London School of Economics and Political Science

Essays on Auctions, Mechanism Design, and Repeated Games

Krittanai Laohakunakorn

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy, July 2019

Declaration

I certify that this thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I confirm that Chapter 2 was jointly co-authored with Professor Gilat Levy and Professor Ronny Razin, and I contributed 33% of this work. A version of this chapter is forthcoming in the *Journal of Economic Theory*, Volume 184, November 2019.

I declare that my thesis consists of approximately 25,000 words.

Abstract

Chapter 1 revisits the classic mechanism design question of when buyers with private information 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. This assumption can be seen as a special case of a general model where buyers have access to some information acquisition technology. We provide necessary and sufficient conditions on the information acquisition technology for the seller to be able to guarantee full rent extraction.

Chapter 2 studies auctions when there is ambiguity over the joint information structures generating the valuations and signals of players. We analyse how two standard auction effects interact with ambiguity. First, a 'competition effect' arises when different beliefs about the correlation between bidders' valuations imply different likelihoods of facing competitive bids. Second, a 'winner's value effect' arises when different beliefs imply different inferences about the winner's value. In private value auctions, only the first effect exists, and the distribution of bids first order stochastically dominates the distribution of bids in the absence of ambiguity. In common value auctions both effects exist, and the seller's revenue decreases with ambiguity.

Chapter 3 characterises the equilibrium payoff set of a repeated game with local interaction and local monitoring. A Nash threats folk theorem holds without any restrictions on the network structure when players are arbitrarily patient, i.e. any feasible payoff above the Nash equilibrium point can be approximated arbitrarily well in sequential equilibrium. When players discount the future, the folk theorem cannot hold unless further restrictions are made either on payoffs or the network structure.

Acknowledgements

I am indebted to my supervisor, Francesco Nava, for excellent advice and guidance, and I am deeply grateful to the rest of the theory faculty at LSE for their generosity and support.

Contents

List of Figures

List of Tables

Chapter 1

Rent Extraction and Information Acquisition

1.1 Introduction

Agents with private information who interact in strategic situations often have incentives to acquire information about each other before making their decisions, and in many relevant environments, it is also plausible that they have the opportunity to do so. 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.

It is well known that in the absence of information acquisition, full rent extraction is possible for generic prior distributions of the buyers' valuations. For example, consider a mechanism where the payment of each buyer consists of two terms. First, each buyer pays the product of her valuation for the object and the probability of winning for each profile of her opponents' types (so that without the second term, she always receives exactly zero payoff ex post). The second term is a 'side bet' with the seller about the types of the other buyers. 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 for each type of buyer 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.

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 a strong incentive to learn about the types of their opponents, and having acquired this information, misreport their own type or drop out of the mechanism altogether. Thus, if buyers are able to acquire information, then this 'side bet' mechanism may fail to fully extract rent. However, this does not necessarily mean that the seller cannot extract rent using some other mechanism that exploits the buyers' information. For example, if each buyer can perfectly learn the other buyers' types, the seller can extract rent using a mechanism in which each buyer reports the types of all the other buyers. With this information acquisition technology, the seller can even extract rent when the underlying valuations are independent. On the other hand, if the information acquisition technology is such that each buyer independently receives a signal with probability $p < 1$, which perfectly reveals the types of the other buyers including whether they have received a signal, then the seller cannot always extract rent. For example, if the seller wishes to implement the efficient allocation rule, then a high type who learns that her opponents are low type and have no signal must earn positive rents.

Thus, the ability to acquire information may help or hurt the buyers in the presence of a seller who wishes to fully extract rent, depending on the information acquisition technology that is available, and so it is natural to ask which types of information acquisition technology can guarantee full rent extraction for the seller. In this paper, we will give an exact answer to this question by characterising the necessary and sufficient conditions on the information acquisition technology such that the seller is able to guarantee full rent extraction.

In our model, 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 both the valuation and the additional signal). After observing their valuations, the buyers are able to choose 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, which is represented by a function $\Sigma : V \times A \mapsto \Delta S$, where V is the set of valuations, A is the set of information acquisition actions, and S is the set of payoff irrelevant signals. V and S are assumed to be finite, but we make no restrictions on A.

We address the question of when the seller can fully extract rent using a mechanism design framework. 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 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 payoffs to each information acquisition strategy is the expectation of the maximum payoff from the mechanism, given the profile information acquisition strategies, and we require that the choice of information is a Nash equilibrium of this game. 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.

Without information acquisition, Crémer and McLean (1988) show that a necessary and sufficient condition for the seller to be able to fully extract rent for any allocation rule is that for each buyer i with valuation v_i, the belief of type v_i about v_{-i} does not lie in the convex hull of the beliefs of any type v_i', such that $v_i' \neq v_i$. When this condition holds, there exists a hyperplane that separates each type's belief from the beliefs of all other types. This is equivalent to the existence of lotteries, one for each type, with zero expected value for that type and strictly positive expected value for every other type. As we discussed previously, this implies that the condition is sufficient for full rent extraction.

Note that the condition depends only on the beliefs, and not on the valuations associated with each belief. This makes necessity less obvious. For example, if the only type whose belief lies in the convex hull of the beliefs of the other types also happens to have the highest valuation, then perhaps the seller can still fully extract rent even though the condition fails. Even though there does not exist a hyperplane separating the belief of the highest type from the beliefs of all other types, the seller does not need such a lottery for the highest type. In any mechanism where the highest type gets zero ex post payoff for every profile of her opponents' types, no type will deviate by pretending to be the highest type. Intuitively, 'side-bets' are required to prevent high types from deviating down, not low types deviating up.

In Crémer and McLean (1988), type v_i's valuation for the object is given by a function $w_i : V_i \mapsto R_+$, and the condition is necessary and sufficient to guarantee rent extraction for every possible w_i, which is the reason why it does not depend on the valuation associated with each belief. However, in our model, since the seller can use a different information acquisition strategy to extract rent for each specification of the payoffs, the conditions on beliefs will retain some dependence on the type that holds each belief.

We show that a necessary and sufficient condition on the information acquisition technology for the seller to be able to guarantee full rent extraction is that for every complete and transitive ordering on $\succeq_i$ on V_i, there exists an information acquisition strategy $\alpha : V \mapsto A$ such that for each type (v_i, s_i) that receives the object with positive probability in equilibrium and does not have the highest valuation under $\succeq$:

1. (v_i, s_i)'s belief is not in the relative interior of $C(\alpha_{-i})$, the convex hull of the beliefs of all types that could arise from any unilateral deviation to an alternative information acquisition strategy, given that the other players are following α_{-i}

2. The smallest exposed face of the closure of $C(\alpha_{-i})$ containing (v_i, s_i)'s belief does not intersect with the closure of the set of beliefs of all types with a strictly higher valuation than v_i under $\succeq_i$ that could arise from any unilateral deviation to an alternative information acquisition strategy, given that the other players are following α_{-i}.

The first condition ensures that for each type, there exists a lottery with zero expected value for that type, and weakly positive expected value for every type that could arise from any information acquisition strategy. The second condition ensures that the lottery can be chosen to have strictly positive expected value that is bounded away from zero for any type that has a strictly higher valuation than v_i under $\succeq_i$. Note that 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 another information acquisition strategy and is a much stronger requirement than in the standard model, since there can be (infinitely) many beliefs that arise from any information acquisition strategy. If this set of beliefs is open (for example, if it is always possible to acquire more information about each type of every opponent) then the seller will not be able to extract rent.

1.2 Related Literature

Crémer and McLean (1988) is the classic reference on full surplus extraction. Several papers have considered mechanism design with information acquisition. For example, Bikhchandani (2011) provides necessary and sufficient conditions for the existence of extraction lotteries that are robust to the possibility of information acquisition. 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 others' valuations, but also about each others' information.

A closely related paper is Obara (2008), which 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. In our setting, the distribution of the payoff relevant type is fixed, but the buyers can determine the distribution of payoff irrelevant signals after learning their payoff relevant type. Bikhchandani and Obara (2017) provides sufficient conditions for efficient implementation and full surplus extraction when buyers are able to acquire information about an unknown payoff relevant state of nature.

Yamashita (2018) considers a related question about the highest revenue the seller can guarantee when the buyers have access to additional information. The seller considers the worst case information structure for each choice of mechanism, and chooses the mechanism to maximise the worst case revenue. One interpretation is the buyers commit to some information structure, and then the seller offers the mechanism. A key difference in our approach is that the seller first commits to a mechanism, and then the buyers choose their information optimally. Thus, the seller is better off.

Finally Heifetz and Neeman (2006) 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 non-convexity 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.

1.3 Model

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 $v_i \in V_i$ and a payoff irrelevant signal $s_i \in S_i$, where V_i and S_i are finite. Let $\theta_i = (v_i, s_i) \in V_i \times S_i \equiv \Theta_i$ be buyer i's type. The profile of valuations $v \in V \equiv V_1 \times \ldots \times V_n$ is distributed according to Π, and buyer i's valuation for the object is given by a function $w_i : V_i \mapsto R_+$. We assume that utility is quasilinear in transfers. After observing v_i, each buyer chooses an information acquisition action $a_i \in A_i$, where A_i can be an arbitrary set. A function $\Sigma : A \times V \mapsto \Delta S$ determines a distribution over signals for each action profile a. Note that the distribution depends on the entire action profile a.

Let $\alpha_i : V_i \mapsto A_i$ be an information acquisition (pure) strategy for buyer i, and define $\Gamma : A^{|V|} \mapsto \Delta(V \times S)$ as $\gamma(v, s|\alpha) = \pi(v)\sigma(s|\alpha(v), v)$. That is, for any profile of information acquisition strategies α, the resulting distribution over $\Theta = V \times S$ is given by $\Gamma(\alpha)$. Let $\Gamma_{\Theta_i}(\alpha)$ denote the marginal distribution on Θ_i given α:

$$\begin{aligned}\gamma_{\Theta_i}(\theta_i|\alpha_i, \alpha_{-i}) &\equiv \sum_{\theta_{-i}} \gamma(\theta_i, \theta_{-i}|\alpha_i, \alpha_{-i}) \\ &= \sum_{v_{-i}} \sum_{s_{-i}} \pi(v_i, v_{-i})\sigma(s_i, s_{-i}|\alpha_i(v_i), \alpha_{-i}(v_{-i}), v_i, v_{-i}).\end{aligned}$$

Finally, let $\Theta_i(\alpha) = \{\theta_i : \gamma_{\Theta_i}(\theta_i|\alpha) > 0\}$ be the set of θ_i that arises with positive probability under α.

Remark 1.3.1. Our framework is general enough to 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 $\Sigma^f : A^f \times V \mapsto \Delta S^f$. Now suppose that each i may randomise among actions in A_i^f. This is equivalent to each i choosing a pure strategy from $A_i = \Delta A_i^f$, and receiving a signal from $S_i = S_i^f \times A_i^f$. Then $\Sigma : A \times V \mapsto \Delta S$ is given by $\sigma(s|a, v) = Pr(a^f|a)\sigma^f(s^f|a^f, v)$.

A direct mechanism (x, t) is an allocation rule $x : V \times S \mapsto \Delta I$ and a transfer rule $t : V \times S \mapsto R^n$. The timing is as follows:

1. Seller commits to a mechanism
2. Each buyer observes the mechanism and v_i
3. Each buyer chooses a_i
4. Each buyer observes s_i and reports $\theta_i = (v_i, s_i)$ to seller
5. 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. This is unlike the approach in the robust mechanism design literature where the seller considers the worst case information structure, which can be interpreted as the result of an optimal choice by buyers who commit to an information structure before the seller chooses the mechanism.

For $\theta_i \in \Theta_i(\alpha)$ 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:

$$U_i(\theta_i', \theta_i; \alpha) = \sum_{\theta_{-i}} \gamma(\theta_{-i}|\theta_i, \alpha) x_i(\theta_i', \theta_{-i}) w_i(v_i) - \sum_{\theta_{-i}} \gamma(\theta_{-i}|\theta_i, \alpha) t_i(\theta_i', \theta_{-i}).$$

Note that $\gamma(\theta_{-i}|\theta_i, \alpha) = \pi(v_{-i}|v_i)\sigma(s_{-i}|s_i, \alpha(v))$. 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 defined only for θ_i such that $\gamma_{\Theta_i}(\theta_i|\alpha) > 0$; hence our restriction to $\theta_i \in \Theta_i(\alpha)$. Finally, define $U_i^*(\theta_i; \alpha) = \max_{\theta_i'} U_i(\theta_i', \theta_i; \alpha)$.

Given a mechanism (x, t), *buyers optimally acquire information* α if α is a Nash equilibrium of the game given by payoffs:

$$\tilde{u}_i(\alpha_i, \alpha_{-i}) = \sum_{\theta_i \in \Theta_i(\alpha_i, \alpha_{-i})} \gamma_{\Theta_i}(\theta_i|\alpha_i, \alpha_{-i}) U_i^*(\theta_i; \alpha_i, \alpha_{-i}).$$

A mechanism is *incentive compatible* given α if for all $\theta_i \in \Theta_i(\alpha)$ and for all $\theta_i' \in \Theta_i$:

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

A mechanism is *individually rational* given α if for all $\theta_i \in \Theta_i(\alpha)$:

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

A mechanism *fully extracts rents* given α if for all $\theta_i \in \Theta_i(\alpha)$:

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

An information structure is $(\Theta, \Pi, (A, \Sigma))$, and for each valuation function $w = (w_1, \ldots, w_n)$, $(\Theta, \Pi, (A, \Sigma), w)$ is the associated allocation problem. A *social choice function* $f : V \mapsto \Delta I$ maps a profile of payoff relevant types to a probability distribution over the set of players.

Definition 1.3.1. The information structure $(\Theta, \Pi, (A, \Sigma))$ *guarantees full rent extraction* for the social choice function $f : V \mapsto \Delta I$ if for all allocation problems $(\Theta, \Pi, (A, \Sigma)), w)$, there exists a mechanism (x, t) and an information acquisition strategy α such that:

- (x, t) is incentive compatible, individually rational, and fully extracts rent given α
- Buyers optimally acquire information α given (x, t)
- $x(v, s) = f(v)$ for all v, s such that $\gamma(v, s|\alpha) > 0$.

Definition 1.3.1 of rent extraction requires that for a given social choice function $f : V \mapsto \Delta I$, the seller is able to fully extract rent for all possible valuation functions w. This approach is standard in the literature and simplifies the conditions on beliefs. Definition 1.3.1 is justified by the revelation principle, which we now prove for our environment.

Let (x', t') be an arbitrary mechanism, where $x' : M \mapsto \Delta I$ and $t' : M \mapsto R^n$, and for each i, let $\mu_i : V_i \times S_i \times A_i \mapsto M_i$ be a strategy in the mechanism. Following the information acquisition strategy α, for $\theta_i \in \Theta_i(\alpha)$, let $U_i'(m_i', \theta_i; \alpha, \mu_{-i})$ be i's payoff when Player i reports m_i', Player i is type θ_i, and the other players are following the strategy μ_{-i} in the mechanism:

$$\begin{aligned} U_i'(m_i', \theta_i; \alpha, \mu_{-i}) = &\sum_{\theta_{-i}} \gamma(\theta_{-i}|\theta_i, \alpha) x_i(m_i', \mu_{-i}(\theta_{-i}, \alpha_{-i}(v_{-i}))) w_i(v_i) \\ &- \sum_{\theta_{-i}} \gamma(\theta_{-i}|\theta_i, \alpha) t_i(m_i', \mu_{-i}(\theta_{-i}, \alpha_{-i}(v_{-i}))). \end{aligned}$$

Define $m^{**}(\theta_i, \alpha, \mu_{-i}) = \arg\max_{m_i'} U_i'(m_i', \theta_i; \alpha, \mu_{-i})$, and $U_i^{**}(\theta_i; \alpha, \mu_{-i}) = U_i'(m_i, \theta_i; \alpha, \mu_{-i})$ for $m_i \in m^{**}(\theta_i, \alpha, \mu_{-i})$.

Proposition 1.3.1. *Suppose that there exists an information acquisition strategy α, a mechanism (x', t'), and for each player i, a strategy μ_i in the mechanism such that for all $\alpha_i' \in A_i^{|V_i|}$:*

1. $\mu_i(\theta_i, \alpha_i(v_i)) \in m^{**}(\theta_i, \alpha, \mu_{-i})$ *for all* $\theta_i \in \Theta_i(\alpha)$
2. $\sum_{\theta_i \in \Theta_i(\alpha)} \gamma_{\Theta_i}(\theta_i|\alpha_i, \alpha_{-i}) U_i^{**}(\theta_i; \alpha, \mu_{-i}) \geq \sum_{\theta_i \in \Theta_i(\alpha_i', \alpha_{-i})} \gamma_{\Theta_i}(\theta_i|\alpha_i', \alpha_{-i}) U_i^{**}(\theta_i; \alpha_i', \alpha_{-i}, \mu_{-i})$.

Then there exists a direct mechanism (x, t) such that (x, t) is incentive compatible given α, buyers optimally acquire information α given (x, t), and $(x(\theta), t(\theta)) = (x'(\mu(\theta, \alpha(v))), t'(\mu(\theta, \alpha(v))))$ for all $\theta \in \Theta$.

Proof. Appendix. □

Remark 1.3.2. Conditions 1 and 2 in Proposition 1.3.1 are equivalent to (α, μ) being a Nash equilibrium of the game $G' = (A^{|V|} \times M^{|\Theta \times A|}, u')$, where

$$u'(\alpha_i, \mu_i; \alpha_{-i}, \mu_{-i}) = \sum_{\theta_i \in \Theta_i(\alpha_i, \alpha_{-i})} \gamma_{\Theta_i}(\theta_i|\alpha_i, \alpha_{-i}) U_i'(\mu_i(\theta_i, \alpha_i(v_i)), \theta_i; \alpha_i, \alpha_{-i}, \mu_{-i}).$$

1.4 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 examples.

Example 1.4.1 (Crémer-McLean)**.** Suppose $S = \emptyset$, so there is no possibility of information acquisition. Then $((V, \emptyset), \Pi, (A, \Sigma))$ guarantees full rent extraction for all social choice functions $f : V \mapsto \Delta I$ if and only if for all v_i:

$$\pi(V_{-i}|v_i) \notin co\{\pi(V_{-i}|v_i') : v_i' \neq v_i\}.^1$$

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

Example 1.4.2. Let $n = 2$, $V_i = \{v^L, v^H\}$, $S_i = \{\emptyset, (v^L, \emptyset), (v^L, s), (v^H, \emptyset), (v^H, s)\}$, $w_i(v^L) = \underline{v} > 0$, $w_i(v^H) = 1$, $\Pi > 0$ (Π can be any full support distribution), $A_i = \{N, Y\}$. If $a_i = N$, then player i receives the null signal. 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 $(v^L, (v^L, \emptyset))$ and $(v^H, (v^L, \emptyset))$ 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 $(v^H, (v^L, \emptyset))$ of player i must receive positive rents. Let $x_i((v^L, (v^L, \emptyset)), (v^L, \emptyset)) \equiv x > 0$. Since type $(v^L, (v^L, \emptyset))$ of player i knows the type of player $-i$ for sure, individual rationality implies $t_i((v^L, (v^L, \emptyset)), (v^L, \emptyset)) \leq x\underline{v}$. Now a possible deviation is for player i to choose $\alpha_i(v^H) = Y$, drop out for all types other than $(v^H, (v^L, \emptyset))$, and for type $(v^H, (v^L, \emptyset))$, report $(v^L, (v^L, \emptyset))$. The utility of this deviation is at least $\pi(v^H)\pi(v^L|v^H)p(1-p)x(1-\underline{v}) > 0$.

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

Example 1.4.3. 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 $\sigma(s_i = v_{-i}, s_{-i} = v_i | v_i, v_{-i}, (Y, Y)) = 1$, $\sigma(s_i = v_{-i}, s_{-i} = v_i | v_i, v_{-i}, (N, Y)) = \frac{1}{2}$, $\sigma(s_i \neq v_{-i}, s_{-i} = v_i | v_i, v_{-i}, (N, Y)) = \frac{1}{2}$. For any x and for any w, the following mechanism is incentive compatible, individually rational, and fully extracts rent given $\alpha_i(v_i) = Y$ for all v_i, and $\alpha_i(v_i) = Y$ for all v_i is optimal for both players:

- $t_i((v_i, s_i), (v_{-i}, s_{-i})) = x_i((v_i, s_i), (v_{-i}, s_{-i}))w_i(v_i)$ if $v_i = s_{-i}$ and $s_i = v_{-i}$
- $t_i((v_i, s_i), (v_{-i}, s_{-i})) = \infty$ otherwise.

1.5 Main Result

In this section, we will characterise the necessary and sufficient conditions for the information structure $(\Theta, \Pi, (A, \Sigma))$ to guarantee full rent extraction.

Let $\gamma(\Theta_{-i}|\theta_i, \alpha) \equiv \left(\gamma(\theta_{-i}^1|\theta_i, \alpha), \gamma(\theta_{-i}^2|\theta_i, \alpha), \ldots, \gamma(\theta_{-i}^{|\Theta_{-i}|}|\theta_i, \alpha)\right)$ be the belief of type θ_i about Θ_{-i}, given the profile of information acquisition strategies α. Let $\succeq_i$ be complete and transitive

[1] $\pi(V_{-i}|v_i) = (\pi(v_{-i}^1|v_i), \ldots, \pi(v_{-i}^{|V_{-i}|}|v_i))$.

binary relation over V_i, and let $\mathcal{R}_i$ be the set of all complete and transitive binary relations over V_i. Let $\succeq = (\succeq_1, \ldots, \succeq_n)$, and $\mathcal{R} = \mathcal{R}_1 \times \ldots \times \mathcal{R}_n$. Define:

- $C(\alpha_{-i}) = co\{\gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) : \alpha_i' \in A_i^{|V_i|}, \theta_i' \in \Theta_i(\alpha_i', \alpha_{-i})\}$
- $D(\theta_i, \alpha_{-i}, \succeq_i) = \{\gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) : \alpha_i' \in A_i^{|V_i|}, \theta_i' \in \Theta_i(\alpha_i', \alpha_{-i}), v_i' \succ_i v_i\}$
- $\Theta_i^+(\alpha, f) = \{\theta_i \in \Theta_i(\alpha) : \sum_{v_{-i}} \sum_{s_{-i}} \gamma(v_{-i}, s_{-i}|\theta_i, \alpha) f_i(v_i, v_{-i}) > 0\}$
- $v_i^*(\succeq_i) = \{v_i : v_i \succeq_i v_i' \text{ for all } v_i' \in V_i\}$.

$C(\alpha_{-i})$ is the convex hull of the beliefs of every type of buyer i that could arise from any information acquisition strategy of buyer i, fixing the information acquisition strategy of the other buyers at α_{-i}. $D(\theta_i, \alpha_{-i}, \succeq_i)$ is the set of beliefs of every type of buyer i with a valuation strictly greater than v_i under $\succeq_i$ that could arise from any information acquisition strategy of buyer i, fixing the information acquisition strategy of the other buyers at α_{-i}. Note that $D(\theta_i, \alpha_{-i}, \succeq_i) \subset C(\alpha_{-i})$. If A_i is finite, then $C(\alpha_{-i})$ is a polytope. If A_i is compact and σ is continuous, then $C(\alpha_{-i})$ and $D(\theta_i, \alpha_{-i}, \succeq_i)$ are compact. $\Theta_i^+(\alpha, f)$ is the set of types of buyer i that that receives the object with positive probability. $v_i^*(\succeq_i)$ is the set containing the greatest elements of V_i under the relation $\succeq_i$.

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

Definition 1.5.1. For any convex set $K \subset R^n$, 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. $F = \emptyset$.

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

Definition 1.5.2. For any convex set $K \subset R^n$ and $\bar{x} \in K$, let $F_K(\bar{x})$ be the intersection of all exposed faces of K containing $\bar{x}$.

Proposition 1.5.1. *The information structure* $(\Theta, \Pi, (A, \Sigma))$ *guarantees full rent extraction for the social choice function* $f : V \mapsto \Delta I$ *if and only if for every* $\succeq \in \mathcal{R}$, *there exists an* α *such that for all* i *and for all* $\theta_i \in \Theta_i^+(\alpha, f) \setminus (v_i^*(\succeq_i) \times S_i)$:

1. $\gamma(\Theta_{-i}|\theta_i, \alpha) \cap \text{ri}(C(\alpha_{-i})) = \emptyset$
2. $F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\theta_i, \alpha)) \cap \overline{D}(\theta_i, \alpha_{-i}, \succeq_i) = \emptyset$.[2]

Proof. Appendix. □

[2]If we assume that A_i is compact and σ is continuous, then $C(\alpha_{-i})$ and $D(\theta_i, \alpha_{-i}, \succeq_i)$ are closed, so we can drop the closure from the conditions. When Condition 1 is satisfied, there is always a hyperplane properly supporting $\overline{C}(\alpha_{-i})$ at $\gamma(\Theta_{-i}|\theta_i, \alpha)$, so $F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\theta_i, \alpha))$ is a proper exposed face of $\overline{C}(\alpha_{-i})$. On the other hand when Condition 1 fails, $F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\theta_i, \alpha)) = \overline{C}(\alpha_{-i})$, and Condition 2 cannot hold.

Condition 1 requires that θ_i's belief does not lie in the relative interior of $C(\alpha_{-i})$, which is the convex hull of the beliefs all of types of buyer i that could arise from any information acquisition strategy of buyer i, fixing the information acquisition strategy of the other buyers at α_{-i}. This is equivalent to the existence of a hyperplane properly supporting $C(\alpha_{-i})$ at $\gamma(\Theta_{-i}|\theta_i, \alpha)$. This requirement is stronger than the analogous condition in the standard model to the extent that there can be (possibly infinitely) many beliefs that arise from any information acquisition strategy, even when there are a small number of types that arise in equilibrium.

Condition 2 requires that the smallest exposed face of the closure of $C(\alpha_{-i})$ that contains $\gamma(\Theta_{-i}|\theta_i, \alpha)$ does not intersect with the closure of the set of beliefs of all types of buyer i with a valuation strictly greater than v_i (i.e. θ_i's valuation) under $\succeq_i$ that could arise from any information acquisition strategy of buyer i, fixing the information acquisition strategy of the other buyers at α_{-i}. By definition, there is a hyperplane that supports $C(\alpha_{-i})$ at $\gamma(\Theta_{-i}|\theta_i, \alpha)$ and intersects $C(\alpha_{-i})$ at exactly this exposed face. Condition 2 is equivalent to this hyperplane strongly separating $\gamma(\Theta_{-i}|\theta_i, \alpha)$ from any belief in $C(\alpha_{-i})$ with a valuation strictly greater than v_i under $\succeq_i$.

For each $\succeq$, we require the existence of an α such that the conditions hold for every type $\theta_i = (v_i, s_i)$ such that s_i arises with positive probability in equilibrium, v_i receives the object with positive probability according to f, and v_i is not the greatest element of V_i under $\succeq_i$. If we want to guarantee full rent extraction for every f, then we can drop the requirement that v_i receives the object with positive probability according to a particular f.

Corollary 1.5.2. *The information structure $(\Theta, \Pi, (A, \Sigma))$ guarantees full rent extraction for every social choice function $f : V \mapsto \Delta I$ if and only if for every $\succeq \in \mathcal{R}$ there exists an α such that for all i and for all $\theta_i \in \Theta_i(\alpha) \setminus (v_i^*(\succeq_i) \times S_i)$:*

1. $\gamma(\Theta_{-i}|\theta_i, \alpha) \cap \text{ri}(C(\alpha_{-i})) = \emptyset$
2. $F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\theta_i, \alpha)) \cap \overline{D}(\theta_i, \alpha_{-i}, \succeq_i) = \emptyset.$

Remark 1.5.1. In the standard case without information acquisition $S = \emptyset$, $\gamma(\Theta_{-i}|\theta_i, \alpha) = \pi(V_{-i}|v_i)$, and hence Corollary 1.5.2 requires:

1. For all $\succeq \in \mathcal{R}$ and for all $v_i \in V_i \setminus v_i^*(\succeq_i)$, $\pi(V_{-i}|v_i) \cap \text{ri}(co\{\pi(V_{-i}|v_i') : v_i' \in V_i\}) = \emptyset$
2. For all $\succeq \in \mathcal{R}$ and for all $v_i \in V_i \setminus v_i^*(\succeq_i)$, $F_{co\{\pi(V_{-i}|v_i'):v_i' \in V_i\}}(\pi(V_{-i}|v_i)) \cap \{\pi(V_{-i}|v_i') : v_i' \succ_i v_i\} = \emptyset$

Note that this is equivalent to $\pi(V_{-i}|v_i) \cap \text{ri}(co\{\pi(V_{-i}|v_i') : v_i' \in V_i\}) = \emptyset$ for all $v_i \in V_i$ and $F_{co\{\pi(V_{-i}|v_i'):v_i' \in V_i\}}(\pi(V_{-i}|v_i)) = \{\pi(V_{-i}|v_i)\}$ for all $v_i \in V_i$. Together they are equivalent to $\pi(V_{-i}|v_i) \cap co\{\pi(V_{-i}|v_i') : v_i' \neq v_i\} = \emptyset$, as we have stated in Example 1.4.1.

Remark 1.5.2. In Definition 1.3.1, the requirement that rent extraction is possible for all w simplifies the conditions on beliefs in the standard model. In particular, as the last remark shows, 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 perturbation 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 perturbation of the valuations.

Proof Intuition

The intuition behind sufficiency is as follows. Condition 1 implies that for each $\theta_i \in \Theta_i^+(\alpha, f) \setminus (v_i^*(\succeq_i) \times S_i)$, there is a hyperplane $\{\gamma : \tau \cdot \gamma = \tau \cdot \gamma(\Theta_{-i}|\theta_i, \alpha)\}$ supporting $C(\alpha_{-i})$ at $\gamma(\Theta_{-i}|\theta_i, \alpha)$. That is, there exists a lottery $\tau(\theta_i)$ that has weakly positive expected value for all types θ_i' arising from any information acquisition strategy. Condition 2 implies that for all θ_i' such that $w(v_i') > w(v_i)$, there exists an $\varepsilon > 0$ such that the expected value of $\tau(\theta_i)$ for type θ_i' is strictly larger than ε—that is, the hyperplane $\tau(\theta_i)$ strongly separates the belief of θ_i from the beliefs of types θ_i' such that $w(v_i') > w(v_i)$.

Now for each $\theta_i \notin \Theta_i^+(\alpha, f)$, we can let $t_i(\theta_i, \Theta_{-i}) = 0$,[3] in which case $U_i(\theta_i, \theta_i', \alpha_i', \alpha_{-i}) = 0$ for all $\alpha_i' \in A_i^{|V_i|}$, $\theta_i' \in \Theta_i(\alpha_i', \alpha_{-i})$. This is saying that no type has an incentive to deviate by pretending to be a type that does not receive the object. For $\theta_i \in (v_i^*(\succeq_i) \times S_i) \cap \Theta_i^+(\alpha, f)$, we can let $t_i(\theta_i, \Theta_{-i}) = w_i(v_i)x_i(\theta_i, \Theta_{-i})$.[4] Intuitively, if types with the highest valuation pay their valuation for the object whenever they receive the object, they receive no rents, and no other type would want to pretend to have the highest valuation as these other types value the object even less.

For types that do receive the object in equilibrium and do not have the highest valuation, i.e. for $\theta_i \in \Theta_i^+(\alpha, f) \setminus (v_i^*(\succeq_i) \times S_i)$, we can add a scaled up version of the lottery $\tau(\theta_i)$ to the payment. So let $t_i(\theta_i, \Theta_{-i}) = w_i(v_i)x_i(\theta_i, \Theta_{-i}) + M\tau(\theta_i)$ for sufficiently large M. Now types θ_i' such that $w_i(v_i') > w_i(v_i)$ following α_i' will not want to pretend to be type θ_i, since the expected payment includes $M\tau(\theta_i) \cdot \gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i})$, which can be made arbitrarily large. For types θ_i' such that $w_i(v_i') \leq w_i(v_i)$, it is possible that $\tau(\theta_i) \cdot \gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) = 0$, but in that case $U_i(\theta_i, \theta_i', \alpha_i', \alpha_{-i}) \leq U_i(\theta_i, \theta_i, \alpha) = 0$.

For necessity, first note that for any θ_i, it is without loss of generality to let $t_i(\theta_i, \Theta_{-i}) = w_i(v_i)x_i(\theta_i, \Theta_{-i}) + M\tau(\theta_i)$ for some $\tau(\theta_i)$. For the mechanism to fully extract rent, we need $\tau(\theta_i) \cdot \gamma(\Theta_{-i}|\theta_i, \alpha) = 0$. If Condition 1 fails, then there exists a type $\theta_i \in \Theta_i^+(\alpha, f) \setminus (v_i^*(\succeq_i) \times S_i)$ such that for any $\tau(\theta_i)$, either $\gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) \cdot \tau(\theta_i) = 0$ for all $\alpha_i' \in A_i^{|V_i|}$, $\theta_i' \in \Theta_i(\alpha_i', \alpha_{-i})$, or there exists $\alpha_i' \in A_i^{|V_i|}$, $\theta_i' \in \Theta_i(\alpha_i', \alpha_{-i})$ such that $\gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) \cdot \tau(\theta_i) < 0$. But incentive compatibility requires that $\gamma(\Theta_{-i}|\theta_i^*, \alpha) \cdot \tau(\theta_i) > 0$ for $\theta_i^* \in v_i^*(\succeq_i) \times S_i$, and for w such that $\max_{v_i'} w_i(v_i')$ is sufficiently large, M must also be large. This means that there must exist $\alpha_i' \in A_i^{|V_i|}$, $\theta_i' \in \Theta_i(\alpha_i', \alpha_{-i})$ such that $\gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) \cdot \tau(\theta_i) < 0$. Then the utility from the following deviation is strictly positive: choose α_i', drop out for all types other than θ_i', and report θ_i when type θ_i'.

If Condition 2 fails, then there exists a type $\theta_i \in \Theta_i^+(\alpha, f) \setminus (v_i^*(\succeq_i) \times S_i)$ such that for any $\tau(\theta_i)$ that supports $C_i(\alpha_{-i})$ at $\gamma(\Theta_{-i}|\theta_i, \alpha)$, for every $\varepsilon > 0$, there exists an information acquisition strategy $\alpha_i' \in A_i^{|V_i|}$ such that for some type $\theta_i' \in \Theta_i(\alpha_i', \alpha_{-i})$ with $w(v_i') > w_i(v_i)$, the expected value of $\tau(\theta_i)$ after having chosen α_i' is less than ε. Then a profitable deviation is to choose α_i', drop out for all types other than θ_i', and report θ_i when type θ_i'.

Simple Conditions

We now give simpler sufficient conditions and necessary conditions that depend only on the beliefs, and not the valuations of the types holding each belief. Define:

- $D^*(\theta_i, \alpha_{-i}) = \cup_{\succeq_i \in \mathcal{R}_i} D(\theta_i, \alpha_{-i}, \succeq_i) = \{\gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) : \alpha_i' \in A_i^{|V_i|}, \theta_i' \in \Theta_i(\alpha_i', \alpha_{-i}), v_i' \neq v_i\}$
- $B(v_i, \alpha_{-i}) = \{\gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) : \alpha_i' \in A_i^{|V_i|}, \theta_i' \in \Theta_i(\alpha_i', \alpha_{-i}), v_i' = v_i\}$.

$D^*(\theta_i, \alpha_{-i})$ is the set of beliefs of every type of buyer i with a valuation different from v_i that could arise from any information acquisition strategy of buyer i, fixing the information acquisition strategy of the other buyers at α_{-i}. $B(v_i, \alpha_{-i})$ is the set of beliefs of all types of buyer i with valuation

[3] $t_i(\theta_i, \Theta_{-i}) = (t_i(\theta_i, \theta_{-i}^1), \ldots, t_i(\theta_i, \theta_{-i}^{|\Theta_{-i}|}))$.

[4] $x_i(\theta_i, \Theta_{-i}) = (x_i(\theta_i, \theta_{-i}^1), \ldots, x_i(\theta_i, \theta_{-i}^{|\Theta_{-i}|}))$.

v_i that could arise from any information acquisition strategy of buyer i, fixing the information acquisition strategy of the other buyers at α_{-i}.

Corollary 1.5.3 (Sufficient Condition)**.** *The information structure* $(\Theta, \Pi, (A, \Sigma))$ *guarantees full rent extraction for every social choice function* $f : V \mapsto \Delta I$ *if there exists an* α *such that for all* i*:*

1. *For all* $\theta_i \in \Theta_i(\alpha)$*,* $\gamma(\Theta_{-i}|\theta_i, \alpha) \cap \mathrm{ri}(C(\alpha_{-i})) = \emptyset$
2. *For all* $\theta_i \in \Theta_i(\alpha)$*,* $F_{\overline{C(\alpha_{-i})}}(\gamma(\Theta_{-i}|\theta_i, \alpha)) \cap \overline{D^*}(\theta_i, \alpha_{-i}) = \emptyset$.

Corollary 1.5.4 (Necessary Condition)**.** *The information structure* $(\Theta, \Pi, (A, \Sigma))$ *guarantees full rent extraction for every social choice function* $f : V \mapsto \Delta I$ *only if there exists an* α *such that for all* i *and for some* v_i'*:*

1. *For all* $\theta_i \in \Theta_i(\alpha) \setminus (\{v_i'\} \times S_i)$*,* $\gamma(\Theta_{-i}|\theta_i, \alpha) \cap \mathrm{ri}(C(\alpha_{-i})) = \emptyset$
2. *For all* $\theta_i \in \Theta_i(\alpha) \setminus (\{v_i'\} \times S_i)$*,* $F_{\overline{C(\alpha_{-i})}}(\gamma(\Theta_{-i}|\theta_i, \alpha)) \cap \overline{B}(v_i', \alpha_{-i}) = \emptyset$.

1.6 Surplus Extraction

A closely related question is when the seller can fully extract surplus.

Definition 1.6.1. The information structure $(\Theta, \Pi, (A, \Sigma))$ *guarantees full surplus extraction* if for all allocation problems $((\Theta, \Pi, (A, \Sigma)), w)$, there exists a mechanism (x, t) and an information acquisition strategy α such that:

- (x, t) is incentive compatible, individually rational, and fully extracts rent given α
- Buyers optimally acquire information given (x, t)
- $x(v, s) = f^*(v)$ for all v, s such that $\gamma(v, s|\alpha) > 0$, where $f^*(v) \in \arg\max_{f \in \Delta I} \sum_{i \in I} f_i w_i(v_i)$.

Clearly, if $(\Theta, \Pi, (A, \Sigma))$ guarantees full rent extraction for every social choice function $f : V \mapsto \Delta I$, then $(\Theta, \Pi, (A, \Sigma))$ guarantees full surplus extraction. However, the converse is not true. The reason is that there is always a buyer such that the lowest type of that buyer values the object at least as much as the lowest type of any other buyer. Then for every other buyer, the seller does not need to allocate to the lowest types. Thus, the conditions are required to hold for a smaller set of types. Let $v_{*i}(\succeq_i) = \{v_i : v_i' \succeq v_i \text{ for all } v_i' \in V_i\}$ be the set containing the smallest elements of V_i under $\succeq_i$. Then:

Proposition 1.6.1. *The information structure* $(\Theta, \Pi, (A, \Sigma))$ *guarantees full surplus extraction if and only if for every* $\succeq \in \mathcal{R}$ *and for every* $i \in I$*, there exists an* α *such that:*

For all $\theta_i \in \Theta_i(\alpha) \setminus (v_i^*(\succeq_i) \times S_i)$*:*

1. $\gamma(\Theta_{-i}|\theta_i, \alpha) \cap \mathrm{ri}(C(\alpha_{-i})) = \emptyset$
2. $F_{\overline{C(\alpha_{-i})}}(\gamma(\Theta_{-i}|\theta_i, \alpha)) \cap \overline{D}(\theta_i, \alpha_{-i}, \succeq_i) = \emptyset$

For all $j \neq i$*, and for all* $\theta_j \in \Theta_j(\alpha) \setminus ((v_j^*(\succeq_j) \cup v_{*j}(\succeq_j)) \times S_j)$*:*

3. $\gamma(\Theta_{-j}|\theta_j, \alpha) \cap \mathrm{ri}(C(\alpha_{-j})) = \emptyset$
4. $F_{\overline{C(\alpha_{-j})}}(\gamma(\Theta_{-j}|\theta_j, \alpha)) \cap \overline{D}(\theta_j, \alpha_{-j}, \succeq_j) = \emptyset$.

Proof. Appendix. □

1.7 Conclusion

In this paper, we have given a full characterisation of information structures that guarantee full rent extraction when the buyers are able to acquire additional information about each other. The standard result in mechanism design with correlated information is that the seller can fully extract surplus. In other words, buyers do not earn any rents from their private information. With information acquisition, buyers may earn positive rents. For example, a buyer who can always learn more about each type of her opponents must earn positive rents whenever the seller wishes to implement any allocation rule in which she receives the object with positive probability.

In our model, the additional signals do not contain any socially useful information since the buyers already know their private valuations. We assume that information acquisition is costless; otherwise, any information acquisition strategy that results in costly information acquisition would be inefficient. It may be interesting to incorporate costly actions into our model. For example, full surplus extraction would then require the existence of an information acquisition strategy which does not involve playing any costly actions such that the beliefs that arise from that strategy do not lie in the convex hull of the beliefs that could arise from any information acquisition strategy, including those with costly actions. That is, full surplus extraction would become even harder for the seller.

Appendix 1.A Omitted Proofs

Proof of Proposition 1.3.1. Assume that Condition 1 and Condition 2 of Proposition 1.3.1 are satisfied for α, (x', t') and μ. Now define the direct mechanism (x, t) as $x(\theta) = x'(\mu(\theta, \alpha(v)))$ and $t(\theta) = t'(\mu(\theta, \alpha(v)))$. First, note that (x, t) is incentive compatible given α since:

$$
\begin{aligned}
U_i(\theta_i', \theta_i; \alpha) &= \sum_{\theta_{-i}} \gamma(\theta_{-i}|\theta_i, \alpha) x_i(\theta_i', \theta_{-i}) w_i(v_i) - \sum_{\theta_{-i}} \gamma(\theta_{-i}|\theta_i, \alpha) t_i(\theta_i', \theta_{-i}) \\
&= \sum_{\theta_{-i}} \gamma(\theta_{-i}|\theta_i, \alpha) x_i'(\mu_i(\theta_i', \alpha_i(v_i'))), \mu_{-i}(\theta_{-i}, \alpha_{-i}(v_{-i}))) w_i(v_i) \\
&\qquad - \sum_{\theta_{-i}} \gamma(\theta_{-i}|\theta_i, \alpha) t_i'(\mu_i(\theta_i', \alpha_i(v_i')), \mu_{-i}(\theta_{-i}, \alpha_{-i}(v_{-i}))) \\
&= U_i'(\mu_i(\theta_i', \alpha_i(v_i')), \theta_i; \alpha, \mu_{-i}) \\
&\leq U_i'(\mu_i(\theta_i, \alpha_i(v_i)), \theta_i; \alpha, \mu_{-i}) \\
&= \sum_{\theta_{-i}} \gamma(\theta_{-i}|\theta_i, \alpha) x_i'(\mu_i(\theta_i, \alpha_i(v_i)), \mu_{-i}(\theta_{-i}, \alpha_{-i}(v_{-i}))) w_i(v_i) \\
&\qquad - \sum_{\theta_{-i}} \gamma(\theta_{-i}|\theta_i, \alpha) t_i'(\mu_i(\theta_i, \alpha_i(v_i)), \mu_{-i}(\theta_{-i}, \alpha_{-i}(v_{-i}))) \\
&= \sum_{\theta_{-i}} \gamma(\theta_{-i}|\theta_i, \alpha) x_i(\theta_i, \theta_{-i}) w_i(v_i) - \sum_{\theta_{-i}} \gamma(\theta_{-i}|\theta_i, \alpha) t_i(\theta_i, \theta_{-i}) \\
&= U_i(\theta_i, \theta_i; \alpha).
\end{aligned}
$$

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

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

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

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

Thus:

$$
\begin{aligned}
\tilde{u}_i(\alpha_i,\alpha_{-i}) &= \sum_{\theta_i\in\Theta_i(\alpha)} \gamma_{\Theta_i}(\theta_i|\alpha_i,\alpha_{-i})U_i^*(\theta_i;\alpha_i,\alpha_{-i}) \\
&= \sum_{\theta_i\in\Theta_i(\alpha)} \gamma_{\Theta_i}(\theta_i|\alpha_i,\alpha_{-i})U_i^{**}(\theta_i;\alpha,\mu_{-i}) \\
&\ge \sum_{\theta_i\in\Theta_i(\alpha_i',\alpha_{-i})} \gamma_{\Theta_i}(\theta_i|\alpha_i',\alpha_{-i})U_i^{**}(\theta_i;\alpha_i',\alpha_{-i},\mu_{-i}) \\
&\ge \sum_{\theta_i\in\Theta_i(\alpha_i',\alpha_{-i})} \gamma_{\Theta_i}(\theta_i|\alpha_i',\alpha_{-i})U_i^*(\theta_i;\alpha_i',\alpha_{-i}) \\
&= \tilde{u}_i(\alpha_i',\alpha_{-i}).
\end{aligned}
$$

□

Proof of Proposition 1.5.1. For sufficiency, take $f : V \mapsto [0,1]^N$, and for any w, take $\succeq$ such that for all i, $v_i \succeq_i v_i'$ if and only if $w_i(v_i) \ge w_i(v_i')$. Then let information acquisition strategy α be such that:

1. For all $\theta_i \in \Theta_i^+(\alpha,f) \setminus (v_i^*(\succeq_i) \times S_i)$, $\gamma(\Theta_{-i}|\theta_i,\alpha) \cap \text{ri}(C(\alpha_{-i})) = \emptyset$
2. For all $\theta_i \in \Theta_i^+(\alpha,f) \setminus (v_i^*(\succeq_i) \times S_i)$, $F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\theta_i,\alpha)) \cap \overline{D}(\theta_i,\alpha_{-i},\succeq_i) = \emptyset$.

