and a function — the information acquisition technology — that determines the joint distribution of signals for each information acquisition decision of the buyers. This approach represents the acquisition of higher order information in a simple and tractable manner and avoids the usual infinite regress problem that arises in a model with learning about the learning of others. We characterise the set of information acquisition technologies that allow the seller to fully extract rent, and we show that this set is small in a topological sense when agents have a large set of information acquisition actions to choose from.

In standard models of auctions with independent private values, buyers earn positive rents. The economic intuition is that such rents are necessary to incentivise the buyers to reveal their private information. Although appealing, this intuition seemingly depends on the assumption that the private information is independently distributed. Indeed, it is well known that in the absence of information acquisition, full rent extraction is possible for generic prior distributions of the buyers' valuations (Crémer and McLean, 1985, 1988), for example, by using a mechanism where each buyer pays her expected valuation for the object plus a 'side bet' with the seller about the types of the other buyers.[1] However, under this 'side bet' mechanism, each buyer's payoff will depend on the types of the other buyers, and the payment can be very large for some types and very small for others. This implies that buyers may have an incentive to learn about the types of their opponents and, having acquired

[1] As long the valuations are correlated, it is possible to design these bets to have zero expected value under the belief of the reported type, and strictly positive expected value under the belief of every other type. As the bets become large, the incentive constraints will be satisfied, since any type that pretends to be another type will have to pay the side bet, which can be made arbitrarily large in expectation.

this information, misreport their own type or drop out of the mechanism altogether. Thus, this 'side bet' mechanism may fail to fully extract rent if buyers are able to acquire information.

Nevertheless, when information acquisition is allowed, the seller may still be able to fully extract rent using some other mechanism that exploits the buyers' ability to acquire information. For example, if each buyer can perfectly learn the other buyers' types, the seller can fully extract rent using a mechanism in which each buyer reports the types of all the other buyers.[2] On the other hand, if the information acquisition technology is such that each buyer, with probability less than one, receives a signal which perfectly reveals the types of the other buyers (including whether each of them has received a signal) and otherwise learns nothing, then the seller cannot always fully extract rent. For example, a high type who learns that her opponents are low types and have no signal has the same beliefs as a low type with the same information and must earn positive rents if the low type is to be allocated the object with positive probability.

Thus, the ability to acquire information may help or hurt the buyers depending on the information acquisition technology that is available. It is then natural to ask which information acquisition technologies can guarantee full rent extraction for the seller. To capture different information acquisition possibilities in a general way, we introduce a framework where the buyers observe not only their private valuations but also some payoff irrelevant signals that may be correlated with their opponents' types (where a type is now both the valuation and the additional signal). After observing their valuations, each buyer covertly chooses an information acquisition action. The profile of the chosen information acquisition actions and the realised valuations then determine the joint distribution of the payoff irrelevant signals according to the *information acquisition technology* $\sigma : A \times V \to \Delta S$, where A is the set of information acquisition actions, V is the set of valuations, and S is the set of payoff irrelevant signals.[3]

[2] Such reports can then be used to discipline each buyer to report her own valuation truthfully.

[3] V and S are assumed to be finite, but we make no restrictions on A. In Section 6, we provide an explicit specification of A and σ which demonstrates the conceptual advantage of allowing A to

The information acquisition technology $\sigma : A \times V \to \Delta S$ determines the information acquisition possibilities in our model. For example, if one wishes to assume that the buyers can learn about each other's valuations but not directly about each other's signals, then one can impose the restriction that the signals are independent conditional on each $(a, v) \in A \times V$, according to $\sigma(a, v) \in \Delta S$.[4] However, the advantage of our formulation is that such an assumption is not necessary and we can also capture the buyers' incentives to learn directly about each other's information. Our concern in this paper is to ask (i) which information acquisition technologies allow full rent extraction and (ii) whether a "generic" information acquisition technology allows full rent extraction. We show that:

(i) An information acquisition technology guarantees full rent extraction if and only if there exists a profile of information acquisition strategies[5] such that the belief of each realised type can be separated from the beliefs of *all possible types* arising from *alternative* information acquisition strategies and *strongly* separated from such beliefs belonging to types with strictly higher valuations.

(ii) If the set of information acquisition actions is sufficiently rich, the information acquisition technology is continuous, and the signal distributions have full support, then the set of information acquisition technologies that allow full rent extraction is nowhere dense, i.e. *full rent extraction is generically impossible.*

Thus, under assumptions which capture the idea that information acquisition is covert (i.e. deviations are never detected for sure) and that buyers are able to make small adjustments to their beliefs if they desire, the possibility of information acquisition reverses the standard result that full rent extraction is generically possible and rescues the economic intuition that positive rents are necessary to incentivise agents to reveal their private information.

We address the question of when the seller can fully extract rent using the tools

be infinite.

[4]I.e. $\sigma(a, v)$ is the product measure $\prod_i \sigma_{S_i}(a, v)$, where $\sigma_{S_i}(a, v)$ is the marginal of $\sigma(a, v)$ on S_i.

[5]For each player, an information acquisition strategy specifies an information acquisition action for each of her valuations.

of mechanism design. However, a nonstandard feature of our setup is that the distribution over the type space is not exogenously given, but determined as the result of buyers' choices. In particular, the buyers choose their information optimally, given the seller's choice of mechanism. Furthermore, the buyers are able to deviate not only by misreporting their types in the mechanism, but also by choosing alternative information acquisition actions. Thus, in designing the mechanism, the seller must take into account the optimal information choices of the buyers.

Following the optimal choice of information, the standard revelation principle implies that we can restrict attention to incentive compatible direct mechanisms. However, when buyers deviate to other information acquisition actions they will not necessarily report truthfully. We define a game between the buyers at the ex ante stage where the payoff to each information acquisition strategy is the expectation of the maximum payoff from the mechanism, given the profile of information acquisition strategies, and we require that the choice of information is a Nash equilibrium of this game. This optimality requirement is justified by a revelation principle for our environment. A seller who wishes to implement a particular allocation rule can fully extract rent if there exists a mechanism that implements that allocation rule such that following the optimal choice of information, the buyers report truthfully and receive zero payoffs.

Our main result, Proposition 1, establishes that the information acquisition technology allows the seller to fully extract rent if and only if there exists an information acquisition strategy such that the belief of each type that arises from the strategy satisfies two conditions. The first condition ensures that there exists a lottery with zero expected value for that type, and *weakly* positive expected value for any type that could arise from a unilateral deviation to an alternative information acquisition strategy. The second condition ensures that this lottery can be chosen to have *strictly* positive expected value that is bounded away from zero for any type *with a strictly higher valuation* that could arise from a unilateral deviation to an alternative information acquisition strategy.

Our conditions involve the beliefs of all types that could arise from any information acquisition strategy. This ensures that there is no profitable deviation to any alternative information acquisition strategy and makes them stronger than the analogous conditions in the standard model, since there can be (infinitely) many beliefs that arise from any information acquisition strategy. The first condition, for example, requires that the beliefs following the optimal choice of information must be on the (relative) boundary of the convex hull of the set of all beliefs that could arise from unilateral deviations to alternative information acquisition strategies. This requirement becomes demanding when there are many beliefs arising from alternative information acquisition strategies. Indeed, Proposition 2 shows that this condition fails "generically" when the set of information acquisition actions is sufficiently rich.

2 Related Literature

The first example of a side bet mechanism that allows the seller to fully extract surplus appears in Myerson (1981). A series of papers, starting with Crémer and McLean (1985) argue that full surplus extraction is generically possible in various mechanism design settings.[6] Subsequently, many papers have considered the limits of these results.[7] Our paper contributes to this line of inquiry by focusing on the incentive effects of large side bets. In particular, we maintain the preferences and the contracting environment of the standard model, but we allow buyers to acquire additional information and require that they do so optimally given the mechanism.[8]

Several papers have considered mechanism design with information acquisition;

[6]Crémer and McLean (1988) provide necessary and sufficient conditions for full surplus extraction to be Bayesian incentive compatible in an auction setting, and argue that these conditions are generically satisfied. McAfee, McMillan, and Reny (1989) show that when buyers have continuous valuations, almost full surplus extraction is possible in a common value auction, and McAfee and Reny (1992) extend their result to general mechanism design problems.

[7]For example, Robert (1991) introduces limited liability and risk aversion, and shows that full surplus extraction fails. Laffont and Martimort (2000) consider Collusion Proofness, and Lopomo, Rigotti, and Shannon (2014) and Renou (2015) show that ambiguity aversion of the buyers prevents full surplus extraction from being a generic property.

[8]Yamashita (2018) considers a related question about the highest revenue the seller can guarantee when the buyers have access to additional information. Rather than assuming that buyers acquire information optimally, his seller considers the worst case information structure for each choice of mechanism; in this case, the second price auction achieves the highest revenue guarantee among efficient mechanisms.

however, most of these only allow agents to learn about the state of the world and not directly about others' beliefs about the state. For example, Bikhchandani (2011) provides necessary and sufficient conditions for the existence of full extraction lotteries that are robust to the possibility of information acquisition.[9] However, in Bikhchandani (2011), buyers can only acquire signals that are independent conditional on the profile of valuations. Our model differs in that the information acquisition stage is modelled as game, and we allow buyers to acquire information not just about each other's valuations, but also about each other's information. This is important because if buyers are able to acquire information only about each other's valuations, the seller can fully extract rent by offering side bets on the information instead. However, in such mechanisms, there is then a strong incentive to acquire information about each other's information.

More recently, some papers have studied more general models of information acquisition. Pernoud and Gleyze (2023) consider a mechanism design setting where agents can learn about their preferences at a cost and show that for typical mechanisms, they will choose to learn about others' preferences as well as their own (since this helps predict others' reports to the designer). Denti (2023) considers a model of costly information acquisition in games where players can learn directly about each other's signals as well as the state, and introduces an equilibrium concept which is robust to such information acquisition. Unlike in these papers, we assume that information acquisition actions are costless (or, equivalently, all have the same cost), but that the joint signal distribution depends on the information choices of all players (through the information acquisition technology); thus, what prevents agents from choosing, for example, to become perfectly informed in our model is that they are constrained by the choices of others (and the specific information acquisition technology).

Beyond information acquisition, buyers' ability to influence the underlying distribution of valuations is another reason why full rent extraction may fail. For exam-

[9] See also Bikhchandani and Obara (2017), who provide sufficient conditions for efficient implementation and full surplus extraction when buyers are able to acquire information about an unknown payoff relevant state of nature; and Bergemann and Välimaäki (2002), who show that with independent private values and information acquisition about own preferences only, the VCG mechanism induces efficient information acquisition.

ple, Obara (2008) extends the Crémer and McLean (1988) result to a setting where agents take some hidden action that determines the distribution of their payoff relevant types. Obara (2008) provides necessary and sufficient conditions for full surplus extraction, and argues that the conditions are not necessarily satisfied when there are many actions to which the agents can deviate. Unlike in Obara (2008), the actions that agents take in our model affect only their information and not their valuations; thus, our model focuses purely on informational incentives. Moreover, we show that when the action space is sufficiently rich, the conditions for full rent extraction fail generically. On a technical level, unlike the papers by Obara (2008), Bikhchandani (2011), and Bikhchandani and Obara (2017), we do not restrict the action space to be finite, which is important for our genericity result.[10]

The rest of the paper proceeds as follows. In Section 3 we introduce the model of information acquisition and motivate our definition of full rent extraction which is justified by a revelation principle. Section 4 contains several examples that illustrate the flexibility of our framework. Our main results are presented in Sections 5 and 7. Section 5 provides a complete characterisation of the set of information structures that guarantee full rent extraction for the seller; these conditions are illustrated in Section 6 for a class of "linear" information structures. Section 7 shows that under natural assumptions, the set of information structures that guarantee full rent extraction is nowhere dense. All proofs are contained in the Appendix.

3 Model

We first specify the underlying auction environment. There is a finite set I of n buyers who may buy a single indivisible object from a seller. Each $i \in I$ has a payoff relevant valuation $v_i \in V_i$, where V_i is a finite set. The profile of valuations

[10] An alternative perspective is provided by Heifetz and Neeman (2006), who argue that the result that full surplus extraction is possible for generic type spaces of a fixed and finite size that admit a common prior hinges on the nonconvexity of the set of priors, and generic priors on the universal type space do not allow for full surplus extraction. In our model, the type space is endogenous and depends on the information acquisition decisions of the buyers.

$v \in V \equiv \prod_{i \in I} V_i$ is distributed according to a full support distribution $\pi \in \Delta V$.[11] The valuations of the buyers are converted to monetary units by a willingness to pay function $w : \cup_{i \in I} V_i \to \mathbb{R}_+$,[12] and utility is quasilinear in transfers; that is, if buyer i has valuation v_i, receives the object with probability x, and pays t to the seller, then buyer i's payoff is $xw(v_i) - t$. Let $\mathcal{W}$ be the set of all valuation functions, i.e. the set of all functions $w : \cup_{i \in I} V_i \to \mathbb{R}_+$.

We now introduce our framework for modelling information acquisition. After observing v_i, each buyer covertly chooses an information acquisition action $a_i \in A_i$, where A_i can be an arbitrary set (i.e. we do not assume that A_i is finite). Following the choice of a_i, each buyer receives a payoff irrelevant signal $s_i \in S_i$, where S_i is a finite set. Let $S \equiv \prod_{i \in I} S_i$ and $A \equiv \prod_{i \in I} A_i$. Our innovation is to introduce a function $\sigma : A \times V \to \Delta S$, which we call the *information acquisition technology*, that determines a distribution over signals $\sigma(a, v) \in \Delta S$ for each $(a, v) \in A \times V$.

Our framework is flexible enough to incorporate many relevant aspects of learning. Since each buyer's signal can be correlated with the valuations of the other buyers, it is possible to learn about these valuations. But since the signals themselves are correlated conditional on the valuations, it is also possible for each buyer to learn about the information that the other buyers have. Moreover, since the distribution depends on the entire action profile a, buyers may be able to take actions (depending on the functional form) that prevent others from learning anything about them or otherwise manipulate their information.

Let $\alpha_i : V_i \to A_i$ be an information acquisition (pure) strategy for buyer i. We refer to $\theta_i = (v_i, s_i) \in V_i \times S_i \equiv \Theta_i$ as buyer i's type. Define $\gamma : A^{|V|} \to \Delta(V \times S)$ such that $\gamma(\alpha)[v, s] = \pi[v]\sigma(\alpha(v), v)[s]$. That is, for any profile of information acquisition strategies α, the resulting distribution over the set of types $\Theta \equiv V \times S$ is given by

[11] Whenever Y is a finite set, ΔY denotes the set of all probability distributions on Y. For each $\mu \in \Delta Y$ and $y \in Y$, we write $\mu[y]$ to denote the probability of y according to μ and $\text{supp}(\mu)$ to denote the support of μ. For each $y \in Y$, we write $1_y \in \Delta Y$ to denote the probability distribution degenerate on y. When $Y = \prod Y_i$ and $\mu \in \Delta Y$, we write $\mu_{Y_i} \in \Delta Y_i$ to denote the marginal on Y_i and, for $y_i \in \text{supp}(\mu_{Y_i})$, we write $\mu(y_i) \in \Delta Y_{-i}$ to denote the distribution conditional on y_i. For convenience, we also treat μ as an element of $\mathbb{R}^{|Y|}$. Thus, enumerating $Y = \{y^1, \ldots, y^{|Y|}\}$, $\mu_i = \mu[y^i]$ and for $x \in \mathbb{R}^{|Y|}$, $\mu \cdot x = \sum_{i=1}^{|Y|} \mu_i x_i$.

[12] As in Crémer and McLean (1988), the relevance of w is that it will vary as the V_is remain fixed.

$\gamma(\alpha)$. Let $\gamma(\theta_i, \alpha)$ denote this distribution conditional on θ_i.

Remark 1 *Our framework can incorporate mixed strategies over a finite set of actions as follows. Fix a finite set of actions A^f, a finite set of signals S^f, and an information acquisition function $\sigma^f : A^f \times V \to \Delta S^f$. Now suppose that each i may randomise among actions in A_i^f. This is equivalent to letting each i choose a pure strategy from $A_i = \Delta A_i^f$, with $S_i = S_i^f \times A_i^f$ and $\sigma : A \times V \to \Delta S$ given by $\sigma(a, v)[s] = \sigma^f(a^f, v)[s^f] a_1[a_1^f] \ldots a_n[a_n^f]$.*

A direct mechanism (α, x, t) is a (recommended) strategy profile $\alpha \in A^{|V|}$, an allocation rule $x : V \times S \to \Delta I$ and a transfer rule $t : V \times S \to \mathbb{R}^n$. The timing is as follows:

1. Seller commits to a mechanism.
2. Each buyer observes v_i and chooses information acquisition action $\alpha_i(v_i)$.
3. Each buyer observes s_i and reports $\theta_i = (v_i, s_i)$ to seller.
4. Seller implements the mechanism and payoffs are realised.

