Aggregation of expert opinions

Published version:

DOI:10.1016/j.geb.2008.02.010

(Article begins on next page)

Aggregation of Expert Opinions*

Dino Gerardi
Yale University

Richard McLean
Rutgers University

Andrew Postlewaite
University of Pennsylvania

February, 2008

Abstract

Conflicts of interest arise between a decision maker and agents who have information pertinent to the problem because of differences in their preferences over outcomes. We investigate how the decision maker can extract the information by distorting the decisions that will be taken. We show that only slight distortions will be necessary when agents' signals are sufficiently accurate or when the number of informed agents becomes large. We argue that the particular mechanisms analyzed are substantially less demanding informationally than those typically employed in implementation and virtual implementation. Further, the mechanisms are immune to manipulation by small groups of agents.

JEL Classification: C72, D71, D82.

*Gerardi: Department of Economics, Yale University (e-mail: donato.gerardi@yale.edu); McLean: Department of Economics, Rutgers University (email: rpmclean@rci.rutgers.edu); Postlewaite: Department of Economics, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6207 (e-mail: apostlew@econ.upenn.edu). We thank an Associate Editor and two anonymous referees of this journal as well as Roberto Serrano, Matt Jackson and seminar participants at LSE, Washington University and Yale University for helpful comments. Postlewaite thanks the National Science Foundation for financial support.

1. Introduction

Consider the problem that an army officer faces in deciding whether or not to send his troops into battle with the enemy. Optimally, his decision will depend on the size of the opposing forces. If the enemy forces are not too strong, he will prefer to engage them, but if they are sufficiently strong he prefers not. He does not know the strength of the enemy, but the various troops in the area have some information regarding the enemy's strength, albeit imperfect. The difficulty the commanding officer faces is that the preferences of the individuals who possess the information regarding whether or not to engage the enemy may be very different from his own preferences. Those with information may exaggerate the strength of the enemy in order to obtain additional resources, or perhaps to avoid engagement entirely. In the extreme, the preferences of those with information may be diametrically opposed to those of the decision maker. When this is the case, those with the information necessary for informed decision making may have a dominant strategy to misrepresent their information, precluding the possibility of nontrivial communication.

Even when there is a conflict between the preferences of the decision maker and the preferences of those who possess information, it may be possible to extract the information with more sophisticated elicitation schemes. Suppose for example that those field officers who report to the commander have highly accurate information regarding whether the enemy is strong or weak. The commander may employ the following scheme. Ask each field officer whether he thinks the enemy is strong or weak, and the action that that officer would most like taken. Then, with probability $1-\varepsilon$, the commander attacks if a majority of field officers report that the enemy is "weak", and does not attack if a majority reports that the enemy is "strong". With probability ε, the commander instead chooses a field officer at random for scrutiny and determines whether his assessment of the enemy strength is "consistent" with the other reports, that is, if the selected officer's report regarding enemy strength agrees with the assessment of a majority of all field officers. If it does, the commander chooses the action that the field officer reported as his first choice, and, if not, the commander chooses a random action.

Truthful reporting on the part of the field officers will be incentive compatible when the officers' signals regarding whether the enemy is strong or weak are highly (but not necessarily perfectly) accurate.[1] When the officers' signals are highly

[1] It should be noted that there may be equilibria in which the informed agents do not report their information truthfully in addition to the truthful reporting equilibrium. We discuss this

accurate and others report truthfully, each maximizes his chance of being in the majority by reporting truthfully, and thereby getting his first choice should the commander randomly choose him for scrutiny. By misreporting that the enemy is strong when in fact an officer has observed a "weak" enemy, an officer's chances of getting his first choice are reduced should he be scrutinized, but he may change the commander's decision in the event that the commander decides the action based on the majority report. However, the probability that any individual field officer will be pivotal goes to zero as the accuracy of the field officers' signals goes to 1. This type of mechanism exhibits an important feature: the commanding officer does not need to know the field officers' preferences which, if known, might provide some information regarding the direction in which an officer might wish to skew the decision.

The commander can thus extract the field officers' information, but the elicitation of the information comes at a cost. With probability ε the commander selects a field officer for scrutiny, and if that officer's announcement is consistent with the majority announcement the outcome will not necessarily be the commander's preferred choice.

The mechanism described above uses the correlation of the officers' signals that naturally arises from the fact that they are making assessments of the same attribute. We will formalize the ideas in the example and provide sufficient conditions under which experts' information can be extracted through small distortions of the decision maker's optimal rule. The basic idea can be seen in the example; when signals are very accurate, no single agent is likely to change the outcome by misreporting his information, hence, small "rewards" will be sufficient to induce truthful announcements. We further show that one can use this basic idea to show that, when the number of informed agents becomes large, one can extract the information at small cost even if each agent's information is *not* accurate. When the number of agents becomes large, the chance that an agent will be pivotal in the decision becomes small even if the signals that agents receive are of low accuracy. This is not enough to ensure that information can be extracted at low cost, since giving each agent a small chance of being a "dictator" might involve a large deviation from the decision maker's optimal rule. Using techniques from McLean and Postlewaite (2002, 2006) we show, however, that an agent's effect on the decision maker's posterior goes to zero faster than the number of agents goes to infinity. Consequently, as the number of agents becomes increasingly large, the decision maker can correspondingly reduce the distortion associated with the

in the last section.

need to scrutinize agents while still inducing the agents to truthfully reveal their private information.

We introduce the model in the next section, and present the results for the case of a finite number of experts with accurate signals in section 3. In section 4 we analyze the case with a large number of experts whose signals may not be accurate. In section 5 we discuss some extensions and further results. Section 6 contains the proofs.

1.1. Related Literature

Our notion of implementation is a weak form of virtual Bayesian implementation (Abreu and Matsushima (1992), Matsushima (1993), Duggan (1997), Serrano and Vohra (2005)), but differs in important ways. First, our focus is not on full implementation in the sense that we do not require that *all* equilibria implement a given social choice function. Second, in the literature on virtual implementation, two social choice functions are close if they specify very similar outcomes in *every* state of the world. This is a natural definition because the probability of every state may be bounded away from zero. In this paper we consider environments in which the probability of some states is vanishing. It is therefore reasonable to use a notion of approximation that requires that two functions are close with arbitrarily high probability. The differences in the two approaches have important ramifications regarding the classes of functions that can be implemented. To be virtually implementable a function must be incentive compatible and satisfy some additional conditions such as measurability (Abreu and Matsushima (1992)), incentive consistency (Duggan (1997)) or virtual monotonicity (Serrano and Vohra (2005)).[2] In contrast, our notion of implementation does not require any of these conditions.

There is an extensive literature on information transmission between informed experts and an uninformed decision maker. The classic reference is Crawford and Sobel (1982) who assume that the decision maker faces a single expert. The literature has also analyzed the case of multiple experts. Of course, if there are at least three experts and they are all perfectly informed (i.e., they possess the same information) the problem of eliciting their information is trivial. The case in which there are two perfectly informed experts has been analyzed by Gilligan and Krehbiel (1989), Krishna and Morgan (2001), and Battaglini (2002).

[2] In some cases, the environment has also to satisfy certain additional assumptions. For example, Matsushima (1993) assumes that side payments are allowed.

Austen-Smith (1993) is the first paper to focus on imperfectly informed experts. Austen-Smith assumes that the decision maker gets advice from two biased experts whose signals about the state are conditionally independent. That paper compares two different communication structures: simultaneous reporting and sequential reporting. Battaglini (2004) extends the analysis to the case in which the number of experts is arbitrary and both the state and the signals are multidimensional. Battaglini exploits the fact that the experts' preferences are different and commonly known and constructs an equilibrium in which every expert truthfully announces (a part of) his signal. If the experts' signals are very accurate or if the number of experts is sufficiently large, the corresponding equilibrium outcome is close to the decision maker's first best. In contrast to Battaglini (2004), we do not impose any restriction on the experts' preferences and, importantly, they can be private information. Furthermore, we provide conditions under which any social choice rule can be approximately implemented.

Wolinsky (2002) analyzes the problem of a decision maker who tries to elicit as much information as possible from a number of experts. The experts share the same preferences which differ from those of the decision maker. The information structure in Wolinsky (2002) is significantly different from ours. In particular, there is no state of the world and the experts' types are independently distributed. Wolinsky first assumes that the decision maker can commit to a choice rule and characterizes the optimal mechanism. He then relaxes the assumption of perfect commitment and shows that it is beneficial for the decision maker to divide the experts in small groups and ask them to send joint reports.

In our paper, as well as in all the articles mentioned above, the experts are affected by the decision maker's choice. One strand of the literature has also studied the case in which the experts are concerned with their reputation for being well informed. Ottaviani and Sørensen (2006a, 2006b) consider a model in which the experts receive a noisy signal about the state of the world and the quality of their information is unknown. Each expert's reputation is updated on the basis on their messages and the realized state. Ottaviani and Sørensen show that the experts generally do not reveal their information truthfully.

Our paper is also related to the recent literature on strategic voting. Feddersen and Pesendorfer (1997, 1998) consider two-candidate elections with privately informed voters. They show that under non-unanimous voting rules, large elections fully aggregate the available information in the sense that the winner is the candidate that would be chosen if all private information were publicly available. This implies that under majority rule, for example, a social planner can implement the outcome preferred by the majority of the voters. In contrast, our asymptotic results show that if the planner has the ability to commit to a mechanism, then he can approximately implement *almost any* social choice rule.

We postpone to section 5 a discussion of the relationship of our results to the notion of informational size introduced in McLean and Postlewaite (2002).

2. The Model

2.1. Information

We will consider a model with $n \geq 3$ experts. If r is a positive integer, let $J_r = \{1, \ldots, r\}$. Experts are in possession of private information of two kinds. First, each expert observes a signal that is correlated with the true but unobservable state of nature θ. The state can directly affect his payoff but the signal does not. Second, each expert knows his own "personal characteristic" which parametrizes his payoff function but has no effect on his beliefs regarding the state. More formally, let $\Theta = \{\theta_1, \ldots, \theta_m\}$ denote the finite set of states of nature, let S_i denote the finite set of possible signals that expert i can receive and let Q_i denote the expert's (not necessarily finite) set of personal characteristics. The set of types of expert i in this setup is therefore $S_i \times Q_i$. Let $S \equiv S_1 \times \cdots \times S_n$ and $S_{-i} \equiv \times_{j \neq i} S_j$. The product sets Q and Q_{-i} are defined in a similar fashion. Let Δ_X denote the set of probability measures on a set X. Let $\delta_x \in \Delta_X$ denote the Dirac measure concentrated on $x \in X$. Each probability measure $P \in \Delta^*_{\Theta \times S}$ is the distribution of an $(n+1)$-dimensional random vector $(\widetilde{\theta}, \tilde{s})$ taking values in $\Theta \times S$ whose dependence on P will be suppressed. Let $\Delta^*_{\Theta \times S}$ denote the subset of $\Delta_{\Theta \times S}$ satisfying the following support conditions:

$$P(\theta) = \text{Prob}\{\widetilde{\theta} = \theta\} > 0 \text{ for each } \theta \in \Theta$$

and

$$P(s) = \text{Prob}\{\widetilde{s}_1 = s_1, \ldots, \widetilde{s}_n = s_n\} > 0 \text{ for each } s \in S.$$

For each $P \in \Delta^*_{\Theta \times S}$ and $s \in S$, let $h(s) = P_\Theta(\cdot|s)$ denote the associated conditional probability on Θ.

In addition, we will make the following conditional independence assumption[3]:

[3] The conditional independence assumption simplifies the presentation but our results will hold under more general circumstances (see section 5 for details).

for each $k \in \{1, \ldots, m\}$ and each $(s_1, \ldots, s_n) \in S$,

$$\text{Prob}\{\tilde{s}_1 = s_1, \ldots, \tilde{s}_n = s_n | \tilde{\theta} = \theta_k\} = \prod_{i \in J_n} P_i(s_i | \theta_k)$$

where

$$P_i(s_i | \theta_k) = \text{Prob}\{\tilde{s}_i = s_i | \tilde{\theta} = \theta_k\}.$$

Let $\Delta^{CI}_{\Theta \times S}$ denote the measures in $\Delta^*_{\Theta \times S}$ satisfying the conditional independence assumption.

The probabilistic relationship between states, signals and characteristics will be defined by a product probability measure $P \otimes \hat{P} \in \Delta_{\Theta \times S \times Q}$ where $P \in \Delta^{CI}_{\Theta \times S}$ and $\hat{P} \in \Delta_Q$. This is a stochastic independence assumption: if $P \otimes \hat{P}$ is the distribution of a $(2n + 1)$-dimensional random vector $(\widetilde{\theta}, \tilde{s}, \tilde{q})$ taking values in $\Theta \times S \times Q$, then

$$\text{Prob}\{\widetilde{\theta} = \theta, \widetilde{s}_1 = s_1, \ldots, \widetilde{s}_n = s_n, (\widetilde{q}_1, \ldots, \widetilde{q}_n) \in C\} = P(\theta, s)\hat{P}(C).$$

for each $(\theta, s_1, \ldots, s_n) \in \Theta \times S$ and each event $C \subseteq Q$.

2.2. The Decision Maker

In addition to the n experts, our model includes a decision maker, or social planner, who is interested in choosing an action a from a finite set of social alternatives A with $|A| = N$. The behavior of the decision maker is described by a function

$$\pi : \Delta_\Theta \to \Delta_A.$$

Loosely speaking, we interpret the function π as a "reduced form" description of the decision maker's behavior: if the probability measure $\rho \in \Delta_\Theta$ represents the decision maker's "beliefs" regarding the state of nature, then the decision maker chooses an action from the set A according to the probability measure $\pi(\cdot|\rho) \in \Delta_A$. For example, suppose that $(a, \theta) \mapsto g(a, \theta)$ is a function describing the payoff to the decision maker if he takes action a and the state is θ. For each vector of beliefs $\rho \in \Delta_\Theta$, we could naturally define $\pi(\cdot|\rho) \in \Delta_A$ so that

$$\pi(a|\rho) > 0 \Rightarrow a \in \arg\max_{a \in A} \sum_{\theta \in \Theta} g(a, \theta)\rho(\theta).$$

Other examples are clearly possible and our reduced form description can accommodate all of these. In particular, suppose that the social choice is made by a committee of individuals with heterogeneous preferences. The committee elicits the information from the experts and then makes a final decision using a certain voting rule. In this case, the function π represents the outcome of the voting.