Let $x(\theta) = f(v)$ for any θ that arises with positive probability under α, and let $x(\theta) = 0$ otherwise. Now we argue that there exists transfers such that (x,t) is incentive compatible, individually rational, and fully extracts rent given α, and α is optimally chosen given (x,t).

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

$$\{\gamma : \tau \cdot \gamma = \tau \cdot \gamma(\Theta_{-i}|\theta_i,\alpha)\}$$

which properly supports $\overline{C}(\alpha_{-i})$ at $\gamma(\Theta_{-i}|\theta_i,\alpha)$. Thus, $F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\theta_i,\alpha))$ is a proper exposed face of $\overline{C}(\alpha_{-i})$. Let $\tau(\theta_i)$ be a lottery such that $F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\theta_i,\alpha)) = \{\gamma : \tau(\theta_i)\cdot\gamma = 0\} \cap \overline{C}(\alpha_{-i})$. This implies that $\tau(\theta_i)\cdot\gamma = 0$ for all $\gamma \in F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\theta_i,\alpha))$. Without loss of generality, let $\tau(\theta_i)\cdot\gamma > 0$ for all $\gamma \in \overline{C}(\alpha_{-i}) \setminus F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\theta_i,\alpha))$.

For $\theta_i \in \Theta_i^+(\alpha,f) \setminus (v_i^*(\succeq_i) \times S_i)$ let $t_i(\theta_i,\Theta_{-i}) = w_i(v_i)x_i(\theta_i,\Theta_{-i}) + M\tau(\theta_i)$,[5] where M is sufficiently large. Note that $\gamma(\Theta_{-i}|\theta_i,\alpha)\cdot\tau(\theta_i) = 0$. For $\theta_i \notin \Theta_i^+(\alpha,f)$, let $t_i(\theta_i,\Theta_{-i}) = 0$. For

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

$\theta_i \in (v_i^*(\succeq_i) \times S_i) \cap \Theta_i^+(\alpha, f)$, let $t_i(\theta_i, \Theta_{-i}) = w_i(v_i)x_i(\theta_i, \Theta_{-i})$. By construction, (x, t) fully extracts rent given α. We now show that (x, t) is incentive compatible given α, and α is chosen optimally.

First we show that for any $\theta_i \in \Theta_i(\alpha)$, $\theta_i' \in \Theta_i$, $U_i(\theta_i', \theta_i; \alpha) \leq U_i(\theta_i, \theta_i; \alpha)$. We will consider the cases where $\theta_i' \in \Theta_i^+(\alpha, f) \setminus (v_i^*(\succeq_i) \times S_i)$, $\theta_i' \notin \Theta_i^+(\alpha, f)$, and $\theta_i' \in (v_i^*(\succeq_i) \times S_i) \cap \Theta_i^+(\alpha, f)$ separately. Note that for $\theta_i' \in \Theta_i^+(\alpha, f) \setminus (v_i^*(\succeq_i) \times S_i)$:

$$\begin{aligned} U_i(\theta_i', \theta_i; \alpha) &= w_i(v_i)\gamma(\Theta_{-i}|\theta_i, \alpha) \cdot x_i(\theta_i', \Theta_{-i}) - \gamma(\Theta_{-i}|\theta_i, \alpha) \cdot (w_i(v_i')x_i(\theta_i', \Theta_{-i}) + M\tau(\theta_i')) \\ &= (w_i(v_i) - w_i(v_i'))\gamma(\Theta_{-i}|\theta_i, \alpha) \cdot x_i(\theta_i', \Theta_{-i}) - M\gamma(\Theta_{-i}|\theta_i, \alpha) \cdot \tau(\theta_i') \\ &\leq 0. \end{aligned}$$

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

For $\theta_i' \notin \Theta_i^+(\alpha, f)$, $U_i(\theta_i', \theta_i; \alpha) = 0$ since $x_i(\theta_i', \Theta_{-i}) = t_i(\theta_i', \Theta_{-i}) = 0$. For $\theta_i' \in v_i^*(\succeq_i) \times S_i$, note that $w_i(v_i') \geq w_i(v_i)$. Then:

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

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

$$\sum_{\theta_i \in \Theta_i(\alpha_i', \alpha_{-i})} \gamma_{\Theta_i}(\theta_i|\alpha_i', \alpha_{-i})U_i^*(\theta_i; \alpha_i', \alpha_{-i}) \leq \sum_{\theta_i \in \Theta_i(\alpha)} \gamma_{\Theta_i}(\theta_i|\alpha)U_i^*(\theta_i; \alpha).$$

Note that for all $\theta_i \in \Theta_i(\alpha_i', \alpha_{-i})$:

$$\begin{aligned} U_i^*(\theta_i; \alpha_i', \alpha_{-i}) &= \max_{\theta_i'} U_i(\theta_i', \theta_i; \alpha_i', \alpha_{-i}) \\ &= \max_{\theta_i'} w_i(v_i)x(\theta_i', \Theta_{-i}) \cdot \gamma(\Theta_{-i}|\theta_i, \alpha_i', \alpha_{-i}) - t(\theta_i', \Theta_{-i}) \cdot \gamma(\Theta_{-i}|\theta_i, \alpha_i', \alpha_{-i}) \\ &= \max \Bigg\{ \max_{\theta_i' \in \Theta_i^+(\alpha, f) \setminus (v_i^*(\succeq_i) \times S_i)} w_i(v_i)x(\theta_i', \Theta_{-i}) \cdot \gamma(\Theta_{-i}|\theta_i, \alpha_i', \alpha_{-i}) \\ &\qquad - (w_i(v_i')x_i(\theta_i', \Theta_{-i}) + M\tau(\theta_i')) \cdot \gamma(\Theta_{-i}|\theta_i, \alpha_i', \alpha_{-i}), 0 \Bigg\} \\ &= \max \Bigg\{ \max_{\theta_i' \in \Theta_i^+(a, f) \setminus (v_i^*(\succeq_i) \times S_i)} (w_i(v_i) - w_i(v_i'))x(\theta_i', \Theta_{-i}) \cdot \gamma(\Theta_{-i}|\theta_i, \alpha_i', \alpha_{-i}) \\ &\qquad - M\tau(\theta_i') \cdot \gamma(\Theta_{-i}|\theta_i, \alpha_i', \alpha_{-i}), 0 \Bigg\} \\ &\leq 0. \end{aligned}$$

The last inequality holds because for any $\alpha_i' \in A_i^{|V_i|}$, $\theta_i \in \Theta_i(\alpha_i', \alpha_{-i})$, and $\theta_i' \in \Theta_i^+(a, f) \setminus (v_i^*(\succeq_i) \times S_i)$, either $\gamma(\Theta_{-i}|\theta_i, \alpha_i', \alpha_{-i}) \cdot \tau(\theta_i') = 0$ and $w(v_i) \leq w(v_i')$, or $\gamma(\Theta_{-i}|\theta_i, \alpha_i', \alpha_{-i}) \cdot \tau(\theta_i') > \varepsilon$ for some $\varepsilon > 0$ (i.e. ε is a uniform lower bound across all α_i').[6] The former was shown when we established

[6]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 $\gamma(\Theta_{-i}|\theta_i, \alpha_i', \alpha_{-i}) \cdot \tau(\theta_i') > 0$ for all α_i', because now there is the possibility that $\gamma(\Theta_{-i}|\theta_i, \alpha_i', \alpha_{-i}) \cdot \tau(\theta_i')$ can be made arbitrarily close to 0.

incentive compatibility. To see the latter, suppose that for every $\varepsilon > 0$, there exists $\alpha_i' \in A_i^{|V_i|}$, $\theta_i \in \Theta_i(\alpha_i', \alpha_{-i})$ where $w_i(v_i) > w_i(v_i')$ such that $\gamma(\Theta_{-i}|\theta_i, \alpha_i', \alpha_{-i}) \cdot \tau(\theta_i') < \varepsilon$. Let $\gamma^\varepsilon \in \overline{D}(\theta_i', \alpha_{-i}, \succeq_i)$ denote the sequence of such beliefs. Since $\overline{D}(\theta_i', \alpha_{-i}, \succeq_i)$ is compact, as $\varepsilon \to 0$, there is a convergent subsequence which converges to $\gamma \in \overline{D}(\theta_i', \alpha_{-i}, \succeq_i)$, 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 $F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\theta_i', \alpha) = \{\gamma : \gamma \cdot \tau(\theta_i') = 0\} \cap \overline{C}(\alpha_{-i})$, $\gamma \in F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\theta_i', \alpha)$ contradicting Condition 2. Thus, for sufficiently large M, the inequality holds.

Now we prove necessity. Suppose that there exists $\succeq$ such that Condition 1 is not satisfied. Define w such that for all i, $w_i(v_i) \geq w_i(v_i')$ if and only if $v_i \succeq_i v_i'$, and for all i, for $v_i \in v_i^*(\succeq_i)$ let $w_i(v_i) = K$, where K is sufficiently large. Take an arbitrary α and let $\hat{\theta}_i \in \Theta_i^+(\alpha, f) \setminus (v_i^*(\succeq_i) \times S_i)$ such that $\gamma(\Theta_{-i}|\hat{\theta}_i, \alpha) \cap \mathrm{ri}(C(\alpha_{-i})) \neq \emptyset$. Suppose for a contradiction that (x, t) is incentive compatible, individually rational, and fully extracts rent given α, α is chosen optimally, and $x_i(\theta_i, \theta_{-i}) = f_i(v_i, v_{-i})$ for all θ such that $\gamma(\theta|\alpha) > 0$. Without loss of generality, let $t_i(\hat{\theta}_i, \Theta_{-i}) = w_i(\hat{v}_i)x_i(\hat{\theta}_i, \Theta_{-i}) + M\tau(\hat{\theta}_i)$ for some $\tau(\hat{\theta}_i)$. For rent extraction, we need $\tau(\hat{\theta}_i) \cdot \gamma(\Theta_{-i}|\hat{\theta}_i, \alpha) = 0$. For $\theta_i \in v_i^*(\succeq_i) \times S_i$, incentive compatibility requires $\gamma(\Theta_{-i}|\theta_i, \alpha) \cdot ((K - w(\hat{v}_i))x_i(\hat{\theta}_i, \Theta_{-i}) - M\tau(\hat{\theta}_i)) \leq 0$, which implies:

$$\gamma(\Theta_{-i}|\theta_i, \alpha) \cdot M\tau(\hat{\theta}_i) \geq \gamma(\Theta_{-i}|\theta_i, \alpha) \cdot (K - w(\hat{v}_i))x_i(\hat{\theta}_i, \Theta_{-i}) > 0.$$

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 $\alpha_i' \in A_i^{|V_i|}$, $\theta_i' \in \Theta_i(\alpha_i', \alpha_{-i})$ such that $\gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) \cdot \tau(\hat{\theta}_i) < 0$. The previous displayed inequality implies that for large K, M must also be large. Then:

$$\begin{aligned}
&\sum_{\theta_i \in \Theta_i(\alpha_i', \alpha_{-i})} \gamma_{\Theta_i}(\theta_i|\alpha_i', \alpha_{-i}) U_i^*(\theta_i; \alpha_i', \alpha_{-i}) \\
&\quad \geq \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) U_i^*(\theta_i'; \alpha_i', \alpha_{-i}) \\
&\quad \geq \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) U_i(\hat{\theta}_i, \theta_i'; \alpha_i', \alpha_{-i}) \\
&\quad = \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) \left(w_i(v_i')x_i(\hat{\theta}_i, \Theta_{-i}) - t(\hat{\theta}_i, \Theta_{-i})\right) \cdot \gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) \\
&\quad = \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) \left((w_i(v_i') - w_i(\hat{v}_i))x_i(\hat{\theta}_i, \Theta_{-i}) - M\tau(\hat{\theta}_i)\right) \cdot \gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) \\
&\quad > 0.
\end{aligned}$$

Suppose that there exists $\succeq$ such that Condition 2 is not satisfied. Then define w as before, take an arbitrary α, and let $\hat{\theta}_i \in \Theta_i^+(\alpha, f) \setminus (v_i^*(\succeq_i) \times S_i)$ be such that $F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\hat{\theta}_i, \alpha)) \cap \overline{D}(\hat{\theta}_i, \alpha_{-i}, \succeq_i) \neq \emptyset$. Suppose for a contradiction that (x, t) is incentive compatible, individually rational, and fully extracts rent given α, α is chosen optimally, and $x_i(\theta_i, \theta_{-i}) = f_i(v_i, v_{-i})$ for all θ such that $\gamma(\theta|\alpha) > 0$. Without loss of generality, let $t_i(\hat{\theta}_i, \Theta_{-i}) = w_i(\hat{v}_i)x_i(\hat{\theta}_i, \Theta_{-i}) + M\tau(\hat{\theta}_i)$ for some $\tau(\hat{\theta}_i)$. For rent extraction, we need $\tau(\hat{\theta}_i) \cdot \gamma(\Theta_{-i}|\hat{\theta}_i, \alpha) = 0$. Assume that $\tau(\hat{\theta}_i) \cdot \gamma(\Theta_{-i}|\hat{\theta}_i, \alpha) \geq 0$ for all $\gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i})$, 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 $F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\hat{\theta}_i, \alpha)) \subset \{\gamma : \gamma \cdot \tau(\hat{\theta}_i) = 0\}$

Since Condition 2 fails, there exists $\gamma \in F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\hat{\theta}_i, \alpha)) \cap \overline{D}(\hat{\theta}_i, \alpha_{-i}, \succeq_i)$. We show that for every $\varepsilon > 0$, we can find $\gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i})$ such that $w_i(v_i') > w_i(\hat{v}_i)$, $\theta_i' \in \Theta_i(\alpha_i', \alpha_{-i})$, and $\tau(\hat{\theta}_i) \cdot \gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) < \varepsilon$. To see this, take any sequence $\gamma^k \in D(\hat{\theta}_i, \alpha_{-i}, \succeq_i)$ converging to γ. As $k \to \infty$, $\tau(\hat{\theta}_i) \cdot \gamma^k \to \tau(\hat{\theta}_i) \cdot \gamma = 0$. Since $\gamma^k \in D(\hat{\theta}_i, \alpha_{-i}, \succeq_i)$, for each k, $\gamma^k = \gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i})$

such that $w_i(v_i') > w_i(\hat{v}_i)$, $\theta_i' \in \Theta_i(\alpha_i', \alpha_{-i})$. Take a sufficiently large k and we are done. Then:

$$\begin{aligned}
&\sum_{\theta_i \in \Theta_i(\alpha_i', \alpha_{-i})} \gamma_{\Theta_i}(\theta_i|\alpha_i', \alpha_{-i}) U_i^*(\theta_i; \alpha_i', \alpha_{-i}) \\
&\quad \geq \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) U_i^*(\theta_i'; \alpha_i', \alpha_{-i}) \\
&\quad \geq \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) U_i(\hat{\theta}_i, \theta_i'; \alpha_i', \alpha_{-i}) \\
&\quad = \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) \left(w_i(v_i') x_i(\hat{\theta}_i, \Theta_{-i}) - t(\hat{\theta}_i, \Theta_{-i})\right) \cdot \gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) \\
&\quad = \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) \left((w_i(v_i') - w_i(\hat{v}_i)) x_i(\hat{\theta}_i, \Theta_{-i}) - M\tau(\hat{\theta}_i)\right) \cdot \gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) \\
&\quad > 0.
\end{aligned}$$

□

Proof of Proposition 1.6.1. This proof is almost identical to the proof of Proposition 1.5.1 and is included for completeness. For sufficiency, take any w, and let $\succeq$ be such that for all i, $v_i \succeq_i v_i'$ if and only if $w_i(v_i) \geq w_i(v_i')$. Let i^* be such that $\min_{v_{i^*}'} w_{i^*}(v_{i^*}') \geq \min_{v_j'} w_j(v_j')$ for all j. Note that we can choose f^* such that for all $i \neq i^*$, and for $v_i \in v_{*i}(\succeq_i)$, $f_i^*(v_i, v_{-i}) = 0$ for all v_{-i}, since $w(v_i) \leq w(v_{i^*})$. Hence for $i \neq i^*$, $\Theta_i^+(\alpha, f^*) \subset \Theta_i(\alpha) \setminus (v_{*i}(\succeq_i) \times S_i)$, and for $i = i^*$, $\Theta_i^+(\alpha, f^*) \subset \Theta_i(\alpha)$. Proposition 1.5.1 then implies that the seller can fully extract rent for the social choice function f^*.

Now we prove necessity. Take an arbitrary α, and suppose that there exists $\succeq$ and a buyer i such that Condition 1 or Condition 2 is not satisfied for some $\theta_i \in \Theta_i(\alpha) \setminus (v_i^*(\succeq_i) \times S_i)$. Define w such that for all $j \in I$, $w_j(v_j) \geq w_j(v_j')$ if and only if $v_j \succeq_j v_j'$, let $\max_{v_j'} w_j(v_j') = K$, and for $j \neq i$, let $\min_{v_j'} w_j(v_j') < \min_{v_i'} w_i(v_i')$. Now we can replicate the steps in the proof of necessity in Proposition 1.5.1.

If Condition 1 is not satisfied, let $\hat{\theta}_i \in \Theta_i(\alpha) \setminus (v_i^*(\succeq_i) \times S_i)$ be such that $\gamma(\Theta_{-i}|\hat{\theta}_i, \alpha) \cap \mathrm{ri}(C(\alpha_{-i})) \neq \emptyset$. Note that $f_i^*(\hat{v}_i, v_{-i}) > 0$ for some v_{-i} (in particular, when for all $j \neq i$, $v_j \in \arg\min_{v_j'} w_j(v_j')$). Suppose for a contradiction that (x, t) is incentive compatible, individually rational, and fully extracts rent given α, α is chosen optimally, and $x_i(\theta_i, \theta_{-i}) = f_i^*(v_i, v_{-i})$ for all θ such that $\gamma(\theta|\alpha) > 0$. Without loss of generality, let $t_i(\hat{\theta}_i, \Theta_{-i}) = w_i(\hat{v}_i) x_i(\hat{\theta}_i, \Theta_{-i}) + M\tau(\hat{\theta}_i)$ for some $\tau(\hat{\theta}_i)$. For rent extraction, we need $\tau(\hat{\theta}_i) \cdot \gamma(\Theta_{-i}|\hat{\theta}_i, \alpha) = 0$. For $\theta_i \in v_i^*(\succeq_i) \times S_i$, incentive compatibility requires $\gamma(\Theta_{-i}|\theta_i, \alpha) \cdot ((K - w(\hat{v}_i)) x_i(\hat{\theta}_i, \Theta_{-i}) - M\tau(\hat{\theta}_i)) \leq 0$, which implies:

$$\gamma(\Theta_{-i}|\theta_i, \alpha) \cdot M\tau(\hat{\theta}_i) \geq \gamma(\Theta_{-i}|\theta_i, \alpha) \cdot (K - w(\hat{v}_i)) x_i(\hat{\theta}_i, \Theta_{-i}) > 0.$$

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 $\alpha_i' \in A_i^{|V_i|}$, $\theta_i' \in \Theta_i(\alpha_i', \alpha_{-i})$ such that $\gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) \cdot \tau(\hat{\theta}_i) < 0$. The previous displayed inequality implies that for large K, M must also be large. Then:

$$\begin{aligned}
&\sum_{\theta_i \in \Theta_i(\alpha_i', \alpha_{-i})} \gamma_{\Theta_i}(\theta_i|\alpha_i', \alpha_{-i}) U_i^*(\theta_i; \alpha_i', \alpha_{-i}) \\
&\quad \geq \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) U_i^*(\theta_i'; \alpha_i', \alpha_{-i}) \\
&\quad \geq \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) U_i(\hat{\theta}_i, \theta_i'; \alpha_i', \alpha_{-i}) \\
&\quad = \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) \left(w_i(v_i') x_i(\hat{\theta}_i, \Theta_{-i}) - t(\hat{\theta}_i, \Theta_{-i})\right) \cdot \gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) \\
&\quad = \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) \left((w_i(v_i') - w_i(\hat{v}_i)) x_i(\hat{\theta}_i, \Theta_{-i}) - M\tau(\hat{\theta}_i)\right) \cdot \gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) \\
&\quad > 0.
\end{aligned}$$

If Condition 2 is not satisfied, let $\hat{\theta}_i \in \Theta_i(\alpha) \setminus (v_i^*(\succeq_i) \times S_i)$ be such that $F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\hat{\theta}_i, \alpha)) \cap \overline{D}(\hat{\theta}_i, \alpha_{-i}, \succeq_i) \neq \emptyset$. Note that $f_i^*(\hat{v}_i, v_{-i}) > 0$ for some v_{-i} (in particular, when for all $j \neq i$, $v_j \in \arg\min_{v_j'} w_j(v_j')$). Suppose for a contradiction that (x, t) is incentive compatible, individually rational, and fully extracts rent given α, α is chosen optimally, and $x_i(\theta_i, \theta_{-i}) = f_i^*(v_i, v_{-i})$ for all θ such that $\gamma(\theta|\alpha) > 0$. Without loss of generality, let $t_i(\hat{\theta}_i, \Theta_{-i}) = w_i(\hat{v}_i)x_i(\hat{\theta}_i, \Theta_{-i}) + M\tau(\hat{\theta}_i)$ for some $\tau(\hat{\theta}_i)$. For rent extraction, we need $\tau(\hat{\theta}_i) \cdot \gamma(\Theta_{-i}|\hat{\theta}_i, \alpha) = 0$. Assume that $\tau(\hat{\theta}_i) \cdot \gamma(\Theta_{-i}|\hat{\theta}_i, \alpha) \geq 0$ for all $\gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i})$, 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 $F_{\overline{C}(\alpha_{-i})}(\gamma(\Theta_{-i}|\hat{\theta}_i, \alpha)) \subset \{\gamma : \gamma \cdot \tau(\hat{\theta}_i) = 0\}$

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

$$
\begin{aligned}
\sum_{\theta_i \in \Theta_i(\alpha_i', \alpha_{-i})} & \gamma_{\Theta_i}(\theta_i|\alpha_i', \alpha_{-i}) U_i^*(\theta_i; \alpha_i', \alpha_{-i}) \\
& \geq \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) U_i^*(\theta_i'; \alpha_i', \alpha_{-i}) \\
& \geq \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) U_i(\hat{\theta}_i, \theta_i'; \alpha_i', \alpha_{-i}) \\
& = \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) \left(w_i(v_i')x_i(\hat{\theta}_i, \Theta_{-i}) - t(\hat{\theta}_i, \Theta_{-i})\right) \cdot \gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) \\
& = \gamma_{\Theta_i}(\theta_i'|\alpha_i', \alpha_{-i}) \left((w_i(v_i') - w_i(\hat{v}_i))x_i(\hat{\theta}_i, \Theta_{-i}) - M\tau(\hat{\theta}_i)\right) \cdot \gamma(\Theta_{-i}|\theta_i', \alpha_i', \alpha_{-i}) \\
& > 0.
\end{aligned}
$$

Now suppose that there exists $\succeq$ and a buyer i such that Condition 3 or Condition 4 is not satisfied for some $j \neq i$, and $\theta_j \in \Theta_j(\alpha) \setminus ((v_j^*(\succeq_j) \cup v_{*j}(\succeq_j)) \times S_j)$. Define w such that for all $j \in I$, $w_j(v_j) \geq w_j(v_j')$ if and only if $v_j \succeq_j v_j'$, and let $\min_{v_j'} w_j(v_j') = \kappa$ and $\max_{v_j'} w_j(v_j') = K$. If Condition 3 is not satisfied, let $\hat{\theta}_j \in \Theta_j(\alpha) \setminus ((v_j^*(\succeq_j) \cup v_{*j}(\succeq_j)) \times S_j)$ be such that $\gamma(\Theta_{-j}|\hat{\theta}_j, \alpha) \cap \mathrm{ri}(C(\alpha_{-j})) \neq \emptyset$. If Condition 4 is not satisfied, let $\hat{\theta}_j \in \Theta_j(\alpha) \setminus ((v_j^*(\succeq_j) \cup v_{*j}(\succeq_j)) \times S_j)$ be such that $F_{\overline{C}(\alpha_{-j})}(\gamma(\Theta_{-j}|\hat{\theta}_j, \alpha)) \cap \overline{D}(\hat{\theta}_j, \alpha_{-j}, \succeq_j) \neq \emptyset$. Note that $f_j^*(\hat{v}_j, v_{-j}) > 0$ for some v_{-j} (in particular, when for all $k \neq j$, $v_k \in \arg\min_{v_k'} w_k(v_k')$). Then by exactly the same argument as before, there can be no (x, t) that is incentive compatible, individually rational, and fully extracts rent given α, where α is chosen optimally, and $x_j(\theta_j, \theta_{-j}) = f_j^*(v_j, v_{-j})$ for all θ such that $\gamma(\theta|\alpha) > 0$. □

Chapter 2

Private and Common Value Auctions with Ambiguity over Correlation[1]

2.1 Introduction

In auctions, as in many other strategic situations, individuals often have a good understanding of their own private information but they might know less about others' information sources. For example, in auctions for drilling rights, a company can understand the test results that it conducted but might be worried that its results are correlated with those of other firms. Similarly the evaluation one gets about a piece of art on sale might be correlated in complex ways to the evaluations other bidders get.

In these situations bidders might worry about their lack of understanding of the correlation structure between their own information and that of other bidders.[2] These considerations are important for their bidding behaviour. In private value auctions, beliefs about correlation influence the assessment of bidders about the competition they might face. In common value auctions, such beliefs also affect the bidders' valuation of the good, which implies an additional effect on the bidders' strategies.

In this paper we analyse private and common value auctions when individuals have ambiguity over the joint information structures generating valuations and signals. Specifically, we assume that individuals know the marginal information structure generating a value or a signal to each bidder, but that they are aware that their information sources might be correlated to a degree, and face ambiguity over the possible correlation scenarios. We propose a simple model to analyse ambiguity over correlation structures that is tailored to the comparison with the standard model.[3] In particular, we use a single parameter, a, to bound the degree of *pointwise mutual information* of the information structures.[4] When an individual receives a signal and contemplates what strategy to play, she faces ambiguity aversion (as in Gilboa and Schmeidler (1989)) about the set of potential joint information structures that are bounded by a.

Specifically, we analyse two-bidder, binary-value and binary-signal auctions (in Section 5 we

[1]This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 681579. We thank Subir Bose and Francesco Nava for helpful comments.

[2]Here we consider sophisticated individuals who entertain the possibility that such correlation might exist. A recent literature looks at naive individuals who are not aware of the correlation between sources of information, i.e., correlation neglect. See Ortoleva and Snowberg (2015), Glaeser and Sunstein (2009), Levy and Razin (2015*a*,*b*), Eyster and Weizsäcker. (2011), Kallir and Sonsino (2009), and Enke and Zimmermann (2017).

[3]Auctions are typically analysed under the assumption of conditionally-independent private information, with the bidders aware of this fact.

[4]See Church and Hanks (1991).

extend the analysis to models with continuous signals and with many bidders). In this model, the set of information structures that the bidders consider is centred around the true canonical case of independent values and signals, and takes the following form. Consider the two bidders, 1 and 2, with values $\mathbf{v} = (v_1, v_2)$ each receiving a signal, s_1 and s_2 respectively. Let $q(\mathbf{s}|\mathbf{v})$ denote a joint probability of the signal vector $\mathbf{s} = (s_1, s_2)$ conditional on the valuations for the good, and $q_1(s_1|\mathbf{v})$ and $q_2(s_2|\mathbf{v})$ denote the marginal probabilities of s_1 and s_2 conditional on $\mathbf{v}$. Similarly, let $p(\mathbf{v})$ denote the prior over $\mathbf{v}$, with $p_i(v_i)$ the marginal prior. We assume that individuals' ambiguity is over the set of joint information structures satisfying $\frac{1}{a} \leq \frac{p(\mathbf{v})q(\mathbf{s}|\mathbf{v})}{\prod_{i=1,2} p_i(v_i) \prod_{i=1,2} q_i(s_i|\mathbf{v})} \leq a$ for some finite parameter $a \geq 1$, for any $\mathbf{s}, \mathbf{v}$.[5] In other words, the (exponential) pointwise mutual information (ePMI) is bounded. The higher is a, the larger is the ambiguity. When $a = 1$ we are back in the standard model with conditionally independent signals and no ambiguity. This framework allows us to analyse ambiguity in the private value auction (where we focus on ambiguity over $p(\mathbf{v})$), and in the common value auction (where we focus on ambiguity over $q(\mathbf{s}|\mathbf{v})$).

In the model, ambiguity over information structures is exogenous but ambiguity over the valuations or the equilibrium bids of opponents is endogenous, and depends on the strategic interaction. An ambiguity averse bidder focuses attention on beliefs that minimise her expected utility. This will be reflected in beliefs that either put weight on more competition, believing the other bidders have similar valuations, or on a low value of the good. The former effect, the "competition effect", is present in both private and common value auctions. The second effect, a "winner's value effect", exists only in common value auctions.

We analyse the equilibria in the second-price and first-price auctions. We first focus on the competition effect. We show how this effect implies overbidding in the private values case; the distribution of bids in the first price auction in the face of ambiguity first-order stochastically dominates that in the standard model. This implies that ambiguity is beneficial for the seller as her revenue increases with a. In contrast, the "winner's value effect" generally implies that bidders should shade their bids. We show that in common-value auctions, in which both effects are present, the "winner's value effect" is stronger and the seller's revenue decreases in ambiguity. In particular, the low type always shades her bids while the high type does so unless ambiguity is small and signals too imprecise. In Section 5 we show that this result (and the intuition behind it) also holds when we consider an environment with many bidders, as well as an environment with continuous signals.

We then turn to consider optimal auction design in the face of ambiguity over correlation structures. We first show that the above results about the seller's revenue also hold for the optimal auction; the seller's optimal revenue increases in ambiguity in the private values case but decreases in the common value case. In the private values case our results are consistent with Bose, Ozdenoren and Pape (2006) who show that the optimal mechanism fully insures the bidders against ambiguity. In Bose et al (2006) bidders believe that valuations are independently drawn, but face ambiguity regarding the particular distributions of valuations that others have. In contrast, in our model bidders know the marginal distribution of valuations but face ambiguity about the joint distribution of valuations. Nevertheless, the full insurance result holds in our setting when values are private.

The full insurance result implies that the competition effect does not arise in equilibrium in the optimal auction. However, there is a sense in which the seller exploits the competition effect to increase her revenue. Since under the worst case belief the high type believes that her opponent

[5]Levy and Razin (2018) show that this restriction provides a meaningful way to constrain the set of ambiguous beliefs.

is likely to also be a high type, she is less willing to deviate than when there is no ambiguity, so that the seller is able to extract more rent from the high type. Thus, the seller's revenue increases in ambiguity.

In the common value case, insuring the buyers against ambiguity is not necessarily optimal. In the standard case without ambiguity, any optimal mechanism must allocate the good with probability 1 and fully extract rent. However, any such mechanism cannot be a full insurance mechanism. To prevent the high type from deviating, the seller has to conduct side bets with the low type. Under ambiguity, these side bets are costly since the expected payment to the seller is smaller under the seller's belief than under the worst case belief of the buyer. However, if ambiguity is small or if the signals are very precise, these side bets remain optimal and the low type is not insured against ambiguity.

Furthermore, even though there is no need to conduct side bets with the high type to satisfy any incentive constraint, it is also not necessarily optimal to fully insure the high type. This is because to fully insure the high type, the allocation rule needs to be independent of the type of her opponent. Otherwise, the winner's value effect implies that the high type will underestimate the value of the object. However, when the ambiguity is small or the signals are precise, the seller will allocate the good to the high type with higher probability when her opponent is a low type in order to slacken the incentive constraint. In this case, the seller provides partial insurance by also asking the high type to pay more when her opponent is low type. Under the worst case belief, the high type does not care about the type of her opponent, but undervalues the object conditional on winning. In other words, the seller insures the buyer against the competition effect but not the winner's value effect.

On the other hand if the signals are very imprecise and the ambiguity large, the seller finds it optimal to fully insure the buyers against the ambiguity, so that the allocation of the good does not depend on their signals. As a result, the high type earns positive rents in equilibrium. Finally, we show that the seller's revenue in the optimal mechanism is decreasing in the amount of ambiguity, as we found in both the first and second price auctions.

2.2 Related Literature

Our paper is related to a recent literature on ambiguity and auctions. As far as we know, our paper is the first to analyse ambiguity in common-value auctions. For private-value auctions, Salo and Weber (1995) and Chen, Katušcák and Ozdenoren (2007) show how ambiguity aversion translates to higher bids as individuals underestimate their winning probabilities.[6] We complement their analysis by defining ambiguity differently, and more importantly, by considering the common-values case and the comparison with the private-values case, both positively and normatively. Bose, Ozdenoren and Pape (2006) analyse optimal auction mechanisms for private-value auctions with ambiguity over other bidders' valuations. They show that the seller will fully insure the buyers against ambiguity. Again, our key contribution is to analyse the common-values case and compare it to the private-values case. Specifically, we show that in the common value auction sometimes only partial insurance arises. Lo (1998) shows that the first-price auction dominates the second-price auction in some environments. He uses a multiple priors approach and shows that equilibrium bids are simply determined as if all players hold the worst-case prior. In our analysis

[6] Chen, Katušcák and Ozdenoren (2007) however provide experimental evidence that bids are lower in the presence of ambiguity in first and second-price auctions with independent private values.

players with different signals focus on different beliefs so the model is not equivalent to one in which players start for example from a mis-specified model with wrong beliefs.

Bose and Renou (2014) study how principals can use ambiguous mechanisms to implement social welfare functions that are not attainable under unambiguous mechanisms. In particular, they construct ambiguous communication mechanisms between the agents and a moderator resulting with agents updating to sets of beliefs. Hanany, Klibanoff and Mukerji (Forthcoming) is a recent contribution to the study of incomplete information games and ambiguity. Finally, Bergemann, Brooks and Morris (2017) consider private values auctions and study the set of achievable utilities when considering, as modellers, the set of different feasible information structures. Our analysis is different as in our approach it is the economic agents, rather than the modeller, who span the possible information structures. In addition, we restrict the set of possible information structures using the notion of pointwise mutual information. We show how this shifts equilibrium behaviour in a non-trivial way.

2.3 Model

We consider a simple symmetric auction with two bidders (1 and 2), each with two possible valuations $v \in \{L, H\}$ where $0 \leq L < H = 1$. For expositional purposes we will focus on the case of $L = 0$ in the body of the paper. Results are easily generalised to $L > 0$ (see the Appendix).

We will consider two cases, private values and common values. In the private value case we will assume that valuations are independent across bidders and distributed uniformly. In the common value case we will assume that the valuations are fully positively correlated with a uniform prior. Bidders know if they are in the private or common values case.

The paper focuses on ambiguity over joint information structures. We model this ambiguity by spanning a set of possible symmetric information structures around the true information structures mentioned above for the private and common value cases. In particular, we assume that bidders know the following aspects of the environment:

1. The state space. A state is a vector of valuations $\mathbf{v} = (v_1, v_2) \in \{L, H\}^2$.

2. Priors about the state. Each bidder knows that the marginal prior distributions are uniform; that is, the probability that each bidder has a low value is a half.

3. Signals. Each bidder $i = 1, 2$ observes a private signal $s_i \in \{l, h\}$.

Bidders entertain a set of possible joint information structures that are consistent with the above three aspects of the environment. A *joint information structure* is $(\{L, H\}^2, \{l, h\}^2, q(\mathbf{s}|\mathbf{v}), p(\mathbf{v}))$ where $q(\mathbf{s}|\mathbf{v})$ is a joint probability on $\{l, h\}^2$ conditional on $\mathbf{v} \in \{L, H\}^2$ and $p(\mathbf{v})$ a distribution over $\{L, H\}^2$ that is consistent with uniform marginals.

To define the level of ambiguity, we use a simple one-parameter characterisation for the set of joint information structures introduced in Levy and Razin (2018). This characterisation uses the exponent of the *pointwise mutual information* (ePMI) to define bounds on the correlation between information structures. Specifically, for each bidder let $q_i(s_i|\mathbf{v})$ denote the marginal conditional probability of receiving the private signal $s_i \in \{l, h\}$ given state $\mathbf{v}$ and let $p_i(v_i)$ denote the marginal prior. We assume the following:

Assumption A1: *There is a parameter* $1 \leq a < \infty$*, such that each bidder only considers joint information structures,* $(\{L,H\}^2, \{l,h\}^2, q(\mathbf{s}|\mathbf{v}), p(\mathbf{v}))$*, so that at any state* $\mathbf{v} \in \{L,H\}^2$ *and for any vector of signals* $\mathbf{s} \in \{l,h\}^2$*,*

$$\frac{1}{a} \leq \frac{p(\mathbf{v})q(\mathbf{s}|\mathbf{v})}{\prod_{i=1,2} p_i(v_i) \prod_{i=1,2} q_i(s_i|\mathbf{v})} \leq a.$$ [7]

The parameter a, the PMI-bound, describes the extent of the ambiguity a bidder faces over the set of correlation scenarios. It is straightforward to see that ambiguity is larger when a is larger. The formulation of the set is general in the sense that it is detail-free in terms of the underlying distribution functions. It also captures the maximal set of joint information structures with correlation bounded by a. Note that a joint information structure which satisfies independence would have $a = 1$ at any point; whenever a joint information structure does not satisfy independence then the ePMI is less than 1 for some $(\mathbf{s}, \mathbf{v})$, and is greater than 1 for some $(\mathbf{s}', \mathbf{v}')$.[8] A higher a implies that bidders consider information structures that are more *concordant;* for more on pointwise mutual information as a measure of correlation see Levy and Razin (2018).

In Section 3 we will analyse first-price and second-price auctions for the private values and common values cases, while Section 4 will provide the optimal auction analysis. In all these cases, an equilibrium is denoted by a pair of bidding strategies for the two players, $(b^1(s_1), b^2(s_2))$, and a symmetric equilibrium has $b^1(.) = b^2(.) \equiv b(.)$. We consider max-min behaviour. Specifically, in equilibrium, given an observed signal, a bidding strategy maximises the utility of the individual under the worst-case information structure.

Our framework is flexible enough to consider different types of underlying correlations, which will be tied to the private/common values environment (indeed in Section 5 we consider environments with continuous signals and with many bidders). To complete the model, we now specify the feasible information structures considered by bidders in these two environments.

Private values

In the private values model bidders know their values (and know they are in the private values model). As signals are fully informative about one's own value, ambiguity will arise about the correlation in the prior distribution $p(\mathbf{v})$. Specifically, each prior distribution must be consistent with the uniform marginal prior distributions, so that the set of functions $p(\mathbf{v})$ for $\mathbf{v} \in \{0,1\}^2$ is represented by:

Table 2.1: Joint information structures for the private value case

$p(\mathbf{v})$	0	1
0	α	$\frac{1}{2} - \alpha$
1	$\frac{1}{2} - \alpha$	α

Ambiguity is then over the parameter α. Given a, the ePMI constraints impose the following restrictions on α:

$$\frac{1}{a} \leq 4\alpha \leq a, \ \frac{1}{a} \leq 4(\frac{1}{2} - \alpha) \leq a$$

It is easy to see that the higher is a, the larger is the set of possible information structures that are considered by bidders. From the above it is also easy to see that each bidder can consider α

[7] All the results can be easily generalised if instead of the lower bound $\frac{1}{a}$ we use some finite $b < 1$.

[8] It is then impossible to consider only priors/information structures with ePMI that is only higher (lower) than 1.

satisfying

$$\frac{1}{4a} = \underline{\alpha}(a) \le \alpha \le \bar{\alpha}(a) = \frac{1}{2}(1 - \frac{1}{2a})$$

Common values

In the common values case, the ambiguity of the bidders will affect their perception about the correlation between their signals and those of their opponents, through $q(\mathbf{s}|\mathbf{v})$. In the previous case we had fixed the marginal priors; here we fix the marginal distributions of the signals. In particular, there are two states of the world, $\mathbf{v} \in \{(1,1),(0,0)\}$. The probability of receiving the signal l in state $\mathbf{v} = (0,0)$, or the signal h in state $\mathbf{v} = (1,1)$, is $q > \frac{1}{2}$ i.e., $q(s = l|(0,0)) = q(s = h|(1,1)) = q$.[9] We assume that the true joint probability distribution, $q(s_1, s_2|\mathbf{v})$, satisfies conditional independence, so that $q(s_1, s_2|\mathbf{v}) = \prod_{i=1,2} q_i(s_i|\mathbf{v})$. However, while individuals know the true marginal probability distribution generating both their signals, they have ambiguity over the set of joint information structures.

With the above specification, we can represent the family of information structures the bidders entertain by:[10]

Table 2.2: Joint information structures for the common value case

$q(\mathbf{s}\|(0,0))$	l	h	$q(\mathbf{s}\|(1,1))$	l	h
l	α_0	$q - \alpha_0$	l	α_1	$1 - q - \alpha_1$
h	$q - \alpha_0$	$1 - 2q + \alpha_0$	h	$1 - q - \alpha_1$	$2q - 1 + \alpha_1$

Under independence ($a = 1$), $\alpha_0 = q^2$ and $\alpha_1 = (1-q)^2$. In this case then, α_0 and α_1 are the parameters over which there is ambiguity, as we define below. It is then easy to derive the bounds for these parameters using the ePMI constraints. Specifically, for a general a, the ePMI constraints imply the following bounds on the values of the αs:

$$\underline{\alpha}_0(a) \le \alpha_0 \le \overline{\alpha}_0(a)$$
$$\underline{\alpha}_1(a) \le \alpha_1 \le \overline{\alpha}_1(a)$$

where:

$$\underline{\alpha}_0(a) = \frac{1}{a}(1-q)^2 + 2q - 1,\ \underline{\alpha}_1(a) = \frac{1}{a}(1-q)^2$$
$$\overline{\alpha}_0(a) = \begin{cases} a(1-q)^2 + 2q - 1 & a \le \frac{q}{1-q} \\ q - \frac{1}{a}q(1-q) & a > \frac{q}{1-q} \end{cases},\ \overline{\alpha}_1(a) = \begin{cases} a(1-q)^2 & a \le \frac{q}{1-q} \\ 1 - q - \frac{1}{a}q(1-q) & a > \frac{q}{1-q} \end{cases}.$$

2.4 The competition and winner's value effects

We now show how ambiguity over correlation can affect bids in two different ways. The *competition effect* arises as ambiguity will play a role in shaping beliefs about the probable bids of the opponent. The *winner's value effect* will arise as ambiguity may play a role in shaping beliefs about one's own valuation given the information held by the opponent. Naturally, only the first effect will arise in the private values case while both will arise in the common values case.

[9] The analysis can be extended to non-symmetric marginal probability distributions.

[10] The table describes an information structure so for each state, all cell entries are non-negative and all entries sum up to one.

Private-values

In the private value case, for each bidder i, the signal $s_i \in \{l, h\}$ fully reveals her value. Ambiguity will then play a role in shaping beliefs about the probable bids of the opponent but not about the value of the good. Below we show that the competition effect implies higher bids. The results in this section are closely related to Salo and Weber (1995) and Chen, Katušcák and Ozdenoren (2007) who analyse auctions with ambiguity about the distribution of values in the case of private values. For concreteness, and in order to compare this environment to that of common values, we replicate these results in our setting.

Consider the first-price auction. When $a = 1$, this is the standard model and the unique equilibrium has the low type submitting a bid of zero, and the high type mixing between bids in $[0, \frac{1}{2}]$ according to the distribution $F(b) = \frac{b}{1-b}$. When $a > 1$, the low type still bids zero—her value—in equilibrium.

We now consider the high type, when $a > 1$. Let p be the belief of the high type of bidder 1 that bidder 2 is a high type. From Table 1 we have that $p = 2\alpha$, and thus given the ePMI constraints for the high type of bidder 1, given a mixed strategy $F(.)$ of the other player, the expected utility for some bid b is given by:

$$\min_{p=2\alpha,\ \underline{\alpha}(a)\leq\alpha\leq\bar{\alpha}(a)} \Pr(\text{Bidder 1 wins})(1-b) = \min_{p=2\alpha,\ \underline{\alpha}(a)\leq\alpha\leq\bar{\alpha}(a)} ((1-p)+pF(b))(1-b)$$

From this expression we see that the expected utility depends on beliefs only through the probability of winning. Moreover, it is easy to see that for any b, the unique minimiser of utility is the highest p feasible, which is easily derived as $p(a) = 2\bar{\alpha}(a) = 1 - \frac{1}{2a}$. Let $\varepsilon(a) \equiv p(a) - \frac{1}{2}$. We can then describe the strategy of the high type in the unique equilibrium as mixing on the interval $[0, \frac{1}{2} + \varepsilon(a)]$, for $\varepsilon(a) = \frac{1}{2}(1 - \frac{1}{a})$, according to the distribution $F_a(b) = \frac{1}{2a-1}\frac{b}{1-b}$ which stochastically dominates $F_a(b)$ for $a = 1$.

Intuitively, ambiguity together with max-min preferences makes the high bidder believe that the other bidder is more likely a high type and so induces the bidder to bid more aggressively. This is the competition effect. Ambiguity over correlation through the prior implies that one can consider different possibilities for the type of the other bidder (and hence her bid). An ambiguity averse bidder will consider the worst case scenario and hence will believe that she faces the toughest possible competition, which will lead her to bid more aggressively, thus increasing the seller's revenue.

In the second-price auction, this effect does not exist as it is still weakly dominant to bid your value. Once the value is known, then the other's bid is not relevant. We then have:

Proposition 2.4.1. *For private value auctions: (i) In the first-price auction, the equilibrium distribution of bids when $a > 1$ first order stochastically dominates that of the case in which $a = 1$, with the seller's revenues increasing in a. (ii) In the second-price auction, the equilibrium distribution of bids when $a > 1$ is the same as that of the case in which $a = 1$.*

Proof. Appendix. □

Common values

We now analyse the common values case. In this case an additional effect of ambiguity emerges, as players learn about their value from the equilibrium behaviour of others. Thus, in addition to minimising their utility by envisaging a low probability of winning, bidders may also choose beliefs that minimise the value of the good conditional on winning. This *winner's value effect* will sometimes induce an opposite incentive compared with the competition effect; the competition effect can potentially induce bidders to bid more aggressively perceiving a tougher competition; the winner's value effect can induce players to bid less aggressively if they perceive lower valuations upon winning. Bidders can then shade their bids even further than they would in the absence of ambiguity.