Note that the seller commits to a mechanism first, and then the buyers choose their information acquisition actions. Below, we impose the requirement on the mechanism that the information acquisition action $\alpha_i(v_i)$ and report θ_i are optimal for each buyer.

For each $\alpha \in A^{|V|}$, $\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha))$ and $\theta_i' \in \Theta_i$, define $U_i(\theta_i', \theta_i; \alpha)$ as i's expected utility when player i reports θ_i', player i is type θ_i, and the players are following the information acquisition strategy α. That is:

$$U_i(\theta_i', \theta_i; \alpha) = \sum_{\theta_{-i}} \gamma(\theta_i, \alpha)[\theta_{-i}] \left(x_i(\theta_i', \theta_{-i}) w(v_i) - t_i(\theta_i', \theta_{-i}) \right).$$

Note that $\gamma(\theta_i, \alpha)[\theta_{-i}] = \frac{\pi(v_i)[v_{-i}]\sigma(\alpha(v),v)[s]}{\sum_{v'_{-i}, s'_{-i}} \pi(v_i)[v'_{-i}]\sigma(\alpha_i(v_i), \alpha_{-i}(v'_{-i}), v_i, v'_{-i})[s_i, s'_{-i}]}$. Thus, $U_i(\theta_i', \theta_i; \alpha)$ depends only on $\alpha_i(v_i)$ and α_{-i}, but to simplify notation we let $U_i(\theta_i', \theta_i; \alpha)$ depend on the entire strategy α. Also note that given α, $U_i(\theta_i', \theta_i; \alpha)$ is well-defined only for

θ_i such that $\gamma_{\Theta_i}(\alpha)[\theta_i] > 0$; hence our restriction to $\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha))$. Finally, define $U_i^*(\theta_i; \alpha) = \max_{\theta_i'} U_i(\theta_i', \theta_i; \alpha)$. Note that the maximum exists since Θ_i is finite.

Given a mechanism (α, x, t), we impose the following requirements:

(i) *Buyers optimally acquire information* if α is a Nash equilibrium of the game given by payoffs:[13]

$$\tilde{u}_i(\alpha_i, \alpha_{-i}) = \sum_{\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha_i, \alpha_{-i}))} \gamma_{\Theta_i}(\alpha_i, \alpha_{-i})[\theta_i] \max\{U_i^*(\theta_i; \alpha_i, \alpha_{-i}), 0\}.$$

(ii) The mechanism is *incentive compatible* if for all $\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha))$ and for all $\theta_i' \in \Theta_i$:

$$U_i(\theta_i, \theta_i; \alpha) \geq U_i(\theta_i', \theta_i; \alpha).$$

(iii) The mechanism *fully extracts rents* if for all $\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha))$:

$$U_i(\theta_i, \theta_i; \alpha) = 0.$$

An *information structure* is $(\Theta, \pi, (A, \sigma))$, and for each valuation function w, $(\Theta, \pi, (A, \sigma), w)$ is the associated allocation problem.

Let $\mathcal{R}$ be the set of all complete and transitive binary relations on $\cup_i V_i$. For $\succeq \in \mathcal{R}$, let $\mathcal{W}(\succeq)$ be the set of valuation functions consistent with $\succeq$, i.e. the set of $w \in \mathcal{W}$ such that $w(v) \geq w(v')$ if and only if $v \succeq v'$, for $v, v' \in \cup_i V_i$.

A *social choice function* $f : V \to \Delta I$ maps a profile of payoff relevant types to a probability distribution over the set of players.

Definition 1 *The information structure* $(\Theta, \pi, (A, \sigma))$ *guarantees full rent extraction for the ordering* $\succeq$ *and the social choice function* $f : V \to \Delta I$ *if for all* $w \in \mathcal{W}(\succeq)$, *in the allocation problem* $(\Theta, \pi, (A, \sigma)), w)$, *there exists a mechanism* (α, x, t) *such that:*

- (α, x, t) *is incentive compatible, fully extracts rent, and buyers optimally acquire information, and*

[13]This definition assumes that buyers may drop out after acquiring information and before reporting in the mechanism.

- $x(v, s) = f(v)$ *for all* $(v, s) \in \text{supp}(\gamma(\alpha))$.

We will say that a mechanism (α, x, t) is *consistent* with f if $x(v, s) = f(v)$ for all $(v, s) \in \text{supp}(\gamma(\alpha))$.

Remark 2 *Note that if (x, t) is the second price auction, then for any α, (α, x, t) is incentive compatible and buyers optimally acquire information. For an arbitrary (x, t), however, it may be that the game given by payoffs $\tilde{u}$ has no Nash equilibrium, and hence the requirement that buyers optimally acquire information cannot be satisfied. We do not view this as a problem since the nonexistence of a Nash equilibrium at the information acquisition stage given (x, t) implies that the seller should not use (x, t) as the allocation and transfer rules if she wishes to guarantee that the buyers get zero payoff.*

Definition 1 of full rent extraction requires that the seller is able to fully extract rent for all possible valuation functions $w \in \mathcal{W}(\succeq)$. This approach is standard in the literature and simplifies the conditions on beliefs. For example, Crémer and McLean (1988) prove that $\pi(v_i) \notin co\{\pi(v_i') : v_i' \neq v_i\}$ for each $v_i \in V_i$ is a necessary and sufficient condition for the seller to fully extract rent from player i. However, for fixed w, the condition is not necessary.[14]

Definition 1 is justified by a revelation principle, which is formally stated in the Appendix. In general, the buyers' reporting strategy may depend on their information acquisition strategy; Proposition A.1 implies that it is without loss to focus on truthful reports following the recommended information acquisition strategy.

[14]Indeed, suppose that $V_1 = \{0, 1\}$, $V_2 = \{2, 3, 4\}$, w is the identity, and we wish to always allocate to player 2 and fully extract rent, i.e. we need to charge each type v_2 exactly $w(v_2)$ in expectation. Suppose that $\pi(2)[1] = 0.3$, $\pi(3)[1] = 0.5$ and $\pi(4)[1] = 0.75$. Note that $\pi(3) \in co\{\pi(2), \pi(4)\}$. However, as we now show, full rent extraction is possible. Suppose that type 4 makes a constant payment of 4, and that type 2 pays 2 in expectation, but with a large payment when $v_1 = 1$ so that that type 2's payment lottery is unattractive to the other types. Consider $\tau \in \Delta V_1$ such that $\tau[0] = 3 - \varepsilon$ and $\tau[1] = 3 + \varepsilon$. Note that $\pi(3) \cdot \tau = 3$ and if $\varepsilon \geq 2$, then $\pi(4) \cdot \tau \geq 4$, and if $\varepsilon \leq 2.5$, then $\pi(2) \cdot \tau \geq 2$. Thus, for $\varepsilon \in [2, 2.5]$, there exists a lottery for type 3 that neither type 2 nor type 4 prefers to her own, and hence full rent extraction is possible. However, as $w(4)$ becomes large, the size of the lottery required to prevent type 4 from deviating becomes larger, and eventually, will cause type 2 to prefer it to her own. Thus, under these beliefs, full rent extraction fails for some w.

4 Examples

In the next section, we will establish necessary and sufficient conditions for the information structure to guarantee full rent extraction. First, we provide some simple examples to illustrate our setup.

Example 1 (Crémer-McLean) *Suppose $S_i = \varnothing$ for each i, so there is no possibility of information acquisition. Then $((V, \varnothing), \pi, (A, \sigma))$ guarantees full rent extraction for all orderings $\succeq$ and all social choice functions $f : V \to \Delta I$ if and only if for all v_i:*

$$\pi(v_i) \notin co\{\pi(v_i') : v_i' \neq v_i\}.$$

Note that Crémer and McLean's (1988) condition is necessary and sufficient for the information structure to guarantee full rent extraction for all orderings $\succeq$. For a fixed ordering $\succeq$, full rent extraction is possible even if $\pi(v_i) \in co\{\pi(v_i') : v_i' \neq v_i\}$ for some types v_i, as long as they satisfy $v_i \succeq v_i'$ for all $v_i' \in V_i$.[15]

With information acquisition, π exhibiting correlation is not sufficient for full rent extraction, as the following example shows:

Example 2 *Let $n = 2$, $V_i = \{v^L, v^H\}$, $S_i = \{\varnothing, (v^L, \varnothing), (v^L, s), (v^H, \varnothing), (v^H, s)\}$, $w(v^L) = \underline{v} > 0$, $w(v^H) = 1$, $\pi > 0$ (π can be any full support distribution), $A_i = \{N, Y\}$. If $a_i = N$, then player i receives the null signal $\varnothing$. If $a_i = Y$, then with probability $p < 1$ player i receives a signal that perfectly reveals the valuation of her opponent and whether her opponent has received a null signal. Players receive signals independently. Note that for player i, types $(v^L, (v^L, \varnothing))$ and $(v^H, (v^L, \varnothing))$ have the same beliefs over Θ_{-i}, namely that player $-i$ has valuation v^L and the null signal. Thus, there exists an allocation rule where type $(v^H, (v^L, \varnothing))$ of player i must*

[15] Consider the example from footnote 14 but with $\pi(2)[1] = 0.4$, $\pi(3)[1] = 0.6$, and $\pi(4)[1] = 0.5$, i.e. $\pi(4) \in co\{\pi(2), \pi(3)\}$. Note that if type 4 has to make a constant payment equal to $w(4)$, neither type 2 nor type 3 will want to report that she is type 4 as long as $w(2) < w(3) < w(4)$. Since there exists a hyperplane separating $\pi(2)$ from $\pi(3)$ and $\pi(4)$ (resp. $\pi(3)$ from $\pi(2)$ and $\pi(4)$), there exists a payment lottery for type 2 (resp. type 3) charging $w(2)$ (resp. $w(3)$) in expectation that can be made arbitrarily unattractive to other types; hence full rent extraction is possible for all w consistent with the ordering $4 \succ 3 \succ 2$. However, by considering orderings where 4's valuation is not the highest, full rent extraction will fail for some w.

receive positive rents. For example, let $x_1((v^L,(v^L,\varnothing)),(v^L,\varnothing)) \equiv x > 0$*. Since type* $(v^L,(v^L,\varnothing))$ *of player* 1 *knows the type of player* 2 *for sure, full rent extraction implies* $t_1((v^L,(v^L,\varnothing)),(v^L,\varnothing)) = x\underline{v}$*. Now a possible deviation is for player* 1 *to choose* $\alpha_1(v^H) = Y$*, drop out for all types other than* $(v^H,(v^L,\varnothing))$*, and for type* $(v^H,(v^L,\varnothing))$*, report* $(v^L,(v^L,\varnothing))$*. The utility of this deviation is at least* $\pi[v^H,v^L]p(1-p)x(1-\underline{v}) > 0$*; thus incentive compatibility requires that player* 1 *gets positive rents.*

On the other hand, with information acquisition, sometimes full rent extraction is possible even when π is independent:

Example 3 *Let* $n = 2$*,* $V_i = \{v^L, v^H\}$*,* $S_i = \{v^L, v^H\}$*,* π *uniform,* $A_i = \{N, Y\}$*. If* $a_i = N$*, then player* i*'s signal is uninformative. If* $a_i = Y$*, player* i*'s signal perfectly reveals the valuation of her opponent. That is, let* $\sigma(v_i, v_{-i}, (Y,Y)) = 1_{(v_{-i},v_i)}$ *and* $\sigma(v_i, v_{-i}, (N,Y)) = \frac{1}{2}1_{(v_{-i},v_i)} + \frac{1}{2}1_{(v'_{-i},v_i)}$*, where* $v'_{-i} \neq v_{-i}$*. For any* x *and for any* w*, the following mechanism, with* $\alpha_i(v_i) = Y$ *for all* v_i*, is incentive compatible, fully extracts rent, and buyers optimally acquire information:*

- $t_i((v_i,s_i),(v_{-i},s_{-i})) = x_i((v_i,s_i),(v_{-i},s_{-i}))w(v_i)$ *if* $v_i = s_{-i}$ *and* $s_i = v_{-i}$*,*
- $t_i((v_i,s_i),(v_{-i},s_{-i})) = \infty$ *otherwise.*[16]

5 Characterisation of Full Rent Extraction

In this section, we will characterise the necessary and sufficient conditions for the information structure $(\Theta, \pi, (A, \sigma))$ to guarantee full rent extraction. First, we introduce the notation required to state our conditions.

For each information acquisition strategy α, social choice function f, and $\theta_i \in \Theta_i$,

[16] More precisely, $t_i((v_i,s_i),(v_{-i},s_{-i})) = K$, for $K \in \mathbb{R}$ sufficiently large (how large K needs to be will depend on x and w).

define:

$$\Lambda_i^f(\theta_i, \alpha) = \{\theta_{-i} \in \Theta_{-i} : \gamma(\alpha)[\theta] f_i(v) > 0\},$$

$$\Theta_i^f(\alpha) = \{\theta_i \in \Theta_i : \Lambda_i^f(\theta_i, \alpha) \neq \varnothing\} = \left\{\theta_i \in \Theta_i : \sum_{\theta_{-i}} \gamma(\alpha)[\theta] f_i(v) > 0\right\}.$$

That is, $\Theta_i^f(\alpha)$ is the subset of $\text{supp}(\gamma_{\Theta_i}(\alpha))$ such that buyer i receives the object with positive probability according to f and, for each $\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha))$, $\Lambda_i^f(\theta_i, \alpha)$ is the subset of $\text{supp}(\gamma(\theta_i, \alpha))$ where buyer i is allocated the object with positive probability according to f.[17] Note that following α, any mechanism that is consistent with f must have $x_i(\theta_i, \theta_{-i}) > 0$ whenever $\theta_{-i} \in \Lambda_i^f(\theta_i, \alpha)$. In other words, $\Theta_i^f(\alpha)$ is the set of θ_i that receives the object and $\Lambda_i^f(\theta_i, \alpha)$ is the set of θ_{-i} when this happens.

For each $i \in N$, $\alpha_{-i} \in A_{-i}^{|V_{-i}|}$, $\theta_i = (v_i, s_i) \in \Theta_i$, $\succeq \in \mathcal{R}$, and $M \subseteq \Theta_{-i}$, let:

- $C(\alpha_{-i}) = co\{\gamma(\theta_i', \alpha_i', \alpha_{-i}) : \alpha_i' \in A_i^{|V_i|}, \theta_i' \in \text{supp}(\gamma_{\Theta_i}(\alpha_i', \alpha_{-i}))\}$,
- $D(\theta_i, \alpha_{-i}, \succeq, M) = \left\{\gamma(\theta_i', \alpha_i', \alpha_{-i}) : \begin{array}{c} \alpha_i' \in A_i^{|V_i|}, \theta_i' \in \text{supp}(\gamma_{\Theta_i}(\alpha_i', \alpha_{-i})), \\ v_i' \succ v_i, \\ \gamma(\theta_i', \alpha_i', \alpha_{-i})[M] > 0 \end{array}\right\}$,
- $V_i^*(\succeq) = \{v_i : v_i \succeq v_i' \text{ for all } v_i' \in V_i\}$.

The set $C(\alpha_{-i})$ is the convex hull of the beliefs of buyer i that can arise from any information acquisition strategy of buyer i, fixing the information acquisition strategy of the other buyers at α_{-i}. The set $D(\theta_i, \alpha_{-i}, \succeq, M)$ is the subset of $C(\alpha_{-i})$ that contains only the beliefs that assign positive probability to M and belong to some type with valuation strictly greater than v_i under $\succeq$. The set $V_i^*(\succeq)$ is the set of greatest elements of V_i under the relation $\succeq$.

For any $X \subseteq \mathbb{R}^m$, let $\overline{X}$ denote the closure of X, and let $\text{ri}(X)$ denote the relative interior of X. Recall that for any $X \subseteq \mathbb{R}^m$ and $\bar{x} \in X$, $p \neq 0$ *supports* X at $\bar{x}$ if $p \cdot y \geq p \cdot \bar{x}$ for all $y \in X$. The hyperplane $\{y \in \mathbb{R}^m : p \cdot y = p \cdot x\}$ is a *supporting hyperplane* for X at $\bar{x}$, and the support is *proper* if $p \cdot y > p \cdot \bar{x}$ for some $y \in X$.

[17] For $\theta_i \notin \text{supp}(\gamma_{\Theta_i}(\alpha))$, $\Lambda_i^f(\theta_i, \alpha) = \varnothing$.

Definition 2 *For any convex set $K \subseteq \mathbb{R}^m$, F is an exposed face of K if F satisfies any of the following:*

1. *There exists a hyperplane H supporting K at some $\bar{x} \in K$ and $F = K \cap H$,*

2. *$F = K$,*

3. *$F = \varnothing$.*

If $F \neq K \neq \varnothing$, then F is a proper exposed face of K.

