

Implementation without Incentive Compatibility: Two Stories with Partially Informed Planners*

Makoto Shimoji[†] Paul Schweinzer[‡]

August 3, 2012

Abstract

We consider implementation problems under incomplete information *without incentive compatibility*. If the social choice functions do not satisfy incentive compatibility, full implementation is unattainable via the existing approaches. By focusing on the actual problems from *Typhoon* by Joseph Conrad and *The Traveler's Dilemma* by Kaushik Basu (1994, 2007), we provide a new approach to such implementation problems. For each problem, we first construct a mechanism which takes advantage of a unique feature of these problems, i.e., the planners possess *some* information regarding the actual state. We then provide a sufficient condition *on players' beliefs* for each problem under which every player has a *unique* rationalizable action. The conditions we identify however depend on the informational structures, suggesting that obtaining a general result within this type of frameworks is nontrivial.

Keywords: Implementation, Rationalizability, Incentive Compatibility, Incomplete Information.

JEL Classification: C72, D82

*This paper supersedes the paper titled “Captain MacWhirr’s Problem Revisited” (2011). We are grateful to Dirk Bergemann, Anindya Bhattacharya, Subir Chattopadhyay, Kim-Sau Chung, Vincent Crawford, Alex Gershkov, Sanjeev Goyal, Matthew Jackson, Kohei Kawamura, Navin Kartik, Gene Mumy, Tymofiy Mylovanov, Ella Segev, Joel Sobel, and Olivier Tercieux for helpful comments, suggestions, and discussions at various stages. We appreciate Ella Segev for mentioning Basu (1994). We would like to thank Pierpaolo Battigalli for his detailed and thoughtful comments which also gave us an opportunity to expand the scope of the paper.

[†]Department of Economics and Related Studies, University of York, Heslington, YO10 5DD, UK, E-mail: makoto.shimoji@york.ac.uk

[‡]Department of Economics and Related Studies, University of York, Heslington, YO10 5DD, UK, E-mail: paul.schweinzer@york.ac.uk

..., when we risque no contradiction, it prompts the tongue to deal in fiction.

– John Gay

The Elephant and the Bookseller in *Fables I*¹

1 Introduction

In a standard implementation problem with incomplete information, the planner’s task is to construct a (direct) mechanism where each player is asked to reveal what she observes. This task is challenging; the planner does not know what the players know and hence is unable to verify whether they are telling the truth. The mechanism the planner constructs should induce the players to tell the truth. To achieve this goal, the corresponding social choice function needs to satisfy *incentive compatibility*; each player’s best response is to tell the truth if the other players tell the truth and the outcome is determined by the social choice function (with the players’ messages). Then, the (desired) truth-telling outcome can be supported as an equilibrium.² Previous studies by Dasgupta, Hammond, and Maskin (1979), Harris and Townsend (1981), Myerson (1979) and more recently by Bergemann and Morris (2009, 2011) show that the (Bayesian or ex-post) incentive compatibility condition is *necessary* for full implementation. This immediately implies that full implementation cannot be achieved without incentive compatibility, at least through the standard framework.

In this paper, we provide a new approach to implementation problems without incentive compatibility. To do so, our focus is on two implementation problems under incomplete information which are sketched in two well-known stories; one is called *Captain MacWhirr’s problem* in *Typhoon* by Joseph Conrad and the other is what we call the *smashed antiques problem* in two papers on the *Traveler’s Dilemma* by Basu (1994, 2007).³ In these stories, individual’s “dishonest” behavior not only could

¹Two volumes of *Fables* were published in 1727 and 1738 (Gay (1969)).

²Exceptions include the paper by Cabrales, Calvó-Armengol, and Jackson (2003) which analyzes a mutual fire insurance called *La Crema* in Spain. Each participant reports the property value. When the property is burnt, the reimbursement only depends on this “announced” value. They show that if the size of the population is large, the truth-telling outcome is supported as an ϵ -Nash equilibrium.

³*Typhoon* was first serialized in Pall Mall Magazine in 1902 (Watts (2008a)). As we will discuss later, we modify the original setting in Basu (1994, 2007).

(potentially) lead to a higher payoff but also would be anonymous or would not be noticed at all. If telling the truth is never a best response for players (as in the original stories), it would be impossible to achieve the desired outcome.

Captain MacWhirr's Problem

The Siamese steamer Nan-Shan transports 200 Chinese workers, who have worked for seven years in the British tropical colonies, from Singapore to their home of Fuchau. Each worker's accumulated savings are stored in individual camphor wood chests aboard the ship. When a typhoon strikes with ferocious force on Christmas Eve, the boxes burst open and the workers' silver dollars are scattered between decks. In the ensuing chaos, the Captain's orders result in the passengers' belongings to be amassed in a coal bunker. As soon as the storm calms down, the Captain intends to return the men's savings to their rightful owners. But the Captain faces an information revelation problem:

You couldn't tell one man's dollars from another's, he said, and if you asked each man how much money he brought on board he was afraid they would lie, and he would find himself a long way short. (Conrad (2008, p. 73))

In the first economic analysis of Captain MacWhirr's problem, Mumy (1981) claims that it is possible to motivate all passengers to truthfully report their entitlements to the Captain through the use of a suitable punishment strategy. The solution identified by Mumy, however, is only one of many possible Nash equilibria of the game.⁴

Smashed Antiques Problem

The story in Basu (1994, 2007) goes as follows: Two travelers on the same flight are carrying antiques, which are smashed during the flight. Knowing that these two antiques are identical, the airline company manager attempts to compensate for these damaged antiques.⁵ The problem is that the manager does not know the actual value.

⁴Saraydar (1983) criticizes this fact and defends MacWhirr's solution of assigning equal shares to each passenger on the grounds of feasibility and transaction costs.

⁵Strictly speaking, it is unclear whether (i) the travelers already know that they are carrying the same antiques, or (ii) they are only informed by the manager. We assume that these travelers have

Given this scenario, Basu (1994) introduced the following game, called the *Traveler's Dilemma*, where each traveler simultaneously announces the value of the antique, which is between 2 and 100. If they announce the same number, each receives the amount she announced. If they are different, the traveler who announced the lower value, v' , receives $v' + 2$ while the other traveler, who announced the higher value, receives $v' - 2$. It is shown that “announcing 2” is a unique rationalizable action for each player independent of the actual value.⁶ Although each announcing a number higher than 5 is Pareto-improving, which may seem to be rather natural, the standard solution concepts predict that both travelers announce the lowest value.

Our interest is in the scenario the manager faces, not the game itself. In Basu (1994, p. 391), the manager “assures the passengers of adequate compensation.”⁷ In this sense, the manager fails to achieve the desired outcome. Just like in Conrad’s story, achieving the desired outcome seems to be a challenging task. Basu (2007, p 68) writes:

Simply asking the travelers for the price is hopeless, he figures, for they will inflate it.

In this paper, we allow for *heterogeneity*. Hence, the realization in the original story (i.e., they are identical) is one of many possible observations. We also assume that the planner can see which antique has a higher value if they are not identical (but not their exact values). Note that even with this modification on the informational assumption, we cannot avoid the incentive problem. We also make it the case that the number of player is at least two.

The social choice functions implied in these stories are indeed the same. That is, each player observes her type (private information); the amount saved in Captain MacWhirr’s problem and the value of the antique in the smashed antiques problem. The planner’s goal is to give each player the amount equal to her type. From the quotations above, it is clear why achieving the desired outcomes in these stories is difficult. Assuming that the planner simply asks each player how much she deserves

not met before and did not know that they are carrying identical antiques.