Recall that bidders perceive correlation structures as depicted in Table 2, and so ambiguity is over the correlation parameters α_0, α_1. Let us consider the low type of bidder 1. Suppose she uses the equilibrium bid of the low type of bidder 2, denoted as $b_a(l)$, so that $b = b_a(l)$. In both the first-price and the second price auction then, her worst-case expected utility is:

$$\min_{(\alpha_0,\alpha_1)} \frac{1}{2} \Pr(l|l,(\alpha_0,\alpha_1))(E_{(\alpha_0,\alpha_1)}(v|l,l) - b_a(l)),$$

where

$$\Pr(l|l,(\alpha_0,\alpha_1)) = \alpha_0 + \alpha_1$$
$$E_{(\alpha_0,\alpha_1)}(v|l,l) = \frac{\alpha_1}{\alpha_1+\alpha_0}$$

And so:

$$\Pr(l|l,(\alpha_0,\alpha_1))(E_{(\alpha_0,\alpha_1)}(v|l,l) - b_a(l)) = (\alpha_0+\alpha_1)(\frac{\alpha_1}{\alpha_0+\alpha_1} - b_a(l))$$

Note that the competition effect will imply that to minimise utility one has to minimise the probability of winning, $\alpha_1 + \alpha_0$. On the other hand, the winner's value effect will demand that $\frac{\alpha_1}{\alpha_0+\alpha_1}$ is minimised, which is best achieved when α_1 is minimised and α_0 is maximised. When we put the two together, the expression becomes

$$\alpha_1(1 - b_a(l)) - \alpha_0 b_a(l)$$

which implies that the worst case scenario is indeed $(\bar{\alpha}_0, \underline{\alpha}_1)$, and so the winner's value effect dominates, as beliefs are chosen to minimise $E_{(\alpha_0,\alpha_1)}(v|l,l)$. In the first-price auction the low type will remain without rent, so that indeed $b = b_a(l) = E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,l)$. This is also the bid that arises in the second-price auction, where the players use the expected valuation that is based on both bidders receiving their own signal. We therefore have derived that the bid of the low type is lower than in the case with no ambiguity, in both types of auctions.

What about the high type? The intuition will be clearer in the second-price auction where pure strategies are used (but the results are qualitatively similar as Proposition 2.4.2 illustrates). Consider the high type of bidder 1. In the second-price auction, her expected utility when she uses her opponent's high type bid, so that $b = b_a(h) > b_a(l)$ is

$$\min_{(\alpha_0,\alpha_1)} \Pr(l|h,(\alpha_0,\alpha_1))(E_{(\alpha_0,\alpha_1)}(v|l,h) - b_a(l)) + (1/2)\Pr(h|h,(\alpha_0,\alpha_1))(E_{(\alpha_0,\alpha_1)}(v|h,h) - b_a(h))$$

where:

$$\Pr(l|h,(\alpha_0,\alpha_1)) = 1 - \alpha_1 - \alpha_0 \qquad E_{(\alpha_0,\alpha_1)}(v|l,h) = \frac{1-q-\alpha_1}{1-\alpha_1-\alpha_0}$$
$$\Pr(h|h,(\alpha_0,\alpha_1)) = \alpha_1 + \alpha_0 \qquad E_{(\alpha_0,\alpha_1)}(v|h,h) = \frac{2q-1+\alpha_1}{\alpha_1+\alpha_0}.$$

Putting these in the expression for the utility, we have:

$$\min_{(\alpha_0,\alpha_1)} [(\alpha_1+\alpha_0)(b_a(l) - \frac{1}{2}b_a(h)) + \frac{1}{2}(2q-1+\alpha_1) + 1 - q - \alpha_1 - b_a(l)].$$

The competition and winner's value effects now incorporate the possibility of winning both against the low type and against another high type, as well as the magnitude of the equilibrium bid of the low type. The terms $\frac{1}{2}(2q-1+\alpha_1)+1-q-\alpha_1$ express the expected valuation upon winning. It is easy to see that to minimise utility, one needs to maximise α_1. The first element, $(\alpha_1+\alpha_0)(b_a(l)-\frac{1}{2}b_a(h))$, denotes the expected payment given the two events in which the high type bidder can win. Recall that $b_a(l)=E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,l)=\frac{\alpha_1}{\alpha_1+\alpha_0}$, which is lower than but related to $\frac{(1-q)^2}{(1-q)^2+q^2}$. Thus, when q is low so that signals are imprecise, $b_a(l)$ can be relatively high, implying that to minimise utility one has to minimise α_1 and α_0, while when q is high, $b_a(l)$ is relatively low, implying that to minimise utility, one has to maximise both. In the case when q is high, both effects are then in tandem implying that the bidder chooses $(\bar{\alpha}_0,\bar{\alpha}_1)$. This means that the bid is $b_a(h)=E_{(\bar{\alpha}_1,\bar{\alpha}_0)}(v|h,h)$, which will be lower compared to the case of no ambiguity. Thus the winner's value upon winning against a high type, a more likely event, is indeed lower. In the case of a low q, where the competition and winner's value effect may clash, the latter dominates implying that the bidder chooses $(\underline{\alpha}_0,\bar{\alpha}_1)$. The effect now is more subtle; while this belief minimises $E_{(\alpha_0,\alpha_1)}(v|l,h)$, the winner's value when she wins against a low type, it can, for a low a, increase $E_{(\alpha_0,\alpha_1)}(v|h,h)$ beyond the canonical case. We then have:

Proposition 2.4.2. *For common value auctions: (i) Second price auctions: In the unique symmetric pure-strategy equilibrium, compared to the canonical model, the low type's bid decreases in a, and the high type's bid decreases in a for either high enough a or high enough q and increases in a when both a and q are low. (ii) First-price auctions: For sufficiently low a, in the unique symmetric equilibrium, compared to the canonical model, the bid of the low type decreases in a, the maximal bid of the high type increases in a, with the average bid of the high type increasing or decreasing depending on q.*

Proof. Appendix. □

It is interesting to note that different types use different joint information structures, and thus the model is not equivalent to one in which individuals simply have the wrong belief over correlation. In other words, ambiguity interacts with equilibrium behaviour in a non-trivial manner. This was harder to see in the private values case where the low type always bids her value; however in the common value case, the low type is utilising a different belief to minimise her utility compared with the high type, and moreover, as the result illustrates, the winner's value effect dominates so that the low type (and in most environments also the high type) end up lowering their bids compared with the canonical model.

Seller's revenue

In the case of private values we had seen that the seller's revenue increases with ambiguity over correlation, as bids increase (at least in the first-price auction). The case of common values yields a different result, as can be gleaned from the fact that bids may decrease with a. As we show in the Appendix, in the common value case, even when the high type increases her bid in the second price auction, compared to the case of no ambiguity, her expected utility evaluated at the true joint probability distribution will be higher compared to the canonical model. This arises as she increases her bid when q is low, which implies she is more likely to encounter the low type and hence pay the low type's bid which is lower than in the canonical model. Thus, considering the utility of bidders and the seller, we have:

Proposition 2.4.3. *For common value auctions: (i) Second price auction: The utility of both the high and the low type increases in a and the seller's revenue decreases with a. (ii) First price auction: For sufficiently low a, the seller's revenue decreases in a. (iii) The seller's revenue is higher in the first-price rather than the second-price auction.*

Proof. Appendix. □

In the common values case, ambiguity implies that bidders decrease their bids; it therefore allows them to gain a higher utility and implies that the seller's revenue decreases with a. As another illustration of this, consider the second-price auction equilibrium in the limit, where all information structures are feasible, and so ambiguity is very large. It is the limit of the unique symmetric pure-strategy equilibria that we characterise in the Appendix, and it is characterised by $b_a(l) = 0$ and $b_a(h) = q$.[11] Specifically, the high type believes that signals are fully correlated and so $E_{(q,1-q)}(v|h,h) = E(v|h)$. These bids are the lowest among all equilibria and the seller's revenue is therefore substantially lower compared with the canonical model.

This result is different than in the private values case. Intuitively, in the former case ambiguity increased the competitiveness of the bids, and thus these increased. In the common value case, ambiguity in general induced the bidders to minimise their utility by minimising the value they expect to gain when they win, implying lower bids at least on average.

The intuition for the result that the seller's revenue is higher in the first-price auction is as follows. In the second-price auction, an individual's payment depends on the other's bid and as a result, there are more elements in her utility in which her beliefs play a role. For the case of no ambiguity, this implies that she conditions her behaviour on more information, which increases the seller's revenue.[12] For the case of ambiguity, this implies that individuals have more possibilities to condition on their worst-case beliefs, which decreases the seller's revenue. Intuitively, the competition effect is more pronounced in the first price auction than in the second price auction. This is because in the first price auction the payment is directly related to the bid.

2.5 Optimal auctions

Our analysis here will focus on two results. First, consistent with results in the previous section, the seller's maximum revenue will increase with ambiguity in the private values case but will decrease with ambiguity in the common values case. Second, the type of optimal auction will also change when we switch from private to common values, and specifically, the seller will not necessarily fully insure the bidders.

Indeed the key issue when considering optimal auctions under ambiguity is the level of insurance provided by the seller to the bidders. Under independent private values, Bose, Ozdenoren and Pape (2006) show that the optimal mechanism fully insures the bidders against ambiguity. In Bose, Ozdenoren and Pape (2006) bidders believe that valuations are independently drawn, but face ambiguity regarding the particular distributions of valuations that others have. In contrast, in our model bidders know the marginal distribution of valuations but face ambiguity about the joint distribution of valuations. In Section 4.2 we replicate the full insurance result in our setting. When reporting truthfully, the buyers are fully insured against ambiguity. However, the competition

[11] This is supported by the low type believing $\alpha_0 = q$ and $\alpha_1 = 0$, and the high type believing $\alpha_0 = q$ and $\alpha_1 = 1 - q$.

[12] With private values and ambiguity over the prior, Lo (1998) shows that the first-price auction dominates the second-price auction in some environments.

effect makes deviations less attractive, and the seller is able to exploit this by asking the high type to pay more compared to the standard case without ambiguity.

In the common value case, however, the seller does not necessarily provide full insurance. We show that when ambiguity is small or the signals sufficiently precise, the optimal mechanism fully extracts rent (i.e. participation constraints are binding), using side bets with the low type to deter the high type from deviating. These side bets expose the low type to ambiguity, and as a result the seller cannot extract full surplus.

Moreover, the seller will allocate the object to the buyer with the highest signal, which exposes the high type to ambiguity as well. As a result, the seller is then able to insure the high type against the competition effect but not the winner's value effect. Specifically, as we show in Section 4.3, when the two players receive different signals, the seller allocates the object to the high type; the winner's value effect then implies that the high type will undervalue the object, but the seller is able to partially insure the high type by also asking her to pay more when her opponent is low type. Under optimal transfers the high type does not care whether her opponent has received the high or the low signal—that is, the high type is insured against the competition effect. On the other hand when ambiguity is large and the signals imprecise, the seller fully insures the buyers, leaving the high type with positive rents.

The seller's problem

We now formalise the seller's problem. A direct mechanism (x, t) is an allocation rule $x : \{l, h\}^2 \mapsto \Delta\{1, 2\}$ and a transfer rule $t : \{l, h\}^2 \mapsto R^2$. Let $U_i^{\boldsymbol{\alpha}}(s', s)$ be i's utility from reporting s' when i's signal is s, given that the information structure is $\boldsymbol{\alpha}$. A direct mechanism is maxmin incentive compatible if for all $s \in \{l, h\}$:

$$\min_{\boldsymbol{\alpha}} U_i^{\boldsymbol{\alpha}}(s, s) \geq \min_{\boldsymbol{\alpha}} U_i^{\boldsymbol{\alpha}}(s', s)$$

for all $s' \in \{l, h\}$. The revelation principle applies in this setting as long as we make the following assumption:

No-hedging: The utility from playing the mixed strategy $\sigma \in \Delta\{l, h\}$ is $\sum_{s'} \sigma(s') \min_{\boldsymbol{\alpha}} U_i^{\boldsymbol{\alpha}}(s', s)$ (as opposed to $\min_{\boldsymbol{\alpha}} \sum_{s'} \sigma(s') U_i^{\boldsymbol{\alpha}}(s', s)$).

This assumption is standard in the literature on mechanism design with maxmin agents (see for example, Bose, Ozdenoren and Pape (2006) or Wolitzky (2016)). In what follows, we restrict attention to maxmin incentive compatible direct mechanisms.

Private values

In the private value case, $q_i(s_i = l | v_i = L) = q_i(s_i = h | v_i = H) = 1$. For each signal $s \in \{l, h\}$, let $v^s \in \{L, H\}$ be the valuation associated with the signal. That is, let $v^l = L$ and $v^h = H$. Let p denote a buyer's belief that her opponent has received the same signal as her. The ePMI constraints imply, following Table 1 in Section 3, that $\frac{1}{2a} \leq p \leq 1 - \frac{1}{2a}$. An optimal mechanism solves the seller's problem:

$$\max_{x_i, t_i} \sum_{i=1}^{2} t_i(l, l) + t_i(l, h) + t_i(h, l) + t_i(h, h)$$

subject to:

$$\min_{p\in[\frac{1}{2a},1-\frac{1}{2a}]} p\left[v^s x_i(s,s)-t_i(s,s)\right]+(1-p)\left[v^s x_i(s,s')-t_i(s,s')\right]$$
$$\geq \min_{p\in[\frac{1}{2a},1-\frac{1}{2a}]} p\left[v^s x_i(s',s)-t_i(s',s)\right]+(1-p)\left[v^s x_i(s',s')-t_i(s',s')\right]$$

$$\min_{p\in[\frac{1}{2a},1-\frac{1}{2a}]} p\left[v^s x_i(s,s)-t_i(s,s)\right]+(1-p)\left[v^s x_i(s,s')-t_i(s,s')\right]\geq 0$$

for all i, and for all $s,s'\in\{l,h\}$.

As in Bose et al (2006), when values are private, the seller finds it optimal to fully insure the buyers against ambiguity and is able to extract full surplus when the ambiguity is large:

Proposition 2.5.1. *When buyers have private values:*

(i) Any optimal mechanism is a full insurance mechanism

(ii) The optimal allocation rule depends on L: if $L<\frac{1}{1+a}$, it is not optimal to sell to buyers with the low valuation, but if $L>\frac{1}{1+a}$, the seller always allocates the object.

(iii) The seller's revenue is increasing in ambiguity, and as $a\to\infty$, converges to full surplus.

Proof. Appendix. □

In the Appendix, we fully characterise the set of optimal mechanisms under private values.

Common values

In the common values case, the possible joint information structures are described in Table 2 in Section 3. The seller's problem is:

$$\max_{x_i,t_i}\frac{1}{2}\left(q^2+(1-q)^2\right)\left[\sum_{i=1}^{2}t_i(l,l)+t_i(h,h)\right]+q(1-q)\left[\sum_{i=1}^{2}t_i(l,h)+t_i(h,l)\right]$$

subject to incentive and participation constraints:

$$\min_{\alpha_0,\alpha_1}\alpha_1 x_i(l,l)+(1-q-\alpha_1)x_i(l,h)-(\alpha_0+\alpha_1)t_i(l,l)-(1-\alpha_0-\alpha_1)t_i(l,h)$$
$$\geq\min_{\alpha_0,\alpha_1}\alpha_1 x_i(h,l)+(1-q-\alpha_1)x_i(h,h)-(\alpha_0+\alpha_1)t_i(h,l)-(1-\alpha_0-\alpha_1)t_i(h,h)$$

$$\min_{\alpha_0,\alpha_1}(1-q-\alpha_1)x_i(h,l)+(2q-1+\alpha_1)x_i(h,h)-(1-\alpha_0-\alpha_1)t_i(h,l)-(\alpha_0+\alpha_1)t_i(h,h)$$
$$\geq\min_{\alpha_0,\alpha_1}(1-q-\alpha_1)x_i(l,l)+(2q-1+\alpha_1)x_i(l,h)-(1-\alpha_0-\alpha_1)t_i(l,l)-(\alpha_0+\alpha_1)t_i(l,h)$$

$$\min_{\alpha_0,\alpha_1}\alpha_1 x_i(l,l)+(1-q-\alpha_1)x_i(l,h)-(\alpha_0+\alpha_1)t_i(l,l)-(1-\alpha_0-\alpha_1)t_i(l,h)\geq 0$$

$$\min_{\alpha_0,\alpha_1}(1-q-\alpha_1)x_i(h,l)+(2q-1+\alpha_1)x_i(h,h)-(1-\alpha_0-\alpha_1)t_i(h,l)-(\alpha_0+\alpha_1)t_i(h,h)\geq 0$$

For a given a, the optimal mechanism will depend on two cutoff values of q, which we now define. Let $q^*(a)$ be the solution to $q^2+(1-q)^2-\underline{\alpha}_0-\underline{\alpha}_1=\underline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_0+\overline{\alpha}_1-1$ that lies between $\frac{1}{2}$

and 1, and let $q^{**}(a)$ be the solution to $q^2 + (1-q)^2 - \underline{\alpha}_0 - \underline{\alpha}_1 = 3q - 2 + \underline{\alpha}_1 + \overline{\alpha}_1$ that lies between $\frac{1}{2}$ and 1. We derive the explicit expressions for $q^*(a), q^{**}(a)$ in the Appendix. For $1 < a < \infty$, $q^*(a) < q^{**}(a)$.

We then have:

Proposition 2.5.2. *When buyers have common values:*

(i) When $q \leq q^*(a)$, an optimal mechanism allocates the good with equal probability for each player disregarding their type, and a transfer of $\frac{1}{2}(1-q)$. The revenue of the seller is $1-q$, both types are fully insured, and the high type earns positive rents.

(ii) When $q \geq q^{**}(a)$, the optimal mechanism allocates the good to the high type and with equal probability to each player if both are of the same type. Transfers are such that the high type is partially insured, the seller makes side bets with the low type, and the buyers earn no rents.

(iii) When $q^*(a) < q < q^{**}(a)$, the optimal mechanism allocates the good to the high type and with equal probability to each player if both are of the same type. There are no side bets with the low type, both types are partially insured, and the high type earns positive rents.

(iv) As $a \to \infty$, both $q^*(a)$ and $q^{**}(a)$ converge to $\frac{1}{2}(3-\sqrt{3})$, and as $a \to 1$, both $q^*(a)$ and $q^{**}(a)$ converge to $\frac{1}{2}$.

(v) Seller's revenue (weakly, and sometimes strictly) decreases with a.

Proof. Appendix. □

When $q \leq q^*(a)$, an implementation of the optimal mechanism is for the seller to first choose each buyer with equal probability, and then sell to the chosen buyer at price $1-q$. Since the decision to sell is not based on the signal realisation, the good is worth $1-q$ to the low type and q to the high type. Thus, the high type earns positive rents in equilibrium. Note that this mechanism is efficient, and that given the seller's design, ambiguity is not relevant or does not arise in equilibrium.

When $q \geq q^{**}(a)$, the participation constraint of the high type is binding: it is optimal to fully extract rent. The seller engages in a side bet with the low type to prevent the high type from deviating. Unlike in the classical case, side bets are costly to the seller, so the seller uses the smallest bet that is sufficient to prevent the high type from deviating. To reduce this cost further, the seller allocates the good to the high type when the players receive different signals, which generates endogenous ambiguity over the expected value of the good. The seller is able to partially insure the high type by asking her to pay more when the other player has received a low signal. In this case, the expected payment from the high type is:

$$T_i^h = \frac{1}{2}(1-(1-q)^2) - \frac{1}{2}(q^2 - \underline{\alpha}_0).$$

The expected payment from the low type is:

$$T_i^l = \frac{\underline{\alpha}_1}{2} - \left[q^2 + (1-q)^2 - \underline{\alpha}_0 - \underline{\alpha}_1\right] \frac{1 - q - \overline{\alpha}_1 - \underline{\alpha}_1}{2(\overline{\alpha}_0 + \overline{\alpha}_1 + \underline{\alpha}_0 + \underline{\alpha}_1 - 1)}.$$

The low type chooses α_0 and α_1 both to minimise the perceived surplus from winning the object and to maximise the perceived value of the transfers. Note that in the optimal mechanism, the belief of the high type regarding α_1 is irrelevant. The high type does not care about the type of her opponent: she gets the same utility from any belief about α_1. In this sense, she is insured against the competition effect. On the other hand she believes that $\alpha_0 = \underline{\alpha}_0$. This gives rise to the winner's value effect: her expectation of the value conditional on winning, $\frac{1-\alpha_1}{2-\underline{\alpha}_0-\alpha_1}$, is lower than true conditional expectation (for any belief $\alpha_1 \in [\underline{\alpha}_1, \overline{\alpha}_1]$). When a converges to 1, only this case remains.

The intuition why participation constraints must bind when a is small is as follows. If the participation constraint of the high type is slack, the seller can achieve a first order increase in revenue by increasing the payment of the high type. In order to ensure that the incentive constraint is not violated, the seller can increase $t_i(l,h)$ and decrease $t_i(l,l)$ in such a way that keeps the low type indifferent, but lowers the high type's utility from deviating. Since the low type may have different beliefs to the seller, these changes in transfers may decrease the seller's revenue; however, as the ambiguity becomes small, this fall in revenue converges to zero. On the other hand, the increase in revenue from increasing the payment of the high type is fixed.

When $q^*(a) < q < q^{**}(a)$, it is optimal to allocate the good to the high type when the players receive different signals, but it is not optimal to conduct side bets with the low type. Instead, the high type earns positive rents in equilibrium in order to satisfy the incentive constraint. The winner's value effect implies that the buyers underestimate the value of the object conditional on winning. However, the seller partially insures both types, so that the competition effect does not arise.[13] This interval shrinks as a goes to either 1 or ∞.

If buyers are fully insured against ambiguity, then the seller's revenue is constant in ambiguity. However, in the cases where buyers are not fully insured, the seller's revenue is strictly decreasing in ambiguity. The intuition is that as ambiguity increases, the larger winner's value effect forces the seller to decrease transfers. Thus, the seller's revenue is always (weakly) decreasing in ambiguity.

Note that Crémer and McLean (1988) show that some of the conclusions from the analysis of optimal auctions with independent private values are not robust. For example, since surplus extraction is possible when signals are correlated, the optimal mechanism is efficient and leaves no rents to the buyers. Proposition 5 shows that these results continue to hold for a close to 1.[14] On the other hand, when a is large, it is possible for buyers to earn positive rents in the optimal mechanism. Note that in this environment, it is always possible to fully extract rent (see Renou (2015)); however, as we have argued, rent extraction conflicts with full insurance and is not necessarily optimal when ambiguity is large.

2.6 Extensions

We now extend the model of Section 3 in two ways. We first consider the case of many bidders and then that of continuous valuations. In each of these extensions we consider different models of correlation which illustrate the flexibility of the framework. The key results of Section 3 extend to these environments as well: seller's revenue decreases with ambiguity for common value auctions.

Many bidders

Our analysis above focused on the case of two bidders. In this section we consider a model with common values and many bidders. Naturally when extending the model to more than two bidders, many correlation patterns can be considered; we now extend the model in the simplest way that

[13]For the low type, any belief about α_0 gives the same utility, but she believes that $\alpha_1 = \underline{\alpha}_1$, so that her expected value of the object conditional on winning, $\frac{\alpha_1}{\alpha_0+\underline{\alpha}_1}$ is smaller than the true conditional expectation for any belief $\alpha_0 \in [\underline{\alpha}_0, \overline{\alpha}_0]$.

[14]The set of optimal mechanisms when $a = 1$ is large. As $a \to 1$, the optimal mechanism described in Proposition 5, which is the unique symmetric mechanism when a is close to 1, converges to an optimal mechanism for the case when $a = 1$.

also maintains symmetry and anonymity when a single bidder considers the correlation among herself and all other bidders.

Specifically, consider a population of n bidders, who consider distributions of the following form: with probability α_0, all bidders receive the same signal in state 0, and with probability α_1, all bidders receive the same signal in state 1. This signal, $s^* \in \{l, h\}$, is drawn from the same marginal distribution as before; that is, $\Pr(s^* = h|v = 1) = \Pr(s^* = l|v = 0) = q > \frac{1}{2}$. With the remaining probability, each bidder draws a (conditionally) independent signal, again, with $\Pr(s_i = h|v = 1) = \Pr(s_i = l|v = 0) = q > \frac{1}{2}$ for all i. Bidders have ambiguity then over α_1 and α_0, as before.

To see that similar considerations are involved, consider a low type. Intuitively, a low type wins only when she faces $n - 1$ low types. Minimising utility implies as before that she will minimise her winning probability in state 1 and maximise it in state 0. This implies that she sets up the correlation among bidders to be the highest in state 0, as positive correlation implies the highest probability of having a vector of identical types. Alternatively, she uses the belief that correlation is lowest in state 1; with independent signals the probability of any specific vector of valuations is the lowest. But of course this means that conditioning on a vector of low types implies that state 0 is more likely, inducing a strictly lower bid, in line with the "winner's value effect". We are then able to show:

Proposition 2.6.1. *When ambiguity is not too large, there exists a symmetric equilibrium in the first-price auction. Moreover, there exists $\bar{n}$, such that for all $n > \bar{n}$, in this equilibrium the high type conditions her bid on the belief $(\alpha_0, \alpha_1) = (0, \bar{\alpha}_1)$, and the low type conditions her bid on the belief $(\alpha_0, \alpha_1) = (\bar{\alpha}_0, 0)$. The (expected) bids of both types are lower than in the independent case, and the seller's revenues decrease with ambiguity.*

Proof. Appendix. □

For the high type (who uses a mixed strategy in equilibrium), the considerations are more complicated as she needs to consider many vectors of types, each affecting both the probability of winning and her valuation. For an h type, perceiving more correlation shifts the weight from the set of events in which some types are also l, to the event in which they are all correlated (all h). Consider for example the low state and hence α_0, and its effect on the expected bid. Increasing α_0 means reducing the probability of paying in all events in which there is at least one l type and increasing the probability of paying when all are h. However the first set of events (where at least one bidder is an l) happens with a much higher probability when n is large; thus increasing α_0 reduces the average bid and as a result increases utility. Thus to minimise utility we have to reduce correlation in the low state. Similar intuition arises in the high state. This belief implies that the expected bid is lower than in the benchmark independent model, and as a result, seller's revenue is lower as well. While we can verify analytically that the above holds also for small n (e.g., $n = 3$), we have derived an analytical proof for large n only.

Note that we focus in this case on a first-price auction as symmetric equilibria in the second-price auction may not exist in a common value environment with discrete types and many bidders.[15]

[15] This is a general issue that does not relate to our specific model.

Continuous signals

We show that the results of Section 3 are robust to the case of a continuum of signals. In the common-value case, let the state of the world be $v \in \{0, 1\}$, with an equal prior. We revert to the case of two bidders, where each now receives a signal $s^i \in [0, 1]$ about the state of the world. The marginal distributions determining the signals given the state of the world, $g_v(s)$ for each player, are known to the players. To simplify, let $g_0(s) = 2(1-s)$ and $g_1(s) = 2s$ (more generally we require in the Appendix that $g_0(s)$ is decreasing and $g_1(s)$ is increasing, so that that $G_0(s)$ is concave and $G_1(s)$ is convex). Note that $G_0(s) > G_1(s)$ for all interior s, and hence MLRP is satisfied too.

Individuals have ambiguity over a set of joint distributions in each state $v \in \{0, 1\}$. We use a simple set of joint distributions, the F-G-M transformation (copula), which was introduced by Morgenstern in 1956. Specifically, given $g_v(s)$, we have:

$$f_v(\mathbf{s}) = [1 + \lambda_v(2G_v(s_1) - 1)(2G_v(s_2) - 1)]g_v(s_1)g_v(s_2).$$

For this to be a distribution, for any v we need $|\lambda_v| \leq 1$.[16] Note that when $\lambda_v > 0$ we have positive correlation of signals in state v while when $\lambda_v < 0$ we have negative correlation. When signals are conditionally independent, we have $\lambda_v = 0$ for all v. Adding ePMI constraints, we then have:

$$\lambda_v \in [\frac{1}{a} - 1, 1 - \frac{1}{a}] \text{ for } v \in \{0, 1\}.$$

We analyse a second-price auction and show very similar results to the one derived in Section 3. Let us first write the utility of a player for each bid b. This is

$$U(s^1, b) \propto \min_{\lambda}(\int_0^z (1 - b(s'))f_1(s^1, s')ds' - \int_0^z b(s')f_0(s^1, s')ds')$$

where $b(s')$ is the bid used by the other player and $z = b^{-1}(b)$. In the Appendix we show that when the level of ambiguity a is small enough, there exists a symmetric equilibrium in which

$$b(s) = E^{\lambda^*(s)}(v|s, s) = \frac{[1 + \lambda_1^*(s)(2G_1(s) - 1)^2]g_1^2(s)}{[1 + \lambda_1^*(s)(2G_1(s) - 1)^2]g_1^2(s) + [1 + \lambda_0^*(s)(2G_0(s) - 1)^2]g_0^2(s)}. \tag{2.1}$$

where $\boldsymbol{\lambda}^*(s)$ minimises the equilibrium utility for each type. To do so, a player needs to consider how correlation affects both the competition and the winner's value effects, at each state.

As before, low types (specifically, below the median of G_0)[17] minimise their utility by postulating the highest possible positive correlation at state 0 and the highest possible negative correlation at state 1, given the strong winner's value effect. They therefore minimise their valuation upon winning (as they only win when the other player has even lower signals). This "winner's value effect" implies underbidding.

For high types, positive correlation in state 1 and in state 0 decreases their utility as it implies paying a higher bid (while their probability of winning is relatively high in any case). This implies that they choose the maximum positive correlation in both states. This was also the case in the discrete case and again implies underbidding as conditional on their signal, and as $G_0(s) > G_1(s) > \frac{1}{2}$ for these types, $f_0(s, s) > f_1(s, s)$ for any fixed positive $\boldsymbol{\lambda}$. Thus $b(s) = E^{\lambda^*(s)}(v|s, s) < E^{(0,0)}(v|s, s)$, where $E^{(0,0)}(v|s, s)$ is the bid in the standard model with conditionally independent signals. We therefore have:

[16] This implies that the highest correlation coefficient in this family is $\frac{1}{3}$ in absolute value. See Schucany, Parr and Boyer (1978).

[17] The median is important as the F-G-M copula is of the form $2G - 1$.

Proposition 2.6.2. *When the level of ambiguity a is small enough, there exists a symmetric equilibrium in the second-price auction in which $b(s)$ is as defined in (2.1), where $\boldsymbol{\lambda}^*(s)$ minimises the utility of type s in equilibrium, and the seller's revenues decrease with ambiguity.*

Proof. Appendix. □

In the Appendix we generalise for other symmetric $g_v(s)$, and characterise a sufficient condition for the marginal distribution function for which revenues decrease more generally. We also consider in the Appendix for completeness the private value case, where we assume that valuations are drawn from $[0, 1]$ according to a uniform distribution, and individuals believe that the joint distribution is $f(\mathbf{v}) = 1 + \lambda(2v_i - 1)(2v_j - 1)$, for $\lambda \in [\frac{1}{a} - 1, 1 - \frac{1}{a}]$. We characterise a symmetric equilibrium and show that bids in the first-price auction are uniformly higher compared to the case without ambiguity (recall that bids in the second-price auction are not affected by ambiguity).

2.7 Conclusion

We have constructed a framework in which we can analyse ambiguity over correlation in information structures of bidders both in the private and in the common value auctions. We have illustrated in this paper that ambiguity over correlation induces different results in the private values versus the common values auctions. Specifically, while seller's revenue increases in the private value auctions, it decreases in the common value auctions. The key insight of the analysis is that in the common value auctions players choose worst-case beliefs that amount to minimising the value of the good conditional on winning. This also leads to the optimal auction providing less than full insurance to bidders.

Appendix 2.A Proofs for Section 2.4

Proof of Proposition 2.4.1. To prove Proposition 2.4.1, we first characterise the equilibrium in the first price auction under private values and show that in this equilibrium, the seller's revenue increases in a. Then we note that in the symmetric equilibrium of the second price auction, under private values, equilibrium behaviour is not affected by ambiguity.

We consider an equilibrium where the low type bids L and the high type mixes on the support $[L, \overline{b}]$, according to a distribution function $F(b)$. It can be shown that this equilibrium is unique.[18] First note that the low type gets zero utility in equilibrium; any bid higher than L yields negative utility, and any bid below L yields 0 utility.

As in the main text, the expected utility from bidding b for the high type is:

$$\min_{p=2\alpha, \underline{\alpha}(a) \leq \alpha \leq \bar{\alpha}(a)} ((1-p) + pF(b))(1-b),$$

[18]For uniqueness, first observe that there cannot be pure strategy equilibria, and that in any equilibrium the low type cannot earn positive rents. In any equilibrium, the low type of both players must bid L. Suppose that the low type of player 1 bids strictly less than L with positive probability; then the low type of player 2 can get positive rents. But this contradicts the fact that the low type cannot earn positive rents in equilibrium. Thus assume that each low type bids at least L (so the utility of the low type is at most zero). Suppose that at least one low type bids strictly more than L with positive probability, and without loss of generality assume that the low type of player 1 wins the auction with positive probability with a bid strictly higher than L (and receiving strictly negative utility). Then the low type of player 1 has a profitable deviation to bidding L with probability 1. Thus assume that both low types bid L in equilibrium. The bottom of the support of the mixed strategy for the high type must be L, otherwise a high type who is supposed to bid just above the bottom of the support can bid just below the support and reduce expected payment without changing the probability of winning. The indifference condition then uniquely determines the equilibrium.

which is minimised by maximising α. As the ePMI constraints are:

$$\frac{1}{4}a = \underline{\alpha}(a) \leq \alpha \leq \bar{\alpha}(a) = \frac{1}{2}(1 - \frac{1}{2}a),$$

utility is minimised at $\bar{\alpha}(a) = \frac{1}{2} - \frac{1}{4a}$ which implies a probability that the opponent has a high value of $1 - \frac{1}{2a}$.

Now we compute the highest bid, $\bar{b}$, in the support of $F(b)$. For the highest bid we have:

$$(1 - \bar{b}) = (\frac{1}{2a})(1 - L) \Leftrightarrow \bar{b} = 1 - \frac{1}{2a}(1 - L).$$

Note that $\bar{b}$ is increasing in a. Computing $F_a(b)$ (making it explicit that $F_a(b)$ depends on a), we have:

$$F_a(b) = \frac{b - L}{(2a - 1)(1 - b)}$$

Note that for $a' > a$, $F_{a'}(b)$ first order stochastically dominates $F_a(b)$ and thus seller's revenues are higher when $a > 1$ and increase in a.

Note that in the second price auction, under private values, it is still a dominant strategy to bid one's value, and hence for all a, $b_a(l) = L$ and $b_a(h) = H$ is a symmetric equilibrium. of the second price auction. □

Proof of Propositions 2.4.2 and 2.4.3. We first characterise the symmetric equilibrium in the second price auction. Then we characterise the equilibrium in the first price auction, and finally we compare the seller's revenue from the two auction formats.[19]

Equilibrium in the second price auction

Lemma 2.A.1. *In the common value second-price auction, the unique symmetric pure-strategy equilibrium satisfies:*

1. *The low type bids $b_a(l) = E_{(\bar{\alpha}_0, \underline{\alpha}_1)}(v|l, l)$, a bid that decreases in a;*
2. *For all $a \leq \frac{q}{1-q}$, there exist cutoffs $\underline{q}, \bar{q}$, with $0.5 < \underline{q} < \bar{q} < 1$, where:*
   a) *For $q \in (0.5, \underline{q})$, the high type bids $b_a(h) = E_{(\underline{\alpha}_0, \bar{\alpha}_1)}(v|h, h)$, a bid that increases with a.*
   b) *For $q \in (\bar{q}, 1)$, the high type bids $b_a(h) = E_{(\bar{\alpha}_0, \bar{\alpha}_1)}(v|h, h)$, a bid that decreases with a.*
   c) *For $q \in (\underline{q}, \bar{q})$, the high type bids $b_a(h) = E_{(\alpha_0, \bar{\alpha}_1)}(v|h, h)$, where α_0 satisfies $E_{(\alpha_0, \bar{\alpha}_1)}(v|h, h) = 2E_{(\bar{\alpha}_0, \underline{\alpha}_1)}(v|l, l)$, a bid that decreases with a.*
3. *For all $a \geq \bar{a}(q) \geq \frac{q}{1-q}$, the high type bids $b_a(h) = E_{(\bar{\alpha}_0, \bar{\alpha}_1)}(v|h, h)$.*[20]

Proof of Lemma 2.A.1. We prove this Lemma by first conjecturing an equilibrium bid for the low type. In all the equilibria we consider, the low type has the same equilibrium beliefs: $(\alpha_0, \alpha_1) = (\overline{\alpha}_0, \underline{\alpha}_1)$, and hence the bid is $b_a(l) = E_{(\overline{\alpha}_0, \underline{\alpha}_1)}[v|l, l]$. Then we characterise the equilibrium beliefs and bidding behaviour for the high type, and verify that neither type has an incentive to deviate. We separate the analysis into two cases: equilibria where there is over-bidding for the high type

[19] For simplicity we set $L = 0$ but the proof is easily extended to $L > 0$.

[20] When q is not too small, $\bar{a}(q) = \frac{q}{1-q}$. When q is sufficiently close to 0.5, $\bar{a}(q) > \frac{q}{1-q}$ and symmetric pure-strategy equilibria may not exist in the region $[\frac{q}{1-q}, \bar{a}(q)]$.

(relative to baseline case of no ambiguity), and equilibria where there is under-bidding for the high type. As in the statement of the Lemma, overbidding arises for small a and small q (which implies that in equilibrium $b_a(l) > \frac{1}{2}b_a(h)$), and underbidding arises for large a or large q (which implies $b_a(l) \leq \frac{1}{2}b_a(h)$).

We consider monotone equilibria where $b_a(l) < b_a(h)$, $b_a(l) = E_{(\alpha_0,\alpha_1)}(v|l,l)$ for some (α_0, α_1), and $b_a(h) = E_{(\alpha_0',\alpha_1')}[v|h,h]$ for some (α_0', α_1').[21]

Let $U_i^{\boldsymbol{\alpha}}(b;s)$ be Player i's utility from bidding b after receiving signal s, given that Player $-i$ is following the equilibrium strategy $(b_a(l), b_a(h))$:

$$U_i^{\boldsymbol{\alpha}}(b;s) = \begin{cases} 0 & \text{if } b < b_a(l) \\ \frac{1}{2}Pr(l|s,\boldsymbol{\alpha})(E_{\boldsymbol{\alpha}}[v|l,s] - b_a(l)) & \text{if } b = b_a(l) \\ Pr(l|s,\boldsymbol{\alpha})(E_{\boldsymbol{\alpha}}[v|l,s] - b_a(l)) & \text{if } b \in (b_a(l), b_a(h)) \\ Pr(l|s,\boldsymbol{\alpha})(E_{\boldsymbol{\alpha}}[v|l,s] - b_a(l)) + \frac{1}{2}Pr(h|s,\boldsymbol{\alpha})(E_{\boldsymbol{\alpha}}[v|h,s] - b_a(h)) & \text{if } b = b_a(h) \\ Pr(l|s,\boldsymbol{\alpha})(E_{\boldsymbol{\alpha}}[v|l,s] - b_a(l)) + Pr(h|s,\boldsymbol{\alpha})(E_{\boldsymbol{\alpha}}[v|h,s] - b_a(h)) & \text{if } b > b_a(h) \end{cases}$$

Consider the following assumption:

No-hedging: The utility for type s of Player i from deviating to the mixed strategy represented by the cdf F is $\int_b \min_{\boldsymbol{\alpha}} U_i^{\boldsymbol{\alpha}}(b,s)dF(b)$ (as opposed to $\min_{\boldsymbol{\alpha}} \int_b U_i^{\boldsymbol{\alpha}}(b,s)dF(b)$).

Under the "no hedging" condition deviations to mixed strategies will have lower utility, and thus equilibria are easier to sustain. We use this to characterise equilibria for large values of a. Note that all equilibria derived without the "no hedging" condition will remain equilibria under the "no hedging" condition.[22]

Consider first the low type. For any bid $b \in [b_a(l), b_a(h))$, we have:

$$\begin{aligned} &\min_{(\alpha_0,\alpha_1)} \rho Pr(l|l,(\alpha_0,\alpha_1)) \left(E_{(\alpha_0,\alpha_1)}[v|l,l] - b_a(l)\right) \\ =& \min_{(\alpha_0,\alpha_1)} \rho\alpha_1 - \rho(\alpha_1+\alpha_0)b_a(l) \end{aligned}$$

where $\rho = 1$ if $b > b_a(l)$ and $\frac{1}{2}$ otherwise. This is minimised by $(\bar{\alpha}_0, \underline{\alpha}_1)$. Thus the conjectured equilibrium bid is $b_a(l) = E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,l)$. This will be the case for all equilibria considered.

Equilibrium with overbidding for the high type

Consider now the high type. Consider the case of an equilibrium that satisfies $b_a(l) > \frac{1}{2}b_a(h)$. Bidding $b = b_a(h)$ yields:

$$\min_{(\alpha_0,\alpha_1)} Pr(l|h,(\alpha_0,\alpha_1)) \left(E_{(\alpha_0,\alpha_1)}[v|l,h] - b_a(l)\right) + \frac{1}{2}Pr(h|h,(\alpha_0,\alpha_1)) \left(E_{(\alpha_0,\alpha_1)}[v|h,h] - b_a(h)\right)$$

[21] Note that in any pure strategy equilibrium where $b_a(l) < b_a(h)$, we must have $b_a(l) = E_{(\alpha_0,\alpha_1)}(v|l,l)$ for some (α_0,α_1) and $b_a(h) = E_{(\alpha_0',\alpha_1')}[v|h,h]$ for some (α_0',α_1'). For example, suppose that $b_a(l) > E_{(\alpha_0,\alpha_1)}(v|l,l)$ for all (α_0,α_1). Then the low type gets negative utility, which contradicts $b_a(l)$ being an equilibrium strategy. Suppose that $b_a(l) < E_{(\alpha_0,\alpha_1)}(v|l,l)$ for all (α_0,α_1). Then the equilibrium utility for the low type is $\min_{(\alpha_0,\alpha_1)} \frac{1}{2}Pr(l|l,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|l,l] - b_a(l)\right) > 0$, but the low type has a profitable deviation to $b_a(l)+\varepsilon$, which yields utility $\min_{(\alpha_0,\alpha_1)} Pr(l|l,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|l,l] - b_a(l) - \varepsilon\right)$. The argument for the high type is similar.

[22] If there does not exist a profitable deviation in the absence of the 'no hedging' condition, there cannot exist a profitable deviation under the 'no hedging' condition, since under the 'no hedging' condition the payoff from the equilibrium (pure) strategy is the same, whereas the payoff to any mixed strategy is weakly lower.

where the optimal (α_0, α_1) is the same as the one that solves

$$\min_{(\alpha_0,\alpha_1)} -\alpha_0 \left(\frac{b_a(h)}{2} - b_a(l)\right) - \alpha_1 \left([1 - b_a(l)] - \frac{[1 - b_a(h)]}{2}\right)$$

Since by assumption $b_a(l) > \frac{1}{2} b_a(h)$, the payoff is minimised by $(\underline{\alpha}_0, \bar{\alpha}_1)$. Thus the conjectured equilibrium bid for the high type is $b_a(h) = E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h, h]$. Note that $E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h, h] = \frac{2q-1+a(1-q)^2}{a(1-q)^2+\frac{1}{a}(1-q)^2+2q-1}$ is increasing in a. The equilibrium will hold then only if $E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l, l) > \frac{1}{2} E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h, h]$.

Note that the equilibrium payoff will be $Pr(l|h, (\alpha_0, \alpha_1)) \left(E_{(\alpha_0,\alpha_1)}[v|l, h] - b_a(l)\right)$. This has to be non negative and thus under $(\underline{\alpha}_0, \bar{\alpha}_1)$, we must have $E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|l, h] \geq b_a(l)$.

We now consider deviations. For the low type, the payoff from any mixed strategy is:

$$\min_{(\alpha_0,\alpha_1)} \rho_0 Pr(l|l, (\alpha_0, \alpha_1)) \left(E_{(\alpha_0,\alpha_1)}[v|l, l] - b_a(l)\right) + \rho_1 Pr(h|l, (\alpha_0, \alpha_1)) \left(E_{(\alpha_0,\alpha_1)}[v|h, l] - b_a(h)\right)$$

where $0 \leq \rho_1 \leq \rho_0 \leq 1$. Under the information structure $(\bar{\alpha}_0, \underline{\alpha}_1)$, the first term is 0. Note that $E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|h, l] < b_a(h)$ is a necessary and sufficient condition for no deviation. In that case, players bid $b = b_a(l)$, use $(\bar{\alpha}_0, \underline{\alpha}_1)$ as the information structure, and the equilibrium payoff is 0.

Let us now consider the high type. As long as the other player is playing the equilibrium (pure) strategy, the payoff from any mixed strategy is:

$$\min_{(\alpha_0,\alpha_1)} \rho_0 Pr(l|h, (\alpha_0, \alpha_1)) \left(E_{(\alpha_0,\alpha_1)}[v|l, h] - b_a(l)\right) + \rho_1 Pr(h|h, (\alpha_0, \alpha_1)) \left(E_{(\alpha_0,\alpha_1)}[v|h, h] - b_a(h)\right)$$

where $0 \leq \rho_1 \leq \rho_0 \leq 1$. Under the information structure $(\underline{\alpha}_0, \bar{\alpha}_1)$, $E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h, h] = b_a(h)$, which implies that the payoff from the deviation is at most $Pr(l|h, (\underline{\alpha}_0, \bar{\alpha}_1)) \left(E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|l, h] - b_a(l)\right)$, which is the equilibrium payoff. Thus, it is not profitable to deviate to any mixed strategy.