Definition 3 *For any convex set $K \subseteq \mathbb{R}^m$ and $\bar{x} \in K$, let $F_K(\bar{x})$ be the intersection of all exposed faces of K containing $\bar{x}$.*

We are now ready to state our main result, which characterises the information structures that guarantee full rent extraction.

Proposition 1 *The information structure $(\Theta, \pi, (A, \sigma))$ guarantees full rent extraction for the ordering $\succeq$ and the social choice function $f : V \to \Delta I$ if and only if there exists an α such that for all i and for all $\theta_i \in \Theta_i^f(\alpha) \setminus (V_i^*(\succeq) \times S_i)$:*

1. *$\gamma(\theta_i, \alpha) \notin \mathrm{ri}(C(\alpha_{-i}))$,*

2. *$F_{\overline{C}(\alpha_{-i})}(\gamma(\theta_i, \alpha)) \cap \overline{D}(\theta_i, \alpha_{-i}, \succeq, \Lambda_i^f(\theta_i, \alpha)) = \varnothing$.*[18]

The proof of Proposition 1 uses type-specific lotteries $\tau(\theta_i)$ familiar from Crémer and McLean (1988) to deter other types from reporting θ_i. But note that these lotteries are not required for types that do not receive the object (following the recommended information acquisition strategy α) or those that have the $\succeq$-greatest valuation in V_i; hence, the conditions are imposed only on $\theta_i \in \Theta_i^f(\alpha) \setminus (V_i^*(\succeq) \times S_i)$.

For certain valuation functions w, the absolute value of these lotteries may need to be very large; Condition 1 is equivalent to the existence of a hyperplane properly supporting $C(\alpha_{-i})$ at $\gamma(\theta_i, \alpha)$, which implies that there are lotteries $\tau(\theta_i)$ that

[18]When Condition 1 is satisfied, there is always a hyperplane properly supporting $\overline{C}(\alpha_{-i})$ at $\gamma(\theta_i, \alpha)$, so $F_{\overline{C}(\alpha_{-i})}(\gamma(\theta_i, \alpha))$ is a proper exposed face of $\overline{C}(\alpha_{-i})$. On the other hand when Condition 1 fails, $F_{\overline{C}(\alpha_{-i})}(\gamma(\theta_i, \alpha)) = \overline{C}(\alpha_{-i})$, and Condition 2 cannot hold.

will not become attractive for other types, however large they become. These other types include all types that can arise from any unilateral deviation to an alternative information acquisition strategy. Hence, Condition 1 is stronger than the analogous condition in the standard model to the extent that there can be (possibly infinitely) many types that arise from any information acquisition strategy, even when there are a small number of types that arise in equilibrium.

Under condition 1, there are hyperplanes that properly support $\overline{C}(\alpha_{-i})$ at $\gamma(\theta_i, \alpha)$; take the one whose intersection with $\overline{C}(\alpha_{-i})$ is minimal, i.e. where the intersection is the smallest exposed face of $\overline{C}(\alpha_{-i})$ containing $\gamma(\theta_i, \alpha)$. Condition 2 is equivalent to this hyperplane strongly separating $\gamma(\theta_i, \alpha)$ from any belief in $C(\alpha_{-i})$ that assigns positive probability to $\Lambda_i^f(\theta_i, \alpha)$ and belongs to a type with valuation strictly greater than v_i. This ensures that there are lotteries $\tau(\theta_i)$ that would deter such types from reporting θ_i. Types with valuation no greater than v_i or those that believe that θ_i will not be allocated the object do not need to be deterred, and $\tau(\theta_i)$ may have zero expected value for such types.

Example 2 illustrates a situation where Condition 2 fails. In that example, types $(l, (l, \varnothing))$ and $(h, (l, \varnothing))$ have the same beliefs — this is a failure of the *beliefs determine preferences* property introduced by Neeman (2004). More generally, as long as $(h, (l, \varnothing))$'s belief is contained in the smallest exposed face of $\overline{C}(\alpha_{-i})$ containing $(l, (l, \varnothing))$'s belief, these beliefs cannot be separated by any supporting hyperplane at $(l, (l, \varnothing))$. Condition 2 can thus be interpreted as a generalisation of the *beliefs determine preferences* property that imposes additional restrictions on types with beliefs on the same exposed face.

If we want to guarantee full rent extraction for every f, we need to impose Conditions 1 and 2 not only on types that receive the object with positive probability according to a particular f, but on all positive probability types (except those with the greatest valuation). Corollary 1 follows from the observation that $\Theta_i^f(\alpha) \subseteq \mathrm{supp}(\gamma_{\Theta_i}(\alpha))$ and $\Lambda_i^f(\theta_i, \alpha) \subseteq \mathrm{supp}(\gamma(\theta_i, \alpha))$, with equality when $f_i(v) > 0$ for all v.

Corollary 1 *The information structure $(\Theta, \pi, (A, \sigma))$ guarantees full rent extraction for the ordering $\succeq$ and every social choice function $f : V \to \Delta I$ if and only if there* *exists an α such that for all i and for all $\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha)) \setminus (V_i^*(\succeq) \times S_i)$:*

1. $\gamma(\theta_i, \alpha) \notin \text{ri}(C(\alpha_{-i}))$,

2. $F_{\overline{C}(\alpha_{-i})}(\gamma(\theta_i, \alpha)) \cap \overline{D}(\theta_i, \alpha_{-i}, \succeq, \text{supp}(\gamma(\theta_i, \alpha))) = \varnothing$.

In Example 3, when $\alpha_i(v_i) = Y$ for all v_i, we have

$$C(\alpha_{-i}) = co\{1_{(v^L,v^L)}, 1_{(v^H,v^L)}, 1_{(v^L,v^H)}, 1_{(v^H,v^H)}\}$$

and $\gamma(\theta_i, \alpha) = 1_{(s_i,v_i)} \in \Delta(V_{-i} \times S_{-i})$ for each $\theta_i = (v_i, s_i)$. Hence Condition 1 holds. In addition, $\{1_{(s_i,v_i)}\}$ is an exposed face of $C(\alpha_{-i})$ and does not intersect $D(\theta_i, \alpha_{-i}, \succeq, \text{supp}(\gamma(\theta_i, \alpha)))$, since the latter contains only beliefs whose marginal on S_{-i} is concentrated on $v_i' \neq v_i$. Hence, Condition 2 holds (for any $\succeq$).[19]

Remark 3 *In the standard case without information acquisition $S = \varnothing$, $\gamma(\theta_i, \alpha) = \pi(v_i)$, and hence to guarantee full rent extraction for every ordering $\succeq$ and every social choice function, Corollary 1 requires:*

1. *For all $\succeq \in \mathcal{R}$ and for all $v_i \in V_i \setminus V_i^*(\succeq)$, $\pi(v_i) \notin \text{ri}(co\{\pi(v_i') : v_i' \in V_i\})$,*

2. *For all $\succeq \in \mathcal{R}$ and for all $v_i \in V_i \setminus V_i^*(\succeq)$, $F_{co\{\pi(v_i'):v_i' \in V_i\}}(\pi(v_i)) \cap \{\pi(v_i') : v_i' \succ v_i\} = \varnothing$.*

Note that this is equivalent to $\pi(v_i) \notin \text{ri}(co\{\pi(v_i') : v_i' \in V_i\})$ for all $v_i \in V_i$ and $F_{co\{\pi(v_i'):v_i' \in V_i\}}(\pi(v_i)) \cap \{\pi(v_i') : v_i' \neq v_i\} = \varnothing$ for all $v_i \in V_i$. Together they are equivalent to the requirement that, for all v_i:

$$\pi(v_i) \notin co\{\pi(v_i') : v_i' \neq v_i\},$$

as we have stated in Example 1.

The requirement that full rent extraction is possible for all w simplifies the conditions on beliefs in the standard model. In particular, as the previous remark shows,

[19] In this example, both $C(\alpha_{-i})$ and $D(\theta_i, \alpha_{-i}, \succeq, \text{supp}(\gamma(\theta_i, \alpha)))$ are closed, which is always the case when A_i is finite.

the conditions on beliefs do not depend on the valuation of the type holding each belief. However, this reduction is not possible in our setting, since for each permutation of the valuations, the seller can induce a different information acquisition strategy to fully extract rent; hence for each information acquisition strategy, the conditions on beliefs do not need to hold for every permutation of the valuations.

6 Illustration

In this section, we introduce and motivate an explicit specification of $(\Theta, \pi, (A, \sigma))$, to which we apply our conditions. This example also demonstrates the advantage of not restricting A_i to be finite.

For simplicity, assume that $n = 2$. Let $V_1 = V_2 = \{v^1, \ldots, v^K\}$, with $K \geq 2$, and let $\succeq$ be such that $v^K \succ \ldots \succ v^1$. For each i, let S_i be such that $|S_i| \geq K$ and let $\hat{S}_i \subseteq S_i$.

For each i, let A_i be the set of all functions $a_i : V_{-i} \to \Delta(S_i \times \hat{S}_{-i})$.[20] The interpretation is that each player i chooses a joint distribution of signals for both players for each of her opponent's valuations, but player i is only allowed to send player $-i$ signals belonging to $\hat{S}_{-i}$, a subset of S_{-i}.

Fix $\beta \in (0, 1)$, and let $\sigma(a, v) \in \Delta S$ be such that, for each $s \in S$,

$$\sigma(a, v)[s] = \beta a_1(v_2)[s] + (1 - \beta) a_2(v_1)[s].$$

We call any $(\Theta, \pi, (A, \sigma))$ satisfying the above a *linear information structure*. The interpretation is that, for each v, each player i wants the joint distribution of signals to be $a_i(v_{-i})$; with probability β, the signals are distributed according to player 1's wishes, and with probability $1-\beta$, they are distributed according to player 2's wishes. We view this as a reduced form representation of a model where the players can take many different actions to manipulate the information structure to their benefit.[21]

[20] Note that A_i is a compact and convex subset of a Euclidean space.

[21] We have used this formulation in Carmona and Laohakunakorn (2023) to model a repeated game where the players choose the monitoring structure in each period.

First consider the case where $\hat{S}_i = S_i$ for each S_i. Consider the following information acquisition strategy α. Partition $S_1 = S_2$ into $\cup_{j=1}^{K} S^j$, and for each $k, l \in \{1, \ldots, K\}$, let $\alpha_i(v^k) = a_i$ such that $a_i(v^l) = \mathcal{U}_{S^l \times S^k}$, where for each finite set Y, $\mathcal{U}_Y$ is the uniform distribution on Y. Note that, for each (v_i, s_i) such that $v_i = v^k$ and $s_i \in S^l$, player i knows that player $-i$ has valuation $v_{-i} = v^l$ and signal $s_{-i} \in S^k$. In the Appendix, we show that, for each $\theta_i \in \Theta_i$, $\gamma(\theta_i, \alpha)$ does not belong to the relative interior of $C(\alpha_{-i})$ and for every belief $\gamma(\theta_i', \alpha_i', \alpha_{-i})$ belonging to a type with $v_i' \neq v_i$, $\gamma(\theta_i', \alpha_i', \alpha_{-i}) \notin F_{\overline{C}(\alpha_{-i})}(\gamma(\theta_i, \alpha))$; thus, full rent extraction is possible by Corollary 1.

Corollary 2 *For a linear information structure, if $\hat{S}_i = S_i$ for all i, then full rent extraction is possible for every social choice function.*

However, when $\hat{S}_i \subsetneq S_i$, each type of buyer has the possibility of becoming perfectly informed about any type of her opponent, making full rent extraction impossible. In particular, fix $(v_{-i}, s_{-i}) \in \Theta_{-i}$ and consider a deviation by player i to choose a_i such that $a_i(v_{-i}) = 1_{(\tilde{s}_i, s_{-i})}$ with $\tilde{s}_i \notin \hat{S}_i$ and $a_i(v_{-i}') = 1_{(\hat{s}_i, s_{-i})}$ with $\hat{s}_i \in \hat{S}_i$ for each $v_{-i}' \neq v_{-i}$. Then, whenever $s_i = \tilde{s}_i$, player i is certain that her opponent's type is (v_{-i}, s_{-i}). Thus $C(\alpha_{-i}) = \Delta\Theta_{-i}$ and in the Appendix, we show that the smallest exposed face of $\Delta\Theta_{-i}$ containing each $\gamma(\theta_i, \alpha)$ contains $1_{\theta_{-i}}$ for each $\theta_{-i} \in \text{supp}(\gamma(\theta_i, \alpha))$. Since, for any α_{-i}, all such $1_{\theta_{-i}}$ belongs to $D(\theta_i, \alpha_{-i}, \succeq, \Lambda_i^f(\theta_i, \alpha))$ when $v_i \neq v^K$ and $\theta_{-i} \in \Lambda_i^f(\theta_i, \alpha)$, by Proposition 1, full rent extraction is impossible if there exists (v_i, v_{-i}) such that $v_i \neq v^K$ and $f(v_i, v_{-i}) > 0$.

Corollary 3 *For a linear information structure, if $\hat{S}_i \subsetneq S_i$ for some i, then full rent extraction is not possible for any social choice function where $f_i(v^k, v_{-i}) > 0$ for some $k < K$ and $v_{-i} \in V_{-i}$.*

In general, whether full rent extraction is possible or not depends on the specific application; for example, we do not take a stand on whether $\hat{S}_i = S_i$ or $\hat{S}_i \subsetneq S_i$ is the appropriate assumption. However, in the next section, we study the extent to which full rent extraction is robust to perturbations of the specific information acquisition technology under consideration. In particular, we will prove a result that implies that for a linear information structure $(\Theta, \pi, (A, \sigma))$, in the case $\hat{S}_i = S_i$, there exists $\hat{\sigma}$ arbitrarily close to σ such that full rent extraction is impossible.

7 Generic Impossibility of Full Rent Extraction

In this section, we will argue that the set of information structures where the seller can fully extract rent is small in a topological sense. We make the following assumptions:

A1. For each $i \in I$, A_i is a (non-singleton) compact and convex subset of a Euclidean space.

A2. $\sigma : A \times V \to \Delta S$ is continuous.

A3. For all $(a, v) \in A \times V$, $\mathrm{supp}(\sigma(a, v)) = S$.

A4. There exists $i \in I$ with $|S_i| > 2$ such that for some $v_i \notin V_i^*(\succeq)$, $f_i(v_i, v_{-i}) > 0$ for some $v_{-i} \in V_{-i}$.

Fix Θ, π, A, $\succeq$ and f satisfying **A1** and **A4**. Let $\mathcal{T}$ be the set of all information acquisition technologies satisfying **A2** and **A3**, i.e. $\mathcal{T}$ is the set of all continuous $\sigma : A \times V \to \Delta S$ such that $\mathrm{supp}(\sigma(a, v)) = S$ for all $(a, v) \in A \times V$. Equip $\mathcal{T}$ with the topology defined by the metric $d(f, g) = \sup_{a,v} \|f(a, v) - g(a, v)\|$, where $\| \cdot \|$ is the usual Euclidean norm. Let $\mathcal{T}^*$ be the subset of $\mathcal{T}$ such that $\sigma \in \mathcal{T}^*$ if and only if $(\Theta, \pi, (A, \sigma))$ guarantees full rent extraction for the ordering $\succeq$ and social choice function f.

Proposition 2 *$\overline{\mathcal{T}^*}$ has empty interior (and hence $\mathcal{T}^*$ is nowhere dense).*

Proposition 2 is equivalent to the statement that under **A1**–**A4**, the set of information structures where the seller cannot fully extract rent contains an open and dense subset of the set of all information structures. In this sense, full rent extraction fails "generically" in a model with information acquisition. Note that in a model where, for example, $|V_i| \geq 2$, $|S_i| > 2$, and $\succeq$ is a strict order on V_i for each i, **A4** must hold for any efficient social choice function; thus, Proposition 2 also implies the generic impossibility of full surplus extraction in that case.

The proof of Proposition 2 proceeds as follows. Let player 1 and v_1^1 be as in **A4**. The first step is to show that for $\sigma \in \overline{\mathcal{T}^*}$, if $C(\alpha_{-1})$ is full dimensional for all α_{-1}, then there exists an information acquisition strategy α such that for all θ_1 with $v_1 = v_1^1$, $\gamma(\theta_1, \alpha) \notin \mathrm{ri}(C(\alpha_{-1}))$. Then for any σ and for any $\varepsilon > 0$, we construct a $\hat{\sigma}$ such that $d(\hat{\sigma}, \sigma) \leq \varepsilon$ and $\hat{\sigma} \in \mathcal{T} \setminus \overline{\mathcal{T}^*}$ by defining $\hat{\sigma}$ so that for every α, $\hat{C}(\alpha_{-1})$ is full dimensional and there exists s_1 such that $\hat{\gamma}((v_1^1, s_1), \alpha) \in \mathrm{ri}(\hat{C}(\alpha_{-1}))$.[22] Intuitively, a small perturbation of σ is enough to ensure that the beliefs following different signals do not simultaneously land on the boundary of $C(\alpha_{-1})$ as α varies. Thus, the closure of the set of information structures that allow full rent extraction has empty interior.