⁶Since each player knows (i) the value of the antique she has and (ii) that the antiques are identical, this game has complete information. It can be shown that for iterative elimination of *strictly* dominated strategies (rationalizability), the use of mixed strategies is essential while it is not the case for iterative elimination of *weakly* dominated strategies.

⁷Basu (2007, p.68) also writes “the manager says that he is happy to compensate them.”

to receive (i.e., direct mechanism), since this is her private information, each player has an incentive to exaggerate it independent of how much the other players are entitled to receive. Thus, the social choice functions implied in the stories above do not satisfy incentive compatibility.

As we stated above, our goal is to show, by focusing on the implementation problems in the stories mentioned above, a feasible approach to implementation problems under incomplete information even when the incentive compatibility condition does not hold. This is possible for two reasons. First, the problems we are interested in have a unique feature which the standard implementation problems do not have; the availability of *some* information to the planner regarding the actual state. Second, our sufficient conditions are not placed on the social choice function. They are rather reflected in players' beliefs about their opponents' types.

One common and unique component in these problems is that the planner can possess some (verifiable) information regarding the actual state – she still does not know the actual state. Namely, the total amount collected in Captain MacWhirr's problem and the observation regarding which has a higher value (if not identical) in the smashed antique problem. In this paper, we allow the planners to embed this private information into the mechanisms. Our approach suggests that even if the available information to the planner is limited, it is still possible for her to rely on such information, which may drastically change the outcome.

Even if the planner possesses such information, it is still not possible for her to see directly whether or not the players are telling the truth, unless she knows the actual state. However, in the contexts we consider, the planner knows which message profiles do *not* match the information she has. This creates an informational advantage for the planner, in particular, if she does not reveal what she observed.⁸ Given the mechanism which embeds the information the planner received, the outcome depends on the players' messages and *their consistency with the information the planner holds*.⁹

Recently, given that the informational requirements in previous studies such as Postlewaite and Schmeidler (1986), Palfrey and Srivastava (1989), and Jackson (1991)

⁸If the players know what the planner observed, the desired outcome would not be achieved. This can be observed as the existence of multiple equilibria in Mury (1981) for Captain MacWhirr's problem. If the values are identical in the smashed antiques problem, there would be multiple equilibria as well.

⁹In other words, inconsistency implies that the planner avoids false negative (Type II error) with the hypothesis that the players are telling the truth.

– relying on the Bayesian framework – may be implausible, Bergemann and Morris (2009, 2011) seek the necessary and sufficient conditions for *robust implementation*, i.e., under a weaker informational requirement.¹⁰ This leads to the employment of a different solution concept, namely *rationalizability*.¹¹ Robust implementation implies that each player has a unique rationalizable action for each (payoff) type with no assumption on the players’ beliefs.

If there is a common belief of the information the planner holds, the mechanism we provide reduces to a simple direct mechanism. In this case, full implementation is not possible.¹² It is hence crucial that the players do not know what the planner observed. This immediately implies that the implementation results in this paper are not “robust” in the sense of Bergemann and Morris (2009, 2011).

Unlike standard approaches which place conditions on the social choice function, given that the social choice functions are already implied in the stories above, we take a different step. To respect the desirability of weaker informational assumptions, we employ the notion of Δ -*rationalizability* by Battigalli and Siniscalchi (2003). In each implementation problem, we provide a sufficient condition on the players’ (first-order) beliefs about the opponents’ types, which corresponds to the prefix “ Δ .” We only assume that there is a common belief of the restriction on players’ beliefs. Given the mechanism which includes the private information the planner observes, we show that truth telling is a unique rationalizable action for each player.

We compare the results for both problems. In addition to the differences in information structures in these problems, our comparison shows (i) that the sufficient conditions are different, and (ii) that their asymptotic properties are also different.¹³ While we show that full implementation is possible for the settings we consider, these observations may give an impression that achieving the general result is nontrivial. Further analyses on this type of environments are certainly needed for our understanding.

¹⁰See also Bergemann and Morris (2005) for *robust mechanism design*.

¹¹The notion of rationalizability is due to Bernheim (1984) and Pearce (1984). For games with incomplete information, see also Battigalli and Siniscalchi (2003) for Δ -rationalizability, which we employ in this paper, and Dekel, Fudenberg, and Morris (2007) for interim correlated rationalizability. Bergemann, Morris and Tercieux (2011) analyze *rationalizable implementation* with complete information.

¹²Exceptions include (i) every player’s type is the lowest in Captain MacWhirr’s problem and (ii) one player has the highest while the other has the second highest in the smashed antiques problem.

¹³As mentioned above, in our analysis on the smashed antiques problem, we allow for an arbitrary number of players.

2 Preliminaries

Let $N = \{1, \dots, n\}$ be the finite set of players with $n \geq 2$. Let $\Theta_i \subset \mathbb{N}$ correspond to the finite set of *types*. We call $\theta = (\theta_1, \dots, \theta_n) \in \Theta = \times_{i \in N} \Theta_i$ a *state*. Each player i simultaneously draws $\theta_i \in \Theta_i$ which is only known to herself. Let $f : \Theta \rightarrow \mathbb{R}^n$ be a social choice function. For each $i \in N$, let $u_i : \mathbb{R}^n \times \Theta \rightarrow \mathbb{R}$ be player i 's utility function.

A social choice function f is said to satisfy *ex-post incentive compatibility* if for each $i \in N$, $\theta_i, \theta'_i \in \Theta_i$, $\theta_{-i} \in \Theta_{-i}$, we have

$$u_i(f(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})) \geq u_i(f(\theta'_i, \theta_{-i}), (\theta_i, \theta_{-i})).$$

The notion of *strict ex-post incentive compatibility* requires the inequality to be strict for $\theta_i \neq \theta'_i$. Let $\rho_i \in \Delta(\Theta)$ be a prior player i has and $\rho_{\theta_i} \in \Delta(\Theta_{-i})$ once player i observes $\theta_i \in \Theta_i$ (following the Bayes' rule). A social choice function is said to satisfy *Bayesian incentive compatibility* if for each $i \in N$, $\theta_i, \theta'_i \in \Theta_i$, we have

$$\sum_{\theta_{-i} \in \Theta_{-i}} \rho_{\theta_i}(\theta_{-i}) u_i(f(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})) \geq \sum_{\theta_{-i} \in \Theta_{-i}} \rho_{\theta_i}(\theta_{-i}) u_i(f(\theta'_i, \theta_{-i}), (\theta_i, \theta_{-i})).$$

The mechanism is $\mathcal{M} = (M_1, \dots, M_n, g)$ where M_i is a message space for each $i \in N$ while $g : M \rightarrow \Delta(\Theta)$ is the outcome function.¹⁴ Our focus is on the *direct* mechanism, that is, $M_i = \Theta_i$ for each $i \in N$.

To respect the desirability of weaker informational assumptions, we provide a sufficient condition on players' (first-order) beliefs in each problem. To do so, we adopt *Δ -rationalizability* from Battigalli and Siniscalchi (2003).¹⁵ The restriction on players' beliefs we propose (sufficient condition) corresponds to the prefix " Δ ", the collection of the sets of players' beliefs satisfying certain (given) conditions. We assume that there is a common belief of Δ in each problem.

When the mechanism \mathcal{M} is played, each player $i \in N$ knows her own type $\theta_i \in \Theta_i$

¹⁴The reason that the set of outcomes is $\Delta(\Theta)$ is that in the mechanism we construct for Captain MacWhirr's problem, the planner uses lotteries when she observes a message profile inconsistent with the information she has. Otherwise, each player receives an element in Θ_i . We will later specify this for each problem.