As we have described above, both s_i and q_i are the private information of expert i. The decision maker cannot observe the experts' characteristics q_i or their signals s_i. Given the function π, the decision maker would like to choose an action using the best available information regarding the state θ. Since the decision maker himself receives no information regarding the state θ, he must ask the experts to report their signals. If $s \in S$ is the experts' reported signal profile, then the measure $h(s) = P_\Theta(\cdot|s)$ defines the decision maker's updated beliefs and he will then choose an action according to the probability measure $\pi(\cdot|h(s))$.

2.3. The Experts

The payoff of expert i depends on the action a chosen by the decision maker, the state of nature θ and the idiosyncratic parameter q_i. Formally, the payoff of expert i is defined by a function

$$u_i : A \times \Theta \times Q_i \to \mathbb{R}.$$

To prove our results, we will need the following definition.

Definition: Let K be a positive number. A function $u_i : A \times \Theta \times Q_i \to \mathbb{R}$ satisfies the *K-strict maximum condition* if

(i) For every $\theta \in \Theta$ and for every $q_i \in Q_i$, the mapping $a \in A \mapsto u_i(a, \theta, q_i)$ has a unique maximizer which we will denote $a_i^*(\theta, q_i)$.

(ii) For every i, for every $\theta \in \Theta$ and for every $q_i \in Q_i$,

$$u_i(a_i^*(\theta, q_i), \theta, q_i) - u_i(a, \theta, q_i) \geq K \text{ for all } a \neq a_i^*(\theta, q_i).$$

Note that (ii) is implied by (i) for some $K > 0$ when Q_i is finite.

2.4. Mechanisms

A mechanism is a mapping $(s, q) \in S \times Q \mapsto \mu(\cdot|s, q) \in \Delta_A$. If (s, q) is the announced profile of signals and characteristics, then $\mu(a|s, q)$ is the probability with which the decision maker chooses action $a \in A$. Obviously, a mechanism induces a game of incomplete information and the decision maker is concerned with the trade-off between the "performance" of the mechanism and its incentive properties. The performance of mechanism μ is measured[4] by

$$\sup_{q\in Q}\sum_{s\in S}||\pi(\cdot|h(s))-\mu(\cdot|s,q)||P(s) \tag{2.1}$$

where

$$||\pi(\cdot|h(s))-\mu(\cdot|s,q)||=\sum_{a\in A}|\pi(a|h(s))-\mu(a|s,q)|.$$

According to this performance criterion, a mechanism μ is "good" if the quantity in expression (2.1) is "small." For an alternative viewpoint of our performance criterion, note that

$$\sum_{s\in S}||\pi(\cdot|h(s))-\mu(\cdot|s,q)||P(s)\leq\varepsilon$$

implies that

$$\text{Prob}\{||\pi(\cdot|h(\tilde{s}))-\mu(\cdot|\tilde{s},q)||\leq\sqrt{\varepsilon}\}>1-\sqrt{\varepsilon}.$$

Thus, a good mechanism has the property that, for each profile $q\in Q$, $\pi(\cdot|h(s))$ and $\mu(\cdot|s,q)$ are close on a set of s profiles of high probability.[5]

Definition: A mechanism μ is *incentive compatible* if for each i, each $(s_i,q_i)\in S_i\times Q_i$ and each $(s_i',q_i')\in S_i\times Q_i$,

$$E_{\tilde{q}}\left[\sum_{s_{-i}\in S_{-i}}\sum_{\theta\in\Theta}\sum_{a\in A}\left[\mu(a|s_{-i},s_i,\tilde{q}_{-i},q_i)-\mu(a|s_{-i},s_i',\tilde{q}_{-i},q_i')\right]u_i(a,\theta,q_i)P(\theta,s_{-i}|s_i)\,|\tilde{q}_i=q_i\right]\geq 0.$$

The mechanisms that we analyze below actually satisfy a stronger notion of incentive compatibility: each expert i has an incentive to report his signal s_i truthfully and, conditional on truthful announcement of the experts' signals, it is a dominant strategy for expert i to announce his preference parameter q_i truthfully. We discuss this in the last section.

[4]Throughout the paper, $||\cdot||$ will denote the ℓ_1 norm and $||\cdot||_2$ will denote the ℓ_2 norm.

[5]We evaluate the performance of a mechanism from an *ex-ante* point of view. It is therefore possible that in some unlikely events the decision maker will implement an outcome which is rather different from the outcome specified by the function π.

3. Finitely Many Experts with Accurate Signals: The Jury Model

We discussed in the introduction the basic idea of how to construct a mechanism to extract experts' information. We next provide a slightly more detailed description of the mechanism. Agents report their private information, which consists of a signal about the state of the world and their utility function. With probability close to 1, the information is used to update the decision maker's beliefs and the desired decision is taken. With small probability an agent is chosen at random and "scrutinized" to see if his report about the state is the same as reported by the majority. The correlation in the agents' signals conditional on the state θ makes this a "statistical test" as to whether the agent reported truthfully, assuming other agents were reporting truthfully. When other agents are reporting truthfully, a given agent is more likely to "pass" this test by truthfully reporting his signal. We reward the agent if he passes the test by distorting the optimal decision rule slightly, and giving the scrutinized agent his optimal decision, contingent on his announced utility function and the announced signals of the other agents.

When agents' signals are very accurate, no single agent's report is likely to have a significant effect on the posterior distribution on Θ, and consequently there is little chance that his report will affect the outcome. Thus, when the signals are highly accurate, the distortion in the decision rule necessary to provide incentives for truthful revelation is small.

3.1. The Setup

In this section, the n experts are interpreted as "jurors" and the decision maker is interpreted as a "judge". Let $\Theta = \{\theta_0, \theta_1\}$ where $\theta_1 = 1$ corresponds to "guilty" and $\theta_0 = 0$ corresponds to "innocent." Suppose that the jurors receive a noisy signal of the state. In particular, let $S_i = \{\sigma_0, \sigma_1\}$ where $\sigma_1 =$ guilty and $\sigma_0 =$ innocent. Let s^0 and s^1 denote the special signal profiles $s^0 = (\sigma_0, \ldots, \sigma_0)$ and $s^1 = (\sigma_1, \ldots, \sigma_1)$. Let $A = \{0, 1\}$ where 0 corresponds to "acquit" and 1 corresponds to "convict."[6] The payoff function of juror i is a mapping $u_i : A \times \Theta \times Q_i \to \mathbb{R}$. We will assume that each Q_i is finite and that u_i satisfies the K_i-strict maximum condition for some $K_i > 0$. We can illustrate the strict maximum condition in an

[6] The results extend to the case in which there are several states, actions and signals in a straightforward manner.

example. For each $i \in J_n$, let

$$u_i(a, \theta, q_i) = a(\theta - q_i).$$

Hence, q_i may be interpreted as an idiosyncratic measure of the "discomfort" experienced by juror i whenever a defendant is convicted, irrespective of guilt or innocence. If $0 < q_i < 1$, then

$$-q_i < 0 < 1 - q_i,$$

so that conviction of a guilty party is preferred to acquittal, and acquittal is preferred to conviction of an innocent party. If each Q_i is a finite subset of $]0, 1[$, then it is easily verified that conditions (i) and (ii) are satisfied.

The behavior of the decision maker is described by a function $\pi : \Delta_\Theta \to \Delta_A$. If $\rho \in \Delta_\Theta$ represents the decision maker's beliefs regarding the defendant's guilt or innocence, then $\pi(1|\rho)$ is the probability of conviction and $\pi(0|\rho)$ is the probability of acquittal. Note that we make no assumptions regarding the properties of π; in particular, π need not be continuous.

For each $P \in \Delta^{CI}_{\Theta \times S}$, let

$$\kappa(P) := \min_{k \in \{0,1\}} \min_{i \in J_n} P_i(\sigma_k | \theta_k).$$

The number $\kappa(P)$ is a measure of the quality of the signals. When the state is θ_k, $k = 0, 1$, each agent observes the correct signal σ_k with probability at least $\kappa(P)$.

3.2. The Jury Result

Proposition 1: Choose $n \geq 3$. Let $K_1, \ldots, K_n$ be given as above, let $K = min_i K_i$ and suppose that $\varepsilon > 0$. There exists a $\overline{\kappa} \in]0, 1[$ (depending on ε and K) such that, for all $\hat{P} \in \Delta_Q$ and for all $P \in \Delta^{CI}_{\Theta \times S}$ satisfying $\kappa(P) > \overline{\kappa}$, there exists an incentive compatible mechanism μ satisfying

$$\max_{q \in Q} \sum_s \|\pi(\cdot|h(s)) - \mu(\cdot|s, q)\| \, P(s) < \varepsilon \ .$$

The proof of Proposition 1 appears in section 6, but we will construct the mechanism and present the idea of the argument here. For $k \in \{0, 1\}$, let

$$\nu_k(s) := \{i \in J_n | s_i = \sigma_k\}$$

denote the set of jurors who observe signal σ_k when the realized signal profile is s. Let C_0 and C_1 be two subsets of S defined as

$$C_0 = \{s \in S| \, |\nu_0(s)| > \frac{n}{2}\}$$

and

$$C_1 = \{s \in S| \, |\nu_1(s)| > \frac{n}{2}\}.$$

When a signal profile $s \in C_k$ is realized the majority of the jurors observe signal σ_k. Next, let

$$\rho_i(s) = \chi_{C_0}(s)\gamma(\theta_0|s_i) + \chi_{C_1}(s)\gamma(\theta_1|s_i)$$

where

$$\begin{aligned} \gamma(\theta_k|s_i) &= 1 \text{ if } s_i = \sigma_k \\ &= 0 \text{ if } s_i \neq \sigma_k. \end{aligned}$$

and χ_{C_k} denotes the indicator function of C_k. Note that $\rho_i(s) = 1$ if and only if s_i is a (strict) majority announcement for the profile s.

Define $\alpha_i(\cdot|s, q_i) \in \Delta_A$ where

$$\alpha_i(a_i^*(\theta_k, q_i)|s, q_i) = 1 \text{ if } s \in C_k \text{ and } k \in \{0, 1\}$$

and

$$\alpha_i(a|s, q_i) = \frac{1}{2} \text{ for each } a \in A \text{ if } s \notin C_0 \cup C_1.$$

Let

$$\begin{aligned} \psi(s) &= s^0 \text{ if } s = s^0 \text{ or } s = (s^0_{-i}, \sigma_1) \text{ for some } i \\ &= s^1 \text{ if } s = s^1 \text{ or } s = (s^1_{-i}, \sigma_0) \text{ for some } i \\ &= s \text{ otherwise.} \end{aligned}$$

Finally, choose $\lambda \in]0, 1[$ and define a mechanism μ as follows: for each $a \in A = \{0, 1\}$,

$$\mu(a|s, q) = (1 - \lambda)\pi(a|h(\psi(s))) + \frac{\lambda}{n}\sum_{j=1}^{n}\left[\rho_j(s)\alpha_j(a|s, q_j) + (1 - \rho_j(s))\frac{1}{2}\right].$$

To interpret the mechanism, suppose that the jurors announce the profile (s, q). With probability $1 - \lambda$, the decision maker will choose action a with probability $\pi(a|h(\psi(s)))$. With probability $\frac{\lambda}{n}$, one of the jurors will be randomly selected for "scrutiny." Suppose that juror i is chosen. If $\rho_i(s) = 0$, the decision maker randomizes uniformly over $a = 0$ and $a = 1$. If $\rho_i(s) = 1$, then a strict majority of the jurors have announced the same signal (either σ_0 or σ_1) and juror i is a member of this strict majority. The decision maker now "rewards" juror i for his majority announcement by choosing action $a_i^*(\theta_0, q_i)$ if $s \in C_0$ or $a_i^*(\theta_1, q_i)$ if $s \in C_1$. The mechanism is designed so that, when the jurors' signals are accurate, the juror who is chosen for scrutiny is rewarded when he truthfully announces his private information. It is important to note that, in our framework, any scheme to "reward" an agent typically must utilize all agents' information. Except in extreme cases, an agent's optimal choice depends nontrivially on his information.

To illustrate the idea of the proof, let μ be the mechanism defined above and suppose that agent i with characteristic q_i observes signal σ_1. To prove incentive compatibility, it suffices to show that, when $\kappa(P) \approx 1$,

$$\sum_{s_{-i} \in S_{-i}} \sum_{\theta \in \Theta} \sum_{a \in A} [\mu(a|s_{-i}, \sigma_1, q_{-i}, q_i) - \mu(a|s_{-i}, s_i', q_{-i}, q_i')] \, u_i(a, \theta, q_i) P(\theta|s_{-i}, \sigma_1) P(s_{-i}|\sigma_1) \geq 0$$

for each $q_{-i} \in Q_{-i}, q_i \in Q_i$, and $(s_i', q_i') \in S_i \times Q_i$. The argument relies very heavily on the fact that

$$P(\theta_1|s^1) \approx 1 \text{ and } P(s_{-i}^1|\sigma_1) \approx 1 \text{ when } \kappa(P) \approx 1.$$

We first claim that by truthfully announcing $s_i' = \sigma_1$, juror i cannot benefit from lying about his characteristic. To see this, first note that, when i truthfully announces his signal, a misreported characteristic has no effect on the action chosen by the decision maker when another juror is chosen for scrutiny. Indeed, a misreported characteristic can only affect the reward that the decision maker will choose for juror i if, after having been chosen for scrutiny, juror i receives a reward. If $\kappa(P) \approx 1$, then $P(\theta_1|s^1) \approx 1$ and $P(s_{-i}^1|\sigma_1) \approx 1$, i.e., juror i believes that, with very high probability, the true state is θ_1 and the other jurors have announced the signal profile s_{-i}^1. Now suppose that juror i is to be rewarded by the decision maker. If i announces (σ_1, q_i), the decision maker will, with high probability, choose the action $a_i^*(\theta_1, q_i)$ that is optimal for (θ_1, q_i). If he announces (σ_1, q_i'), then the decision maker will, with high probability, choose the action $a_i^*(\theta_1, q_i')$ that is optimal for q_i'. Assumptions (i) and (ii) guarantee that such a lie cannot be profitable if q_i is the true characteristic of juror i. We next show that, by misreporting $s_i' = \sigma_0$, juror i still cannot benefit from lying about his characteristic. Since $(s_{-i}^1, \sigma_0) \in C_1$ and $(s_{-i}^1, \sigma_1) \in C_1$, we conclude that neither i's misreported signal nor i's misreported type can have any effect on the action chosen by the decision maker when $s_{-i} = s_{-i}^1$ and juror i is not chosen for scrutiny. As in the previous case, $\kappa(P) \approx 1$ implies that $P(\theta_1|s^1) \approx 1$ and $P(s_{-i}^1|\sigma_1) \approx 1$ so, again, juror i is concerned with the consequences of his announcement when the state is θ_1 and the other jurors have announced the signal profile s_{-i}^1. If i announces (σ_1, q_i) and if i is chosen for scrutiny, then the decision maker will choose action $a_i^*(\theta_1, q_i)$ when $s_{-i} = s_{-i}^1$ since i agrees with the majority and the majority has announced σ_1. If i announces (σ_0, q_i') and if i is chosen for scrutiny, then the decision maker will randomize over 0 and 1 when $s_{-i} = s_{-i}^1$ since i does not agree with the majority. Again, conditions (i) and (ii) guarantee that i cannot benefit from lying.