We now discuss the necessity of assumption **A4**. If $|S_i| = 1$ for all $i \in I$, then signals are uninformative and we are in the standard model, where full rent extraction is generically possible. If there is no $i \in I$ with $v_i \notin V_i^*(\succeq)$ and $f_i(v_i, v_{-i}) > 0$ for some v_{-i}, then each buyer i gets the good only when her valuation $\succeq$-greatest in V_i, i.e. when it belongs to $V_i^*(\succeq)$; thus, the seller can fully extract rent by charging types such that $v_i \in V_i^*(\succeq)$ exactly their value for the object whenever they receive it.[23] When $|S_i| > 2$ for some $i \in I$ such that $v_i \notin V_i^*(\succeq)$ and $f_i(v_i, v_{-i}) > 0$ for some v_{-i}, Proposition 2 implies that full rent extraction is generically impossible; we conjecture that Proposition 2 remains true if we relax the requirement on the cardinality of the signals to $|S_i| > 1$.[24]

Finally, consider the linear information structure $(\Theta, \pi, (A, \sigma))$ from Section 6 which satisfies **A1** and **A2**. If the assumption of Corollary 2 holds, then full rent extraction is possible for every social choice function. We will now argue that for any f

[22] Where $\hat{\gamma}(\theta_1, \alpha)$ and $\hat{C}(\alpha_{-1})$ denote the beliefs and the convex hull of the beliefs under $\hat{\sigma}$.

[23] More explicitly, seller can fully extract rent by letting $x_i(\theta) = f_i(v)$ and $t_i(\theta) = w(v_i) f_i(v)$ for each $\theta = (v, s)$. In this case, for type θ_i, reporting $\hat{\theta}_i$ such that $\hat{v}_i \notin V_i^*(\succeq)$ results in zero payoff and reporting θ_i^* such that $v_i^* \in V_i^*(\succeq)$ results in zero payoff if $v_i \in V_i^*(\succeq)$ and negative payoff if $v_i \notin V_i^*(\succeq)$.

[24] In our construction of $\hat{\sigma}$, we use the player i satisfying **A4** and require that for all α, either type (v_i, s_i^1)'s belief or type (v_i, s_i^2)'s belief is in the interior of $\hat{C}(\alpha_{-i})$ for $s_i^1, s_i^2 \in S_i$ and $v_i \notin V_i^*(\succeq)$. The assumption that $|S_i| > 2$ makes such construction easy because, for each s_{-i}, the probabilities assigned to (s_i^1, s_{-i}) and (s_i^2, s_{-i}) can be varied independently.

such that **A4** is satisfied, Proposition 2 implies the existence of a nearby information acquisition technology such that full rent extraction is impossible. For any $\varepsilon > 0$, define σ^ε such that for each (a, v, s), $\sigma^\varepsilon(a, v)[s] = \left(1 - \frac{\varepsilon}{2\sqrt{|S|}}\right)\sigma(a, v)[s] + \frac{\varepsilon}{2\sqrt{|S|}}\frac{1}{|S|}$ and note that σ^ε satisfies **A2** and **A3**, i.e. $\sigma^\varepsilon \in \mathcal{T}$, and that $d(\sigma, \sigma^\varepsilon) \leq \frac{\varepsilon}{2}$. Then Proposition 2 implies that there exists $\hat{\sigma} \in \mathcal{T} \setminus \overline{\mathcal{T}^*}$ such that $d(\hat{\sigma}, \sigma^\varepsilon) \leq \frac{\varepsilon}{2}$, and hence $d(\hat{\sigma}, \sigma) \leq \varepsilon$. Thus, there exists an information acquisition technology arbitrarily close to σ such that full rent extraction is impossible.

References

Aliprantis, C., and K. Border (2006): *Infinite Dimensional Analysis*. Springer, Berlin, 3rd edn.

Bergemann, D., and J. Välimaäki (2002): "Information Acquisition and Efficient Mechanism Design," *Econometrica*, 70.

Bikhchandani, S. (2011): "Information acquisition and full surplus extraction," *Journal of Economic Theory*, 145(6), 2282–2308.

Bikhchandani, S., and I. Obara (2017): "Mechanism design with information acquisition," *Economic Theory*, 63(3), 783–812.

Carmona, G., and K. Laohakunakorn (2023): "The folk theorem for the prisoner's dilemma with endogenous private monitoring," *Journal of Economic Theory*, 213, 105731.

Crémer, J., and R. P. McLean (1985): "Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent," *Econometrica*, 53(2), 1247–1257.

——— (1988): "Full extraction of the surplus in Bayesian and dominant strategy auctions," *Econometrica*, 56(6), 1247–1257.

Denti, T. (2023): "Unrestricted Information Acquisition," *Theoretical Economics*, 18.

Heifetz, A., and Z. Neeman (2006): "On the generic (im)possibility of full surplus extraction in mechanism design," *Econometrica*, 74(1), 213–233.

Laffont, J.-J., and D. Martimort (2000): "Mechanism design with collusion and correlation," *Econometrica*, 68(2), 309–342.

Lopomo, G., L. Rigotti, and C. Shannon (2014): "Uncertainty in mechanism design," Working Paper.

McAfee, R. P., J. McMillan, and P. J. Reny (1989): "Extracting the surplus in the common-value auction," *Econometrica*, 57(6), 1451–1459.

McAfee, R. P., and P. J. Reny (1992): "Correlated information and mecanism design," *Econometrica*, 60(2), 395–421.

Myerson, R. B. (1981): "Optimal auction design," *Mathematics of operations research*, 6(1), 58–73.

Neeman, Z. (2004): "The relevance of private information in mechanism design," *Journal of Economic Theory*, 117(1), 55–77.

Obara, I. (2008): "The full surplus extraction theorem with hidden actions," *B.E. Journal of Theoretical Economics*, 8(1), 1935–1704.

Pernoud, A., and S. Gleyze (2023): "Informationally simple incentives," *Journal of Political Economy*, 131.

Renou, L. (2015): "Rent extraction and uncertainty," Working paper.

Robert, J. (1991): "Continuity in auction design," *Journal of Economic Theory*, 55(1), 169–179.

Yamashita, T. (2018): "Revenue guarantees in auctions with a (correlated) common prior and additional information," Working paper.

A Appendix

A.1 Revelation Principle

Let $(\alpha, \hat{x}, \hat{t})$ be an arbitrary mechanism, where α is an information acquisition strategy and each i can send messages from some arbitrary set M_i, so that $\hat{x} : M \to \Delta I$ and $\hat{t} : M \to \mathbb{R}^n$. For each i, let $\mu_i : V_i \times S_i \times A_i \to M_i$ be a general reporting strategy (which may depend on a_i). Following information acquisition strategy α, for $\theta_i \in \operatorname{supp}(\gamma_{\Theta_i}(\alpha))$, let $\hat{U}_i(m_i', \theta_i; \alpha, \mu_{-i})$ be i's payoff when player i reports m_i', player i is type θ_i, and the other players are following the general reporting strategy μ_{-i}:

$$\begin{aligned}\hat{U}_i(m_i', \theta_i; \alpha, \mu_{-i}) = \sum_{\theta_{-i}} \gamma(\theta_i, \alpha)[\theta_{-i}] \Big(&\hat{x}_i(m_i', \mu_{-i}(\theta_{-i}, \alpha_{-i}(v_{-i}))) w(v_i) \\ &- \hat{t}_i(m_i', \mu_{-i}(\theta_{-i}, \alpha_{-i}(v_{-i}))) \Big).\end{aligned}$$

Proposition A.1 *Suppose that there exists a mechanism $(\alpha, \hat{x}, \hat{t})$, and for each player i, a general reporting strategy μ_i such that for all $\alpha_i' \in A_i^{|V_i|}$:*

1. $\mu_i(\theta_i, \alpha_i'(v_i)) \in \arg\max_{m_i'} \hat{U}_i(m_i', \theta_i; \alpha_i', \alpha_{-i}, \mu_{-i})$ *for all* $\theta_i \in \operatorname{supp}(\gamma_{\Theta_i}(\alpha_i', \alpha_{-i}))$,
2. $\hat{U}_i(\mu_i(\theta_i, \alpha_i(v_i)), \theta_i; \alpha, \mu_{-i}) \geq 0$ *for all* $\theta_i \in \operatorname{supp}(\gamma_{\Theta_i}(\alpha))$,
3. $\sum_{\theta_i \in \operatorname{supp}(\gamma_{\Theta_i}(\alpha))} \gamma_{\Theta_i}(\alpha)[\theta_i] \hat{U}_i(\mu_i(\theta_i, \alpha_i(v_i)), \theta_i; \alpha, \mu_{-i})$
$\geq \sum_{\theta_i \in \operatorname{supp}(\gamma_{\Theta_i}(\alpha_i', \alpha_{-i}))} \gamma_{\Theta_i}(\alpha_i', \alpha_{-i})[\theta_i] \max\{\hat{U}_i(\mu_i(\theta_i, \alpha_i'(v_i)), \theta_i; \alpha_i', \alpha_{-i}, \mu_{-i}), 0\}$.

Then there exists an incentive compatible direct mechanism (α, x, t) such that buyers optimally acquire information and $(x(\theta), t(\theta)) = (\hat{x}(\mu(\theta, \alpha(v))), \hat{t}(\mu(\theta, \alpha(v))))$ for all $\theta \in \Theta$.

Condition 1 requires that the reporting strategy is optimal after all information acquisition strategies, Condition 2 is individual rationality, and Condition 3 requires that the recommended information acquisition strategy is optimal.

Proof. Define $U_i^{**}(\theta_i; \alpha, \mu_{-i}) = \sup_{m_i'} \hat{U}_i(m_i', \theta_i; \alpha, \mu_{-i})$ and assume that Conditions 1 and 2 and 3 are satisfied for $(\alpha, \hat{x}, \hat{t})$ and μ. Now define the direct mechanism

(α, x, t) as $x(\theta) = \hat{x}(\mu(\theta, \alpha(v)))$ and $t(\theta) = \hat{t}(\mu(\theta, \alpha(v)))$. First, note that (α, x, t) is incentive compatible since:

$$\begin{aligned}
U_i(\theta_i', \theta_i; \alpha) &= \sum_{\theta_{-i}} \gamma(\theta_i, \alpha)[\theta_{-i}] \left(x_i(\theta_i', \theta_{-i})w(v_i) - t_i(\theta_i', \theta_{-i})\right) \\
&= \sum_{\theta_{-i}} \gamma(\theta_i, \alpha)[\theta_{-i}] \Big(\hat{x}_i(\mu_i(\theta_i', \alpha_i(v_i'))), \mu_{-i}(\theta_{-i}, \alpha_{-i}(v_{-i})))w(v_i) \\
&\qquad\qquad - \hat{t}_i(\mu_i(\theta_i', \alpha_i(v_i')), \mu_{-i}(\theta_{-i}, \alpha_{-i}(v_{-i})))\Big) \\
&= \hat{U}_i(\mu_i(\theta_i', \alpha_i(v_i')), \theta_i; \alpha, \mu_{-i}) \leq \hat{U}_i(\mu_i(\theta_i, \alpha_i(v_i)), \theta_i; \alpha, \mu_{-i}) \\
&= \sum_{\theta_{-i}} \gamma(\theta_i, \alpha)[\theta_{-i}] \Big(\hat{x}_i(\mu_i(\theta_i, \alpha_i(v_i)), \mu_{-i}(\theta_{-i}, \alpha_{-i}(v_{-i})))w(v_i) \\
&\qquad\qquad - \hat{t}_i(\mu_i(\theta_i, \alpha_i(v_i)), \mu_{-i}(\theta_{-i}, \alpha_{-i}(v_{-i})))\Big) \\
&= \sum_{\theta_{-i}} \gamma(\theta_i, \alpha)[\theta_{-i}] \left(x_i(\theta_i, \theta_{-i})w(v_i) - t_i(\theta_i, \theta_{-i})\right) = U_i(\theta_i, \theta_i; \alpha).
\end{aligned}$$

To see that buyers optimally acquire information given (α, x, t), note that for $\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha_i', \alpha_{-i}))$:

$$\begin{aligned}
U_i^*(\theta_i; \alpha_i', \alpha_{-i}) &= \max_{\theta_i'} U_i(\theta_i', \theta_i; \alpha_i', \alpha_{-i}) \\
&\leq \max_{m_i'} \hat{U}_i(m_i', \theta_i; \alpha_i', \alpha_{-i}, \mu_{-i}) = U_i^{**}(\theta_i; \alpha_i', \alpha_{-i}, \mu_{-i})
\end{aligned}$$

where the inequality follows because for all θ_i', $U_i(\theta_i', \theta_i; \alpha_i', \alpha_{-i}) = \hat{U}_i(\mu_i(\theta_i', \alpha_i(v_i')), \theta_i; \alpha_i', \alpha_{-i}, \mu_{-i})$. Also note that for $\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha))$:

$$\begin{aligned}
U_i^*(\theta_i; \alpha) &= U_i(\theta_i, \theta_i; \alpha) \\
&= \hat{U}_i(\mu_i(\theta_i, \alpha_i(v_i)), \theta_i; \alpha, \mu_{-i}) = U_i^{**}(\theta_i; \alpha, \mu_{-i}).
\end{aligned}$$

Thus:

$$
\begin{aligned}
\tilde{u}_i(\alpha_i, \alpha_{-i}) &= \sum_{\theta_i \in \operatorname{supp}(\gamma_{\Theta_i}(\alpha))} \gamma_{\Theta_i}(\alpha_i, \alpha_{-i})[\theta_i] \max\{U_i^*(\theta_i; \alpha_i, \alpha_{-i}), 0\} \\
&= \sum_{\theta_i \in \operatorname{supp}(\gamma_{\Theta_i}(\alpha))} \gamma_{\Theta_i}(\alpha_i, \alpha_{-i})[\theta_i] \max\{U_i^{**}(\theta_i; \alpha, \mu_{-i}), 0\} \\
&\geq \sum_{\theta_i \in \operatorname{supp}(\gamma_{\Theta_i}(\alpha))} \gamma_{\Theta_i}(\alpha_i, \alpha_{-i})[\theta_i] U_i^{**}(\theta_i; \alpha, \mu_{-i}) \\
&\geq \sum_{\theta_i \in \operatorname{supp}(\gamma_{\Theta_i}(\alpha_i', \alpha_{-i}))} \gamma_{\Theta_i}(\alpha_i', \alpha_{-i})[\theta_i] \max\{U_i^{**}(\theta_i; \alpha_i', \alpha_{-i}, \mu_{-i}), 0\} \\
&\geq \sum_{\theta_i \in \operatorname{supp}(\gamma_{\Theta_i}(\alpha_i', \alpha_{-i}))} \gamma_{\Theta_i}(\alpha_i', \alpha_{-i})[\theta_i] \max\{U_i^*(\theta_i; \alpha_i', \alpha_{-i}), 0\} = \tilde{u}_i(\alpha_i', \alpha_{-i}).
\end{aligned}
$$

■

A.2 Proofs

Proof of Proposition 1. For sufficiency, fix $f : V \to \Delta I$ and $\succeq$ let $w \in \mathcal{R}(\succeq)$. Then let information acquisition strategy α be such that:

1. For all $\theta_i \in \Theta_i^f(\alpha) \setminus (V_i^*(\succeq) \times S_i)$, $\gamma(\theta_i, \alpha) \notin \operatorname{ri}(C(\alpha_{-i}))$,
2. For all $\theta_i \in \Theta_i^f(\alpha) \setminus (V_i^*(\succeq) \times S_i)$, $F_{\overline{C}(\alpha_{-i})}(\gamma(\theta_i, \alpha)) \cap \overline{D}(\theta_i, \alpha_{-i}, \succeq, \Lambda_i^f(\theta_i, \alpha)) = \varnothing$.

Let $x(\theta) = f(v)$ for any θ that arises with positive probability given α, and let $x(\theta) = 0$ otherwise. Note that $x_i(\theta_i, \theta_{-i}) > 0$ if and only if $\theta_{-i} \in \Lambda_i^f(\theta_i, \alpha)$. Now we argue that there exists transfers such that (α, x, t) is incentive compatible, fully extracts rent, and buyers optimally acquire information.