¹⁵See also Battigalli (1999, 2003), Battigalli and Siniscalchi (2007), and Battigalli and Prestipino (2012). For the relationship among three different notions of rationalizability (belief-free, Δ , and interim correlated), see Battigalli, Di Tillo, Grillo, and Penta (2011).

and forms a belief about the other players' types $\theta_{-i} \in \Theta_{-i}$ as well as their messages $m_{-i} \in M_{-i}$. Let $q_{\theta_i} \in \Delta(\Theta_{-i} \times M_{-i})$ denote player i 's belief about the other players' types $\theta_{-i} \in \Theta_{-i}$ as well as messages $m_{-i} \in M_{-i}$ when her type is θ_i . Let $p_{\theta_i} \in \Delta(\Theta_{-i})$ be player i 's belief about the other players' types θ_{-i} obtained from q_{θ_i} (that is, $p_{\theta_i} = \text{marg}_{\Theta_{-i}} q_{\theta_i}$). Our sufficient conditions are placed on p_{θ_i} for each $i \in N$, corresponding to Δ (and hence not on the opponents' behavior).

Given the collection of the sets of players' beliefs satisfying this restriction, Δ , after the realization of her type, each player eliminates messages which cannot be best responses to any of her beliefs satisfying the condition. Then, with the reduced set of type-message combinations, each player again eliminates messages which cannot be best responses to any of her beliefs which satisfy the condition and take into account the fact that some type-message combinations would not arise for the other players. This continues until there is no more message eliminated for each possible type of each player. The remaining type-message combinations give the set of rationalizable actions for each possible type of each player.

Given \mathcal{M} , $\theta_i \in \Theta_i$ and $q_{\theta_i} \in \Delta(\Theta_{-i} \times M_{-i})$, let

$$BR_i(\theta_i, q_{\theta_i}) = \arg \max_{m'_i \in M_i} \sum_{m_{-i}, \theta_{-i}} u_i(g(m'_i, m_{-i}), (\theta_i, \theta_{-i})) q_{\theta_i}(\theta_{-i}, m_{-i})$$

be the set of best responses for player i with type θ_i .

Let Δ_{θ_i} be the set of beliefs for each θ_i satisfying the restriction and $\Delta = ((\Delta_{\theta_i})_{\theta_i \in \Theta_i})_{i \in N}$. The following iterative procedure eliminates messages for each θ_i for each $i \in N$.¹⁶

- $R_i^0 = M_i \times \Theta_i$; that is, the procedure starts with all possible combinations of messages and types for each player $i \in N$.
- At the k -th iteration where $k > 0$, for each $i \in N$ and $\theta_i \in \Theta_i$, a message m_i is a best response to a belief for the player with type θ_i which satisfies the condition Δ_{θ_i} and whose support only includes the type-message combinations of the others which have survived so far;

$$R_i^k = \left\{ (\theta_i, m_i) \in \Theta_i \times M_i \mid \begin{array}{l} \text{there exists } q_{\theta_i} \in \Delta_{\theta_i} \cap \Delta(R_{-i}^{k-1}) \\ \text{such that } m_i \in BR_i(\theta_i, q_{\theta_i}) \end{array} \right\}.$$

¹⁶Our result requires iterative elimination of strictly dominated strategies, excluding the possibility of implementation in dominant strategies (at least in our framework).

- For each $i \in N$, let $R_i = \bigcap_{k=0}^{\infty} R_i^k$.

For each $(\theta_i, m_i) \in R_i$, we say that a message m_i is *rationalizable* to player i with θ_i . Let $R_i(\theta_i) = \{m_i \in M_i \mid (\theta_i, m_i) \in R_i\}$ and $R(\theta) = \times_{j \in N} R_j(\theta_j)$. We say that a social choice function f is fully implemented via \mathcal{M} if for every $\theta \in \Theta$ and $m \in R(\theta)$, $g(m) = f(\theta)$.¹⁷

2.1 Social Choice Function and Incentive Compatibility

In Captain MacWhirr’s problem, the planner (Captain MacWhirr) wants to return the exact amount each worker saved while in the smashed antiques problem, the planner (airline company manager) wants to compensate for the damaged antiques. In each case, the amount the planner intends to give to each player, i.e., each player’s type, is the private information only the corresponding player holds. This means that we have a very specific social choice function which applies to both problems described above. That is, the planner’s goal is to give θ_i to each $i \in N$, which implies that for each $\theta \in \Theta$, $f(\theta) = \theta$.

We impose the following natural assumption: For each $i \in N$ and $\theta_i \in \Theta_i$, player i ’s utility function is such that $u_i(f(\theta'_i, \theta_{-i}), (\theta_i, \theta_{-i}))$ is strictly increasing in θ'_i for each $\theta_{-i} \in \Theta_{-i}$. In other words, the higher the amount she receives, the higher utility she obtains independent of the actual state. This implies that for each $i \in N$ and $\theta_i, \theta'_i \in \Theta_i$ with θ_i being the actual type and $\theta_i < \theta'_i$, we have

$$u_i(f(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})) < u_i(f(\theta'_i, \theta_{-i}), (\theta_i, \theta_{-i}))$$

for every $\theta_{-i} \in \Theta_{-i}$. This simply means that ex-post incentive compatibility and Bayesian incentive compatibility do not hold.¹⁸ Moreover, this observation is consistent with the fact that the each player has an incentive to “exaggerate.” Since the incentive compatibility condition is necessary for full implementation, this simply implies that the desired goal is not attainable for the implementation problems we consider with the standard approaches.

¹⁷Our result shows that for every $i \in N$ and $\theta_i \in \Theta_i$, $R_i(\theta_i) = \{\theta_i\}$ and hence $R(\theta) = \{\theta\}$.

¹⁸Hence, this also implies that the social choice function does not satisfy *robust monotonicity* in Bergemann and Morris (2011, Lemma 1). This also implies that *Bayesian monotonicity* does not hold since, irrelevant of what the other players’ messages (deceptions), player i ’s deception “ $\bar{\theta}$ independent of the actual θ_i ” leads to a payoff of $\bar{\theta}$.

2.2 Setup

We assume that the planner acquires *some* information regarding the actual state without incurring any cost. Let Λ be a partition of Θ with $\lambda \in \Lambda$ being a typical element. Our assumption on the planner's knowledge regarding the actual state θ means that the planner observes an element $\lambda \in \Lambda$ where $\theta \in \lambda$. For example, if the planner observes $\lambda \in \Lambda$ with $\theta \in \lambda$ and λ is a singleton, the planner knows the actual state. We assume (i) that Λ is exogenously given and commonly known to the players, and (ii) that each $\lambda \in \Lambda$ is verifiable.

In *Typhoon*, Captain MacWhirr observes some information regarding the workers' savings, that is, the *total amount* of their savings. However, he does not utilize this fact, and rather simply allocates the money *equally* among the workers at the end.¹⁹ In our mechanism, the planner does not reveal this private information and designs the outcome function so that the outcome depends on not only the message profile but also its consistency with the information she holds.

In Basu (1994, 2007), the information that these antiques are identical is shared by the airline company manager and the travelers (on which the traveler's dilemma game is built). We rather assume that, after checking the antiques, the planner does not reveal information on whether they are identical. In other words, we introduce a possibility of *heterogeneity*. Moreover, we also assume that the planner sees which has a higher value (but not their actual values). Again, in our mechanism, the planner designs the outcome function so that the outcome depends on the message profile and its consistency with the fact she withholds.

For each implementation problem, given that the planner observes $\lambda \in \Lambda$, we require that $g(\theta) = f(\theta)$ for each $\theta \in \lambda$. That is, this only applies to the message profiles which are consistent with what the planner observes. We provide the complete description of the outcome functions later.