Bringing together all the conditions, we now have:

$$E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|l, h] > E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l, l]$$

$$E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|h, l] < E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h, h]$$

$$E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l, l] > \tfrac{1}{2} E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h, h].$$

For $a \leq \frac{q}{1-q}$, these conditions are:

(1) $\frac{1-q-a(1-q)^2}{1-(\frac{1}{a}(1-q)^2+2q-1)-a(1-q)^2} - \frac{\frac{1}{a}(1-q)^2}{\frac{1}{a}(1-q)^2+a(1-q)^2+2q-1} > 0$

(2) $\frac{2q-1+a(1-q)^2}{a(1-q)^2+\frac{1}{a}(1-q)^2+2q-1} - \frac{1-q-\frac{1}{a}(1-q)^2}{1-(a(1-q)^2+2q-1)-\frac{1}{a}(1-q)^2} > 0$

(3) $\frac{\frac{1}{a}(1-q)^2}{\frac{1}{a}(1-q)^2+a(1-q)^2+2q-1} - \frac{1}{2}\frac{2q-1+a(1-q)^2}{a(1-q)^2+\frac{1}{a}(1-q)^2+2q-1} > 0$

Conditions (1) and (2) are satisfied for all q, while condition (3) is satisfied for all $q < \underline{q}(a)$.

For $a \geq \frac{q}{1-q}$, condition (3), now $\frac{\frac{1}{a}(1-q)^2}{\frac{1}{a}(1-q)^2+q-\frac{1}{a}q(1-q)} - \frac{1}{2}\frac{2q-1+1-q-\frac{1}{a}q(1-q)}{1-q-\frac{1}{a}q(1-q)+\frac{1}{a}(1-q)^2+2q-1} > 0$, is not satisfied for a which is above a cutoff $\bar{a}$. Note that allowing for no hedging will not affect the existence of this equilibrium for high a.

Equilibria with underbidding for the high type

Next consider the case $b_a(l) < \frac{1}{2}b_a(h)$. Consider the high type, and assume that the other player is playing the equilibrium strategy $(b_a(l), b_a(h))$.

Bidding $b = b_a(h)$ yields:

$$\min_{(\alpha_0,\alpha_1)} Pr(l|h,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|l,h] - b_a(l)\right) + \frac{1}{2}Pr(h|h,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|h,h] - b_a(h)\right)$$

which is like solving

$$\begin{aligned}&\min_{(\alpha_0,\alpha_1)} -\alpha_1 + (\alpha_1+\alpha_0)b_a(l) + \frac{1}{2}\alpha_1 - \frac{1}{2}(\alpha_1+\alpha_0)b_a(h)\\ =&\min_{(\alpha_0,\alpha_1)} -\alpha_1(\frac{1}{2} - b_a(l) + \frac{1}{2}b_a(h)) + \alpha_0(b_a(l) - \frac{1}{2}b_a(h))\end{aligned}$$

For $b_a(l) < \frac{1}{2}b_a(h)$, this is minimised by $(\bar{\alpha}_0, \bar{\alpha}_1)$.

The equilibrium payoff is then $Pr(l|h,(\bar{\alpha}_0,\bar{\alpha}_1))\left(E_{(\bar{\alpha}_0,\bar{\alpha}_1)}[v|l,h] - b_a(l)\right)$.

We now consider deviations. Let us consider first the high type. The payoff from any mixed strategy is:

$$\min_{(\alpha_0,\alpha_1)} \rho_0 Pr(l|h,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|l,h] - b_a(l)\right) + \rho_1 Pr(h|h,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|h,h] - b_a(h)\right)$$

where $0 \leq \rho_1 \leq \rho_0 \leq 1$. Under the information structure $(\bar{\alpha}_0, \bar{\alpha}_1)$, $E_{(\bar{\alpha}_0,\bar{\alpha}_1)}[v|h,h] = b_a(h)$. Note that this bid decreases with a.

Note that in equilibrium we must have $\left(E_{(\bar{\alpha}_0,\bar{\alpha}_1)}[v|l,h] - E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l]\right) \geq 0$, and that the equilibrium maximises the probability of winning against the low type.

Consider now the low type. Under no "no hedging", we have that the payoff from any mixed strategy is:

$$\min_{(\alpha_0,\alpha_1)} \rho_0 Pr(l|l,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|l,l] - b_a(l)\right) + \rho_1 Pr(h|l,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|h,l] - b_a(h)\right)$$

Under the information structure $(\bar{\alpha}_0, \underline{\alpha}_1)$, the first term is 0. Thus a necessary and sufficient condition for the low type not to deviate is $E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|h,l] < E_{(\bar{\alpha}_0,\bar{\alpha}_1)}[v|h,h]$.

The equilibrium conditions as described above are therefore:

(4) $E_{(\bar{\alpha}_0,\bar{\alpha}_1)}[v|l,h] > E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l]$

(5) $E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|h,l] < E_{(\bar{\alpha}_0,\bar{\alpha}_1)}[v|h,h]$

(6) $E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l] < \frac{1}{2}E_{(\bar{\alpha}_0,\bar{\alpha}_1)}[v|h,h]$.

Conditions (4) and (5) are satisfied for $a \leq \frac{q}{1-q}$, while condition (6) is satisfied for $q > \bar{q}(a)$.

To consider $a > \frac{q}{1-q}$, consider deviations of the low type under the "no hedging" condition. Her utility from a mixed strategy which wins against the low type only with probability β and with probability $1-\beta$ wins against the low type with probability 1 as well as against the high type with probability $\frac{1}{2}$ is:

$$\begin{aligned}&\beta \min_{(\alpha_0,\alpha_1)} Pr(l|l,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|l,l] - b_a(l)\right)\\ &+ (1-\beta)\min_{(\alpha_0,\alpha_1)}\Big(Pr(l|l,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|l,l] - b_a(l)\right)\\ &+ \frac{1}{2}Pr(h|l,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|h,l] - b_a(h)\right)\Big)\end{aligned}$$

Note that $\arg\min_{(\alpha_0,\alpha_1)} Pr(l|l,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|l,l]-b_a(l)\right)$ is $(\overline{\alpha}_0,\underline{\alpha}_1)$, and thus this part of the utility is 0, and that

$$\arg\min_{(\alpha_0,\alpha_1)} \Pr(l|l,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|l,l]-b_a(l)\right)+\frac{1}{2}Pr(h|l,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|h,l]-b_a(h)\right)$$
$$=(\underline{\alpha}_0,\underline{\alpha}_1)$$

A necessary and sufficient condition under the no hedging condition is for the above utility to be lower than 0, the equilibrium utility.

Thus, for $a>\frac{q}{1-q}$, the equilibrium conditions are:

(4) $E_{(\bar{\alpha}_0,\bar{\alpha}_1)}[v|l,h]>E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l]$

(5*) $\Pr(l|l,(\underline{\alpha}_0,\underline{\alpha}_1))\left(E_{\underline{\alpha}_0,\underline{\alpha}_1}[v|l,l]-b_a(l)\right)+\frac{1}{2}Pr(h|l,(\underline{\alpha}_0,\underline{\alpha}_1))\left(E_{(\underline{\alpha}_0,\underline{\alpha}_1)}[v|h,l]-b_a(h)\right)<0$

(6) $E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l]<\frac{1}{2}E_{(\bar{\alpha}_0,\bar{\alpha}_1)}[v|h,h]$.

This equilibrium exists when $a>\bar{a}(q)\geq\frac{q}{1-q}$, where $\bar{a}(q)>\frac{q}{1-q}$ for a low enough q but $\bar{a}(q)=\frac{q}{1-q}$ otherwise.

Note also that the equilibrium converges to the equilibrium in the limit where all information structures are allowed. To see the limit equilibrium, suppose that $b_a(l)=0$. For the low type we minimise α_1 at 0 and set $\alpha_0=q$ (note that the low type gets the same utility under any belief about α_0) and hence $E_{(q,0)}(v|l,l)=0$. We are therefore in the case in which $b_a(l)<\frac{1}{2}b_a(h)$ and hence the high type uses $\alpha_0=q$ and $\alpha_1=1-q$. As a result, $b_a(h)=q=E(v|h)<E(v|h,h)$. This yields the lowest revenue to the seller.

Finally, consider the case $b_a(l)=\frac{1}{2}b_a(h)$. We will show that this equilibrium holds for $a<\frac{q}{1-q}$, for values $\underline{q}(a)<q<\bar{q}(a)$. Let a and q satisfy: $\frac{1}{2}E_{(\bar{\alpha}_0,\bar{\alpha}_1)}[v|h,h]<E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l]<\frac{1}{2}E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h,h]$.

Consider the high type, and assume that the other player is playing the equilibrium strategy $(b_a(l),b_a(h))$. Bidding $b=b_a(h)$ yields:

$$\min_{(\alpha_0,\alpha_1)} Pr(l|h,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|l,h]-b_a(l)\right)+\frac{1}{2}Pr(h|h,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|h,h]-b_a(h)\right)$$

Since $b_a(l)=\frac{1}{2}b_a(h)$, for any α_0', $(\alpha_0',\bar{\alpha}_1)$ achieves the minimum payoff.

The payoff from any mixed strategy is:

$$\min_{(\alpha_0,\alpha_1)} \rho_0 Pr(l|h,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|l,h]-b_a(l)\right)+\rho_1 Pr(h|h,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|h,h]-b_a(h)\right)$$

Since $Pr_{(\alpha_0,\alpha_1)}(l|h)\left(E_{(\alpha_0,\alpha_1)}[v|l,h]-b_a(l)\right)>0$ for any (α_0,α_1) and increasing ρ_0 relaxes the constraint on ρ_1, it is without loss to set $\rho_0=1$. Using the fact that $b_a(l)=\frac{1}{2}b_a(h)$, the payoff becomes:

$$\min_{(\alpha_0,\alpha_1)} Pr(l|h,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|l,h]-b_a(l)\right)+\rho_1 Pr(h|h,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|h,h]-2b_a(l)\right)$$
$$=\min_{(\alpha_0,\alpha_1)} 1-b_a(l)-q+\rho_1(2q-1)+\alpha_0 b_a(l)(1-2\rho_1)+\alpha_1\left(\rho_1[1-2b_a(l)]-[1-b_a(l)]\right)$$

The payoff is minimised by $(\underline{\alpha}_0,\bar{\alpha}_1)$ when $\rho_1\leq\frac{1}{2}$ and $(\bar{\alpha}_0,\bar{\alpha}_1)$ when $\rho_1\geq\frac{1}{2}$.

Suppose that $\rho_1>\frac{1}{2}$. Then under $(\bar{\alpha}_0,\bar{\alpha}_1)$, the payoff is lower than when $\rho_1=\frac{1}{2}$, since $E_{(\bar{\alpha}_0,\bar{\alpha}_1)}[v|h,h]<b_a(h)$. If $\rho_1<\frac{1}{2}$, then under $(\underline{\alpha}_0,\bar{\alpha}_1)$, the payoff is lower than when $\rho_1=\frac{1}{2}$, since

$E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h,h] > b_a(h)$. Thus, for $\rho_1 \neq \frac{1}{2}$, the payoff must be lower than when $\rho_1 = \frac{1}{2}$, which is the equilibrium payoff.

Now consider the low type. As before, the equilibrium payoff is 0. The payoff from any mixed strategy is:

$$\min_{(\alpha_0,\alpha_1)} \rho_0 Pr(l|l,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|l,l] - b_a(l)\right) + \rho_1 Pr(h|l,(\alpha_0,\alpha_1))\left(E_{(\alpha_0,\alpha_1)}[v|h,l] - b_a(h)\right)$$

Under the information structure $(\bar{\alpha}_0,\underline{\alpha}_1)$, the first term is 0 and the second term is negative if $E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|h,l] < b_a(h)$. So we need $\frac{1}{2}E_{(\bar{\alpha}_0,\bar{\alpha}_1)}[v|h,h] < E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l] < \frac{1}{2}E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h,h]$ and $\frac{1}{2}E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|h,l] < E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l]$, which is satisfied for the range of qs considered. □

Equilibrium in the first price auction

Lemma 2.A.2. *In the unique symmetric equilibrium of the first price auction, the low type bids $b_a(l)$ and the high type mixes on the support $[b_a(l), \bar{b}_a(h)]$ according to the distribution $F_a(b)$, where:*

$$\begin{aligned}
b_a(l) &= E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,l)\\
\bar{b}_a(h) &= \Pr(h|h,(\underline{\alpha}_0,\bar{\alpha}_1))E_{(\underline{\alpha}_0,\bar{\alpha}_1)}(v|h,h) + \Pr(l|h,(\underline{\alpha}_0,\bar{\alpha}_1))E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,l)\\
F_a(b) &= \frac{\Pr(l|h,(\underline{\alpha}_0,\bar{\alpha}_1))(b - b_a(l))}{\Pr(h|h,(\underline{\alpha}_0,\bar{\alpha}_1))(E_{(\underline{\alpha}_0,\bar{\alpha}_1)}(v|h,h) - b)}.
\end{aligned}$$

Proof of Lemma 2.A.1. We first conjecture the bid of the low type, and then we solve for the equilibrium beliefs and mixed strategy of the high type, checking that neither type has an incentive to deviate. We show that for the high type, the maximum bid increases in a, and that the minimum bid decreases in a, for a close to 1.

Consider a low type. For any $b_a(l)$, this type's expected utility is perceived as

$$\begin{aligned}
&\min_{(\alpha_0,\alpha_1)} (\alpha_0+\alpha_1)\left(\frac{\alpha_1}{\alpha_0+\alpha_1} - b_a(l)\right)\\
&= \min_{(\alpha_0,\alpha_1)} \alpha_0(-b_a(l)) + \alpha_1(1 - b_a(l))
\end{aligned}$$

which is resolved by setting α_0 to be the highest possible value and α_1 to be the lowest possible value, given the ePMI constraints. Therefore for a sufficiently close to 1, the solution is $(\bar{\alpha}_0,\underline{\alpha}_1)$.

Note that the low type cannot earn positive rents in equilibrium, and thus we conjecture:

$$b_a(l) = E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,l) = \frac{\frac{1}{a}(1-q)^2}{\frac{1}{a}(1-q)^2 + a(1-q)^2 + 2q - 1} < \frac{(1-q)^2}{(1-q)^2+q^2}$$

Since the derivative of $E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,l)$ with respect to a is negative, the bid of the low type decreases with a. We will establish later that this type will not want to use any other bid given the behaviour of the high type.

Now let us consider the high type. Without loss of generality we can consider a mixed strategy with support on $[b_a(l), \bar{b}_a(h)]$, as bidding less than $b_a(l)$ will provide a zero utility.

First let us consider a bid just above $b_a(l)$ which allows the individual to win against the low type only. We then need to solve the following:

$$\begin{aligned}
&\min_{(\alpha_0,\alpha_1)} \Pr(l|h,(\alpha_0,\alpha_1))(E_{(\alpha_0,\alpha_1)}(v|h,l) - b_a(l))\\
&= \min_{(\alpha_0,\alpha_1)} (q-\alpha_0)(-b_a(l)) + (1-q-\alpha_1)(1 - b_a(l))
\end{aligned}$$

which yields the need to maximise α_1 and to minimise α_0. The solution is $(\underline{\alpha}_0, \bar{\alpha}_1)$. Note that this bid provides a utility of $\Pr(l|h,(\underline{\alpha}_0,\bar{\alpha}_1))(E_{(\underline{\alpha}_0,\bar{\alpha}_1)}(v|h,l) - E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,l)) > 0$, and that $\overline{\alpha}_0 + \underline{\alpha}_1 = \underline{\alpha}_0 + \overline{\alpha}_1$.

We now consider the highest bid in the support, $\bar{b}_a(h)$. Such bid implies winning for sure and thus unambiguous gain of $E(v|h)$. To be indifferent, this bid has to satisfy:

$$E(v|h) - \bar{b}_a(h) = \Pr(l|h,(\underline{\alpha}_0,\overline{\alpha}_1))(E_{(\underline{\alpha}_0,\overline{\alpha}_1)}(v|h,l) - E_{(\overline{\alpha}_0,\underline{\alpha}_1)}(v|l,l))$$

Thus:

$$\begin{aligned}
\bar{b}_a(h) &= \Pr(h|h,(\underline{\alpha}_0,\bar{\alpha}_1))E_{(\underline{\alpha}_0,\bar{\alpha}_1)}(v|h,h) + \Pr(l|h,(\underline{\alpha}_0,\bar{\alpha}_1))E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,l) \\
&= 2q - 1 + \overline{\alpha}_1 + (1 - \underline{\alpha}_0 - \overline{\alpha}_1)\frac{\underline{\alpha}_1}{\overline{\alpha}_0 + \underline{\alpha}_1} \\
&= 2q - 1 + \overline{\alpha}_1 + \frac{\underline{\alpha}_1}{\overline{\alpha}_0 + \underline{\alpha}_1} - \underline{\alpha}_1 \\
&= 2q - 1 + a(1-q)^2 + \frac{\frac{1}{a}(1-q)^2}{a(1-q)^2 + 2q - 1 + \frac{1}{a}(1-q)^2} - \frac{1}{a}(1-q)^2 \\
&= 2q - 1 + (1-q)^2\left(a - \frac{1}{a} + \frac{\frac{1}{a}}{(a + \frac{1}{a} - 1)(1-q)^2 + q^2}\right)
\end{aligned}$$

Note that the derivative of $a - \frac{1}{a} + \frac{\frac{1}{a}}{(\frac{1}{a}+a-1)(1-q)^2+q^2}$, evaluated at $a = 1$, is $\frac{(2q-1)^2}{2q^2-2q+1} > 0$. Thus the maximum bid increases in a.

We now continue to characterise the equilibrium distribution. Let us consider the worst case scenario in terms of utility for some distribution $F(b)$ with density $f(b)$. The expected utility is

$$\begin{aligned}
&\int_b f(b)[\Pr(l|h,(\alpha_0,\alpha_1))(E_{(\alpha_0,\alpha_1)}(v|h,l) - b) + \Pr(h|h,(\alpha_0,\alpha_1))F(b)(E_{(\alpha_0,\alpha_1)}(v|h,h) - b)]db \\
= &\int_b f(b)[E(v|h) - b - (1 - F(b))\Pr(h|h,(\alpha_0,\alpha_1))(E_{(\alpha_0,\alpha_1)}(v|h,h) - b)]db
\end{aligned}$$

To choose the information structure to minimise utility, we maximise

$$\Pr(h|h,(\alpha_0,\alpha_1))(E_{(\alpha_0,\alpha_1)}(v|h,h) - b) = (2q-1) + (\alpha_0(-b) + \alpha_1(1-b))$$

and the solution is therefore, for all b, to maximise α_1 and to minimise α_0.

$F(b)$ is characterised by using the indifference condition under the belief $(\underline{\alpha}_0, \bar{\alpha}_1)$:

$$\begin{aligned}
&\Pr(l|h,(\underline{\alpha}_0,\bar{\alpha}_1))(E_{(\underline{\alpha}_0,\bar{\alpha}_1)}(v|h,l) - b) + \Pr(h|h,(\underline{\alpha}_0,\bar{\alpha}_1))F(b)(E_{(\underline{\alpha}_0,\bar{\alpha}_1)}(v|h,h) - b) \\
= &\Pr(l|h,(\underline{\alpha}_0,\bar{\alpha}_1))(E_{(\underline{\alpha}_0,\bar{\alpha}_1)}(v|h,l) - b_a(l))
\end{aligned}$$

implying that

$$F_a(b) = \frac{\Pr(l|h,(\underline{\alpha}_0,\bar{\alpha}_1))(b - b_a(l))}{\Pr(h|h,(\underline{\alpha}_0,\bar{\alpha}_1))(E_{(\underline{\alpha}_0,\bar{\alpha}_1)}(v|h,h) - b)}.$$

We complete the equilibrium characterisation by showing that given the strategy of the high type, the low type will not deviate.

For the low type, bidding any b above $b_a(l)$, we choose the belief to minimise expected utility:

$$\begin{aligned}
&\min_{(\alpha_1,\alpha_0)} \Pr(l|l,(\alpha_0,\alpha_1))(E_{(\alpha_0,\alpha_1)}(v|l,l) - b) + \Pr(h|l,(\alpha_0,\alpha_1))F_a(b)(E_{(\alpha_0,\alpha_1)}(v|l,h) - b) \\
= &\min_{(\alpha_0,\alpha_1)} (-b)(\alpha_0(1 - F_a(b)) + F_a(b)q) + (1-b)(\alpha_1(1 - F_a(b)) + F_a(b)(1-q))
\end{aligned}$$

As $F_a(b) \leq 1$, the solution is $(\bar{\alpha}_0, \underline{\alpha}_1)$. This gives us a utility of:

$$\begin{aligned}
&\Pr(l|l,(\bar{\alpha}_0,\underline{\alpha}_1))(E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,l)-b)\\
&+\Pr(h|l,(\bar{\alpha}_0,\underline{\alpha}_1))\frac{\Pr(l|h,(\underline{\alpha}_0,\bar{\alpha}_1))(b-E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,l))}{\Pr(h|h,(\underline{\alpha}_0,\bar{\alpha}_1))(E_{(\underline{\alpha}_0,\bar{\alpha}_1)}(v|h,h)-b)}(E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,h)-b)\\
&=\Pr(h|l,(\bar{\alpha}_0,\underline{\alpha}_1))(b-E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,l))(\frac{\Pr(l|h,(\underline{\alpha}_0,\bar{\alpha}_1))}{\Pr(h|h,(\underline{\alpha}_0,\bar{\alpha}_1))}\frac{(E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,h)-b)}{(E_{(\underline{\alpha}_0,\bar{\alpha}_1)}(v|h,h)-b)}-\frac{\Pr(l|l,(\bar{\alpha}_0,\underline{\alpha}_1))}{\Pr(h|l,(\bar{\alpha}_0,\underline{\alpha}_1))}).
\end{aligned}$$

Note that $E_{(\bar{\alpha}_0,\underline{\alpha}_1)}(v|l,h) = \frac{1-q-\frac{1}{a}(1-q)^2}{1-\bar{\alpha}_0-\underline{\alpha}_1} < \frac{2q-1+a(1-q)^2}{\underline{\alpha}_0+\bar{\alpha}_1} = E_{(\underline{\alpha}_0,\bar{\alpha}_1)}(v|h,h)$, for a sufficiently close to 1,[23] and that $\frac{\Pr_{(\underline{\alpha}_0,\bar{\alpha}_1)}(l|h)}{\Pr_{(\underline{\alpha}_0,\bar{\alpha}_1)}(h|h)} = \frac{1-\underline{\alpha}_0-\bar{\alpha}_1}{\underline{\alpha}_0,\bar{\alpha}_1} < \frac{(\bar{\alpha}_0+\underline{\alpha}_1)}{1-(\bar{\alpha}_0+\underline{\alpha}_1)} = \frac{\Pr(l|l,(\bar{\alpha}_0,\underline{\alpha}_1))}{\Pr(h|l,(\bar{\alpha}_0,\underline{\alpha}_1))}$, as $\bar{\alpha}_0+\underline{\alpha}_1 = \bar{\alpha}_1+\underline{\alpha}_0 > \frac{1}{2}$ for all a. Thus the utility is negative and the low type does not deviate. □

Seller's revenue in the first price auction

Now we show that the seller's revenue in the first price auction is decreasing in a, for sufficiently small a. We prove this using five preliminary facts that allow us to rewrite the revenue in a convenient form, and then we take the derivative with respect to a, evaluated at $a = 1$, and show that it is negative.

The expected payment to seller, Π, is given by the linear combination of receiving the bid of the low type (when both are l), the expected bid of the high type (when only one is h), and the maximum bid of the two h types:

$$\Pi = \Pr(l,l)E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l] + 2\Pr(l,h)E[b_a(h)] + \Pr(h,h)E[\max_{i=1,2} b_a^i(h)]$$

For expositional purposes we write this as

$$\Pi = \frac{1}{2}\gamma y + (1-\gamma)\int_y^{\alpha x+(1-\alpha)y} b f_a(b)db + \frac{\gamma}{2}\int_y^{\alpha x+(1-\alpha)y} b2f_a(b)F_a(b)db,$$

where $\gamma = \Pr(l|l)$, according to the true (independent) information structure, $\alpha = \Pr(l|l,(\underline{\alpha}_0,\bar{\alpha}_1))$ according to the belief of the high bidder, $(\underline{\alpha}_0,\bar{\alpha}_1)$, $x = E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h,h]$, $y = E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l] = b_a(l)$. We therefore also have $\bar{b}_a(h) = \alpha x + (1-\alpha)y$, $F_a(b) = \frac{1-\alpha}{\alpha}\frac{b-y}{x-b}$ and $f_a(b) = \frac{1}{\alpha}\frac{1-\alpha}{(x-b)^2}\left(x-y\right)$.

We start by some preliminary results:

Fact 1

$$\frac{\partial x}{\partial a} = -\frac{\partial y}{\partial a} > 0$$

Proof. Note that:

$$\begin{aligned}
x = E_{(\underline{\alpha}_0,\bar{\alpha}_1)}(v|h,h) &= \frac{2q-1+a(1-q)^2}{q^2-(1-\frac{1}{a})(1-q)^2+a(1-q)^2}\\
\frac{\partial x}{\partial a} &= \frac{\partial}{\partial a}(\frac{2q-1+a(1-q)^2}{q^2-(1-\frac{1}{a})(1-q)^2+a(1-q)^2})\\
&= \left(q-1\right)^2\frac{2a+2q+2aq^2-4aq-1}{\left(a^2q^2-2a^2q+a^2+2aq-a+q^2-2q+1\right)^2} > 0\\
\frac{\partial y}{\partial a} &= \frac{\partial}{\partial a}\frac{\frac{1}{a}(1-q)^2}{2q-1+a(1-q)^2+\frac{1}{a}(1-q)^2}\\
&= -\left(q-1\right)^2\frac{2a+2q+2aq^2-4aq-1}{\left(a^2q^2-2a^2q+a^2+2aq-a+q^2-2q+1\right)^2}.
\end{aligned}$$

□

[23] In fact, this is true for all a.

Fact 2

$$\frac{\partial\alpha}{\partial a}_{|a=1} = 0$$

Proof. Note that:

$$\alpha = \Pr_{(\underline{\alpha}_0,\bar{\alpha}_1)}(l|l) = \underline{\alpha}_0 + \overline{\alpha}_1 = a(1-q)^2 + \frac{1}{a}(1-q)^2 + 2q - 1$$

$$\frac{\partial\alpha}{\partial a}_{|a=1} = (1-q)^2\frac{\partial(a+\frac{1}{a})}{\partial a}_{|a=1} = (1-q)^2(1-\frac{1}{a^2})_{|a=1} = 0.$$

□

Fact 3

(i) $E[b_a(h)] = x(1+\frac{1-\alpha}{\alpha}\ln(1-\alpha)) - y\frac{1-\alpha}{\alpha}\ln(1-\alpha)$.

(ii) At $a = 1$, the expected bid of the high type decreases in a for low q and increases in a for high q.

Proof.

(i) Note that:

$$\int \frac{b}{(x-b)^2}db = \frac{1}{x-b}\left(x - b\ln(x-b) + x\ln(x-b)\right) = \frac{x-(b-x)\ln(x-b)}{x-b}.$$

Therefore:

$$\int_y^{\alpha x+(1-\alpha)y} \frac{b}{(b-x)^2}db = \frac{\alpha x}{(1-\alpha)(x-y)} + \ln(1-\alpha).$$

Hence:

$$E[b_a(h)] = \frac{1-\alpha}{\alpha}\left(x-y\right)\int_y^{\alpha x+(1-\alpha)y} \frac{b}{(b-x)^2}db = x(1+\frac{1-\alpha}{\alpha}\ln(1-\alpha)) - y\frac{1-\alpha}{\alpha}\ln(1-\alpha).$$

(ii) As $\frac{\partial\alpha}{\partial a}_{|a=1} = 0$ and $\frac{\partial x}{\partial a}_{|a=1} = -\frac{\partial y}{\partial a}_{|a=1}$:

$$\begin{aligned}\frac{\partial E[b_a(h)]}{\partial a}_{|a=1} &= \frac{\partial x}{\partial a}_{|a=1}(1+2\frac{1-\alpha}{\alpha}\ln(1-\alpha))_{|a=1} \\ &= \frac{\partial x}{\partial a}_{|a=1}(1+2\frac{1-2(1-q)^2-2q+1}{2(1-q)^2+2q-1}\ln(1-2(1-q)^2-2q+1)).\end{aligned}$$

For $q > \frac{1}{2}$, the expression $(1+2\frac{1-2(1-q)^2-2q+1}{2(1-q)^2+2q-1}\ln(1-2(1-q)^2-2q+1))$ is strictly increasing, negative for $q < q^*$ and positive for $q > q^*$ for some $q^* \in (0.5, 1)$. As $\frac{\partial x}{\partial a}_{|a=1} > 0$, we are done. □

Fact 4

(i) $E[\max_{i=1,2} b_a^i(h)] = -\left(x-y\right)2((\frac{1-\alpha}{\alpha})^2\ln(1-\alpha) + \frac{1-\alpha}{\alpha}) + x$

(ii) The expectation of the maximal bid when both are high types decreases in a for low q and increases in a for high q.

Proof.

(i) $\int \frac{b(b-y)}{(x-b)^3} db = \frac{1}{2(x-b)^2}\left(2b^2 \ln(x-b) + 2x^2 \ln(x-b) - 4bx + 2by - xy + 3x^2 - 4bx\ln(x-b)\right) = -\ln(x-b) - \frac{-4bx+2by-xy+3x^2}{2(x-b)^2}$. Therefore, $\int_y^{\alpha x+(1-\alpha)y} \frac{b(b-y)}{(x-b)^3} db = -\ln(1-\alpha) - \frac{\alpha(2x-2y-3x\alpha+2y\alpha)}{2(1-\alpha)^2(x-y)}$. Hence, $E[\max_{i=1,2} b_a^i(h)] = 2(\frac{1-\alpha}{\alpha})^2 (x-y) \int_y^{\alpha x+(1-\alpha)y} \frac{b(b-y)}{(x-b)^3} db = -2(\frac{1-\alpha}{\alpha})^2 (x-y)(\ln(1-\alpha) + \frac{\alpha(2x-2y-3x\alpha+2y\alpha)}{2(1-\alpha)^2(x-y)}) = -(x-y)\, 2((\frac{1-\alpha}{\alpha})^2 \ln(1-\alpha) + \frac{1-\alpha}{\alpha}) + x$.

(ii) Recalling that $\frac{\partial x}{\partial a} = -\frac{\partial y}{\partial a}$ and that $\frac{\partial \alpha}{\partial a} = 0$ we have,
$\frac{\partial E[\max_{i=1,2} b_a^i(h)]}{\partial a} = \frac{\partial x}{\partial a}(-4(\frac{1-\alpha}{\alpha})^2 \ln(1-\alpha) - 4\frac{1-\alpha}{\alpha} + 1)$, and
$\frac{\partial E[\max_{i=1,2} b_a^i(h)]}{\partial a}_{|a=1} = \frac{\partial x}{\partial a}(-4(\frac{2q(1-q)}{q^2+(1-q)^2})^2 \ln(1-\alpha) - 4\frac{2q(1-q)}{q^2+(1-q)^2} + 1)$
For the expression $(-4(\frac{2q(1-q)}{q^2+(1-q)^2})^2 \ln(2q(1-q)) - 4\frac{2q(1-q)}{q^2+(1-q)^2} + 1)$ there is a $\bar{q} \in (0.5, 1)$ such that the expression is negative for $q < \bar{q}$ and positive for $q > \bar{q}$. As $\frac{\partial x}{\partial a} > 0$, we are done. □

Given the above facts we can write the profit function as:

$$\begin{aligned}
\Pi &= \frac{1}{2}\gamma y + (1-\gamma)x + \frac{1-\alpha}{\alpha}(x-y)(1-\gamma)\ln(1-\alpha) \\
&\quad + -\frac{1-\alpha}{\alpha}^2 (x-y)\gamma \ln(1-\alpha) - \frac{1}{\alpha}\gamma(\frac{2x-2y-3x\alpha+2y\alpha}{2}) \\
&= x(\frac{\gamma}{2} + \frac{\alpha-\gamma}{\alpha}(\frac{1-\alpha}{\alpha}\ln(1-\alpha)+1)) + y(1 - \frac{\gamma}{2} - \frac{\alpha-\gamma}{\alpha}(\frac{1-\alpha}{\alpha}\ln(1-\alpha)+1))
\end{aligned}$$

Taking a derivative with respect to a, recalling that $\frac{d\alpha}{da}_{|a=1} = 0$ and that $\frac{\partial x}{\partial a} = -\frac{\partial y}{\partial a}$ we get:

$$\begin{aligned}
\frac{\partial \Pi}{\partial a}_{|a=1} &= \frac{\partial x}{\partial a}(\gamma - 1 + 2\frac{\alpha-\gamma}{\alpha}(\frac{1-\alpha}{\alpha}\ln(1-\alpha)+1)) \\
&= \frac{\partial x}{\partial a}(\gamma - 1) < 0.
\end{aligned}$$

Seller's revenue in the second price auction and a comparison across auctions

We now consider the seller's revenue in the second price auction and show that the profits of the seller are higher in the first-price auction, which is part (iii) of Proposition 2.4.3. We also show that the seller's revenue in the second price auction is decreasing in a, and that the expected utility of both types of bidders in the second price auction, evaluated at the true joint probability distribution, is increasing in a, which is part (i) of Proposition 2.4.3.

In the second price auction, the low type always bids $E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l]$, and the high type bids at most $E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h,h]$ (either $b_a(h) = E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h,h]$ for low a and q, where $E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h,h] > E_{(\bar{\alpha}_0\bar{\alpha}_1)}[v|h,h] = b_a(h)$ for higher q and a, or where $E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h,h] > 2E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l] = b_a(h)$). Let $\bar{R}^{SPA}$ be the revenue from a virtual auction where the low type bids $E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l]$ and the high type bids $E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h,h]$. Then the actual revenue in the second price auction must be weakly less than $\bar{R}^{SPA}$.

The seller's revenue in the second price auction satisfies:

$$R^{SPA} \leq \bar{R}^{SPA} = x\left(\frac{\gamma}{2}\right) + y\left(1 - \frac{\gamma}{2}\right)$$

where $x = E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h,h]$, $y = E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l]$, and $\gamma = Pr(l|l)$. The seller's revenue in the first price auction is:

$$R^{FPA} = x\left[\frac{\gamma}{2} + \frac{\alpha-\gamma}{\alpha}\left(\frac{1-\alpha}{\alpha}\ln(1-\alpha)+1\right)\right] + y\left[1 - \frac{\gamma}{2} - \frac{\alpha-\gamma}{\alpha}\left(\frac{1-\alpha}{\alpha}\ln(1-\alpha)+1\right)\right]$$

where $\alpha = Pr(l|l,(\underline{\alpha}_0, \bar{\alpha}_1))$. Thus, the difference in revenue between the two auctions is:

$$R^{FPA} - R^{SPA} \geq R^{FPA} - \bar{R}^{SPA} = (x-y)\left[\frac{\alpha-\gamma}{\alpha}\left(\frac{1-\alpha}{\alpha}\ln(1-\alpha)+1\right)\right] > 0.$$

Finally, to see that R^{SPA} is decreasing in a, first note that when a and q are low:

$$\frac{\partial R_1^{SPA}}{\partial a}_{|a=1} = \frac{\partial x}{\partial a}\left(\frac{\gamma}{2}\right) + \frac{\partial y}{\partial a}\left(1-\frac{\gamma}{2}\right) = -(1-\gamma)\frac{\partial x}{\partial a} < 0$$

Let $x' = E_{(\bar{\alpha}_0, \bar{\alpha}_1)}[v|h,h]$. Then the revenue when q is high and a is high:

$$R_2^{SPA} = x'\left(\frac{\gamma}{2}\right) + y\left(1-\frac{\gamma}{2}\right)$$

Therefore:

$$\frac{\partial R_2^{SPA}}{\partial a}_{|a=1} = \frac{\partial x'}{\partial a}\left(\frac{\gamma}{2}\right) + \frac{\partial y}{\partial a}\left(1-\frac{\gamma}{2}\right) < 0$$

since $\frac{\partial x'}{\partial a} < 0$. Finally in the last case: $\frac{\partial R_3^{SPA}}{\partial a}_{|a=1} = \frac{\partial y}{\partial a}\left(1+\frac{\gamma}{2}\right) < 0$.

We have argued that the seller's revenue in the second price auction is lower than in the first price auction and is decreasing in a. Now we argue that in the second price auction, the expected utility of both types of bidders, evaluated at the true joint distribution, is increasing in a. Since the allocation does not depend on a, it is sufficient to show that the expected payment of each type, evaluated at the true joint distribution, is decreasing in a. Note that for low q and low a (the case where the high type's bid is increasing in a) $E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l] > \frac{1}{2}E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h,h]$, and:

$$\begin{aligned} E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l] > \frac{1}{2}E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h,h] &\iff \frac{\underline{\alpha}_1}{\bar{\alpha}_0+\underline{\alpha}_1} > \frac{1}{2}\frac{2q-1+\bar{\alpha}_1}{\underline{\alpha}_0+\bar{\alpha}_1} \\ &\iff 2\underline{\alpha}_1 - \bar{\alpha}_1 + 1 - 2q > 0 \end{aligned}$$

where the second line follows from the fact that $\overline{\alpha}_0 + \underline{\alpha}_1 = \underline{\alpha}_0 + \overline{\alpha}_1$.The derivative of $2\underline{\alpha}_1 - \bar{\alpha}_1 + 1 - 2q$ with respect to a is negative; thus a necessary condition for $2\underline{\alpha}_1 - \bar{\alpha}_1 + 1 - 2q > 0$ to hold is $(1-q)^2 + 1 - 2q > 0$,or $q < 2-\sqrt{2}$. But $\Pr(l|h) > \frac{1}{2}\Pr(h|h) \Leftrightarrow q < \frac{1}{6}\left(3+\sqrt{3}\right)$. Thus, $E_{(\bar{\alpha}_0,\underline{\alpha}_1)}[v|l,l] > \frac{1}{2}E_{(\underline{\alpha}_0,\bar{\alpha}_1)}[v|h,h]$ implies $\Pr(l|h) > \frac{1}{2}\Pr(h|h)$. $\Pr(l|h) > \frac{1}{2}\Pr(h|h)$ and $\frac{\partial x}{\partial a} = -\frac{\partial y}{\partial a}$ imply that the overall expected payment (under the true joint distribution) for the high type, $\Pr(l|h)y + \frac{1}{2}\Pr(h|h)x$, is decreasing in a. In all other cases, the high type reduces her bid with a, which implies that her expected payment is decreasing in a. Finally the bid and expected payment of the low type is always decreasing in a. □

Appendix 2.B Proofs for Section 2.5

We now prove two Propositions which are more detailed versions of Propositions 2.5.1 and 2.5.2.

Proposition 2.B.1. *Private value case: If $L < \frac{1}{1+a}$, (x,t) is an optimal mechanism if and only if it satisfies:*

- $x_i(l,l) = x_i(l,h) = t_i(l,l) = t_i(l,h) = 0$
- $x_i(h,l) = 1$, $\sum_{i=1}^{2} x_i(h,h) = 1$
- $t_i(h,l) = 1$
- $t_i(h,h) = x_i(h,h)$

That is, if both buyers have the low valuation, then the seller does not allocate the object. If a single buyer has the high valuation, then she gets the object with probability 1. If both buyers have the high valuation, the seller can allocate to either buyer. Transfers are such that both buyers are fully insured against ambiguity and receive zero rents.

If $L > \frac{1}{1+a}$, (x,t) is an optimal mechanism if and only if it satisfies:

- $x_i(l,h) = t_i(l,h) = 0$
- $\sum_{i=1}^{2} x_i(l,l) = 1$
- $t_i(l,l) = Lx_i(l,l)$
- $x_i(h,l) = 1$, $\sum_{i=1}^{2} x_i(h,h) = 1$
- $t_i(h,h) = x_i(h,h) - \frac{1}{2a}(1-L)x_i(l,l)$
- $t_i(h,l) = 1 - \frac{1}{2a}(1-L)x_i(l,l)$

That is, the seller allocates the object with probability 1. If the buyers have different valuations, the seller allocates to the one with the high valuation. Otherwise the seller may allocate to either buyer. Low types receive zero rents. Transfers for high types are pinned down by the binding incentive constraint and the full insurance condition.

The revenue of the seller is:

$$\frac{1}{4}\sum_{i=1}^{2} t_i(l,l) + t_i(l,h) + t_i(h,l) + t_i(h,h) = \begin{cases} \frac{3}{4} & \text{if } L < \frac{1}{1+a} \\ \frac{1}{2} + \frac{L}{2} + \frac{(a-1)(1-L)}{4a} & \text{if } L \geq \frac{1}{1+a} \end{cases}$$

which is increasing in ambiguity. As $a \to \infty$, the revenue converges to $\frac{3}{4} + \frac{L}{4}$, which is full surplus.

Proof. We prove Proposition 2.B.1 in two steps. First, we argue that any optimal mechanism must be a full insurance mechanism, and then we solve the seller's problem by finding the best full insurance mechanism.

First, we argue that any optimal mechanism must be a full insurance mechanism. Define a full insurance mechanism as one in which the ex post payoff of each buyer is independent of the type of her opponent:

Definition. The mechanism (x,t) is a full insurance mechanism if for each player i and for each $s, s' \in \{l, h\}$: $v^s x_i(s,s) - t_i(s,s) = v^s x_i(s,s') - t_i(s,s')$.

Suppose that (x,t) is an optimal mechanism, but $v^s x_1(s,s) - t_1(s,s) > v^s x_1(s,s') - t_1(s,s')$. Then the equilibrium belief for type s of player 1 is that her opponent is type s with probability $\frac{1}{2a}$ and type s' with probability $1 - \frac{1}{2a}$. Consider the alternative mechanism (x', t') where:

- $t'_1(s,s) = t_1(s,s) + \delta$
- $t'_1(s,s') = t_1(s,s') - \frac{1}{2a-1}\delta$
- $x'_1(s',\cdot) = x_1(s',\cdot)$ and $t'_1(s',\cdot) = t_1(s',\cdot)$
- $x'_2(\cdot,\cdot) = x_2(\cdot,\cdot)$ and $t'_2(\cdot,\cdot) = t_2(\cdot,\cdot)$

For sufficiently small δ:

- Type s of player 1 gets the same utility in (x', t') and (x, t) for every report
- Type s' of player 1 gets the same utility in (x', t') and (x, t) from reporting s'
- Type s' of player 1 gets weakly lower utility in (x', t') than in (x, t) from reporting s
- Both types of player 2 get the same utility in (x', t') and (x, t) for every report
- The seller gets a strictly higher revenue in (x', t') than in (x, t)

Thus, the new mechanism is incentive compatible, individually rational, and yields strictly higher revenue to the seller.

Next, note that $x_i(l, h) = 0$. For example, suppose that $x_1(l, h) > 0$. Consider the alternative mechanism (x', t') that differs from (x, t) in the following way:

- $x'_1(l, h) = x_1(l, h) - \delta$
- $t'_1(l, h) = t_1(l, h) - \delta L$
- $x'_2(h, l) = x_2(h, l) + \delta$
- $t'_2(h, l) = t_2(h, l) + \delta$

Note that (x', t') is incentive compatible, individually rational, and yields strictly higher revenue to the seller.

We can ignore the participation constraint of the high type as it is implied by the incentive constraint. The participation constraint of the low type must be binding, and full insurance then implies that the transfers for the low type satisfy:

$$t_i(l, l) = Lx_i(l, l) \tag{2.2}$$

$$t_i(l, h) = 0 \tag{2.3}$$

We ignore the incentive constraint of the low type and check ex post that it is satisfied. Then the incentive constraint of the high type must bind:

$$\min_{p \in \Pi} p\,[x_i(h, h) - t_i(h, h)] + (1 - p)\,[x_i(h, l) - t_i(h, l)] = \min_{p \in \Pi}(1 - p)\,[x_i(l, l) - t_i(l, l)]$$

Since $t_i(l, l) = Lx_i(l, l)$, $\min_{p\in\Pi}(1-p)\,[x_i(l, l) - t_i(l, l)] = \frac{1}{2a}(1-L)x_i(l, l)$. Full insurance implies that $x_i(h, h) - t_i(h, h) = x_i(h, l) - t_i(h, l)$, so we have:

$$t_i(h, h) = x_i(h, h) - \frac{1}{2a}(1 - L)x_i(l, l) \tag{2.4}$$

$$t_i(h, l) = x_i(h, l) - \frac{1}{2a}(1 - L)x_i(l, l) \tag{2.5}$$

Thus, we can write the seller's problem as:

$$\max_{x_i, t_i} \sum_{i=1}^{2} Lx_i(l, l) + x_i(h, h) + x_i(h, l) - \frac{1}{a}(1 - L)x_i(l, l)$$

Clearly, it is optimal to set $x_i(h, l) = 1$, $\sum_{i=1}^{2} x_i(h, h) = 1$, and $\sum_{i=1}^{2} x_i(l, l)$ to be either 1 or 0, depending on the sign of $L - \frac{1}{a}(1 - L)$.

Therefore, (x, t) solves the seller's problem if and only if the allocation rule satisfies:

$$x_i(l,h) = 0$$
$$x_i(h,l) = 1$$
$$\sum_{i=1}^{2} x_i(h,h) = 1$$
$$\sum_{i=1}^{2} x_i(l,l) = \begin{cases} 0 & \text{if } L < \frac{1}{1+a} \\ \text{any } w \in [0,1] & \text{if } L = \frac{1}{1+a} \\ 1 & \text{if } L > \frac{1}{1+a} \end{cases}$$

and transfers are given by equations 2.2, 2.3, 2.4, and 2.5. □

Proposition 2.B.2. *Common value case:*

(i) When $q \leq q^*(a)$, an optimal mechanism is, for $i \in \{1,2\}$:

- $x_i(l,l) = x_i(h,h) = x_i(l,h) = x_i(h,l) = \frac{1}{2}$
- $t_i(h,l) = t_i(h,h) = t_i(l,h) = t_i(l,l) = \frac{1}{2}(1-q)$

and the revenue of the seller is $1 - q$*.*

(ii) When $q^*(a) < q < q^{**}(a)$, an optimal mechanism is:

- $x_i(l,l) = x_i(h,h) = \frac{1}{2}$*,* $x_i(l,h) = 0$*, and* $x_i(h,l) = 1$
- $t_i(l,l) = t_i(l,h) = \frac{1}{2}\underline{\alpha}_1$
- $t_i(h,h) = \frac{q+\underline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_1-1}{2}$
- $t_i(h,l) = \frac{q+\underline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_1}{2}$

and the revenue of the seller is $\frac{1}{2}\left(\underline{\alpha}_0 + 2\underline{\alpha}_1 + \overline{\alpha}_1 + q - q^2 - (1-q)^2\right)$.