For each $\theta_i \in \Theta_i^f(\alpha) \setminus (V_i^*(\succeq) \times S_i)$, Condition 1 implies that there is a hyperplane:

$$\{\gamma : \tau \cdot \gamma = \tau \cdot \gamma(\theta_i, \alpha)\}$$

which properly supports $\overline{C}(\alpha_{-i})$ at $\gamma(\theta_i, \alpha)$. Thus, $F_{\overline{C}(\alpha_{-i})}(\gamma(\theta_i, \alpha))$ is a proper exposed face of $\overline{C}(\alpha_{-i})$. Let $\tau(\theta_i)$ be a lottery such that $F_{\overline{C}(\alpha_{-i})}(\gamma(\theta_i, \alpha)) = \{\gamma : \tau(\theta_i) \cdot \gamma =$

$0\} \cap \overline{C}(\alpha_{-i})$. This implies that $\tau(\theta_i) \cdot \gamma = 0$ for all $\gamma \in F_{\overline{C}(\alpha_{-i})}(\gamma(\theta_i, \alpha))$. Without loss of generality, let $\tau(\theta_i) \cdot \gamma > 0$ for all $\gamma \in \overline{C}(\alpha_{-i}) \setminus F_{\overline{C}(\alpha_{-i})}(\gamma(\theta_i, \alpha))$.

For $\theta_i \in \Theta_i^f(\alpha) \setminus (V_i^*(\succeq) \times S_i)$ let $t_i(\theta_i) = w(v_i)x_i(\theta_i) + K\tau(\theta_i)$,[25] where K is sufficiently large. Note that $\gamma(\theta_i, \alpha) \cdot \tau(\theta_i) = 0$. For $\theta_i \notin \Theta_i^f(\alpha)$, let $t_i(\theta_i) = 0$. For $\theta_i \in (V_i^*(\succeq) \times S_i) \cap \Theta_i^f(\alpha)$, let $t_i(\theta_i) = w(v_i)x_i(\theta_i)$. By construction, (α, x, t) fully extracts rent. We now show that (α, x, t) is incentive compatible and buyers optimally acquire information.

First we show that for any $\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha))$, $\theta_i' \in \Theta_i$, $U_i(\theta_i', \theta_i; \alpha) \leq U_i(\theta_i, \theta_i; \alpha)$. We will consider the cases where $\theta_i' \in \Theta_i^f(\alpha) \setminus (V_i^*(\succeq) \times S_i)$, $\theta_i' \notin \Theta_i^f(\alpha)$, and $\theta_i' \in (V_i^*(\succeq) \times S_i) \cap \Theta_i^f(\alpha)$ separately. Note that for $\theta_i' \in \Theta_i^f(\alpha) \setminus (V_i^*(\succeq) \times S_i)$:

$$\begin{aligned} U_i(\theta_i', \theta_i; \alpha) &= w(v_i)\gamma(\theta_i, \alpha) \cdot x_i(\theta_i') - \gamma(\theta_i, \alpha) \cdot (w(v_i')x_i(\theta_i') + K\tau(\theta_i')) \\ &= (w(v_i) - w(v_i'))\gamma(\theta_i, \alpha) \cdot x_i(\theta_i') - K\gamma(\theta_i, \alpha) \cdot \tau(\theta_i') \leq 0. \end{aligned}$$

The last inequality follows because $\gamma(\theta_i, \alpha) \in C(\alpha_i)$, so $\gamma(\theta_i, \alpha) \cdot \tau(\theta_i') \geq 0$, and when $\gamma(\theta_i, \alpha) \cdot \tau(\theta_i') = 0$, either $w(v_i) \leq w(v_i')$ or $\gamma(\theta_i, \alpha) \cdot x_i(\theta_i') = 0$ by Condition 2. To see this, first note that if $\gamma(\theta_i, \alpha) \cdot \tau(\theta_i') = 0$, then $\gamma(\theta_i, \alpha) \in F_{\overline{C}(\alpha_{-i})}(\gamma(\theta_i', \alpha))$, since $F_{\overline{C}(\alpha_{-i})}(\gamma(\theta_i', \alpha)) = \overline{C}(\alpha_{-i}) \cap \{\gamma : \tau(\theta_i') \cdot \gamma = 0\}$. But then Condition 2 requires that $\gamma(\theta_i, \alpha) \notin D(\theta_i', \alpha_{-i}, \succeq, \Lambda_i^f(\theta_i', \alpha))$. Since $\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha))$, this implies that θ_i must be such that $w(v_i) \leq w(v_i')$ or $\gamma(\theta_i, \alpha) \cdot x_i(\theta_i') = 0$.

For $\theta_i' \notin \Theta_i^f(\alpha)$, $U_i(\theta_i', \theta_i; \alpha) = 0$ since $x_i(\theta_i') = t_i(\theta_i') = 0$. For $\theta_i' \in V_i^*(\succeq) \times S_i$, note that $w(v_i') \geq w(v_i)$. Then:

$$\begin{aligned} U_i(\theta_i', \theta_i; \alpha) &= w(v_i)\gamma(\theta_i, \alpha) \cdot x_i(\theta_i') - \gamma(\theta_i, \alpha) \cdot (w(v_i')x_i(\theta_i')) \\ &= (w(v_i) - w(v_i'))\gamma(\theta_i, \alpha) \cdot x_i(\theta_i') \leq 0. \end{aligned}$$

Now we show that buyers optimally acquire information. That is, for any $\alpha_i' \in A_i^{|V_i|}$:

$$\sum_{\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha_i', \alpha_{-i}))} \gamma_{\Theta_i}(\alpha_i', \alpha_{-i})[\theta_i] \max\{U_i^*(\theta_i; \alpha_i', \alpha_{-i}), 0\} \leq 0.$$

[25] Recall that $t_i(\theta_i) = (t_i(\theta_i, \theta_{-i}^1), \ldots, t_i(\theta_i, \theta_{-i}^{|\Theta_{-i}|}))$ and $x_i(\theta_i) = (x_i(\theta_i, \theta_{-i}^1), \ldots, x_i(\theta_i, \theta_{-i}^{|\Theta_{-i}|}))$.

Note that for all $\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha'_i, \alpha_{-i}))$:

$$\begin{aligned}
U_i^*(\theta_i; \alpha'_i, \alpha_{-i}) &= \max_{\theta'_i} U_i(\theta'_i, \theta_i; \alpha'_i, \alpha_{-i}) \\
&= \max_{\theta'_i} w(v_i)x(\theta'_i) \cdot \gamma(\theta_i, \alpha'_i, \alpha_{-i}) - t(\theta'_i) \cdot \gamma(\theta_i, \alpha'_i, \alpha_{-i}) \\
&\leq \max \Bigg\{ \max_{\theta'_i \in \Theta_i^f(\alpha) \backslash (V_i^*(\succeq) \times S_i)} w(v_i)x(\theta'_i) \cdot \gamma(\theta_i, \alpha'_i, \alpha_{-i}) \\
&\qquad\qquad - (w(v'_i)x_i(\theta'_i) + K\tau(\theta'_i)) \cdot \gamma(\theta_i, \alpha'_i, \alpha_{-i}), 0 \Bigg\} \\
&= \max \Bigg\{ \max_{\theta'_i \in \Theta_i^f(\alpha) \backslash (V_i^*(\succeq) \times S_i)} (w(v_i) - w(v'_i))x(\theta'_i) \cdot \gamma(\theta_i, \alpha'_i, \alpha_{-i}) \\
&\qquad\qquad - K\tau(\theta'_i) \cdot \gamma(\theta_i, \alpha'_i, \alpha_{-i}), 0 \Bigg\} \\
&= 0.
\end{aligned}$$

The last equality holds because for any $\alpha'_i \in A_i^{|V_i|}$, $\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha'_i, \alpha_{-i}))$, and $\theta'_i \in \Theta_i^f(a) \backslash (V_i^*(\succeq) \times S_i)$, if $\gamma(\theta_i, \alpha'_i, \alpha_{-i}) \cdot \tau(\theta'_i) = 0$, then $w(v_i) \leq w(v'_i)$ or $\gamma(\theta_i, \alpha'_i, \alpha_{-i}) \cdot x_i(\theta'_i) = 0$, and if $w(v_i) > w(v'_i)$ and $\gamma(\theta_i, \alpha'_i, \alpha_{-i}) \cdot x_i(\theta'_i) > 0$, then $\gamma(\theta_i, \alpha'_i, \alpha_{-i}) \cdot \tau(\theta'_i) > \varepsilon$ for some $\varepsilon > 0$ (i.e. ε is a uniform lower bound across all α'_i).[26] The former was shown when we established incentive compatibility. To see the latter, suppose that for every $\varepsilon > 0$, there exists $\alpha'_i \in A_i^{|V_i|}$, $\theta_i \in \text{supp}(\gamma_{\Theta_i}(\alpha'_i, \alpha_{-i}))$ where $w(v_i) > w(v'_i)$ and $\gamma(\theta_i, \alpha'_i, \alpha_{-i}) \cdot x_i(\theta'_i) > 0$ such that $\gamma(\theta_i, \alpha'_i, \alpha_{-i}) \cdot \tau(\theta'_i) < \varepsilon$. Let $\gamma^\varepsilon \in \overline{D}(\theta'_i, \alpha_{-i}, \succeq, \Lambda_i^f(\theta'_i, \alpha))$ denote the sequence of such beliefs. Since $\overline{D}(\theta'_i, \alpha_{-i}, \succeq, \Lambda_i^f(\theta'_i, \alpha))$ is compact, as $\varepsilon \to 0$, there is a convergent subsequence which converges to $\gamma \in \overline{D}(\theta'_i, \alpha_{-i}, \succeq, \Lambda_i^f(\theta'_i, \alpha))$, with $\gamma \cdot \tau(\theta'_i) = 0$. Note that $\gamma^\varepsilon \in \overline{C}(\alpha_{-i})$, so $\gamma \in \overline{C}(\alpha_{-i})$, and $\gamma \in \{\gamma : \gamma \cdot \tau(\theta'_i) = 0\}$. Since $F_{\overline{C}(\alpha_{-i})}(\gamma(\theta'_i, \alpha)) = \{\gamma : \gamma \cdot \tau(\theta'_i) = 0\} \cap \overline{C}(\alpha_{-i})$, $\gamma \in F_{\overline{C}(\alpha_{-i})}(\gamma(\theta'_i, \alpha))$ contradicting Condition 2. Thus, for sufficiently large K, the equality holds.

Now we prove necessity. Suppose that Condition 1 is not satisfied. Consider $w \in$

[26]When we established incentive compatibility, we did not need this step because the set of θ_i is finite. However, since there can be infinitely many α'_i, it is not enough that $\gamma(\theta_i, \alpha'_i, \alpha_{-i}) \cdot \tau(\theta'_i) > 0$ for all α'_i, because now there is the possibility that $\gamma(\theta_i, \alpha'_i, \alpha_{-i}) \cdot \tau(\theta'_i)$ can be made arbitrarily close to 0.

$\mathcal{W}(\succeq)$ such that for all i and for $v_i \in V_i^*(\succeq)$, $w(v_i) = L$, where L is sufficiently large. Take an arbitrary α and let $\hat{\theta}_i \in \Theta_i^f(\alpha) \setminus (V_i^*(\succeq) \times S_i)$ be such that $\gamma(\hat{\theta}_i, \alpha) \in \mathrm{ri}(C(\alpha_{-i}))$. Suppose for a contradiction that (α, x, t) is incentive compatible, fully extracts rent, buyers optimally acquire information, and $x_i(\theta_i, \theta_{-i}) = f_i(v_i, v_{-i})$ for all $\theta \in \mathrm{supp}(\gamma(\alpha))$. Without loss of generality, let $t_i(\hat{\theta}_i) = w(\hat{v}_i)x_i(\hat{\theta}_i) + K\tau(\hat{\theta}_i)$ for some $\tau(\hat{\theta}_i)$ with $\|\tau(\hat{\theta}_i)\| \leq 1$ and $K > 0$. For full rent extraction, we need $\tau(\hat{\theta}_i) \cdot \gamma(\hat{\theta}_i, \alpha) = 0$. Note that there must exist $\theta_i^* \in V_i^*(\succeq) \times S_i$ such that $\gamma(\theta_i^*, \alpha) \cdot x_i(\hat{\theta}_i) > 0$, otherwise Condition 1 must hold; to see this, suppose that $\gamma(\theta_i, \alpha) \cdot x_i(\hat{\theta}_i) = 0$ for all θ_i such that $v_i \in V_i^*(\succeq)$. Then any $\tau(\hat{\theta}_i)$ that is strictly positive for and only for components corresponding to θ'_{-i} with $f_i(\hat{v}_i, v'_{-i}) > 0$ and $\gamma(\hat{\theta}_i, \alpha)[\theta'_{-i}] = 0$ properly supports $C(\alpha_{-i})$ at $\gamma(\hat{\theta}_i, \alpha)$,[27] which implies that $\gamma(\hat{\theta}_i, \alpha) \notin \mathrm{ri}(C(\alpha_{-i}))$, a contradiction. Incentive compatibility then requires $\gamma(\theta_i^*, \alpha) \cdot ((L - w(\hat{v}_i))x_i(\hat{\theta}_i) - K\tau(\hat{\theta}_i)) \leq 0$, which implies:

$$\gamma(\theta_i^*, \alpha) \cdot K\tau(\hat{\theta}_i) \geq \gamma(\theta_i^*, \alpha) \cdot (L - w(\hat{v}_i))x_i(\hat{\theta}_i) > 0.$$

Since Condition 1 is not satisfied, there must exist $\gamma \in C(\alpha_{-i})$ such that $\gamma \cdot \tau(\hat{\theta}_i) < 0$. Thus, there must exists $\alpha'_i \in A_i^{|V_i|}$, $\theta'_i \in \mathrm{supp}(\gamma_{\Theta_i}(\alpha'_i, \alpha_{-i}))$ such that $\gamma(\theta'_i, \alpha'_i, \alpha_{-i}) \cdot \tau(\hat{\theta}_i) < 0$. The previous displayed inequality implies that for large L, K must also be large. Then:

$$\begin{aligned}
&\sum_{\theta_i \in \mathrm{supp}(\gamma_{\Theta_i}(\alpha'_i, \alpha_{-i}))} \gamma_{\Theta_i}(\alpha'_i, \alpha_{-i})[\theta_i] \max\{U_i^*(\theta_i; \alpha'_i, \alpha_{-i}), 0\} \\
&\qquad \geq \gamma_{\Theta_i}(\alpha'_i, \alpha_{-i})[\theta'_i] U_i^*(\theta'_i; \alpha'_i, \alpha_{-i}) \\
&\qquad \geq \gamma_{\Theta_i}(\alpha'_i, \alpha_{-i})[\theta'_i] U_i(\hat{\theta}_i, \theta'_i; \alpha'_i, \alpha_{-i}) \\
&\qquad = \gamma_{\Theta_i}(\alpha'_i, \alpha_{-i})[\theta'_i] \left(w(v'_i)x_i(\hat{\theta}_i) - t(\hat{\theta}_i)\right) \cdot \gamma(\theta'_i, \alpha'_i, \alpha_{-i}) \\
&\qquad = \gamma_{\Theta_i}(\alpha'_i, \alpha_{-i})[\theta'_i] \left((w(v'_i) - w(\hat{v}_i))x_i(\hat{\theta}_i) - K\tau(\hat{\theta}_i)\right) \cdot \gamma(\theta'_i, \alpha'_i, \alpha_{-i}) > 0.
\end{aligned}$$

[27] Since $\hat{\theta}_i \in \Theta_i^f(\alpha)$, there exists v'_{-i} such that $f_i(\hat{v}_i, v'_{-i}) > 0$, and for some $\theta_i \in \mathrm{supp}(\gamma_{\Theta_i}(\alpha))$ such that $v_i \in V_i^*(\succeq)$, we must have $\gamma(\theta_i, \alpha)[v'_{-i}, s'_{-i}] > 0$ for some s'_{-i}. Since $\gamma(\theta_i, \alpha) \cdot x_i(\hat{\theta}_i) = 0$, the requirement that the mechanism is consistent with f implies $\gamma(\hat{\theta}_i, \alpha)[v'_{-i}, s'_{-i}] = 0$.