We further make the following assumptions. First, we assume that the set of possible types is identical for each player. That is, for any $i, j \in N$, $\Theta_i = \Theta_j$. For each $i \in N$, let $\underline{\theta}$ be the lowest possible type while $\bar{\theta}$ is the highest with $0 < \underline{\theta} < \bar{\theta}$.²⁰ Second, we assume that the players' preferences are *state-independent* and only depend on how

¹⁹Watts (2008b) mentions that "Equitable Division" was considered as a title by the author.

²⁰This means that each player earned some money in Captain MacWhirr's problem. In the smashed antiques problem, this means that their antiques have some values.

much she receives.²¹ That is, let $u_i : \mathbb{N} \rightarrow \mathbb{R}$ be player i 's utility function which only depends on (and is strictly increasing in) how much she receives. We further assume that u_i is linear in its argument.²²

Both problems have the following common structure:

1. Each player observes her type $\theta_i \in \Theta_i$. Remember that for each $i, j \in N$, $\Theta_i = \Theta_j = \{\underline{\theta}, \dots, \bar{\theta}\}$.
2. The planner observes $\lambda \in \Lambda$.²³ Note that Λ itself is commonly known among the players while λ is the planner's private information. We will define the structure of Λ for each problem later.
3. The social choice function f is such that for each $\theta \in \Theta$, $f(\theta) = \theta$.
4. The players play the game $\mathcal{M} = (M_1, \dots, M_n, g)$. Each player's message space is $M_i = \Theta_i$.
5. If $m = \theta \in \lambda$, then $g(\theta) = f(\theta)$. The complete description of g for each problem will be given later.

For each problem, we provide the description of Λ , the corresponding (direct) mechanism, and a sufficient condition, and establish the uniqueness result in the next two sections.

3 Captain MacWhirr's Problem

In Captain MacWhirr's problem, the planner observes the total amount of the players' savings. Let $T = \{n\underline{\theta}, n\underline{\theta}+1, \dots, n\bar{\theta}-1, n\bar{\theta}\}$ be the set of possible numbers the planner observes. This implies $|\Lambda| = |T| = n(\bar{\theta} - \underline{\theta}) + 1$. Given $t \in T$, let $\lambda(t) \in \Lambda$ be the set

²¹See Ben-Porath and Lipman (2011) for the implication of state-independent preferences in implementation problems under complete information.

²²The linearity assumption is for exposition purpose. As long as the utility functions are strictly increasing, our result on the smashed antiques problem works without any modification. For Captain MacWhirr's problem, although the result does not change qualitatively, it does change the expression of the sufficient condition; that is, the elements of the right hand side expression in Condition 1 are measured in utility terms.

²³In Captain MacWhirr's problem, the planner not only observes t , but also holds the actual amount.

of states where the total of the players' savings is equal to t . For each $t \in T$, we have

$$\lambda(t) = \left\{ \theta \in \Theta \mid \sum_{j \in N} \theta_j = t \right\}.$$

Note that for each $t, t' \in T$ with $t \neq t'$, $\lambda(t) \cap \lambda(t') = \emptyset$ and $\cup_{t \in T} \lambda(t) = T$. In the original story, the planner does not take advantage of her private information. Here, we assume that the planner keeps t (and hence $\lambda(t)$) as her private information and embeds this variable in the mechanism. This information will be revealed only after the realization of m .

Given the profile of messages $m \in M = \Theta$ and the planner's private information $\lambda(t)$ for $t \in T$, the outcome function g follows the rule below:

1. If $\sum_{j \in N} m_j \leq t$, the planner gives m_i to each player i (and keeps the rest).
2. If $\sum_{j \in N} m_j > t$, each player receives $r(z)$ where $z = \sum_{j \in N} m_j - t$ (and the planner keeps the rest).

In general, the planner constructs $r(\cdot)$ with the following properties: (i) $r(z) \in [0, \underline{\theta}]$ for $z \in \{1, \dots, n(\bar{\theta} - \underline{\theta})\}$, and (ii) $r(z)$ is *strictly decreasing* and *linear* in $z \in \{1, \dots, n(\bar{\theta} - \underline{\theta})\}$. Each player faces a lottery with two possible outcomes $\{0, \underline{\theta}\}$.²⁴ With the chance $\frac{z}{n(\bar{\theta} - \underline{\theta})}$, she receives 0 while she receives $\underline{\theta}$ with probability $\frac{n(\bar{\theta} - \underline{\theta}) - z}{n(\bar{\theta} - \underline{\theta})}$. Hence, a player's expected payoff for $z > 0$ is

$$r(z) = \left(\frac{n(\bar{\theta} - \underline{\theta}) - z}{n(\bar{\theta} - \underline{\theta})} \right) \underline{\theta}.$$

Under the specification above, given the other players' messages m_{-i} , if $z > 0$, announcing a higher number would lead to a strictly lower payoff. In such scenarios, the marginal cost of announcing a higher number is

$$d = r(z) - r(z + 1) = \frac{\underline{\theta}}{n(\bar{\theta} - \underline{\theta})}$$

²⁴We use lotteries since we assume the discreteness of money. If $\underline{\theta}$ is not sufficiently high, the planner cannot assign an adequate punishment (where each player receives some amount lower than $\underline{\theta}$) when she observes the total higher than t . If instead $\underline{\theta}$ is sufficiently high, there is no need of using lotteries, as we discuss below.

which plays a crucial role in the sufficient condition we provide. Note (i) that d is decreasing in $\bar{\theta}$, and (ii) that d is increasing in $\underline{\theta}$.

For a sufficiently large $\underline{\theta}$, the planner can instead construct $r(\cdot)$ without lotteries such that (i) $r(\cdot)$ is strictly decreasing in z , (ii) the image of $r(z)$ for $z \in \{1, \dots, n(\bar{\theta} - \underline{\theta})\}$ is a subset of $\{0, \dots, \underline{\theta} - 1\}$, and (iii) the following expression (i.e., the minimum of the marginal cost of announcing a higher number when $z > 0$) is sufficiently large:

$$\min_{z \in \{1, \dots, n(\bar{\theta} - \underline{\theta}) - 1\}} [r(z) - r(z + 1)] = d.$$

We now specify a sufficient condition. The condition below states that for each $i \in N$ and $\theta_i \in \Theta_i$, player i believes that the chance of everyone else's type being the lowest is sufficiently high. Note that if this is indeed the true state (i.e., every player's type is $\underline{\theta}$) and commonly known, given the outcome function specified above, stating the actual type is a strictly dominant strategy for each player. The condition below implies that as long as there is a common belief that each player believes it is very likely that the other players have the lowest type, the planner can achieve the desired outcome with the direct mechanism described above.

Take $p_{\theta_i} = \text{marg}_{\Theta_{-i}} q_{\theta_i}$, the marginal on the type profiles Θ_{-i} obtained from q_{θ_i} . We assume that each player i 's first-order belief satisfies the following condition.

Condition 1 For each $i \in N$ and $\theta_i \in \Theta_i \setminus \{\bar{\theta}\}$,

$$p_{\theta_i}(\theta_j = \underline{\theta} \text{ for all } j \neq i) > \frac{\bar{\theta} - \theta_i}{(\bar{\theta} - \theta_i) + d}.$$

We also assume that there is a common belief that the players' beliefs satisfy this sufficient condition.²⁵

If n is sufficiently large (i.e., d is small), it is harder for Condition 1 to be satisfied. Instead, if $\bar{\theta} - \underline{\theta}$ is relatively small and $\underline{\theta}$ is sufficiently large, the condition may be easily satisfied.²⁶ Intuitively, the former implies that deviation is less profitable while the latter implies severe punishment.