Finally, we turn to the performance of the mechanism. For sufficiently small λ,

$$\sum_s \|\pi(\cdot|h(s)) - \mu(\cdot|s,q)\| \, P(s) \approx \sum_s \|\pi(\cdot|h(s)) - \pi(\cdot|h(\psi(s)))\| \, P(s)$$

for each $q \in Q$. Now s and $\psi(s)$ differ only when all but one juror have announced the same signal, an event of small probability when $\kappa(P) \approx 1$. Consequently,

$$\sum_s \|\pi(\cdot|h(s)) - \pi(\cdot|h(\psi(s)))\| \, P(s) \le 2 \sum_{i \in J_n} \left[P(s_{-i}^1, \sigma_0) + P(s_{-i}^0, \sigma_1) \right]$$

and, since

$$\sum_{i \in J_n} \left[P(s_{-i}^1, \sigma_0) + P(s_{-i}^0, \sigma_1) \right] \approx 0 \text{ when } \kappa(P) \approx 1,$$

we obtain the desired result.

4. The Case of Many Experts

We turn next to the case in which there is an increasing number of experts. Similar to the case with a fixed finite number of experts, the basic idea here plays off the fact that, as the number of agents increases, any single agent will have little chance of affecting the decision. As in the previous case, the correlation in the agents' signals allows us to construct a statistical test of the announcement submitted by a randomly chosen agent, which he will more likely pass by announcing truthfully if other agents are announcing truthfully. Again we reward an agent who passes the test by distorting the decision in favor of that agent's optimal decision with small probability. There is one new difficulty here, however. As the number of agents increases, the chance that any given agent is chosen for scrutiny (and the chance to get his more preferred decision) goes to zero as the number of agents increases. Therefore, the reward that provides the incentive for truthful revelation goes to zero. We show, however, that the basic idea can still be employed by showing that an agent's expected effect on the decision goes to zero more quickly than does the expected reward for truthful announcement.

4.1. The Setup

Throughout this section, $\Theta = \{\theta_1, .., \theta_m\}$ will denote a fixed finite state space and A a fixed finite action space with $|A| = N$. In addition, let T be a fixed finite set that will serve as the set of signals common to all experts (i.e., $S_i = T$ for all i) and let $\Delta^{**}_{\Theta\times T}$ denote the set of probability measures β on $\Theta \times T$ satisfying the following condition: for every $\theta, \hat{\theta}$ with $\theta \neq \hat{\theta}$, there exists $s \in T$ such that $\beta(s|\theta) \neq \beta(s|\hat{\theta})$. Next, let T^n denote the cartesian product of n copies of T.

Definition: Let K, L and M be positive numbers, let n be a positive integer and let $\beta \in \Delta^{**}_{\Theta\times T}$. An *aggregation problem* $\Pi(n, K, M, \beta, L)$ is a collection consisting of the following objects:

(i) For each $i \in J_n$, a space of characteristics Q^n_i and a probability measure $\hat{P}^n \in \Delta_{Q^n}$ ($Q^n := Q^n_1 \times \cdots \times Q^n_n$).

(ii) For each $i \in J_n$, a function

$$u^n_i : A \times \Theta \times Q^n_i \to \mathbb{R}$$

satisfying the K-strict maximum condition and bounded by M, i.e., $|u^n_i(a, \theta, q_i)| \leq M$ for each $(a, \theta, q_i) \in A \times \Theta \times Q^n_i$.

(iii) A probability measure $P^n \in \Delta^{CI}_{\Theta\times T^n}$ satisfying the following conditional independence condition: for each $(\theta, s_1, \dots, s_n) \in \Theta \times T^n$,

$$P^n(s_1, \dots, s_n, \theta) = \text{Prob}\{\widetilde{s}^n_1 = s_1, \widetilde{s}^n_2 = s_2, \dots, \widetilde{s}^n_n = s_n, \tilde{\theta} = \theta\} = \beta(\theta) \prod_{i=1}^{n} \beta(s_i|\theta).$$

(iv) Let

$$\hat{\gamma}_i(\theta|s_i, q_i^n) = \left[\max_{a\in A}\left[u_i^n(a,\theta,q_i^n)\right] - \frac{1}{N}\sum_a u_i^n(a,\theta,q_i^n)\right]\beta(\theta|s_i). \tag{4.1}$$

Then

$$\left|\left|\frac{\hat{\gamma}_i(\cdot|s_i', q_i'^n)}{||\hat{\gamma}_i(\cdot|s_i', q_i'^n)||_2} - \frac{\hat{\gamma}_i(\cdot|s_i, q_i^n)}{||\hat{\gamma}_i(\cdot|s_i, q_i^n)||_2}\right|\right|_2 \geq L$$

for all $i \in \{1, \ldots, n\}$ and all $(s_i, q_i^n), (s_i', q_i'^n) \in T \times Q_i^n$.

Condition (iii) is a conditional independence assumption. Condition (iv) is a nondegeneracy assumption that says that $\hat{\gamma}_i(\cdot|s_i', q_i'^n)$ is not a scalar multiple of $\hat{\gamma}_i(\cdot|s_i, q_i^n)$ and that the normalized vectors are "uniformly" bounded away from each other for all n. In the simple jury example in which (1) $\Theta = \{\theta_0, \theta_1\}$, (2) $S_i = \{\sigma_0, \sigma_1\}$, (3) for all n, $Q_i^n \subseteq C$ for some finite set $C \subseteq [0,1]$ and (4) for all n, $u_i^n(a,\theta,q_i) = a(\theta - q_i)$, condition (iv) is satisfied for some positive number L if and only if for each $(s_i, q_i), (s_i', q_i') \in T \times Q_i^n$ with $(s_i, q_i) \neq (s_i', q_i')$,

$$\frac{q_i\beta(\theta_0|s_i)}{(1-q_i)\beta(\theta_1|s_i)} \neq \frac{q_i'\beta(\theta_0|s_i')}{(1-q_i')\beta(\theta_1|s_i')}.$$

This condition is "generic" in the following sense: for each $\beta \in \Delta^{**}_{\Theta\times T}$, the nondegeneracy condition is satisfied for all $(q_1, \ldots, q_{|C|}) \in [0,1]^{|C|}$ outside a closed set of Lebesgue measure zero.

An aggregation problem corresponds to an instance of our general model in which the signal sets of the experts are identical and the stochastic structure exhibits symmetry.[7] We are assuming, as always, that the profile of experts characteristics is a realization of a random vector $\tilde{q}^n$ and that $(\tilde{\theta}, \tilde{s}^n)$ and $\tilde{q}^n$ are stochastically independent so that the joint distribution of the state, signals and characteristics is given by the product probability measure $P^n \otimes \hat{P}^n$.

Let $\pi : \Delta_\Theta \to \Delta_A$ denote the planner's choice function. We will assume that π is continuous at each of the vertices of Δ_Θ, i.e., π is continuous at δ_θ for each $\theta \in \Theta$.

[7]Other more general structures are possible. For example, we could allow for a replica model in which each replica consists of r cohorts, each of which contains n experts as in McLean and Postlewaite (2002).

4.2. The Asymptotic Result

Proposition 2: Let K, M, L and ε be positive numbers and let $\beta \in \Delta^{**}_{\Theta\times T}$. There exists an $\hat{n}$ such that, for all $n > \hat{n}$ and for each aggregation problem $\Pi(n, K, M, \beta, L)$, there exists an incentive compatible mechanism μ^n satisfying

$$\sup_{q^n\in Q^n} \sum_{s^n\in S^n} \|\pi(\cdot|h(s^n)) - \mu^n(\cdot|s^n, q^n)\| \, P(s^n) \le \varepsilon.$$

The proof of Proposition 2 is deferred to section 6, but we will provide an informal construction of the mechanism and present the idea of the argument for the asymptotic version of the jury problem. The proof depends crucially on the fact that, for all sufficiently large n, we can partition T^n into $m+1$ disjoint sets $B^n_0, B^n_1, \ldots, B^n_m$ such that for each $i \in J_n$,

$$\text{Prob}\{\tilde{s}^n \in B^n_0 | \tilde{s}^n_i = s_i\} \approx 0 \text{ for all } s_i \in T$$

$$\text{Prob}\{\tilde{\theta} = \theta_k | \tilde{s}^n = s^n\} \approx 1 \text{ for all } k = 1, \ldots, m \text{ and all } s^n \in B^n_k$$

$$\text{Prob}\{\tilde{s}^n \in B^n_k | \tilde{s}^n_i = s_i\} \approx \beta(\theta_k|s_i) \text{ for all } k = 1, \ldots, m \text{ and all } s_i \in T$$

and

$$\text{Prob}\{(\tilde{s}^n_{-i}, s_i) \in B^n_k \text{ and } (\tilde{s}^n_{-i}, s'_i) \in B^n_k | \tilde{s}^n_i = s_i\} \approx \text{Prob}\{(\tilde{s}^n_{-i}, s_i) \in B^n_k | \tilde{s}^n_i = s_i\}$$

for each $k = 1, \ldots, m$ and each $s_i, s'_i \in T$. These sets are used to define the mechanism. For each i, $q_i \in Q^n_i$ and θ, let $a^n_i(\theta, q^n_i) \in A$ denote the optimal action for expert i in state θ when i's characteristic is q^n_i. Formally, let

$$\{a^n_i(\theta, q^n_i)\} = \arg\max_{a\in A} u^n_i(a, \theta, q^n_i).$$

For each $k = 1, \ldots, m$ and each $s^n \in B^n_k$, define $\alpha^n_i(\cdot|s^n, q^n_i) \in \Delta_A$ where

$$\begin{aligned} \alpha^n_i(a|s^n, q^n_i) &= 1 \text{ if } a = a^n_i(\theta_k, q^n_i) \\ &= 0 \text{ if } a \ne a^n_i(\theta_k, q^n_i) \end{aligned}$$

If $s^n \in B^n_0$, define $\alpha^n_i(\cdot|s^n, q^n_i) \in \Delta_A$ where

$$\alpha^n_i(a|s^n, q^n_i) = \frac{1}{N} \text{ for all } a \in A.$$

Let

$$\rho_i^n(s^n, q_i^n) = \sum_{k=1}^{m} \chi_{B_k^n}(s^n)\gamma_i(\theta_k|s_i^n, q_i^n)$$

where

$$\gamma_i(\theta|s_i^n, q_i^n) = \frac{\hat{\gamma}_i(\theta|s_i^n, q_i^n)}{||\hat{\gamma}_i(\cdot|s_i^n, q_i^n)||_2}$$

and $\hat{\gamma}_i(\theta|s_i^n, q_i^n)$ is defined in equation (4.1). Note that $\gamma_i(\cdot|s_i^n, q_i^n)$ is not generally a probability measure on Θ but it is the case that $0 \leq \gamma_i(\theta|s_i^n, q_i^n) \leq 1$ for each θ.

Next, define

$$\begin{aligned} \varphi(s^n) &= \delta_{\theta_k} \text{ if } s^n \in B_k^n \text{ and } k = 1, \ldots, m \\ &= h(s^n) \text{ if } s^n \in B_0^n. \end{aligned}$$

Finally, define a mechanism where for each $a \in A$ and each $(s^n, q^n) \in T^n \times Q^n$,

$$\mu^n(a|s^n, q^n) = (1-\lambda)\pi(a|\varphi(s^n)) + \frac{\lambda}{n}\sum_{j=1}^{n}\left[\rho_j^n(s^n, q_j^n)\alpha_j^n(a|s^n, q_j^n) + (1 - \rho_j^n(s^n, q_j^n))\frac{1}{N}\right].$$

The mechanism has a flavor similar to that of the jury model presented above. With probability $1 - \lambda$, the decision maker will choose action a with probability $\pi(a|\varphi(s^n))$. With probability $\frac{\lambda}{n}$, one of the experts will be randomly selected for "scrutiny." Suppose that expert i is chosen. If $s^n \in B_k^n$, then the decision maker behaves as if the true state is θ_k. In this case, $\rho_i^n(s^n, q_i^n) = \gamma_i(\theta_k|s_i^n, q_i^n)$. If $s^n \in B_k^n$, then the decision maker will randomize uniformly over the actions $a \in A$ with probability $1 - \gamma_i(\theta_k|s_i^n, q_i^n)$ while, with probability $\gamma_i(\theta_k|s_i^n, q_i^n)$, he will choose action $a_i^n(\theta_k, q_i^n)$ which is the best possible action for expert i in state θ_k if his true characteristic is q_i^n. The mechanism is designed so that, in the presence of many experts, the expert who is chosen for scrutiny is rewarded when he truthfully announces his private information. Since we need to provide incentives for truthful announcements with many alternatives, the mechanism requires a more complex randomizing scheme than that of the jury model where the analogue of $\gamma_i(\theta_k|s_i^n, q_i^n)$ simply takes the value 0 or 1, depending on whether or not a juror's announcement agrees with the majority. To illustrate the idea of the proof, let μ be the mechanism defined above and suppose that expert i with characteristic $q_i^n \in Q_i^n$ observes signal $s_i^n = s_i \in T$. To prove incentive compatibility, it suffices to show that for each $s_i' \in T$, for each $q_i^{n\prime} \in Q_i^n$, and for each $q_{-i}^n \in Q_{-i}^n$,

$$\sum_{s_{-i}^n} \sum_{a \in A} \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i, q_i^n) - \mu(a|s_{-i}^n, q_{-i}^n, s_i', q_i^{n\prime})\right] \left[\sum_k u_i(a, \theta_k, q_i^n) P(\theta_k|s_{-i}^n, s_i)\right] P(s_{-i}^n|s_i) \geq 0. \tag{4.2}$$

The properties of the partition enumerated above imply that the expression on the LHS of inequality (4.2) is approximately equal to

$$\sum_k \sum_{\substack{s_{-i}^n \\ :(s_{-i}^n, s_i) \in B_k^n \\ (s_{-i}^n, s_i') \in B_k^n}} \sum_{a \in A} \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i, q_i^n) - \mu^n(a|s_{-i}^n, q_{-i}^n, s_i', q_i^{n\prime})\right] \left[u_i(a, \theta_k, q_i^n)\right] P(s_{-i}^n|s_i) \tag{4.3}$$

so it actually suffices to prove that expression (4.3) is positive.