Suppose that Condition 2 is not satisfied. Then define w as before, take an arbitrary α, and let $\hat{\theta}_i \in \Theta_i^f(\alpha) \setminus (V_i^*(\succeq) \times S_i)$ be such that $F_{\overline{C}(\alpha_{-i})}(\gamma(\hat{\theta}_i, \alpha)) \cap \overline{D}(\hat{\theta}_i, \alpha_{-i}, \succeq, \Lambda_i^f(\theta_i, \alpha)) \neq \varnothing$. Suppose for a contradiction that (α, x, t) is incentive compatible, fully extracts rent, buyers optimally acquire information, and $x_i(\theta_i, \theta_{-i}) = f_i(v_i, v_{-i})$ for all $\theta \in \mathrm{supp}(\gamma(\alpha))$. Without loss of generality, let $t_i(\hat{\theta}_i) = w(\hat{v}_i)x_i(\hat{\theta}_i) + K\tau(\hat{\theta}_i)$ for some $\tau(\hat{\theta}_i)$ with $\|\tau(\hat{\theta}_i)\| \leq 1$ and $K > 0$. For full rent extraction, we need $\tau(\hat{\theta}_i) \cdot \gamma(\hat{\theta}_i, \alpha) = 0$. Assume that $\tau(\hat{\theta}_i) \cdot \gamma(\theta_i', \alpha_i', \alpha_{-i}) \geq 0$ for all $\alpha_i' \in A_i^{|V_i|}$ and $\theta_i' \in \mathrm{supp}(\gamma_{\Theta_i}(\alpha_i', \alpha_{-i}))$, otherwise by the same argument as in the previous paragraph, the strategy α_i' must yield strictly positive rents. Thus, $\tau(\hat{\theta}_i)$ is a supporting hyperplane for $\overline{C}(\alpha_{-i})$, which implies that $F_{\overline{C}(\alpha_{-i})}(\gamma(\hat{\theta}_i, \alpha)) \subseteq \{\gamma : \gamma \cdot \tau(\hat{\theta}_i) = 0\}$.

Since Condition 2 fails, there exists $\gamma \in F_{\overline{C}(\alpha_{-i})}(\gamma(\hat{\theta}_i, \alpha)) \cap \overline{D}(\hat{\theta}_i, \alpha_{-i}, \succeq, \Lambda_i^f(\hat{\theta}_i, \alpha))$. We show that for every $\varepsilon > 0$, we can find $\alpha_i' \in A_i^{|V_i|}$ and $\theta_i' \in \mathrm{supp}(\gamma_{\Theta_i}(\alpha_i', \alpha_{-i}))$ such that $w(v_i') > w(\hat{v}_i)$, $\gamma(\theta_i', \alpha_i', \alpha_{-i}) \cdot x_i(\hat{\theta}_i) > 0$, and $\tau(\hat{\theta}_i) \cdot \gamma(\theta_i', \alpha_i', \alpha_{-i}) < \varepsilon$. To see this, take any sequence $\gamma^k \in D(\hat{\theta}_i, \alpha_{-i}, \succeq, \Lambda_i^f(\hat{\theta}_i, \alpha))$ converging to $\gamma \in \overline{D}(\hat{\theta}_i, \alpha_{-i}, \succeq, \Lambda_i^f(\hat{\theta}_i, \alpha))$. As $k \to \infty$, $\tau(\hat{\theta}_i) \cdot \gamma^k \to \tau(\hat{\theta}_i) \cdot \gamma = 0$. Since $\gamma^k \in D(\hat{\theta}_i, \alpha_{-i}, \succeq, \Lambda_i^f(\hat{\theta}_i, \alpha))$, for each k, there exists $\alpha_i' \in A_i^{|V_i|}$ and $\theta_i' \in \mathrm{supp}(\gamma_{\Theta_i}(\alpha_i', \alpha_{-i}))$ such that $\gamma^k = \gamma(\theta_i', \alpha_i', \alpha_{-i})$, $w(v_i') > w(\hat{v}_i)$, and $\gamma(\theta_i', \alpha_i', \alpha_{-i}) \cdot x_i(\hat{\theta}_i) > 0$. Take a sufficiently large k and we are done. Then:

$$\begin{aligned}
&\sum_{\theta_i \in \mathrm{supp}(\gamma_{\Theta_i}(\alpha_i', \alpha_{-i}))} \gamma_{\Theta_i}(\alpha_i', \alpha_{-i})[\theta_i] \max\{U_i^*(\theta_i; \alpha_i', \alpha_{-i}), 0\} \\
&\qquad \geq \gamma_{\Theta_i}(\alpha_i', \alpha_{-i})[\theta_i'] U_i^*(\theta_i'; \alpha_i', \alpha_{-i}) \\
&\qquad \geq \gamma_{\Theta_i}(\alpha_i', \alpha_{-i})[\theta_i'] U_i(\hat{\theta}_i, \theta_i'; \alpha_i', \alpha_{-i}) \\
&\qquad = \gamma_{\Theta_i}(\alpha_i', \alpha_{-i})[\theta_i'] \left(w(v_i') x_i(\hat{\theta}_i) - t(\hat{\theta}_i) \right) \cdot \gamma(\theta_i', \alpha_i', \alpha_{-i}) \\
&\qquad = \gamma_{\Theta_i}(\alpha_i', \alpha_{-i})[\theta_i'] \left((w(v_i') - w(\hat{v}_i)) x_i(\hat{\theta}_i) - K\tau(\hat{\theta}_i) \right) \cdot \gamma(\theta_i', \alpha_i', \alpha_{-i}) > 0.
\end{aligned}$$

■

Proof of Corollary 2. Note that for each $\theta_i \in \Theta_i$ such that $v_i = v^k$ and $s_i \in S^l$,

$$\gamma(v_i, s_i, \alpha)[\{v^l\} \times S^k] = 1.$$

Thus, $\gamma(v_i, s_i, \alpha) \notin \mathrm{ri}(C(\alpha_{-i}))$ since $\gamma(v_i, s_i, \alpha)$ is a support point of $C(\alpha_{-i})$ (e.g. by Aliprantis and Border (2006, Theorem 7.36)). That $\gamma(v_i, s_i, \alpha)$ is a support point of $C(\alpha_{-i})$ can be seen by considering the supporting hyperplane defined by $x \in \Delta\Theta_{-i}$ where $x[v_{-i}, s_{-i}] = 0$ for all $v_{-i}, s_{-i} \in \mathrm{supp}(\gamma(v_i, s_i, \alpha)) = \{v^l\} \times S^k$ and $x[v_{-i}, s_{-i}] > 0$ otherwise. Note that $\gamma(v_i, s_i, \alpha) \cdot x = 0$, $\gamma \cdot x \geq 0$ for all $\gamma \in C(\alpha_{-i})$, and for (v_i', s_i) such that $v_i' \neq v_i$, $\gamma(v_i', s_i, \alpha) \in C(\alpha_{-i})$, $\gamma(v_i', s_i, \alpha)[\{v^l\} \times (S_{-i} \setminus S^k)] > 0$ and hence $\gamma(v_i', s_i, \alpha) \cdot x > 0$; thus, the hyperplane defined by x properly supports $C(\alpha_{-i})$.

Now consider $F_{\overline{C}(\alpha_{-i})}(\gamma(v_i, s_i, \alpha))$, which is the smallest exposed face of $\overline{C}(\alpha_{-i})$ containing $\gamma(v_i, s_i, \alpha)$. This must be a subset of the hyperplane defined by $x \in \Delta\Theta_{-i}$ where $x[v_{-i}, s_{-i}] = 0$ if $(v_{-i}, s_{-i}) \in \{v^l\} \times S^k$ and $x[v_{-i}, s_{-i}] > 0$ otherwise. Note that, for any $v_i' \neq v^k$, s_i', and α_i', $\gamma(v_i', s_i', \alpha_i', \alpha_{-i})[V_{-i} \times (S_{-i} \setminus S^k)] > 0$; hence $\gamma(v_i', s_i', \alpha_i', \alpha_{-i}) \cdot x > 0$ and $\gamma(v_i', s_i', \alpha_i', \alpha_{-i})$ does not lie on the smallest exposed face containing $\gamma(v_i, s_i, \alpha)$, and hence $F_{\overline{C}(\alpha_{-i})}(\gamma(v_i, s_i, \alpha)) \cap \overline{D}(\theta_i, \alpha_{-i}, \succeq, \Theta_{-i}) = \varnothing$. ■

Proof of Corollary 3. We have $C(\alpha_{-i}) = \Delta\Theta_{-i}$ and, for each $M \subseteq \Theta_{-i}$, $\{1_{\theta_{-i}} : \theta_{-i} \in M\} \subseteq \overline{D}(\theta_i, \alpha_{-i}, \succeq, M)$ for each θ_i such that $v_i \neq v^K$. We show that for each θ_i, $1_{\theta_{-i}} \in F_{\Delta\Theta_{-i}}(\gamma(\theta_i, \alpha))$ for each $\theta_{-i} \in \mathrm{supp}(\gamma(\theta_i, \alpha))$. This follows because $\Delta\Theta_{-i}$ is a polytope with vertices $\{1_{\theta_{-i}} : \theta_{-i} \in \Theta_{-i}\}$ and $\gamma(\theta_i, \alpha) = \sum_{\theta_{-i}} \gamma(\theta_i, \alpha)[\theta_{-i}] 1_{\theta_{-i}}$. Thus, $co\{1_{\theta_{-i}} : \theta_{-i} \in \mathrm{supp}(\gamma(\theta_i, \alpha))\}$ is the smallest (exposed) face containing $\gamma(\theta_i, \alpha)$.[28] Thus, for θ_i such that $v_i \neq v^K$ and $f(v_i, v_{-i}) > 0$ for some v_{-i}, $F_{\Delta\Theta_{-i}}(\gamma(\theta_i, \alpha)) \cap \overline{D}(\theta_i, \alpha_{-i}, \succeq, \Lambda_i^f(\theta_i, \alpha)) \neq \varnothing$. ■

Proof of Proposition 2. Without loss of generality, let player 1 be such that $|S_1| > 2$ and for some $v_1 \notin V_1^*(\succeq)$, $f_1(v_1, v_{-1}) > 0$ for some v_{-1}. Let $V_1 = \{v_1^1, \ldots, v_1^{|V_1|}\}$ and $S_1 = \{s_1^1, \ldots, s_1^{|S_1|}\}$, and let v_1^1 be such that $v_1^1 \notin V_1^*(\succeq)$ and $f_1(v_1^1, v_{-1}) > 0$ for some v_{-1}. In what follows, we assume that $|\Theta_{-1}| > 1$, since for $|\Theta_{-1}| = 1$, trivially $\mathcal{T}^* = \varnothing$ and hence $\mathcal{T}^*$ is nowhere dense.

We treat $C(\alpha_{-1})$ as a subset of the $|\Theta_{-1}| - 1$ dimensional simplex and we say that $C(\alpha_{-1})$ is *full dimensional* if its affine hull has dimension $|\Theta_{-1}| - 1$.

Lemma 1 *For any $\sigma \in \overline{\mathcal{T}^*}$, if $C(\alpha_{-1})$ is full dimensional for all α_{-1}, then there exists an α such that for all θ_1 with $v_1 = v_1^1$, $\gamma(\theta_1, \alpha) \notin \mathrm{ri}(C(\alpha_{-1}))$.*

[28] Every face of a polytope is an exposed face.

Proof. Let $\sigma^n \to \sigma$ so that $\sigma^n \in \mathcal{T}^*$ for all n, and let α^n be such that for all θ_1 with $v_1 = v_1^1$, $\gamma^n(\theta_1, \alpha^n) \notin \text{ri}(C^n(\alpha^n_{-1}))$.[29] Such an α^n must exist by Proposition 1. Since $A^{|V|}$ is compact, we can assume that α^n converges to α. Note that **A3** implies that $\gamma_{\Theta_1}(\alpha')$ and $\gamma^n_{\Theta_1}(\alpha')$ have full support for all α' (and hence $\gamma^n(\theta_1, \alpha') \to \gamma(\theta_1, \alpha')$ for all θ_1 and α') and **A2** implies that $\alpha' \mapsto \gamma(\theta_1, \alpha')$ and $\alpha' \mapsto \gamma^n(\theta_1, \alpha')$ are continuous for all θ_1. Suppose for a contradiction that $C(\alpha_{-1})$ is full dimensional for all α_{-1} and for some θ_1 with $v_1 = v_1^1$, $\gamma(\theta_1, \alpha) \in \text{ri}(C(\alpha_{-1}))$. For large enough n, $\gamma(\theta_1, \alpha) \in \text{ri}(C^n(\alpha^n_{-1}))$,[30] and hence $\gamma^n(\theta_1, \alpha^n) \in \text{ri}(C^n(\alpha^n_{-1}))$, a contradiction. Thus for each θ_1 with $v_1 = v_1^1$, $\gamma(\theta_1, \alpha) \notin \text{ri}(C(\alpha_{-1}))$. ■

By Lemma 1, to show that $\overline{\mathcal{T}^*}$ has empty interior, it suffices to show that for any continuous $\sigma : A \times V \to \Delta S$ such that $\text{supp}(\sigma(a, v)) = S$ for all $(a, v) \in A \times V$ and for any $\varepsilon > 0$, there exists a continuous $\hat{\sigma} : A \times V \to \Delta S$ such that $\text{supp}(\hat{\sigma}(a, v)) = S$ for all $(a, v) \in A \times V$, and for every α, $\hat{C}(\alpha_{-1})$ is full dimensional and there exists s_1 such that $\hat{\gamma}((v_1^1, s_1), \alpha) \in \text{ri}(\hat{C}(\alpha_{-1}))$,[31] where $d(\sigma, \hat{\sigma}) \equiv \sup_{a,v} \|\hat{\sigma}(a, v) - \sigma(a, v)\| \leq \varepsilon$ (i.e. for every σ and every $\varepsilon > 0$, we can find $\hat{\sigma} \in \mathcal{T} \setminus \overline{\mathcal{T}^*}$ such that $d(\sigma, \hat{\sigma}) \leq \varepsilon$).

As mentioned in the main text, $\gamma(\theta_1, \alpha_1, \alpha_{-1})$ depends only on θ_1, $\alpha_1(v_1)$, and α_{-1}, but for notational simplicity we included α_1 as an argument. In this proof, it will be more convenient to explicitly restrict the argument to $\alpha_1(v_1)$, so with abuse of notation, we will write $\gamma(\theta_1, a_1, \alpha_{-1})$.

Since σ is continuous and A is compact, σ is uniformly continuous on A. Fix $\varepsilon > 0$ and let δ be such $\|a - a'\| \leq \delta \implies \sup_v \|\sigma(a, v) - \sigma(a', v)\| \leq \frac{\varepsilon\eta}{2|\Theta_{-1}|^2}$, where $\eta = \min_{v_{-1}} \pi(v_1^1)[v_{-1}]$.

Let q be the dimension of A_1. First, we define a tiling $\mathcal{H}$ of $\mathbb{R}^q$ using hypercubes with side length $\delta/\sqrt{q}$, i.e. let e_k denote the k-unit vector,[32] let $H = [0, 1]^q$ denote the unit hypercube and let $\mathcal{H} = \left\{ \frac{\delta}{\sqrt{q}}(H + \sum_{k=1}^q l_k e_k) : l \in \mathbb{N}^q \right\}$ denote the collection of hypercubes with side length $\delta/\sqrt{q}$, disjoint interiors and whose union is $\mathbb{R}^q$. Let

[29] Where $\gamma^n(\theta_1, \alpha^n)$ and $C^n(\alpha^n_{-1})$ denote the beliefs and the convex hull of the beliefs under σ^n.

[30] Since $C(\alpha_{-1})$ is full dimensional and $\gamma(\theta_1, \alpha) \in \text{ri}(C(\alpha_{-1}))$, there must exist $(\theta_1^k, \alpha_1^k)_{k=1}^K$ such that $co\{\gamma(\theta_1^k, \alpha_1^k, \alpha_{-1}) : k \in \{1, \ldots, K\}\}$ contains an open ball around $\gamma(\theta_1, \alpha)$. Then for large n $co\{\gamma^n(\theta_1^k, \alpha_1^k, \alpha^n_{-1}) : k \in \{1, \ldots, K\}\} \subseteq C^n(\alpha^n_{-1})$ also contains an open ball around $\gamma(\theta_1, \alpha)$ which implies that $\gamma(\theta_1, \alpha) \in \text{ri}(C^n(\alpha^n_{-1}))$.

[31] Where $\hat{\gamma}(\theta_1, \alpha)$ and $\hat{C}(\alpha_{-1})$ denote the beliefs and the convex hull of the beliefs under $\hat{\sigma}$.

[32] I.e. the kth coordinate of e_k is 1.

$A_1 = \cup_k A_{1,k}$, where each $A_{1,k} = A_1 \cap h$ for some $h \in \mathcal{H}$. Since A_1 is bounded, there exists K such that $A_1 = \cup_{k=1}^{K} A_{1,k}$. Moreover, let K be minimal for this purpose and note that the convexity of A_1 implies each $A_{1,k}$ has dimension q. Also note that $a_1, a_1' \in A_{1,k}$ implies that $\|a_1 - a_1'\| \leq \delta$.