As a simple example, consider the case where $n = 3$, $\bar{\theta} = 220$ and $\underline{\theta} = 210$ (and

²⁵Take player i 's belief regarding the other players' types given θ_i as a conditional probability. It is important to note that we exclude neither correlation nor independence of types. This also applies to the smashed antiques problem.

²⁶Even in Conrad's story, this might have been the case since the workers had "worked in the same place and for the same length of time" (Conrad (2008, p.73)).

hence $\bar{\theta} - \underline{\theta} = 10$). Then, the planner can have $r(z) = 210 - 7z$ where $z \in \{1, \dots, 30\}$, which implies $d = 7$. Here, the planner does not use lotteries. The sufficient condition above requires each player i 's belief that $\theta_j = \underline{\theta}$ for all $j \neq i$ to be strictly higher than

$$\frac{\bar{\theta} - \theta_i}{\bar{\theta} - \theta_i + d} = \frac{220 - \theta_i}{227 - \theta_i} \leq \frac{10}{17} \approx 0.588.$$

Under Condition 1, we have the following result:

Proposition 1 *Given Condition 1, for each $i \in N$ and each $\theta_i \in \Theta_i$, $m_i = \theta_i$ is a unique rationalizable action.*

The proof for the result is given in Appendix A. Here, we outline the proof for the result. Compare two strategies $m_i = \theta_i$ and $m'_i > \theta_i$. The net *loss* from m_i compared to m'_i is simply $m'_i - \theta_i$ if $t \geq \sum_{j \in N} m_j$. However, there are cases where $t < \sum_{j \in N} m_j$ for which there are two possible reasons; (a) $\sum_{j \neq i} m_j$ is too high, or (b) t is too low. There is at least one scenario to which (b) applies. If $\theta_j = \underline{\theta}$ for each $j \neq i$, independent of m_{-i} , m'_i leads to an expected payoff which is strictly lower than $\underline{\theta}$. In this scenario, the net *gain* from m_i compared to m'_i is *at least* $(m'_i - \theta_i) + d$. If Condition 1 holds, this net expected gain always exceeds the net expected loss for any $m'_i > \theta_i$ independent of m_{-i} . Hence, $m_i = \theta_i$ strictly dominates $m'_i > \theta_i$ for each $i \in N$ and $\theta_i \in \Theta_i \setminus \{\bar{\theta}\}$. This means that no one claims more than her type, and, hence, each player receives what she claims. Since claiming less than her type simply lowers a player's payoff, $m_i = \theta_i$ strictly dominates $m''_i < \theta_i$ for each $\theta_i \in \Theta \setminus \{\underline{\theta}\}$. Hence, $m_i = \theta_i$ is a unique rationalizable action for each $i \in N$ and $\theta_i \in \Theta$.

4 Smashed Antiques Problem

The structure of Λ here is different from that of Captain MacWhirr's problem. For comparison purpose, we keep the assumption that there are $n \geq 2$ players while there are only two in the original story. In the original story of Basu (1994), the planner as well as the players know that the antiques are identical, that is for each $i, j \in N$, $\theta_i = \theta_j$. Instead, we assume that each player does not know the values of the other players' antiques. Given this, we assume that the planner can check whether the players' antiques are identical or not. For any $i, j \in N$, if their antiques are not

identical ($\theta_i \neq \theta_j$), we assume that the planner can see which antique has a higher value (but not their actual values), that is, either $\theta_i > \theta_j$ or $\theta_i < \theta_j$.

Assumption 1 *For any $i, j \in N$, the planner observes either $\theta_i = \theta_j$, $\theta_i > \theta_j$ or $\theta_i < \theta_j$.*

Hence, each element of $\lambda \in \Lambda$ contains every $\theta \in \Theta$ which maintains the same order. In other words, for each $\lambda \in \Lambda$,

$$\text{sgn}(\theta_i - \theta_j) = \text{sgn}(\theta'_i - \theta'_j)$$

for every $\theta, \theta' \in \lambda$ and $i, j \in N$ where $\text{sgn}(\cdot)$ is the sign function. Again, we assume that the planner keeps λ as her private information and reveals this only after the game is played.

The mechanism, which is simpler than the one in the previous problem, works as follows:

1. For each $i \in N$, if there exists $j \in N$ ($j \neq i$) such that $\theta_j > \theta_i$ and $m_j \leq m_i$, player i receives 0.
2. Otherwise, player i receives m_i .

Hence, the only case where player i does not receive what she claims (and only receives zero) is when there exists another player j whose type is strictly higher than hers, and their messages are not consistent with this fact.

A sufficient condition we need to achieve the desired outcome is the following:

Condition 2 *For each $i \in N$, $\theta_i, \tilde{\theta}_i \in \Theta_i \setminus \{\bar{\theta}\}$ where θ_i is the actual type of player i and $\tilde{\theta}_i > \theta_i$,*

$$\theta_i > p_{\theta_i}(\text{for each } j \neq i, \theta_j \leq \theta_i \text{ or } \theta_j > \tilde{\theta}_i) \tilde{\theta}_i.$$

Again, we assume that there is a common belief that the players' beliefs satisfy the sufficient condition. It is important to note that stating $\tilde{\theta}_i > \theta_i$ does not lead to punishment under the event “for each $j \neq i$, $\theta_j \leq \theta_i$ or $\theta_j > \tilde{\theta}_i$.”

Under Condition 2, we have the following result.

Proposition 2 *Given Condition 2, for each $i \in N$ and $\theta_i \in \Theta$, announcing θ_i is a unique rationalizable action.*

The proof, which is given in Appendix B, uses induction. Given $\theta_i \in \Theta_i$, assume that for every $\hat{\theta}_i > \theta_i$, $m_i = \hat{\theta}_i$ is a unique rationalizable action. Given this assumption and our mechanism, $m_i = \theta_i$ guarantees a payoff of θ_i for player i with θ_i . First, any $m'_i < \theta_i$ simply leads to a strictly lower payoff (namely m'_i) and hence is strictly dominated by $m_i = \theta_i$. Second, $m''_i > \theta_i$ leads to a payoff strictly higher than θ_i if there is no one whose type is strictly higher than θ_i and weakly lower than m''_i . Note that if there is a player whose type falls into this interval, she tells the truth (by the previous assumption) and hence player i receives a payoff of zero. Her payoff therefore depends on her beliefs about the other players' signals, which is reflected in Condition 2. The initial step does not require any condition since for any player with $\theta_i = \bar{\theta}$, it is a strictly dominant action for her to choose $m_i = \bar{\theta}$.

5 Comments

We consider two implementation problems sketched in two stories, *Typhoon* by Conrad and Basu (1994, 2007). In each problem, the planner acquires some information regarding the actual state. Although the social choice functions implied in these stories are identical, the planner receives different information; in the former, it is the total amount of their savings while it is the order of antiques' values for the latter (with our extension). As a consequence, we establish two distinct mechanisms (outcome functions) for these problems. Our results rely on (i) the availability of some information to the planner and (ii) conditions on the players' beliefs regarding the opponents' types.

As for the treatments of the players' beliefs, there are other studies which take similar routes. In first-price auction settings, Battigalli and Siniscalchi (2003) and Dekel and Wolinsky (2003) also utilize Δ -rationalizability.²⁷ In particular, Dekel and Wolinsky (2003) show the uniqueness result for large auctions (with discrete values and bids) which relies on the assumptions on the players' first-order beliefs regarding the opponents' values. In this sense, Dekel and Wolinsky (2003) is an inspiration to our study.