If $(s_{-i}^n, s_i) \in B_k^n$ and $(s_{-i}^n, s_i') \in B_k^n$, then $\alpha_j^n(a|s_{-i}^n, s_i, q_j^n) = \alpha_j^n(a|s_{-i}^n, s_i', q_j^n)$ for $j \neq i$ and it follows that

$$\sum_k \sum_{\substack{s_{-i}^n \\ :(s_{-i}^n, s_i) \in B_k^n \\ (s_{-i}^n, s_i') \in B_k^n}} \sum_{a \in A} \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i, q_i^n) - \mu(a|s_{-i}^n, q_{-i}^n, s_i', q_i^{n\prime})\right] \left[u_i(a, \theta_k, q_i^n)\right] P(s_{-i}^n|s_i)$$

$$\geq \tfrac{\lambda}{n} \sum_k \left[\gamma_i(\theta_k|s_i, q_i^n) - \gamma_i(\theta_k|s_i', q_i^{n\prime})\right] \left[u_i(a_i^n(\theta_k, q_i^n), \theta_k, q_i^n) - \tfrac{1}{N} \sum_a u_i^n(a, \theta_k, q_i^n)\right] \sum_{\substack{s_{-i}^n \\ :(s_{-i}^n, s_i) \in B_k^n \\ (s_{-i}^n, s_i') \in B_k^n}} P(s_{-i}^n|s_i)$$

$$\approx \tfrac{\lambda}{n} \sum_k \left[\gamma_i(\theta_k|s_i, q_i^n) - \gamma_i(\theta_k|s_i', q_i^{n\prime})\right] \left[u_i(a_i^n(\theta_k, q_i^n), \theta_k, q_i^n) - \tfrac{1}{N} \sum_a u_i^n(a, \theta_k, q_i^n)\right] \beta(\theta_k|s_i)$$

$$= \tfrac{\lambda}{n} \sum_k \left[\gamma_i(\theta_k|s_i, q_i^n) - \gamma_i(\theta_k|s_i', q_i^{n\prime})\right] \hat{\gamma}_i(\theta_k|s_i, q_i^n).$$

The nondegeneracy condition guarantees that this last expression is positive and the mechanism is incentive compatible.

In the jury case,

$$\begin{aligned} \hat{\gamma}_i(\theta_0|s_i, q_i) &= \frac{q_i}{2}\beta(\theta_0|s_i) \\ \hat{\gamma}_i(\theta_1|s_i, q_i) &= \frac{1 - q_i}{2}\beta(\theta_1|s_i). \end{aligned}$$

Consequently,

$$\sum_k [\gamma_i(\theta_k|s_i,q_i) - \gamma_i(\theta_k|s_i',q_i')]\,\hat{\gamma}_i(\theta_k|s_i,q_i) =$$

$$[\gamma_i(\theta_0|s_i,q_i) - \gamma_i(\theta_0|s_i',q_i')]\,\tfrac{q_i}{2}\beta(\theta_0|s_i) + [\gamma_i(\theta_1|s_i,q_i) - \gamma_i(\theta_1|s_i',q_i')]\,\tfrac{1-q_i}{2}\beta(\theta_1|s_i)$$

and this equation has the following interpretation. In a jury model with many jurors, the mechanism designer will learn the true state with probability close to 1. Let $r(\theta, s_i, q_i)$ denote the probability with which the designer chooses $a = 1$ when the estimated state is θ and agent i announces (s_i, q_i). When all other agents are truthful, agent i has little effect on the outcome if no agent is chosen for scrutiny. Furthermore, agent i has little effect on the outcome when another agent is chosen for scrutiny. Hence, when the number of experts is sufficiently large, expert i will have a strict incentive to tell the truth if he has a strict incentive to tell the truth when he is chosen for scrutiny. Conditional on being chosen, i's expected payoff when he truthfully announces (s_i, q_i) will be

$$-q_i r(\theta_0, s_i, q_i)\beta(\theta_0|s_i) + (1-q_i)r(\theta_1, s_i, q_i)\beta(\theta_1|s_i)$$

while his expected payoff when when he deceitfully announces (s_i', q_i') will be

$$-q_i r(\theta_0, s_i', q_i')\beta(\theta_0|s_i) + (1-q_i)r(\theta_1, s_i', q_i')\beta(\theta_1|s_i).$$

Consequently, he will have a strict incentive to tell the truth if

$$-q_i\,[r(\theta_0, s_i, q_i) - r(\theta_0, s_i', q_i')]\,\beta(\theta_0|s_i) + (1-q_i)\,[r(\theta_1, s_i, q_i) - r(\theta_1, s_i', q_i')]\,\beta(\theta_1|s_i) > 0.$$

In terms of the mechanism μ, $\gamma_i(\theta_0|s_i, q_i)$ represents the probability that juror i is "rewarded" when the decision maker believes that the true state is θ_k. Consequently,

$$\begin{aligned} r(\theta_0, s_i, q_i) &= \gamma_i(\theta_0|s_i,q_i)(0) + [1 - \gamma_i(\theta_0|s_i,q_i)]\,\frac{1}{2} \\ r(\theta_1, s_i, q_i) &= \gamma_i(\theta_1|s_i,q_i)(1) + [1 - \gamma_i(\theta_1|s_i,q_i)]\,\frac{1}{2} \end{aligned}$$

with similar expressions for $r(\theta_0, s_i', q_i')$ and $r(\theta_1, s_i', q_i')$. Substituting the definitions of $r(\theta_0, s_i, q_i)$, $r(\theta_1, s, q)$, $r(\theta_0, s_i', q_i')$ and $r(\theta_1, s_i', q_i')$ above, we conclude that agent i will have a strict incentive to tell the truth if

$$[\gamma_i(\theta_0|s_i,q_i) - \gamma_i(\theta_0|s_i',q_i')]\,\frac{q_i}{2}\beta(\theta_0|s_i) + [\gamma_i(\theta_1|s_i,q_i) - \gamma_i(\theta_1|s_i',q_i')]\,\frac{1-q_i}{2}\beta(\theta_1|s_i) > 0$$

which is precisely what we need.

5. Discussion

Informational requirements of the mechanism

In constructing a mechanism of the sort we analyze, the decision maker needs to know some, but not all the data of the problem. Importantly, the decision maker does *not* need to know the experts' biases, that is, their preferences: these are elicited by the mechanism. The experts have an incentive to truthfully announce that part of their private information independently of whether or not they truthfully announce their information about the state. To employ the mechanism the decision maker needs only to set the probability that he will scrutinize a randomly chosen agent. He would like to choose the smallest probability that will provide each agent with the incentive to reveal truthfully. When agents' signals have varying precision, the decision maker needs to know the minimum precision of the signals to determine the optimal probability of scrutiny. If the decision maker believes that the minimum precision is low, he will need to scrutinize with higher probability than if he believes the minimum precision to be high. The decision maker does not need to know which expert has that minimum precision or the distribution of the precisions of the signals. In other words, the decision maker can always be conservative and scrutinize with sufficiently high probability so that the experts will have incentives to truthfully reveal their private information. Higher probabilities of scrutiny will still provide incentives to reveal truthfully, but at a higher cost of distorting his decisions. In summary, the decision maker needs to know nothing about the agents' preferences, and very little about their information in order to employ the sort of mechanism we analyze.

The mechanism is similarly informationally undemanding on the informed agents. Truthful revelation gives an agent increased probability of getting his preferred outcome should he be scrutinized, while misreporting his information gives the agent a chance of affecting the decision maker's choice in the absence of scrutiny. To weigh these, an agent needs to know, roughly, the number of agents and the minimum precision of other agents' signals. The chance that he is scrutinized depends on the number of other agents, and the chance that he passes scrutiny depends on the minimum precision of their signals. The chance that he will affect the outcome in the absence of scrutiny similarly depends on the number of other agents and the precision of their signals. He needs to know neither of these perfectly, and most importantly, he does not need to know *anything* about other agents' preferences.

Informational size

Our results bear a resemblance to those in McLean and Postlewaite (2002) (MP). That paper considered allocations in pure exchange economies in which agents had private information. The paper introduced a notion of informational size and showed (roughly) that when agents were informationally small, efficient allocations could be approximated by incentive compatible mechanisms. Those results are somewhat similar to our results in that we show that a decision rule that depends on private information can be approximated by an incentive compatible mechanism in some circumstances. If one used the notion of informational size, the experts are informationally small in the circumstances that a decision rule can be closely approximated.

While there is a resemblance between the results in MP and the current paper, there are important differences. First, MP deals with pure exchange economies, so agents can be given incentives to reveal truthfully private information through transfers of goods. In the current paper there do not exist goods that can be used for transfers; incentives have to be provided by distorting the choice rule.

More importantly, in the current paper experts have private information about their own preferences, as did agents in the pure exchange economies in MP. There is an important difference, however. In a pure exchange economy with monotonic preferences, the mechanism designer knows that independent of preferences, he can construct outcomes that reward truthful revelation: simply give an agent strictly more of all goods. The preferences in the current paper are not restricted in this way and for a mechanism to reward or punish an expert, the expert's utility function must be elicited. Finally, MP show only that efficient outcomes can be approximated by incentive compatible allocations. Restriction to efficient outcomes is relatively innocuous when all relevant parties are included in the efficiency calculation. However, we treat the case in which a non-participant – the decision-maker – is not a disinterested party. In our motivating example of the commanding officer eliciting information from his field officers, the choice function of interest was to attack the enemy if their strength was not too great. This may be inefficient from the field officers' perspective since they may prefer not attack under any circumstances. The current paper provides guidance to a decision maker who has a stake in the outcome, while MP does not.

Conditionally dominant strategy mechanisms

The mechanisms that we construct in Propositions 1 and 2 actually satisfy an incentive compatibility requirement that is stronger than the "traditional" notion of interim Bayesian incentive compatibility as defined above. In particular, our mechanism μ satisfies the following condition: for each i, each $(s_i, q_i) \in S_i \times Q_i$, each $(s_i', q_i') \in S_i \times Q_i$, and each function $b_{-i} : S_{-i} \to Q_{-i}$,

$$\sum_{s_{-i}\in S_{-i}} \sum_{\theta\in\Theta} \sum_{a\in A} \left[\mu(a|s_{-i}, s_i, b_{-i}(s_{-i}), q_i) - \mu(a|s_{-i}, s_i', b_{-i}(s_{-i}), q_i')\right] u_i(a, \theta, q_i) P(\theta, s_{-i}|s_i) \geq 0. \tag{5.1}$$

Since the random vectors $\left(\tilde{\theta}, \tilde{s}\right)$ and $\tilde{q}$ are stochastically independent, it is clear that this condition is stronger than interim incentive compatibility. In our model, we formulate a revelation game in which expert i announces his type, i.e., a pair, (s_i, q_i), where s_i is his signal about the state and q_i is the expert's personal characteristic. The equilibrium of the mechanism requires, as usual, that no expert's expected utility would increase by announcing an incorrect signal when other experts are truthfully announcing their signals. The equilibrium requirement for the expert's characteristic, however, is stronger: it requires that, conditional on truthful announcement of signals, truthful announcement of his characteristic be optimal regardless of other experts' announcements.

There are many problems in which agents have multidimensional private information and it is useful to know whether their information can be decomposed into distinct parts, with some parts being more easily extracted than other parts. It is well-understood that mechanisms for which truthful revelation is a dominant strategy are preferable to those for which truthful revelation is only Bayesian incentive compatible. Much of mechanism design uses the weaker notion of incentive compatibility only because of the nonexistence of dominant strategy mechanisms that will accomplish similar goals. Dominant strategy mechanisms have many advantages: they are not sensitive to the distribution of players' characteristics, players have no incentive to engage in espionage to learn other players' characteristics and players need to know nothing about other players' strategies in order to determine their own optimal play. If mechanisms exist for which truthful announcement of some component of a player's information is a dominant strategy, these advantages will accrue at least to certain parts of a player's information.

Formally, consider a revelation game Γ with n players whose (finite) type sets are $(T_i)_{i=1}^n$. As usual, $T = T_1 \times \cdots \times T_n$. We say that (T_i^1, T_i^2) is a *decomposition of* T_i if $T_i = T_i^1 \times T_i^2$, and that $\{(T_i^1, T_i^2)\}_{i=1}^n$ is a *decomposition of* T if (T_i^1, T_i^2) is a decomposition of T_i for $i = 1, \ldots, n$. Let $x : T \to A$ be a mechanism and $\{(T_i^1, T_i^2)\}_{i=1}^n$ be a decomposition of T, and consider functions $\{d_i : T_i \to T_i^2\}_{i=1}^n$; denote by $d_{-i}(t_{-i})$ the collection $\{d_j(t_j)\}_{j\neq i}$. We say x is a *conditionally dominant* *strategy mechanism with respect to* $T^2 := T_1^2 \times \cdots \times T_n^2$ if for each i, for each $(t_i^1, t_i^2) \in T_i$, for each $(\hat{t}_i^1, \hat{t}_i^2) \in T_i$, and for each $\{d_j(\cdot)\}_{j \neq i}$,

$$\sum_{t_{-i}} \left[u_i(x(t_{-i}^1, d_{-i}(t_{-i}), t_i^1, t_i^2); t_i^1, t_i^2, t_{-i}) - u_i(x(t_{-i}^1, d_{-i}(t_{-i}), \hat{t}_i^1, \hat{t}_i^2); t_i^1, t_i^2, t_{-i}) \right] P(t_{-i}|t_i^1, t_i^2) \geq 0.$$

If $T_i = T_i^2$ for all i, then the notion of conditional dominant strategy coincides with the notion of dominant strategy for games of incomplete information (see, for example, the discussion in Cremer and McLean (1985), pp349-350.)

It is easy to verify that in our setup μ is a conditionally dominant strategy mechanism with respect to Q if and only if it satisfies inequality (5.1). This result follows from the fact that the utility of expert i does not depend on his opponents' personal characteristics q_{-i} and the random vectors $(\tilde{\theta}, \tilde{s})$ and $\tilde{q}$ are stochastically independent.