(iii) When $q \geq q^{**}(a)$, an optimal mechanism is:

- $x_i(l,l) = x_i(h,h) = \frac{1}{2}$*,* $x_i(l,h) = 0$*, and* $x_i(h,l) = 1$
- $t_i(h,l) = \dfrac{1}{2}(1+\underline{\alpha}_0)$
- $t_i(h,h) = \dfrac{\underline{\alpha}_0}{2}$
- $t_i(l,h) = \dfrac{\underline{\alpha}_1}{2} + (\underline{\alpha}_0 + \underline{\alpha}_1)\dfrac{1-q-\overline{\alpha}_1-\underline{\alpha}_1}{2(\overline{\alpha}_0+\overline{\alpha}_1+\underline{\alpha}_0+\underline{\alpha}_1-1)}$
- $t_i(l,l) = \dfrac{\underline{\alpha}_1}{2} - (1-\underline{\alpha}_0-\underline{\alpha}_1)\dfrac{1-q-\overline{\alpha}_1-\underline{\alpha}_1}{2(\overline{\alpha}_0+\overline{\alpha}_1+\underline{\alpha}_0+\underline{\alpha}_1-1)}$

and the revenue of the seller is $\frac{1}{2}\left(1-(q^2+(1-q)^2-\underline{\alpha}_0-\underline{\alpha}_1)\dfrac{\underline{\alpha}_0+\overline{\alpha}_0-q}{\overline{\alpha}_0+\overline{\alpha}_1+\underline{\alpha}_0+\underline{\alpha}_1-1}\right)$.

Proof. The proof of Proposition 2.B.2 proceeds as follows. First we solve for the optimal symmetric mechanism for the seller, and then we show that the optimal symmetric mechanism is fully optimal.

Recall from the main text that $q^*(a)$ is the solution to $q^2+(1-q)^2-\underline{\alpha}_0-\underline{\alpha}_1 = \underline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_0+\overline{\alpha}_1-1$ that lies between $\frac{1}{2}$ and 1, and $q^{**}(a)$ is the solution to $q^2+(1-q)^2-\underline{\alpha}_0-\underline{\alpha}_1 = 3q-2+\underline{\alpha}_1+\overline{\alpha}_1$ that lies between $\frac{1}{2}$ and 1. The expressions for $q^*(a)$ and $q^{**}(a)$ are as follows:

$$q^*(a) = \begin{cases} \frac{1}{2}\left(\frac{6-7a+2a^2}{3-2a+a^2} + \sqrt{\frac{-12a+17a^2-4a^3}{(3-2a+a^2)^2}}\right) & 1 < a \leq \frac{1}{2}(-1+\sqrt{13}) \\ \frac{1}{4}\left(\frac{-7+6a}{-2+a} + \sqrt{\frac{1-12a+12a^2}{(-2+a)^2}}\right) & \frac{1}{2}(-1+\sqrt{13}) < a < 2 \\ \frac{3}{5} & a = 2 \\ \frac{1}{4}\left(\frac{-7+6a}{-2+a} - \sqrt{\frac{1-12a+12a^2}{(-2+a)^2}}\right) & a > 2 \end{cases}$$

$$q^{**}(a) = \begin{cases} \frac{2-2a+a^2}{2-a+a^2} + \frac{\sqrt{\frac{-2a+3a^2-a^3)}{(2-a+a^2)^2}}}{\sqrt{2}} & 1 < a \leq \frac{1}{2}(-1+\sqrt{17}) \\ \frac{1}{2}\left(\frac{-5+3a}{-3+a} + \sqrt{\frac{1-4a+3a^2}{(-3+a)^2}}\right) & \frac{1}{2}(-1+\sqrt{17}) < a < 3 \\ \frac{5}{8} & a = 3 \\ \frac{1}{2}\left(\frac{-5+3a}{-3+a} - \sqrt{\frac{1-4a+3a^2}{(-3+a)^2}}\right) & a > 3 \end{cases}$$

First we find the optimal symmetric mechanism which will depend on the cutoff values of q defined above. We will ignore the incentive constraint of the low type and check ex post that it is satisfied. Therefore, the participation constraint of the low type must be binding.

Let U_i^h be the utility of the high type in equilibrium and U_i^l be the utility of the low type in equilibrium, that is:

$$U_i^l \equiv \min_{\alpha_0,\alpha_1} \alpha_1 x_i(l,l) + (1-q-\alpha_1)x_i(l,h) - (\alpha_0+\alpha_1))t_i(l,l) - (1-\alpha_0-\alpha_1)t_i(l,h)$$
$$U_i^h \equiv \min_{\alpha_0,\alpha_1} (1-q-\alpha_1)x_i(h,l) + (2q-1+\alpha_1)x_i(h,h) - (1-\alpha_0-\alpha_1))t_i(h,l) - (\alpha_0+\alpha_1)t_i(h,h)$$

Note that it is optimal to set $U_i^l = 0$. The incentive constraint of the high type is:

$$\min_{\alpha_0,\alpha_1} (1-q-\alpha_1)x_i(l,l) + (2q-1+\alpha_1)x_i(l,h) - (\alpha_0+\alpha_1))t_i(l,h) - (1-\alpha_0-\alpha_1)t_i(l,l) \leq U_i^h$$

The participation constraint of the low type is:

$$\min_{\alpha_0,\alpha_1} \alpha_1 x_i(l,l) + (1-q-\alpha_1)x_i(l,h) - (\alpha_0+\alpha_1))t_i(l,l) - (1-\alpha_0-\alpha_1)t_i(l,h) = 0$$

We can subtract the latter from the former to get:

$$\frac{(1-q-\underline{\alpha}_1-\overline{\alpha}_1)x_i(l,l) + (3q-2+\underline{\alpha}_1+\overline{\alpha}_1)x_i(l,h) - U_i^h}{\underline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_0+\overline{\alpha}_1-1} \leq t_i(l,h) - t_i(l,l)$$

Define:

$$\Delta t_{l,i} \equiv t_i(l,h) - t_i(l,l)$$
$$\Delta t_{h,i} \equiv t_i(h,l) - t_i(h,h)$$

We can write the expected transfers to the seller from each type as:

$$T_i^l = \min_{\alpha_0,\alpha_1} \alpha_1 x_i(l,l) + (1-q-\alpha_1)x_i(l,h) - \left(q^2 + (1-q)^2 - (\alpha_0+\alpha_1)\right)\Delta t_{l,i}$$
$$T_i^h = \min_{\alpha_0,\alpha_1} (1-q-\alpha_1)x_i(h,l) + (2q-1+\alpha_1)x_i(h,h) - \left(q^2 + (1-q)^2 - (\alpha_0+\alpha_1)\right)\Delta t_{h,i} - U_i^h$$

Let α_0^l, α_1^l and α_0^h, α_1^h be solutions to these minimisation problems. The seller chooses $x_i(l,l) \in [0,\frac{1}{2}]$, $x_i(h,h) \in [0,\frac{1}{2}]$, $x_i(l,h) \in [0,1]$, $x_i(h,l) \in [0, 1-x_i(l,h)]$, $\Delta t_{l,i} \in R$, $\Delta t_{h,i} \in R$, and $U_i^h \geq 0$ to maximise $T_i^l + T_i^h$ subject to:

$$\frac{(1-q-\underline{\alpha}_1-\overline{\alpha}_1)x_i(l,l) + (3q-2+\underline{\alpha}_1+\overline{\alpha}_1)x_i(l,h) - U_i^h}{\underline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_0+\overline{\alpha}_1-1} \leq \Delta t_{l,i}$$

Clearly, it is optimal to set $x_i(h,h) = \frac{1}{2}$ and $x_i(h,l) = 1 - x_i(l,h)$. Thus, the seller's problem is:

$$\max_{x_i(l,l),x_i(l,h),\Delta t_{l,i},\Delta t_{h,i},U_i^h} \Bigg\{ \min_{\alpha_0,\alpha_1} \alpha_1 x_i(l,l) + (1-q-\alpha_1)x_i(l,h) - \left(q^2+(1-q)^2-(\alpha_0+\alpha_1)\right)\Delta t_{l,i}$$
$$+ \min_{\alpha_0,\alpha_1} (1-q-\alpha_1)(1-x_i(h,l)) + \frac{1}{2}(2q-1+\alpha_1) - \left(q^2+(1-q)^2-(\alpha_0+\alpha_1)\right)\Delta t_{h,i} - U_i^h \Bigg\}$$

subject to:

$$\frac{(1-q-\underline{\alpha}_1-\overline{\alpha}_1)x_i(l,l) + (3q-2+\underline{\alpha}_1+\overline{\alpha}_1)x_i(l,h) - U_i^h}{\underline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_0+\overline{\alpha}_1-1} \leq \Delta t_{l,i}$$
$$0 \leq x_i(l,l) \leq \frac{1}{2}$$
$$0 \leq x_i(l,h) \leq 1$$

Define:

$$\Delta t_{l,i}^*(x_i(l,l), x_i(l,h), U_i^h) \equiv \max\Bigg\{0, \frac{(1-q-\underline{\alpha}_1-\overline{\alpha}_1)x_i(l,l) + (3q-2+\underline{\alpha}_1+\overline{\alpha}_1)x_i(l,h) - U_i^h}{\underline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_0+\overline{\alpha}_1-1}, x_i(l,h) - x_i(l,l)\Bigg\}$$

We show that $\Delta t_{l,i} = \Delta t_{l,i}^*(x_i(l,l), x_i(l,h), U_i^h)$ is optimal. Note that $\Delta t_{l,i} < 0$ implies $\alpha_0^l = \overline{\alpha}_0$, which implies that $q^2 + (1-q)^2 - \alpha_0^l - \alpha_1^l < 0$, so it is profitable to increase $\Delta t_{l,i}$ (which also slackens IC_h). Similarly when $\Delta t_{l,i} < x_i(l,h) - x_i(l,l)$, $\alpha_1^l = \overline{\alpha}_1$, which implies that $q^2+(1-q)^2 - \alpha_0^l - \alpha_1^l < 0$, so it is profitable to increase $\Delta t_{l,i}$. If $\Delta t_{l,i} > x_i(l,h) - x_i(l,l) \geq 0$, then it is profitable to decrease $\Delta t_{l,i}$, which is possible if $\Delta t_{l,i} > \frac{(1-q-\underline{\alpha}_1-\overline{\alpha}_1)x_i(l,l)+(3q-2+\underline{\alpha}_1+\overline{\alpha}_1)x_i(l,h)-U_i^h}{\underline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_0+\overline{\alpha}_1-1}$.

Define:

$$\Delta t_{h,i}^*(x_i(l,h)) \equiv \max\left\{0, \frac{1}{2} - x_i(l,h)\right\}$$

Similarly, it is optimal to set $\Delta t_{h,i} = \Delta t_{h,i}^*(x_i(l,h))$. Thus, the problem becomes:

$$\max_{x_i(l,l),x_i(l,h),U_i^h} R(x_i(l,l), x_i(l,h), U_i^h)$$

where:

$$\begin{aligned} R(x_i(l,l), x_i(l,h), U_i^h) \equiv \min_{\alpha_1} &\Big\{\alpha_1 x_i(l,l) + (1-q-\alpha_1)x_i(l,h) \\ &- \left(q^2 + (1-q)^2 - (\underline{\alpha}_0 + \alpha_1)\right)\Delta t^*_{l,i}(x_i(l,l), x_i(l,h), U_i^h)\Big\} \\ &+ \min_{\alpha_0,\alpha_1}\Big\{(1-q-\alpha_1)(1-x_i(l,h)) + \frac{2q-1+\alpha_1}{2} \\ &- \left(q^2 + (1-q)^2 - (\alpha_0 + \alpha_1)\right)\Delta t^*_{h,i}(x_i(l,h)) - U_i^h\Big\} \end{aligned}$$

Now we show that $x_i(l,l) = \frac{1}{2}$ is optimal. Note that:

$$\frac{\partial R}{\partial x_i(l,l)} = \begin{cases} \underline{\alpha}_1 - \frac{(q^2+(1-q)^2-\underline{\alpha}_0-\underline{\alpha}_1)(1-q-\underline{\alpha}_1-\overline{\alpha}_1)}{\underline{\alpha}_0+\overline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_1-1} > 0 & x_i(l,l) > \max\left\{x_i(l,h), \frac{-x_i(l,h)(3q-2+\underline{\alpha}_1+\overline{\alpha}_1)+U_i^h}{1-q-\underline{\alpha}_1-\overline{\alpha}_1}\right\} \\ \underline{\alpha}_1 > 0 & x_i(l,h) < x_i(l,l) < \frac{-x_i(l,h)(3q-2+\underline{\alpha}_1+\overline{\alpha}_1)+U_i^h}{1-q-\underline{\alpha}_1-\overline{\alpha}_1} \\ q^2 + (1-q)^2 - \underline{\alpha}_0 > 0 & \frac{-x_i(l,h)(3q-1-\underline{\alpha}_0-\overline{\alpha}_0)+U_i^h}{\underline{\alpha}_0+\overline{\alpha}_0-q} < x_i(l,l) < x_i(l,h) \\ \overline{\alpha}_1 - \frac{(q^2+(1-q)^2-\underline{\alpha}_0-\overline{\alpha}_1)(1-q-\underline{\alpha}_1-\overline{\alpha}_1)}{\underline{\alpha}_0+\overline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_1-1} > 0 & x_i(l,l) < \min\left\{x_i(l,h), \frac{-x_i(l,h)(3q-1-\underline{\alpha}_0-\overline{\alpha}_0)+U_i^h}{\underline{\alpha}_0+\overline{\alpha}_0-q}\right\} \end{cases}$$

Thus, the problem becomes:

$$\max_{x_i(l,h), U_i^h} R(x_i(l,h), U_i^h)$$

where:

$$\begin{aligned} R(x_i(l,h), U_i^h) \equiv \min_{\alpha_1} &\Big\{\frac{\alpha_1}{2} + (1-q-\alpha_1)x_i(l,h) \\ &- \left(q^2 + (1-q)^2 - (\underline{\alpha}_0 + \alpha_1)\right)\Delta t^*_{l,i}(x_i(l,h), U_i^h)\Big\} \\ &+ \min_{\alpha_1}\Big\{(1-q-\alpha_1)(1-x_i(l,h)) + \frac{2q-1+\alpha_1}{2} \\ &- \left(q^2 + (1-q)^2 - (\underline{\alpha}_0 + \alpha_1)\right)\Delta t^*_{h,i}(x_i(l,h)) - U_i^h\Big\} \end{aligned}$$

Now we find the optimal $x_i(l,h)$ as a function of U_i^h:

$$\frac{\partial R(x_i(l,h), U_i^h)}{\partial x_i(l,h)} = \begin{cases} -\frac{(q^2+(1-q)^2-\underline{\alpha}_0-\underline{\alpha}_1)(3q-2+\underline{\alpha}_1+\overline{\alpha}_1)}{\underline{\alpha}_0+\overline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_1-1} < 0 & x_i(l,h) > \max\left\{\frac{-\frac{1}{2}(\underline{\alpha}_0+\overline{\alpha}_0-q)+U_i^h}{3q-1-\underline{\alpha}_0-\overline{\alpha}_0}, \frac{1}{2}\right\} \\ -(q^2 + (1-q)^2 - \underline{\alpha}_0 - \underline{\alpha}_1) < 0 & \frac{1}{2} < x_i(l,h) < \frac{-\frac{1}{2}(\underline{\alpha}_0+\overline{\alpha}_0-q)+U_i^h}{3q-1-\underline{\alpha}_0-\overline{\alpha}_0} \\ -\frac{(q^2+(1-q)^2-\underline{\alpha}_0-\underline{\alpha}_1)(3q-1-\underline{\alpha}_0-\overline{\alpha}_0)}{\underline{\alpha}_0+\overline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_1-1} < 0 & \frac{-\frac{1}{2}(1-q-\underline{\alpha}_1-\overline{\alpha}_1)+U_i^h}{3q-2+\underline{\alpha}_1+\overline{\alpha}_1} < x_i(l,h) < \frac{1}{2} \\ q^2 + (1-q)^2 - \underline{\alpha}_0 - \underline{\alpha}_1 > 0 & x_i(l,h) < \min\left\{\frac{1}{2}, \frac{-\frac{1}{2}(1-q-\underline{\alpha}_1-\overline{\alpha}_1)+U_i^h}{3q-2+\underline{\alpha}_1+\overline{\alpha}_1}\right\} \end{cases}$$

Therefore:

$$x_i^*(l,h)(U^h) = \begin{cases} 0 & U_i^h \le \frac{1}{2}(1-q-\underline{\alpha}_1-\overline{\alpha}_1) \\ \frac{-\frac{1}{2}(1-q-\underline{\alpha}_1-\overline{\alpha}_1)+U_i^h}{3q-2+\underline{\alpha}_1+\overline{\alpha}_1} & \frac{1}{2}(1-q-\underline{\alpha}_1-\overline{\alpha}_1) < U_i^h < \frac{1}{2}(2q-1) \\ \frac{1}{2} & U_i^h \ge \frac{1}{2}(2q-1) \end{cases}$$

We can now write the problem just in terms of U_i^h:

$$\max_{U_i^h} R(U_i^h)$$

where:

$$
\begin{aligned}
R(U_i^h) \equiv \frac{\underline{\alpha}_1}{2} &+ (1-q-\underline{\alpha}_1)x_i^*(l,h)(U_i^h) - \left(q^2+(1-q)^2-(\underline{\alpha}_0+\underline{\alpha}_1)\right)\Delta t^*_{l,i}(U_i^h) \\
&+ (1-q-\underline{\alpha}_1)(1-x_i^*(l,h)(U_i^h)) + \frac{2q-1+\underline{\alpha}_1}{2} \\
&- \left(q^2+(1-q)^2-(\underline{\alpha}_0+\underline{\alpha}_1)\right)\Delta t^*_{h,i}(x^*(l,h)(U_i^h)) - U_i^h
\end{aligned}
$$

$$
\frac{\partial R(U_i^h)}{\partial U_i^h} = \begin{cases} \frac{q^2+(1-q)^2-\underline{\alpha}_0-\underline{\alpha}_1}{\underline{\alpha}_0+\overline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_1-1} - 1 & U_i^h < \frac{1}{2}(1-q-\underline{\alpha}_1-\overline{\alpha}_1) \\ \frac{q^2+(1-q)^2-\underline{\alpha}_0-\underline{\alpha}_1}{3q-2+\underline{\alpha}_1+\underline{\alpha}_1} - 1 & \frac{1}{2}(1-q-\underline{\alpha}_1-\overline{\alpha}_1) < U_i^h < \frac{1}{2}(2q-1) \\ -1 & U_i^h > \frac{1}{2}(2q-1) \end{cases}
$$

Thus,

$$
U^h = \begin{cases} 0 & q^2+(1-q)^2-\underline{\alpha}_0-\underline{\alpha}_1 \leq \underline{\alpha}_0+\overline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_1-1 \\ \frac{1}{2}(1-q-\underline{\alpha}_1-\overline{\alpha}_1) & \underline{\alpha}_0+\overline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_1-1 < q^2+(1-q)^2-\underline{\alpha}_0-\underline{\alpha}_1 < 3q-2+\underline{\alpha}_1+\underline{\alpha}_1 \\ \frac{1}{2}(2q-1) & q^2+(1-q)^2-\underline{\alpha}_0-\underline{\alpha}_1 \geq 3q-2+\underline{\alpha}_1+\underline{\alpha}_1 \end{cases}
$$

This is equivalent to:

$$
U^h = \begin{cases} 0 & q \geq q^{**}(a) \\ \frac{1}{2}(1-q-\underline{\alpha}_1-\overline{\alpha}_1) & q^*(a) < q < q^{**}(a) \\ \frac{1}{2}(2q-1) & q \leq q^*(a) \end{cases}
$$

Thus, in the optimal symmetric mechanism, $x_i(l,l) = x_i(h,h) = \frac{1}{2}$, and:

$$
x_i(l,h) = \begin{cases} 0 & q \geq q^{**}(a) \\ \frac{1}{2} & q^*(a) < q < q^{**}(a) \\ \frac{1}{2} & q \leq q^*(a) \end{cases}
$$

$$
\Delta t_{l,i} = \begin{cases} \frac{1-q-\underline{\alpha}_1-\overline{\alpha}_1}{2(\underline{\alpha}_0+\underline{\alpha}_1+\overline{\alpha}_0+\overline{\alpha}_1-1)} & q \geq q^{**}(a) \\ 0 & q^*(a) < q < q^{**}(a) \\ 0 & q \leq q^*(a) \end{cases}
$$

$$
\Delta t_{h,i} = \begin{cases} \frac{1}{2} & q \geq q^{**}(a) \\ \frac{1}{2} & q^*(a) < q < q^{**}(a) \\ 0 & q \leq q^*(a) \end{cases}
$$

To recover the transfers $t_i(l,l)$, $t_i(l,h)$, $t_i(h,l)$, and $t_i(h,h)$, use:

$$
\min_{\alpha_0,\alpha_1} \alpha_1 x_i(l,l) + (1-q-\alpha_1)x_i(l,h) - (\alpha_0+\alpha_1))t_i(l,l) - (1-\alpha_0-\alpha_1)(t_i(l,l)+\Delta t_{l,i}) = 0
$$

$$
\begin{aligned}
\min_{\alpha_0,\alpha_1} \Big\{ &(1-q-\alpha_1)x_i(h,l) + (2q-1+\alpha_1)x_i(h,h) \\
&- (1-\alpha_0-\alpha_1))(t_i(h,h)+\Delta t_{h,i}) - (\alpha_0+\alpha_1)t_i(h,h) \Big\} = U_i^h
\end{aligned}
$$

Now we show that the optimal symmetric mechanism is fully optimal. Suppose that there exists an asymmetric mechanism (x, t) that is optimal. Define:

$$\overline{x}(\cdot,\cdot) \equiv \frac{1}{2}x_1(\cdot,\cdot) + \frac{1}{2}x_2(\cdot,\cdot)$$
$$\overline{t}(\cdot,\cdot) \equiv \frac{1}{2}t_1(\cdot,\cdot) + \frac{1}{2}t_2(\cdot,\cdot)$$

Consider the following symmetric mechanism:

$$x_i'(\cdot,\cdot) = \overline{x}(\cdot,\cdot)$$
$$t_i'(l,\cdot) = \overline{t}(l,\cdot) + \min_{\alpha_0,\alpha_1}\Big\{\alpha_1\overline{x}(l,l) + (1-q-\alpha_1)\overline{x}(l,h) - (\alpha_0+\alpha_1)\overline{t}(l,l) - (1-\alpha_0-\alpha_1)\overline{t}(l,h) - \frac{1}{2}\sum_{i=1}^{2}U_i^l\Big\}$$
$$t_i'(h,\cdot) = \overline{t}(h,\cdot) + \min_{\alpha_0,\alpha_1}\Big\{(1-q-\alpha_1)\overline{x}(h,l) + (2q-1+\alpha_1)\overline{x}(h,h) - (1-\alpha_0+\alpha_1)\overline{t}(h,l) - (\alpha_0+\alpha_1)\overline{t}(h,h) - \frac{1}{2}\sum_{i=1}^{2}U_i^h\Big\}$$

By construction, the high type gets $\frac{1}{2}\sum_{i=1}^{2}U_i^h$ in equilibrium and the low type gets $\frac{1}{2}\sum_{i=1}^{2}U_i^l$ in equilibrium; therefore both participation constraints are satisfied.

Define $\Delta\overline{x}_l \equiv \overline{x}(l,h) - \overline{x}(l,l)$, $\Delta\overline{t}_l \equiv \overline{t}(l,h) - \overline{t}(l,l)$, $\Delta x_{l,i} \equiv x_i(l,h) - x_i(l,l)$, and $\Delta t_{l,i} \equiv t_i(l,h) - t_i(l,l)$. To see that IC_h is satisfied, first note that:

$$\min_{\alpha_0,\alpha_1}(1-q-\alpha_1)\overline{x}(l,l) + (2q-1+\alpha_1)\overline{x}(l,h) - (1-\alpha_0-\alpha_1)\overline{t}(l,l) - (\alpha_0+\alpha_1)\overline{t}(l,h)$$
$$= \frac{1}{2}\sum_{i=1}^{2}\min_{\alpha_0,\alpha_1}(1-q-\alpha_1)x_i(l,l) + (2q-1+\alpha_1)x_i(l,h) - (1-\alpha_0-\alpha_1)t_i(l,l) - (\alpha_0+\alpha_1)t_i(l,h)$$
$$+ \min_{\alpha_0,\alpha_1}\alpha_1\Delta\overline{x}_l - (\alpha_0+\alpha_1)\Delta\overline{t}_l - \frac{1}{2}\sum_{i=1}^{2}\min_{\alpha_0,\alpha_1}\alpha_1\Delta x_{l,i} - (\alpha_0+\alpha_1)\Delta t_{l,i}$$

and by definition:

$$\frac{1}{2}\sum_{i=1}^{2}\min_{\alpha_0,\alpha_1}\alpha_1 x_i(l,l) + (1-q-\alpha_1)x_i(l,h) - (\alpha_0+\alpha_1)t_i(l,l) - (1-\alpha_0-\alpha_1)t_i(l,h) - U_i^l = 0$$

Therefore:

$$\min_{\alpha_0,\alpha_1} (1-q-\alpha_1)\overline{x}(l,l) + (2q-1+\alpha_1)\overline{x}(l,h) - (1-\alpha_0-\alpha_1)\overline{t}(l,l) - (\alpha_0+\alpha_1)\overline{t}(l,h)$$

$$-\left(\min_{\alpha_0,\alpha_1} \alpha_1\overline{x}(l,l) + (1-q-\alpha_1)\overline{x}(l,h) - (\alpha_0+\alpha_1)\overline{t}(l,l) - (1-\alpha_0-\alpha_1)\overline{t}(l,h)\right) + \frac{1}{2}\sum_{i=1}^{2} U_i^l$$

$$= \frac{1}{2}\sum_{i=1}^{2} \min_{\alpha_0,\alpha_1} (1-q-\alpha_1)x_i(l,l) + (2q-1+\alpha_1)x_i(l,h) - (1-\alpha_0-\alpha_1)t_i(l,l) - (\alpha_0+\alpha_1)t_i(l,h)$$

$$+ \min_{\alpha_0,\alpha_1} \alpha_1\Delta\overline{x}_l - (\alpha_0+\alpha_1)\Delta\overline{t}_l - \frac{1}{2}\sum_{i=1}^{2}\min_{\alpha_0,\alpha_1}\alpha_1\Delta x_{l,i} - (\alpha_0+\alpha_1)\Delta t_{l,i}$$

$$+ \frac{1}{2}\sum_{i=1}^{2} \min_{\alpha_0,\alpha_1} \alpha_1 x_i(l,l) + (1-q-\alpha_1)x_i(l,h) - (\alpha_0+\alpha_1)t_i(l,l) - (1-\alpha_0-\alpha_1)t_i(l,h) - U_i^l$$

$$-\left(\min_{\alpha_0,\alpha_1} \alpha_1\overline{x}(l,l) + (1-q-\alpha_1)\overline{x}(l,h) - (\alpha_0+\alpha_1)\overline{t}(l,l) - (1-\alpha_0-\alpha_1)\overline{t}(l,h)\right) + \frac{1}{2}\sum_{i=1}^{2} U_i^l$$

$$= \frac{1}{2}\sum_{i=1}^{2} \min_{\alpha_0,\alpha_1} (1-q-\alpha_1)x_i(l,l) + (2q-1+\alpha_1)x_i(l,h) - (1-\alpha_0-\alpha_1)t_i(l,l) - (\alpha_0+\alpha_1)t_i(l,h)$$

$$+ \min_{\alpha_0,\alpha_1} \alpha_1\Delta\overline{x}_l - (\alpha_0+\alpha_1)\Delta\overline{t}_l - \frac{1}{2}\sum_{i=1}^{2}\min_{\alpha_0,\alpha_1}\alpha_1\Delta x_{l,i} - (\alpha_0+\alpha_1)\Delta t_{l,i}$$

$$+ \frac{1}{2}\sum_{i=1}^{2} \min_{\alpha_0,\alpha_1} -\alpha_1\Delta x_{l,i} + (\alpha_0+\alpha_1)\Delta t_{l,i} - \min_{\alpha_0,\alpha_1} -\alpha_1\Delta\overline{x}_l + (\alpha_0+\alpha_1)\Delta\overline{t}_l \leq \frac{1}{2}\sum_{i=1}^{2} U_i^h$$

The last inequality follows because:

$$\min_{\alpha_0,\alpha_1} \alpha_1\Delta\overline{x}_l - (\alpha_0+\alpha_1)\Delta\overline{t}_l - \frac{1}{2}\sum_{i=1}^{2}\min_{\alpha_0,\alpha_1}\alpha_1\Delta x_{l,i} - (\alpha_0+\alpha_1)\Delta t_{l,i}$$

$$+ \frac{1}{2}\sum_{i=1}^{2} \min_{\alpha_0,\alpha_1} -\alpha_1\Delta x_{l,i} + (\alpha_0+\alpha_1)\Delta t_{l,i} - \min_{\alpha_0,\alpha_1} -\alpha_1\Delta\overline{x}_l + (\alpha_0+\alpha_1)\Delta\overline{t}_l$$

$$= \min_{\alpha_0,\alpha_1} \alpha_1\Delta\overline{x}_l - (\alpha_0+\alpha_1)\Delta\overline{t}_l - \frac{1}{2}\sum_{i=1}^{2}\min_{\alpha_0,\alpha_1}\alpha_1\Delta x_{l,i} - (\alpha_0+\alpha_1)\Delta t_{l,i}$$

$$+ \max_{\alpha_0,\alpha_1} \alpha_1\Delta\overline{x}_l - (\alpha_0+\alpha_1)\Delta\overline{t}_l - \frac{1}{2}\sum_{i=1}^{2}\max_{\alpha_0,\alpha_1}\alpha_1\Delta x_{l,i} - (\alpha_0+\alpha_1)\Delta t_{l,i}$$

$$= (\underline{\alpha}_1 + \overline{\alpha}_1)\Delta\overline{x}_l - (\underline{\alpha}_0 + \overline{\alpha}_0 + \underline{\alpha}_1 + \overline{\alpha}_1)\Delta\overline{t}_l - \frac{1}{2}\sum_{i=1}^{2}(\underline{\alpha}_1 + \overline{\alpha}_1)\Delta x_{l,i} - (\underline{\alpha}_0 + \overline{\alpha}_0 + \underline{\alpha}_1 + \overline{\alpha}_1)\Delta t_{l,i} = 0$$

The proof that IC_l is satisfied is analogous. Define $\Delta\overline{x}_h \equiv \overline{x}(h,l) - \overline{x}(h,h)$, $\Delta\overline{t}_h \equiv \overline{t}(h,l) - \overline{t}(h,h)$,

$\Delta x_{h,i} \equiv x_i(h,l) - x_i(h,h)$, and $\Delta t_{h,i} \equiv t_i(h,l) - t_i(h,h)$. Then:

$$
\begin{aligned}
&\min_{\alpha_0,\alpha_1} \alpha_1\overline{x}(h,l) + (1-q-\alpha_1)\overline{x}(h,h) - (\alpha_0+\alpha_1)\overline{t}(h,l) - (1-\alpha_0-\alpha_1)\overline{t}(h,h) \\
&- \left(\min_{\alpha_0,\alpha_1} (1-q-\alpha_1)\overline{x}(h,l) + (2q-1+\alpha_1)\overline{x}(h,h) - (1-\alpha_0-\alpha_1)\overline{t}(h,l) - (\alpha_0+\alpha_1)\overline{t}(h,h)\right) \\
&+ \frac{1}{2}\sum_{i=1}^{2} U_i^h \\
&= \frac{1}{2}\sum_{i=1}^{2} \min_{\alpha_0,\alpha_1} \alpha_1 x_i(h,l) + (1-q-\alpha_1)x_i(h,h) - (\alpha_0+\alpha_1)t_i(h,l) - (1-\alpha_0-\alpha_1)t_i(h,h) \\
&+ \min_{\alpha_0,\alpha_1} \alpha_1\Delta\overline{x}_h - (\alpha_0+\alpha_1)\Delta\overline{t}_h - \frac{1}{2}\sum_{i=1}^{2}\min_{\alpha_0,\alpha_1}\alpha_1\Delta x_{h,i} - (\alpha_0+\alpha_1)\Delta t_{h,i} \\
&+ \frac{1}{2}\sum_{i=1}^{2}\Bigg(\min_{\alpha_0,\alpha_1} (1-q-\alpha_1)x_i(h,l) + (2q-1+\alpha_1)x_i(h,h) - (1-\alpha_0-\alpha_1)t_i(h,l) - (\alpha_0+\alpha_1)t_i(h,h) \\
&- U_i^h\Bigg) \\
&- \left(\min_{\alpha_0,\alpha_1} (1-q-\alpha_1)\overline{x}(h,l) + (2q-1+\alpha_1)\overline{x}(h,h) - (1-\alpha_0-\alpha_1)\overline{t}(h,l) - (\alpha_0+\alpha_1)\overline{t}(h,h)\right) \\
&+ \frac{1}{2}\sum_{i=1}^{2} U_i^h \\
&= \frac{1}{2}\sum_{i=1}^{2} \min_{\alpha_0,\alpha_1} \alpha_1 x_i(h,l) + (1-q-\alpha_1)x_i(h,h) - (\alpha_0+\alpha_1)t_i(h,l) - (1-\alpha_0-\alpha_1)t_i(h,h) \\
&+ \min_{\alpha_0,\alpha_1} \alpha_1\Delta\overline{x}_h - (\alpha_0+\alpha_1)\Delta\overline{t}_h - \frac{1}{2}\sum_{i=1}^{2}\min_{\alpha_0,\alpha_1}\alpha_1\Delta x_{h,i} - (\alpha_0+\alpha_1)\Delta t_{h,i} \\
&+ \frac{1}{2}\sum_{i=1}^{2}\min_{\alpha_0,\alpha_1} -\alpha_1\Delta x_{h,i} + (\alpha_0+\alpha_1)\Delta t_{h,i} - \min_{\alpha_0,\alpha_1} -\alpha_1\Delta\overline{x}_h + (\alpha_0+\alpha_1)\Delta\overline{t}_h \le \frac{1}{2}\sum_{i=1}^{2} U_i^l
\end{aligned}
$$

Finally, note that $\frac{1}{2}\sum_{i=1}^{2} t_i'(l,l) \ge \overline{t}(l,l)$, $\frac{1}{2}\sum_{i=1}^{2} t_i'(l,h) \ge \overline{t}(l,h)$, $\frac{1}{2}\sum_{i=1}^{2} t_i'(h,l) \ge \overline{t}(h,l)$, and $\frac{1}{2}\sum_{i=1}^{2} t_i'(h,h) \ge \overline{t}(h,h)$, so the symmetric mechanism (x',t') is incentive compatible and yields weakly greater revenue to the seller than (x,t). □

Appendix 2.C Proofs for Section 2.6

Many bidders

Each of n bidders has a common valuation $v \in \{0,1\}$ for an object, where $\Pr(v=0) = \frac{1}{2}$. A public signal $s^* \in \{l,h\}$ is drawn so that $\Pr(s^* = h|v=1) = \Pr(s^* = l|v=0) = q > \frac{1}{2}$. With probability α_v, all bidders receive the signal s^* in state v: that is, with probability α_v each bidder's signal is $s_i = s^*$. Otherwise each bidder's signal is drawn independently from the same distribution as the public signal so that $\Pr(s_i = h|v=1) = \Pr(s_i = l|v=0) = q > \frac{1}{2}$. The bidders consider as possible the following values of α_v:

$$
\begin{aligned}
0 \le \alpha_0 \le \overline{\alpha}_0 \\
0 \le \alpha_1 \le \overline{\alpha}_1
\end{aligned}
$$

Let $\overline{\alpha}_0 = \overline{\alpha}_1 \equiv a$. Define $P(k\ l, n-k-1\ h|h,(\alpha_0,\alpha_1))$ as the probability of k l signals (ordered), and $n-k-1$ h signals, conditional on an h signal under beliefs (α_0,α_1), and let $E_{(\alpha_0,\alpha_1)}[v|k\ l, n-k\ h]$ be the corresponding conditional expectation. That is:

$$P(k\ l, n-k-1\ h|h,(\alpha_0,\alpha_1)) = (1-q)((1-\alpha_0)q^k(1-q)^{n-k-1} + \alpha_0\mathcal{I}(k=0)) \\ + q((1-\alpha_1)q^{n-k-1}(1-q)^k + \alpha_1\mathcal{I}(k=0))$$

and $E_{(\alpha_0,\alpha_1)}[v|k\ l, n-k\ h] = \frac{q((1-\alpha_1)q^{n-k-1}(1-q)^k+\alpha_1\mathcal{I}(k=0))}{(1-q)((1-\alpha_0)q^k(1-q)^{n-k-1}+\alpha_0\mathcal{I}(k=0))+q((1-\alpha_1)q^{n-k-1}(1-q)^k+\alpha_1\mathcal{I}(k=0))}$

Proposition 2.6.1 will be implied by Propositions 2.C.1 and 2.C.3.

Proposition 2.C.1. *When ambiguity is not too large, there exists a symmetric equilibrium of the first price auction where for sufficiently large n:*

- *The low type has beliefs $(\alpha_0,\alpha_1) = (\overline{\alpha}_0, 0)$ and bids $b_a(l) = E_{(\overline{\alpha}_0,0)}[v|l,\ldots,l]$.*
- *The high type has beliefs $(\alpha_0,\alpha_1) = (0,\overline{\alpha}_1)$ and mixes on the support $[b_a(l), \overline{b}_F]$ according to distribution $F(b)$, where:*

$$\overline{b}_F = \sum_{k=0}^{n-2}\binom{n-1}{k}P(k\ l, n-k-1\ h|h,(0,\overline{\alpha}_1))E_{(0,\overline{\alpha}_1)}[v|k\ l, n-k\ h] + P(n-1\ l|h,(0,\overline{\alpha}_1))b_a(l)$$

and for $b \in [b_a(l), \overline{b}_F]$, $F(b)$ solves:

$$\sum_{k=0}^{n-2}\binom{n-1}{k}P(k\ l, n-k-1\ h|h,(0,\overline{\alpha}_1))(E_{(0,\overline{\alpha}_1)}[v|k\ l, n-k\ h] - b)F(b)^{n-k-1} \\ = P(n-1\ l|h,(0,\overline{\alpha}_1))(b - b_a(l))$$

Proof. For each type we first find the equilibrium beliefs and then show that there does not exist a profitable deviation.

Note that $\Pr(l,\ldots,l|l,(\alpha_0,\alpha_1)) = (1-q)(\alpha_1 + (1-\alpha_1)(1-q)^{n-1}) + q(\alpha_0 + (1-\alpha_0)q^{n-1})$, and $E_{(\alpha_0,\alpha_1)}(v|l,\ldots,l) = \frac{(1-q)(\alpha_1+(1-\alpha_1)(1-q)^{n-1})}{(1-q)(\alpha_1+(1-\alpha_1)(1-q)^{n-1})+q(\alpha_0+(1-\alpha_0)q^{n-1})}$. Thus, the equilibrium utility of the low type is:

$$\min_{\alpha_0,\alpha_1}(1-q)(\alpha_1 + (1-\alpha_1)(1-q)^{n-1})(1-b_a(l)) - q(\alpha_0 + (1-\alpha_0)q^{n-1})b_a(l)$$

Note that the derivative with respect to α_0 has the same sign as $-q(1-q^{n-1}) < 0$ and the derivative with respect to α_1 has the same sign as $(1-q)(1-(1-q)^{n-1}) > 0$. Thus, in equilibrium, the low type's beliefs are given by $(\alpha_0,\alpha_1) = (\overline{\alpha}_0, 0)$. Under the equilibrium bid, the low type gets zero utility. Note that any bid below $b_a(l)$ yields zero utility, and any bid above $b_a(l)$ yields negative utility under the equilibrium beliefs (for sufficiently small a). Therefore any deviation is also not profitable evaluated under the utility minimising beliefs.

The equilibrium utility of the high type (under no 'no hedging') is:

$$\min_{\alpha_0,\alpha_1}\sum_{k=0}^{n-1}\frac{1}{n-k}\binom{n-1}{k}P(k\ l, n-k-1\ h|h,(\alpha_0,\alpha_1))\left[E_{(\alpha_0,\alpha_1)}[v|k\ l, n-k\ h] - \left(\int_{b_a(l)}^{\overline{b}_F} b dF^{n-k}(b)\right)\right]$$

The derivative of the utility with respect to α_0 is:

$$\sum_{k=0}^{n-1}\frac{1}{n-k}\binom{n-1}{k}\Bigg[-\frac{\partial P(k\,l,n-k-1\,h|h,(\alpha_0,\alpha_1))}{\partial\alpha_0}\left(\int_{b_a(l)}^{\overline{b}_F} bdF^{n-k}(b)\right)\Bigg]$$
$$=\sum_{k=0}^{n-1}\frac{1}{n-k}\binom{n-1}{k}\Bigg[-(1-q)\left(-(1-q)^{n-k-1}q^k+\mathcal{I}(k=0)\right)\left(\int_{b_a(l)}^{\overline{b}_F} bdF^{n-k}(b)\right)\Bigg]$$
$$=\sum_{k=1}^{n-1}\frac{1}{n-k}\binom{n-1}{k}\Bigg[(1-q)^{n-k}q^k\left(\int_{b_a(l)}^{\overline{b}_F} bdF^{n-k}(b)\right)\Bigg]-\frac{1}{n}(1-q)\left(1-(1-q)^{n-1}\right)\left(\int_{b_a(l)}^{\overline{b}_F} bdF^{n}(b)\right).$$

In Lemma 2.C.2 below we show that the above derivative is positive. The derivative of the utility with respect to α_1 is:

$$=\min_{\alpha_0,\alpha_1}\sum_{k=0}^{n-1}\frac{1}{n-k}\binom{n-1}{k}\Bigg[\frac{\partial P(k\,l,n-k-1\,h|h,(\alpha_0,\alpha_1))E_{(\alpha_0,\alpha_1)}[v|k\,l,n-k\,h]}{\partial\alpha_1}$$
$$-\frac{\partial P(k\,l,n-k-1\,h|h,(\alpha_0,\alpha_1))}{\partial\alpha_1}\left(\int_{b_a(l)}^{\overline{b}_F} bdF^{n-k}(b)\right)\Bigg]$$
$$=\sum_{k=0}^{n-1}\frac{1}{n-k}\binom{n-1}{k}\Bigg[\frac{\partial P(k\,l,n-k-1\,h|h,(\alpha_0,\alpha_1))}{\partial\alpha_1}\left(1-\int_{b_a(l)}^{\overline{b}_F} bdF^{n-k}(b)\right)\Bigg]$$
$$=\sum_{k=0}^{n-1}\frac{1}{n-k}\binom{n-1}{k}\Bigg[q\left(-(1-q)^{k}q^{n-k-1}+\mathcal{I}(k=0)\right)\left(1-\int_{b_a(l)}^{\overline{b}_F} bdF^{n-k}(b)\right)\Bigg]<0$$
$$=-\sum_{k=1}^{n-1}\frac{1}{n-k}\binom{n-1}{k}\Bigg[(1-q)^{k}q^{n-k}\left(1-\int_{b_a(l)}^{\overline{b}_F} bdF^{n-k}(b)\right)\Bigg]+\frac{1}{n}q\left(1-q^{n-1}\right)\left(1-\int_{b_a(l)}^{\overline{b}_F} bdF^{n}(b)\right)\Bigg]$$

Note that by Lemma 2.C.2 this derivative is negative.

Note now that F is increasing, $F(b_a(l))=0$, and $F(\overline{b}_F)=1$. Moreover, $b_a(l)<E[v|n-1\ l,h,(0,\overline{\alpha}_1)]$ (which is true when ambiguity is not too large) ensures that the utility of high type is greater than 0. Any bid below $b_a(l)$ gives zero utility, and any bid above $\overline{b}_F$ gives strictly lower utility than bidding $\overline{b}_F$. By construction, under beliefs $(\alpha_0,\alpha_1)=(0,\overline{\alpha}_1)$, any bid in the support $[b_a(l),\overline{b}_F]$ gives the same utility, and any other mixed strategy can give at most this utility. □

Lemma 2.C.2. *There exists an $\bar{n}$ such that for all $n>\bar{n}$,*

$$\sum_{k=1}^{n-1}\frac{1}{n-k}\binom{n-1}{k}(1-q)^{n-k}q^k\left(\int_{b_a(l)}^{\overline{b}_F} bdF^{n-k}(b)\right)-\frac{1}{n}(1-q)\left(1-(1-q)^{n-1}\right)\left(\int_{b_a(l)}^{\overline{b}_F} bdF^{n}(b)\right)>0.$$

Proof. Below we will focus just on some of the expressions in the sum above that are close to $k=nq$. These expressions will have the term $\int_{b_a(l)}^{\overline{b}_F} bdF^{n(1-q)}(b)$ in them. We next prove that this term will be bounded away from zero.

Claim: $\lim_{n\to\infty}\left(\int_{b_a(l)}^{\overline{b}_F} bdF^{n(1-q)}(b)\right)>0$ and $\lim_{n\to\infty}\left(\int_{b_a(l)}^{\overline{b}_F} bdF^{n}(b)\right)>0$.