Let:

$$
\begin{aligned}
V_{-1} &= \left\{ v_{-1}^1, \ldots v_{-1}^{|V_{-1}|} \right\} \\
S_{-1} &= \left\{ s_{-1}^1, \ldots s_{-1}^{|S_{-1}|} \right\}.
\end{aligned}
$$

For $1 \leq l \leq |V_{-1}|$ and $1 \leq m \leq |S_{-1}|$, let $\phi_{l,m} \equiv (l-1)|S_{-1}| + m$. That is:

$$
\Theta_{-1} = \left\{ \theta_{-1}^{\phi_{1,1}}, \ldots, \theta_{-1}^{\phi_{1,|S_{-1}|}}, \ldots, \theta_{-1}^{\phi_{|V_{-1}|,1}}, \ldots, \theta_{-1}^{\phi_{|V_{-1}|,|S_{-1}|}} \right\},
$$

where $\theta_{-1}^{\phi_{l,m}} = (v_{-1}^l, s_{-1}^m)$.

For each $k \in K$, let $a_{1,k} \in \mathrm{int}(A_{1,k})$, let $B_k \subseteq A_{1,k}$ be an open ball around $a_{1,k}$ such that $\overline{B}_k \subseteq \mathrm{int}(A_{1,k})$, and let $\{b_{k,j}^1\}_{j=1}^{2|\Theta_{-i}|}$ and $\{b_{k,j}^2\}_{j=1}^{2|\Theta_{-i}|}$ be collections of distinct $4|\Theta_{-i}|$ points in B_k. For each k, j, let $B_{k,j}^1$ and $B_{k,j}^2$ be the open ball of radius r around $b_{k,j}^1$ and $b_{k,j}^2$ respectively. Assume that r is sufficiently small so that for each $k \in \{1, \ldots, K\}$, $j, j' \in \{1, \ldots, 2|\Theta_{-i}|\}$ and $l, l' \in \{1, 2\}$, $B_{k,j}^1 \subseteq B_k$, $B_{k,j}^2 \subseteq B_k$, and $B_{k,j}^l \cap B_{k,j'}^{l'} = \varnothing$.

For each a_1 in the boundary of $A_{1,k}$, let $\mathrm{proj}_{\overline{B}_k}(a_1) \in \overline{B}_k$ be the closest point in $\overline{B}_k$ to a_1, which exists and is unique since $\overline{B}_k$ is closed and convex, and let $l(a_1)$ be the line connecting a_1 and $\mathrm{proj}_{\overline{B}_k}(a_1)$. Note that each $a_1 \in A_{1,k} \setminus \overline{B}_k$ lies on a unique line $l(a_1')$ for some a_1' in the boundary of $A_{1,k}$. For each $a_1 \in A_{1,k} \setminus \overline{B}_k$, let $a_1 \mapsto g(a_1)$ be such that $g(a_1)$ is on the boundary of $A_{1,k}$ and $a_1 \in l(g(a_1))$.

For a set Y, let $\mathcal{I}_Y$ denote the indicator function of Y, i.e. $\mathcal{I}_Y(y) = 1$ if $y \in Y$ and $\mathcal{I}_Y(y) = 0$ if $y \notin Y$. We define $\hat{\sigma}$ as follows.

1. For any (a, v) such that $v_1 \neq v_1^1$, let $\hat{\sigma}(a, v) = \sigma(a, v)$.

2. For (v_1^1, s_1^1), (v_{-1}^l, s_{-1}^m), for each $a_1 \in A_{1,k}$, and for all a_{-1}:

$$
\begin{aligned}
&\hat{\sigma}(a_1, a_{-1}, v_1^1, v_{-1}^l)[s_1^1, s_{-1}^m] \\
&= \sigma(a_{1,k}, a_{-1}, v_1^1, v_{-1}^l)[s_1^1, s_{-1}^m] \\
&+ \mathcal{I}_{B^1_{k,2\phi_{l,m}-1}}(a_1) \left(1 - \frac{\|a_1 - b^1_{k,2\phi_{l,m}-1}\|}{r}\right) \frac{\varepsilon\eta}{2\pi(v_1^1)[v_{-1}^l]\sqrt{|\Theta_{-1}|}} \\
&- \mathcal{I}_{B^1_{k,2\phi_{l,m}}}(a_1) \left(1 - \frac{\|a_1 - b^1_{k,2\phi_{l,m}}\|}{r}\right) \frac{\varepsilon\eta}{2\pi(v_1^1)[v_{-1}^l]\sqrt{|\Theta_{-1}|}} \\
&+ \mathcal{I}_{A_{1,k}\setminus\overline{B}_k}(a_1) \frac{\|a_1 - \mathrm{proj}_{\overline{B}_k}(g(a_1))\|}{\|g(a_1) - \mathrm{proj}_{\overline{B}_k}(g(a_1))\|} \\
&\qquad \times \left(\sigma(g(a_1), a_{-1}, v_1^1, v_{-1}^l)[s_1^1, s_{-1}^m] - \sigma(a_{1,k}, a_{-1}, v_1^1, v_{-1}^l)[s_1^1, s_{-1}^m]\right).
\end{aligned}
$$

3. For (v_1^1, s_1^2), (v_{-1}^l, s_{-1}^m), for each $a_1 \in A_{1,k}$, and for all a_{-1}:

$$
\begin{aligned}
&\hat{\sigma}(a_1, a_{-1}, v_1^1, v_{-1}^l)[s_1^2, s_{-1}^m] \\
&= \sigma(a_{1,k}, a_{-1}, v_1^1, v_{-1}^l)[s_1^2, s_{-1}^m] \\
&+ \mathcal{I}_{B^2_{k,2\phi_{l,m}-1}}(a_1) \left(1 - \frac{\|a_1 - b^2_{k,2\phi_{l,m}-1}\|}{r}\right) \frac{\varepsilon\eta}{2\pi(v_1^1)[v_{-1}^l]\sqrt{|\Theta_{-1}|}} \\
&- \mathcal{I}_{B^2_{k,2\phi_{l,m}}}(a_1) \left(1 - \frac{\|a_1 - b^2_{k,2\phi_{l,m}}\|}{r}\right) \frac{\varepsilon\eta}{2\pi(v_1^1)[v_{-1}^l]\sqrt{|\Theta_{-1}|}} \\
&+ \mathcal{I}_{A_{1,k}\setminus\overline{B}_k}(a_1) \frac{\|a_1 - \mathrm{proj}_{\overline{B}_k}(g(a_1))\|}{\|g(a_1) - \mathrm{proj}_{\overline{B}_k}(g(a_1))\|} \\
&\qquad \times \left(\sigma(g(a_1), a_{-1}, v_1^1, v_{-1}^l)[s_1^2, s_{-1}^m] - \sigma(a_{1,k}, a_{-1}, v_1^1, v_{-1}^l)[s_1^2, s_{-1}^m]\right).
\end{aligned}
$$

4. For (v_1^1, s_1^3), (v_{-1}^l, s_{-1}^m), for each $a_1 \in A_{1,k}$, and for all a_{-1}:

$$
\begin{aligned}
&\hat{\sigma}(a_1, a_{-1}, v_1^1, v_{-1}^l)[s_1^3, s_{-1}^m] \\
&= \sigma(a_{1,k}, a_{-1}, v_1^1, v_{-1}^l)[s_1^3, s_{-1}^m] \\
&- \mathcal{I}_{B^1_{k,2\phi_{l,m}-1}}(a_1) \left(1 - \frac{\|a_1 - b^1_{k,2\phi_{l,m}-1}\|}{r}\right) \frac{\varepsilon\eta}{2\pi(v_1^1)[v_{-1}^l]\sqrt{|\Theta_{-1}|}} \\
&+ \mathcal{I}_{B^1_{k,2\phi_{l,m}}}(a_1) \left(1 - \frac{\|a_1 - b^1_{k,2\phi_{l,m}}\|}{r}\right) \frac{\varepsilon\eta}{2\pi(v_1^1)[v_{-1}^l]\sqrt{|\Theta_{-1}|}} \\
&- \mathcal{I}_{B^2_{k,2\phi_{l,m}-1}}(a_1) \left(1 - \frac{\|a_1 - b^2_{k,2\phi_{l,m}-1}\|}{r}\right) \frac{\varepsilon\eta}{2\pi(v_1^1)[v_{-1}^l]\sqrt{|\Theta_{-1}|}} \\
&+ \mathcal{I}_{B^2_{k,2\phi_{l,m}}}(a_1) \left(1 - \frac{\|a_1 - b^2_{k,2\phi_{l,m}}\|}{r}\right) \frac{\varepsilon\eta}{2\pi(v_1^1)[v_{-1}^l]\sqrt{|\Theta_{-1}|}} \\
&+ \mathcal{I}_{A_{1,k}\setminus\overline{B}_k}(a_1) \frac{\|a_1 - \mathrm{proj}_{\overline{B}_k}(g(a_1)\|}{\|g(a_1) - \mathrm{proj}_{\overline{B}_k}(g(a_1))\|} \\
&\quad \times \left(\sigma(g(a_1), a_{-1}, v_1^1, v_{-1}^l)[s_1^3, s_{-1}^m] - \sigma(a_{1,k}, a_{-1}, v_1^1, v_{-1}^l)[s_1^3, s_{-1}^m]\right).
\end{aligned}
$$

5. For (v_1^1, s_1^j), $j \in \{4, \ldots, |S_1|\}$, (v_{-1}^l, s_{-1}^m), for each $a_1 \in A_{1,k}$, and for all a_{-1}:

$$
\begin{aligned}
&\hat{\sigma}(a_1, a_{-1}, v_1^1, v_{-1}^l)[s_1^j, s_{-1}^m] \\
&= \sigma(a_{1,k}, a_{-1}, v_1^1, v_{-1}^l)[s_1^j, s_{-1}^m] \\
&+ \mathcal{I}_{A_{1,k}\setminus\overline{B}_k}(a_1) \frac{\|a_1 - \mathrm{proj}_{\overline{B}_k}(g(a_1))\|}{\|g(a_1) - \mathrm{proj}_{\overline{B}_k}(g(a_1))\|} \\
&\times \left(\sigma(g(a_1), a_{-1}, v_1^1, v_{-1}^l)[s_1^j, s_{-1}^m] - \sigma(a_{1,k}, a_{-1}, v_1^1, v_{-1}^l)[s_1^j, s_{-1}^m]\right).
\end{aligned}
$$

Note that by construction, for each $a_1 \in A_{1,k}$ and each $(a_{-1}, v_{-1}) \in A_{-1} \times V_{-1}$,

$$
\begin{aligned}
&\sum_s \hat{\sigma}(a_1, a_{-1}, v_1^1, v_{-1})[s] \\
&= \sum_s \sigma(a_{1,k}, a_{-1}, v_1^1, v_{-1})[s] + \mathcal{I}_{A_{1,k}\setminus\overline{B}_k}(a_1) \frac{\|a_1 - \mathrm{proj}_{\overline{B}_k}(g(a_1))\|}{\|g(a_1) - \mathrm{proj}_{\overline{B}_k}(g(a_1))\|} \\
&\quad \times \sum_s \left(\sigma(g(a_1), a_{-1}, v_1^1, v_{-1})[s] - \sigma(a_{1,k}, a_{-1}, v_1^1, v_{-1})[s]\right) = 1.
\end{aligned}
$$

For each $\ell \in \{1,2\}$, define $\hat{y}_\ell(a_1, \alpha_{-1}) \in \mathbb{R}^{|\Theta_{-1}|}$ such that its $\phi_{l,m}$-th component is $\pi(v_1^1)[v_{-1}^l]\hat{\sigma}(a_1, \alpha_{-1}(v_{-1}^1), v_1^1, v_{-1}^l)[s_1^\ell, s_{-1}^m]$, i.e. $\frac{\hat{y}_\ell(a_1,\alpha_{-1})}{\|\hat{y}_\ell(a_1,\alpha_{-1})\|_1} = \hat{\gamma}((v_1^1, s_1^\ell), a_1, \alpha_{-1})$, and define $y_\ell(a_1, \alpha_{-1})$ analogously but with $\hat{\sigma}$ replaced by σ. Note that $\hat{\sigma}$ is defined so that the convex hull of the image of B_k under $a_1 \mapsto \hat{y}_\ell(a_1, \alpha_{-1})$ is a $|\Theta_{-1}|$ dimensional cross polytope with centre $y_\ell(a_{1,k}, \alpha_{-1})$ and circumradius $\frac{\varepsilon\eta}{2\sqrt{|\Theta_{-1}|}}$, as:

- $\hat{y}_\ell(\alpha_1, \alpha_{-1}) = y_\ell(a_{1,k}, \alpha_{-1})$ when $a_1 \in B_k \setminus \cup B_{k,j}^\ell$
- $\hat{y}_\ell(a_1, \alpha_{-1}) = y_\ell(a_{1,k}, \alpha_{-1}) + (\frac{\varepsilon\eta}{2\sqrt{|\Theta_{-1}|}}, 0, 0, 0, \ldots)$ at $a_1 = b_{k,1}^\ell$
- $\hat{y}_\ell(a_1, \alpha_{-1}) = y_\ell(a_{1,k}, \alpha_{-1}) - (\frac{\varepsilon\eta}{2\sqrt{|\Theta_{-1}|}}, 0, 0, 0, \ldots)$ at $a_1 = b_{k,2}^\ell$
- $\hat{y}_\ell(a_1, \alpha_{-1}) = y_\ell(a_{1,k}, \alpha_{-1}) + (0, \frac{\varepsilon\eta}{2\sqrt{|\Theta_{-1}|}}, 0, 0, \ldots)$ at $a_1 = b_{k,3}^\ell$, etc,

and for each $a_1 \in B_{k,j}^\ell \setminus \{b_{k,j}^\ell\}$, $\hat{y}_\ell(a_1, \alpha_{-1})$ is a strict convex combination of $y_\ell(a_{1,k}, \alpha_{-1})$ and $\hat{y}_\ell(b_{k,j}^\ell, \alpha_{-1})$.

Define:

$$
\begin{aligned}
A_{1,k}^* &= \{b_{k,j}^1 : 1 \leq j \leq 2|\Theta_{-i}|\} \\
A_{1,k}^{**} &= \{b_{k,j}^2 : 1 \leq j \leq 2|\Theta_{-i}|\}.
\end{aligned}
$$

Lemma 2 *For each α_{-1}, $\hat{C}(\alpha_{-1})$ is full dimensional.*

Proof. Fix α_{-1} and note that $\hat{\gamma}((v_1^1, s_1^1), a_1, \alpha_{-1}) = \frac{\hat{y}_1(a_1,\alpha_{-1})}{\|\hat{y}_1(a_1,\alpha_{-1})\|_1}$. Since, for example, $\text{co}\{\hat{y}_1(a_1', \alpha_{-1}) : a_1' \in A_{1,k}^*\}$ is $|\Theta_{-1}|$ dimensional, $\text{co}\{\hat{\gamma}((v_1^1, s_1^1), a_1', \alpha_{-1}) : a_1' \in A_{1,k}^*\} = \text{co}\{\frac{\hat{y}_1(a_1',\alpha_{-1})}{\|\hat{y}_1(a_1',\alpha_{-1})\|_1} : a_1' \in A_{1,k}^*\}$ is $|\Theta_{-1}| - 1$ dimensional and $\text{co}\{\hat{\gamma}((v_1^1, s_1^1), a_1', \alpha_{-1}) : a_1' \in A_{1,k}^*\} \subseteq \hat{C}(\alpha_{-1})$. ■

Lemma 3 *For each $k \in \{1, \ldots, K\}$, $a_1 \in A_{1,k}$, and for all α_{-1}:*

1. *If $\hat{y}_1(a_1, \alpha_{-1}) \in \text{int}(co\{\hat{y}_1(a_1', \alpha_{-1}) : a_1' \in A_{1,k}^*\})$, $\hat{\gamma}((v_1^1, s_1^1), a_1, \alpha_{-1}) \in \text{ri}(\hat{C}(\alpha_{-1}))$.*
2. *If $\hat{y}_2(a_1, \alpha_{-1}) \in \text{int}(co\{\hat{y}_2(a_1', \alpha_{-1}) : a_1' \in A_{1,k}^{**}\})$, $\hat{\gamma}((v_1^1, s_1^2), a_1, \alpha_{-1}) \in \text{ri}(\hat{C}(\alpha_{-1}))$.*

Proof. First, note that if $\hat{y}_1(a_1, \alpha_{-1}) = \sum_{a_1' \in A_{1,k}^*} \lambda(a_1')\hat{y}_1(a_1', \alpha_{-1})$, where $\lambda(a_1') \in (0, 1)$ for each $a_1' \in A_{1,k}^*$ and $\sum_{a_i' \in A_{1,k}^*} \lambda(a_1') = 1$, then $\hat{\gamma}((v_1^1, s_1^1), a_1, \alpha_{-1}) =$