Mechanisms satisfying the conditional dominant strategy property with respect to some part of the asymmetric information are less sensitive to the informational assumptions underlying Bayes equilibria. For this reason, the "maximal" decomposition (that is, the decomposition that makes T^2 as "large" as possible) for which there exist incentive compatible mechanisms that are conditionally dominant strategy with respect to T^2 is of interest.

Group manipulation

This paper uses Bayes equilibrium as the solution concept, as does much of the literature on implementation in asymmetric information games. A drawback of many of the games that employ Bayes equilibrium to implement, or virtually implement, social choice functions is that they are susceptible to manipulation by coalitions: even a pair of agents can gain dramatically by colluding.[8] The mechanism used in this paper is not immune to coalitional manipulation, but it is far less sensitive to it. The probability that an agent can get his most desired alternative if he is scrutinized offsets the probability that he can alter the decision maker's choice in the absence of scrutiny. When there is a fixed finite number of agents, the probability that an agent can affect the decision maker's choice in the absence of scrutiny becomes arbitrarily small as signals become increasingly accurate, which allows the decision maker to choose the probability of scrutiny to be small. The probability that any coalition with fewer than a majority of the agents can affect the outcome will similarly be vanishingly small

[8] See, for example, Jackson (2001) and Abreu and Sen (1991).

as signals become sufficiently accurate. Consequently, even small probabilities of scrutiny will make manipulation unprofitable for coalitions with fewer than half the agents when signals are very accurate. Similarly, when the number of agents gets large, the minimal size coalition that will find coordinated deviations from truthful announcement increases without bound.

Conditional independence of experts' information

In both the finite case and the large numbers case we assumed that experts' information was conditionally independent. This is primarily for pedagogical reasons and the logic underlying the mechanism does not depend crucially on the assumption. Suppose that the accuracy of the experts' signal is fixed. The expected number of experts who receive the same signal is higher for the typical case of conditionally correlated signals than when the signals are conditionally independent. It follows that in the case of correlated signals the probability that any single expert will be pivotal in the final decision decreases relative to the case of conditional independence. Hence the expert's gain from misreporting his signal decreases. At the same time the probability that the expert's signal is in the majority increases. And this makes truthful revelation more profitable. To sum up, allowing correlation across signals increases the benefits of truthful revelation and decreases the benefits from misreporting, thus permitting the decision maker to decrease the probability of scrutiny.[9]

Uniform convergence

To simplify the exposition we assumed that the experts know the social choice rule π. However, our results extend to the case in which the experts are uncertain about the rule that the decision maker is implementing. For example, suppose that the social planner has a personal characteristic that affects his payoff, but is unknown to the experts. Clearly, the planner would like to condition the choice of the choice rule π on his private information.

In Propositions 1 and 2 we provided conditions under which there exists an incentive compatible mechanism that converges to the social choice rule π. Notice that these conditions do not depend on π. In other words, we have a uniform convergence over the set of social choice rules. This implies that the planner can approximately implement a set of rules, one for each of his types.

[9]Of course, there are cases in which the correlation among the signals helps the agents make profitable coalitional deviations. For example, suppose that there are two groups of experts and the signals are perfectly correlated within each group. It is possible that all the agents in a certain group have an incentive to lie about their signals.

Costly information acquisition

Agents in our model are exogenously informed. There may be a serious problem with the type of mechanism we analyze if agents must make investments to acquire information. For both the finite population model and the increasing number of agents model, it is agents' minimal effect on the posterior beliefs that determine the decision that allows the extraction of agents' information at low cost. However, agents have little or no incentive to invest in information if that information will have minimal effect on the posterior beliefs.

Multiple equilibria and weak Virtual Bayesian Implementation

As we mentioned in our discussion of the related literature, our approach to the aggregation of expert opinions is very much related to virtual Bayesian implementation. To clarify this relationship, define a social choice function $f : S \to \Delta(A)$ where $f(s) = \pi(\cdot|h(s))$. Following the definitions in Serrano and Vohra (2005), the rule f is virtually Bayesian implementable if for every $\varepsilon > 0$, there exists a social choice function $f^\varepsilon : S \times Q \to \Delta(A)$ such that f^ε is exactly Bayesian implementable and

$$\sup_{q \in Q} \max_{s \in S} ||\pi(\cdot|h(s)) - f^\varepsilon(\cdot|s,q)|| < \varepsilon.$$

To be precise, Serrano and Vohra stipulate that $\sup_{q\in Q} \max_{s\in S} \max_{a\in A} |f(s)(a) - f^\varepsilon(s,q)(a)| < \varepsilon$, but this is inconsequential. To say that f^ε is exactly Bayesian implementable means that there exists a mechanism consisting of message spaces $M_1, .., M_n$ and an outcome function $G : M_1 \times \cdots \times M_n \to \Delta(A)$ with the following property: every Bayes-Nash equilibrium of the associated game of incomplete information induces an outcome distribution on A that coincides with f^ε. Hence, the rule f is virtually Bayesian implementable if *every* Bayes-Nash equilibrium of the game of incomplete information associated with f^ε induces an outcome distribution on A that is close to f *for all* profiles in S. A weaker notion of virtual Bayesian implementation would only require that *every* Bayes-Nash equilibrium of the game of incomplete information associated with f^ε induce an outcome distribution on A that is close to f *on a subset* of $\hat{S} \subseteq S$ with $P(\hat{S}) \approx 1$. A third, still weaker notion would only require that *there exists* a Bayes-Nash equilibrium of the game of incomplete information associated with f^ε that induces an outcome distribution on A that is close to f *on a subset* of $\hat{S} \subseteq S$ with $P(\hat{S}) \approx 1$.

Recall that

$$\sum_{s \in S} ||\pi(\cdot|h(s)) - \mu(\cdot|s,q)||P(s) \leq \varepsilon$$

implies that

$$\text{Prob}\{||\pi(\cdot|h(\tilde{s})) - \mu(\cdot|\tilde{s}, q)|| \leq \sqrt{\varepsilon}\} > 1 - \sqrt{\varepsilon}.$$

Hence, our notion of implementation corresponds precisely to this third, weakest notion of virtual Bayesian implementation where $f(s)$ corresponds to $\pi(\cdot|h(s))$ and $\mu(\cdot|s, q)$ corresponds to $f^{\varepsilon}(\cdot|s, q)$. This form of weak virtual Bayesian implementation has certain strengths and weaknesses relative to the stronger extant definition of virtual Bayesian implementation. Our weak implementation concept does not require that the social choice rule satisfy Bayesian monotonicity, measurability, virtual monotonicity or related assumptions. Instead, agents need only be informationally small as defined in McLean and Postlewaite (2002), a feature of the probability structure that is not related to the properties of π. On the other hand, there may be several equilibria associated with the mechanism μ, not all of which are good approximations of π in our sense. Indeed, these other equilibria may be preferred by the experts to the truthful revelation equilibrium. Consider the example in the introduction with a commander trying to extract information from his field officers. All field officers reporting that the enemy is strong might be an equilibrium preferred by all field officers to the truthful equilibrium. It is often the case that mechanisms of the type we analyze can be augmented so nontruthful announcements will no longer be equilibria while the truthful equilibria remain.[10] Whether or not this is possible in our framework is interesting but beyond the scope of the present paper.

Commitment

In our mechanism, the decision maker elicits truthful announcement and then uses the announced types to choose an element of A. Our mechanism is quite standard in that, for each (s, q) profile, we construct a random variable (e.g., a "spinner") taking values in A and whose distribution is precisely $\mu(\cdot|s, q)$. However, a potential problem of commitment may arise: will the decision maker, after eliciting the type profile (s, q), actually choose the outcome using the measure $\mu(\cdot|s, q)$, rather than his "ideal" measure $\pi(\cdot|h(s)) = \pi(\cdot|P_{\Theta}(\cdot|s))$? This commitment question will also arise in the virtual Bayesian implementation context described above: will the decision maker, after eliciting the type profile s, actually choose the outcome using the approximating measure f^{ε}, rather than his "ideal" measure f? Perhaps it is helpful to decompose the commitment question into two stages. The decision maker must build the "right" spinner, and then abide by the

[10] See, for example, Postlewaite and Schmeidler (1986).

spinner's realization. The first stage could be eliminated by allowing the agents perform a jointly controlled lottery (Aumann, Maschler and Stearns (1968)) that mimics μ. In this way, the decision maker does not make random choices. However, the decision maker must still commit to choosing the outcome of the jointly controlled lottery.

We assume the decision maker can commit to outcomes he does not like ex post (e.g., an outcome that is optimal for one of the experts.) This ability to commit is crucial, since as pointed out in the introduction, experts may have a dominant strategy to report a given signal in the absence of commitment. However, for some problems (such as the jury problem) it may be natural that an outcome rule is chosen prior to the experts receiving information, which essentially implements the necessary commitment. The sort of mechanism we analyze might also be used by a single decision maker for a sequence of decision problems with groups of informed agents who play only one time (for example the jury problem). In such cases reputational concerns might provide the decision maker with the incentive to follow the mechanism's prescribed outcome but this more complex strategic formulation requires an analysis that is beyond the scope of this paper.

6. Proofs

6.1. Proof of Proposition 1

Choose $\varepsilon > 0$ and let μ be the mechanism defined after the statement of Proposition 1 with $0 < \lambda < \frac{\varepsilon}{4}$.

Part 1: There exists $\overline{\kappa} \in]0, 1[$ such that, for all P satisfying $\kappa(P) > \overline{\kappa}$,

$$\sum_s \|\pi(\cdot|h(s)) - \mu(\cdot|s, q)\| \, P(s) < \varepsilon \text{ for all } q \in Q.$$

Proof: Let

$$H(a|s, q) = \sum_{j=1}^{n} \left[\rho_j(s) \alpha_j(a|s, q_j) + (1 - \rho_j(s)) \frac{1}{2} \right]$$

so that

$$\mu(\cdot|s, q) = (1 - \lambda)\pi(\cdot|h(\psi(s))) + \frac{\lambda}{n} H(\cdot|s, q).$$

Therefore,

$$\sum_s \|\pi(\cdot|h(s)) - \mu(\cdot|s,q)\| P(s) \leq$$

$$\sum_s \|\pi(\cdot|h(s)) - \pi(\cdot|h(\psi(s)))\| P(s) + \lambda \sum_s \left\|\pi(\cdot|h(\psi(s))) - \tfrac{1}{n} H(\cdot|s,q)\right\| P(s).$$

Next, observe that

$$\lambda \sum_s \left\|\pi(\cdot|h(\psi(s))) - \frac{1}{n} H(\cdot|s,q)\right\| P(s) \leq 2\lambda < \frac{\varepsilon}{2}$$

and that

$$\sum_s \|\pi(\cdot|h(s)) - \pi(\cdot|h(\psi(s)))\| P(s)$$

$$= \sum_{i \in J_n} \left\|\pi(\cdot|h(s^1_{-i}, \sigma_0)) - \pi(\cdot|h(s^1))\right\| P(s^1_{-i}, \sigma_0) + \sum_{i \in J_n} \left\|\pi(\cdot|h(s^0_{-i}, \sigma_1)) - \pi(\cdot|h(s^0))\right\| P(s^0_{-i}, \sigma_1)$$

$$\leq 2 \sum_{i \in J_n} \left[P(s^1_{-i}, \sigma_0) + P(s^0_{-i}, \sigma_1)\right].$$

Since $\sum_{i \in J_n} \left[P(s^1_{-i}, \sigma_0) + P(s^0_{-i}, \sigma_1)\right] \to 0$ as $\kappa(P) \to 1$, it follows that there exists $\overline{\kappa} \in]0,1[$ such that $\sum_s \|\pi(\cdot|h(s)) - \pi(\cdot|h(\psi(s)))\| P(s) < \frac{\varepsilon}{2}$ whenever $\kappa(P) > \overline{\kappa}$. Therefore, $\kappa(P) > \overline{\kappa}$ implies that

$$\sum_s \|\pi(\cdot|h(s)) - \mu(\cdot|s,q)\| P(s) < \varepsilon.$$

Part 2: In this part, we establish incentive compatibility. Suppose that juror i observes signal σ_1. A mirror image argument can be applied when juror i observes signal σ_0.

Step 1: There exists a κ'_i such that, for all P satisfying $\kappa(P) > \kappa'_i$ and for all $q_{-i} \in Q_{-i}$ and all $q_i, q'_i \in Q_i$,

$$\sum_{s_{-i} \in S_{-i}} \sum_{\theta \in \Theta} \sum_{a \in A} \left[\mu(a|s_{-i}, \sigma_1, q_{-i}, q_i) - \mu(a|s_{-i}, \sigma_1, q_{-i}, q'_i)\right] u_i(a, \theta, q_i) P(\theta|s_{-i}, \sigma_1) P(s_{-i}|\sigma_1) \geq 0.$$

Proof: For each $a \in A$,

$$\sum_{\substack{s_{-i}\in S_{-i}\\ :(s_{-i},\sigma_1)\in C_1}} \sum_{\theta\in\Theta} \left[u_i(a_i^*(\theta_1,q_i),\theta,q_i) - u_i(a,\theta,q_i)\right] P(\theta|s_{-i},\sigma_1)P(s_{-i}|\sigma_1)$$

$$\approx \left[u_i(a_i^*(\theta_1,q_i),\theta_1,q_i) - u_i(a,\theta_1,q_i)\right]$$

when $\kappa(P) \approx 1$. Hence, there exists a κ_i' such that, for all P satisfying $\kappa(P) > \kappa_i'$,

$$\sum_{\substack{s_{-i}\in S_{-i}\\ :(s_{-i},\sigma_1)\in C_1}} \sum_{\theta\in\Theta} \left[u_i(a_i^*(\theta_1,q_i),\theta,q_i) - u_i(a,\theta,q_i)\right] P(\theta|s_{-i},\sigma_1)P(s_{-i}|\sigma_1) \geq \frac{K}{2}$$

whenever $a \neq a_i^*(\theta_1,q_i)$. Next, note that

$$\mu(a|s_{-i},\sigma_1,q_{-i},q_i) - \mu(a|s_{-i},\sigma_1,q_{-i},q_i') = \frac{\lambda}{n}\rho_i(s_{-i},\sigma_1)\left[\alpha_i(a|s_{-i},\sigma_1,q_i) - \alpha_i(a|s_{-i},\sigma_1,q_i')\right].$$