Proof of Claim. Assume that a is very small and consider the equation for the high type to be indifferent between any bid in $[b_a(l), \overline{b}_F]$: for any $b \in [b_a(l), \overline{b}_F]$, $F(b)$ solves:

$$\sum_{k=0}^{n-1} \binom{n-1}{k} P(k\ l, n-k-1\ h|h,(0,\overline{\alpha}_1))(E_{(0,\overline{\alpha}_1)}[v|k\ l, n-k\ h]-b)F(b)^{n-k-1}$$
$$= P(n-1\ l|h,(0,\overline{\alpha}_1))(b-b_a(l))$$
$$\Longleftrightarrow$$
$$(*)\quad q\overline{\alpha}_1(1-b)F(b)^{n-1} - (1-q)^n F(b)^{n-1} b + (1-\bar{\alpha}_1) q^n F(b)^{n-1}(1-b)$$
$$+\sum_{k=1}^{n-1} \binom{n-1}{k} P(k\ l, n-k-1\ h|h,(0,\overline{\alpha}_1))(E_{(0,\overline{\alpha}_1)}[v|k\ l, n-k\ h]-b)F(b)^{n-k-1}$$
$$= P(n-1\ l|h,(0,\overline{\alpha}_1))(b-b_a(l))$$

Fact: For any $b < \overline{b}_F$, either $F(b)^n \to 0$ or $F(b)^n \to \zeta(b) < 1$.

Proof of Fact. Suppose that for some b, $\overline{b}_F > b > b_a(l)$, $F(b)^n \to \zeta(b) > 0$. Note that as n grows large, by the law of large numbers, the LHS of (*) converges to

$$q\overline{\alpha}_1(1-b)\zeta(b) + q(1-\overline{\alpha}_1)(1-b)(\zeta(b))^q - (1-q)b(\zeta(b))^{1-q}$$

while the RHS converges to zero. Therefore, (*) implies

$$\overline{\alpha}_1(\zeta(b))^q + (1-\overline{\alpha}_1)(\zeta(b))^{2q-1} = \frac{1-q}{q}\frac{b}{1-b}$$

As $\overline{b}_F = E[V|h] = q$, if $b < \overline{b}_F$ the solution will be $\zeta(b) < 1$. □

Given the fact, for any b either $F(b)^n \to 0$ or $F(b)^n \to \zeta(b) < 1$. Note that $\frac{P(v=0|h,(0,\overline{\alpha}_1)))}{P(v=1|h,(0,\overline{\alpha}_1)))}\frac{b}{1-b}$ is increasing in b and reaches its maximum for $b = \overline{b}_F$ where the expression is equal 1. Therefore, we conclude that for any $b < \overline{b}_F$ $\lim_{n\to\infty} F(b)^{n(2q-1)} < 1$ and so $\lim_{n\to\infty}\left(\int_{b_a(l)}^{\overline{b}_F} b dF^{n(1-q)}(b)\right) > 0$ and $\lim_{n\to\infty}\left(\int_{b_a(l)}^{\overline{b}_F} b dF^{n}(b)\right) > 0$. This completes the proof of the Claim. □

For now look at the comparison between
$\sum_{k=1}^{n-1}\frac{1}{n-k}\binom{n-1}{k}(1-q)^{n-k}q^k\left(\int_{b_a(l)}^{\overline{b}_F} b dF^{n-k}(b)\right)$ and $\frac{1}{n}(1-q)\left(1-(1-q)^{n-1}\right)\left(\int_{b_a(l)}^{\overline{b}_F} b dF^{n}(b)\right)$ as n grows large.

Above we have seen that $\left(\int_{b_a(l)}^{\overline{b}_F} b dF^{n}(b)\right)$ is bounded from zero, therefore the second term is the same magnitude as $\frac{1}{n}$.

We will consider only part of the first term. Consider the expression for $k = nq$. By Stirling's approximation:

$$\binom{n-1}{k} = \frac{(n-1)!}{(n-k-1)!k!} \simeq \frac{\sqrt{2\pi n}(\frac{n}{e})^n}{\sqrt{2\pi n(1-q)}(\frac{n(1-q)}{e})^{n(1-q)}\sqrt{2\pi nq}(\frac{nq}{e})^{nq}} = \sqrt{\frac{1}{nq(1-q)}}\frac{1}{(1-q))^{n(1-q)}(q)^{nq}}$$

so that the term in the sum relating to $k = nq$ becomes: $\frac{1}{n-k}\binom{n-1}{k}\left[(1-q)^{n-k}q^k\right] \simeq \frac{1}{n(1-q)}\sqrt{\frac{1}{nq(1-q)}}$.
For any n take a sequence n_1 such that $\frac{n_1}{n} \to 0$ and $\frac{n_1}{\sqrt{n}} \to \infty$. Consider all the terms in the sum

$k \in \{nq - n_1, ..., nq + n_1\}$, note that for all these we have that:

$$\sum_{k\in\{nq-n_1,...,nq+n_1\}} \frac{1}{n-k}\binom{n-1}{k}\left[(1-q)^{n-k}q^k\right] \simeq 2n_1\frac{1}{n(1-q)}\sqrt{\frac{1}{nq(1-q)}}$$

where $\frac{2n_1\frac{1}{n(1-q)}\sqrt{\frac{1}{nq(1-q)}}}{\frac{1}{n}} = \frac{2n_1\frac{1}{(1-q)}\sqrt{\frac{1}{q(1-q)}}}{\sqrt{n}} \to \infty$. As $\left(\int_{b_a(l)}^{\bar{b}_F} bdF^{n(1-q)}(b)\right)$ is bounded from zero we are done. □

Proposition 2.C.3. *In the equilibrium described in Proposition 6A, the seller's revenue is decreasing in a.*

Proof. We show that the expected payment of each type, evaluated under the true joint probability distribution, is decreasing in a. For the low type, the expected payment is $P(l,\ldots,l|l)E_{(\overline{\alpha}_0(a),0)}[v|l,\ldots,l]$ which is decreasing in a.

To make clear that the equilibrium strategy depends on n and a, denote the bidding strategy for the high type by $F(b;n,a)$. The expected payment of the high type is:

$$\sum_{k=0}^{n-1}\frac{1}{n-k}\binom{n-1}{k}P(k\ l, n-k-1\ h|h)\int_{b_a(l)}^{\bar{b}_F} bdF^{n-k}(b;n,a)$$

As in the proof of Lemma 6A, consider the equation defining $F(b;n,a)$:

$$\begin{aligned}
&q\overline{\alpha}_1(1-b)F(b;n,a)^{n-1} - (1-q)^n F(b;n,a)^{n-1}b + (1-\bar{\alpha}_1)q^n F(b;n,a)^{n-1}(1-b)\\
&+\sum_{k=1}^{n-1}\binom{n-1}{k}P(k\ l, n-k-1\ h|h,(0,\overline{\alpha}_1))(E_{(0,\overline{\alpha}_1(a))}[v|k\ l, n-k\ h]-b)F(b;n,a)^{n-k-1}\\
&= P(n-1\ l|h,(0,\overline{\alpha}_1))(b-b_a(l))
\end{aligned}$$

By the law of large numbers, for large n, the LHS is close to:

$$q\overline{\alpha}_1(1-b)F(b;n,a)^{n-1} + q(1-\overline{\alpha}_1)(1-b)F(b;n,a)^{q(n-1)} - (1-q)bF(b;n,a)^{(1-q)(n-1)}$$

and the RHS is close to 0. Thus, $F(b;n,a)$ is close to $F^*(b;n,a)$, where for $b \in [0,q]$, $F^*(b;n,a)$ is the unique positive solution to:

$$q\overline{\alpha}_1(1-b)F(b;n,a)^{n-1} + q(1-\overline{\alpha}_1)(1-b)F(b;n,a)^{q(n-1)} - (1-q)bF(b;n,a)^{(1-q)(n-1)} = 0$$

Note that $F^*(b;n,a)$ is increasing in a. Now when n is large, $\int_{b_a(l)}^{\bar{b}_F} bdF^{n-k}(b;n,a)$ is close to $\int_0^q bdF^{*n-k}(b;n,a)$. Take $a' > a$; since $F^{*n-k}(b;n,a)$ first order stochastically dominates $F^{*n-k}(b;n,a')$, $\int_0^q bdF^{*n-k}(b;n,a') < \int_0^q bdF^{*n-k}(b;n,a)$, which implies $\int_{b_a(l)}^{\bar{b}_F} bdF^{n-k}(b;n,a') < \int_{b_a(l)}^{\bar{b}_F} bdF^{n-k}(b;n,a)$ for sufficiently large n. Thus, the expected payment of the high type is decreasing in a. □

Continuous signals

Here we consider a general set of marginal distributions, $g_v(s)$. Specifically assume that g_1 is increasing and g_0 is decreasing, and assume symmetry so that $g_1(s) = g_0(1-s)$.

Let us first write the utility of a player with signal s^1 for each bid b, when the other player is following an increasing bidding strategy $b(s')$. This is

$$U(s^1,b) \propto \min_{\lambda}(\int_0^z (1-b(s'))f_1(s^1,s')ds' - \int_0^z b(s')f_0(s^1,s')ds')$$

where $z = b^{-1}(b)$. Thus for each bid b, each player minimises utility by choosing a vector λ, given the strategy of the other player. Recall that s_v, for $v \in \{0,1\}$, is the median of the cdf $G_v()$.

Lemma 2.C.4. *Consider an equilibrium in which* $b(s)$ *is increasing. Let* $\boldsymbol{\lambda}^*(s)$ *denote the information structure which minimises the utility of the player for each* s*. Then:*

(i) $(\lambda_0^*(s), \lambda_1^*(s)) = (\lambda_{\max}, \lambda_{\min})$ *for all* $s < s_0$.

(ii) $(\lambda_0^*(s), \lambda_1^*(s)) = (\lambda_{\min}, \lambda_{\min})$ *for all* $s \in [s_0, \min\{\hat{s}, s_1\}]$.

(iii) $(\lambda_0^*(s), \lambda_1^*(s)) = (\lambda_{\min}, \lambda_{\max})$ *for* $s \in [s_1, \hat{s}]$ *if* $s_1 < \hat{s}$ *and* $(\lambda_0^*(s), \lambda_1^*(s)) = (\lambda_{\max}, \lambda_{\min})$ *for* $s \in [\hat{s}, s_1]$ *otherwise.*

(iv) $(\lambda_0^*(s), \lambda_1^*(s)) = (\lambda_{\max}, \lambda_{\max})$ *for all* $s > \max\{s_1, \hat{s}\}$,

and $\hat{s} < 1$ *satisfies*

$$\int_0^{\hat{s}} b(s')g_0(s')(2G_0(s') - 1))ds' = 0.$$

That is, $\boldsymbol{\lambda}^*(s)$ changes with s, so the behaviour as described cannot be rationalised with a unique a priori $\boldsymbol{\lambda}$.

Proof of Lemma 2.C.4 and Proposition 2.6.2. We first show in Claims 1-3 that $\boldsymbol{\lambda}^*(s)$ minimises the equilibrium utility of each player for each s. We then show that the bidding function $b(s) = E^{\boldsymbol{\lambda}^*(s)}(v|s, s)$ is an equilibrium. Finally, we identify a sufficient condition under which the seller's revenue is decreasing in ambiguity.

Define:

$$\begin{aligned} I_1(s) &= \int_0^s (1 - b(s'))g_1(s')g_1(s)(2G_1(s) - 1)(2G_1(s') - 1)ds' \\ I_0(s) &= -\int_0^s b(s')g_0(s')g_0(s)(2G_0(s) - 1)(2G_0(s') - 1))ds', \end{aligned}$$

Thus:

Claim 1: In equilibrium, $\lambda_v^*(s) = \lambda_{\min}$ $(\lambda_{\max})$ iff $I_v(s) > (<) \, 0$.

Proof. $I_v(s)$ is the derivative of the expected utility with respect to λ_v at $b = b(s)$. Given max-min behaviour, the statement follows. □

Claim 2: (i) $I_1(s) > 0$ for $s < s_1$, $I_1(s) < 0$ for all $s > s_1$; (ii) $I_0(s) < 0$ for $s < s_0$, $I_0(s) > 0$ for all $s \in (s_0, \hat{s})$, $I_0(s) < 0$ for all $s > \hat{s}$.

Proof. (i) $I_1(s)$: This function must be strictly positive for $s < s_1$ as $(2G_1(s) - 1)(2G_1(s') - 1) > 0$ for $s, s' < s_1$. Note that $I_1(s_1) = 0$, and that

$$\begin{aligned} \frac{\partial I_1(s)}{\partial s}|_{s=s_1} &= \frac{\partial g_1(s)(2G_1(s) - 1)}{\partial s}|_{s=s_1} \int_0^{s_1} (1 - b(s'))g_1(s')(2G_1(s') - 1)ds' \\ &= 2(g_1(s_1))^2 \int_0^{s_1} (1 - b(s'))g_1(s')(2G_1(s') - 1)ds' < 0 \end{aligned}$$

More generally:

$$\begin{aligned} \frac{\partial I_1(s)}{\partial s} &= (g_1'(s)(2G_1(s) - 1) + 2(g_1(s))^2) \int_0^s (1 - b(s'))g_1(s')(2G_1(s') - 1)ds' \\ &\quad +(1 - b(s))g_1(s)g_1(s)(2G_1(s) - 1)(2G_1(s) - 1) \\ &= (\frac{g_1'(s)}{g_1(s)} + \frac{2(g_1(s))}{(2G_1(s) - 1)})I_1(s) + 2(g_1(s))^2)(1 - b(s))(2G_1(s) - 1)(2G_1(s) - 1)) \end{aligned}$$

So whenever $I_1(s) > 0$ and $s > s_1$ we have that $\frac{\partial I_1(s)}{\partial s} > 0$ as $g_1(s)$ is increasing and $2(g_1(s))^2)(1 - b(s))(2G_1(s) - 1)(2G_1(s) - 1)) > 0$. So now it suffices to check that $I_1(1) < 0$:

$$\begin{aligned}
I_1(1) &= g_1(1)\int_0^1 (1 - b(s'))g_1(s')(2G_1(s') - 1)ds' \\
&= g_1(1)\int_0^{s_1} (1 - b(s'))g_1(s')(2G_1(s') - 1)ds' + g_1(1)\int_{s_1}^1 (1 - b(s))g_1(s')(2G_1(s') - 1)ds' \\
&< g_1(1)\int_0^{s_1} (1 - b(s_1))g_1(s')(2G_1(s') - 1)ds' + g_1(1)\int_{s_1}^1 (1 - b(s_1))g_1(s')(2G_1(s') - 1)ds' \\
&= g_1(1)\int_0^{s_1} (1 - b(s_1))g_1(s')(2G_1(s') - 1)ds' = 0,
\end{aligned}$$

where the last inequality follows as $b(s')$ is increasing, $(2G_1(s') - 1) > 0$ (< 0) whenever $s > s_1$ $(s < s_1)$. The last equality follows from $\int_0^1 g_1(s')(2G_1(s') - 1)ds' = 0$.

(ii) $I_0(s)$: This function must be strictly negative for $s < s_0$ as $(2G_0(s) - 1)(2G_0(s') - 1) > 0$ for $s, s' < s_0$. Note that $I_0(s_0) = 0$. Moreover,

$$\begin{aligned}
\frac{\partial I_0(s)}{\partial s}|_{s=s_0} &= -\frac{\partial g_0(s)(2G_0(s) - 1)}{\partial s}|_{s=s_0}\int_0^{s_0} b(s')g_0(s')(2G_0(s') - 1)ds' \\
&\quad -b(s')g_0(s')g_0(s_0)(2G_0(s_0) - 1)(2G_0(s') - 1)) \\
&= -2(g_0(s_0))^2\int_0^{s_0} b(s')g_0(s')(2G_0(s') - 1)ds' > 0
\end{aligned}$$

So $I_0(s) < 0$ for $s \geq s_0$. Note that $-\int_0^s b(s')g_0(s')(2G_0(s') - 1))ds'$ is decreasing for $s > s_0$. Thus if $I_0(1) < 0$, we have the result. But

$$\begin{aligned}
|I_0(1)| &= g_0(1)\int_0^1 b(s')g_0(s')(2G_0(s') - 1)ds' \\
&> g_0(1)\int_0^{s_0} b(s_0)g_0(s')(2G_0(s') - 1)ds' + g_0(1)\int_{s_0}^1 b(s_0)g_0(s')(2G_0(s') - 1)ds' \\
&= g_0(1)b(s_0)\int_0^1 g_0(s')(2G_0(s') - 1)ds' = 0.
\end{aligned}$$

Thus we know there exists $\hat{s} < 1$ such that:

$$\int_0^{\hat{s}} b(s')g_0(s')(2G_0(s') - 1))ds' = 0,$$

and we can conclude that $I_0(s) > 0$ for $s \in (s_0, \hat{s})$ and that $I_0(s) < 0$ for $s > \hat{s}$. □

Consider now the bidding function $E^{\lambda^*(s)}(v|s, s)$. Note that overbidding, compared to the canonical model, arises when

$$\frac{[1 + \lambda_1^*(s)(2G_1(s) - 1)^2]g_1^2(s)}{[1 + \lambda_1^*(s)(2G_1(s) - 1)^2]g_1^2(s) + [1 + \lambda_0^*(s)(2G_0(s) - 1)^2]g_0^2(s)} > \frac{g_1^2(s)}{g_1^2(s) + g_0^2(s)}$$

which holds if and only if:

$$\frac{[1 + \lambda_1^*(s)(2G_1(s') - 1)^2]}{[1 + \lambda_0^*(s)(2G_0(s') - 1)^2]} > 1.$$

We then have:

Claim 3: When $b(s) = E^{\lambda^*(s)}(v|s,s)$, a necessary condition for overbidding compared to the canonical model is $\hat{s} > 0.5$, that is:

$$\int_0^{0.5} b(s')g_0(s')(2G_0(s') - 1)ds' < 0$$

If this holds, there is overbidding in the region $[s_0, \hat{s}]$, and underbidding for any other s. Otherwise, all types underbid compared to the canonical model.

Proof. Given Claims 1 and 2, we can then deduce the different values of λ_v^* in equilibrium and consider when overbidding/underbidding arises compared to the canonical model when the bidding function is as described in the Proposition.

(a) $(\lambda_0^*(s), \lambda_1^*(s)) = (\lambda_{\max}, \lambda_{\min})$ for all $s < s_0$. As a result, if this is an equilibrium, we would have underbidding as
$$\frac{[1 + \lambda_{\min}(2G_1(s') - 1)^2]}{[1 + \lambda_{\max}(2G_0(s') - 1)^2]} < 1,$$
which is indeed the case as $\lambda_{\min} < 0 < \lambda_{\max}$.

(b) $(\lambda_0^*(s), \lambda_1^*(s)) = (\lambda_{\min}, \lambda_{\min})$ for all $s \in [s_0, \min\{\hat{s}, s_1\}]$. We have underbidding iff:
$$\frac{[1 + \lambda_{\min}(2G_1(s') - 1)^2]}{[1 + \lambda_{\min}(2G_0(s') - 1)^2]} < 1$$
If $\min\{\hat{s}, s_1\} > 0.5$, then we would have overbidding because in the region above 0.5, as $(2G_1(0.5) - 1)^2 = (2G_0(0.5) - 1)^2$ by symmetry, but because of convexity (concavity) of G_1 (G_0), the fraction would be greater than 1, as we would have $(2G_1(s') - 1)^2 < (2G_0(s') - 1)^2$ just above 0.5.

(c) $(\lambda_0^*(s), \lambda_1^*(s)) = (\lambda_{\max}, \lambda_{\max})$ for all $s > \max\{s_1, \hat{s}\}$. In this case we also have underbidding as $[1 + \lambda_{\max}(2G_1(s') - 1)^2] < [1 + \lambda_{\max}(2G_0(s') - 1)^2]$, because $\frac{1}{2} < G_1(s') < G_0(s')$.

(d) If $0.5 < s_1 < \hat{s}$: in the region $[s_1, \hat{s}]$ we have $(\lambda_0^*(s), \lambda_1^*(s)) = (\lambda_{\min}, \lambda_{\max})$. In this case we have overbidding as:
$$\frac{[1 + \lambda_{\max}(2G_1(s') - 1)^2]}{[1 + \lambda_{\min}(2G_0(s') - 1)^2]} > 1$$
For this we need $s_1 < \hat{s}$, implying that $0.5 < \hat{s}$.

(e) If $\hat{s} < s_1$: Then we have $(\lambda_0^*(s), \lambda_1^*(s)) = (\lambda_{\max}, \lambda_{\min})$ in this region between the two values. Then we have underbidding as:
$$\frac{[1 + \lambda_{\min}(2G_1(s') - 1)^2]}{[1 + \lambda_{\max}(2G_0(s') - 1)^2]} < 1. \qquad \square$$

Thus the structure of the equilibrium is therefore as above. So for overbidding we need:

$$\begin{aligned}
&\int_0^{0.5} b(s')g_0(s')(2G_0(s') - 1)ds' \\
= &\int_0^{s_0} \frac{[1 + \lambda_{\min}(2G_1(s') - 1)^2]g_1^2(s')}{[1 + \lambda_{\min}(2G_1(s') - 1)^2]g_1^2(s') + [1 + \lambda_{\max}(2G_0(s') - 1)^2]g_0^2(s')} g_0(s')(2G_0(s') - 1)ds' \\
&+ \int_{s_0}^{0.5} \frac{[1 + \lambda_{\min}(2G_1(s') - 1)^2]g_1^2(s')}{[1 + \lambda_{\min}(2G_1(s') - 1)^2]g_1^2(s') + [1 + \lambda_{\min}(2G_0(s') - 1)^2]g_0^2(s')} g_0(s')(2G_0(s') - 1)ds' \\
< \ & 0
\end{aligned}$$

which is analogous to what is in the Proposition. Finally we need to show that the construction above is an equilibrium:

Claim 4: The bidding function $b(s) = E^{\lambda^*(s)}(v|s,s)$ is a symmetric equilibrium when a is low enough.

Proof. We now show it is optimal for player 1 with signal s to bid $b(s)$, when player 2 is following the same strategy $b(s')$.

Let $\hat{\lambda}$ equal $\lambda^*(s)$ and consider the virtual utility for player 1 with signal s from bidding $b(z)$:

$$\begin{aligned}\hat{U}(s,z) &= \int_0^z (E^{\hat{\lambda}}(v|s,s') - b(s'))dF^{\hat{\lambda}}(s,s') \\ &= \frac{1}{2}(\int_0^z ((1-b(s'))f_1(\hat{\lambda},s,s') - b(s')f_0(\hat{\lambda},s,s'))ds'\end{aligned}$$

This is not player 1's utility as it is evaluated at $\hat{\lambda}$ for all z. However, player 1's utility from bidding $b(s)$ is indeed $\hat{U}(s,s)$, and player 1's utility from bidding $b(z)$, $z \neq s$ is at most $\hat{U}(s,z)$. Thus if $z = s$ maximises $\hat{U}(s,z)$, then bidding $b(s)$ is optimal for player 1 with signal s. Note that when $s' = s$, then the integrand is zero. To see that the integrand equals 0 note that, as $\hat{\lambda} = \lambda^*(s)$,

$$(1-b(s))f_1(\hat{\lambda},s,s) = b(s)f_0(\hat{\lambda},s,s)$$

iff

$$\begin{aligned}&[1+\lambda_0^*(s)(2G_0(s)-1)^2]g_0^2(s)[1+\lambda_1^*(s)(2G_1(s)-1)(2G_1(s)-1)]g_1(s)g_1(s) \\ =\ &[1+\lambda_1^*(s)(2G_1(s)-1)^2]g_1^2(s)[1+\lambda_0^*(s)(2G_0(s)-1)(2G_0(s)-1)]g_0(s)g_0(s)\end{aligned}$$

which holds.

The integrand when $s' = s$ is the derivative of $\hat{U}(s,z)$ with respect to z evaluated at $z = s$; thus, the first order condition with respect to z is satisfied for $z = s$. Moreover as we now show, the second order condition evaluated at this point is negative, thus $z = s$ is a maximum. To see this, suppose that we have a z for which $\hat{U}_z(s,z) = 0$. Taking a second derivative w.r.t. z we get: $-b'(z)(f_1(\hat{\lambda},s,z) + f_0(\hat{\lambda},s,z)) + (1-b(z))f_1'(\hat{\lambda},s,z) - b(z)f_0'(\hat{\lambda},s,z)$. As $\hat{U}_z(s,z) = 0$, this implies that $(1-b(z)) = \frac{b(z)f_0(\hat{\lambda},s,z)}{f_1(\hat{\lambda},s,z)}$, and thus the second order derivative at that z is

$$\begin{aligned}&-b'(z)(f_1(\hat{\lambda},s,z) + f_0(\hat{\lambda},s,z)) + \frac{b(z)f_0(\hat{\lambda},s,z)}{f_1(\hat{\lambda},s,z)}f_1'(\hat{\lambda},s,z) - b(z)f_0'(\hat{\lambda},s,z) \\ =\ &-b'(z)(f_1(\hat{\lambda},s,z) + f_0(\hat{\lambda},s,z)) + b(z)(\frac{f_0(\hat{\lambda},s,z)}{f_1(\hat{\lambda},s,z)}f_1'(\hat{\lambda},s,z) - f_0'(\hat{\lambda},s,z))\end{aligned}$$

Note that the first element is always negative. The second element is negative iff: $\frac{g_1'(z)g_1(s)(1+\hat{\lambda}_1(2G_1(z)-1)(2G_1(s)-1))+g_1(z)g_1(s)\hat{\lambda}_1 2g_1(z)(2G_1(s)-1))}{g_0'(z)g_0(s)(1+\hat{\lambda}_0(2G_0(z)-1)(2G_0(s)-1))+g_0(z)g_0(s)\hat{\lambda}_0 2g_0(z)(2G_0(s)-1))} < \frac{g_1(z)g_1(s)(1+\hat{\lambda}_1(2G_1(z)-1)(2G_1(s)-1))}{g_0(z)g_0(s)(1+\hat{\lambda}_0(2G_0(z)-1)(2G_0(s)-1))}$
Note that when $\hat{\lambda}$ is small enough, this is always the case as the LHS is negative. Thus a solution to the first order condition is unique.

But the above implies that player 1 can achieve this utility above and cannot improve upon it when using other bids $b(z)$, $z \neq s$.

So we know that the player bids up to the point where the integrand becomes negative, so, written differently, until $E^{\hat{\lambda}}(v|s,s) = b(s)$, which gives us the equilibrium bidding function. □

We now consider the seller's revenue and show that it is decreasing in a, under the sufficient condition identified. Consider the case when $\hat{s} > s_1$ (in the other cases, all types underbid compared to the canonical model, and hence revenue is also lower). Let

$$w(s') = (1 - G_1(s'))g_1(s') + (1 - G_0(s'))g_0(s')$$

The seller's revenue can be written as:

$$\begin{aligned} R(a) &= \int_0^{s_0} b(s', \lambda_{\max}, \lambda_{\min})w(s')ds' + \int_{s_0}^{s_1} b(s', \lambda_{\min}, \lambda_{\min})w(s')ds' + \\ &\quad \int_{s_1}^{\hat{s}} b(s', \lambda_{\min}, \lambda_{\max})w(s')ds' + \int_{\hat{s}}^{1} b(s', \lambda_{\max}, \lambda_{\max})w(s')ds' \end{aligned}$$

The derivative w.r.t. a is:

$$\begin{aligned} \frac{\partial R(a)}{\partial a} &= \int_0^{s_0} \frac{\partial}{\partial a} b(s', \lambda_{\max}, \lambda_{\min})w(s')ds' + \int_{s_0}^{s_1} \frac{\partial}{\partial a} b(s', \lambda_{\min}, \lambda_{\min})w(s')ds' + \\ &\quad \int_{s_1}^{\hat{s}} \frac{\partial}{\partial a} b(s', \lambda_{\min}, \lambda_{\max})w(s')ds' + \int_{\hat{s}}^{1} \frac{\partial}{\partial a} b(s', \lambda_{\max}, \lambda_{\max})w(s')ds' + \\ &\quad \frac{\partial \hat{s}}{\partial a}(b(\hat{s}, \lambda_{\min}, \lambda_{\max}) - b(\hat{s}, \lambda_{\max}, \lambda_{\max}))w(\hat{s}) \end{aligned}$$

We note that

$$\begin{aligned} \frac{\partial}{\partial a} b(s', \lambda_{\max}, \lambda_{\min})|_{a=1} &= \frac{g_1(s')^2 g_0(s')^2[-(2G_1(s') - 1)^2 - (2G_0(s') - 1)^2]}{(g_1(s')^2 + g_0(s')^2)^2} \\ \frac{\partial}{\partial a} b(s', \lambda_{\min}, \lambda_{\min})|_{a=1} &= \frac{g_1(s')^2 g_0(s')^2[-(2G_1(s') - 1)^2 + (2G_0(s') - 1)^2]}{(g_1(s')^2 + g_0(s')^2)^2} \\ \frac{\partial}{\partial a} b(s', \lambda_{\min}, \lambda_{\max})|_{a=1} &= \frac{g_1(s')^2 g_0(s')^2[(2G_1(s') - 1)^2 + (2G_0(s') - 1)^2]}{(g_1(s')^2 + g_0(s')^2)^2} \\ \frac{\partial}{\partial a} b(s', \lambda_{\max}, \lambda_{\max})|_{a=1} &= \frac{g_1(s')^2 g_0(s')^2[(2G_1(s') - 1)^2 - (2G_0(s') - 1)^2]}{(g_1(s')^2 + g_0(s')^2)^2} \end{aligned}$$

So that:

$$\begin{aligned} \frac{\partial}{\partial a} b(s', \lambda_{\max}, \lambda_{\min})|_{a=1} &= -\frac{\partial}{\partial a} b(s', \lambda_{\min}, \lambda_{\max})|_{a=1} = -\frac{\partial}{\partial a} b(1 - s', \lambda_{\min}, \lambda_{\max})|_{a=1} \\ \frac{\partial}{\partial a} b(s', \lambda_{\max}, \lambda_{\max})|_{a=1} &= -\frac{\partial}{\partial a} b(s', \lambda_{\min}, \lambda_{\min})|_{a=1} = \frac{\partial}{\partial a} b(1 - s', \lambda_{\min}, \lambda_{\min})|_{a=1} \end{aligned}$$

And therefore we can write $\frac{\partial R(a)}{\partial a}|_{a=1}$ as:

$$\begin{aligned} \frac{\partial R(a)}{\partial a}|_{a=1} &= -\int_0^{1-\hat{s}} \frac{\partial}{\partial a} b(s', \lambda_{\min}, \lambda_{\max})w(s')ds' - \int_{1-\hat{s}}^{s_0} \frac{\partial}{\partial a} b(s', \lambda_{\min}, \lambda_{\max})[w(s') - w(1 - s')]ds' \\ &\quad - \int_{s_0}^{0.5} \frac{\partial}{\partial a} b(s', \lambda_{\max}, \lambda_{\max})[w(s') - w(1 - s')]ds' - \int_{\hat{s}}^{1} \frac{\partial}{\partial a} b(s', \lambda_{\min}, \lambda_{\min})w(s')ds'. \end{aligned}$$

Thus, a sufficient condition for revenue to be decreasing in a is for $w(s')$ to be decreasing over $[\hat{s}, 1 - \hat{s}]$. Note that this sufficient condition is satisfied for $g_0(s) = 2(1 - s)$ and $g_1(s) = 2s$. □

Private values: Consider the following model where each buyer has private valuation $v_i \in [0, 1]$, independently distributed according to a uniform distribution. Suppose that the buyers believe that the joint distribution is given by the F-G-M copula $f(v_i, v_j; \lambda) = 1 + \lambda(2v_i - 1)(2v_j - 1)$, where $\lambda \in \Lambda = [\underline{\lambda}, \bar{\lambda}]$. Assume that $0 \in \Lambda$, and define:

$$G_\lambda(z|v_i) = \int_0^z (1 + \lambda(2v_i - 1)(2v_j - 1))dv_j = z + \lambda z(z - 1)(2v_i - 1)$$

That is, $G_\lambda(\cdot|v_i)$ is the distribution of v_j conditional on v_i, when the belief about the joint distribution is given by λ. Let $g_\lambda(\cdot|v_i)$ denote the corresponding density. Define:

$$L_\lambda(y|v_i) = \exp\left(-\int_y^{v_i} \frac{g_\lambda(t|t)}{G_\lambda(t|t)}dt\right)$$

Proposition 2.C.5. *When ambiguity is not too large, a symmetric equilibrium in the first price auction is:*

$$b(v_i) = \begin{cases} \int_0^{v_i} y dL_{\underline{\lambda}}(y|v_i) & v_i \leq 0.5 \\ \int_0^{v_i} y dL_{\bar{\lambda}}(y|v_i) + \int_0^{0.5} y d(L_{\underline{\lambda}}(y|0.5) - L_{\bar{\lambda}}(y|0.5)) & v_i > 0.5 \end{cases}$$

Bids in the first price auction are uniformly higher than the case without ambiguity.

Proof. Assume that player j is following the equilibrium strategy. Player i's utility from bidding $b(z) \in [0, v_i]$ is:

$$\min_{\lambda \in \Lambda} U_\lambda(z|v_i) \equiv G_\lambda(z|v_i)(v_i - b(z))$$

Note that $G_\lambda(z|v_i) = z + \lambda z(z - 1)(2v_i - 1)$. Thus, $\underline{\lambda}$ minimises this utility for types $v_i \leq 0.5$, and $\bar{\lambda}$ minimises this utility for types $v_i > 0.5$. Thus, the utility from bidding $b(z)$ is $U_{\underline{\lambda}}(z|v_i)$ for $v_i \leq 0.5$ and $U_{\bar{\lambda}}(z|v_i)$ for $v_i > 0.5$. First order conditions imply:

$$g_{\underline{\lambda}}(v_i|v_i)(v_i - b(v_i)) - G_{\underline{\lambda}}(v_i|v_i)b'(v_i) = 0 \qquad v_i \leq 0.5$$
$$g_{\bar{\lambda}}(v_i|v_i)(v_i - b(v_i)) - G_{\bar{\lambda}}(v_i|v_i)b'(v_i) = 0 \qquad v_i > 0.5$$

Using boundary conditions $b(0) = 0$ and $b(0.5) = \int_0^{0.5} y dL_{\underline{\lambda}}(y|0.5)$, the solution to these differential equations is:

$$b(v_i) = \begin{cases} \int_0^{v_i} y dL_{\underline{\lambda}}(y|v_i) & v_i \leq 0.5 \\ \int_0^{v_i} y dL_{\bar{\lambda}}(y|v_i) + \int_0^{0.5} y d(L_{\underline{\lambda}}(y|0.5) - L_{\bar{\lambda}}(y|0.5)) & v_i > 0.5 \end{cases}$$

Consider type $v_i \leq 0.5$ deviating to a bid $b(z)$.[24] Note that for $0.5 > z > v_i$:

$$\begin{aligned} \frac{\partial U_{\underline{\lambda}}(z|v_i)}{\partial z} &= G_{\underline{\lambda}}(z|v_i)\left[\frac{g_{\underline{\lambda}}(z|v_i)}{G_{\underline{\lambda}}(z|v_i)}(v_i - b(z)) - b'(z)\right] \\ &< G_{\underline{\lambda}}(z|v_i)\left[\frac{g_{\underline{\lambda}}(z|z)}{G_{\underline{\lambda}}(z|z)}(z - b(z)) - b'(z)\right] = 0 \end{aligned}$$

The inequality follows because $z > v_i$, and when $\underline{\lambda}$ is close to zero, $\frac{g_{\underline{\lambda}}(z|v_i)}{G_{\underline{\lambda}}(z|v_i)}$ and $\frac{g_{\underline{\lambda}}(z|z)}{G_{\underline{\lambda}}(z|z)}$ are close.[25] A similar argument shows that for $z < v_i$, $\frac{\partial U_{\underline{\lambda}}(z|v_i)}{\partial z} > 0$.

When $z > 0.5$, note that if $\bar{\lambda}$ and $\underline{\lambda}$ are close, then $\frac{\partial U_{\bar{\lambda}}(z|v_i)}{\partial z}$ and $\frac{\partial U_{\underline{\lambda}}(z|v_i)}{\partial z}$ are close. Thus for some $\varepsilon > 0$,

[24] The argument for $v > 0.5$ is analogous.

[25] The derivative of $\frac{g_\lambda(z|v)}{G_\lambda(z|v)}$ with respect to v is $\frac{2\lambda z^2}{G_\lambda(z|v)^2}$.

$$\begin{aligned}\frac{\partial U_{\underline{\lambda}}(z|v_i)}{\partial z} &< \frac{\partial U_{\bar{\lambda}}(z|v_i)}{\partial z} + \varepsilon \\ &= G_{\bar{\lambda}}(z|v_i)\left[\frac{g_{\bar{\lambda}}(z|v_i)}{G_{\bar{\lambda}}(z|v_i)}(v_i - b(z)) - b'(z)\right] + \varepsilon \\ &< G_{\bar{\lambda}}(z|v_i)\left[\frac{g_{\bar{\lambda}}(z|z)}{G_{\bar{\lambda}}(z|z)}(z - b(z)) - b'(z)\right] = 0\end{aligned}$$

where the last inequality follows because $z > v_i$, and since $\bar{\lambda} > 0$, $\frac{g_{\bar{\lambda}}(z|z)}{G_{\bar{\lambda}}(z|z)} > \frac{g_{\bar{\lambda}}(z|v_i)}{G_{\bar{\lambda}}(z|v_i)}$ (see footnote 25).

To see that bids are uniformly higher under ambiguity, note that in the case without ambiguity, the bid can be written as $b_0(v_i) = \int_0^{v_i} y dL_0(y|v_i)$. When $v_i \leq 0.5$, $\int_y^{v_i} \frac{g_\lambda(t|t)}{G_\lambda(t|t)} dt$ is decreasing in λ since the integrand $\frac{g_\lambda(t|t)}{G_\lambda(t|t)}$ is decreasing in λ.[26] Thus, $L_{\underline{\lambda}}$ first order stochastically dominates L_0, implying that for $v_i \leq 0.5$, $\int_0^{v_i} y dL_{\underline{\lambda}}(y|v_i) \geq \int_0^{v_i} y dL_0(y|v_i)$. A similar argument implies that for $v_i > 0.5$, $\int_{0.5}^{v_i} y dL_{\bar{\lambda}}(y|v_i) \geq \int_{0.5}^{v_i} y dL_0(y|v_i)$.[27] Note that $\int_0^{0.5} y dL_{\bar{\lambda}}(y|v_i) \geq \int_0^{0.5} y dL_{\bar{\lambda}}(y|0.5)$ and $\int_0^{0.5} y dL_{\underline{\lambda}}(y|0.5) \geq \int_0^{0.5} y dL_0(y|0.5)$, implying the result. □

[26]The derivative of $\frac{g_\lambda(t|t)}{G_\lambda(t|t)}$ with respect to λ is $\frac{t^2(2t-1)}{G_\lambda(t|t)^2}$.

[27]Integrating by parts and using the fact that for $v_i > y > 0.5$, $L_\lambda(y|v_i)$ is decreasing in λ:

$$\begin{aligned}&\int_{0.5}^{v_i} y dL_{\bar{\lambda}}(y|v_i) - \int_{0.5}^{v_i} y dL_0(y|v_i) \\ &= 0.5(L_0(0.5|v_i) - L_{\bar{\lambda}}(0.5|v_i)) + \int_{0.5}^{v_i} L_0(y|v_i) - L_{\bar{\lambda}}(y|v_i) dy \geq 0.\end{aligned}$$

Chapter 3

A Nash Threats Folk Theorem for Repeated Games with Local Monitoring

3.1 Introduction

In this paper we will prove a Nash threats folk theorem for infinitely repeated games with local monitoring and interaction. We assume that the monitoring and the interaction structure are both determined by an undirected network. This means that for each player, her stage game payoffs only depend on the actions of a subset of players, and these actions are the only actions that she observes. We also assume that players are patient, and repeated game payoffs are given by the Banach-Mazur limit of the sequence of average stage game payoffs. This latter assumption is crucial for our results. Indeed, we show that under discounting, a folk theorem cannot hold in our setting without further assumptions on the network structure or the payoffs. In particular, we show without further assumptions on payoffs, a necessary condition for a folk theorem to hold under discounting is that every connected component of the network is 2-connected.

Local monitoring prevents standard results from applying. For example, in a star network, the core player may not wish to punish a single defecting peripheral player if that involves ending cooperation in all relationships forever. Thus, grim trigger strategies may not be an equilibrium. One way to view this problem is as follows. With local monitoring, it can be difficult for players to distinguish between punishment and defection. Suppose that Player j defects, and Player j's only neighbour is Player i. Since Player i's neighbours do not observe whether Player j has defected, if Player i punishes Player j, then Player i's neighbours will punish Player i. Now if Player i had strict incentives to play on-path, she may prefer not to punish Player j to avoid getting punished herself. A natural approach would be to construct strategies such that cooperating is only slightly preferred to defecting. Indeed, in the equilibria we construct, players are indifferent between cooperating and any finite sequence of defections.

The proof will be constructive: for any payoff above the Nash equilibrium point, we will construct a strategy profile that achieves the target payoff. The strategies will be stable in the sense that after any arbitrary history, everyone goes back to playing on-path after a finite number of periods. Since players are arbitrarily patient, short term incentives are irrelevant, which means that they can always achieve their equilibrium payoff by following the strategy. Thus, the challenge is to show that no deviation can achieve a payoff greater than the equilibrium payoff (note that the one-shot deviation principle does not apply).

Stability is not a trivial requirement under local monitoring because different players may have different beliefs about when a punishment phase is supposed to end. We circumvent this problem by constructing strategies that exploit the common knowledge of time—that is, rather than requiring players to punish deviations for a fixed number of periods, we require punishment to last until certain dates. Roughly speaking, we will divide the repeated game into T period blocks, and each T period block into 2 parts. Deviations in the first part of a block are punished until the end of the block, and deviations in the second part of the block are punished until the end of the next block. As long as the second part is long enough, deviations will be punished sufficiently harshly, and as long as the first part is long enough, the strategies will be stable.

Here is the intuition behind stability. Suppose that we are trying to sustain mutual cooperation in the prisoners' dilemma, so that C is played on-path and D is played if a deviation is observed (by both the deviating player and her neighbours). At the end of the block following an arbitrary history, it may be the case that some players are supposed to play C, whereas others have to play D for one more block. But those who are supposed to play C and observe D from a neighbour will start playing D themselves. In this way, the punishment will spread throughout the network, and as long as the first part of the block is long enough, everyone will have seen D by the end of the first part, and respond by playing D until the end of the block. Thus, everyone will end punishment at the same time.

A different type of coordination problem is that even though players are punished effectively when all of their neighbours play the Nash equilibrium action, they may benefit from being 'punished' by only some of their neighbours. In order to prevent players from exploiting this, our strategies ensure that the number of periods in each length T block in which a player is punished by only a proper subset of her neighbours is small relative to the number of periods in which she is punished by every neighbour. This is achieved by requiring punishment to continue for an extra block in some cases, and to end prematurely in others.

One final difficulty that could arise under local monitoring is that after observing a deviation, players may try to infer the spread of the deviation in the network and the beliefs of others about future play. We circumvent this difficult by constructing strategies that are optimal for every belief that players might have about play outside their neighbourhood.

3.2 Related Literature

This paper continues the tradition of providing limiting results for repeated games on networks. Nava (2016) provides a survey of of this literature, which has largely found that very weak conditions on the network structure are required for a folk theorem to hold. In this paper, we impose no restrictions on the network structure. An earlier and related strand of literature, which focused on a random matching environment, was pioneered by Kandori (1992) and Ellison (1994). They establish that even when players are unable to recognise their opponents, cooperation can be sustained as a sequential equilibrium supported by contagious punishments. In our setting, interaction takes place on a stable network, and players know their neighbours and observe their actions.

Ben-Porath and Kahneman (1996) showed that under the assumption that public communication is possible, a necessary and sufficient condition for a folk theorem to hold is that every player is observed by at least two other players. Renault and Tomala (1998) and Tomala (2011) establish Nash folk theorems under similar conditions without any explicit communication, and Laclau (2012) and Laclau (2014) establish Nash and sequentially rational folk theorems, respectively, under different assumptions about the communication possibilities. In this paper, we consider a restricted class of games, which allows us to establish a sequentially rational folk theorem for any network structure.

Cho (2011) and Cho (2014) construct stable and sequentially rational equilibria that sustain cooperation in the repeated prisoners' dilemma. Cho (2011) assumes that players have access to a public randomisation device, and Cho (2014) allows players to communicate with their neighbours. In this paper, players do not have access to a public randomisation device or any method of communication.

This paper builds on Nava and Piccione (2014), who show that mutual cooperation is possible for a broad class of two-action games. They also consider discounted payoffs as well as the case of arbitrarily patient players. However, we extend their results to a folk theorem, and we construct different strategies to support the equilibria.

The model is described in Section 3.3. Our main result, that any feasible payoff above the Nash equilibrium point can be (approximately) sustained in sequential equilibrium, is stated as Proposition 3.4.1 in Section 3.4. The strategies used to prove Proposition 3.4.1 are described in Section 3.5, and the proof of Proposition 3.4.1 appears in Section 3.6. Section 3.7 concludes.

3.3 Model

A set N contains n players who interact according to an undirected graph (N, G). We assume that (N, G) is common knowledge. For each $i \in N$, define $N_i = \{j \in N \setminus \{i\} : ij \in G\}$. Note that $j \in N_i$ if and only if $i \in N_j$. We interpret N_i as the neighbourhood of Player i. Define a path as $(j_1, \dots, j_m)$ such that $j_{k+1} \in N_{j_k}$, $k = 1, \dots, m-1$, and $j_k \neq j_l$ for $k \neq l$.