Artemov, Kunimoto, and Serrano (2011) also employ Δ -rationalizability for their analysis of robust virtual implementation. In their analysis, the planner has some information regarding the players' beliefs, i.e., Δ , and there is a common belief of

²⁷Robles and Shimoji (2012) extend the results of Dekel and Wolinsky (2003).

this. They introduce the notion of *first-order incentive compatibility* for social choice functions, which is concerned with the restrictions. Artemov, Kunitomo, and Serrano (2011) show that first-order incentive compatibility is necessary for robust virtual implementation. Just like Bayesian incentive compatibility, the social choice function considered in the current paper does not satisfy this condition.

The current paper is also related to a recent literature on implementation problems with the presence of *evidence*. Ben-Porath and Lipman (2011) and Kartik and Tercieux (2011) consider implementation problem under complete information where the players can decide which evidence to submit. Kartik and Tercieux (2011) introduce the condition called *evidence-monotonicity* which is similar to Maskin monotonicity but takes into account the cost of evidence provision. By considering the possibility of state-independent preferences, Ben-Porath and Lipman (2011) introduce the condition called *measurability*, which says that, for any pair of states, if the sets of evidence are identical for each player, the social choice function chooses the same outcome for these two states. Although the availability of “evidence” is crucial in Ben-Porath and Lipman (2011) and Kartik and Tercieux (2011) as well as in the current paper, the approaches are completely different. Note that in the current paper, the players have *no* discretion regarding the availability of information related to the actual state. Instead, the planner simply acquires information related to the actual state without incurring any cost.

Baliga, Corchon, and Sjöström (1997) consider an implementation problem where the planner is one of the players (under complete information among the other players). Our setting has a similar flavor as theirs in the sense that the planner responds to the players’ action profiles. In their paper, the planner maximizes her expected payoff (in a dynamic game) while in our study the social choice function is given to the planner who hence commits herself to it.

For the two implementation problems we consider, the sufficient conditions are significantly different. In Captain MacWhirr’s problem, the condition looks at a probability assigned to a certain type profile – everyone else has the lowest type – which is in fact a subset of the type profiles in consideration for smashed antiques problems. Moreover, it is possible to have two extreme cases where only one condition holds while the other does not.²⁸

²⁸This is possible by making $p_{\theta_i}(\theta_j = \underline{\theta})$ for all $j \neq i$ in Condition 1 close to one or p_{θ_i} (for each $j \neq i, \theta_j \leq \theta_i$ or $\theta_j > \tilde{\theta}_i$) in Condition 2 close to zero.

Assuming that the types are i.i.d., these sufficient conditions show distinct asymptotic properties. In Captain MacWhirr's problem, let \underline{p} be the probability that a player's type is $\underline{\theta}$. Condition 1 implies

$$\underline{p}^{n-1} > \frac{\bar{\theta} - \theta_i}{(\bar{\theta} - \theta_i) + d}.$$

Hence, for a sufficiently large n , the inequality above does not hold. In the smashed antiques problem, given $\theta_i \in \Theta_i$ and $\tilde{\theta}_i > \theta_i$, let $p(\theta_i, \tilde{\theta}_i)$ be the probability that player j 's type θ_j is such that $\theta_i < \theta_j \leq \tilde{\theta}_i$. Then, Condition 2 becomes

$$\theta_i > \left[1 - p(\theta_i, \tilde{\theta}_i)\right]^{n-1} \tilde{\theta}_i.$$

Contrary to the case above, for any set of parameter values, there exists \bar{n} such that for any $n > \bar{n}$, the inequality above holds.²⁹

These highlighted differences suggest that obtaining general results in this type of environments with different types of Λ 's is not a trivial exercise. We hope that the results in this paper would help our understanding of implementation problems even when incentive compatibility does not hold.

A Proof for Proposition 1

We first show that for each $i \in N$ and $\theta_i < \bar{\theta}$, any $m_i > \theta_i$ is strictly dominated by $m_i = \theta_i$, and hence $m_i > \theta_i$ is not a best response for each $i \in N$ and $\theta_i < \bar{\theta}$. Then, we show that $m_i = \theta_i$ is a unique rationalizable strategy for each $i \in N$ and $\theta_i \in \Theta$.

Take $\theta_i \in \Theta \setminus \{\bar{\theta}\}$. Let $t^i = \sum_{j \neq i} \theta_j$ and $m^i = \sum_{j \neq i} m_j$. For the following expressions in A, let m_i be a message such that $m_i > \theta_i$ (we later show that m_i is strictly dominated by θ_i). The expected payoff from announcing θ_i can be computed as follows – player i receives θ_i as long as $m^i \leq t^i$;

$$A = \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=(n-1)\underline{\theta}}^{t^i} q_{\theta_i}(t^i, m^i) \theta_i + \sum_{m^i=t^i+1}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i) r(m^i - t^i) \right\}$$

²⁹In smashed antiques problems, as n becomes sufficiently large, the planner eventually obtains the exact order over θ_i 's, which makes the problem trivial.

$$= \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} \left\{ \sum_{m^i=(n-1)\underline{\theta}}^{t^i} q_{\theta_i}(t^i, m^i)\theta_i + \sum_{m^i=t^i+1}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i)r(m^i - t^i) \right\} \quad (1)$$

$$+ \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=(n-1)\underline{\theta}}^{t^i-(m_i-\theta_i)} q_{\theta_i}(t^i, m^i)\theta_i \right\} \quad (2)$$

$$+ \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=t^i-(m_i-\theta_i)+1}^{t^i} q_{\theta_i}(t^i, m^i)\theta_i \right\} \quad (3)$$

$$+ \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=t^i+1}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i)r(m^i - t^i) \right\}. \quad (4)$$

For the first summations in (2), (3), and (4),

$$(n-1)\bar{\theta} - [(n-1)\underline{\theta} + (m_i - \theta_i)] \geq (n-1)(\bar{\theta} - \underline{\theta}) - (\bar{\theta} - \underline{\theta}) = (n-2)(\bar{\theta} - \underline{\theta}) \geq 0$$

For the second summation in (2), note that

$$[t^i - (m_i - \theta_i)] - (n-1)\underline{\theta} = t^i - [(n-1)\underline{\theta} + (m_i - \theta_i)] \geq 0$$

since $t^i \in \{(n-1)\underline{\theta} + (m_i - \theta_i), \dots, (n-1)\bar{\theta}\}$ in (2). In addition, note that even if we change the upper bound of the first summation in (4) from $(n-1)\bar{\theta}$ to $(n-1)\bar{\theta} - 1$, it does not affect the expression.³⁰

The expected payoff from announcing $m_i \in \{\theta_i + 1, \dots, \bar{\theta}\}$ can be computed as follows – the players will be punished independent of m^i if $t^i < (n-1)\underline{\theta} + (m_i - \theta_i)$ while player i receives m_i if $m^i \leq t^i - (m_i - \theta_i)$:

$$\begin{aligned} B = & \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} \left\{ \sum_{m^i=(n-1)\underline{\theta}}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i)r((m^i - t^i) + (m_i - \theta_i)) \right\} \\ & + \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=(n-1)\underline{\theta}}^{t^i-(m_i-\theta_i)} q_{\theta_i}(t^i, m^i)m_i \right\} \\ & + \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=t^i-(m_i-\theta_i)+1}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i)r((m^i - t^i) + (m_i - \theta_i)) \right\} \end{aligned}$$

³⁰This also applies to the first summation in (8) below.