$\sum_{a_1'} \lambda'(a_1')\hat{\gamma}((v_1^1, s_1^1), a_1', \alpha_{-1})$ for $\lambda'(a_1') \in (0,1)$ for each $a_1' \in A_{1,k}^*$ and $\sum_{a_i' \in A_{1,k}^*} \lambda'(a_1') = 1$. To see this, note that:

$$
\begin{aligned}
&\hat{\gamma}((v_1^1, s_1^1), a_1, \alpha_{-1})[\theta_{-1}] \\
&= \frac{\sum_{a_1' \in A_{1,k}^*} \lambda(a_1')\pi(v_1^1)[v_{-1}]\hat{\sigma}(a_1', \alpha_{-1}(v_{-1})v_1^1, v_{-1})[s_1^1, s_{-1}]}{\sum_{a_1' \in A_{1,k}^*} \lambda(a_1')\left(\sum_{v_{-1}', s_{-1}'} \pi(v_1^1)[v_{-1}']\hat{\sigma}(a_1', \alpha_{-1}(v_{-1}'), v_1^1, v_{-1}')[s_1^1, s_{-1}']\right)} \\
&= \sum_{a_1' \in A_{1,k}^*} \lambda'(a_1')\hat{\gamma}((v_1^1, s_1^1), a_1', \alpha_{-1})[\theta_{-1}]
\end{aligned}
$$

where:

$$
\lambda'(a_1') = \frac{\lambda(a_1')\sum_{v_{-1}', s_{-1}'} \pi(v_1^1)[v_{-1}']\hat{\sigma}(a_1', \alpha_{-1}(v_{-1}'), v_1^1, v_{-1}')[s_1^1, s_{-1}']}{\sum_{a_1' \in A_{1,k}^*} \lambda(a_1')\sum_{v_{-1}', s_{-1}'} \pi(v_1^1)[v_{-1}']\hat{\sigma}(a_1', \alpha_{-1}(v_{-1}'), v_1^1, v_{-1}')[s_1^1, s_{-1}']}.
$$

Now, if $\hat{y}_1(a_1, \alpha_{-1}) \in \mathrm{int}(co\{\hat{y}_1(a_1', \alpha_{-1}) : a_1' \in A_{1,k}^*\})$, then $\hat{y}_1(a_1, \alpha_{-1})$ can be written as a strict convex combination of $\hat{y}_1(a_1', \alpha_{-1})$ for $a_1' \in A_{1,k}^*$. This implies that $\hat{\gamma}((v_1^1, s_1^1), a_1, \alpha_{-1})$ is a strict convex combination of $\hat{\gamma}((v_1^1, s_1^1), a_1', \alpha_{-1})$ for $a_1' \in A_{1,k}^*$, and since $co\{\hat{\gamma}((v_1^1, s_1^1), a_1', \alpha_{-1}) : a_1' \in A_{1,k}^*\} = co\{\frac{\hat{y}_1(a_1', \alpha_{-1})}{\|\hat{y}_1(a_1', \alpha_{-1})\|_1} : a_1' \in A_{1,k}^*\}$ has dimension $|\Theta_{-1}| - 1$, $\hat{\gamma}((v_1^1, s_1^1), a_1, \alpha_{-1}) \in \mathrm{ri}(co\{\hat{\gamma}((v_1^1, s_1^1), a_1', \alpha_{-1}) : a_1' \in A_{1,k}^*\}) \subseteq \mathrm{ri}(\hat{C}(\alpha_{-1}))$.

Analogous arguments establish the second part of the Lemma. ■

Lemma 4 *For any $k \in \{1, \ldots, K\}$, and for all α_{-1}:*

1. *For all $a_1 \in A_{1,k} \setminus A_{1,k}^*$, $\hat{y}_1(a_1, \alpha_{-1}) \in \mathrm{int}(co\{\hat{y}_1(a_1, \alpha_{-1}) : a_1 \in A_{1,k}^*\})$.*

2. *For all $a_1 \in A_{1,k} \setminus A_{1,k}^{**}$, $\hat{y}_2(a_1, \alpha_{-1}) \in \mathrm{int}(co\{\hat{y}_2(a_1, \alpha_{-1}) : a_1 \in A_{1,k}^{**}\})$.*

Proof. Note that the convex hull of the image of B_k under $a_1 \mapsto \hat{y}_1(a_1, \alpha_{-1})$ is a cross polytope with centre $y_1(a_{1,k}, \alpha_{-1})$ and circumradius $\frac{\varepsilon\eta}{2\sqrt{|\Theta_{-1}|}}$ that is equal to $co\{\hat{y}_1(a_1, \alpha_{-1}) : a_1 \in A_{1,k}^*\}$. We will show that for $a_1 \in A_{1,k} \setminus A_{1,k}^*$, $\hat{y}_1(a_1, \alpha_{-1}) \in \mathrm{int}(co\{\hat{y}_1(a_1, \alpha_{-1}) : a_1 \in A_{1,k}^*\})$.

For each $a_1 \in B_{k,j}^1 \setminus \{b_{k,j}^1\}$, $\hat{y}_1(a_1, \alpha_{-1})$ is a strict convex combination of $y_1(a_{1,k}, \alpha_{-1})$ and $\hat{y}_1(b_{k,j}^1, \alpha_{-1})$; thus, $\hat{y}_1(a_1, \alpha_{-1}) \in \mathrm{int}(co\{\hat{y}_1(a_1, \alpha_{-1}) : a_1 \in A_{1,k}^*\})$. For $a_1 \in$

$\overline{B}_k \setminus \cup_j B^1_{k,j}$, $\hat{y}_1(a_1, \alpha_{-1}) = y_1(a_{1,k}, \alpha_{-1}) \in \text{int}(co\{\hat{y}_1(a_1, \alpha_{-1}) : a_1 \in A^*_{1,k}\})$. Thus, for $a_1 \in \overline{B}_k \setminus A^*_{1,k}$, $\hat{y}_1(a_1, \alpha_{-1}) \in \text{int}(co\{\hat{y}_1(a_1, \alpha_{-1}) : a_1 \in A^*_{1,k}\})$. Analogous arguments establish that for $a_1 \in \overline{B}_k \setminus A^{**}_{1,k}$, $\hat{y}_2(a_1, \alpha_{-1}) \in \text{int}(co\{\hat{y}_2(a_1, \alpha_{-1}) : a_1 \in A^{**}_{1,k}\})$.

For $a_1 \in A_{1,k} \setminus \overline{B}_k$, note that the convex hull of the image of B_k under $a_1 \mapsto \hat{y}_1(a_1, \alpha_{-1})$ is a cross polytope that contains a hypersphere with radius $\frac{\varepsilon\eta}{2|\Theta_{-1}|}$ centred around $y_1(a_{1,k}, \alpha_{-1})$. Since $\|\sigma(a_{1,k}, a_{-1}, v) - \sigma(a_1, a_{-1}, v)\| \leq \frac{\varepsilon\eta}{2|\Theta_{-1}|^2}$, for each (a_{-1}, v_{-1}, s_{-1}), $|\sigma(a_{1,k}, a_{-1}, v^1_1, v_{-1})[s^1_1, s_{-1}] - \sigma(a_1, a_{-1}, v^1_1, v_{-1})[s^1_1, s_{-1}]| \leq \frac{\varepsilon\eta}{2|\Theta_{-1}|^2}$. It follows that, for each a_1 in the boundary of $A_{1,k}$, $\|y_1(a_{1,k}, \alpha_{-1}) - y_1(a_1, \alpha_{-1})\| \leq \frac{\varepsilon\eta}{2|\Theta_{-1}|}$, and hence, for each $a_1 \in A_{1,k} \setminus \overline{B}_k$, $y_1(g(a_1), \alpha_{-1}) \in \text{int}(co\{\hat{y}_1(a_1, \alpha_{-1}) : a_1 \in A^*_{1,k}\})$. Since for $a_1 \in A_{1,k} \setminus \overline{B}_k$, $\hat{y}_1(a_1, \alpha_{-1})$ is a convex combination of $y_1(a_{1,k}, \alpha_{-1})$ and $y_1(g(a_1), \alpha_{-1})$, it follows that $\hat{y}_1(a_1, \alpha_{-1}) \in \text{int}(co\{\hat{y}_1(a_1, \alpha_{-1}) : a_1 \in A^*_{1,k}\})$. An analogous argument establishes that for $A_{1,k} \setminus \overline{B}_k$ $\hat{y}_2(a_1, \alpha_{-1}) \in \text{int}(co\{\hat{y}_2(a_1, \alpha_{-1}) : a_1 \in A^{**}_{1,k}\})$. ■

Lemma 5 *For each l, m, s_1, and a_{-1}, $a_1 \mapsto \hat{\sigma}(a_1, a_{-1}, v^1_1, v^l_{-1})[s_1, s^m_{-1}]$ is continuous.*

Proof. First, let $a_1 \in B_k \setminus (B^1_{k,2\phi_{l,m}-1} \cup B^1_{k,2\phi_{l,m}} \cup B^2_{k,2\phi_{l,m}-1} \cup B^2_{k,2\phi_{l,m}})$, and consider $\{a^n_1\}^\infty_{n=1}$ such that $a^n_1 \to a_1$. If $a^n_1 \in B_k \setminus (B^1_{k,2\phi_{l,m}-1} \cup B^1_{k,2\phi_{l,m}} \cup B^2_{k,2\phi_{l,m}-1} \cup B^2_{k,2\phi_{l,m}})$ then $\hat{\sigma}(a_1, a_{-1}, v^1_1, v^l_{-1})[s_1, s^m_{-1}] = \hat{\sigma}(a^n_1, a_{-1}, v^1_1, v^l_{-1})[s_1, s^m_{-1}]$.

If, for $\ell \in \{1, 2\}$, $j \in \{2\phi_{l,m} - 1, 2\phi_{l,m}\}$ $a^n_1 \in B^\ell_{k,j}$ for infinitely many n, then $a_1 \in \overline{B}^\ell_{k,j}$, i.e. $\|a_1 - b^\ell_{k,j}\| = r$. Thus, $\left(1 - \frac{\|a^n_1 - b^\ell_{k,j}\|}{r}\right) \to 0$ and $\hat{\sigma}(a^n_1, a_{-1}, v^1_1, v^l_{-1})[s_1, s^m_{-1}] \to \sigma(a_{1,k}, a_{-1}, v^1_1, v^l_{-1})[s_1, s^m_{-1}] = \hat{\sigma}(a_1, a_{-1}, v^1_1, v^l_{-1})[s_1, s^m_{-1}]$.

Now let $a_1 \in B^\ell_{k,j}$, for $\ell \in \{1, 2\}$, $j \in \{2\phi_{l,m} - 1, 2\phi_{l,m}\}$, and consider $a_1 \in B^\ell_{k,j}$. Note that for any $a'_1 \in B^\ell_{k,j}$,

$$\begin{aligned} &|\hat{\sigma}(a'_1, a_{-1}, v^1_1, v^l_{-1})[s_1, s^m_{-1}] - \hat{\sigma}(a_1, a_{-1}, v^1_1, v^l_{-1})[s_1, s^m_{-1}]| \\ \leq\ &\frac{\varepsilon}{2r\sqrt{|\Theta_{-1}|}} \left| \|a_1 - b^\ell_{k,j}\| - \|a'_1 - b^\ell_{k,j}\| \right| \leq \frac{\varepsilon}{2r\sqrt{|\Theta_{-1}|}} \|a_1 - a'_1\|. \end{aligned}$$

Thus, for $\{a^n_1\}^\infty_{n=1}$ such that $a^n_1 \to a_1$, we have that $\hat{\sigma}(a^n_1, a_{-1}, v^1_1, v^l_{-1})[s_1, s^m_{-1}] \to \hat{\sigma}(a_1, a_{-1}, v^1_1, v^l_{-1})[s_1, s^m_{-1}]$.

Next let $a_1 \in \overline{B}_k \setminus B_k$ and consider $\{a^n_1\}^\infty_{n=1}$ such that $a^n_1 \to a_1$. If $a^n_1 \in \overline{B}_k$

then $\hat{\sigma}(a_1, a_{-1}, v_1^1, v_{-1}^l)[s_1, s_{-1}^m] = \hat{\sigma}(a_1^n, a_{-1}, v_1^1, v_{-1}^l)[s_1, s_{-1}^m]$. If $a_1^n \notin \overline{B}_k$, then since $\|a_1^n - \text{proj}_{\overline{B}_k}(g(a_1^n))\| \to 0$, $\hat{\sigma}(a_1^n, a_{-1}, v_1^1, v_{-1}^l)[s_1, s_{-1}^m] \to \sigma(a_{1,k}, a_{-1}, v_1^1, v_{-1}^l)[s_1, s_{-1}^m] = \hat{\sigma}(a_1, a_{-1}, v_1^1, v_{-1}^l)[s_1, s_{-1}^m]$.

Finally, let $a_1 \in A_{1,k} \setminus \overline{B}_k$. Then

$$\begin{aligned}\hat{\sigma}(a_1, a_{-1}, v_1^1, v_{-1}^l)[s_1, s_{-1}^m] = &\frac{\|a_1 - \text{proj}_{B_k}(a_1)\|}{\|g(a_1) - \text{proj}_{\overline{B}_k}(a_1)\|}\sigma(g(a_1), a_{-1}, v_1^1, v_{-1}^l)[s_1, s_{-1}^m] \\ &+ \left(1 - \frac{\|a_1 - \text{proj}_{B_k}(a_1)\|}{\|g(a_1) - \text{proj}_{B_k}(a_1)\|}\right)\sigma(a_{1,k}, a_{-1}, v_1^1, v_{-1}^l)[s_1, s_{-1}^m]\end{aligned}$$

which is continuous in a_1 since since $\text{proj}_{\overline{B}_k}$, g and $a_1 \mapsto \sigma(a_1, a_{-1}, v_1^1, v_{-1}^l)[s_1, s_{-1}^m]$ are continuous.

Thus, $a_1 \mapsto \hat{\sigma}(a_1, a_{-1}v_1^1, v_{-1}^l)[s_1, s_{-1}^m]$ is continuous over $A_{1,k}$. Since this is true for every k, the lemma follows. ■

Lemma 6 *$\hat{\sigma}$ is continuous, has full support, and $d(\hat{\sigma}, \sigma) \leq \varepsilon$.*

Proof. To see that $\hat{\sigma}$ is continuous, we consider the case $v_1 = v_1^1$ (otherwise $\hat{\sigma}(a, v) = \sigma(a, v)$ and continuity follows from **A2**). Note that for each s, v_{-1}, and a_{-1}, $a_1 \mapsto \hat{\sigma}(a_1, a_{-1}, v_1^1, v_{-1})[s]$ is continuous by Lemma 5. Moreover, for $a_1 \in A_{1,k}$, $\hat{\sigma}(a_1, a_{-1}, v_1^1, v_{-1})[s]$ is the sum of $\sigma(a_{1,k}, a_{-1}, v_1^1, v_{-1})[s]$, $\sigma(g(a_1), a_{-1}, v_1^1, v_{-1})[s]$, and functions that are constant (and hence continuous) in a_{-1}. Thus, $\hat{\sigma}$ is continuous. Note that for $a_1 \in A_{1,k}$, for all a_{-1} and v, $|\hat{\sigma}(a_1, a_{-1}, v)[s] - \sigma(a_{1,k}, a_{-1}, v)[s]| \leq \frac{\varepsilon}{2\sqrt{|\Theta_{-1}|}}$ for all s. Thus, $\|\hat{\sigma}(a_1, a_{-1}, v) - \sigma(a_1, a_{-1}, v)\| \leq \|\hat{\sigma}(a_1, a_{-1}, v) - \sigma(a_{1,k}, a_{-1}, v)\| + \|\sigma(a_1, a_{-1}, v) - \sigma(a_{1,k}, a_{-1}, v)\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon\eta}{2|\Theta_{-1}|^2} \leq \varepsilon$. Thus, $d(\sigma, \hat{\sigma}) \leq \varepsilon$. Finally, for sufficiently small ε, full support of σ implies the full support of $\hat{\sigma}$. ■

Lemmas 2–4 imply that $\hat{C}(\alpha_{-1})$ is full dimensional for every α_{-1} and for every $a_1 \in A_{1,k}$, either $\hat{\gamma}((v_1^1, s_1^1), a_1, \alpha_{-1}) \in \text{ri}(\hat{C}(\alpha_{-1}))$ or $\hat{\gamma}((v_1^1, s_1^2), a_1, \alpha_{-1}) \in \text{ri}(\hat{C}(\alpha_{-1}))$. Since this is true for all α_{-1}, it follows that for every α, there exists θ_1 such that $\hat{\gamma}(\theta_1, \alpha) \in \text{ri}(\hat{C}(\alpha_{-1}))$. Combined with Lemma 1 and Lemma 6, this implies that $\hat{\sigma} \in \mathcal{T} \setminus \overline{\mathcal{T}^*}$ and $d(\sigma, \hat{\sigma}) \leq \varepsilon$. Since σ and ε were arbitrary, this implies that $\overline{\mathcal{T}^*}$ has empty interior. ■