For $M \subset N$, define $G_M = \{ij \in G : i, j \in M\}$. That is, (M, G_M) is the subgraph with vertices in M. (M, G_M) is connected if for each $j_1, j_m \in M$, there is a path $(j_1, \dots, j_m)$ such that $j_{k+1} \in N_{j_k} \cap M$. (M, G_M) is 2-connected if for each $i \in M$, $(M \setminus \{i\}, G_{M\setminus\{i\}})$ is connected. For $M \subset N$, (M, G_M) is a connected component of (N, G) if (M, G_M) is connected and for all $i \in N \setminus \{M\}$, $(M \cup \{i\}, G_{M\cup\{i\}})$ is not connected.

The action set of Player i is a finite set A_i. Let $A_M = \times_{j\in M} A_j$. The stage game payoff for Player i is given by the function $v_i : A_{N_i \cup \{i\}} \mapsto R$, and we denote an element of $A_{N_i\cup\{i\}}$ by a_i, a_{N_i}.

Assumption 1. *For each A_i, there exists an action $D \in A_i$ such that $v_i(D, D, \dots, D) > v_i(a_i', D, \dots, D)$ for all $a_i' \in A_i \setminus \{D\}$.*

This assumption says that the strategy profile $(D, \dots, D)$ is a strict Nash equilibrium of the stage game. Without loss of generality, let $v_i(D, \dots, D) = 0$ for all $i \in N$.

The stage game is repeated infinitely many times. Each player observes the past play of her neighbours. The set of possible histories for Player i is:

$$H_i = \{\emptyset\} \cup \left\{ \cup_{t=1}^{\infty} \left[\times_{s=1}^{t} A_{N_i \cup \{i\}} \right] \right\}.$$

A (pure) strategy for Player i is a function $\sigma_i : H_i \mapsto A_i$. The set of all strategies available to Player i is given by Σ_i.

Given a strategy profile $\sigma_N = (\sigma_1, \dots, \sigma_n)$, let $\{a_N^t\}_{t=1}^{\infty}$ be the sequence of actions generated by σ_N, and let $\{v_i(a_i^t, a_{N_i}^t)\}_{t=1}^{\infty}$ be the sequence of stage game payoffs for Player i. Players discount the future without common discount factor $\delta \leq 1$, and repeated game payoffs are defined as:

$$\mathcal{V}_i(\sigma_N) = \begin{cases} \Lambda\left(\left\{ \frac{1}{t} \sum_{s=1}^{t} v_i(a_i^s, a_{N_i}^s) \right\}_{t=1}^{\infty} \right) & \text{when } \delta = 1 \\ (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} v_i(a_i^t, a_{N_i}^t) & \text{when } \delta < 1, \end{cases}$$

where $\Lambda(\cdot)$ denotes the Banach-Mazur limit of a sequence. A Banach-Mazur limit is a positive linear functional $\Lambda : \ell_\infty \to \mathbb{R}$, where ℓ_∞ is the space of all bounded sequences, such that $\Lambda(\mathbf{e}) = 1$, where $\mathbf{e} = (1, 1, \ldots)$, and $\Lambda(x_1, x_2, \ldots) = \Lambda(x_2, x_3, \ldots)$ for each $(x_1, x_2, \ldots) \in \ell_\infty$ (see Aliprantis and Border (2006), pp. 550-551).

The set of histories for the entire game is:

$$H = \{\varnothing\} \cup \left\{\cup_{t=1}^{\infty}\left[\times_{s=1}^{t} A_N\right]\right\}.$$

For each observed history $\bar{h}_i$, we can define the information set of Player i as $\mathcal{I}(\bar{h}_i) = \{h \in H : h_i = \bar{h}_i\}$. Let $\beta(h|h_i)$ denote the belief of Player i that the history is h, conditional on observing h_i.

3.4 Folk Theorem

Let $v(a_N) = (v_1(a_1, a_{N_1}), \ldots, v_n(a_n, a_{N_n}))$. The set of feasible payoffs is $F = \text{co}\{v(a_N) : a_N \in A_N\}$, and the set of feasible payoffs such that each player receives strictly more than her stage game Nash equilibrium payoff is $F^{IR} = F \cap \{v : v_i > 0, \forall i\}$. Let F^* be the subset of F^{IR} where the weights used in the convex combinations are rational numbers.

Proposition 3.4.1. *If $\delta = 1$, any $v^* \in F^*$ can be supported as a sequential equilibrium payoff.*

In the next section, we will construct a strategy profile that supports v^* as a sequential equilibrium payoff for any v^*. For the remainder of this section, we discuss the necessity of the assumption that $\delta = 1$.

Remark 3.4.1. The folk theorem does not hold in this setting when payoffs are discounted. In order for a folk theorem to hold under discounting, additional restrictions must be made either on payoffs or the network structure. The following example shows that when $\delta < 1$, there exists a stage game satisfying Assumption 1, a network structure, and a payoff $v^* \in F^*$ such that v^* is not the payoff in any sequential equilibrium.

Example 3.4.1. Let four players connected on a line play the following stage game with common action space $A = \{C, D\}$, and payoff function:

- $v(D, D, D, D) = (0, 0, 0, 0)$
- $v(C, C, C, C) = (1, 1, 1, 1)$
- $v(C, D, C, C) = (-1, 2, 1, 1)$
- $v(D, D, C, C) = (0, 2, 1, 1)$
- $v(D, C, C, C) = (1, 1, 1, 1)$

Otherwise the payoffs are such that Player 3 gets at most zero, Assumption 1 is satisfied (i.e. (D, D, D, D) is a strict Nash equilibrium of the stage game), and for each i, Player i's payoffs depends only on the actions of Player i and Player i's neighbours. A complete specification of the stage game payoffs are given in the Appendix.

The efficient payoff $(1, 1, 1, 1)$ (which requires Player 2, Player 3, and Player 4 to play C) cannot be sustained in any sequential equilibrium under discounting. To see this, note that in any equilibrium in which Player 3 and Player 4 play C forever on the equilibrium path, Player 3 can guarantee the payoff 1 by always playing C. Moreover, 1 is the highest payoff Player 3 can get in the stage game. In order to prevent Player 2 from deviating from C, Player 2 must be punished for playing D. However, the only way Player 2 can be punished for playing D is if Player 3 plays D (since if Player 3 plays C, Player 2 gets at least 1). But in this case, Player 3 will get strictly less than 1, so Player 3 will not find it optimal to trigger this punishment. Thus, Player 2 will have an incentive to deviate from C, and there cannot exist an equilibrium in which Player 2, Player 3, and Player 4 play C forever.

Note that under perfect monitoring, the payoff $(1, 1, 1, 1)$ can be sustained in sequential equilibrium for high enough δ, for example by using grim trigger strategies. In this case, the problem under network monitoring is that Player 4 does not observe Player 2, and fixing the action of Player 4 to be C, Player 3 can guarantee her maximum stage game payoff (which is 1). In other words, the feasible and individually rational payoff set of Player 2's neighbourhood has empty interior when Player 4's action is fixed at C.

This example generalises to the following Proposition, which establishes that a necessary condition on the network structure for the folk theorem to hold under discounting is that every connected component is 2-connected.

Proposition 3.4.2. *For any (N, G) such that there exists a connected component that is not 2-connected, there exists a stage game satisfying Assumption 1 and a payoff $v^* \in F^*$ such that v^* is not the payoff in any sequential equilibrium for any $\delta < 1$.*

Proof. Without loss of generality, assume that the entire network (N, G) is connected (if not, we can just consider each connected component separately). Now suppose that the network is not 2-connected. Let i^* be a player such that $(N \setminus \{i^*\}, G_{N\setminus\{i^*\}})$ is not connected. First, we will define $N^L_{i^*}$ and $N^R_{i^*}$ as two nonempty, disjoint subsets of N_{i^*} with the property that for $i \in N^L_{i^*}$ and $j \in N^R_{i^*}$, every path $(i, \dots, j)$ contains i^*. Let (N^L, G_{N^L}) and (N^R, G_{N^L}) be two distinct connected components of $(N \setminus \{i^*\}, G_{N\setminus\{i^*\}})$, and let $N^L_{i^*} = N^L \cap N_{i^*}$ and $N^R_{i^*} = N^R \cap N_{i^*}$. We argue that $N^L_{i^*}$ and $N^L_{i^*}$ are non empty. By the assumption that (N, G) is connected, for $i \in N^L \subset N$ and $j \in N^R \subset N$, there must be a path $(i, \dots, j)$, where each element of the path is in N. Since (N^L, G_{N^L}) and (N^R, G_{N^L}) are two disjoint connected components of $(N \setminus \{i^*\}, G_{N\setminus\{i^*\}})$, there is no path from i to j that does not contain i^*. Therefore, there must exist a path $(i, \dots, i^L, i^*, i^R, \dots, j)$. Now we show that $i^L \in N^L \cap N_{i^*}$. The argument that $i^R \in N^R \cap N_{i^*}$ is analogous. By definition $i^L \in N_{i^*}$. If $i^L = i \in N^L$, then we are done. Otherwise, there is a path $(i, i_2, \dots, i_{m-1}, i^L)$, where each element is in $N \setminus \{i^*\}$. Since (N^L, G_{N^L}) is a connected component of $(N \setminus \{i^*\}, G_{N\setminus\{i^*\}})$, (N^L, G_{N^L}) is connected. Since $i \in N^L$ and there is a path (i, i_2), $(N^L \cup \{i_2\}, G_{N^L\cup\{i_2\}})$ is also connected. But the definition of a connected component then implies that $i_2 \notin N \setminus (\{i^*\} \cup N^L)$. Since $i_2 \neq i^*$, this implies that $i_2 \in N^L$. Repeating this argument yields $i^L \in N^L$.

Now we define a stage game satisfying Assumption 1 such that there exists a payoff $v^* \in F^*$ that is not the payoff in any sequential equilibrium for any $\delta < 1$. For each player, let $A_i = \{C, D\}$. Without loss of generality, let $v_i(D, \dots, D) = 0$ and $v_i(C, D, \dots, D) = -1$ for all i.

For all $i \notin \{i^*\} \cup N^L_{i^*} \cup N^R_{i^*}$:

- $v_i(D, a_{N_i}) = 0$ for $a_{N_i} = (D, \dots, D)$
- $v_i(C, a_{N_i}) = -1$ for $a_{N_i} = (D, \dots, D)$
- $v_i(C, a_{N_i}) = 1$ for all $a_{N_i} \neq (D, \dots, D)$
- $v_i(D, a_{N_i}) = 1$ for all $a_{N_i} \neq (D, \dots, D)$.

For $i \in N_{i^*}^L \cup N_{i^*}^R$:

- $v_i(D, a_{N_i}) = 0$ for $a_{N_i} = (D, \ldots, D)$
- $v_i(C, a_{N_i}) = -1$ for $a_{N_i} = (D, \ldots, D)$
- $v_i(C, a_{N_i}) = 1$ whenever $a_{i^*} = C$
- $v_i(C, a_{N_i}) = -1$ whenever $a_{i^*} = D$
- $v_i(D, a_{N_i}) = 2$ whenever $a_{i^*} = C$
- $v_i(D, a_{N_i}) = 1$ otherwise.

For i^*, let

- $v_i(D, a_{N_i}) = 0$ for $a_{N_i} = (D, \ldots, D)$
- $v_i(C, a_{N_i}) = -1$ for $a_{N_i} = (D, \ldots, D)$
- $v_{i^*}(D, a_{N_{i^*}}) = 0$ for all $a_{N_{i^*}}$
- $v_{i^*}(C, a_{N_{i^*}}) = 1$ whenever $a_{N_{i^*}^R} = (C, \ldots, C)$
- $v_{i^*}(C, a_{N_{i^*}}) = -1$ otherwise.

Note that for each player i, $0 < 1 \leq \max_{(a_i, a_{N_i})} v_i(a_i, a_{N_i})$. Suppose for a contradiction that there exists a sequential equilibrium σ_N such that $\mathcal{V}_i(\sigma_N) = 1$ for all i. This requires that for every on-path history h_i, $\sigma_i(h_i) = C$ for all $i \in \{i^*\} \cup N_{i^*}^R$, since $v_{i^*}(a_{i^*}, a_{N_{i^*}}) < 1$ for all $(a_{i^*}, a_{N_{i^*}})$ such that $(a_{i^*}, a_{N_{i^*}^R}) \neq (C, \ldots, C)$. As long as no player $i \in \{i^*\} \cup N_{i^*}^R$ has deviated, Player i^* can guarantee the stage game payoff 1 in every period by always playing C. Now, following a history in which no player has deviated, consider the one shot deviation for some player $i \in N_{i^*}^L$ to D. For $i \in N_{i^*}^L$, the stage game payoff to D when i^* is playing C is 2, and $i \in N_{i^*}^L$ can only receive a stage game payoff less than 1 if player i^* plays D.

Without loss of generality, let $1 \in N_{i^*}^L$. Let h be the one period history where $a_i = \sigma_i(\varnothing)$ for each $i \neq 1$, and $a_1 = D$. Let σ_N^h be the strategy profile induced by the σ_N after history h. Let σ_1' be the strategy profile where Player 1 plays D in period 1, and follows σ_1 from period 2 onwards, and let $\{\bar{a}_N^t\}_{t=1}^{\infty}$ be the sequence of action profiles induced by $(\sigma_1', \sigma_{N\setminus\{1\}})$. Note that in any sequential equilibrium, $\beta(h|h_1) = 1$. Then Player 1's payoff is:

$$V_1(\sigma_1', \sigma_{N\setminus\{1\}}) = (1-\delta)2 + \delta \mathcal{V}_1(\sigma_N^h)$$

For this one shot deviation not to be profitable, we need $\mathcal{V}_1(\sigma_N^h) < 1$, which involves Player i^* playing D at least once following the deviation. To see this, suppose that $a_{i^*}^t \neq D$ for all $t \geq 2$. Then $v_i(\bar{a}_1^t, \bar{a}_{N_1}^t) \geq 1$ for all $t \geq 2$ and:

$$\begin{aligned}\mathcal{V}_1(\sigma_N^h) &= (1-\delta)\sum_{t=2}^{\infty} \delta^{t-2} v_i(\bar{a}_1^t, \bar{a}_{N_1}^t) \\ &\geq (1-\delta)\sum_{t=2}^{\infty} \delta^{t-2} \geq 1\end{aligned}$$

But consider the strategy σ'_{i^*} where Player i^* plays C in every period. In this case, all $i \in N^R_{i^*}$ will also play C and:

$$\mathcal{V}_{i^*}(\sigma'^h_{i^*}, \sigma^h_{N\backslash\{i^*\}}) = 1$$

On the other hand, if $a^t_{i^*} = D$ for some $t \geq 2$, then $v_i(\bar{a}^t_{i^*}, \bar{a}^t_{N_{i^*}}) \leq 1$ for all t with a strict inequality for the t such that $a^t_{i^*} = D$. Then:

$$\begin{aligned}\mathcal{V}_{i^*}(\sigma^h_N) &= (1-\delta)\sum_{t=2}^{\infty}\delta^{t-2}v_i(\bar{a}^t_{i^*}, \bar{a}^t_{N_{i^*}}) \\ &< (1-\delta)\sum_{t=2}^{\infty}\delta^{t-2} = 1\end{aligned}$$

Therefore, it must be the case that $\bar{a}^t_{i^*} \neq D$ for all $t \geq 2$, and so the one shot deviation must be profitable for Player 1. This contradicts the assumption that σ_N is a sequential equilibrium. □

3.5 Strategies

For any payoff $v^* \in F^*$, construct a finite deterministic sequence of stage game actions $\{a^s_N\}^{T_*}_{s=1}$ such that $\frac{1}{T_*}\sum_{s=1}^{T_*} v(a^s_N) = v^*$. There must be a finite sequence that achieves this payoff because of the assumption that the weights used in the convex combinations are rational numbers. Now for $T > nT_*$, let $\{a^s_N\}^T_{s=1}$ be T/T_* repetitions of the length T_* sequence (we will require $T - nT_*$ to be sufficiently large). Note that this implies that T is a multiple of T_*, i.e. $T \mod T_* = 0$.

Define $s(t) \equiv ((t-1) \mod T) + 1$ (i.e. $s(t) = t \mod T$ except when $t \mod T = 0$, in which case $s(t) = T$). In any period t, let $C_{is(t)}$ denote the i-th element of the vector a^s_N that is s-th term of the sequence $\{a^s_N\}^T_{s=1}$. We will refer to $C_{is(t)}$ as the on-path action for player i in period t. By construction $C_{is(t)} = C_{i(s(t)+T_*)}$ (as long as $s(t) + T_* \leq T$, otherwise $C_{i(s(t)+T_*)}$ is not defined), which ensures that on-path actions are repeated at least every T_* periods. For $M \subset N$, let $C_{Ms(t)} = (C_{js(t)})_{j\in M}$.

The strategy profile $\zeta_i : H_i \mapsto A_i$

The strategy profile $\zeta_i : H_i \mapsto A_i$ can be described with the help of:

- A set of states $\mathcal{S} = \{\mathcal{A}, \mathcal{B}(1), \ldots, \mathcal{B}(2T), \mathcal{A}^1, \ldots, \mathcal{A}^n\}$
- For each state, an output function $f_i : \mathcal{S} \times \mathbb{N} \mapsto A_i$, where:

$$\begin{aligned}f_i(\mathcal{A}, t) &= C_{is(t)} \\ f_i(\mathcal{B}(p), t) &= D \\ f_i(\mathcal{A}^j, t) &= C_{is(t)}\end{aligned}$$

- For each history h_i, a state function $P_i : H_i \mapsto \mathcal{S}$

In the next subsection, we will define $P_i(h_i)$ recursively. Then for each h_i of length $t-1$:

$$\zeta_i(h_i) = f_i(P_i(h_i), t)$$

That is, if Player i is in state $\mathcal{A}$ or $\mathcal{A}^j$ at time t, then Player i plays $C_{is(t)}$. If Player i is in state $\mathcal{B}(p)$, then Player i plays D. Note that our state function depends on the history, unlike a conventional transition function that depends on the state and the action profile. Thus, our 'states' are not true states. Moreover, our output function depends on time. Of course it is possible to describe the strategy profile using a standard automaton representation, but we would need many more states.

The function P_i

After the initial history:

$$P_i(\varnothing) = \mathcal{A}$$

Let h_i^t denote a length t history for Player i. Suppose that $P_i(h_i^{t-1}) = \mathcal{A}$.

- If $a_j^t = C_{js(t)}$ for all $j \in N_i \cup \{i\}$, then $P_i(h_i^t) = \mathcal{A}$
- Otherwise $P_i(h_i^t) = \mathcal{B}(p(t))$, where:

$$p(t) \equiv \begin{cases} T - s(t) & \text{for } s(t) \in \{1, \ldots, nT_*\} \\ 2T - s(t) & \text{for } s(t) \in \{nT_* + 1, \ldots, T\} \end{cases}$$

Suppose that $P_i(h_i^{t-1}) = \mathcal{B}(p)$. $P_i(h_i^t)$ depends on two functions $d_{1i} : H_i \mapsto \{0, 1\}$ and $d_{2i} : H_i \mapsto \{0, 1\}$, which we define in a footnote to preserve continuity.[1]

- If $d_{1i}(h_i^t) = 1$, then $P_i(h_i^t) = \mathcal{A}^j$
- If $d_{2i}(h_i^t) = 1$, then $P_i(h_i^t) = \mathcal{B}(2T - s(t))$
- Otherwise, $P_i(h_i^t) = \begin{cases} \mathcal{B}(p-1) & \text{if } p > 1 \\ \mathcal{A} & \text{if } p = 1 \end{cases}$

Suppose that $P_i(h_i^{t-1}) = \mathcal{A}^j$.

- If $a_j^t = C_{js(t)}$ and $s(t) < T$, then $P_i(h_i^t) = \mathcal{A}^j$
- If $a_j^t = C_{js(t)}$ and $s(t) = T$, then $P_i(h_i^t) = \mathcal{A}$
- Otherwise, $P_i(h_i^{t-1}) = B(2T - s(t))$

[1] Define $d_{1i} : H_i \mapsto \{0, 1\}$ and $d_{2i} : H_i \mapsto \{0, 1\}$ as follows:

$d_{1i}(h_i^t) = 1$ if and only if:

a) $s(t) = nT_*$
b) $a_i^\tau \neq C_{is(\tau)}$ for some $\tau \in \{t - s(t) + 1, \ldots, t - s(t) + (n-1)T_*\}$
c) There is a unique player $j \in N_i$ such that:
   i) $a_j^\tau = C_{js(\tau)}$ for all $\tau \in \{t - s(t) + 1, \ldots, t - s(t) + nT_*\}$,
   ii) $a_j^\tau \neq C_{js(\tau)}$ for all $\tau \in \{t - s(t) - T + 1, \ldots, t - s(t) - T + nT_*\}$

$d_{2i}(h_i^t) = 1$ if and only if:

a) $s(t) = T$
b) There is a player $j \in N_i$ such that:
   i) $a_j^\tau = C_{js(\tau)}$ for all $\tau \in \{t - s(t) + 1, \ldots, t - s(t) + nT_*\}$
   ii) $a_j^\tau \neq C_{js(t)}$ for some $\tau \in \{t - s(t) + nT_* + 1, \ldots, t - s(t) + T\}$

Intuition

After a history h_i^{t-1} in which Player i is in state $\mathcal{A}$, if Player i and her neighbours all play the on-path action $C_{js(t)}$, then Player i remains in state $\mathcal{A}$ after history h_t. If anyone deviates, then after history h^t Player i will be in state $\mathcal{B}(T - s(t))$ if t is in the first nT_* periods of a length T block or $\mathcal{B}(2T - s(t))$ if t is in the last $T - nT_*$ periods of a length T block. An interpretation of $\mathcal{B}(p)$ is that Player i should play D for p periods; however, this interpretation is not strictly correct because after a history where $d_{1i}(h^t) = 1$, which is possible only when $s(t) = nT_*$, Player i could 'transition' from $\mathcal{B}(T - nT_* + 1)$ to $\mathcal{A}^j$ (by 'transition', we mean that $P_i(h_i^{t-1}) = \mathcal{B}(T - nT_* + 1)$ and $P_i(h_i^t) = \mathcal{A}^j$, and after a history where $d_{2i}(h^t) = 1$, which is possible only when $s(t) = T$, Player i could transition from $\mathcal{B}(1)$ to $\mathcal{B}(T)$.

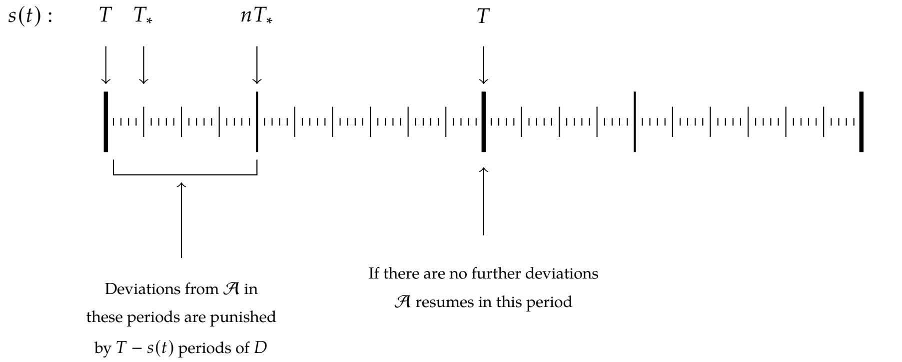

Figure 3.1: Transitions in $\mathcal{A}$, $s(t) \in \{1, \ldots, nT_*\}$

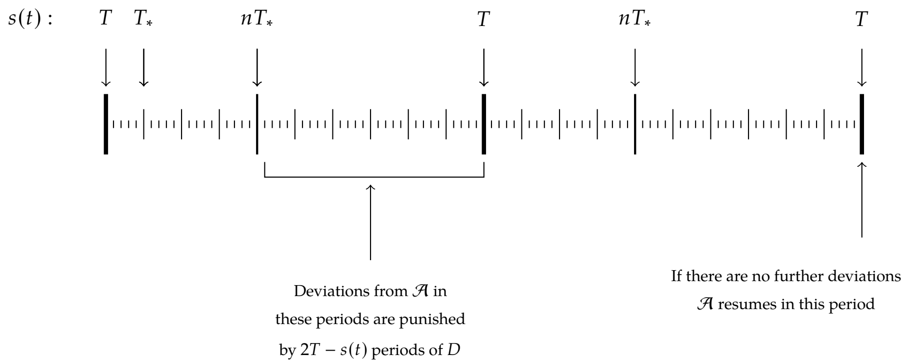

Figure 3.2: Transitions in $\mathcal{A}$, $s(t) \in \{nT_* + 1, \ldots, T\}$

Loosely speaking, deviations from $\mathcal{A}$ in the first part of each length T block are punished until the end of the block, and deviations in the second part of each length T block are punished until the end of the next block. The idea is that any player who deviates from $\mathcal{A}$ is punished by D for a sufficiently long time (at least $T - nT_*$ periods). The reason that first part of the length T block is nT_* periods is to ensure that ζ_N is stable, in the sense that if every player follows the strategy after an arbitrary history h, after finitely many periods, every player will return to state $\mathcal{A}$.

To see the intuition behind this result, suppose that the network is connected and for all i, $C_{is(t)} \neq D$ for some $s(t)$ (that is, every player has to play an action other than D in some period on the equilibrium path). If Player 1 is in state $\mathcal{B}(T)$ at the start of a length T block, then she will play D at least for the first nT_* periods. To see this, note that for any h_i^t such that $s(t) \notin \{nT_*, T\}$,

$d_{1i}(h_i^t) = 0$ and $d_{2i}(h_i^t) = 0$, and so if $P(h_i^{t-1}) = \mathcal{B}(p)$, then $P(h_i^t) = \mathcal{B}(p-1)$. Thus, if Player 1 is in state $\mathcal{B}(T)$ at the start of a length T block, after the first period of the length T block (when $s(t) = 1$), she will be in state $\mathcal{B}(T-1)$, and after after $nT_* - 1$ periods, she will be in state $\mathcal{B}(T - nT_* + 1)$. So she will play D in at least the first nT_* periods.

Since the on-path actions are repeated every T_* periods, after T_* periods Player 1 must have played D in some period t where $D \neq C_{1s(t)}$, and hence her neighbours will also be in state $\mathcal{B}(T - T_*)$. By the same argument as before, they will also play D until at least the nT_*-th period of the length T block. Then after $2T_*$ periods, all of their neighbours will be in state $\mathcal{B}(T - T_*)$, and after $(n-1)T_*$ periods, everyone in the network will be in state $\mathcal{B}(T - (n-1)T_*)$.

After the nT_*-th period of the block, it is possible that some Player i may be in state $\mathcal{A}^j$ for some $j \in N_i$; in the Appendix, we will deal with this case, but under our current assumption that for all i, $C_{is(t)} \neq D$ for some $s(t)$, it is easy to verify that part (i) of condition (c) in the definition of d_{1j} will not be satisfied by any $j \in N_i$ since everyone in the network is playing D after $(n-1)T_*$ periods. Thus, everyone will continue playing D until the end of the length T block.

After the last period of the length T block, when $s(t) = T$, everyone will be in state $\mathcal{A}$ unless d_{2j} is equal to 1 after that history for some i. However, since everyone was playing D after $(n-1)T_*$ periods, part (i) of condition (b) in the definition of d_{2j} cannot be satisfied; hence after T periods, everyone will be in state $\mathcal{A}$. In the next subsection we will discuss stability in more detail.

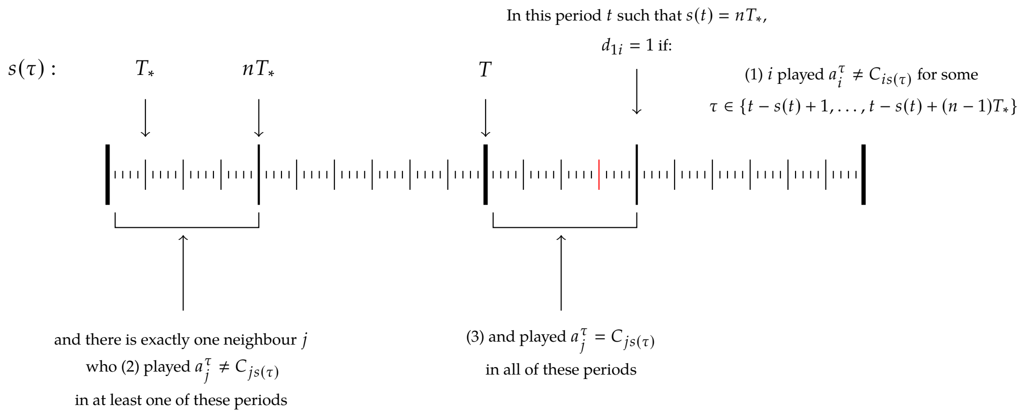

Figure 3.3: Transitions in $\mathcal{B}$: d_1

Now we discuss the purpose of d_{1i} and d_{2i}. If we consider the strategy without the transitions out of $\mathcal{B}(p)$ defined by these two functions, there are two types of deviations that are possible for certain network structures that could be profitable. If Player j deviates towards the end of the first part of a length T block, it is possible that in the next length T block some neighbours are in state $\mathcal{A}$ and some neighbours are in state $\mathcal{B}(p)$ in the second part of that block, and having some neighbours play C and others play D may yield a high payoff to Player j. Then repeatedly deviating in this way could improve Player j's repeated game payoff (note that the one shot deviation principle does not apply since $\delta = 1$, and any finite sequence of deviations will have no effect on repeated game payoffs owing to stability).

For a concrete example of this type of deviation, consider 4 players connected on a line. Let $A_i = \{C, D\}$, and suppose that on the equilibrium path, C is played in every period (so we can let $T_* = 1$, $nT_* = 4$). Assume that every player $i \neq 2$ will play according to the strategy profile ζ_i. Now suppose that there are no deviations up to period 3, so that each player is in state $\mathcal{A}$ after history h^3, and Player 2 deviates for the first time in period 4, so that Players 1, 2, and 3 are in state

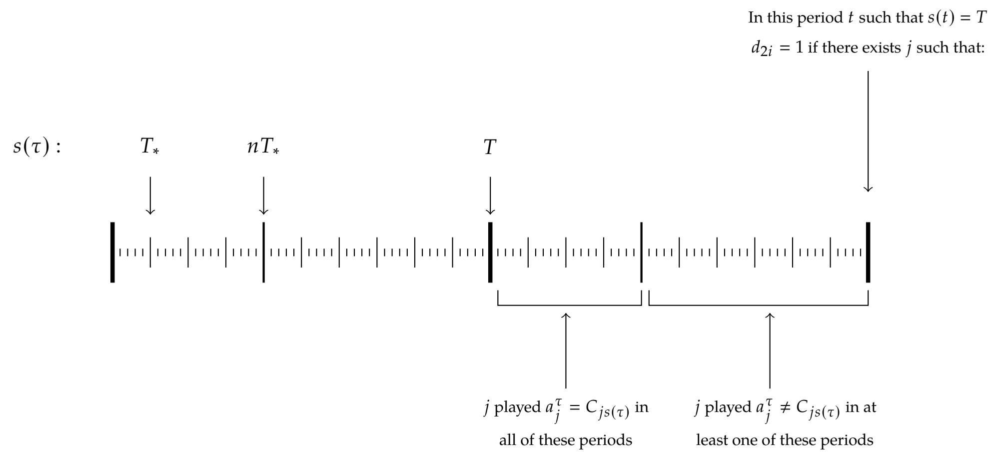

Figure 3.4: Transitions in $\mathcal{B}$: d_2

$\mathcal{B}(T-4)$ after history h^4. Note that Player 4 then does not observe D until period 5, and so Player 4 is in state $\mathcal{B}(2T-5)$ after h^5. Thus, after period T, Players 1, 2, and 3 are in state $\mathcal{A}$, but Player 4 is in state $\mathcal{B}(T)$.

In period $T+1$, Player 4 will play D, so Player 3 will play D from period $T+2$ onwards. Note however, that if Player 2 does not play D, Player 1 will remain in state $\mathcal{A}$. Thus, from periods $T+2$ to $2T$, Player 1 will play C and if Player 3 remains in state $\mathcal{B}(p)$ she will will play D, which may be beneficial for Player 2. The transition out of $\mathcal{B}(p)$ when d_{1i} is equal to 1 is designed to rule out exactly this type of deviation. In this case, if Player 2 plays C in every period from $T+1$ to $T+4$, then $d_{1_3}(h_3^{T+4}) = 1$, and Player 3 will be in state $\mathcal{A}^2$ after h^{T+4}, and revert to playing C (unless Player 2 deviates to D again). Intuitively, if a player has already been punished for a deviation, and refuses to match another player's punishment, then the latter player should end the punishment in case it is actually benefiting the former player.

The second type of deviation occurs if a player finds herself being punished by one group of neighbours but not another. By deviating near the end of the length T block, she can reverse the pattern, which may also be profitable. Using the same set up as the previous example, suppose that after some arbitrary history, at the start of a length T block, Player 1 is in state $\mathcal{B}(T)$, and Player 3 is in state $\mathcal{A}$. Now as long as Player 2 plays C in this length T block, Player 1 will play D and Player 3 will play C, which may be beneficial for Player 2. It may also be beneficial for Player 2 if Player 1 plays C and Player 3 plays D. Suppose that Player 2 plays D in the last period of the length T block. Then Player 3 will be in state $\mathcal{B}(T)$ after the last period of the length T block, and play D in every period in the next block. Note that Player 1 is in state $\mathcal{B}(1)$ in the penultimate period of the length T block. If she is in state $\mathcal{A}$ after the last period of the length T block, she will play C in the next block. The transition out of $\mathcal{B}(p)$ when $d_{2i} = 1$ is designed to ensure that in this situation, Player 1 will instead be in state $\mathcal{B}(T)$ after the last period of the length T block and play D for an additional block. Intuitively, if a player deviates in the second part of a block, playing on-path actions in the first part, all of her neighbours should punish her until the end of the next block even if they were already in state $\mathcal{B}(p)$. This ensures that even when some neighbours are in state $\mathcal{B}(p)$, any deviation from the on-path actions will be punished with D *by every neighbour* for at least $T - nT_*$ periods.

If Player j is in state $\mathcal{A}^k$, for $k \neq i$, then Player j will not punish any deviation by Player i. However, it is shown in the Appendix that if all players other than i have been playing according to $\zeta_{N\setminus\{i\}}$ for a sufficiently long time, then Player j cannot be in state $\mathcal{A}^k$, for $k \neq i$.

Stability

If everyone plays according to the strategy ζ_N, $C_{Ns(t)}$ will be played for all $t \geq 1$. An important feature of the strategy profile ζ_N is that if every player plays according to ζ_N^h after an arbitrary history h of length z, after at most $2T$ periods, every player will play as if there are no deviations in h. This is established in the following definition and Lemma.

Definition 3.5.1. For any arbitrary history h of length z, let $\hat{h}^t$, $t \geq z$, be the length t history generated by ζ_N after h. The strategy profile ζ_N is *stable* if for any h there exists a T such that $\zeta_N(\hat{h}^t) = C_{Ns(t)}$ for all $t > T$.

Lemma 3.5.1. *The strategy profile ζ_N is stable. Moreover, for any h of length z, $\zeta_N(\hat{h}^t) = C_{Ns(t)}$ for all $t > z + 2T - s(z)$.*

Proof of Lemma 3.5.1. Appendix. □

The proof proceeds as follows. Let N_C be the set of all players such that $C_{is(t)} \neq D$ for some $s(t)$, and let $(N_C^1, \dots, N_C^L)$ be a partition of N_C into connected components, i.e. each $(N_C^l, G_{N_C^l})$ is a connected component of (N_C, G_{N_C}).

First it is shown that for any arbitrary history of length z, for any $i \in N$, either $P_i(\hat{h}_i^{z+T-s(z)}) = \mathcal{A}$ or $P_i(\hat{h}_i^{z+T-s(z)}) = \mathcal{B}(T)$. If for all $i \in N_C^l$, $P_i(\hat{h}_i^{z+T-s(z)}) = \mathcal{A}$, then in period $z + 2T - s(z)$, for each $i \in N_C^l$, $P_i(\hat{h}_i^{z+T-s(z)}) = \mathcal{A}$.

Suppose that in period $z + T - s(z)$ there is at least one $i \in N_C^l$ such that $P_i(\hat{h}_i^{z+T-s(z)}) = \mathcal{B}(T)$. Using the fact that the on-path actions are repeated every T_* periods, it is then shown that by period $z+T-s(z)+nT_*$, for all $i \in N_C^l$, either $P_i(\hat{h}_i^{z+T-s(z)+nT_*}) = \mathcal{B}(T-nT_*)$ or $P_i(\hat{h}_i^{z+T-s(z)+nT_*}) = \mathcal{A}^j$ for some $j \notin N_C$. In either case, it follows that in period $z + 2T - s(z)$, $P_i(\hat{h}_i^{z+2T-s(z)}) = \mathcal{A}$ for all $i \in N_C^l$.

Since this is true for any N_C^l and $N_C = \cup_{l=1}^L N_C^l$, it follows that for any $i \in N_C$, i will play $\hat{a}_i^t = C_{is(t)}$ for all $t > z + 2T - s(z)$. Since any $i \notin N_C$ always plays $C_{is(t)} = D$ regardless of history, it follows that for all $i \in N$, $\hat{a}_i^t = C_{is(t)}$ for all $t > z + 2T - s(z)$.

3.6 Proof of Proposition 3.4.1

Now we will prove Proposition 3.4.1 by showing that the strategy profile ζ_N is a sequential equilibrium and supports the desired payoff v^*.

Proof of Proposition 3.4.1. First, we show that for any $i \in N$, $\mathcal{V}_i(\zeta_N) = v_i^*$. Let $\{a_N^t\}_{t=1}^\infty$ be the sequence of stage game actions generated by ζ_N, and note that:

$$\begin{aligned}\mathcal{V}_i(\zeta_N) &= \Lambda\left(\left\{\frac{1}{K}\sum_{t=1}^K v_i(a_i^t, a_{N_i}^t)\right\}_{K=1}^\infty\right)\\ &= \lim_{K\to\infty}\frac{1}{K}\sum_{t=1}^K v_i(a_i^t, a_{N_i}^t)\\ &= v_i^*.\end{aligned}$$

The second equality follows from the fact that $\Lambda(x) = \lim_{n\to\infty} x_n$ for each $x \in c$, the space of all convergent sequences (Lemma 16.45 in Aliprantis and Border (2006) p. 550). To see that the limit of $\frac{1}{K}\sum_{t=1}^{K} v_i(a_i^t, a_{N_i}^t)$ exists and is indeed equal to v_i^*, note that for $K > T$ the following expression is valid:

$$\begin{aligned}
&\left|\frac{1}{K}\sum_{t=1}^{K} v_i(a_i^t, a_{N_i}^t) - v_i^*\right| \\
&= \left|\frac{1}{K}\sum_{t=1}^{K-s(K)} v_i(C_{is(t)}, C_{N_i s(t)}) - v_i^* + \frac{1}{K}\sum_{t=K-s(K)+1}^{K} v_i(C_{is(t)}, C_{N_i s(t)})\right| \\
&\leq \left|\left(\frac{1}{K} - \frac{1}{K-s(K)}\right)\sum_{t=1}^{K-s(K)} v_i(C_{is(t)}, C_{N_i s(t)}) + \frac{1}{K}\sum_{t=K-s(K)+1}^{K} v_i(C_{is(t)}, C_{N_i s(t)})\right| \\
&< \frac{KT \max_{a_i, a_{N_i}} v_i(a_i, a_{N_i})}{K(K-T)} + \frac{T \max_{a_i, a_{N_i}} v_i(a_i, a_{N_i})}{K},
\end{aligned}$$

which can be made arbitrarily small for large K. In the third line, we have used the fact that when on-path actions are played over complete length T blocks, the average payoff over those blocks is v^*, i.e. $\frac{1}{K-s(K)}\sum_{t=1}^{K-s(K)} v_i(C_{is(t)}, C_{N_i s(t)}) = v_i^*$, which is true by construction.

Given a strategy profile $\sigma_N = (\sigma_1, \ldots, \sigma_n)$ and a history h, let $\sigma_N^h = (\sigma_1^h, \ldots, \sigma_n^h)$ be the profile induced by the history h. We show that for any history h and any strategy $\theta_i \in \Sigma_i$,

$$\mathcal{V}_i(\zeta_N^h) \geq \mathcal{V}_i(\theta_i, \zeta_{-i}^h).$$

Since this is true for any history h, it is optimal for player i to follow the strategy at each of her information sets $\mathcal{I}(h_i)$, whatever her beliefs about h conditional on h_i may be.

Consider any history $h \in H$ of length z. Let $\{\hat{a}_N^t\}_{t=z+1}^{\infty}$ be the sequence of stage game actions generated by ζ_N^h after history h. By Lemma 3.5.1 and the properties of Banach-Mazur limits, $\mathcal{V}_i(\zeta_N^h) = v_i^*$. To see this, first note that for $K \geq 2T - s(z)$:

$$\frac{1}{K}\sum_{t=z+1}^{z+K} v_i(\hat{a}_i{}^t, \hat{a}_{N_i}^t) = \frac{1}{K}\sum_{t=z+1}^{z+2T-s(z)} v_i(\hat{a}_i{}^t, \hat{a}_{N_i}^t) + \frac{1}{K}\sum_{t=z+2T-s(z)+1}^{z+K} v_i(\hat{a}_i{}^t, \hat{a}_{N_i}^t).$$

Then by the property that $\Lambda(x_1, x_2, \ldots) = \Lambda(x_2, x_3, \ldots)$ for each $(x_1, x_2, \ldots) \in \ell_\infty$:

$$\begin{aligned}
\mathcal{V}_i(\zeta_N^h) &= \Lambda\left(\left\{\frac{1}{K}\sum_{t=z+1}^{z+K} v_i(\hat{a}_i{}^t, \hat{a}_{N_i}^t)\right\}_{K=2T-s(z)}^{\infty}\right) \\
&= \Lambda\left(\left\{\frac{1}{K}\sum_{t=z+1}^{z+2T-s(z)} v_i(\hat{a}_i{}^t, \hat{a}_{N_i}^t) + \frac{1}{K}\sum_{t=z+2T-s(z)+1}^{z+K} v_i(\hat{a}_i{}^t, \hat{a}_{N_i}^t)\right\}_{K=2T-s(z)}^{\infty}\right) \\
&= \lim_{K\to\infty}\left(\frac{1}{K}\sum_{t=z+1}^{z+2T-s(z)} v_i(\hat{a}_i{}^t, \hat{a}_{N_i}^t) + \frac{1}{K}\sum_{t=z+2T-s(z)+1}^{z+K} v_i(\hat{a}_i{}^t, \hat{a}_{N_i}^t)\right) \\
&= 0 + v_i^*,
\end{aligned}$$

where third equality follows since the sequence converges (which we will show), and the fourth equality follows from Lemma 3.5.1 (which implies that for all $t > z + 2T - s(z)$, $\hat{a}_j^t = C_{js(t)}$ for all $j \in N$) and the fact that $\lim_{K\to\infty}\frac{1}{K}\sum_{t=z+2T-s(z)+1}^{z+K} v_i(C_{is(t)}, C_{N_i s(t)}) = v_i^*$. To see this, note that for

$K > 3T - s(z)$ the following expression is valid:

$$\begin{aligned}
&\left| \frac{1}{K} \sum_{t=z+2T-s(z)+1}^{z+K} v_i(C_{is(t)}, C_{N_i s(t)}) - v_i^* \right| \\
&= \left| \frac{1}{K} \sum_{t=z+2T-s(z)+1}^{z+K-s(z+K)} v_i(C_{is(t)}, C_{N_i s(t)}) - v_i^* + \frac{1}{K} \sum_{t=z+K-s(z+K)+1}^{z+K} v_i(C_{is(t)}, C_{N_i s(t)}) \right| \\
&= \left| \left(\frac{1}{K} - \frac{1}{K-\kappa} \right) \sum_{t=z+2T-s(z)+1}^{z+K-s(z+K)} v_i(C_{is(t)}, C_{N_i s(t)}) + \frac{1}{K} \sum_{t=z+K-s(z+K)+1}^{z+K} v_i(C_{is(t)}, C_{N_i s(t)}) \right| \\
&< \frac{3KT \max_{a_i, a_{N_i}} v_i(a_i, a_{N_i})}{K(K-3T)} + \frac{T \max_{a_i, a_{N_i}} v_i(a_i, a_{N_i})}{K},
\end{aligned}$$

where $\kappa \equiv s(z+K) + 2T - s(z)$. For the last inequality, we multiply the maximum number of terms in each summation by the maximum value of the summand and replace κ by its maximum value, noting that since $\kappa < K$, $1/K - 1/(K-\kappa)$ is less than zero and decreasing in κ. This bound can be made arbitrarily small for large K. Since $z + 2T - s(z) + 1$ is the beginning of a length T block, $\frac{1}{K-\kappa} \sum_{t=z+2T-s(z)+1}^{z+K-s(z+K)} v_i(C_{is(t)}, C_{N_i s(t)}) = v_i^*$, which explains the third line.