If $(s_{-i},\sigma_1) \notin C_0 \cup C_1$, then

$$\frac{\lambda}{n}\rho_i(s_{-i},\sigma_1)\left[\alpha_i(a|s_{-i},\sigma_1,q_i) - \alpha_i(a|s_{-i},\sigma_1,q_i')\right] = \frac{\lambda}{n}\rho_i(s_{-i},\sigma_1)\left[\frac{1}{2} - \frac{1}{2}\right] = 0.$$

If $(s_{-i},\sigma_1) \in C_0$, then

$$\rho_i(s_{-i},\sigma_1) = 0.$$

If $(s_{-i},\sigma_1) \in C_1$, then

$$\rho_i(s_{-i},\sigma_1) = 1$$

Therefore,

$$\sum_{s_{-i}\in S_{-i}} \sum_{\theta\in\Theta} \sum_{a\in A} \left[\mu(a|s_{-i},\sigma_1,q_{-i},q_i) - \mu(a|s_{-i},\sigma_1,q_{-i},q_i')\right] u_i(a,\theta,q_i)P(\theta|s_{-i},\sigma_1)P(s_{-i}|\sigma_1)$$

$$= \tfrac{\lambda}{n} \sum_{\substack{s_{-i}\in S_{-i}\\ :(s_{-i},\sigma_1)\in C_1}} \sum_{\theta\in\Theta} \sum_{a\in A} \left[\alpha_i(a|s_{-i},\sigma_1,q_i) - \alpha_i(a|s_{-i},\sigma_1,q_i')\right] u_i(a,\theta,q_i)P(\theta|s_{-i},\sigma_1)P(s_{-i}|\sigma_1)$$

$$= \tfrac{\lambda}{n} \sum_{\substack{s_{-i}\in S_{-i}\\ :(s_{-i},\sigma_1)\in C_1}} \sum_{\theta\in\Theta} \left[u_i(a_i^*(\theta_1,q_i),\theta,q_i) - u_i(a_i^*(\theta_1,q_i'),\theta,q_i)\right] P(\theta|s_{-i},\sigma_1)P(s_{-i}|\sigma_1) \geq 0.$$

Step 2: There exists a κ_i'' such that, for all P satisfying $\kappa(P) > \kappa_i''$ and for all $q_{-i} \in Q_{-i}$ and all $q_i, q_i' \in Q_i$,

$$\sum_{s_{-i}\in S_{-i}} \sum_{\theta\in\Theta} \sum_{a\in A} [\mu(a|s_{-i},\sigma_1,q_{-i},q_i) - \mu(a|s_{-i},\sigma_0,q_{-i},q_i')]\, u_i(a,\theta,q_i)P(\theta|s_{-i},\sigma_1)P(s_{-i}|\sigma_1) \geq 0.$$

Proof: If $\kappa(P) \approx 1$, then $P(\theta_1|s^1) \approx 1$ and $P(s^1_{-i}|\sigma_1) \approx 1$. Since $\kappa(P) \approx 1$ implies that

$$\sum_{s_{-i}\in S_{-i}} \sum_{\theta\in\Theta} \sum_{a\in A} [\mu(a|s_{-i},\sigma_1,q_{-i},q_i) - \mu(a|s_{-i},\sigma_0,q_{-i},q_i')]\, u_i(a,\theta,q_i)P(\theta|s_{-i},\sigma_1)P(s_{-i}|\sigma_1)$$

$$\approx \sum_{a\in A} \left[\mu(a|s^1_{-i},\sigma_1,q_{-i},q_i) - \mu(a|s^1_{-i},\sigma_0,q_{-i},q_i')\right] u_i(a,\theta_1,q_i)$$

for all $q_i, q_i' \in Q_i$ and all $q_{-i} \in Q_{-i}$, it suffices to prove that

$$\sum_{a\in A} \left[\mu(a|s^1_{-i},\sigma_1,q_{-i},q_i) - \mu(a|s^1_{-i},\sigma_0,q_{-i},q_i')\right] u_i(a,\theta_1,q_i) > 0$$

for all $q_i, q_i' \in Q_i$ and all $q_{-i} \in Q_{-i}$. Since $(s^1_{-i},\sigma_0) \in C_1$ and $(s^1_{-i},\sigma_1) \in C_1$, it follows that, for all $j \neq i$,

$$\rho_j(s^1_{-i},\sigma_0) = 1 = \rho_j(s^1_{-i},\sigma_1)$$

and that

$$\alpha_j(\cdot|s^1_{-i},\sigma_1,q_j) = \alpha_j(\cdot|s^1_{-i},\sigma_0,q_j).$$

Therefore,

$$\begin{aligned}&\tfrac{\lambda}{n}\left[\rho_j(s^1_{-i},\sigma_1)\alpha_j(a|s^1_{-i},\sigma_1,q_j) + (1-\rho_j(s^1_{-i},\sigma_1))\tfrac{1}{2}\right]\\ &-\tfrac{\lambda}{n}\left[\rho_j(s^1_{-i},\sigma_0)\alpha_j(a|s^1_{-i},\sigma_0,q_j) + (1-\rho_j(s^1_{-i},\sigma_0))\tfrac{1}{2}\right]\\ &= \tfrac{\lambda}{n}\left[\alpha_j(a|s^1_{-i},\sigma_1,q_j) - \alpha_j(a|s^1_{-i},\sigma_0,q_j)\right] = 0\end{aligned}$$

whenever $j \neq i$. Next, note that

$$\rho_i(s^1_{-i},\sigma_0) = 0 \text{ and } \rho_i(s^1_{-i},\sigma_1) = 1.$$

Combining these observations, we obtain

$$\mu(a|s^1_{-i},\sigma_1,q_{-i},q_i)-\mu(a|s^1_{-i},\sigma_0,q_{-i},q_i')$$
$$=\tfrac{\lambda}{n}\left[\rho_i(s^1_{-i},\sigma_1)\alpha_i(a|s^1_{-i},\sigma_1,q_i)+(1-\rho_i(s^1_{-i},\sigma_1))\tfrac{1}{2}\right]$$
$$-\tfrac{\lambda}{n}\left[\rho_i(s^1_{-i},\sigma_0)\alpha_i(a|s^1_{-i},\sigma_0,q_i')+(1-\rho_i(s^1_{-i},\sigma_0))\tfrac{1}{2}\right]$$
$$=\tfrac{\lambda}{n}\left[\alpha_i(a|s^1_{-i},\sigma_1,q_i)-\tfrac{1}{2}\right]$$

so that

$$\sum_{a\in A}\left[\mu(a|s^1_{-i},\sigma_1,q_{-i},q_i)-\mu(a|s^1_{-i},\sigma_0,q_{-i},q_i')\right]u_i(a,\theta_1,q_i)$$
$$=\tfrac{\lambda}{n}\left[u_i(a_i^*(\theta_1,q_i),\theta_1,q_i)-\tfrac{1}{2}\sum_{a\in A}u_i(a,\theta_1,q_i)\right]\geq\tfrac{\lambda}{n}\tfrac{K}{2}.$$

6.2. Proof of Proposition 2

The proof of Proposition 2 relies on the following technical result whose proof is a summary of results found in McLean and Postlewaite (2002) and (2006). To ease the burden on the reader, we provide a self contained proof of the lemma in the appendix.

Lemma 1: For every $\zeta>0$, there exists an $\hat{n}>0$ such that, for all $n>\hat{n}$, there exists a partition $B^n_0,B^n_1,\ldots,B^n_m$ of T^n such that

(i) For each $i\in J_n$ and $s_i\in T$,

$$\sum_{\substack{s^n_{-i}\\:(s^n_{-i},s_i)\in B^n_0}}P(s^n_{-i}|s_i)\leq n^{-2}$$

(ii) For each $i\in J_n$ and $s_i,s_i'\in T$,

$$\sum_{k=1}^{m}\sum_{\substack{s^n_{-i}\\:(s^n_{-i},s_i)\in B^n_k\\(s^n_{-i},s_i')\notin B^n_k}}P(s^n_{-i}|s_i)\leq n^{-2}$$

(iii) For each $k = 1, \ldots, m$, and for each $s^n \in B_k^n$,

$$||\delta_{\theta_k} - h(s^n)|| \leq \zeta.$$

Choose $\varepsilon > 0$. Let η and λ be positive numbers satisfying

$$\lambda < \frac{\varepsilon}{4},$$

$$K||\beta(\cdot|s_i)||_2 \frac{L^2}{2} - 4m\eta M > 0,$$

and[11] for each $k = 1, \ldots, m$,

$$||h(s^n)) - \delta_{\theta_k}|| \leq \eta \Rightarrow \|\pi(\cdot|h(s^n)) - \pi(\cdot|\delta_{\theta_k})\| \leq \frac{\varepsilon}{2}.$$

Let $\hat{n}$ be an integer such that, for all $n > \hat{n}$, the following three conditions are satisfied :

$$\frac{1}{n^2} < \frac{\eta}{2}$$

$$\left(\frac{N-1}{N}\right)\left(K||\beta(\cdot|s_i)||_2 \frac{L^2}{2} - 4mM\eta\right) - \frac{4M}{\lambda n} > 0$$

and there exists a collection $B_0^n, B_1^n, \ldots, B_m^n$ of disjoint subsets of S^n satisfying the conditions of the lemma with $\zeta = \eta$.

We now define the mechanism. For each i, $q_i^n \in Q_i^n$ and θ, let $a_i^n(\theta, q_i^n) \in A$, $\alpha_i^n(\cdot|s^n, q_i^n) \in \Delta_A$, $\gamma_i(\theta_k|s_i^n, q_i^n), \hat{\gamma}_i(\theta_k|s_i^n, q_i^n)$, $\rho_i^n(s^n, q_i^n)$, and $\varphi(s^n)$ be defined as they are in Section 4.2. As in Section 4.2, define a mechanism μ^n as follows: for each $a \in A$ and each $(s^n, q^n) \in T^n \times Q^n$,

$$\mu^n(a|s^n, q^n) = (1-\lambda)\pi(a|\varphi(s^n)) + \frac{\lambda}{n}\sum_{j=1}^{n}\left[\rho_j^n(s^n, q_j^n)\alpha_j^n(a|s^n, q_j^n) + (1 - \rho_j^n(s^n, q_j^n))\frac{1}{N}\right].$$

First, we record a few facts that will be used throughout the proof.

Fact 1: For all $s_{-i}^n, q_{-i}^n, s_i^n, q_i^n, s_i^{n\prime}$ and $q_i^{n\prime}$,

$$\sum_a \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i^n, q_i^n) - \mu(a|s_{-i}^n, q_{-i}^n, s_i^{n\prime}, q_i^{n\prime})\right]\left[\sum_{\theta\in\Theta} u_i^n(a, \theta, q_i^n)P(\theta|s_{-i}^n, s_i^n)\right] \geq -2M.$$

[11]Recall that the mapping $\pi : \Delta_\Theta \to \Delta_A$ is continuous at δ_θ for each $\theta \in \Theta$.

This follows from the observation that

$$\sum_a |\left[\mu^n(a|s^n_{-i}, q^n_{-i}, s^n_i, q^n_i) - \mu(a|s^n_{-i}, q^n_{-i}, s^{n\prime}_i, q^{n\prime}_i)\right]\left[\sum_{\theta\in\Theta} u^n_i(a,\theta,q^n_i)P(\theta|s^n_{-i}, s^n_i)\right]|$$

$$\leq \sum_a |\mu^n(a|s^n_{-i}, q^n_{-i}, s^n_i, q^n_i) - \mu(a|s^n_{-i}, q^n_{-i}, s^{n\prime}_i, q^{n\prime}_i)| \sum_{\theta\in\Theta} |u^n_i(a,\theta,q^n_i)|P(\theta|s^n_{-i}, s^n_i) \leq 2M$$

Fact 2: For all s^n_i and $s^{n\prime}_i$,

$$\sum_{k=1}^m |\beta(\theta_k|s^n_i) - \text{Prob}\{(\tilde{s}^n_{-i}, s^n_i) \in B^n_k, (\tilde{s}^n_{-i}, s^{n\prime}_i) \in B^n_k|\tilde{s}^n_i = s^n_i\}| \leq 2\eta$$

To see this, simply duplicate the calculations in the proof of Claim 1, p.2444, in McLean and Postlewaite (2002), then use parts (i) and (ii) of Lemma 1 to deduce that

$$\sum_{k=1}^m |\beta(\theta_k|s^n_i) - \text{Prob}\{\tilde{s}^n \in B^n_k|\tilde{s}^n_i = s^n_i\}| \leq \eta + \frac{1}{n^2}.$$

Consequently, part (ii) of Lemma 1 and the assumption that $n^{-2} < \eta/2$ imply that

$$\sum_{k=1}^m |\beta(\theta_k|s^n_i) - \text{Prob}\{(\tilde{s}^n_{-i}, s^n_i) \in B^n_k, (\tilde{s}^n_{-i}, s^{n\prime}_i) \in B^n_k|\tilde{s}^n_i = s^n_i\}|$$

$$\leq \eta + \tfrac{1}{n^2} + \sum_{k=1}^m \text{Prob}\{(\tilde{s}^n_{-i}, s^n_i) \in B^n_k, (\tilde{s}^n_{-i}, s'_i) \notin B^n_k|\tilde{s}^n_i = s^n_i\}$$

$$\leq \eta + \tfrac{2}{n^2} < 2\eta.$$

Fact 3 : For all $s^n_i, q^n_i, s^{n\prime}_i$ and $q^{n\prime}_i$,

$$\sum_k \left[\gamma_i(\theta_k|s^n_i, q^n_i) - \gamma_i(\theta_k|s^{n\prime}_i, q^{n\prime}_i)\right]\hat{\gamma}_i(\theta_k|s^n_i, q^n_i) \geq \left(\frac{N-1}{N}\right) K||\beta(\cdot|s^n_i)||_2 \frac{L^2}{2}$$

First, note that

$$\hat{\gamma}_i(\theta|s^n_i, q^n_i) = \left[u^n_i(a^n_i(\theta, q^n_i), \theta, q^n_i) - \tfrac{1}{N}\sum_a u^n_i(a,\theta,q^n_i)\right]\beta(\theta|s^n_i)$$

$$\geq \left(\tfrac{N-1}{N}\right) K\beta(\theta|s^n_i)$$

for each $\theta \in \Theta$. Therefore,

$$||\hat{\gamma}_i(\cdot|s_i^n, q_i^n)||_2 \geq \left(\frac{N-1}{N}\right) K||\beta(\cdot|s_i^n)||_2.$$