$$= \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} \left\{ \sum_{m^i=(n-1)\underline{\theta}}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i) r((m^i - t^i) + (m_i - \theta_i)) \right\} \quad (5)$$

$$+ \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=(n-1)\underline{\theta}}^{t^i-(m_i-\theta_i)} q_{\theta_i}(t^i, m^i) m_i \right\} \quad (6)$$

$$+ \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=t^i-(m_i-\theta_i)+1}^{t^i} q_{\theta_i}(t^i, m^i) r((m^i - t^i) + (m_i - \theta_i)) \right\} \quad (7)$$

$$+ \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=t^i+1}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i) r((m^i - t^i) + (m_i - \theta_i)) \right\} \quad (8)$$

We now show that $A > B$ under Condition 1 independent of m_{-i} . Note that (3) \geq (7) and (4) \geq (8). In addition, consider (2) $-$ (6):

$$\begin{aligned} & (2) - (6) \\ &= -[m_i - \theta_i] \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=(n-1)\underline{\theta}}^{t^i-(m_i-\theta_i)} q_{\theta_i}(t^i, m^i) \right\} \\ &= -[m_i - \theta_i] \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=(n-1)\underline{\theta}}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i) - \sum_{m^i=t^i-(m_i-\theta_i)+1}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i) \right\} \\ &= [m_i - \theta_i] \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=t^i-(m_i-\theta_i)+1}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i) \right\} - [m_i - \theta_i] \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} p_{\theta_i}(t^i) \\ &= [m_i - \theta_i] \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=t^i-(m_i-\theta_i)+1}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i) \right\} \\ &\quad - [m_i - \theta_i] \left\{ \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\bar{\theta}} p_{\theta_i}(t^i) - \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} p_{\theta_i}(t^i) \right\} \\ &= [m_i - \theta_i] \sum_{t^i=(n-1)\underline{\theta}+(m_i-\theta_i)}^{(n-1)\bar{\theta}} \left\{ \sum_{m^i=t^i-(m_i-\theta_i)+1}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i) \right\} \\ &\quad - [m_i - \theta_i] \left\{ 1 - \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} p_{\theta_i}(t^i) \right\} \end{aligned}$$

$$\geq -[m_i - \theta_i] + [m_i - \theta_i] \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} p_{\theta_i}(t^i).$$

Remember the definition of d – the lower bound for the marginal cost of choosing a higher message when $\sum_{j \in N} m_j > \sum_{j \in N} \theta_j$. Take $A - B$:

$$\begin{aligned} & A - B \\ & \geq [(1) - (5)] + [(2) - (6)] \\ & \geq \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} \left\{ \sum_{m^i=(n-1)\underline{\theta}}^{t^i} q_{\theta_i}(t^i, m^i) \theta_i + \sum_{m^i=t^i+1}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i) r(m^i - t^i) \right\} \\ & \quad - \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} \left\{ \sum_{m^i=(n-1)\underline{\theta}}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i) r((m^i - t^i) + (m_i - \theta_i)) \right\} \\ & \quad - [m_i - \theta_i] + [m_i - \theta_i] \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} p_{\theta_i}(t^i) \\ & = \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} \left\{ \sum_{m^i=(n-1)\underline{\theta}}^{t^i} q_{\theta_i}(t^i, m^i) \{ \theta_i - r((m^i - t^i) + (m_i - \theta_i)) \} \right\} \\ & \quad + \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} \left\{ \sum_{m^i=t^i+1}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i) \{ r(m^i - t^i) - r((m^i - t^i) + (m_i - \theta_i)) \} \right\} \\ & \quad - [m_i - \theta_i] + [m_i - \theta_i] \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} p_{\theta_i}(t^i) \\ & \geq d \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} \left\{ \sum_{m^i=(n-1)\underline{\theta}}^{t^i} q_{\theta_i}(t^i, m^i) + \sum_{m^i=t^i+1}^{(n-1)\bar{\theta}} q_{\theta_i}(t^i, m^i) \right\} \\ & \quad - [m_i - \theta_i] + [m_i - \theta_i] \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} p_{\theta_i}(t^i) \\ & = d \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} p_{\theta_i}(t^i) - [m_i - \theta_i] + [m_i - \theta_i] \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} p_{\theta_i}(t^i) \\ & = -[m_i - \theta_i] + [m_i - \theta_i + d] \sum_{t^i=(n-1)\underline{\theta}}^{(n-1)\underline{\theta}+(m_i-\theta_i)-1} p_{\theta_i}(t^i) \end{aligned}$$

$$\begin{aligned}
&\geq -[m_i - \theta_i] + [m_i - \theta_i + d]p_{\theta_i}(t^i = (n-1)\underline{\theta}) \\
&= -[m_i - \theta_i] + [m_i - \theta_i + d]p_{\theta_i}(\theta_j = \underline{\theta} \text{ for all } j \neq i) \\
&= [m_i - \theta_i + d] \left\{ p_{\theta_i}(\theta_j = \underline{\theta} \text{ for all } j \neq i) - \frac{m_i - \theta_i}{m_i - \theta_i + d} \right\} \\
&\geq [m_i - \theta_i + d] \left\{ p_{\theta_i}(\theta_j = \underline{\theta} \text{ for all } j \neq i) - \frac{\bar{\theta} - \theta_i}{\bar{\theta} - \theta_i + d} \right\} \\
&> 0
\end{aligned}$$

where the last inequality comes from Condition 1. This implies that $A - B > 0$ independent of m_{-i} . Hence, we have the following result.

Lemma 1 *Given Condition 1, for any $i \in N$, any $m_i > \theta_i$ is strictly dominated by $m_i = \theta_i$ for each $\theta_i < \bar{\theta}$.*

Given Lemma 1, each player knows that no opponent chooses $m_i > \theta_i$. Hence, in each possible outcome, each player with her endowment equal to θ_i receives an amount at most θ_i . Given this, it is easy to see that choosing $m_i < \theta_i$ is strictly dominated since such an announcement strictly lowers the payoff.

Lemma 2 *Given Condition 1, for each $i \in N$, any $m_i < \theta_i$ is strictly dominated by $m_i = \theta_i$ for each $\theta_i > \underline{\theta}$.*

The results above imply Proposition 1.

B Proof for Proposition 2

First, $m_i = \bar{\theta}$ strictly dominates $m'_i < \bar{\theta}$ for each $i \in N$ and $\theta_i = \bar{\theta}$. Note that the payoff from $m_i = \bar{\theta}$ is simply $\bar{\theta}$ while the payoff from $m'_i < \bar{\theta}$ is m'_i which is strictly lower than $\bar{\theta}$. It is important to note that this argument does not depend on the players' beliefs regarding the other players' types when the type is $\theta_i = \bar{\theta}$.

Next, take $i \in N$ and $\theta_i = \bar{\theta} - 1$. Given the argument above, it is clear that $m'_i < \bar{\theta} - 1$ is strictly dominated by $m_i = \bar{\theta} - 1$ since the former gives the payoff of m'_i while the latter gives $\bar{\theta} - 1 (> m'_i)$. In addition, $m''_i = \bar{\theta}$ is strictly dominated by $m_i = \bar{\theta} - 1$ for each $i \in N$ and $\theta_i = \bar{\theta} - 1$. The expected payoff from $m''_i = \bar{\theta}$ is

$$p_{\theta_i}(\text{for each } j \neq i, \theta_j \leq \bar{\theta} - 1)\bar{\theta} \tag{9}$$

while the expected payoff from $m_i = \bar{\theta} - 1$ is simply $\bar{\theta} - 1$. Then, the difference from the payoff from announcing $\bar{\theta} - 1$ and the expected payoff from announcing $\bar{\theta}$ (the only higher type) is

$$\begin{aligned} & (\bar{\theta} - 1) - (9) \\ &= (\bar{\theta} - 1) - p_{\theta_i}(\text{for each } j \neq i, \theta_j \leq \bar{\theta} - 1)\bar{\theta} \end{aligned}$$

which is strictly positive if

$$(\bar{\theta} - 1) > p_{\theta_i}(\text{for each } j \neq i, \theta_j \leq \bar{\theta} - 1)\bar{\theta}.$$

Note that this inequality holds given Condition 2.