Thus, we need to show that for any player $i \in N$ and any strategy $\theta_i \in \Sigma_i$,

$$v_i^* \geq \mathcal{V}_i(\theta_i, \zeta_{-i}^h).$$

Let $\{\bar{a}_N^t\}_{t=z+1}^{\infty}$ be the sequence of stage game actions generated by (θ_i, ζ_{-i}^h) after history h, and let $\bar{h}^t$, $t \geq z$, be the length t history generated by (θ_i, ζ_{-i}^h) after h. Note the payoff for player i satisfies:

$$\begin{aligned}
\sum_{t=z+1}^{z+K} v_i(\bar{a}_i^t, \bar{a}_{N_i}^t) &= \sum_{t=z+1}^{z+K} v_i(C_{is(t)}, C_{N_i s(t)}) + \sum_{t=z+1}^{z+K} \left[v_i(\bar{a}_i^t, \bar{a}_{N_i}^t) - v_i(C_{is(t)}, C_{N_i s(t)}) \right] \\
&= \sum_{t=z+1}^{z+K} v_i(C_{is(t)}, C_{N_i s(t)}) + \Pi(K)
\end{aligned}$$

where:

$$\begin{aligned}
\Pi(K) &\equiv \sum_{t=z+1}^{z+K} \left[v_i(\bar{a}_i^t, \bar{a}_{N_i}^t) - v_i(C_{is(t)}, C_{N_i s(t)}) \right] \\
&\leq 6T \max_{a_i a_{N_i}, a_i' a_{N_i}'} |v_i(a_i, a_{N_i}) - v_i(a_i', a_{N_i}')|
\end{aligned} \tag{3.1}$$

for all $K \geq 1$ (when $T - nT_*$ is chosen to be sufficiently large). We establish inequality 3.1 as a Lemma:

Lemma 3.6.1. *For an appropriately chosen value of T, $\Pi(K) \leq 6TM$ for all $K \geq 1$, where $M \equiv \max_{a_i a_{N_i}, a_i' a_{N_i}'} |v_i(a_i, a_{N_i}) - v_i(a_i', a_{N_i}')|$.*

Proof. Appendix. □

Note also:

$$\lim_{K \to \infty} \frac{1}{K} \sum_{t=z+1}^{z+K} v_i(C_{is(t)}, C_{N_i s(t)}) = v_i^*.$$

To see this, note that for $K > 2T$ the following expression is valid:

$$\begin{aligned}
&\left|\frac{1}{K}\sum_{t=z+1}^{z+K} v_i(C_{is(t)}, C_{N_i s(t)}) - v_i^*\right| \\
\leq &\left|\frac{1}{K}\sum_{t=z+1}^{z+T-s(z)} v_i(C_{is(t)}, C_{N_i s(t)}) + \frac{1}{K}\sum_{t=z+K-s(z+K)+1}^{z+K} v_i(C_{is(t)}, C_{N_i s(t)})\right| \\
&+\left|\frac{1}{K}\sum_{t=z+T-s(z)+1}^{z+K-s(z+K)} v_i(C_{is(t)}, C_{N_i s(t)}) - v_i^*\right| \\
= &\left|\frac{1}{K}\sum_{t=z+1}^{z+T-s(z)} v_i(C_{is(t)}, C_{N_i s(t)}) + \frac{1}{K}\sum_{t=z+K-s(z+K)+1}^{z+K} v_i(C_{is(t)}, C_{N_i s(t)})\right| \\
&+\left|\left(\frac{1}{K} - \frac{1}{K-\nu}\right)\sum_{t=z+T-s(z)+1}^{z+K-s(z+K)} v_i(C_{is(t)}, C_{N_i s(t)})\right| \\
< &\frac{2T \max_{a_i, a_{N_i}} v_i(a_i, a_{N_i})}{K} + \frac{2KT \max_{a_i, a_{N_i}} v_i(a_i, a_{N_i})}{K(K-2T)},
\end{aligned}$$

where $\nu \equiv s(z+K) + T - s(z)$. This last line can be made arbitrarily small for large K. The argument here is similar to when we showed that $\lim_{K\to\infty} \frac{1}{K}\sum_{t=1}^{K} v_i(a_i^t, a_{N_i}^t) = v_i^*$, the difference being that now we have to 'trim' the beginning as well as the end of the summation.

Combining the previous results, we have that:

$$\limsup_{K\to\infty} \frac{1}{K}\sum_{t=z+1}^{z+K} v_i(\bar{a}_i^t, \bar{a}_{N_i}^t) \leq v_i^*.$$

Since $\Lambda(x) \leq \limsup x_n$ for any $x = (x_1, x_2, \ldots) \in \ell_\infty$ (Lemma 16.45 in Aliprantis and Border (2006) p. 550), $\mathcal{V}_i(\theta_i, \zeta_{-i}^h) = \Lambda\left(\left\{\frac{1}{K+1}\sum_{t=z}^{z+K} v_i(\bar{a}_i^t, \bar{a}_{N_i}^t)\right\}_{K=0}^{\infty}\right) \leq v_i^*$. □

3.7 Conclusion

We have shown that it is possible to sustain as a sequential equilibrium any payoff that can be achieved by a finite deterministic sequence of stage game actions such that each player receives strictly more than her Nash equilibrium payoff. We provide incentives to play the equilibrium strategies by punishing deviations for at least $T - nT_*$ periods. In order to achieve stability, the exact number of periods for which each deviation is punished depends on when it occurs. Similar strategies can be used to prove a folk theorem without the assumption of local interaction as long as the entire network is a connected component. In this case, we would also need a signalling action to ensure that deviations spread to the entire network. It may be possible to extend the result to payoffs that require infinite sequences using a trick similar to Fudenberg and Maskin (1991). For stability, we would need the non-punishing action to be played sufficiently often, so that punishment can spread to the entire network quickly enough, and it seems plausible that sequential rationality would be maintained if the continuation payoff at each date remained close to the target payoff.

Appendix 3.A Proof of Lemma 3.5.1

Lemma 3.A.1. *For an arbitrary history h_i^t of length t:*

1. *If $P_i(h_i^t) = \mathcal{B}(p)$, $p > 0$, then $p = T - s(t)$ or $p = 2T - s(t)$*
2. *If $s(t) \in \{1, \dots, nT_*\}$, then $P_i(h_i^{t-1}) \neq \mathcal{A}^j$*

Proof. We prove part 1 by induction on the history length. For the empty history, the antecedent is false, and so the statement is true. Now suppose that for all h_i^{t-1}, $P_i(h_i^{t-1}) = \mathcal{B}(p)$ implies $p = T - s(t-1)$ or $p = 2T - s(t-1)$. We show that $P_i(h_i^t) = \mathcal{B}(p)$ implies $p = \mathcal{B}(T - s(t))$ or $p = \mathcal{B}(2T - s(t))$. If $P_i(h_i^{t-1}) = \mathcal{A}$ or $P_i(h_i^{t-1}) = \mathcal{A}^j$, and $P_i(h_i^t) = B(p)$, then $p = p(t) = T - s(t)$ or $p = p(t) = 2T - s(t)$ by definition. Suppose that $P_i(h_i^{t-1}) = \mathcal{B}(T - s(t-1))$ or $P_i(h_i^{t-1}) = \mathcal{B}(2T - s(t-1))$. Note that $P_i(h_i^t)$ is either $\mathcal{A}^j$, $\mathcal{B}(2T - s(t))$, or $\mathcal{B}(p-1)$. For $s(t-1) < T$, $s(t) = s(t-1) + 1$, so $p - 1 = T - s(t)$ or $p - 1 = 2T - s(t)$. For $s(t-1) = T$, $p - 1 = T - 1 = T - s(t)$. Thus, in each case, $P_i(h_i^t) = \mathcal{B}(p)$ implies $p = \mathcal{B}(T - s(t))$ or $p = \mathcal{B}(2T - s(t))$.

For part 2, note that $P_i(h_i^t) = \mathcal{A}^j$ and $P_i(h_i^{t-1}) \neq \mathcal{A}^j$ if and only if $t = nT_*$, and when $s(t) = T$, $P_i(h_i^t) \neq \mathcal{A}^j$. □

Lemma 3.A.2. *Take any h, hh' of lengths t and $t + t'$ such that $s(t), s(t+t') \in \{1, \dots, nT_* - 1\}$. For $\tau \leq t$, let a_N^τ be the period τ action profile in h. Then:*

1. *For any $i \in N$, $P_i(h) = \mathcal{A}$ or $P_i(h) = \mathcal{B}(T - s(t))$*
2. *If $P_i(h) = \mathcal{B}(T - s(t))$, then $P_i(hh') = \mathcal{B}(T - s(t + t'))$*
3. *If for any $\tau \in \{t - s(t) + 1, \dots, t\}$, $(a_i^\tau, a_{N_i}^\tau) \neq (C_{is(\tau)}, C_{N_i s(\tau)})$, $P_i(h) = \mathcal{B}(T - s(t))$*

Proof. For any history h of length t such that $s(t) = T$, if $P_i(h_i) = \mathcal{B}(p)$, then p is at most T by Lemma 3.A.1. For any h of length t such that $s(t) \in \{1, \dots, nT_* - 1\}$, if $P_i(h_i^{t-1}) = \mathcal{A}$, $P_i(h_i) = \mathcal{A}$ or $P_i(h_i) = \mathcal{B}(T - s(t))$, and if $P_i(h_i^{t-1}) = \mathcal{B}(p)$, $P_i(h_i) = \mathcal{B}(p-1)$. Thus, if $P_i(h_i) = \mathcal{B}(p)$, $p < T$. By Lemma 3.A.1, $P_i(h_i) = \mathcal{A}$, $P_i(h_i) = \mathcal{B}(T - s(t))$, or $P_i(h) = \mathcal{B}(2T - s(t))$, but by the previous claim, it cannot be $\mathcal{B}(2T - s(t))$.

Part 2 follows immediately from the definition of P_i after noting that for any h_i^{t-1} such that $s(t) \in \{1, \dots, nT_* - 1\}$, if $P_i(h_i^{t-1}) = \mathcal{B}(p)$ then $P_i(h_i^t) = \mathcal{B}(p-1)$. For part 3, the definition of P_i implies that $P_i(h^\tau) \neq \mathcal{A}$, and the claim follows from parts 1 and 2. □

Proof of Lemma 3.5.1. For an arbitrary history h of length t, let $\hat{h}^\tau$, $\tau \geq t$ be the length τ history generated by (ζ_N^h) after h, and let $\{\hat{a}_N^\tau\}_{\tau=t}^\infty$ be the corresponding sequence of stage game action profiles. Let N_C denote the set of players who have to play an action other than D in some period on the equilibrium path. Note that for each player $i \notin N_C$, $\hat{a}_i^\tau = C_{is(\tau)}$ for all $\tau \geq t + 1$. Thus, we need to show that for all $i \in N_C$, $P_i(\hat{h}^\tau) = \mathcal{A}$ for all $\tau \geq t + 2T - s(t)$, which implies that $\hat{a}_i^\tau = C_{is(\tau)}$ for all $\tau > t + 2T - s(t)$ for all $i \in N_C$.

Let $(N^1, \dots, N^L)$ be a partition of N_C into L connected components. That is, each $(N_C^l, G_{N_C^l})$ is connected component of (N_C, G_{N_C}). Note that a player $i \in N_C^l$ cannot have a neighbour $j \in N_C^k$, $k \neq l$, because in that case $(N_C^l \cup \{j\}, G_{N_C^l \cup \{j\}})$ would be connected, contradicting the definition of a connected component. We will consider the connected component N_C^l, where l is arbitrary, and show that for all $i \in N_C^l$, $P_i(\hat{h}_i^{t+2T-s(t)}) = \mathcal{A}$.

First, we argue that if for all $i \in N_C^l$, $P_i(\hat{h}_i^k) = \mathcal{A}$ for some period k, then for all $i \in N_C^l$, $P_i(\hat{h}_i^\tau) = \mathcal{A}$ for all $\tau \geq k$. To see this, consider the set $B \equiv N_C^l \cup \{j : j \in N_i$ for some $i \in N_C^l\}$, and note that if $P_i(\hat{h}_i^{\tau-1}) = \mathcal{A}$ for all $i \in N_C^l$, then $P_i(\hat{h}_i^{\tau}) = \mathcal{A}$ for all $i \in N_C^l$. This is because each $j \in B$ is either in N_C^l or not in N_C (since N_C^l is a connected component of N_C). Each $j \in N_i \cap N_C^l$ will play $C_{js(\tau)}$ since $P_j(\hat{h}_j^{\tau-1}) = \mathcal{A}$, each $j \notin N_C$ will play $D = C_{js(\tau)}$ regardless of $P_j(\hat{h}_j^{\tau-1})$, and i will play $C_{is(\tau)}$ since $P_i(\hat{h}_i^{\tau-1}) = \mathcal{A}$. Thus, $\hat{a}_j^{\tau} = C_{js(\tau)}$ for all $j \in N_i \cup \{i\}$, and so $P_i(\hat{h}_i^{\tau}) = \mathcal{A}$.

By Lemma 3.A.1, for any $i \in N_C^l$, we have either $P_i(\hat{h}_i^{t+T-s(t)}) = \mathcal{A}$ or $P_i(\hat{h}_i^{t+T-s(t)}) = \mathcal{B}(T)$. If $P_i(\hat{h}_i^{t+T-s(t)}) = \mathcal{A}$ for all $i \in N_C$, then for all $i \in N_C^l$, $P_i(\hat{h}_i^{\tau}) = \mathcal{A}$ for all $\tau \geq t + T - s(t)$, so assume that for some $i \in N_C^l$, $P_i(\hat{h}_i^{t+T-s(t)}) = \mathcal{B}(T)$.

Without loss of generality, let Player 1 belong to N_C^l and let $P_1(\hat{h}_l^{t+T-s(t)}) = \mathcal{B}(T)$. First, we argue that for all $i \in N_C^l$, $P_i(\hat{h}_i^{t+T-s(t)+nT_*-1}) = \mathcal{B}(T - nT_* + 1)$. For any player $i \in N_C^l$, there exists a path $(1, j_2, \ldots, j_{m-1}, i)$ such that each player in the path belongs to N_C^l (since N_C^l is connected), and $m \leq n$. Note that player 1 will play D in every period $\tau \in \{t + T - s(t) + 1, \ldots, t + T - s(t) + T_*\}$ according to ζ_1^h. Since player 1 belongs to N_C, it must be the case that $C_{1s(\tau)} \neq D$ for at least one $\tau \in \{t + T - s(t) + 1, \ldots, t + T - s(t) + T_*\}$ (recall that the sequence of on-path actions repeat every T_* periods). By Lemma 3.A.2, $P_{j_2}(\hat{h}_{j_2}^{t+T-s(t)+T_*}) = B(T - T_*)$. Now j_2 will play D in every period $\tau \in \{t+T-s(t)+T_*+1, \ldots, t+T-s(t)+2T_*\}$, and since $j_2 \in N_C$, the previous argument implies that $P_{j_3}(\hat{h}_{j_3}^{t+T-s(t)+2T_*}) = B(T-2T_*)$. Thus, by period $t+T-s(t)+(m-1)T_*$, we have $P_i(\hat{h}_i^{t+T-s(t)+(m-1)T_*}) = B(T - (m-1)T_*)$. Since, $m \leq n$, Lemma 3.A.2 implies that $P_i(\hat{h}_i^{t+T-s(t)+nT_*-1}) = B(T - nT_* + 1)$.

Now for each $i \in N_C^l$, $P_i(\hat{h}_i^{t+T-s(t)+nT_*})$ depends on whether $d_{1i}(\hat{h}_i^{t+T-s(t)+nT_*}) = 1$. If there does not exist a unique j satisfying parts (i) and (ii) of condition (c) in the definition of d_{1i}, then $P_i(\hat{h}_i^{t+T-s(t)+nT_*}) = \mathcal{B}(T - nT_*)$. Now we argue that $d_{2i}(h_i^{\tau}) = 0$ for all $\tau \in \{t + T - s(t) + nt_* + 1, \ldots, t + 2T - s(t)\}$, and hence $P_i(\hat{h}_i^{t+2T-s(t)}) = \mathcal{A}$. First note that condition (a) in the definition of d_{2i} implies that $d_{2i}(h_i^{\tau}) = 0$ for all τ such that $s(\tau) \neq T$. Thus, we only need to show that $d_{2i}(h_i^{t+2T-s(t)}) = 0$. To see this, note that for any $j \in N_i$, there is a path $(1, j_2, \ldots, j_{m-1}, j)$, where $m \leq n$, which implies that $P_j(\hat{h}_j^{t+T-s(t)+(n-1)T_*}) = B(T - (n-1)T_*)$, and thus, $\hat{a}_j^{\tau} = D$ for all $\tau \in \{t + T - s(t) + (n-1)T_* + 1, \ldots, t + T - s(t) + nT_*\}$. Thus, for j to satisfy part (i) of condition (b) in the definition of d_{2i}, it must be the case that $j \notin N_C$. But this means that j will play $D = C_{js(\tau)}$ for each $\tau \in \{t + T - s(t) + nT_* + 1, \ldots, t + 2T - s(t)\}$, and so will not satisfy part (ii) of the condition. Therefore, for each $\tau \in \{t + T - s(t) + nt_* + 1, \ldots, t + 2T - s(t)\}$, if $P_i(h_i^{\tau-1}) = \mathcal{B}(p)$, then $P_i(h_i^{\tau}) = \mathcal{B}(p-1)$ for $p > 1$, and $P_i(h_i^{\tau}) = \mathcal{A}$ for $p = 1$. Since $P_i(\hat{h}_i^{t+T-s(t)+nT_*}) = \mathcal{B}(T - nT_*)$, $P_i(\hat{h}_i^{t+2T-s(t)}) = \mathcal{A}$.

If for some i, $d_{1i}(h_i^{t+T-s(t)+nT_*}) = 1$, then $P_i(\hat{h}_i^{t+T-s(t)+nT_*}) = A^{j^*}$, where j^* is the unique neighbour satisfying condition (c) in the definition of d_{1i}. Note that j^* cannot be in N_C, and therefore j^* will play $\hat{a}_{j^*}^{\tau} = C_{j^*s(\tau)}$ for each $\tau \in \{t + T - s(t) + nT_* + 1, \ldots, t + 2T - s(t)\}$, which implies that $P_i(\hat{h}_i^{t+2T-s(t)}) = \mathcal{A}$. □

Appendix 3.B Proof of Lemma 3.6.1

Let $\zeta_i'(h_i) = C_{is(t)}$ for any h_i of length $t-1$. That is, ζ_i' is the strategy where player i plays the on-path actions after every history. For any arbitrary history h of legnth t, let $\tilde{h}_i^{\tau}$, $\tau \geq t$, be the length τ history generated by $(\zeta_i', \zeta_{N\setminus\{i\}})$ after h.

Lemma 3.B.1. *The strategy profile $(\zeta_i', \zeta_{N\setminus\{i\}})$ is stable. Moreover for all $i \in N_C \setminus \{i\}$, $P_i(\tilde{h}_i^{\tau}) = \mathcal{A}$ for all $\tau \geq t + 2T - s(t)$.*

Proof. Replace N_C with $N_C \setminus \{i\}$ in the proof of Lemma 3.5.1. □

Recall that h is an arbitrary history of length z, $\{\bar{a}^t_N\}_{t=z+1}^{\infty}$ is the sequence of stage game actions generated by (θ_i, ζ^h_{-i}) after history h, and $\bar{h}^t$, $t \geq z$ is the length t history generated by (θ_i, ζ^h_{-i}) after h.

Lemma 3.B.2. *Suppose that for $\bar{h}^t_j$, $j \in N_i$, $t > z + 2T - s(t)$, conditions (a) and (b) in the definition of d_{1j} are satisfied. Then $d_{1j}(\bar{h}^t_j) = 1$ if and only if Player i satisfies parts (i) and (ii) of condition (c) in the definition of d_{1j}.*

Proof. We will show that no $k \in N_j \setminus \{i\}$ can satisfy parts (i) and (ii) of condition (c) in the definition of d_{1j}, when conditions (a) and (b) are satisfied. Assume that $s(t) = nT_*$, and take any $k \in N_j \setminus \{i\}$, and note that for k to satisfy (i), $\bar{a}^\tau_k \neq C_{ks(\tau)}$ for some $\tau \in \{t - s(t) - T + 1, \ldots, t - s(t) - T + nT_*\}$. But this implies that $k \in N_C$. Condition (b) requires that $\bar{a}^\tau_j \neq C_{js(\tau)}$ for some $\tau \in \{t - s(t) + 1, \ldots, t - s(t) + (n-1)T_*\}$, which implies that then $\bar{a}^\tau_k = D$ for all $\tau \in \{t - s(t) + (n-1)T_* + 1, \ldots, t - s(t) + nT_*\}$, and in at least one of these periods $C_{ks(\tau)} \neq D$, and so k cannot satisfy condition (ii). Thus, if Player i satisfies parts (i) and (ii) of condition (c), then Player i is the unique player satisfying this condition, and $d_{1j}(\bar{h}^t_j) = 1$. If Player i does not satisfy parts (i) and (ii) of condition (c), then no player satisfies the condition, and $d_{1j}(\bar{h}^t_j) = 0$. □

Lemma 3.B.3 says that when i's opponents play D in every period in a length T_* block, the value of $\Pi(K)$ decreases by at least $T_* v^*_i$.

Lemma 3.B.3. *Assume that for all $j \in N_i$, $\bar{a}^t_j = D$ for all $t \in \{k+1, \ldots, k+T_*\}$, where k is such that k* mod $T_* = 0$. *Then:*

$$\Pi(k - z + T_*) - \Pi(k - z) \leq -T_* v^*_i.$$

Proof.

$$\begin{aligned}
&\Pi(k - z + T_*) - \Pi(k - z) \\
&= \sum_{t=z+1}^{k+T_*} \left[v_i(\bar{a}^t_i, \bar{a}^t_{N_i}) - v_i(C_{is(t)}, C_{N_i s(t)}) \right] - \sum_{t=z+1}^{k} \left[v_i(\bar{a}^t_i, \bar{a}^t_{N_i}) - v_i(C_{is(t)}, C_{N_i s(t)}) \right] \\
&= \sum_{t=k+1}^{k+T_*} \left[v_i(\bar{a}^t_i, \bar{a}^t_{N_i}) - v_i(C_{is(t)}, C_{N_i s(t)}) \right] \\
&= \sum_{t=k+1}^{k+T_*} \left[v_i(\bar{a}^t_i, D, \ldots, D) - v_i(C_{is(t)}, C_{N_i s(t)}) \right] \\
&\leq - \sum_{t=k+1}^{k+T_*} v_i(C_{is(t)}, C_{N_i s(t)}) \\
&= -T_* v^*_i.
\end{aligned}$$

□

Lemma 3.B.4 says that if each player other than i has been playing according to $\zeta_{N \setminus \{i\}}$ for a sufficiently long time, and i plays an action other than $C_{is(t)}$ in some period t in the first part of a length T block, then all of i's neighbours will punish i with D until the end of the length T block.

Lemma 3.B.4. *For any $k_m \geq z + 2T - s(z)$ such that $s(k_m) = T$, if $\bar{a}^{t^*}_i \neq C_{is(t^*)}$, where $t^* \in \{k_m + 1, k_m + nT_*\}$, then for all $j \in N_i$, $\bar{a}^t_j = D$ for all $t \in \{t^* + 1 \ldots, k_m + T\}$.*

Proof. In period $t^* - 1$, for each $j \in N_i$, either $P_j(\bar{h}_j^{t^*-1}) = \mathcal{A}$ or $P_j(\bar{h}_j^{t^*-1}) = \mathcal{B}(T - s(t^* - 1))$ by Lemma 3.A.2. Since $\bar{a}_i^{t^*} \neq C_{is(t^*)}$, in both cases $P_j(\bar{h}_j^{t^*}) = \mathcal{B}(T - s(t^*))$. Since $\bar{a}_i^{t^*} \neq C_{is(t^*)}$, Lemma 3.B.2 implies that $d_j(\bar{h}_j^{k_m+nT_*}) = 0$, and so $P_j(\bar{h}_j^{k_m+nT_*}) = \mathcal{B}(T - nT_*)$. Thus, for all $t \in \{t^* + 1 \ldots, k_m + T\}$, $P_j(\bar{h}_j^{t-1}) = \mathcal{B}(T - s(t-1))$ or $P_j(\bar{h}_j^{t-1}) = \mathcal{B}(2T - s(t-1))$, depending on whether $d_{2i}(\bar{h}_j^{t-1}) = 1$, but in either case for all $j \in N_i$, $\bar{a}_j^t = D$ for all $t \in \{t^* + 1 \ldots, k_m + T\}$. □

Lemma 3.B.5 says that if each player other than i has been playing according to $\zeta_{N \setminus \{i\}}$ for a sufficiently long time, and i plays an action other than $C_{is(t)}$ in some period t in the first part of a length T block, then plays $C_{is(t)}$ in every period t in the first part of the next length T block, then for every period t starting from the beginning of the second part of that length T block, each of of i's neighbours, $j \in N_i$, will play $C_{js(t)}$ until i plays something other than $C_{is(t)}$.

Lemma 3.B.5. *For any $k_m \geq z + 2T - s(z)$ such that $s(k_m) = T$, if $\bar{a}_i^t \neq C_{is(t)}$ for some $t \in \{k_m - T + 1, k_m - T + nT_*\}$ and $\bar{a}_i^t = C_{is(t)}$ for all $t \in \{k_m + 1, \ldots, t^* - 1\}$, where $t^* > k_m + nT_*$, then for all $j \in N_i$, $\bar{a}_j^t = C_{js(t)}$ for all $t \in \{k_m + nT_* + 1 \ldots, t^*\}$.*

Proof. Take an arbitrary $j \in N_i$. Let M be the maximal connected component of $N_C \setminus \{i\}$ containing j. As in the proof of Lemma 3.5.1, replacing N_C with $N_C \setminus \{i\}$, if for all $k \in M$, $P_k(\bar{h}_i^{k_m}) = \mathcal{A}$, then for all $k \in M$, $P_k(\bar{h}_i^t) = \mathcal{A}$, for all $t \in \{k_m, \ldots, t^*\}$.

If for some $k \in M$, $P_k(\bar{h}_i^{k_m}) = \mathcal{B}(T)$, then after period $k_m + (n-2)T_*$, $P_j(\bar{h}_i^{k_m+(n-2)T_*}) = \mathcal{B}(T - (n-2)T_*)$. This means that $\bar{a}_j^t \neq C_{js(t)}$ for some $t \in \{k_m + (n-2)T_* + 1, \ldots, k_m + (n-1)T_*\}$, and therefore j satisfies condition (b) in the definition of d_{1j}. Note that i satisfies parts (i) and (ii) of condition (c) in the definition of d_{1j} after history $\bar{h}_j^{k_m+nT_*}$. Since, $s(k_m + nT_*) = nT_*$, condition (a) of the definition of d_{1j} is also satisfied, and by Lemma 3.B.2, for all $j \in N_i$, $P_j(\bar{h}_j^{k_m+nT_*}) = \mathcal{A}^i$.

□

Lemma 3.B.6 says that if each player other than i has been playing according to $\zeta_{N \setminus \{i\}}$ for a sufficiently long time, and i plays $C_{is(t)}$ in every period t in the first part of a length T block, and then plays an action other than $C_{is(t)}$ in some period t in the second part of a length T block, then all of i's neighbours will punish i with D until the end of the following length T block.

Lemma 3.B.6. *For any $k_m \geq z + 2T - s(z)$ such that $s(k_m) = T$, if $\bar{a}_i^t = C_{is(t)}$ for all $t \in \{k_m + 1, \ldots, t^* - 1\}$ and $\bar{a}_i^{t^*} \neq C_{is(t^*)}$, where $t^* \in \{k_m + nT_* + 1, \ldots, k_m + T\}$, then for all $j \in N_i$, $\bar{a}_j^t = D$ for all $t \in \{t^* + 1, \ldots, k_m + 2T\}$.*

Proof. Take an arbitrary $j \in N_i$. If $P_j(\bar{h}_j^{t^*-1}) = \mathcal{A}$ or $P_j(\bar{h}_j^{t^*-1}) = \mathcal{A}^i$, then $P_j(\bar{h}_j^{t^*}) = \mathcal{B}(2T - s(t))$ (note that by Lemma 3.B.2, $P_j(\bar{h}_j^{t^*-1}) \neq \mathcal{A}^k$ for $k \in N_j \setminus \{i\}$) and $\bar{a}_j^t = D$ for all $t \in \{t^* + 1, \ldots, k_m + 2T\}$, unless $d_{1j}(\bar{h}_j^{k_m+T+nT_*}) = 1$. Note that i cannot satisfy (i) and (ii) of condition (c) in the definition of d_{1j} because $\bar{a}_i^t = C_{is(t)}$ for all $t \in \{k_m + 1, \ldots, k_m + nT_*\}$, so by Lemma 3.B.2, $d_{1j}(\bar{h}_j^{k_m+T+nT_*}) = 0$.

If $P_j(\bar{h}_j^{t^*-1}) = \mathcal{B}(p)$, $p > 0$, then we immediately have $\bar{a}_j^t = D$ for all $t \in \{t^* + 1, \ldots, k_m + T\}$. Note, however, that i satisfies (i) and (ii) of condition (b) in the definition of d_{2j} after history $\bar{h}_j^{k_m+T}$ and $s(k_m + T) = T$, so $P_j(\bar{h}_j^{k_m+T}) = \mathcal{B}(T)$. Thus $\bar{a}_j^t = D$ for all $t \in \{t^* + 1, \ldots, k_m + 2T\}$, since $d_{1j}(\bar{h}_j^{k_m+T+nT_*}) = 0$ by the same argument as in the previous paragraph. □

For Lemmas 3.B.7 and 3.B.8, assume that T is chosen such that $Tv_i^* > (n+1)T_*M$.

Lemma 3.B.7. *Take some $\Pi(k_m - z)$, $k_m \geq z + 2T - s(z)$, such that $s(k_m) = T$, and $P_j(\bar{h}_j^{k_m}) = \mathcal{A}$ for all $j \in N_C \setminus \{i\}$. For either $k_{m+1} = k_m + T$ or $k_{m+1} = k_m + 2T$, it is either the case that:*

1. $\Pi(k_{m+1} - z) \leq \Pi(k_m - z)$

2. $P_j(\bar{h}_j^{k_{m+1}}) = \mathcal{A}$ *for all* $j \in N_C \setminus \{i\}$

or

1. $\Pi(k_{m+1} - z) \leq \Pi(k_m - z) - (T - nT_*)v_i^* + T_*M$

Proof. Assume that $k_m \geq z + 2T - s(z)$, $s(k_m) = T$, and $P_j(\bar{h}_j^{k_m}) = \mathcal{A}$ for all $j \in N_C \setminus \{i\}$. We need to show that for either $k_{m+1} = k_m + T$ or $k_{m+1} = k_m + 2T$, it is the case that either $\Pi(k_{m+1} - z) \leq \Pi(k_m - z)$ and $P_j(\bar{h}_j^{k_m}) = \mathcal{A}$ for all $j \in N_C \setminus \{i\}$, or $\Pi(k_{m+1} - z) \leq \Pi(k_m - z) - (T - nT_*)v_i^* + T_*M$.

Case 1: $\bar{a}_i^t = C_{is(t)}$ for all $t \in \{k_m + 1, \ldots, k_m + T\}$. Note that in this case, $(\bar{a}_i^t, \bar{a}_{N_i}^t) = (C_{is(t)}, C_{N_i s(t)})$ for all $t \in \{k_m + 1, \ldots, k_m + T\}$. Therefore, there will be no change in $\Pi(K)$. By Lemma 3.B.1, $P_j(\bar{h}_j^{k_m+T}) = \mathcal{A}$ for all $j \in N \setminus \{i\}$. Thus, $k_{m+1} = k_m + T$ will do.

Case 2: $\bar{a}_i^t = C_{is(t)}$ for all $t \in \{k_m + 1, \ldots, k_m + nT_*\}$, $\bar{a}_i^t \neq C_{is(t)}$ for some $t \in \{k_m + nT_* + 1, k_m + T\}$. Let t^* be the first $t \in \{k_m + nT_* + 1, k_m + T\}$ such that $\bar{a}_i^t \neq C_{is(t)}$. There will be no change in $\Pi(K)$ over the periods in $\{k_m + 1, \ldots, t^* - 1\}$. By Lemma 3.B.6, all of i's neighbours will play D in every period $t \in \{t^* + 1, k_m + 2T\}$. In the length T_* block containing t^*, $\Pi(K)$ can can go up by at most T_*M. By Lemma 3.B.3, $\Pi(K)$ must go down by at least Tv_i^* over the periods in $\{k_m + T + 1, k_m + 2T\}$. So as long as $Tv_i^* > T_*M$, $k_{m+1} = k_m + 2T$ will do.

Case 3: $\bar{a}_i^t \neq C_{is(t)}$ for some $t \in \{k_m + 1, \ldots, k_m + nT_*\}$. $\Pi(K)$ can go up by at most T_*M over the periods in $\{k_m + 1, \ldots, k_m + nT_*\}$. Lemma 3.B.4 and Lemma 3.B.3 imply that $\Pi(K)$ must go down by $(T - nT_*)v_i^*$ over the periods in $\{k_m + nT_* + 1, \ldots, k_m + T\}$. So as long as $(T - nT_*)v_i^* > T_*M$, $k_{m+1} = k_m + T$ will do. □

Lemma 3.B.8. *Take some* $\Pi(k_m - z)$, $k_m \geq z + 2T - s(z)$, *such that* $s(k_m) = T$, *and* $P_j(\bar{h}_j^{k_m}) = \mathcal{B}(T)$ *for some* $j \in N_C \setminus \{i\}$. *For either* $k_{m+1} = k_m + T$ *or* $k_{m+1} = k_m + 2T$, *it is either the case that:*

1. $\Pi(k_{m+1} - z) \leq \Pi(k_m - z)$

or

1. $\Pi(k_{m+1} - z) \leq \Pi(k_m - z) + nT_*M$

2. $P_j(\bar{h}_j^{k_{m+1}}) = \mathcal{A}$ *for all* $j \in N_C \setminus \{i\}$

Proof. Assume that $k_m \geq z + 2T - s(z)$, $s(k_m) = T$, and $P_j(\bar{h}_j^{k_m}) = \mathcal{B}(T)$ for some $j \in N_C \setminus \{i\}$. We need to show that for either $k_{m+1} = k_m + T$ or $k_{m+1} = k_m + 2T$, it is the case that either $\Pi(k_{m+1} - z) \leq \Pi(k_m - z)$, or $\Pi(k_{m+1} - z) \leq \Pi(k_m - z) + nT_*M$ and $P_j(\bar{h}_j^{k_{m+1}}) = \mathcal{A}$ for all $j \in N_C \setminus \{i\}$.

Note that if $\bar{a}_i^t = C_{is(t)}$ for all $t \in \{k_m - T + 1, \ldots, k_m\}$, then by Lemma 3.B.1 $P_j(\bar{h}_j^{k_m}) = \mathcal{A}$ for all $j \in N_C \setminus \{i\}$. Thus, we only have to consider what happens when $\bar{a}_i^t \neq C_{is(t)}$ for some $t \in \{k_m - T + 1, \ldots, k_m\}$.

First, suppose that $\bar{a}_i^t = C_{is(t)}$ for all $t \in \{k_m - T + 1, \ldots, k_m - T + nT_*\}$, but $\bar{a}_i^t \neq C_{is(t)}$ for some $t \in \{k_m - T + nT_* + 1, \ldots, k_m\}$. Thus, $P_i(\bar{h}_i^{k_m}) = \mathcal{B}(T)$ for each $j \in N_i$, and all of i's neighbours will play D in every period $t \in \{k_m + 1, \ldots, k_m + T\}$, as long as for all $j \in N_i \cap N_C$, $d_{1j}(\bar{h}_j^{k_m+nT_*}) = 0$. Take any $j \in N_i \cap N_C$. Note that $\bar{a}_i^t = C_{is(t)}$ for all $t \in \{k_m - T + 1, \ldots, k_m - T + nT_*\}$, and so i does not satisfy (i) and (ii) of condition (c) in the definition of d_{1j}. Then Lemma 3.B.2 implies that $d_{1j}(\bar{h}_j^{k_m+nT_*}) = 0$.

By Lemma 3.B.3, $\Pi(K)$ must fall by at least Tv_i^* over the periods in $\{k_m + 1, \ldots, k_m + T\}$, and so $k_{m+1} = k_m + T$ will do.

Now suppose that $\bar{a}_i^t \neq C_{is(t)}$ for some $t \in \{k_m - T + 1, \ldots, k_m - T + nT_*\}$. There are three cases.

Case 1: $\bar{a}_i^t = C_{is(t)}$ for all $t \in \{k_m + 1, \ldots, k_m + T\}$. By Lemma 3.B.5, each $j \in N_i$ will play $C_{js(t)}$ for each $t \in \{k_m + nT_* + 1, \ldots, k_m + T\}$. Note that if every $j \in N_i \cup \{i\}$ plays $C_{js(t)}$ in period t, the value of $\Pi(K)$ does not change in that period. Thus, $\Pi(K)$ cannot increase over the periods in $\{k_m + nT_* + 1, \ldots, k_m + T\}$. The most $\Pi(K)$ can go up by over the periods in $\{k_m + 1, \ldots, k_m + nT_*\}$ is nT_*M. By Lemma 3.B.1, $P_j(\bar{h}_j^{k_m+T}) = \mathcal{A}$ for all $j \in N_C \setminus \{i\}$. Thus, $k_{m+1} = k_m + T$ will do.

Case 2: $\bar{a}_i^t = C_{is(t)}$ for all $t \in \{k_m+1, \ldots, k_m+nT_*\}$, $\bar{a}_i^t \neq C_{is(t)}$ for some $t \in \{k_m+nT_*+1, \ldots, k_m+T\}$. Let t^* be the first $t \in \{k_m + nT_* + 1, \ldots, k_m + T\}$ such that $\bar{a}_i^t \neq C_{is(t)}$. By Lemma 3.B.5, each $j \in N_i$ will play $C_{js(t)}$ for each $t \in \{k_m + nT_* + 1, \ldots, t^*\}$. By Lemma 3.B.6, all of i's neighbours will play D in every period $t \in \{t^* + 1, \ldots, k_m + 2T\}$. This implies that $\Pi(K)$ can increase by at most $(n + 1)T_*M$ over the periods in $\{k_m + 1, \ldots, k_m + T\}$. By Lemma 3.B.3, $\Pi(K)$ must fall by at least Tv^* over the periods in $\{k_m + T + 1, \ldots, k_m + 2T\}$. Thus, $k_{m+1} = k_m + 2T$ will do

Case 3: $\bar{a}_i^t \neq C_{is(t)}$ for some $t \in \{k_m + 1, \ldots, k_m + nT_*\}$. The maximum $\Pi(K)$ can go up by over the periods $t \in \{k_m+1, \ldots, k_m+nT_*\}$ is nT_*M, but Lemma 3.B.4 implies that every $j \in N_i$ will play D in every period in $\{k_m+nT_*+1, \ldots, k_m+T\}$. By Lemma 3.B.3, $\Pi(K)$ must go down by $(T-nT_*)v_i^*$ over the periods in $\{k_m + nT_* + 1, \ldots, k_m + T\}$. For sufficiently large T, $(T - nT_*)v_i^* > nT_*M$, and so $k_{m+1} = k_m + T$ will do. □

Proof of Lemma 3.6.1. First, note that for any $K \in \{1, \ldots, 2T - s(z)\}$, $\Pi(K)$ can be at most $2TM$. Let $k_0 = z + 2T - s(z)$. Lemmas 3.B.8 and 3.B.7 imply that we can find a sequence $(k_1, k_2, \ldots)$, $k_m < k_{m+1} \leq k_m + 2T$ such that $\Pi(k_m - z) < 4TM$ for all $m \geq 0$ (for sufficiently large T). Since the maximum amount $\Pi(K)$ can change in $2T$ periods is $2TM$, this implies that $6TM$ is an upper bound for $\Pi(K)$ for all $K \geq 1$. □

Appendix 3.C Complete payoffs and proof for Example 3.4.1

$v(D,D,D,D) = (0,0,0,0)$

$v(C,D,D,D) = (-\varepsilon,0,0,0)$

$v(D,C,D,D) = (0,-\varepsilon,0,0)$

$v(D,D,C,D) = (0,2,-\varepsilon,0)$

$v(D,D,D,C) = (0,0,0,-\varepsilon)$

$v(C,C,D,D) = (1,0,0,0)$

$v(D,C,C,D) = (1,1,0,0)$

$v(D,D,C,C) = (0,2,1,1)$

$v(C,D,C,D) = (-1,2,-\varepsilon,0)$

$v(D,C,D,C) = (0,-\varepsilon,0,-\varepsilon)$

$v(C,D,D,C) = (-\varepsilon,0,0,-\varepsilon)$

$v(C,C,C,D) = (1,1,0,0)$

$v(C,C,C,C) = (1,1,1,1)$

$v(D,C,C,C) = (1,1,1,1)$

$v(C,C,D,C) = (1,0,0,-\varepsilon)$

$v(C,D,C,C) = (-1,2,1,1)$

Proposition 3.C.1. *If $\delta < 1$, there exists a stage game satisfying Assumption 1, a network structure, and a payoff $v^* \in F^*$ such that v^* is not the payoff in any sequential equilibrium.*

Proof. Suppose for a contradiction that there exists a sequential equilibrium σ_N such that $\mathcal{V}_i(\sigma_N) = 1$ for all i. This requires that for every on-path history h_i, $\sigma_i(h_i) = C$. As long as Player 3 and Player 4 have not deviated, Player 3 can guarantee the payoff 1 by always playing C. Now consider the one shot deviation where Player 2 plays D. For this not to be profitable, Player 3 must play D at least once following the deviation. But if Player 3 plays D, her payoff is strictly less than 1, which cannot be optimal. Thus, the one shot deviation must be profitable, contradicting the assumption that σ_N is a sequential equilibrium. □

Bibliography

Aliprantis, Charalambos D. and Kim C. Border. 2006. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. 3 ed. Springer-Verlag.

Ben-Porath, Elchanan and Michael Kahneman. 1996. "Communication in Repeated Games with Private Monitoring." *Journal of Economic Theory* 70(2):281–297.

Bergemann, Dick, Benjamin Brooks and Stephen Morris. 2017. "First Price Auctions with General Information Structures: Implications for Bidding and Revenue." *Econometrica* 85(1):107–143.

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

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

Bose, Subir, Emre Ozdenoren and Andreas Pape. 2006. "Optimal auctions with ambiguity." *Theoretical Economics* 1(4):411–438.

Bose, Subir and Ludovic Renou. 2014. "Mechanism Design with Ambiguous Communication Devices." *Econometrica* 82(5):1853–1872.

Chen, Y., P. Katušcák and E. Ozdenoren. 2007. "Sealed bid auctions with ambiguity: Theory and experiments." *Journal of Economic Theory* 136(1):513–535.

Cho, Myeonghwan. 2011. "Public randomization in the repeated prisoner's dilemma game with local interaction." *Economics Letters* 112(3):280–282.

Cho, Myeonghwan. 2014. "Cooperation in the repeated prisoner's dilemma game with local interaction and local communication." *International Journal of Economic Theory* 10(3):235–262.

Church, K.W and P. Hanks. 1991. "Word association norms, mutual information, and lexicography." *Computational Linguistics* 16(1):22–29.

Crémer, J. and R. P. McLean. 1988. "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions." *Econometrica* 56(6):1247–1257.

Ellison, Glenn. 1994. "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching." *Review of Economic Studies* 61(3):567–88.

Enke, B. and F. Zimmermann. 2017. "Correlation neglect in belief formation." *Review of Economic Studies* 86(1):313–332.

Eyster, E. and G. Weizsäcker. 2011. "Correlation Neglect in Financial Decision-Making." Discussion Papers (1104). DIW Berlin.

Fudenberg, Drew and Eric Maskin. 1991. "On the dispensability of public randomization in discounted repeated games." *Journal of Economic Theory* 53(2):428–438.

Gilboa, I. and D. Schmeidler. 1989. "Maxmin expected utility with non-unique prior." *Journal of Mathematical Economics* 18(2):141–153.

Glaeser, E. and Cass R. Sunstein. 2009. "Extremism and social learning." *Journal of Legal Analysis* 1(1):262–324.

Hanany, E., P. Klibanoff and S. Mukerji. Forthcoming. "Incomplete Information Games with Ambiguity Averse Players." *American Economic Journal: Microeconomics* .

Heifetz, Aviad and Zvika Neeman. 2006. "On the Generic (Im)Possibility of Full Surplus Extraction in Mechanism Design." *Econometrica* 74(1):213–233.

Kallir, I. and D. Sonsino. 2009. "The Perception of Correlation in Investment Decisions." *Southern Economic Journal* 75(4):1045–66.

Kandori, Michihiro. 1992. "Social Norms and Community Enforcement." *Review of Economic Studies* 59(1):63–80.

Laclau, Marie. 2012. "A folk theorem for repeated games played on a network." *Games and Economic Behavior* 76(2):711–737.

Laclau, Marie. 2014. "Communication in repeated network games with imperfect monitoring." *Games and Economic Behavior* 87:136–160.

Levy, Gilat and Ronny Razin. 2015*a*. "Correlation neglect, voting behaviour and information aggregation." *American Economic Review* 105(4):1634–1645.

Levy, Gilat and Ronny Razin. 2015*b*. "Does polarisation of opinions lead to polarisation of platforms? The case of correlation neglect." *Quarterly Journal of Political Science* 10(3):321–355.

Levy, Gilat and Ronny Razin. 2018. "Combining Forecasts in the Presence of Ambiguity over Correlation Structures." Working paper.

Lo, K. C. 1998. "Sealed bid auctions with uncertainty averse bidders." *Economic Theory* 12(1):1–20.

Nava, Francesco. 2016. Repeated Games and Networks. In *Oxford Handbook on the Economics of Networks*, ed. Yann Bramoulle, Brian Rogers and Andrea Galleotti. Oxford University Press.

Nava, Francesco and Michele Piccione. 2014. "Efficiency in repeated games with local interaction and uncertain local monitoring." *Theoretical Economics* 9(1):279–312.

Obara, Ichiro. 2008. "The Full Surplus Extraction Theorem with Hidden Actions." *B.E. Journal of Theoretical Economics* 8(1):1935–1704.

Ortoleva, P. and E. Snowberg. 2015. "Overconfidence in political economy." *American Economic Review* 105(2):504–535.

Renault, Jerome and Tristan Tomala. 1998. "Repeated Proximity Games." *International Journal of Game Theory* 27(4):539–559.

Renou, L. 2015. "Rent Extraction and Uncertainty." Working paper.

Salo, A. S. and M. Weber. 1995. "Ambiguity aversion in first-price sealed-bid auctions." *Journal of Risk and Uncertainty* 11(2):123–137.

Schucany, W., W.C. Parr and J.E. Boyer. 1978. "Correlation structure in Farlie-Gumbel-Morgenstern distributions." *Biometrika* 65(3):650–653.

Tomala, Tristan. 2011. "Fault Reporting in Partially Known Networks and Folk Theorems." *Operations Research* 59(3):754–763.

Wolitzky, A. 2016. "Mechanism design with maxmin agents: Theory and an application to bilateral trade." *Theoretical Economics* 11(3):971–1004.

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