To complete the argument, observe that

$$\begin{aligned}
& \sum_k \left[\gamma_i(\theta_k|s_i^n, q_i^n) - \gamma_i(\theta_k|s_i^{n\prime}, q_i^{n\prime})\right] \hat{\gamma}_i(\theta_k|s_i^n, q_i^n) \\
= \quad & ||\hat{\gamma}_i(\cdot|s_i^n, q_i^n)||_2 \sum_k \left[\gamma_i(\theta_k|s_i^n, q_i^n) - \gamma_i(\theta_k|s_i^{n\prime}, q_i^{n\prime})\right] \gamma_i(\theta_k|s_i^n, q_i^n) \\
= \quad & \frac{||\hat{\gamma}_i(\cdot|s_i^n, q_i^n)||_2}{2} \left[||\gamma_i(\theta_k|s_i^n, q_i^n) - \gamma_i(\theta_k|s_i^{n\prime}, q_i^{n\prime})||_2\right]^2 \\
\geq \quad & \left(\frac{N-1}{N}\right) K||\beta(\cdot|s_i^n)||_2 \frac{L^2}{2}
\end{aligned}$$

Part 1: First we will prove that the mechanism is incentive compatible. For each a, s_{-i}^n, s_i and q_i^n, let

$$v_i(a, s_{-i}^n, s_i, q_i^n) = \sum_{\theta \in \Theta} u_i^n(a, \theta, q_i^n) P(\theta|s_{-i}^n, s_i).$$

We will show that for each $s_i, s_i' \in T$ and for each $q_i^n, q_i^{n\prime} \in Q_i^n$,

$$\sum_{s_{-i}^n} \sum_{a \in A} \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i, q_i^n) - \mu(s_{-i}, q_{-i}^n, s_i', q_i^{n\prime})\right] \left[v_i(a, s_{-i}^n, s_i, q_i^n)\right] P(s_{-i}^n|s_i) \geq 0.$$

Claim 1:

$$\sum_{\substack{s_{-i}^n \\ :(s_{-i}^n, s_i) \in B_0^n}} \sum_a \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i, q_i^n) - \mu(a|s_{-i}^n, q_{-i}^n, s_i', q_i^{n\prime})\right] \left[v_i(a, s_{-i}^n, s_i, q_i^n)\right] P(s_{-i}^n|s_i) \geq -\frac{2M}{n^2}$$

Proof of Claim 1: Applying Fact 1 and (i) of the main Lemma, we conclude that

$$\sum_{\substack{s_{-i}^n \\ :(s_{-i}^n, s_i) \in B_0^n}} \sum_a \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i, q_i^n) - \mu(a|s_{-i}^n, q_{-i}^n, s_i', q_i^{n\prime})\right] \left[v_i(a, s_{-i}^n, s_i, q_i^n)\right] P(s_{-i}^n|s_i)$$

$$\geq -2M \left[\sum_{\substack{s_{-i}^n \\ :(s_{-i}^n, s_i) \in B_0^n}} P(s_{-i}^n|s_i)\right] \geq \frac{-2M}{n^2}$$

Claim 2:

$$\sum_{k} \sum_{\substack{s^n_{-i} \\ :(s^n_{-i}, s_i) \in B^n_k \\ (s^n_{-i}, s'_i) \in B^n_k}} \sum_{a \in A} \left[\mu^n(a|s^n_{-i}, q^n_{-i}, s_i, q^n_i) - \mu(s^n_{-i}, q^n_{-i}, s'_i, q^{n\prime}_i)\right] \left[v_i(a, s^n_{-i}, s_i, q^n_i)\right] P(s^n_{-i}|s_i)$$

$$\geq \tfrac{\lambda}{n} \left(\tfrac{N-1}{N}\right) \left[K||\beta(\cdot|s_i)||_2 \tfrac{L^2}{2} - 4Mm\eta\right]$$

Proof of Claim 2: Suppose that $(s^n_{-i}, s_i) \in B^n_k$ and $(s^n_{-i}, s'_i) \in B^n_k$, then

$$\mu^n(a|s^n_{-i}, s_i, q^n_{-i}, q^n_i) = (1-\lambda)\pi(a|\delta_{\theta_k})$$
$$+\tfrac{\lambda}{n} \textstyle\sum_{j \neq i} \left[\gamma_j(\theta_k|s^n_j, q^n_j)\alpha^n_j(a|\theta_k, q^n_j) + (1 - \gamma_j(\theta_k|s^n_j, q^n_j))\tfrac{1}{N}\right]$$
$$+\tfrac{\lambda}{n} \left[\gamma_i(\theta_k|s_i, q^n_i)\alpha^n_i(a|\theta_k, q^n_i) + (1 - \gamma_i(\theta|s_i, q^n_i))\tfrac{1}{N}\right]$$

and

$$\mu^n(a|s^n_{-i}, s'_i, q^n_{-i}, q^{n\prime}_i) = (1-\lambda)\pi(a|\delta_{\theta_k})$$
$$+\tfrac{\lambda}{n} \textstyle\sum_{j \neq i} \left[\gamma_j(\theta_k|s^n_j, q^n_j)\alpha^n_j(a|\theta_k, q^n_j) + (1 - \gamma_j(\theta_k|s^n_j, q^n_j))\tfrac{1}{N}\right]$$
$$+\tfrac{\lambda}{n} \left[\gamma_i(\theta_k|s'_i, q^{n\prime}_i)\alpha^n_i(a|\theta_k, q^{n\prime}_i) + (1 - \gamma_i(\theta|s'_i, q^{n\prime}_i))\tfrac{1}{N}\right]$$

so that

$$\mu^n(a|s^n_{-i}, s_i, q^n_{-i}, q^n_i) - \mu^n(a|s^n_{-i}, s'_i, q^n_{-i}, q^{n\prime}_i)$$
$$= \tfrac{\lambda}{n}\gamma_i(\theta_k|s_i, q^n_i) \left[\alpha^n_i(a|\theta_k, q^n_i) - \tfrac{1}{N}\right] - \tfrac{\lambda}{n}\gamma_i(\theta_k|s'_i, q^{n\prime}_i) \left[\alpha^n_i(a|\theta_k, q^{n\prime}_i) - \tfrac{1}{N}\right]$$
$$= \tfrac{\lambda}{n} \left[\gamma_i(\theta_k|s_i, q^n_i) - \gamma_i(\theta_k|s'_i, q^{n\prime}_i)\right] \left[\alpha^n_i(a|\theta_k, q^n_i) - \tfrac{1}{N}\right]$$
$$+\tfrac{\lambda}{n}\gamma_i(\theta_k|s'_i, q^{n\prime}_i) \left[\alpha^n_i(a|\theta_k, q^n_i) - \alpha^n_i(a|\theta_k, q^{n\prime}_i)\right].$$

Letting

$$\overline{u}_i(\theta_k, q^n_i) = \frac{1}{N} \sum_a u^n_i(a, \theta_k, q^n_i),$$

it follows that

$$\begin{aligned}
&\sum_a \left[\mu^n(a|s_{-i}^n, s_i, q_{-i}^n, q_i^n) - \mu^n(a|s_{-i}^n, s_i', q_{-i}^n, q_i^{n\prime})\right] u_i(a, \theta_k, q_i^n) \\
&= \tfrac{\lambda}{n}\left[\gamma_i(\theta_k|s_i, q_i^n) - \gamma_i(\theta_k|s_i', q_i^{n\prime})\right]\left[u_i(a_i^n(\theta_k, q_i^n), \theta_k, q_i^n) - \overline{u}_i(\theta_k, q_i^n)\right] \\
&\quad + \tfrac{\lambda}{n}\gamma_i(\theta_k|s_i', q_i^{n\prime})\left[u_i(a_i^n(\theta_k, q_i^n), \theta_k, q_i^n) - u_i(a_i^n(\theta_k, q_i^{n\prime}), \theta_k, q_i^n)\right] \\
&\geq \tfrac{\lambda}{n}\left[\gamma_i(\theta_k|s_i, q_i^n) - \gamma_i(\theta_k|s_i', q_i^{n\prime})\right]\left[u_i(a_i^n(\theta_k, q_i^n), \theta_k, q_i^n) - \overline{u}_i(\theta_k, q_i^n)\right].
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\sum_k \sum_{\substack{s_{-i}^n \\ :(s_{-i}^n, s_i)\in B_k^n \\ (s_{-i}^n, s_i')\in B_k^n}} \sum_a \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i, q_i^n) - \mu(s_{-i}^n, q_{-i}^n, s_i', q_i^{n\prime})\right]\left[u_i(a, \theta_k, q_i^n)\right] P(s_{-i}^n|s_i) \\
&\geq \sum_k \sum_{\substack{s_i^n \\ :(s_{-i}^n, s_i)\in B_k^n \\ (s_{-i}^n, s_i')\in B_k^n}} \tfrac{\lambda}{n}\left[\gamma_i(\theta_k|s_i, q_i) - \gamma_i(\theta_k|s_i', q_i')\right]\left[u_i(a_i^n(\theta_k, q_i), \theta_k, q_i) - \overline{u}_i(\theta_k, q_i)\right] P(s_{-i}^n|s_i) \\
&= \tfrac{\lambda}{n}\sum_k \left[\gamma_i(\theta_k|s_i, q_i) - \gamma_i(\theta_k|s_i', q_i')\right]\left[u_i(a_i^n(\theta_k, q_i), \theta_k, q_i) - \overline{u}_i(\theta_k, q_i)\right] \sum_{\substack{s_i^n \\ :(s_{-i}^n, s_i)\in B_k^n \\ (s_{-i}^n, s_i')\in B_k^n}} P(s_{-i}^n|s_i) \\
&= \tfrac{\lambda}{n}\sum_k \left[\gamma_i(\theta_k|s_i, q_i) - \gamma_i(\theta_k|s_i', q_i')\right]\left[u_i(a_i^n(\theta_k, q_i), \theta_k, q_i) - \overline{u}_i(\theta_k, q_i)\right]\beta(\theta_k|s_i) \\
&\quad + \tfrac{\lambda}{n}\sum_k \left[\gamma_i(\theta_k|s_i, q_i) - \gamma_i(\theta_k|s_i', q_i')\right]\left[u_i(a_i^n(\theta_k, q_i), \theta_k, q_i) - \overline{u}_i(\theta_k, q_i)\right] \times \\
&\qquad \left[\sum_{\substack{s_i^n \\ :(s_{-i}^n, s_i)\in B_k^n \\ (s_{-i}^n, s_i')\in B_k^n}} P(s_{-i}^n|s_i) - \beta(\theta_k|s_i)\right] \\
&\geq \tfrac{\lambda}{n}\left[\sum_k \left[\gamma_i(\theta_k|s_i, q_i) - \gamma_i(\theta_k|s_i', q_i')\right]\hat{\gamma}_i(\theta_k|s_i, q_i) - 4mM\left(\tfrac{N-1}{N}\right)\eta\right] \\
&\geq \tfrac{\lambda}{n}\left[\left(\tfrac{N-1}{N}\right)K||\beta(\cdot|s_i)||_2\tfrac{L^2}{2} - 4mM\left(\tfrac{N-1}{N}\right)\eta\right]
\end{aligned}$$

where the last two inequalities follow from Facts 2 and 3.

Claim 3:

$$\sum_{k} \sum_{\substack{s_i^n \\ :(s_{-i}^n, s_i) \in B_k^n \\ (s_{-i}^n, s_i') \notin B_k^n}} \sum_{a} \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i, q_i^n) - \mu(a|s_{-i}^n, q_{-i}^n, s_i', q_i^{n\prime})\right] \left[v_i(a, s_{-i}^n, s_i, q_i^n)\right] P(s_{-i}^n|s_i) \geq -\frac{2M}{n^2}$$

Proof of Claim 3: This follows from Fact 1 and part (ii) of the main Lemma since

$$\sum_{k} \sum_{\substack{s_{-i}^n \\ :(s_{-i}^n, s_i) \in B_k^n \\ (s_{-i}^n, s_i') \notin B_k^n}} \sum_{a} \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i, q_i^n) - \mu(a|s_{-i}^n, q_{-i}^n, s_i', q_i^{n\prime})\right] v_i(a, s_{-i}^n, s_i, q_i^n) P(s_{-i}^n|s_i)$$

$$\geq -2M \sum_{k} \sum_{\substack{s_{-i}^n \\ :(s_{-i}^n, s_i) \in B_k^n \\ (s_{-i}^n, s_i') \notin B_k^n}} P(s_{-i}^n|s_i) \geq -\tfrac{2M}{n^2}.$$

Final step of the proof: Combining these claims, we conclude that

$$\sum_{s_{-i}^n} \sum_{a \in A} \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i, q_i^n) - \mu(s_{-i}, q_{-i}, s_i', q_i^{n\prime})\right] \left[v_i(a, s_{-i}^n, s_i, q_i^n)\right] P(s_{-i}^n|s_i)$$

$$= \sum_{\substack{s_{-i}^n \\ :(s_{-i}^n, s_i) \in B_0^n}} \sum_{a} \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i, q_i^n) - \mu(a|s_{-i}^n, q_{-i}^n, s_i', q_i^{n\prime})\right] \left[v_i(a, s_{-i}^n, s_i, q_i^n)\right] P(s_{-i}^n|s_i)$$

$$+ \sum_{k} \sum_{\substack{s_{-i}^n \\ :(s_{-i}^n, s_i) \in B_k^n \\ (s_{-i}^n, s_i') \in B_k^n}} \sum_{a \in A} \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i, q_i^n) - \mu(s_{-i}^n, q_{-i}^n, s_i', q_i^{n\prime})\right] \left[v_i(a, s_{-i}^n, s_i, q_i^n)\right] P(s_{-i}^n|s_i)$$

$$+ \sum_{k} \sum_{\substack{s_{-i}^n \\ :(s_{-i}^n, s_i) \in B_k^n \\ (s_{-i}^n, s_i') \notin B_k^n}} \textstyle\sum_{a \in A} \left[\mu^n(a|s_{-i}^n, q_{-i}^n, s_i, q_i^n) - \mu(s_{-i}^n, q_{-i}^n, s_i', q_i^{n\prime})\right] \left[v_i(a, s_{-i}^n, s_i, q_i^n)\right] P(s_{-i}^n|s_i)$$

$$\geq \tfrac{-2M}{n^2} + \tfrac{\lambda}{n} \left[\left(\tfrac{N-1}{N}\right) K ||\beta(\cdot|s_i)||_2 \tfrac{L^2}{2} - 2mM \left(\tfrac{N-1}{N}\right) 2\eta \right] - \tfrac{2M}{n^2}$$

$$= \tfrac{\lambda}{n} \left[\left(\tfrac{N-1}{N}\right) \left(K ||\beta(\cdot|s_i)||_2 \tfrac{L^2}{2} - 4mM\eta \right) - \tfrac{4M}{\lambda n} \right] > 0.$$