Consider $i \in N$ and $\theta_i \in \Theta \setminus \{\bar{\theta}\}$. Suppose that for each $i \in N$ and $\tilde{\theta}_i \in \{\theta_i + 1, \dots, \bar{\theta}\}$, $m_i = \tilde{\theta}_i$ is a unique rationalizable strategy. First, $m'_i < \theta_i$ is strictly dominated by $m_i = \theta_i$; the former gives a payoff of m'_i while the latter gives $\theta_i (> m'_i)$. In addition, $m_i = \theta_i$ strictly dominates $m''_i > \theta_i$ for θ_i . The payoff from $m_i = \theta_i$ is simply θ_i while the expected payoff from $m''_i > \theta_i$ is

$$p_{\theta_i}(\text{for each } j \neq i, \theta_j \leq \theta_i \text{ or } \theta_j > m''_i)m''_i \tag{10}$$

Then, the difference between the payoff from announcing θ_i and the expected payoff from announcing $m''_i > \theta_i$ is

$$\begin{aligned} & \theta_i - (10) \\ &= \theta_i - p_{\theta_i}(\text{for each } j \neq i, \theta_j \leq \theta_i \text{ or } \theta_j > m''_i)m''_i \end{aligned}$$

which is strictly positive if

$$\theta_i > p_{\theta_i}(\text{for each } j \neq i, \theta_j \leq \theta_i \text{ or } \theta_j > m''_i)m''_i.$$

Again, the inequality holds given Condition 2. Induction leads to Proposition 2.

References

- [1] Artemov, G., T. Kunimoto, and R. Serrano (2011): “Robust Virtual Implementation with Incomplete Information: Towards a Reinterpretation of the Wilson Doctrine,” Working Paper.
- [2] Baliga, S., L. C. Corchon, and T. Sjöström (1997): “The Theory of Implementation When the Planner is a Player,” *Journal of Economic Theory*, 77, 15–33.
- [3] Battigalli, P. (1999): “Rationalizability in Incomplete Information Games,” Working Paper (ECO No. 99/17), European University Institute
- [4] Battigalli, P. (2003): “Rationalizability in Infinite, Dynamic Games with Incomplete Information,” *Research in Economics*, 57, 1–38.
- [5] Battigalli, P., A. Di Tillio, E. Grillo, and A. Penta (2011): “Interactive Epistemology and Solution Concepts for Games with Asymmetric Information,” *The B.E. Journal of Theoretical Economics*, 1 (*Advances*), Article 6.
- [6] Battigalli, P. and A. Prestipino (2012): “Transparent Restrictions on Beliefs and Forward-Induction Reasoning in Games with Asymmetric Information,” Working Paper.
- [7] Battigalli, P. and M. Siniscalch (2003): “Rationalization and Incomplete Information,” *Advances in Theoretical Economics*, 3, Article 3.
- [8] Battigalli, P. and M. Siniscalchi (2003): “Rationalizable Bidding in First-Price Auctions,” *Games and Economic Behavior*, 45, 38–72.
- [9] Battigalli, P. and M. Siniscalchi (2007): “Interactive Epistemology in Games with Payoff Uncertainty,” *Research in Economics*, 61, 165–184.
- [10] Basu, K. (1994): “The Traveler’s Dilemma: Paradoxes of Rationality in Game Theory,” *American Economic Review*, 84, 391–395.
- [11] Basu, K. (2007): “The Traveler’s Dilemma,” *Scientific American*, 296, 68–73.
- [12] Ben-Porath, E. and B. L. Lipman (2012): “Implementation with Partial Provability,” *Journal of Economic Theory*, forthcoming.

- [13] Bergemann, D. and S. Morris (2005): “Robust Mechanism Design,” *Econometrica*, 73, 1711–1813.
- [14] Bergemann, D. and S. Morris (2009): “Robust Implementation in Direct Mechanisms,” *Review of Economic Studies*, 76, 1175–1204.
- [15] Bergemann, D. and S. Morris (2011): “Robust Implementation in General Mechanism,” *Games and Economic Behavior*, 71, 261–281.
- [16] Bergemann, D., S. Morris, and O. Tercieux (2011): “Rationalizable Implementation,” *Journal of Economic Theory*, 146, 1253–1274.
- [17] Bernheim, D. (1984): “Rationalizable Strategic Behavior,” *Econometrica*, 52, 1007–1028.
- [18] Cabrales, A., A. Calvó-Armengol, and M. O. Jackson (2003): “*La Crema*: A Case Study of Mutual Fire Insurance,” *Journal of Political Economy*, 111, 425–458.
- [19] Conrad, J. (2008): *Typhoon and Other Tales*, Oxford: Oxford University Press.
- [20] Dasgupta, P. S., P. J. Hammond, and E. S. Maskin (1979): “The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility,” *Review of Economic Studies*, 46, 185–216.
- [21] Dekel, E., D. Fudenberg, and S. Morris (2007): “Interim Correlated Rationalizability,” *Theoretical Economics*, 2, 15–40.
- [22] Dekel, E. and A. Wolinsky (2003): “Rationalizable Outcomes of Large Private-Value First-Price Auctions,” *Games and Economic Behavior*, 43, 175–188.
- [23] Gay, J. (1969): *Fables*, Menston: Scolar Press Limited.
- [24] Harris, M. and R. M. Townsend (1981): “Resource Allocation under Asymmetric Information,” *Econometrica*, 49, 33–64.
- [25] Jackson, M. O. (1991): “Bayesian Implementation,” *Econometrica*, 59, 461–477.
- [26] Kartik, N. and O. Tercieux (2012): “Implementation with Evidence,” *Theoretical Economics*, 7, 323–355.

- [27] Mumy, G. E. (1981): “A Superior Solution to Captain MacWhirr’s Problem: An Illustration of Information Problems and Entitlement Structures,” *Journal of Political Economy*, 89, 1039–1043.
- [28] Myerson, R. B. (1979): “Incentive Compatibility and the Bargaining Problem,” *Econometrica*, 61–73.
- [29] Palfrey, T. R., and S. Srivastava (1989): “Implementation with Incomplete Information in Exchange Economies,” *Econometrica*, 57, 115–134.
- [30] Pearce, D. (1984): “Rationalizable Strategic Behavior and the Problem of Perfection,” *Econometrica*, 52, 1029–1050.
- [31] Postlewaite, A. and D. Schmeidler (1986): “Implementation in Differential Information Economies,” *Journal of Economic Theory*, 39, 14–33.
- [32] Robles, J. and M. Shimoji (2012): “On Rationalizability and Beliefs in Discrete Private-Value First-Price Auctions,” *The B.E. Journal of Theoretical Economics*, 12 (Contributions), Article 16.
- [33] Saraydar, E. (1983): “Transaction Costs and the Solution to Captain MacWhirr’s Problem,” *Journal of Political Economy*, 91, 312–15.
- [34] Schweinzer, P. and M. Shimoji (2011): “Captain MacWhirr’s Problem Revised,” Discussion Paper (11-12), Department of Economics and Related Studies, University of York.
- [35] Watts, C. (2008a): Note on the Text to *Typhoon and Other Tales* by Joseph Conrad, Oxford University Press.
- [36] Watts, C. (2008b): Explanatory Notes to *Typhoon and Other Tales* by Joseph Conrad, Oxford University Press.