Part 2: We now show that , for all $n > \hat{n}$,

$$\sup_{q^n \in Q^n} \sum_{s^n \in T^n} \|\pi(\cdot|h(s^n)) - \mu^n(\cdot|s^n, q^n)\| \, P(s^n) < \varepsilon.$$

Fixing $q^n \in Q^n$, it follows that

$$\pi(a|h(s^n)) - \mu^n(a|s^n, q^n) = \pi(a|h(s^n)) - \pi(a|\varphi(s^n))$$

$$+\lambda\left(\pi(a|\varphi(s^n)) - \tfrac{1}{n}\sum_{j=1}^{n}\left[\rho_j^n(s^n, q_j^n)\alpha_j^n(a|s^n, q_j^n) + (1 - \rho_j^n(s^n, q_j^n))\tfrac{1}{N}\right]\right).$$

Since

$$\sum_{s^n \in B_0^n} \|\pi(\cdot|h^n(s^n)) - \pi(\cdot|\varphi(s^n))\| \, P(s^n) = 0$$

and

$$s^n \in B_k^n \Rightarrow ||h^n(s^n)) - \delta_{\theta_k}|| \leq \eta \Rightarrow \|\pi(\cdot|h^n(s^n)) - \pi(\cdot|\delta_{\theta_k})\| \leq \frac{\varepsilon}{2}$$

for each $k = 1, \ldots, m$, we conclude that

$$\begin{aligned}
&\sum_{s^n \in T^n} \|\pi(\cdot|h(s^n)) - \pi(\cdot|\varphi(s^n))\| \, P(s^n) \\
= &\sum_{s^n \in B_0^n} \|\pi(\cdot|h(s^n)) - \pi(\cdot|\varphi(s^n))\| \, P(s^n) \; + \sum_{k=1}^{m} \sum_{s^n \in B_k^n} \|\pi(\cdot|h(s^n)) - \pi(\cdot|\delta_{\theta_k})\| \, P(s^n) \\
\leq &\frac{\varepsilon}{2}.
\end{aligned}$$

Therefore, we obtain

$$\sum_{s^n \in T^n} \|\pi(\cdot|h(s^n)) - \mu^n(\cdot|s^n, q^n)\| \leq \frac{\varepsilon}{2} + 2\lambda < \varepsilon.$$

7. Appendix

Proof of Lemma 1: For each $s^n \in T^n$, let $f(s^n)$ denote the "empirical frequency distribution" that s^n induces on T. More formally, $f(s^n)$ is a probability measure on T defined for each $\hat{s} \in T$ as follows:

$$f(s^n)(\hat{s}) = \frac{|\{i \in J_n | s_i^n = \hat{s}\}|}{n}$$

(We suppress the dependence of f on n for notational convenience.)

Choose $\zeta > 0$. For each $\gamma > 0$ and $1 \leq k \leq m$, let

$$B_k^\gamma = \{s^n | ||f(s^n) - \beta(\cdot|\theta_k)|| < \gamma\}$$

where $\beta(\cdot|\theta_k)$ denotes the conditional distribution on T given θ_k. (We suppress the dependence of B_k^γ on n for notational convenience.) Applying the argument in the appendix to Gul and Postlewaite(1992) (see the analysis of their equation (9)), it follows that there exists $\alpha > 0$ and an integer n_1 such that $B_1^\alpha, \ldots, B_m^\alpha$ are disjoint (because the conditional distributions $\beta(\cdot|\theta_1), \ldots, \beta(\cdot|\theta_m)$ are distinct) and for all $n > n_1$,

$$s^n \in B_k^\alpha \Rightarrow ||P_\Theta(\cdot|s^n) - \delta_{\theta_k}|| < \zeta \text{ for all } k \geq 1.$$

Furthermore, there exists an n_2 such that, for all $n > n_2$ and for each i, each $s^n \in T^n$ and each $s' \in T$,

$$s^n \in B_k^{\frac{\alpha}{4}} \Rightarrow (s_{-i}^n, s') \in B_k^{\frac{\alpha}{2}}$$

and

$$s^n \in B_k^{\frac{\alpha}{2}} \Rightarrow (s_{-i}^n, s') \in B_k^\alpha.$$

Finally, there exists an n_3, such that for all $n > n_3$,

$$z_n \equiv 2|T| \exp[\frac{-n\alpha^2}{8|T|^2}] < \frac{1}{n^2}.$$

Let $\hat{n} = max\{n_1, n_2, n_3\}$ and suppose that $n > \hat{n}$. Define $B_0^\alpha = T^n \backslash [B_1^\alpha \cup \cdots \cup B_m^\alpha]$.

Claim 1: For each i and for each $\hat{s} \in T$,

$$\text{Prob}\{\tilde{s}^n \in B_0^\alpha | \tilde{s}_i^n = \hat{s}\} \leq z_n$$

Proof of Claim 1: First, note that

$$\text{Prob}\{\tilde{s}^n \in B_0^\alpha | \tilde{s}_i^n = \hat{s}\} = 1 - \sum_{k=1}^m \text{Prob}\{\tilde{s}^n \in B_k^\alpha | \tilde{s}_i^n = \hat{s}\}$$

since the sets $B_0^\alpha, B_1^\alpha, \ldots, B_m^\alpha$ partition T^n. Fixing k, it follows that

$$\text{Prob}\{\tilde{s}^n \in B_k^\alpha | \tilde{s}_i^n = \hat{s}\} = \sum_{\ell=1}^m \text{Prob}\{\tilde{s}^n \in B_k^\alpha | \tilde{s}_i^n = \hat{s}, \tilde{\theta} = \theta_\ell\} \beta(\theta_\ell | \hat{s})$$

We will now borrow an argument from McLean and Postlewaite (2006) to bound the RHS of this equality using a classic large deviations result due to Hoeffding. For each k, the conditional independence assumption implies that

$$\text{Prob}\{\tilde{s}^n \in B_k^{\frac{\alpha}{2}} | \tilde{s}_i^n = \hat{s}, \tilde{\theta} = \theta_k\} = \text{Prob}\{(\tilde{s}_{-i}^n, \hat{s}) \in B_k^{\frac{\alpha}{2}} | \tilde{\theta} = \theta_k\}$$

$$\geq \text{Prob}\{\tilde{s}^n \in B_k^{\frac{\alpha}{4}} | \tilde{\theta} = \theta_k\} \geq 1 - z_n$$

where the last inequality is an application of Theorems 1 and 2 in Hoeffding (1963). Combining these observations, we deduce that

$$\text{Prob}\{\tilde{s}^n \in B_k^{\alpha} | \tilde{s}_i^n = \hat{s}\} = \textstyle\sum_{\ell=1}^m \text{Prob}\{\tilde{s}^n \in B_k^{\alpha} | \tilde{s}_i^n = \hat{s}, \tilde{\theta} = \theta_\ell\}\beta(\theta_\ell|\hat{s})$$

$$\geq \text{Prob}\{\tilde{s}^n \in B_k^{\alpha} | \tilde{s}_i^n = \hat{s}, \tilde{\theta} = \theta_k\}\beta(\theta_k|\hat{s})$$

$$\geq \text{Prob}\{\tilde{s}^n \in B_k^{\frac{\alpha}{2}} | \tilde{s}_i^n = \hat{s}, \tilde{\theta} = \theta_k\}\beta(\theta_k|\hat{s})$$
$$\geq (1 - z_n)\beta(\theta_k|\hat{s}).$$

Therefore,

$$\text{Prob}\{\tilde{s}^n \in B_0^{\alpha} | \tilde{s}_i^n = \hat{s}\} = 1 - \sum_{k=1}^m \text{Prob}\{\tilde{s}^n \in B_k^{\alpha} | \tilde{s}_i^n = \hat{s}\}$$

$$\leq 1 - \sum_{k=1}^m (1 - z_n)\beta(\theta_k|\hat{s}) = z_n.$$

Claim 2: For each i and for each $\hat{s}, s' \in T$,

$$\sum_{k=1}^m \text{Prob}\{\tilde{s}^n \in B_k^{\alpha}, (\tilde{s}_{-i}^n, s') \notin B_k^{\alpha} | \tilde{s}_i^n = \hat{s}\} \leq z_n$$

Proof of Claim 2: Note that

$$\sum_{k=1}^m \text{Prob}\{\tilde{s}^n \in B_k^{\alpha}, (\tilde{s}_{-i}^n, s') \notin B_k^{\alpha} | \tilde{s}_i^n = \hat{s}\}$$

$$= \sum_{k=1}^m \text{Prob}\{\tilde{s}^n \in B_k^{\frac{\alpha}{2}}, (\tilde{s}_{-i}^n, s') \notin B_k^{\alpha} | \tilde{s}_i^n = \hat{s}\}$$

$$+ \sum_{k=1}^m \text{Prob}\{\tilde{s}^n \in B_k^{\alpha} \backslash B_k^{\frac{\alpha}{2}}, (\tilde{s}_{-i}^n, s') \notin B_k^{\alpha} | \tilde{s}_i^n = \hat{s}\}$$

$$= \sum_{k=1}^{m} \text{Prob}\{\tilde{s}^n \in B_k^{\alpha} \backslash B_k^{\frac{\alpha}{2}}, (\tilde{s}_{-i}^n, s') \notin B_k^{\alpha} | \tilde{s}_i^n = \hat{s}\}$$

$$\leq \sum_{k=1}^{m} \text{Prob}\{\tilde{s}^n \notin B_k^{\frac{\alpha}{2}} | \tilde{s}_i^n = \hat{s}\}$$

$$= \sum_{k=1}^{m} \sum_{\ell=1}^{m} \text{Prob}\{\tilde{s}^n \notin B_k^{\frac{\alpha}{2}} | \tilde{s}_i^n = \hat{s}, \tilde{\theta} = \theta_\ell\} \beta(\theta_\ell | \hat{s})$$

$$\leq \sum_{k=1}^{m} \text{Prob}\{\tilde{s}^n \notin B_k^{\frac{\alpha}{2}} | \tilde{s}_i^n = \hat{s}, \tilde{\theta} = \theta_k\} \beta(\theta_k | \hat{s})$$

$$\leq \sum_{k=1}^{m} z_n \beta(\theta_k | \hat{s})$$

$$= z_n.$$

References

[1] Abreu, D. and H. Matsushima (1992), "Virtual Implementation in Iteratively Undominated Strategies: Incomplete Information," Mimeo, Princeton University and the University of Tsukuba.

[2] Abreu, D. and A. Sen (1991), "Virtual Implementation in Nash Equilibrium," *Econometrica* 59, 997-1021.

[3] Aumann, R. J., M. Maschler and R. E. Stearns (1968): "Repeated Games of Incomplete Information: An Approach to the Non-zero Sum Case," in Aumann and Maschler (1995): *Repeated Games with Incomplete Information.* Cambridge, MA: The MIT Press.

[4] Austen-Smith, D. (1993), "Interested Experts and Policy Advice: Multiple Referrals under Open Rule," *Games and Economic Behavior* 5, 3-43.

[5] Battaglini, M. (2002), "Multiple Referrals and Multidimensional Cheap Talk," *Econometrica* 70, 1379-401.

[6] Battaglini, M. (2004), "Policy Advice with Imperfectly Informed Experts," *Advances in Theoretical Economics* 4, Article 1.

[7] Crawford, P. V. and J. Sobel (1982), "Strategic Information Transmission," *Econometrica* 50, 1431-51.

[8] Cremer, J. and McLean, R. P. (1985), "Optimal Selling Strategies under Uncertainty for a Discriminatory Monopolist when Demands Are Interdependent," *Econometrica* 53, 345-61.

[9] Duggan, J. (1997), "Virtual Bayesian Implementation," *Econometrica* 65, 1175-1199.

[10] Feddersen, T. and W. Pesendorfer (1987), "Voting Behavior and Information Aggregation In Elections with Private Information," *Econometrica* 65, 1029-58.

[11] Feddersen, T. and W. Pesendorfer (1988), "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," *American Political Science Review* 92, 23-35.

[12] Gilligan, T. W. and K. Krehbiel (1989), "Asymmetric Information and Legislative Rules with a Heterogeneous Committee," *American Journal of Political Science* 33, 459-90.

[13] Gul, F. and A. Postlewaite (1992), "Asymptotic Efficiency in Large Economies wit Asymmetric Information," *Econometrica* 60, 1273-1292.

[14] Hoeffding, W.(1963), "Probability Inequalities for Sums of Bounded Random Variables," *Journal of the American Statistical Association* 58, 13-30.

[15] Jackson, M. (2001), "A Crash Course in Implementation Theory," *Social Choice and Welfare* 18, 655-708.

[16] Krishna, V. and J. Morgan (2001), "Asymmetric Information and Legislative Rules: Some Amendments," *American Political Science Review* 95, 435-52.

[17] Matsushima, H. (1993), "Bayesian Monotonicity with Side Payments," *Journal of Economic Theory* 59, 107-121.

[18] McLean, R. and A. Postlewaite (2002), "Informational Size and Incentive Compatibility," *Econometrica* 70, 2421-2454.

[19] McLean, R. and A. Postlewaite (2004), "Informational Size and Efficient Auctions," *Review of Economic Studies* 71, 809-827.

[20] McLean, R., and A. Postlewaite (2006), "Implementation with Interdependent Valuations," mimeo, University of Pennsylvania.

[21] Ottaviani, M. and P. N. Sørensen (2006a), "Reputational Cheap Talk," *Rand Journal of Economics* 37, 155-175.

[22] Ottaviani, M. and P. N. Sørensen (2006b), "Professional Advice," *Journal of Economic Theory*, 126, 120-142.

[23] Postlewaite, A. and D. Schmeidler (1986), "Implementation in Differential Information Economies," *Journal of Economic Theory* 39, 14-33.

[24] Serrano, R. and R. Vohra (2005), "A Characterization of Virtual Bayesian Implementation," *Games and Economic Behavior* 50, 312-331.

[25] Wolinsky, A. (2002), "Eliciting Information from Multiple Experts," *Games and Economic Behavior* 41, 141-60.